



UNIVERSITAT DE VALÈNCIA

FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES
DEPARTAMENTO DE DIRECCIÓN DE EMPRESAS
(ADMINISTRACIÓN Y MARKETING)



**Modelos de determinación de la
cobertura y la distribución de
contactos en la planificación de
medios publicitarios impresos**

Facultad de Ciencias Económicas y Empresariales

Fecha de Entrada *29 de junio 1998*

Fecha de Lectura *30 septiembre 1998*

Calificación *sobresaliente con laud por unanimidad*

TESIS DOCTORAL

Presentada por
JOAQUÍN ALDÁS MANZANO

Dirigida por el
Dr. D. J. ENRIQUE BIGNÉ ALCANIZ
*Catedrático de Comercialización e Investigación de Mercados
Universitat Jaume I Castellón*

VALENCIA, JUNIO DE 1998

UMI Number: U607276

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U607276

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

UNIVERSITAT DE VALÈNCIA
CC SOCIALS
BIBLIOTECA
Nº Registre 2079
DATA 19/4/99
SIGNATURA CIDT 650
Nº LIBIS: 802434

Nº Dòcs 802414



Al profesor Kike Bigné:

“A mi dadme, os lo ruego, españoles sin trampa ni disfraz. Los que sin mesianismo y sin aparato trabajan lo mejor que pueden en la biblioteca, el laboratorio, el taller o el pegujal. Los que saben conversar, reír o llorar con sencillez, y a través de sus palabras, sus risas o sus lágrimas os dejan ver, allá en lo hondo, esa impagable realidad que solemos llamar “una persona”. Los que por hombría de bien, cristiana o no cristiana, saben ver y tratar como personas, como verdaderas personas, a quienes con ellos conviven. Los que frente a la jactancia ajena dicen “No será para tanto” y ante la desgracia propia saben decir “No importa”. Tantos y tantos así, entre los que todavía andan y esperan por las avenidas estruendosas o por las silenciosas callejuelas de España. Para que el vivir en mi tierra me sea de cuando en cuando consuelo o regalo, a mí dadme, os lo ruego, españoles sin trampa ni disfraz.”

Pedro Laín Entralgo (1970): *A qué llamamos España*. Madrid: Plaza y Janés.

Agradecimientos

Pensar que una tesis doctoral es fruto del exclusivo esfuerzo de su autor sería, además de una vanidad carente de fundamento, una falsedad. Muchas son las personas e instituciones que han contribuido, animado y favorecido este trabajo y, llegados a su conclusión, es de justicia dejar patente mi deuda con ellos.

Quiero, en primer lugar, agradecer a la *Universitat de València* la formación que, a lo largo de todos estos años, he recibido gracias a ella y a sus profesores, los que están y los que nos dejaron, pero no sin antes impregnar nuestro alma del ansia de conocimiento, el mayor activo que podremos nunca atesorar. No puedo olvidarme de todos mis *compañeros del Departamento de Dirección de Empresas* que han asumido gran parte de mis obligaciones durante este curso, permitiéndome así liberar un tiempo precioso para la conclusión de esta tesis doctoral. Quiero simbolizar esta gratitud en la persona del profesor *Salvador Miquel* del cual, cualquier alabanza que pudiera escribir, sería un pobre intento de reflejar su humanidad, cualidad que, por encima de muchas otras que también posee, le adorna.

A la *Asociación para la Investigación de Medios de Comunicación (AIMC)* y a *Luis Pistoni* y *Sergio Guerrero* de *ODEC*, les debo el acceso a los datos del Estudio General de Medios, pieza clave en el desarrollo de este trabajo. A la *Universitat Jaume I de Castellón*, deseo agradecer la hospitalidad demostrada al recibirme como profesor visitante en los últimos meses del desarrollo de esta tesis doctoral. Pocas instituciones habrán dado tanto a cambio de tan poco.

De *mis padres*, de *Mónica*, mi mujer y de *mis amigos* debo valorar la comprensión demostrada por el tiempo que este trabajo les ha robado, dado que, pocas veces, como en este periodo de nuestra vida, se está tan lejos de los que tenemos más cerca.

Y finalmente, quiero referirme a mi director, el profesor *Kike Bigné*. A él quiero dedicarle esta tesis doctoral porque, el valor que tenga, lo tiene gracias a él. No creo que nunca llegue a conocer a una persona capaz, como él, de unir a una capacidad de trabajo desbordante y una calidad científica indiscutible, el trato humano, cordial, exigente cuando era necesario, pero siempre cercano, siempre disponible, independientemente del día y la hora (por ello, gracias también, *Ana*). Como reza la dedicatoria, gracias por hacer que este periodo haya sido para mi, en todo momento, consuelo y regalo.

JOAQUÍN ALDÁS MANZANO
Valencia, junio de 1998

Índice

INTRODUCCIÓN	iv
1. LA MODELIZACIÓN EN LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS	1
1.1 LAS DECISIONES DE COMUNICACIÓN EN LA EMPRESA Y LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS	1
1.2 LA CRECIENTE IMPORTANCIA DE LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS	18
1.3 EL PAPEL DE LA COBERTURA Y LA FRECUENCIA DE CONTACTO EN LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS: NECESIDAD DE MODELIZACIÓN	29
1.4 EVOLUCIÓN DE LA LITERATURA SOBRE MODELIZACIÓN EN LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS	43
1.4.1 El peso de la modelización en los trabajos sobre planificación de medios publicitarios	43
1.4.2 La literatura sobre modelización en la planificación de medios	51
1.4.2.1 <i>Modelos de exposición</i>	54
1.4.2.2 <i>Modelos de planificación de medios publicitarios</i>	59
1.5 LA MODELIZACIÓN EN LA PRÁCTICA DE LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS	72
2. TIPOLOGÍA DE LOS MODELOS DE EXPOSICIÓN	79
2.1 INTRODUCCIÓN	79
2.2 MODELOS DE INDIVIDUACIÓN	82
2.2.1 Procedimientos de probabilización	82
2.2.2 Modelos de distribución exacta	89
2.2.3 Simplificaciones del cálculo exacto	93
2.2.3.1 <i>Reducción a un soporte</i>	93
2.2.3.2 <i>Reducción a un soporte con ajuste de cobertura</i>	94
2.2.3.3 <i>Método TOM</i>	95
2.2.4 Métodos de simulación	99
2.3 MODELOS DE AGREGACIÓN	103

3.	REVISIÓN DE LOS MODELOS DE EXPOSICIÓN	112
3.1	MODELOS DE ACUMULACIÓN DE AUDIENCIAS	112
3.1.1	Modelos <i>ad hoc</i>	112
3.1.1.1	<i>Modelo de acumulación logarítmica</i>	113
3.1.1.2	<i>Modelo de acumulación exponencial</i>	114
3.1.1.3	<i>Modelo de acumulación de Agostini</i>	116
3.1.1.4	<i>Modelo de acumulación de Hofmans</i>	118
3.1.2	Modelos estocásticos	121
3.1.2.1	<i>Modelo binomial de acumulación</i>	122
3.1.2.2	<i>Modelo de acumulación de dos segmentos</i>	123
3.1.2.3	<i>Modelo de acumulación de tres segmentos</i>	125
3.1.2.4	<i>Modelo de acumulación de Politz</i>	128
3.1.2.5	<i>Modelo de acumulación beta binomial</i>	132
3.1.2.6	<i>Modelo de acumulación de la binomial negativa</i>	135
3.1.2.7	<i>Modelo de acumulación de Broadbent</i>	138
3.1.2.8	<i>Modelo de acumulación de Morgensztern</i>	141
3.1.2.9	<i>Modelo de acumulación beta binomial modificado</i>	147
3.1.2.10	<i>Modelo de acumulación binomial de Markov</i>	149
3.2	MODELOS DE DUPLICACIÓN DE AUDIENCIAS	154
3.2.1	Modelos <i>ad hoc</i>	154
3.2.1.1	<i>Modelo de duplicación de Agostini</i>	154
3.2.1.2	<i>Modelo de duplicación de Hofmans</i>	156
3.2.1.3	<i>Modelo de duplicación de Little y Lodish</i>	158
3.2.1.4	<i>Fórmula de duplicación de Starch</i>	159
3.2.1.5	<i>Modelo de Lozano</i>	160
3.2.2	Modelos estocásticos	161
3.2.2.1	<i>Modelo binomial</i>	162
3.2.2.2	<i>Modelo de Sainsbury</i>	164
3.2.2.3	<i>Modelo de Kwerel</i>	166
3.2.2.4	<i>Fórmula de Sainsbury modificada</i>	168
3.2.2.5	<i>Modelo de agregación secuencial con duplicación constante</i>	171
3.2.2.6	<i>Modelo de agregación secuencial con duplicación variable</i>	173
3.2.2.7	<i>Modelo de independencia condicional</i>	174
3.2.2.8	<i>Fórmula de Hofmans ampliada</i>	176
3.2.2.9	<i>Fórmula de duplicación de Metheringham</i>	178
3.3	MODELOS DE AUDIENCIA NETA ACUMULADA	181
3.3.1	Modelos estocásticos	182
3.3.1.1	<i>Modelo binomial</i>	182
3.3.1.2	<i>Modelo binomial multivariable con independencia de soportes</i>	184
3.3.1.3	<i>Modelo multinomial</i>	186
3.3.1.4	<i>Modelo de Markov Bernouilli</i>	190
3.3.1.5	<i>Modelo de Metheringham</i>	196
3.3.1.6	<i>Modelo Dirichlet Multinomial</i>	199
3.3.1.7	<i>Modelo de la distribución Poisson Binomial</i>	203
3.3.1.8	<i>Modelo de la distribución Kwerel Geométrica</i>	205
3.3.1.9	<i>Modelo de la distribución Hofmans Geométrica</i>	207
3.3.1.10	<i>Modelo de la distribución Hofmans Beta Binomial</i>	208
3.3.1.11	<i>Modelo de la distribución Hofmans Poisson Binomial</i>	209
3.3.2	Modelos <i>ad hoc</i>	211
3.3.2.1	Modelo de Hofmans	211
3.3.2.2	Modelo de Morgensztern	212

3.3.2.3	Modelo promedio Metheringham Dirichlet Multinomial	214
---------	--	-----

4.	OBJETIVOS Y DISEÑO METODOLÓGICO	215
-----------	--	------------

4.1	OBJETIVOS: FORMULACIÓN Y JUSTIFICACIÓN	215
4.2	DISEÑO DE LA INVESTIGACIÓN	221
4.2.1	Fuente de datos: el Estudio General de Medios	225
4.2.2	Diseño de los planes de medios	243
4.2.3	Definición del error	254
4.2.3.1	Medidas del error en la estimación de la cobertura	255
4.2.3.2	Medidas del error en la estimación de la distribución de contactos	258
4.2.4	Factores de error	265
4.2.5	Técnicas de análisis de datos	269
4.2.5.1	Evaluación y jerarquización de los modelos de exposición	269
4.2.5.2	Determinación de los factores de error	275

5.	ANÁLISIS DE RESULTADOS E IMPLICACIONES	277
-----------	---	------------

5.1	INTRODUCCIÓN	277
5.2	MODELOS DE ACUMULACIÓN DE AUDIENCIAS	278
5.2.1	Error en la estimación de la cobertura	278
5.2.2	Error en la estimación de la distribución de contactos	290
5.2.3	Factores de error de los modelos de acumulación de audiencias	298
5.3	MODELOS DE DUPLICACIÓN DE AUDIENCIAS	305
5.3.1	Error en la estimación de la cobertura	305
5.3.2	Error en la estimación de la distribución de contactos	314
5.3.3	Factores de error de los modelos de duplicación de audiencias	322
5.4	MODELOS DE AUDIENCIA NETA ACUMULADA	330
5.4.1	Error en la estimación de la cobertura	331
5.4.2	Error en la estimación de la distribución de contactos	341
5.4.3	Factores de error de los modelos de audiencia neta acumulada	352
5.4.3	Corolario: propuesta de un nuevo modelo de audiencia neta acumulada	360

CONCLUSIONES	366
---------------------	------------

BIBLIOGRAFÍA	384
---------------------	------------

Introducción

El fuerte incremento que, en la última década, ha experimentado la inversión publicitaria en nuestro país, ha venido a acentuar la natural preocupación de los anunciantes, por obtener algún tipo de garantía de que estos recursos crecientes estén siendo invertidos de una manera eficaz y eficiente, a través de las agencias de publicidad y las centrales de compra de medios. Si bien es cierto que esta eficiencia ha de derivar de múltiples aspectos: adecuación de la campaña a los objetivos de la organización, integración correcta de las distintas herramientas comunicacionales, estrategia creativa acertada, no es menos cierto que un elemento clave en esta labor, lo ha de jugar la selección acertada de los medios y soportes que han de integrar el plan de medios de la campaña de publicidad, máxime cuando aproximadamente el 85% del presupuesto de publicidad va destinado a difusión. En términos generales, el problema

de la planificación de medios, es un problema económico clásico: unos recursos escasos, el presupuesto de publicidad, que debe ser distribuido entre usos alternativos, los soportes, de tal manera que se maximice una función objetivo determinada, en suma, la asignación eficiente de recursos escasos.

La labor del planificador de medios, sin embargo, ha ganado en complejidad toda vez que, en esos mismos diez años, se han producido numerosos cambios en el entorno mediático que han contribuido a aumentar la turbulencia del mismo. El importante incremento en el número de soportes, cuyo ejemplo arquetípico ha sido la irrupción de las televisiones privadas y autonómicas, unido a las desconexiones publicitarias regionales, se ha traducido en la fragmentación de la audiencia, la reducción de la fidelidad de ésta y, sobre todo, en unas políticas comerciales muy agresivas por parte de los medios que han redundado en un fuerte abaratamiento de los espacios publicitarios, el incremento de la publicidad emitida y, como consecuencia, en una previsible reducción de la eficacia de la publicidad emitida en estos entornos saturados.

El efecto combinado de estos dos hechos, fuerte incremento de la inversión, y complejidad creciente de la labor de la planificación de medios, deriva en la necesidad de, o bien encontrar nuevos instrumentos analíticos que, a través de la formalización, faciliten la tarea del planificador, o bien encontrar criterios para determinar cuál es la precisión de los ya existentes y, de esta forma, seleccionar los más adecuados. Y a esta labor, bajo nuestro punto de vista, ha de contribuir de una manera fundamental, la investigación llevada a cabo en el ámbito académico.

Uno de los elementos centrales dentro de estos instrumentos analíticos son, sin lugar a dudas, los modelos que han de estimar cuál es la cobertura y la distribución de contactos que se logra con un plan de medios determinado, en la nomenclatura que seguiremos en esta tesis doctoral, los *modelos de exposición*. Su importancia deriva de que estas variables, son fundamentales a la hora de valorar los distintos soportes para

poder seleccionar aquellos que son los más adecuados para formar parte del plan de medios y, también, para elegir entre planes de medios alternativos.

La atención prestada a estos modelos en nuestro país, sin embargo, no ha sido excesiva y, en la mayoría de los casos, se ha centrado en explicar el funcionamiento de los modelos, más que en evaluar la precisión de las estimaciones de los mismos. En el intento de contribuir a mitigar esta laguna, se inscribe la realización de esta tesis doctoral que, en términos generales, busca *dos grandes objetivos*. Por un lado, desde una perspectiva positiva, se pretende llevar a cabo la evaluación y jerarquización de los principales modelos de exposición desarrollados hasta el presente, atendiendo al error que cometen en la estimación de la cobertura y la distribución de contactos. Por otro lado, y bajo un enfoque normativo, derivar de los resultados anteriores las directrices que deben marcar el futuro desarrollo de nuevos modelos de exposición, para asegurar que sus estimaciones sean lo más precisas posible.

Para la realización de esta tarea, se ha dividido la presente tesis doctoral en cinco capítulos. En el *capítulo primero*, se sitúa la planificación de medios en el ámbito de la toma de decisiones de comunicación en la empresa, haciendo especial hincapié en destacar la influencia que, sobre la realización del plan de medios y soportes, tienen las decisiones adoptadas en cada una de las etapas anteriores de este proceso. A continuación, se analiza la forma en que la mencionada evolución del entorno mediático ha acentuado la dificultad y, paralelamente, la importancia de la labor del planificador de medios, justificando la necesidad de formalizar mediante modelización esta labor. Tras una revisión de la literatura existente sobre los modelos de exposición, en la cual se enfatiza el análisis del peso relativo, tanto cuantitativo como cualitativo, de estos trabajos en el ámbito de la modelización en comunicación, el capítulo finaliza con un estudio de la situación de la modelización en la práctica empresarial de la planificación de medios.

El *capítulo segundo*, por su lado, lleva a cabo una tipología de los modelos de exposición que van a ser objeto del proceso evaluador, prestándose especial atención en precisar las hipótesis en que se sustentan cada uno de ellos y que sirven de base para la clasificación que se ofrece.

En el *capítulo tercero* se procede a desarrollar los modelos de exposición, derivando las expresiones que permiten obtener la cobertura y la distribución de contactos en cada uno de ellos. Este capítulo se ha estructurado en tres grandes bloques, atendiendo al tipo de plan de medios para el que se ha diseñado cada modelo. Así, el primer bloque corresponde a los modelos de acumulación de audiencias, que son los concebidos para evaluar planes de medios formados por n inserciones en un sólo soporte. El segundo desarrolla los modelos de duplicación de audiencias, basados en la realización de una sola inserción en n soportes distintos y, finalmente, el tercer bloque es objeto de los modelos de carácter más general, los de audiencia neta acumulada, caracterizados por aplicarse a los planes de medios resultantes de realizar n inserciones en m soportes diferentes. En conjunto, se describen los cuarenta y dos modelos que han sido objeto de evaluación en esta tesis doctoral.

En el *capítulo cuarto* se explicitan los objetivos de la tesis doctoral. Básicamente, se pretende evaluar los modelos de exposición atendiendo al error que cometen en la estimación de la cobertura y la distribución de contactos, intentando derivar, en ese proceso, directrices para la construcción futura de nuevos modelos. A continuación, se describe el diseño experimental realizado para la evaluación de los modelos de exposición descritos en el capítulo anterior, que pasa por la comparación de las estimaciones de los modelos frente a los resultados obtenidos mediante tabulación de los datos del Estudio General de Medios. De la revisión de la literatura sobre los trabajos empíricos con objetivos análogos al nuestro, se desprenden las directrices que en el diseño metodológico han de seguirse para asegurar la comparabilidad de los resultados obtenidos. Así, se describe la base de datos utilizada dedicando especial atención a describir las transformaciones que ha sido necesario introducir en la misma.

Tras explicar el proceso seguido en la construcción de los planes de medios que se utilizarán en la evaluación, y explicar las razones que llevan a seleccionar el tamaño muestral empleado, se justifica la elección de las medidas de error de las estimaciones de la cobertura y la distribución de contactos que van a servir de base para jerarquizar a los distintos modelos según su precisión. El capítulo termina señalando las técnicas de análisis de datos aplicadas para el logro de los objetivos propuestos.

En el *capítulo cinco* se presentan y analizan los resultados obtenidos en el proceso de evaluación y jerarquización de modelos. Esta presentación de resultados se estructura de nuevo en tres grandes apartados que corresponden a los tres tipos de modelos descritos y, en cada uno de ellos, se establece una clasificación de los modelos atendiendo a la precisión en sus estimaciones, se señala la influencia que sobre el error ha tenido la introducción de hipótesis más reales en su construcción y que, por ello, han añadido complejidad al modelo y se acaba determinando qué relación existe entre las características estructurales del plan de medios, en términos de audiencia bruta y nivel de acumulación y duplicación de la audiencia y el error cometido por los modelos. De todos estos resultados se derivan y presentan una serie de directrices que parece recomendable seguir en la construcción de futuros modelos de exposición.

Como *corolario* al capítulo cinco, se propone un nuevo modelo de exposición, que hemos denominado *modelo Morgensztern beta binomial*, y que se ha construido atendiendo a las directrices propuestas. Se pretende de esta manera analizar la validez de las mismas dado que, si estas son acertadas, el nuevo modelo debe ofrecer una mayor precisión en la estimación de la cobertura y la distribución de contactos que los modelos precedentes.

1

La modelización en la planificación de medios publicitarios

1.1 LAS DECISIONES DE COMUNICACIÓN EN LA EMPRESA Y LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS

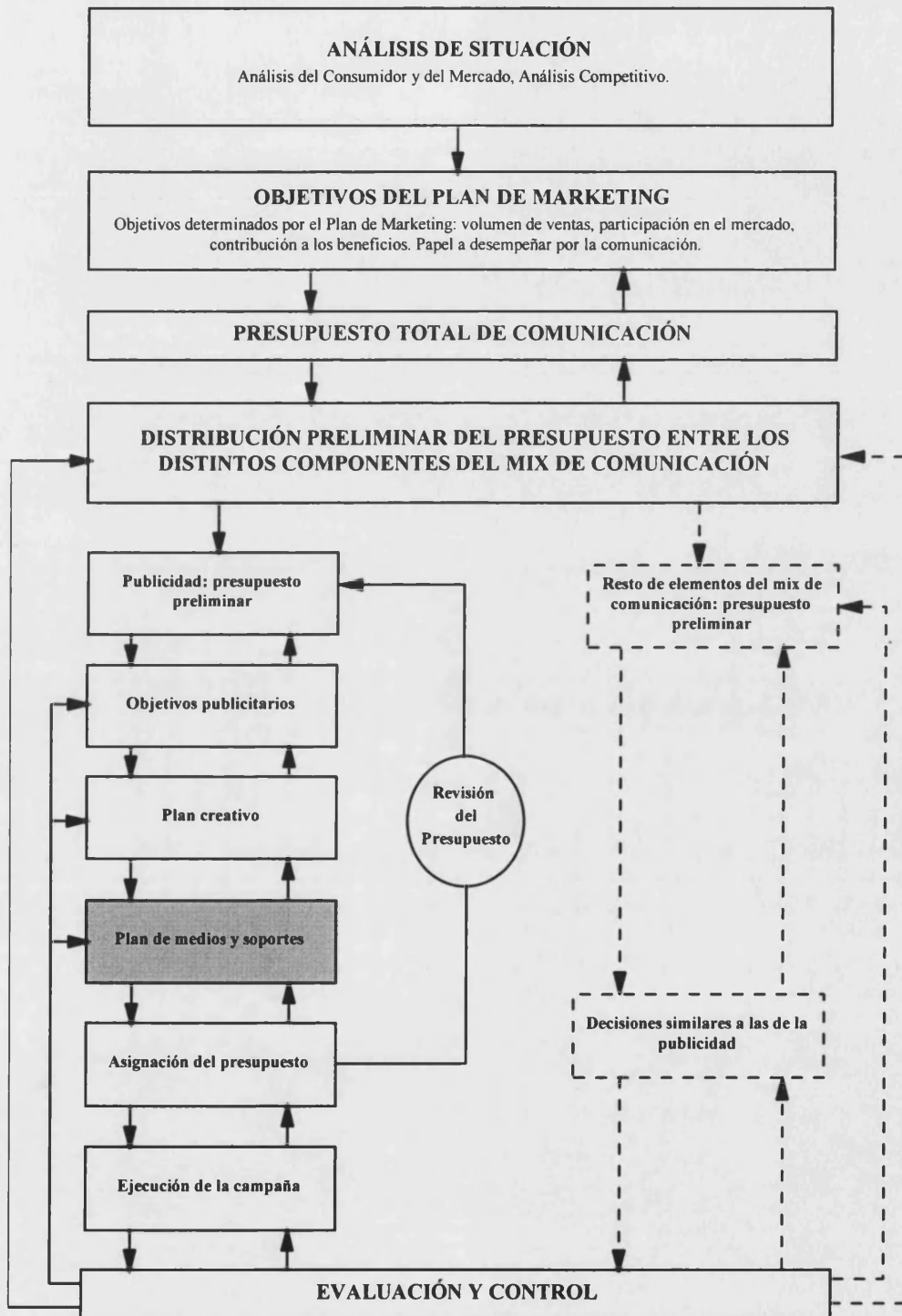
Existe un acuerdo generalizado en la consideración de que ni la publicidad ni, por tanto, cualquiera de sus componentes, como la planificación de medios, puede abordarse fuera del ámbito integrado de las decisiones de comunicación en marketing (Sissors y Bumba, 1996; Barban, Cristol y Kopec, 1994). En este sentido, adoptando un enfoque sistémico de la empresa por aquel entonces innovador y, por ello, meritorio, Martín Armario (1980) ya integraba sucesivamente los planes de medios y la creatividad en el plan de publicidad, éste en el de marketing y todos ellos

en el plan general de la empresa, poniendo de manifiesto el riesgo en el que puede incurrirse al cometer el error de considerar a la publicidad una actividad aislada y no subordinada al logro de los objetivos de marketing y generales de la organización.

El objetivo de este apartado reside, precisamente, en situar adecuadamente la planificación de medios en el contexto más general de las decisiones de marketing, intentando, por una parte, que la especialización inherente a una tesis doctoral no provoque la pérdida de la necesaria visión de conjunto que lleve a confundir la parte con el todo. Por contra, dado el carácter necesariamente generalista de los aspectos que se abordarán, tampoco se ha creído pertinente extenderse en demasía en unos aspectos que si bien tienen, como se verá, fuertes implicaciones con el objeto principal de este trabajo, no son el núcleo del mismo. Bajo este espíritu, se abordan secuencialmente el conjunto de decisiones que han de desembocar en una política de comunicación determinada, haciendo hincapié, especialmente, en las implicaciones que de cada una de ellas se desprenden para la planificación de medios.

Para Ray (1982), el hecho de que la comunicación sea la herramienta más importante y, entre las puestas a disposición de los mandos intermedios del departamento de marketing, la única con efectos a corto plazo, provoca que en muchos casos, ésta se inicie sin prestar la debida atención a los objetivos de marketing preestablecidos y, lo que es más importante, sin analizar las interacciones y posibles sinergias entre las distintas herramientas promocionales. Evitar esta tentación pasa, necesariamente, por ser conscientes del lugar que ocupan las herramientas de comunicación en el proceso completo de toma de decisiones de marketing que recoge la figura 1.1.

Figura 1.1. El marco de la planificación de medios.



Fuente: Elaboración propia a partir de Ray (1982)

Batra, Myers y Aaker (1996) atribuyen gran importancia al hecho de que el responsable de comunicación ha de ser conocedor del análisis de la situación que se llevó a cabo para el establecimiento del plan de marketing y en el que probablemente participó, únicamente, como miembro de un equipo. Este análisis, en el que se proporcionaron respuestas a decenas de preguntas relativas a la naturaleza y tamaño de la demanda, competencia, factores del entorno e internos a la empresa, han de ser analizadas de nuevo antes de abordar las decisiones propiamente comunicacionales. Como muestra el cuadro 1.1., informaciones tales como las fuentes principales de información acerca del producto que utiliza el consumidor, sujeto que toma la decisión de compra e influencias personales y familiares que recibe, y características sociodemográficas del consumidor actual o potencial, tendrán una importancia decisiva a la hora de proporcionar criterios adecuados para la selección de los medios y soportes más correcta. La relevancia de hacer uso de esta información, ha sido puesta de manifiesto por Weilbacher (1990) quien señala que el procedimiento más efectivo y eficiente de llegar a los consumidores potenciales pasa por ser capaz de relacionar adecuadamente sus estilos de vida con los medios, coincidiendo en este planteamiento con la necesidad expuesta por Katz (1990) de conocer más acerca de los hábitos individuales y cotidianos de los consumidores respecto a esos mismos medios.

Cuadro 1.1. Temas y elementos de un análisis de situación

A. Naturaleza de la demanda

1. Descripción de los patrones de comportamiento y actitudes de los compradores
2. Segmentación del mercado: Criterios sociodemográficos, económicos...

B. Tamaño de la demanda

1. Tamaño actual del mercado (unidades y pesetas) y predicción del tamaño futuro
2. Cuotas de mercado actuales
3. ¿Es mejor analizar el mercado en su conjunto o segmentado?

C. Naturaleza de la competencia

1. Estructura actual y futura de la competencia (número competidores, cuotas mercado, recursos...)
2. Programas de marketing actuales de los competidores
3. Análisis de la existencia de oportunidades para nuevos competidores
4. Actuaciones esperadas por parte de la competencia

D. Entorno

1. Tendencias sociales, políticas, económicas y tecnológicas
2. Evaluación de las tendencias: oportunidades o problemas

E. Etapa del ciclo de vida del producto

1. Etapa de ciclo de vida de la categoría de producto
2. Características del mercado en que se basa la clasificación del producto en una fase del ciclo

F. Estructura de costes de la empresa

1. Coste marginal de incrementar la producción y composición del mismo

G. Capacidades de la empresa

1. Análisis de la existencia de capacidades para mantenerse en el negocio
2. Comparación de las capacidades respecto a las de la competencia

H. Recursos financieros de la empresa

1. Existencia de recursos para apoyar un programa de marketing efectivo
2. Procedencia de los fondos y momento del tiempo en que estarán disponibles

I. Estructura de la distribución

1. Canales existentes y cómo acceder a ellos
2. Coste frente a beneficio de cada canal
3. Posibilidad de uso de diversos canales simultáneamente
4. Naturaleza y grado de competencia inter e intra canales
5. Tendencias futuras de la estructura de canales
6. Exigencias de los distintos canales respecto a promoción y márgenes
7. Análisis del beneficio para el canal de distribuir el producto de la empresa

Fuente: Batra, Myers y Aaker (1996)

Por otro lado, el grado de notoriedad actual de la marca, cuotas actuales de mercado, naturaleza de la competencia o ciclo de vida del producto deben ser determinantes en decisiones como la fijación del presupuesto publicitario o elección de la estrategia creativa más adecuada. En este sentido, Evans (1990) enfatiza la necesidad, bajo la perspectiva de un “pensamiento estratégico” en la planificación de medios, de que los planes profundicen en el análisis de la influencia que la publicidad de la competencia ejerce en un mercado en el que todos compiten por el mismo público objetivo. Bajo esta misma perspectiva, Bell (1982), desarrolla distintas alternativas publicitarias para cada uno de los elementos de la cartera de productos de la empresa, según en qué celda de la matriz de crecimiento / cuota de mercado del Boston Consulting Group se encuentren.

También en la misma línea, Carson y Rice (1989) subrayan la importancia del análisis del entorno, desarrollando un sistema experto que incorpora el uso de la información acerca de la competencia, en el proceso de toma de decisiones publicitarias. La mayoría de los factores que forman el marco de decisión de este sistema, están estrechamente conectados con los del cuadro 1.1: Intensidad de la actividad publicitaria, etapa del ciclo de vida en la que se encuentra el producto, otras actividades de marketing de los competidores, posición competitiva de la empresa, recursos disponibles y cultura corporativa.

Otros autores, como Coe (1981), demuestran que el incremento en la presión competitiva del entorno, requiere que la empresa modifique radicalmente su filosofía publicitaria, la estructura organizativa que marca las relaciones entre el departamento de marketing y los responsables de comunicación, los objetivos que se atribuyen a la publicidad, sus métodos de fijación de los presupuestos publicitarios e incluso se ha de producir un cambio en sus estrategias creativas. Denota, con todo ello, que las implicaciones son tan importantes que este análisis de situación no puede obviarse en modo alguno.

Finalmente, Barban, Cristol y Kopec (1994), establecen una serie de recomendaciones sobre qué características deben cumplir los planes de medios en función de la etapa del ciclo de vida en que se encuentra el producto, derivada del análisis situacional que estamos comentando. Estos autores consideran que en la etapa de introducción, la misión fundamental de la publicidad es crear notoriedad, siendo en esta etapa más importante lograr cobertura que no frecuencia. En etapas de madurez, la publicidad suele encaminarse a mantener la notoriedad lograda, recomendando estos autores sacrificar la cobertura en aras a una mayor frecuencia.

Si importante es, como hemos tratado de destacar, que el responsable de comunicación revise la información proporcionada por el análisis situacional previo, no lo es menos que, al traducirse las conclusiones obtenidas en objetivos concretos del plan de marketing, éstos sean tomados como punto principal de referencia, dado que se descompondrán en otros más precisos y delimitados de carácter comunicacional que deben inspirar los planes publicitarios, de ventas, relaciones públicas, promoción de ventas y *publicity* (Martín Armario, 1993). La creciente dificultad de integrar estos objetivos adecuadamente en una campaña coordinada ha sido puesta de manifiesto por Wang y Petrison (1991). Estos autores señalan que la menguante fidelidad a la marca, la continua fragmentación de las audiencias de los medios tradicionales, la gran cantidad de categorías de productos que están alcanzando su fase de madurez y la notoria orientación hacia el corto plazo de los consumidores, están provocando un creciente traslado de recursos desde la publicidad hacia herramientas alternativas como la promoción de ventas o el marketing directo.

La necesidad de integración de las diferentes herramientas y su subordinación y coordinación con los objetivos generales del marketing que se ha señalado, ha sido recogida bajo el término de *Comunicación Integrada en el Marketing* (IMC)¹

¹ Dada la extensión de su uso, se ha optado por mantener las siglas correspondientes al término en inglés IMC: *Integrated Marketing Communications*

inicialmente presentado por Caywood, Schultz y Wang (1991), y que ha dado lugar a una gran cadena de libros y publicaciones entre los que destacan los trabajos de Burnett y Moriarty (1998), Ogden (1998) o Sirgy (1998). Caywood, Schultz y Wang (1991), lo definen como un concepto que reconoce el valor añadido de un plan global que valore el papel estratégico de las distintas herramientas de comunicación y las combine para lograr la claridad, consistencia y el máximo impacto comunicativo. La importancia de este tipo de actuaciones es reconocida por las agencias de publicidad estadounidenses, la mayoría de las cuales (Schultz y Kitchen, 1997) declaran dedicar al menos el 25% del tiempo empleado en cada cliente a los programas de IMC. Con ello se logra, siempre según el mencionado estudio, incrementar el impacto de las herramientas de comunicación, lograr que la creatividad sea más efectiva y conferir mayor consistencia a la política de comunicación.

McArthur y Griffin (1997) en un estudio sobre 121 de los mil primeros anunciantes norteamericanos, obtuvieron que, la IMC es uno de los temas de mayor relevancia en sus políticas de marketing, dejando mayoritariamente su diseño y aplicación, en manos de las agencias de publicidad de servicios plenos. Las principales actividades comunicacionales objeto de la coordinación son: la creatividad y los medios, la promoción de ventas, el marketing directo, la *publicity* y las relaciones públicas.

Las consecuencias para los planificadores de la implantación de estas políticas de gestión de la comunicación, se derivan del hecho de que el trabajo en equipo es uno de sus principales activos (Cook, 1997). La implantación de la IMC, como hemos indicado, aumenta la necesidad de coordinación y está cambiando el rol que deben jugar los departamentos de planificación de medios en las agencias. La principal consecuencia de este requerimiento, ha de ser que los responsables de la planificación de medios se vean envueltos mucho antes de lo que lo venían haciendo tradicionalmente en el proceso publicitario, contribuyendo a asesorar en la decisión de la combinación de medios más adecuada para los objetivos de comunicación que

se hayan establecido (Wang y Petrison, 1991). Su labor se convierte, por tanto, en más importante y compleja al exigírseles un conocimiento profundo de una variedad creciente de medios que deben integrar con los objetivos establecidos.

En esta misma línea, González y Carrero (1997) indican que el planificador ha de ser consciente que los problemas de medios son problemas de marketing y que, consecuentemente, debe erradicarse la visión de la planificación como una disciplina aislada, para asumir que los medios no son sino una herramienta más para coadyuvar al logro de los objetivos de marketing, armonizando sus acciones con la distribución, el precio, los objetivos de venta, los hábitos y actitudes de los consumidores, las tensiones del mercado y la actividad de la competencia. Esta preocupación se ha trasladado, incluso, a algunos creadores de *software* para la planificación de medios (Martins y Coons, 1991) quienes señalan que bajo sus programas subyace la filosofía de crear *MDMP's*, eso es, planes de medios guiados por el marketing². Por lo tanto, el perfil del planificador ha cambiado, ya no basta (González Lobo, 1994) con que dominen el mundo de los medios de comunicación y sepan manejar las estadísticas, ajustar los presupuestos, utilizar las fuentes de datos e incluso negociar la compra de medios con habilidad, se les exige además que sepan interpretar los problemas de marketing y sean capaces de analizar adecuadamente las necesidades de comunicación derivadas de los mismos.

La evolución en la estructura de las agencias de publicidad no está, sin embargo, favoreciendo este proceso. Si se exceptúan las grandes agencias, la mayor parte de las planificaciones se realizan en las centrales de medios, más alejadas del conocimiento del problema de marketing que ha de resolver la campaña de publicidad e, incluso, de sus características creativas. El peligro, según González y Carrero (1997; pág. 17), radica en que “se olviden los aspectos de marketing y su influencia en la selección final de soportes que son, en última instancia, los que deben presidir todas las actuaciones”. Sin embargo, este proceso de especialización

² Traducción propia del original *Marketing Driven Media Plans*.

funcional, resulta coincidente con la evolución y especialización que se viene produciendo en las modernas economías.

Siguiendo el proceso secuencial definido por Ray (1982) se llega en este punto a la toma de una de las decisiones que más importancia va a tener sobre la planificación de medios: el establecimiento del presupuesto tentativo para las herramientas de comunicación y su reparto entre las mismas. No se trata aquí tanto de realizar una revisión profunda de la literatura sobre el tema, como de señalar las implicaciones que para la planificación de medios se derivan de esta fase del proceso. En todo caso, los trabajos de Banks (1973), o más recientemente los de Piercy (1986, 1987), ofrecen una buena visión de la influencia que la complejidad del proceso organizacional ejerce sobre la fijación del presupuesto publicitario. Trabajos como los de McMeekin (1988) o Zufryden (1989) pretenden establecer criterios normativos para la determinación del presupuesto, aunque la mayor parte de la literatura³ se centra en analizar los métodos que en la práctica aplican las empresas para fijar el gasto en publicidad. Así el caso norteamericano es analizado en Patti y Blasko (1981), Lancaster y Stern (1983) y Blasko y Patti (1984). Los trabajos de Hooley y Lynch (1985); Lynch y Hooley (1987), Piercy (1987), Keown, Synodinos y Jacobs (1989) y Lynch y Hooley, (1989; 1990), están centrados en Europa, mayoritariamente en Gran Bretaña, el de Hung y West (1991) analiza las diferencias entre Estados Unidos, Canadá y Gran Bretaña y, finalmente, Ortega (1986) y Díez de Castro, Barreiro, Aguirre e Iturbe (1989) revisan el caso español.

La importancia para la planificación de medios del método que se utilice en la determinación del presupuesto, se deriva de que, según cuál sea éste, el planificador se encontrará ante una variable instrumental, sobre la cual puede tomar decisiones o, por contra, será una restricción no sujeta a revisión. Diversos estudios, tómesese como ejemplo Lynch y Hooley (1990) o Bigné (1995), demuestran que, con

³ Una buena revisión de la misma puede encontrarse en Bigné (1991a) y en Sissors y Surmanek (1982)

el tiempo, los métodos de fijación del presupuesto publicitario, han evolucionado desde la preeminencia de aquellos con un componente fundamentalmente subjetivo: método arbitrario, porcentaje sobre ventas, paridad competitiva, entre otros, hacia un desarrollo en el uso del presupuesto basado en los objetivos y tareas.

Esta evolución hacia la utilización de métodos donde el planificador puede jugar un papel más significativo no ha sido tan marcada en el caso español. Así, el estudio realizado entre las agencias de publicidad por Bigné (1989), señala que el presupuesto viene, casi siempre, fijado por el anunciante. A esta misma conclusión llegan desde su experiencia profesional González y Carrero (1997; p. 283) cuando afirman que aunque "... el determinar el presupuesto de medios puede ser tarea del planificador, para lo cual se basará en los objetivos de mercado y en la actividad publicitaria de las marcas competidoras [...] esto no suele ser frecuente, sino que lo normal es que el planificador reciba la cifra a invertir como un dato que él no puede cambiar". La principal consecuencia que se deriva de esta realidad para el planificador, es que al actuar el presupuesto como una restricción inamovible, éste dedique más tiempo a especializarse en la "compra de medios" que a cuidar el resto de elementos que conforman la verdadera "planificación de medios" (Priemer; 1990).

Respecto a la fijación de los objetivos publicitarios, siguiente etapa en el esquema de Ray (1982) y recogido en la figura 1.1, del proceso de decisión publicitaria, no sería correcto afirmar que los objetivos que se atribuyan a una campaña publicitaria concreta, influyan sobre el plan de medios, más bien, y en buena lógica, deben ser su determinante fundamental (Wilson y Machleit, 1985). Aunque diversos autores señalan que fijar los objetivos publicitarios en términos de ventas es bastante habitual (Batra, Myers y Aaker, 1996; Arens, 1996 o Schultz, Martin y Brown, 1990), e incluso son muchos los que consideran que el objetivo final de la publicidad son las ventas (Hopkins, 1980; Preston, 1985), existe hoy en día un acuerdo bastante generalizado respecto a que los objetivos de una campaña no

pueden plantearse en términos de incremento en ventas (Ray, 1982; Batra, Myers y Aaker, 1996 ; Díez de Castro y Martín Armario, 1993; González Lobo, 1994) debido, fundamentalmente, a que siendo éstas consecuencia de la influencia simultánea de todos y cada uno de los elementos del *mix* de marketing, resultaría imposible aislar el efecto causal debido exclusivamente a la publicidad y valorar sus efectos retardados sobre las ventas (Arens, 1996; Farbey, 1994; Hague y Jackson, 1992). Este planteamiento ha sido sustituido por la atribución de objetivos de carácter comunicacional, entendiendo que la publicidad no puede vender (resultado comportamental) si previamente no ha logrado respuestas comunicacionales como: conseguir el recuerdo o reconocimiento de la marca, fomentar un cambio de las percepciones respecto al producto o alguno de sus atributos, buscar una modificación de las creencias sobre el producto o sus características, construir o reforzar la imagen corporativa o de línea de producto o recuerdo de uso de producto, entre otros (Batra, Myers y Aaker, 1996; Guiltinan y Gordon, 1994; Kotler, 1996; Ortega, 1997, 1987).

Percy y Rossiter (1980) distinguen dos niveles de respuesta del individuo a la publicidad en cada uno de los cuales pueden establecerse objetivos distintos que tienen diferentes implicaciones para la planificación de medios. El primero de estos niveles es el del anuncio individual y la primera respuesta que el anuncio debe lograr, y por ende su primer objetivo, debe ser la exposición al mismo. Las implicaciones para la planificación de medios son obvias y, de hecho, uno de los elementos fundamentales de la misma pasa por ser capaz de estimar correctamente la cobertura de los planes, esto es, el número de individuos distintos efectivamente expuestos al anuncio. La segunda respuesta buscada en este nivel, es el adecuado procesamiento del mensaje, que implica la captación de la atención del individuo y la decodificación del mensaje en los términos previstos por el anunciante. Aunque esta respuesta parece afectar más a la fase creativa de la publicidad, no debe olvidarse que, según Craik y Lockhart (1972) la atención reside en la llamada “memoria a corto plazo” que tiene una capacidad limitada, por lo que el anuncio debe desplazar otro tipo de pensamientos del receptor. Lo que no logre la creatividad, deberá ser

conseguido por la reiteración en la exposición, entrando aquí en el concepto de frecuencia de exposición, cuya importancia en la planificación de medios es, de nuevo, hartamente evidente (Naples, 1979). Recientemente, Sánchez Franco (1998), ha revisado las distintas teorías sobre la adquisición de información publicitaria y, además de la ya señalada de Craik y Lockhart (1972), en la mayoría de ellas (Broadbent, 1958; Atkinson y Shiffrin, 1968; Shallice y Warrington, 1970, Anderson, 1983) se hipotetiza la existencia de almacenes de memoria a corto plazo, en la activación de los cuales, además de los aspectos creativos de la publicidad, la repetición debe jugar un papel determinante.

El segundo de los niveles al que, todavía según Percy y Rossiter (1980), pueden establecerse los objetivos publicitarios, es el nivel más amplio de comunicación del producto o servicio. En este nivel, el objetivo de la estrategia publicitaria puede centrarse en lograr *notoriedad* para el producto, crear o modificar *creencias* acerca del mismo (aspecto cognitivo), buscar determinadas respuestas de carácter *afectivo*, entendiendo por tal la evaluación emocional que hace el receptor de la publicidad de un producto o servicio y, finalmente, incidir sobre las respuestas de tipo *conativo*, es decir, los planes de actuación del receptor respecto a lo anunciado. Resulta evidente que las estrategias publicitarias, y por ello el plan de medios, serán muy distintas según (1) el producto tenga mayor o menor notoriedad para el público objetivo; (2), si teniendo esta notoriedad sus creencias acerca del producto son unas u otras; (3) si es necesario o no crear determinados sentimientos positivos hacia el producto; y, finalmente, (4) qué es lo que se espera que el receptor haga respecto al producto: probarlo, recomprar, recomendarlo, etc. Así, Percy y Rossiter (1980) señalan que, probablemente, si el objetivo es incrementar la notoriedad de una nueva marca, la acción inmediata es proponer planes de medios de elevada cobertura mientras que si, por el contrario, se pretende explicar la necesidad de la misma (aspecto cognitivo), lo recomendable sería maximizar la frecuencia de contactos.

En esta misma línea Rodríguez del Bosque, De la Ballina y Santos (1997), derivan también implicaciones para la planificación de medios, de determinados objetivos publicitarios, considerando que el logro de la cobertura es deseable cuando se lanzan al mercado productos nuevos o como apoyo a grandes promociones, mientras que el logro de niveles elevados de repetición, por el contrario, sería aconsejable para productos de compra frecuente y con débil fidelidad a la marca.

La influencia de los objetivos de una campaña sobre la planificación de medios no termina, sin embargo, en las variables de cobertura y frecuencia. Así, aunque la literatura a este respecto es escasa, algunos estudios llegan a establecer que determinados medios son más adecuados que otros para el logro de ciertos objetivos ofreciendo de esta modo una herramienta adicional al planificador. En este sentido Krugman (1972), Wright (1974) y Coffin (1975) postulan que los medios impresos favorecen, en relación a la radio o la televisión, el logro de respuestas de tipo cognitivo por parte del individuo.

Desde una perspectiva mucho más global y con pretensiones normativas, Bell (1982) relaciona la fijación de los objetivos publicitarios que hay que seguir en cada unidad de negocio, según la situación de cuota de mercado y tasa de crecimiento en que se encuentra cada una de ellas, de acuerdo con el análisis clásico de la matriz de negocios del Boston Consulting Group⁴. Así, ante productos “dilema”, esto es, con baja cuota de mercado, pero elevadas tasas de crecimiento, el objetivo de marketing debe ser conseguir rápidamente cuota de mercado. La publicidad se convierte, según este autor, en la principal herramienta para lograr reconocimiento de la marca e incitar a la prueba del producto, debiendo desembocar la fuerte inversión a realizar en planes de medios con elevadas coberturas y frecuencias de exposición, conclusión idéntica a la que llega Evans (1990) para este tipo de estrategia. Algunos autores han

⁴ Dado que es ésta una herramienta clásica y suficientemente conocida, no entraremos en su descripción. Una visión general de la misma puede encontrarse en Menguzzato y Renau (1991), mientras que sus implicaciones para el marketing se encuentran bien tratadas en Day (1977).

constatado (Hulbert, Lehmann y Winer, 1983) que en la práctica, las empresas asocian los incrementos en las inversiones publicitarias a las estrategias de búsqueda de crecimiento de la cuota de mercado.

Cuando el producto tiene elevadas cuotas de mercado y crecimiento, nos encontramos ante una “estrella” cuyas tendencias deben ser mantenidas atrayendo nuevos usuarios e incrementando el consumo entre los consumidores medios y los poco frecuentes. Se postula en este caso que la inversión en publicidad debe mantenerse dado que es una herramienta que sirve bien a esos objetivos y, por otro lado, Bell (1982) afirma que es generalmente aceptado entre las empresas, la táctica de igualar el porcentaje de presupuesto publicitario de cada producto con su cuota de mercado para lograr este objetivo. Evans (1990) argumenta que la mejor forma de que los planificadores de medios coadyuven a una política de mantenimiento de cuota, es desarrollar planes que aseguren la presencia continuada en los medios, pero con bajo nivel de *Gross Rating Points* para evitar inversiones no sostenibles por la empresa. Una buena forma de operativizar esta política es, para esta autora, que los planes aumenten la participación de medios de vida prolongada, como las revistas.

Cuando la “estrella” decae se convierte, según este análisis, en una “vaca”, con bajas tasas de crecimiento pero, todavía, una elevada cuota de mercado. El mercado comienza a estar más fragmentado y la publicidad deja de ser la mejor herramienta de comunicación. Se opta entonces por desviar parte del presupuesto hacia otras herramientas promocionales como la promoción de ventas. Finalmente, los “pesos muertos”, nunca han alcanzado una elevada cuota de mercado y su tasa de crecimiento es baja. La política a seguir con estos productos es, por lo general, eliminarlos antes de que absorban demasiados recursos. Sin embargo, dado que el mercado se ha vuelto excesivamente fraccionado, pueden encontrarse nichos de mercado donde existan consumidores con necesidades muy concretas que puedan ser atendidas. La publicidad basada en planes de medios no masivos, sino que utilizan soportes que se ajustan al perfil de estos clientes fieles, sin pretender llegar a los no

consumidores, son un elemento útil para mantener esta lealtad y asegurar que, en su última etapa, el producto puede aportar todavía beneficios a la empresa, sobre todo si los competidores comienzan a abandonar el mercado. La propuesta de Weilbacher (1990) de una mayor investigación de los estilos de vida de estos segmentos y sus hábitos de exposición a los medios, para conseguir planes más adecuados, es especialmente acertada en estos casos. Asimismo, Barban, Cristol y Kopec (1994) apuestan por reforzar en los planes de medios la frecuencia sobre el segmento en cuestión frente a la cobertura, como procedimiento más adecuado para establecerse en el nicho.

Siguiendo siempre el esquema planteado por Ray (1982) que recoge la figura 1.1 y puesta de manifiesto la importancia que tiene para una correcta planificación de medios, elementos tales como el análisis de situación, el establecimiento de los objetivos y el presupuesto, nos resta por abordar someramente, antes profundizar en el plan de medios y soportes, las implicaciones sobre éste del plan creativo. Para Sissors y Bumba (1996), las consideraciones creativas y del producto, determinan en muchas ocasiones la selección de los medios. Si un anuncio utiliza la demostración como estrategia creativa, difícilmente se podrá recurrir a otro medio que no sea la televisión. González Lobo (1994) indica que en el lanzamiento de un producto es fundamental apreciar su diseño y por ello es imperativo recurrir a un medio con imagen, el movimiento puede ser fundamental para productos que también se mueven (v.g. coches) y la calidad de la reproducción, para productos de prestigio en cuya imagen se quiere insistir (v.g. joyas, alta costura, productos de diseño, azulejos, etc.). Shimp (1993), por su parte, también señala a la televisión como un medio adecuado cuando se trata de realizar demostraciones del producto o cuando el objetivo fundamental es impactar al receptor. Las revistas destacan por su capacidad para mostrar la elegancia, belleza, prestigio y tradición. Los periódicos son particularmente adecuados cuando se incide sobre ofertas puntuales, mientras que la radio, encaja perfectamente cuando la creatividad del anuncio pretende que la imaginación del oyente juegue un papel importante. Una buena revisión de las

características cualitativas de los medios a las que estamos haciendo referencia para su selección publicitaria, se puede encontrar en Bigné (1990a) o en Greenburg (1972).

Se ha pretendido, en estas páginas iniciales, situar la planificación de medios en el marco de las decisiones comunicacionales del marketing y enfatizar, simultáneamente, la fuerte interrelación existente entre cada una de las etapas del proceso y la planificación de medios. En los siguientes apartados, tras destacar la creciente importancia que los cambios en el entorno, están otorgando a la esta disciplina, se abordará el concepto de planificación de medios, señalando el papel que juega la modelización en esta actividad. A continuación se revisará la evolución de la literatura concerniente a este tema, y se finalizará con un análisis del papel que juegan los modelos en la práctica de la planificación de medios.

1.2 LA CRECIENTE IMPORTANCIA DE LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS

En un reciente documento de trabajo del Departamento de Publicidad de la Universidad de Texas en Austin (1995; pág. 13), se señala que los profundos cambios que están ocurriendo en la industria de la publicidad en general y en el mundo de los medios en particular, están provocando que la planificación de medios juegue un papel cada vez más importante en este contexto:

“El número de nuevos medios, la explosión del número de soportes dentro de los medios tradicionales, la fragmentación de las audiencias, están haciendo que sea cada vez más difícil planificar la compra de medios. Es necesario emplear más tiempo en identificar las nuevas alternativas de medios y analizar la forma más eficiente de llegar al público objetivo. Debido a esta creciente necesidad de patrones analíticos, la planificación de medios se ha convertido en una labor de ejecutivos, cuando antes lo era de oficinistas.”

El objetivo de este epígrafe, es poner de manifiesto que esos cambios y otros que se señalarán, también se han producido en nuestro país, de tal forma que la creciente importancia del papel de la planificación de medios señalada con anterioridad, es una conclusión también válida en nuestro entorno.

El primero de estos hechos, es el *importante crecimiento de la inversión publicitaria*. Desde la comienzos de la década de los ochenta hasta nuestros días, la inversión publicitaria total, se ha multiplicado por nueve en pesetas corrientes (cuadro 1.2) y por 5'5 en pesetas constantes de 1986 (cuadro 1.3). Aunque este crecimiento no ha afectado de la misma forma a los medios convencionales: diarios, suplementos y dominicales, revistas, radio, cine, televisión y publicidad exterior, donde la inversión publicitaria se ha multiplicado por 5'6 (3'4 en pesetas constantes), que a los medios no convencionales: marketing directo, telemarketing, publicidad en el punto de venta, ferias y exposiciones, patrocinio, entre otros, los cuales han visto multiplicada la inversión realizada en ellos más de 22 veces (13 en términos reales), no es menos cierto que ante la magnitud del crecimiento de las cifras de inversión,

esté aumentando la preocupación de las empresas porque tal volumen de inversión se esté asignando eficientemente (Kirkham, 1984).

Cuadro 1.2. Inversión publicitaria en España 1981-1995
Millardos de pesetas corrientes

Año	Inversión real estimada			Índice			Tasa anual de crecimiento			% Conv sobre no Conv
	Total	Medios		Total	Medios		Total	Medios		
		Conven-cionales	Otros		Conven-cionales	Otros		Conven-cionales	Otros	
1.981	132,2	103,8	28,4	100	100	100	-	-	-	78,5
1.982	181,7	142,1	39,6	137	137	139	37,4	36,9	39,4	78,2
1.983	215,0	167,0	48,0	163	161	169	18,3	17,5	21,2	77,7
1.984	260,0	199,0	61,0	197	192	215	20,9	19,2	27,1	76,5
1.985	315,0	240,0	75,0	238	231	264	21,2	20,6	23,0	76,2
1.986	417,0	317,0	100,0	315	305	352	32,4	32,1	33,3	76,0
1.987	540,0	413,0	127,0	408	398	447	29,5	30,3	27,0	76,5
1.988	685,0	521,0	164,0	518	502	577	26,9	26,2	29,1	76,1
1.989	860,0	647,0	213,0	651	623	750	25,5	24,2	29,9	75,2
1.990	1.055,0	780,0	275,0	798	751	968	22,7	20,6	29,1	73,9
1.991	1.177,0	782,6	394,0	890	754	1.389	11,6	0,3	43,4	66,5
1.992	1.265,0	746,9	518,0	957	720	1.824	7,5	-4,6	31,4	59,0
1.993	1.157,0	597,1	560,0	875	575	1.972	-8,5	-20,1	8,1	51,6
1.994	1.168,0	600,9	567,0	884	579	1.997	0,9	0,6	1,2	51,4
1.995	1.215,0	588,2	627,0	919	567	2.208	4,0	-2,1	10,6	48,4

Fuente: 1981-1990: J. Walter Thompson (1991); 1991-1995: Infoadex (1996) ;
Elaboración propia

Cuadro 1.3. Inversión publicitaria en España 1981-1995.
Millardos de pesetas constantes de 1986

Año	Inversión real estimada			Índice			Tasa anual de crecimiento			% Conv sobre no Conv
	Total	Medios		Total	Medios		Total	Medios		
		Conven- cionales	Otros		Conven- cionales	Otros		Conven- cionales	Otros	
1.981	224,8	176,5	48,3	100	100	100	-	-	-	78,5
1.982	271,1	212,0	59,1	205	204	208	20,6	20,1	22,4	78,2
1.983	287,0	223,0	64,1	217	215	226	5,9	5,2	8,5	77,7
1.984	311,0	238,0	73,0	235	229	257	8,4	6,8	13,9	76,5
1.985	349,8	266,5	83,3	265	257	293	12,5	12,0	14,2	76,2
1.986	417,0	317,0	100,0	315	305	352	19,2	18,9	20,1	76,0
1.987	510,2	390,2	120,0	386	376	422	22,3	23,1	20,0	76,5
1.988	612,5	465,9	146,7	463	449	516	20,1	19,4	22,2	76,1
1.989	718,1	540,2	177,9	543	520	626	17,2	16,0	21,3	75,2
1.990	820,9	606,9	214,0	621	585	753	14,3	12,3	20,3	73,9
1.991	855,1	568,5	286,5	647	548	1.009	4,3	-6,3	33,9	66,5
1.992	859,9	507,7	352,2	650	489	1.240	0,6	-10,7	22,9	59,0
1.993	753,9	389,0	364,9	570	375	1.285	-12,3	-23,4	3,6	51,6
1.994	731,8	376,5	355,3	554	363	1.251	-2,9	-3,2	-2,6	51,4
1.995	726,4	351,6	374,8	549	339	1.320	-0,7	-6,6	5,5	48,4

Fuente: Elaboración propia

Algunos autores como Abrams (1991) exigen reacciones en el mundo de la planificación de medios ante esta responsabilidad, afirmando que mientras que los investigadores no logren mejoras en los procedimientos de medición de audiencias y las agencias y los medios las incorporen, la planificación y compra de medios acertada, será más una fantasía que una realidad. En un sentido similar, Sylvester (1991) afirma que los anunciantes exigen, cada día más, pruebas de que sus inversiones en medios están funcionando eficazmente. Esto obliga a las agencias, y por ende a los planificadores, a prestar más atención al papel de la publicidad en la generación de una respuesta en ventas, lo cual, si hace algún tiempo era de difícil operativización, hoy en día es factible gracias, según esta autora, a los sistemas de datos de fuente única. La importancia para el planificador de recurrir al uso de estas herramientas ha sido también expuesta por Sperling (1996), señalando que le permitiría, entre otras cosas, analizar la eficacia prevista de distintas combinaciones de soportes o diferentes repartos temporales de las inserciones.

Otro factor que ha ayudado a añadir complejidad al panorama de los medios en España y, consecuentemente, a la labor del planificador de medios, es el continuo *crecimiento de la oferta de medios y de soportes dentro cada medio*. Bigné (1991c) ya adelantaba que la televisión por satélite y la televisión por cable, iban a provocar un aumento en la oferta de canales muy especializados, con la consiguiente disminución, fragmentación y especialización de las audiencias que iba a tener que desembocar, necesariamente, en mejoras en la medición de la audiencia de los programas y los anuncios en dos dimensiones: exactitud y rapidez, procediendo parte de esa mejora de la realización de estudios cualitativos para analizar cómo se ve la televisión.

Sobre este mismo tema, Moix y Medinaveitia (1996) también coinciden en la necesidad de adaptar los métodos de medición de audiencias a estos nuevos medios. Aunque consideran que el Estudio General de Medios y el Audímetro de Sofres A.M., pueden ser válidos para conocer las audiencias globales de un canal de satélite, serán insuficientes cuando se pretenda medir con una mayor precisión el reparto entre los distintos programas. En este caso, apuntan, sería necesario diseñar un panel de hogares que tengan acceso a estas programaciones (hogares con antena parabólica particular o colectiva, conectados a redes de cable o televisiones locales que redifundan la señal del satélite) y recurrir a una ampliación de los audímetros o, en una primera fase más modesta, al diario de escucha.

Un medio que, como muestra el cuadro 1.4, está irrumpiendo con fuerza en el panorama mediático y que tiene importantes implicaciones para la planificación de medios es Internet. Según Leckenby y Hong (1996), la publicidad en Internet permite a las empresas llegar a poblaciones objetivo tremendamente segmentadas, así como abrir líneas de comunicación instantáneas con clientes reales o potenciales. Estos autores consideran que, sin embargo, es necesario redefinir y adaptar a la especificidad de este medio, la mayoría de conceptos básicos de la planificación de

medios. Así, un *hit* en una determinada página *web*, no es directamente equivalente a un contacto publicitario como tradicionalmente se entiende, en cuanto que no representan audiencia real: dos *hits* pueden ser dos personas distintas que entran en una página o una que lo hace dos veces, asimismo, como muchas direcciones contienen varias páginas, en la medida en que el navegante va pasando por ellas, va acumulando *hits* pese a ser el mismo individuo. En estudios posteriores, estos mismos autores (Leckenby y Hong, 1997), han demostrado que tampoco los modelos de exposición tradicionales pueden ser utilizados para la medición de cobertura y frecuencias en Internet sin adaptación previa recomendando, incluso, que lo más adecuado sería el desarrollo de modelos específicos.

Cuadro 1.4. Evolución de Internet en España

	Feb. / Marzo 1996	Abril / Mayo 1996	Oct. / Nov. 1996	Feb. / Marzo 1997	Abril / Mayo 1997	Oct. / Nov. 1997	Crecimiento Oct - Nov 97 / Oct - Nov 96
Usan Ordenador	6.208 <i>18,4%</i>	6.543 <i>19,4%</i>	6.946 <i>20,6%</i>	7.215 <i>21,2%</i>	7.429 <i>21,9%</i>	7.477 <i>22,0%</i>	7,6%
Usan ordenador habitualmente	4.179 <i>12,4%</i>	4.473 <i>13,2%</i>	4.696 <i>13,9%</i>	4.947 <i>14,6%</i>	5.014 <i>14,8%</i>	5.355 <i>15,8%</i>	14%
Tienen acceso a Internet	487 <i>1,4%</i>	684 <i>2,0%</i>	802 <i>2,4%</i>	1.073 <i>3,2%</i>	1.342 <i>3,9%</i>	1.455 <i>4,3%</i>	81,4%
Usaron Internet en el último mes	242 <i>0,7%</i>	277 <i>0,8%</i>	526 <i>1,6%</i>	765 <i>2,3%</i>	919 <i>2,7%</i>	1.110 <i>3,3%</i>	111%

Las cifras en negrita son miles de individuos.

Las cifras en cursiva, representan el porcentaje de penetración sobre la población de 14 años o más.

Fuente: AIMC

Pero, en el caso concreto de España, más que la aparición de nuevos medios, el factor que más ha contribuido a la turbulencia del panorama de los medios, ha sido el crecimiento de soportes dentro de uno de ellos: el medio televisión. Domínguez (1993) y González (1993), señalan que con la aparición de las televisiones privadas en 1990, la atmósfera publicitaria se transformó considerablemente, con una serie de consecuencias que afectan de una manera importante a la planificación de medios: fragmentación de la audiencia, políticas comerciales agresivas de venta de espacios, e incremento de la saturación publicitaria que merecen ser analizadas con detenimiento.

Como muestra el cuadro 1.5, en 1987, la oferta de televisión estaba concentrada en torno a las dos cadenas de Televisión Española. En 1997, el panorama es radicalmente distinto, dado que las audiencias se reparten casi equitativamente entre las tres cadenas abiertas de ámbito nacional: TVE-1, Antena 3 y Tele 5, pero donde, además, existen un importante número de televisiones autonómicas con audiencia significativas, hecho que además de ampliar la oferta, la regionaliza (Yarza, 1989).

Cuadro 1.5. Evolución de la audiencia acumulada diaria de las cadenas de televisión.

Porcentaje del total de la población mayor de 14 años

CADENA	1.987	1.988	1.989	1.990	1.991	1.992	1.993	1.994	1.995	1.996	1.997
TVE-1	82,6	81,4	78,4	71,0	69,9	60,1	56,6	51,2	52,1	50,0	47,0
TVE-2	26,0	30,6	33,7	34,5	22,9	23,3	17,7	18,5	19,0	18,3	18,1
ANTENA 3	-	-	-	9,1	17,4	24,7	36,0	46,0	47,2	46,9	43,0
TELE 5	-	-	-	14,5	27,3	35,7	37,5	34,1	33,9	38,6	40,8
CANAL PLUS	-	-	-	1,0	2,3	4,6	5,2	5,7	7,5	7,4	8,0
TV3	8,8	8,0	7,5	7,1	7,3	7,2	6,7	7,6	8,1	8,4	8,3
CANAL 33	-	-	0,8	0,9	1,6	2,3	2,4	2,1	1,6	2,8	2,8
ETB-1	0,6	0,3	0,5	0,7	0,7	0,7	0,6	0,6	0,8	0,8	1,0
ETB-2	0,6	0,8	1,2	1,5	1,1	1,3	1,3	1,8	2,3	2,3	2,2
TVG	1,5	2,1	3,0	2,5	2,0	2,1	2,0	2,2	2,5	2,5	2,5
TELE MADRID	-	-	3,2	3,1	3,4	4,6	5,0	4,4	5,3	5,6	5,1
CANAL SUR	-	-	6,8	8,2	7,7	7,5	7,5	6,5	6,6	6,9	7,4
CANAL 9	-	-	3,2	4,1	4,7	4,6	4,2	3,7	3,9	4,2	4,9
TV LOCAL	-	-	-	-	-	-	-	-	-	-	1,6
RESTO TV'S	-	-	-	-	-	-	-	-	-	-	1,9

Fuente: AIMC

Como señalan González y Carrero (1997), este incremento en la oferta ha desembocado en la modificación de los hábitos de comportamiento de las audiencias. Cuando solamente existían dos canales, la población actuaba según unos hábitos establecidos a lo largo de los años, que le llevaban a conectar siempre el primer canal y sólo cambiaba cuando se sentía atraído por algún programa en concreto de la

segunda cadena. Con la aparición de las cadenas privadas, los hábitos han cambiado: por una parte, se han *fragmentado las audiencias* (González, 1993), ya no es posible llegar, como ocurría en los setenta, a catorce millones de individuos con un sólo anuncio; lo habitual es que los espacios con más éxito tengan en torno a cinco millones de espectadores. Para Katz (1990) y Weilbacher (1990), la única solución que puede aportar la planificación de medios a este problema, reside en analizar los hábitos de audiencia de los individuos, prestando más atención a cómo utilizan los medios, en lugar de centrarlo todo en la segmentación basada en las clásicas variables demográficas y sicográficas, cuya eficacia es cada día más cuestionada (Block y Brezen, 1990).

Por otra parte, como consecuencia de estos cambios, el planificador se encuentra con que, en la búsqueda de un público más numeroso, las cadenas recurren a la *contraprogramación* (Domínguez, 1993), cambiando sin aviso unos programas por otros similares a los que están obteniendo éxito en otras cadenas, sin dar tiempo a reaccionar a los planificadores y creando malestar entre anunciantes y centrales ante una obligatoria planificación cambiante. La falta de un estilo permanente en el tipo de programación, produce una enorme *ausencia de fidelidad* por parte de la audiencia, que salta de cadena en cadena apenas desciende, aunque sea momentáneamente, el interés del programa que se está viendo (González y Carrero, 1997) provocando graves problemas en la eficacia de la publicidad y dificultando el estudio de las audiencias mediante el sistema de entrevista, donde el hábito juega un papel central.

La última gran consecuencia para la planificación de medios, del incremento de soportes en el medio televisión que supuso la creación de las cadenas privadas en 1990, es la enorme *agresividad comercial* en la venta de espacios publicitarios que ocasionó. El cuadro 1.6, muestra la diferencia entre la inversión publicitaria a precios de tarifa y la inversión real estimada para un año intermedio en el periodo que va desde la aparición de las cadenas privadas hasta nuestros días: 1994. Como puede

comprobarse, esta política es general en todas las cadenas y especialmente acentuada en las privadas abiertas: Tele 5 y Antena 3, derivándose de la misma, un descenso continuado en el coste por minuto de la publicidad en televisión, como se aprecia en la figura 1.2.



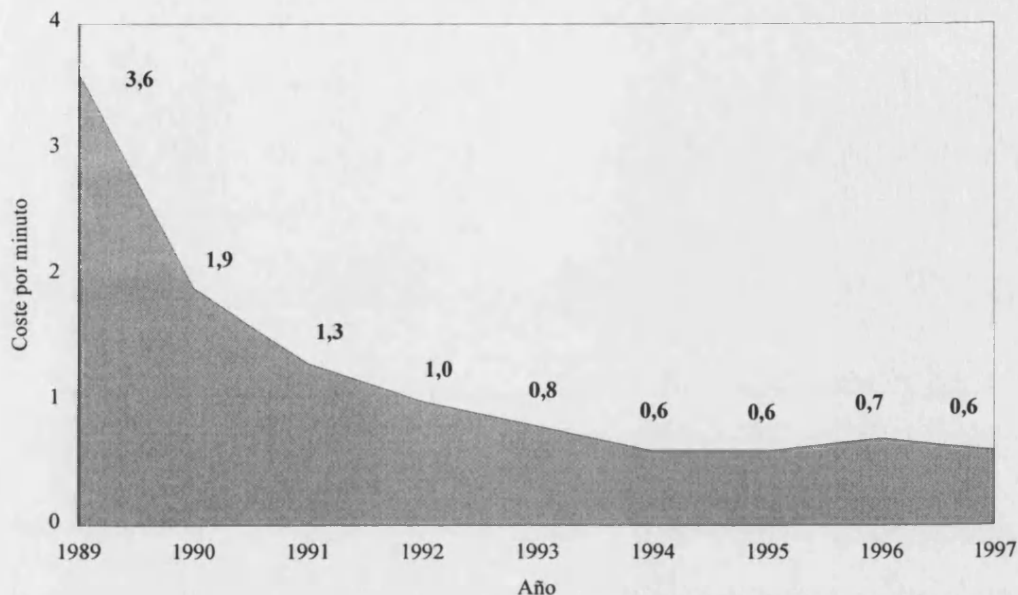
Cuadro 1.6. Política de descuentos en el medio televisión.
Año 1994. Datos en miles de pesetas.

Soporte	Inversión publicitaria		
	A precios de tarifa	Real Estimada	% de descuento
TVE 1	169.646,7	51.403	69,7
La 2	44.819,8	13.580	69,7
Tele 5	302.279,3	41.412	86,3
Antena 3	315.286,9	63.057	80,0
Canal Plus	5.592,9	1.996	64,3
Autonómicas	128.597,7	29.379	77,2
Circuitos regionales	26.824,7	8.128	69,7
TVE en Canarias	6.900,1	2.090	69,7
Total Medio TV	999.948,1	211.045	78,9

Fuente: Infoadex (1995) y elaboración propia

Las consecuencias para la planificación de medios del descenso de tarifas y la aparición de los paquetes de contratación, puede centrarse en dos aspectos. Por un lado, el descenso del precio por minuto ha provocado un mayor número de anunciantes pero, lo que es más grave, un crecimiento del índice de saturación (Domínguez, 1993) y, en consecuencia, más rechazo, menos rendimiento y menos credibilidad de la publicidad (Castellanos, 1994).

Figura 1.2. Evolución del coste por minuto de la publicidad en televisión.
Millones de pesetas



Fuente: Media Planning (1997)

Por otro lado, la agresividad comercial ha provocado un importante cambio en las relaciones medio, agencia y anunciante (González y Carrero, 1997). Según estos autores, el mercado en España no es transparente, se desconocen los baremos de lo que se puede considerar una compra cara o barata, provocando una fuerte especulación y dificultando las negociaciones. Ello lleva a que la evaluación de los planes de medios haya comenzado a realizarse atendiendo a criterios exclusivamente económicos, en detrimento de valores cualitativos como puedan ser la afinidad entre el programa y el producto. González (1993), señala que ante el abaratamiento progresivo del medio, algunos anunciantes dejan de delegar la negociación de sus campañas en las agencias y centrales, asumiéndola de manera directa, rompiéndose la relación clásica del anunciante con su central o agencia y el medio. Se llega, incluso, a que el anunciante reparta su campaña entre distintas empresas: una que realiza la creatividad, otra la compra de medios, con las peligrosas consecuencias

que, para la necesaria coordinación entre marketing, creatividad y medios, pueden derivarse.

Mientras tanto, en otros medios, la evolución comentada del medio televisión, provoca reacciones ante la transformación rápida del mercado (Domínguez, 1993). En el medio prensa, ante el estancamiento de la inversión publicitaria en los mismos, se ha optado por una mayor difusión mediante promociones de regalos o nuevos suplementos. Los descuentos al comprador y las mencionadas promociones, provocan reducciones considerables de los márgenes que se intentan paliar con la pequeña publicidad (clasificada, local...) o con la potenciación de fórmulas publicitarias distintas a la inserción clásica (encartes, reportajes...). También es de destacar, dentro de estos cambios estructurales en el mundo de los medios, la política de regionalización de las principales cabeceras de diarios, que han optado por realizar ediciones específicas en muchas comunidades autónomas (El País) e incluso descendiendo a nivel provincial (El Mundo), con el objetivo de incrementar la difusión gracias al acercamiento temático a los públicos periféricos. Muchos diarios regionales han reaccionado, consecuentemente, con la comarcalización de sus ediciones, como puede ser el caso de Las Provincias y Levante en la Comunidad Valenciana.

En las revistas el estancamiento ha sido más importante, y se ha traducido en descuentos cada vez mayores y buscando nuevos anunciantes con una oferta que englobe varios soportes integrados en una editorial. Esta práctica viene favorecida por el fenómeno de concentración en grandes grupos multimedia que se ha producido en nuestro país (González y Carrero, 1997) donde, a modo de ejemplo, el grupo Correo controla o participa, además de diarios como El Correo Español, Diario Vasco, La Verdad, Sur, Ideal, Hoy o El Comercio, suplementos de los mismos como Suplemento Semanal o Suplemento TV, emisoras de radio o televisión por cable y, recientemente, ha entrado a formar parte del accionariado de Gestevisión (Tele 5).

En la radio, la incidencia de la oferta televisiva ha sido menor, siempre según Domínguez (1993). El medio se ha vuelto hacia la publicidad local como un método de mejorar sus cuentas de resultados, a la vez que se producía un importante movimiento de búsqueda de contratación de las llamadas estrellas de la radio que aseguraran, con su carisma, una audiencia y unos ingresos publicitarios considerables.

A modo de resumen, concluiremos este epígrafe como lo iniciamos: indicando que la turbulencia del mercado mediático ha complicado la labor de la planificación de medios hasta tal punto, que la disciplina en general, y la figura del planificador en particular, está viendo incrementada su importancia en el proceso publicitario. Se acentúa la necesidad de crear herramientas que ayuden a establecer los niveles adecuados de frecuencia y cobertura en un entorno cambiante (González y Carrero, 1997) y, por ello, la revisión de los mecanismos clásicos y la valoración de su capacidad de estimación, devienen en fundamentales. En este marco se trata de inscribir esta tesis doctoral, donde se evaluará la capacidad estimativa de la cobertura y la distribución de contactos de más de cuarenta modelos, jerarquizando los mismos en base al error cometido, ofreciendo así, criterios objetivos para la selección del procedimiento más adecuado para estimar las mencionadas variables, claves, como se verá, en el proceso de planificación de medios.

1.3 EL PAPEL DE LA COBERTURA Y LA FRECUENCIA DE CONTACTO EN LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS: NECESIDAD DE MODELIZACIÓN

Cuando en la búsqueda de una definición adecuada de la planificación de medios, se reflexiona sobre el problema que con la misma se pretende resolver, se llega, con Adams (1977) a la conclusión de que nos encontramos ante un problema de economía clásica. En palabras de Simon y Thiel (1980; pág. 25): unos recursos limitados, el presupuesto de publicidad que debe ser distribuido entre usos alternativos, los soportes, de tal manera que se maximice una función objetivo determinada, en suma, la asignación eficiente de recursos.

Bajo esta perspectiva, González Lobo (1994) define la planificación de medios como “una decisión acerca de los medios y soportes a utilizar en una campaña para cumplir los objetivos de la misma”. Señalando que, a partir de un presupuesto dado, una duración determinada de la campaña y del conocimiento de cuáles son los requerimientos creativos, el planificador ha de decidir qué medios van a utilizarse (decisiones inter-medios) y qué soportes serán elegidos dentro de cada uno de ellos (decisiones intra-medios), así como su programación temporal.

Martín Armario (1980) da un contenido análogo al concepto, pero introduce una precisión terminológica al distinguir entre “estrategia de medios” para referirse a la selección de los medios publicitarios más idóneos para la campaña y dejando el término “planificación de medios” para referirse a aquellas actividades consistentes en la selección de soportes concretos dentro de cada medio. Señala, sin embargo, que este último término es utilizado por otros autores para referirse a ambas fases: la estrategia y la planificación.

Ray (1982) amplía las definiciones anteriores al considerar que la planificación de medios debe determinar también la pauta temporal de las inserciones comportando, en definitiva, las siguientes decisiones:

- a. Qué medio o medios van a utilizarse en la campaña publicitaria.
- b. Qué soportes van a considerarse dentro de cada medio.
- c. Cuál va a ser el esquema temporal de las inserciones, es decir, en qué época del año debe desarrollarse y si debe realizarse con una intensidad uniforme o interesa saturar determinados meses.

Batra, Myers y Aaker (1996), incluyen en este concepto un elemento más propio del plan creativo, lo que denominan *opciones de medios*. Tras señalar que el plan de medios detalla la programación de medios a utilizar, especificando el modo en que va a distribuirse el presupuesto disponible, apuntan que deberá decidirse:

- a. Qué medios se van a utilizar, lo que denominan *clase* de medios.
- b. Qué soportes dentro de cada medio.
- c. *Opciones* de medios a emplear: que es una descripción detallada de las características del anuncio, considerando, entre otras, el tamaño (media página, página completa), el color, la duración y la colocación (interior, portada).
- d. Esquema temporal de las inserciones o *intensidad* de la programación.

Rust (1986) amplía el término hablando de *proceso* de planificación de medios, e indica que incluye cuatro componentes:

- a. Operativización de la estrategia global del marketing, señalando que, siendo la publicidad parte de éste, debe ser coherente con sus objetivos globales.
- b. Plan de medios, que incluye según este autor, la determinación del presupuesto y la elección de los medios.
- c. Evaluación de los soportes de cada medio elegido, en términos de alcance y frecuencia.
- d. Selección de los mejores soportes uniendo la variable constante a las variables señaladas en la fase anterior.

Sissors y Bumba (1996) consideran que la planificación de medios, consiste en una serie de decisiones tomadas para contestar a la pregunta general que hacen los anunciantes de: “¿Cuál es el mejor modo de hacer llegar los anuncios de mi producto o servicio a los clientes potenciales?”. Conscientes del carácter demasiado genérico de esta pregunta, desdoblan su enunciado en las siguientes cuestiones:

- a. A cuántos clientes potenciales necesito llegar.
- b. En qué medio debe colocarse el anuncio.
- c. Cuántas veces debe llegarse al cliente potencial.
- d.Cuál es el momento del tiempo en que el anuncio debe aparecer.
- e. En qué regiones del país debe verse el anuncio.
- f. Qué cantidad del presupuesto global debe invertirse en cada medio.

Barban, Cristol y Kopec (1994), definen la planificación de medios como el proceso de diseño de un curso de acción que muestra cómo debe utilizarse el tiempo y el espacio de la publicidad para lograr los objetivos de marketing. De nuevo, esta cuestión genérica es concretada en una serie de etapas más concretas del proceso:

- a. Análisis externo e interno de la empresa.
- b. Fijación de los objetivos.
- c. Definición del público objetivo.
- d. Selección de la combinación adecuada de medios.
- e. Consideraciones acerca del esquema temporal de las inserciones.

De una manera integradora y, a la vez, mucho más precisa, Lilien y Kotler (1983) acotan el problema de la definición de la planificación de medios en los siguientes términos: Dado un presupuesto para medios, un mensaje publicitario y la creatividad, un conjunto de medios alternativos y datos de audiencia y costes de estos medios, se trata de decidir: (1) qué soportes utilizar, (2) el número de inserciones en cada uno y el momento del tiempo en el que se realizan las mismas y (3) el tamaño y el color de los formatos publicitarios de forma tal que las decisiones maximicen el efecto, medido de alguna manera, del presupuesto publicitario.

Tomando como base las definiciones anteriores, entenderemos en este estudio por planificación de medios: el proceso de selección eficiente de medios y soportes que van a utilizarse en una campaña de publicidad, la distribución de los recursos entre ellos y la determinación del esquema temporal de las decisiones, siempre de acuerdo con los objetivos generales del plan de marketing y de manera que se maximice la difusión del mensaje entre la población objetivo. El cuadro 1.7 resume las fases integrantes de este proceso y las variables que han de considerarse en cada una de ellas.

Cuadro 1.7. Fases de la planificación de medios.

FASES	VARIABLES A CONSIDERAR	MODELOS
FASE PREVIA Consideración de los condicionantes impuestos por el plan de marketing	<ul style="list-style-type: none"> - Objetivos del plan de marketing - Objetivos del plan de publicidad - Montante del presupuesto - Aspectos creativos - Horizonte temporal de la planificación de la campaña - Público objetivo 	
FASE I Selección de los medios a utilizar	<ul style="list-style-type: none"> - Características del medio (presencia de imagen, color, sonido). - Adecuación del medio al producto. - Población objetivo. - Alcance del medio. - Disponibilidad. - Restricciones legales. - Utilización por la competencia. - Estacionalidad. 	
FASE II Valoración cuantitativa y cualitativa de los soportes y de combinaciones de los mismos dentro de cada medio	<ul style="list-style-type: none"> - Frecuencia de aparición. - Credibilidad y prestigio. - Audiencia. - Coste por soporte. - Perfil de la audiencia. 	<ul style="list-style-type: none"> - Modelos de exposición: cobertura y distribución de contactos. - Modelos de función de respuesta.
FASE III Elección de la combinación de soportes y del esquema temporal de las inserciones	<ul style="list-style-type: none"> - Cobertura y distribución de contactos de diversas combinaciones de soportes e indicadores derivados: GRP's OTS. - Indicadores de rentabilidad: CPM, Coste total. 	<ul style="list-style-type: none"> - Modelos de planificación de medios.

Fuente: Elaboración propia a partir de Rust (1986)

La presente tesis doctoral, se encuentra enmarcada en la valoración cuantitativa de los soportes de la fase II. Así, quedan fuera de su ámbito de estudio las decisiones inter-medios de la fase I y, pese a su relevancia (Gensch, 1970a), las consideraciones cualitativas de la valoración de los soportes, tales como sus características visuales y / o auditivas (Houston y Scott, 1984; Diamond, 1968) o el efecto fuente⁵ así como las decisiones implicadas en la fase III.⁶ No se trata de minusvalorar la importancia de los factores cualitativos en la selección de soportes pero, como señala Martínez Ramos (1992), cuando se trata de buscar indicadores que permitan discriminar entre los soportes y faculten al planificador para determinar qué soporte es mejor que otro, los más importantes son los criterios cuantitativos, fundamentalmente, de un lado, su audiencia y la frecuencia de contacto y de otro el coste. Este mismo planteamiento ya estaba implícito en el modelo que la *Advertising Research Foundation* (1961) elaboró para objetivar el proceso de comparación de los soportes y en el cual tres de los seis niveles en los que debe compararse los soportes: difusión del soporte, exposiciones al mismo y exposición al anuncio, guardan una íntima relación con las variables reseñadas. Aunque es cierto que se está señalando la necesidad de introducir modificaciones en este modelo (Banks, 1989; Chook, 1989; Sprague, 1989; Phelps, 1989; 1993; 1994) y algunos autores como Harvey, (1997) han propuesto un modelo revisado, no es menos cierto que la mayoría de los cambios se centran en etapas distintas a las señaladas.

Tampoco se hará referencia en esta tesis a consideraciones mucho más recientes sobre la planificación de medios que se centran en el bloque publicitario en el que aparecerá el anuncio. Nos estamos refiriendo a aspectos tales como la repetición del anuncio dentro del bloque, la duración del mismo y su posición dentro

⁵ Pueden consultarse las revisiones sobre el tema efectuadas por Percy y Rossiter (1980) Batra, Myers y Aaker (1996) o Bigné (1990a).

⁶ Los modelos tendentes a determinar las pautas temporales más adecuadas de inserción de los anuncios (Ehrenberg, 1974; Simon, 1982), junto con los principales modelos completos de planificación de medios (Aaker, 1968; Little y Lodish, 1969; Zufryden, 1981, 1982) se abordarán sólo tangencialmente en la revisión de la literatura sobre la evolución de la modelización en la planificación de medios.

del bloque. Trabajos como el de Bigné y Miquel Romero (1996) demuestran que el hecho de repetir un anuncio dentro de un bloque publicitario, aumenta la posibilidad de que el individuo lo recuerde aunque, por otro lado, algunos autores advierten acerca de los efectos negativos cuando se alcance un nivel de saturación (Cacioppo y Petty, 1979; Krugman, 1975; Rethans, Swasy y Marks, 1986). Respecto a la duración, muchos estudios han demostrado que un anuncio de corta duración siempre es menos eficaz que una misma versión del mismo anuncio, creativamente hablando, de mayor duración (Fabian, 1986; Lonning, 1988; Mord y Gilson, 1985; Patzer, 1991). Finalmente, en lo que se refiere a la posición dentro del bloque, Pieters y Bijmolt (1994) han encontrado que los anuncios situados en los extremos del bloque publicitario tienen mayor probabilidad de atención por parte del individuo. Recientemente Miquel Romero (1997) ha sometido a revisión, para el caso español, todas estas consideraciones. Sánchez Franco (1998), por su parte, también ha analizado la influencia variables como el tamaño del anuncio y el color sobre la eficacia de la publicidad, concluyendo que los tamaños grandes son siempre más adecuados cuando no existe familiaridad con lo anunciado y, respecto al color, la inclusión de éste siempre ejerce efectos positivos sobre la eficacia publicitaria en situaciones de baja motivación de los consumidores, mientras que el efecto en situaciones de elevada motivación es contradictorio.

La importancia que posee la determinación, con una fiabilidad suficiente, de la cobertura y la distribución de contactos de un plan de medios, descansa, además de lo ya expuesto, en las implicaciones que se derivan para la estimación de la eficacia de la publicidad. Leckenby y Kim (1992) señalan acertadamente que el mensaje publicitario y los soportes que lo transportan, interactúan para restar o sumar eficacia a la publicidad. Si bien es cierto que un mensaje mal diseñado no funcionará aunque se inserte en un plan de medios con elevada cobertura y frecuencia, también es obvio que el mejor mensaje publicitario carecerá de eficacia si es incapaz de llegar el número adecuado de veces a la suficiente cantidad de público objetivo.

Bajo la evidencia anteriormente expuesta, subyace el concepto de frecuencia efectiva, definido por Sissors y Bumba (1996) como la frecuencia, o rango de frecuencias, necesarios para que los anuncios sean efectivos en términos comunicacionales⁷. En palabras de Miles y Arnold (1990), es muy abundante la literatura que constata que, en un mercado competitivo, existe un intervalo de la frecuencia de contactos, en el cual incrementos en la frecuencia se traducen en incrementos, medido de distintas maneras, de la efectividad publicitaria (Naples, 1979; Sissors, 1986; Turk, 1987; Wenzel y Speetzen, 1987; Goldring, 1988). Estaríamos ante el tramo en que el anuncio tiene efecto positivo sobre el individuo: efecto *wear-in* (Pechmann y Stewart, 1988). Por otro lado, también se ha constatado (Appel, 1971; Greenberg y Suttoni, 1973; Ackoff y Emshoff, 1975; Craig, Sternthal y Leavitt, 1976; Calder y Sternthal, 1980; Burke y Edell, 1986; Schultz y Block, 1986; Cannon, 1987; Turk, 1987; Goldring, 1988), que cuando el nivel de frecuencia de contactos excede el extremo superior del intervalo mencionado, el anuncio ya no produce ningún efecto positivo adicional e, incluso, empieza a tener efectos negativos sobre su receptor: efecto *wear-out*⁸ (Pechmann y Stewart, 1988).

Son numerosos los estudios empíricos que soportan las afirmaciones anteriores. Ackoff y Emshoff (1975), encuentran que, bajo determinadas condiciones, incrementos en la frecuencia de exposición, estaban correlacionadas con descensos en las ventas. Craig, Sternthal y Leavitt (1976), centraron su estudio en la relación entre frecuencia y recuerdo del nombre de la marca, constando el efecto *wear-out* a niveles altos de frecuencia, aunque lo atribuyen a la falta de atención del individuo o a reacciones contra la publicidad. Calder y Sternthal (1980), utilizaron la teoría del procesamiento de la información para establecer el impacto de la frecuencia sobre el recuerdo de la marca. Llegaron a la conclusión de que el *wear-out* existe y que puede presentarse, incluso cuando se diseñan estrategias específicas para

⁷ Murray y Jenkins (1992) definen el concepto afín de *cobertura efectiva* y derivan su cálculo a partir de la cobertura tradicional de los planes de medios.

⁸ Dado lo extendido de su uso, hemos preferido mantener el término original en inglés a alternativas como efecto *desgaste* u otras..

mantener la atención. Aaker, Carman y Jacobson (1982), recurrieron al análisis econométrico para estudiar la relación entre publicidad y ventas para seis marcas de cereales y no encontraron relación significativa entre inversión en publicidad y nivel de ventas, atribuyendo este hecho a que la frecuencia de exposición ha alcanzado el nivel de saturación en esta industria. Zielske (1986) obtuvo que siempre existen rendimientos marginales decrecientes ante cada nueva exposición, concluyendo que la famosa proposición de Krugman (1972) de las tres exposiciones, y cualquier regla numérica sobre el nivel de exposición, debe ser considerada con suma prudencia. Blair (1988), finalmente, constató la existencia de ambos efectos, concluyendo que a corto plazo, los niveles elevados de frecuencia pueden ser la herramienta más adecuada en el logro de la respuesta deseada.

La importancia teórica del concepto de frecuencia efectiva, deviene en relevancia práctica cuando se analizan los resultados de diversos estudios sobre el uso de la frecuencia efectiva en la planificación de medios por parte de las agencias de publicidad estadounidenses. Leckenby y Kishi (1982a) concluyen que es utilizado en el 59% de los planes de medios, Lancaster, Pelati y Cho (1991) por el 37 %, Kreshel, Lancaster y Toomey (1985) en el 86% de los mismos y Turk (1988) en el 100% de los planes realizados.

Aunque como muy bien han señalado Sissors (1982) o Kreshel, Lancaster y Toomey (1985), la frecuencia media de un plan de medios — frecuencia media de exposición a un soporte — no coincide con el concepto de frecuencia que manejan los estudios sobre eficacia publicitaria (v.g. Naples, 1979) — frecuencia media de exposición al anuncio —, sí que hay numerosos estudios que tratan de vincular la exposición al soporte con algún indicador de eficacia publicitaria, como el de Zielske y Henry (1980) que, tras analizar los resultados de diecisiete estudios de seguimiento⁹ sobre recuerdo espontáneo de anuncios televisivos, obtiene una expresión que permite derivar la variación de este variable a partir de los *rating*

⁹ Traducción propia del término original *tracking*.

points logrados por el plan de medios. o bien, el trabajo de Kamin (1978) que le permite establecer que, un rango de frecuencias con el soporte entre 5 y 20 contactos con el soporte, aproximarían bien el valor de tres contactos con el anuncio establecido por Krugman (1972)¹⁰. En cualquier caso, el estudio y medida de la eficacia publicitaria, requiere del conocimiento de la exposición al soporte. En palabras de Leckenby y Kim (1992; pág. 101): “No puede haber avances en la planificación de medios, sin mejoras en la estimación de la exposición al soporte, dado que esta variable es la piedra angular para el desarrollo de cualquier aproximación a la frecuencia efectiva”. En este mismo sentido se manifiesta Metheringham (1977)¹¹ cuando afirma que “... lo más importante es ser capaz de calcular la distribución de frecuencias tanto de las exposiciones como de los impactos. Cualquier otra función se deriva de ésta principal y, si se resuelve este problema, el resto es solamente una cuestión de organización y método”. Así, todos los estudios sobre el procedimiento en que las agencias de publicidad introducen en sus planes el concepto de frecuencia efectiva, constatan que éstas optan por ponderar de alguna manera el valor de frecuencia media de exposición al soporte. Así, Kreshel, Lancaster y Toomey (1985) encuentran que el 52% de las agencias establecen algún tipo de ponderación, mientras que Lancaster, Kreshel y Harris (1986) obtienen un porcentaje del 30,8%, de las cuales el 67,9% opta por algún tipo de ponderación de carácter subjetivo para este fin, un 35,7% tienen establecida una fórmula para ello y un 25% cuentan con una tabla estándar de pesos establecida por su agencia (la pregunta era de respuesta múltiple). En España, Sánchez-Fabrés (1993) avanzaba un proyecto de la agencia Tiempo-BBDO en la línea de cualificar las cifras de frecuencia en el medio revistas. La excepción a esta necesidad de derivar de alguna manera la exposición al anuncio de la exposición al soporte, la constituye el medio televisión cuando se recurre a los datos del audímetro, dado que, al ser un

¹⁰ En general cualquier modelo de planificación de medios ha de incorporar en su función de respuesta un nexo entre las dos variables señaladas: exposición al soporte con exposición al anuncio (Gensch, 1969; Little y Lodish, 1969; Gensch, 1970a; Rust, 1985; Rust y Stout, 1989).

¹¹ Cita textual recogida por Leckenby y Boyd (1984; pág. 2).

estudio coincidental, se es capaz de determinar si el individuo estaba o no sentado delante del televisor en el momento en que se emitió un *spot* concreto.

Se ha puesto de manifiesto la utilidad del conocimiento para el planificador de la cobertura y distribución de contactos de un plan de medios. Asimismo es evidente, que sería deseable tener acceso a los datos poblacionales, es decir, conocer si todos y cada uno de los individuos de la población objetivo han resultado o no alcanzados por el anuncio y cuántas veces. Sin embargo hay dos razones que examinaremos a continuación y que lo hacen inviable (Leckenby y Boyd, 1984a).

En primer lugar, el coste, sea cual fuere el procedimiento de recogida de información del estudio de investigación de audiencias, sería prohibitivo dado que el número de posibles combinaciones de soportes, aún dentro de un mismo medio, es prácticamente ilimitado, lo que se traduce en una complejidad de cálculo insoslayable que desemboca en la necesidad de formalizar los procesos mediante modelos matemáticos susceptibles de ser informatizados. En palabras de Martínez Ramos (1992; pág. 51) resulta imposible intentar obtener los valores de cobertura y distribución de contactos “sin más útiles de trabajo que el lápiz, el papel y los datos de un estudio de investigación de audiencias” afirmación que trataremos de ilustrar inmediatamente. Surgen así los *modelos de exposición* objeto de esta tesis doctoral.

La complejidad señalada deriva de dos aspectos fundamentales. De un lado el número de combinaciones de soportes incluidos en cualquier plan de medios, cada una de las cuales debe ser evaluada en términos de cobertura y frecuencia para elegir la más adecuada y, de otro lado, las necesidades de información que se derivan de los fenómenos de acumulación y duplicación de las audiencias.

Para ilustrar el primero de los casos, siguiendo a Martínez Ramos (1992), consideremos un plan de medios formado por N soportes, en cada uno de los cuales

pueden realizarse m inserciones. Considerando una única inserción en cada uno de ellos, el número total de combinaciones posibles es:

$$\binom{N}{1} + \binom{N}{2} + \binom{N}{3} + \dots + \binom{N}{N}$$

Incorporando ahora el número de inserciones, cada pareja de soportes puede combinarse de $m \times m$ maneras, cada trío de $m \times m \times m$ maneras y así sucesivamente. Generalizando, el número total de combinaciones es:

$$\binom{N}{0}m^0 + \binom{N}{1}m^1 + \binom{N}{2}m^2 + \dots + \binom{N}{N}m^N$$

que al ser un binomio de Newton es igual a:

$$(m + 1)^N$$

Suponiendo un plan de medios sencillo, formado por 10 soportes en cada uno de los cuales se realizan 5 inserciones, basta sustituir en las expresiones anteriores, para obtener que las combinaciones que habría que valorar en términos de cobertura y frecuencia, serían 60.466.176, a todas luces inabordable.

La segunda razón que justifica la introducción de los modelos de exposición, deriva, como hemos señalado, de las necesidades de información. Cuando un plan de medios está formado por una única inserción en un sólo soporte, el problema de la determinación de la cobertura es sencillo. A través de una encuesta se puede conocer cuántos individuos están habitualmente expuestos al soporte, esto es, se puede conocer la audiencia del mismo y, consecuentemente, sería innecesaria cualquier modelización adicional. Sin embargo, cuando se realiza más de una inserción, o en el plan de medios se recurre a más de un soporte, el problema gana en complejidad. Esto es consecuencia de la interacción simultánea de los fenómenos de acumulación de las audiencias (un mismo individuo puede resultar expuesto a varias de las

inserciones que se han realizado en un soporte dado) o duplicación de las mismas (un mismo sujeto puede resultar expuesto a inserciones realizadas en soportes distintos). En general, como ilustran Leckenby y Ju (1990) o Chandon (1985), la teoría de la probabilidad demuestra que la probabilidad de exposición R a al menos una inserción en un plan de medios determinado (definición de cobertura del plan) viene dada por la expresión:

$$R = S_1 - S_2 + S_3 - \dots \pm S_N$$

donde S_1 es la suma de las audiencias de los soportes considerados, S_2 es la suma de las audiencias duplicadas, S_3 de las triplicadas y así sucesivamente.

Por su parte, como se verá en el capítulo tercero de esta tesis doctoral, el cálculo de la distribución de contactos (número de individuos expuestos a una, dos, tres, ... k inserciones de las N realizadas: E_k^N , también requiere estimar las audiencias duplicadas, triplicadas ... N -plicadas, al calcularse mediante el Teorema de Waring (Waring, 1792) como sigue:

$$E_k^N = \sum_{i=k}^N (-1)^{i-k} \binom{i}{k} S_i \quad 0 \leq k \leq N \quad S_0 = 1$$

donde S_i es la señalada suma de las audiencias i -plicadas.

El principal problema reside en que los estudios de audiencia proporcionan, a lo sumo, las duplicaciones entre los soportes¹² y, por tanto, se hace necesario encontrar métodos para estimar de una manera precisa las duplicaciones de orden superior a partir de estos datos limitados. Para llevar a cabo ese proceso de cálculo nacen los modelos de exposición que se desarrollaran en esta tesis doctoral que, una

¹² Este problema se agudiza cuando la población objetivo son segmentos de población muy especializados, dado que ni siquiera estos datos de duplicación de audiencias entre soportes están disponible en Estados Unidos salvo para unas variables sociodemográficas muy generales (Cannon, 1983)

vez obtenidas esas duplicaciones, permitirán derivar la cobertura y distribución de contactos de los planes de medios.

En síntesis, se ha tratado de poner de manifiesto el interés para el marketing de el tema abordado en esta tesis doctoral. Entendemos, al igual que Leckenby y Kim, (1992, pág. 100) que “los modelos de estimación de la cobertura y la distribución de contactos son y, probablemente seguirán siendo, un elemento importante de la planificación de medios al ser piezas fundamentales para la toma de decisiones publicitarias. Diversos estudios indican que hay un fuerte interés en que se mejoren los modelos de exposición (Leckenby y Kishi, 1982a; [Leckenby y Boyd, 1984b; Leckenby y Kim, 1994a]; Kreshel, Lancaster y Toomey, 1985; Lancaster, Kreshel y Harris, 1986). En efecto, el incremento del coste de los medios, la competencia de los “nuevos medios” y la saturación de los mismos, hacen esta mejora en la estimación de las exposiciones, incluso más crítica que en el pasado”. La constatación de esta realidad, también para el caso español, fue objeto del epígrafe 1.2.

1.4 EVOLUCIÓN DE LA LITERATURA SOBRE MODELIZACIÓN EN LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS

1.4.1 La importancia de la modelización en los trabajos sobre planificación de medios publicitarios

El objetivo de este epígrafe, es situar la importancia que los trabajos sobre la modelización en la planificación de medios han tenido en la literatura sobre publicidad en marketing. Para ello hemos procedido en dos fases. En la primera de ellas, se compara el peso de estos trabajos en el contexto de los realizados sobre modelización en la herramienta de comunicación en general, poniéndolos en relación con otros realizados en publicidad, promoción de ventas, relaciones públicas y marketing directo. En la segunda fase, se analiza el peso de la modelización sobre el conjunto de trabajos realizados sobre planificación de medios, incluyan o no modelos. Se pretende, de esta forma, dar un argumento adicional más en la justificación de la realización de esta tesis doctoral, en base al peso que ha tenido su objeto de estudio en la literatura de nuestro área de conocimiento.

Stautamoyer y Lindquist (1994) revisan toda la literatura existente desde los años sesenta hasta el momento de realización su trabajo, sobre los modelos de decisión en el campo de la comunicación en marketing y, entre las diversas tipologías que utilizan en su clasificación, basan una de ellas en los elementos del *mix* de comunicación en que se centra cada modelo. De sus resultados, sintetizados en el cuadro 1.8a, puede concluirse en primer lugar que la modelización en los temas de comunicación se ha centrado, principalmente en la publicidad. Así, para el conjunto del periodo analizado 1960-1993, casi el 80% de los trabajos versaban sobre modelos en esta disciplina. No es menos cierto que, con el paso de los años, el

peso de la investigación sobre modelización en publicidad ha disminuido sensiblemente, dando lugar a trabajos más numerosos en otras herramientas comunicacionales, fundamentalmente la promoción de ventas. De este modo, si en los años sesenta, el 92,9% de todos los trabajos estaban centrados en la publicidad, durante el periodo 1980 a 1993, este porcentaje se había reducido al 71,2%, pasando de no existir ningún trabajo sobre modelización en la promoción de ventas, a representar éstos casi el 20% del total. Merecen ser destacados los de Varadarajan (1985), quien analizó las prácticas de promoción de ventas de las pequeñas empresas, concluyendo que la mayoría de las mismas no formalizan los planes de promoción de ventas, ni guardan información de pasadas promociones. Chapman (1986), quien estableció la influencia de los cupones sobre las ventas a corto y largo plazo en el negocio de las pizzas, Allaway, Mason y Brown (1987), quienes diseñaron un modelo de decisión que, además de publicidad, incorporaba la promoción de ventas como herramienta comunicacional. Finalmente, Walters (1989), usando un modelo logit, examinó el apoyo del minorista a las reducciones de precios y uso de expositores del producto promovidos por el fabricante.

Cuadro 1.8a. Evolución de los trabajos sobre modelización en comunicación
Porcentajes verticales

TEMA DEL MODELO	PERIODO CONSIDERADO			
	1960-69 (N = 28)	1970-79 (N = 43)	1980-93 (N = 66)	1960-93 (N = 137)
Publicidad	92,9	86,0	71,2	79,9
<i>Planificación de medios</i>	80,8	43,2	55,3	56,8
<i>Otros temas</i>	19,2	56,8	44,7	43,2
Promoción de ventas	0,0	11,6	19,7	12,9
Relaciones públicas	0,0	0,0	1,5	0,7
Marketing directo	0,0	0,0	4,5	2,2
Mix de comunicación	7,1	2,3	3,0	4,3
Totales	100	100	100	100

Fuente: Elaboración propia a partir de Stautamoyer y Lindquist (1994)

La composición interna de los trabajos en publicidad, recogida en el cuadro 1.8b, además, también ha sufrido cambios sustanciales. Si durante los años 60 la mayoría de los modelos versaban sobre planificación de medios (80,8%), durante la década de los setenta su peso se había reducido a la mitad (43,2%), debido a dos factores. Por un lado, al desarrollo de gran número de trabajos sobre gestión publicitaria, como el libro de Simon (1971) sobre la toma de decisiones publicitarias, los artículos de Doyle y Fenwick (1975a; 1975b) o Doyle (1977) sobre determinación de la estrategia publicitaria más eficaz. Por otro lado, había pasado la etapa más creativa de la planificación de medios, los años sesenta, con los trabajos de Agostini (1961, 1962), Metheringham (1964), Schreiber (1969), Day (1962), Engel y Warshaw (1964), Kotler (1964), para reducirse, incluso en términos absolutos, la producción de modelos sobre este tema, destacando los de Gensch (1970b), Sissors (1971), Fleck (1973) y Aaker (1975) sobre selección de medios publicitarios.

Cuadro 1.8b. Evolución de los trabajos sobre modelización en publicidad
Porcentajes verticales

TEMA DEL MODELO	PERIODO CONSIDERADO			
	1960-69 (N = 28)	1970-79 (N = 43)	1980-93 (N = 66)	1960-93 (N = 137)
Planificación de medios	100	100	100	100
<i>Audiencia de medios</i>	19,0	0,0	3,8	7,9
<i>Duplicación de audiencia</i>	0,0	6,3	3,8	3,2
<i>Mix de medios</i>	4,8	6,3	0,0	3,2
<i>Selección de medios</i>	57,1	50,0	26,9	42,9
<i>Esquema temporal</i>	14,3	6,3	11,5	11,1
<i>Selección de agencias</i>	0,0	6,3	0,0	1,6
<i>Eficacia de los medios</i>	4,8	25,0	50,0	28,6
<i>Revisión de los modelos</i>	0,0	0,0	3,8	1,6
Otros temas	100	100	100	100
<i>Gestión publicitaria</i>	0,0	28,6	9,5	16,7
<i>Asignación de recursos</i>	60,0	38,1	38,1	41,7
<i>Creatividad y contenido</i>	20,0	9,5	19,0	14,6
<i>Eficacia publicitaria</i>	20,0	23,8	33,3	27,1

Fuente: Elaboración propia a partir de Stautamoyer y Lindquist (1994)

En el último de los periodos analizados, el peso de los trabajos sobre planificación de medios ha vuelto a incrementarse sobre el resto de trabajos de modelización en publicidad, suponiendo más de la mitad del total (55,3%). La explicación cabe buscarla en el hecho de que la creación de modelos sobre otros temas publicitarios se ha estancado, podemos destacar: los dedicados a analizar la eficacia de la publicidad, como los de O'Herlihy (1980), Farley y Tapiero (1981) o Wierenga (1981). Por contra, en la planificación de medios, abundan los trabajos relativos a crear nuevos modelos de exposición o a comparar los existentes, muchos de los cuales serán analizados en capítulos posteriores: Rust y Klompmaker (1981), Leckenby y Kishi (1984), Rust y Leone (1984), Leckenby y Rice (1985b; 1986) o Danaher (1988b; 1989; 1991).

En síntesis, del análisis del trabajo de Stautamoyer y Lindquist (1994) podemos establecer las siguientes conclusiones. En primer lugar, que la modelización en el área de la comunicación, se ha centrado fundamentalmente en publicidad y, dentro de ella, en la planificación de medios. En segundo lugar, que los trabajos sobre otras herramientas de comunicación ha crecido sustancialmente, principalmente los modelos que versan sobre promoción de ventas y finalmente que, aunque el peso de la planificación se ha reducido, sigue abarcando más de la mitad de la producción de modelos en publicidad.

En un reciente artículo publicado en el *Journal of Current Issues and Research in Advertising*, Pasadeos, Barban, Yi y Kim (1997) revisan la evolución de la literatura sobre planificación de medios durante las tres últimas décadas (1962-1991). En el mismo, se pone de manifiesto el papel destacado que los trabajos sobre modelización han ocupado en el conjunto de la producción científica sobre este tema. Del análisis del cuadro 1.9 se derivan claramente tres resultados:

- a. En el horizonte de los treinta años revisados, sólo los artículos sobre aspectos estratégicos de la planificación de medios superan en número a los

que tienen a la modelización como tema fundamental y, en todo caso, prácticamente uno de cada cinco artículos se centra en esta materia.

- b. La época más creativa de esta disciplina tuvo lugar durante las dos primeras décadas analizadas, produciéndose una progresiva pérdida de interés por parte de los investigadores en el último decenio que, quizás, quepa atribuir a que las aportaciones que se van produciendo no mejoran significativamente la capacidad predictiva de los modelos ya precedentes.
- c. Dentro de los modelos de selección de medios, son los modelos de exposición quienes han gozado de mayor atención por parte de los investigadores, en detrimento de los modelos completos de planificación de medios.

Cuadro 1.9. Temas principales de los artículos sobre planificación de medios
Porcentajes verticales *

Tema del artículo	1962-71 (N=80)	1972-81 (N=98)	1982-91 (N=188)	1962-91 (N=366)
Valoración general de la Planificación de Medios	6,3	7,1	11,2	9,0
Métodos de investigación	26,6	13,3	15,4	17,3
Los medios y su relación con el Marketing	3,8	-	7,4	4,7
Modelos de selección de medios	32,9	20,4	11,6	18,6
<i>Modelos de exposición</i>	13,9	8,2	7,4	9,0
<i>Modelos de Planificación de Medios</i>	-	-	0,5	0,3
<i>Ambos tipos</i>	19,0	12,2	3,7	9,3
Selección cualitativa de medios	5,1	2,0	3,7	3,6
Aspectos estratégicos	19,0	33,7	21,4	24,0
<i>Esquema temporal inserciones</i>	5,1	10,2	5,9	6,8
<i>Determinación público objetivo</i>	2,5	14,3	3,7	6,3
<i>Mix de medios</i>	3,8	3,1	5,3	4,4
<i>Funciones de respuesta</i>	-	3,1	2,7	2,2
<i>Presupuesto</i>	6,3	1,0	1,1	2,2
<i>Otros / Combinación</i>	1,3	2,0	2,7	2,1
Aspectos tácticos	2,5	-	0,5	0,8
Aspectos macro	1,3	1,0	1,6	1,4
Compra de medios	1,3	10,2	8,5	7,4
Comportamiento del suministrador	-	1,0	2,7	1,6
Características de los medios	1,3	6,1	5,9	4,9
Otros	-	5,1	9,1	6,0
Combinación	-	-	1,1	0,5
Totales	100	100	100	100

* Los porcentajes no suman el 100% debido al redondeo

Fuente: Pasadeos, Barban, Yi y Kim (1997)

Este análisis, sin embargo, es meramente cuantitativo. Un enfoque más cualitativo pasa por tratar de determinar qué trabajos han ejercido una influencia más relevante en la literatura sobre planificación de medios. La técnica bibliográfica habitualmente utilizada para este fin, es la revisión de las veces que han sido citados en otros artículos (Culnan, 1986; Pasadeos, 1985; Tankard, Chang y Tsan, 1984). En este sentido, el trabajo de Pasadeos, Barban, Yi y Kim (1997), considera solamente los cuatro últimos años de cada periodo de diez, para asegurarse de que tienen la posibilidad de ser citados el mayor número posible de artículos en cada periodo. Asimismo, para este análisis, se han eliminado las autocitas.

Como muestra el cuadro 1.10, donde para cada periodo aparecen ordenados los artículos según el número de citas que han recibido, la importancia de la modelización en planificación de medios tiene incluso un peso más destacado cuando se realizan consideraciones de carácter cualitativo acerca de la producción literaria. Nueve de los dieciocho artículos diferentes listados, tienen que ver directamente con modelos de selección y otros cinco guardan una relación indirecta con la modelización. Cronológicamente, sin embargo, el peso de las citas, se ha desplazado de aquellos trabajos que desarrollaban teóricamente nuevos modelos, hacia la investigación sobre el uso que hacen de los mismos las agencias de publicidad.

Llama la atención que el trabajo de Metheringham (1964) sea uno de los más citados en los tres periodos analizados. La explicación cabe buscarla en que como señalaron Little y Lodish (1969), los procedimientos descritos por Metheringham (1964) resultan idénticos a los que se derivan del uso de la distribución beta binomial compuesta, y esta distribución, como se verá en capítulos posteriores, es utilizada, de una u otra manera, en la mayoría de los modelos desarrollados en los Estados Unidos (Leckenby y Ju, 1990).

Cuadro 1.10. Artículos sobre planificación de medios más citados.

Periodo	Artículo	Relación	Tema
1968-71	1. Agostini (1961)	[**]	Modelos de selección de medios
	2. Metheringham (1964)	[**]	Modelos de selección de medios
	3. Agostini (1962)	[**]	Modelos de selección de medios
	4. Broadbent (1966)	[**]	Modelos de selección de medios
	5. Hofmans (1966)	[**]	Modelos de selección de medios
	6. Little y Lodish (1969)	[**]	Modelos de selección de medios
	7. Zielske (1959)	[-]	Funciones de respuesta
1978-81	1. Little y Lodish (1969)	[**]	Modelos de selección de medios
	2. Metheringham (1964)	[**]	Modelos de selección de medios
	3. Swanson (1967)	[-]	Características de los medios
	4. Aaker (1975)	[**]	Modelos de selección de medios
	5. Frank, Becknell y Clokey (1971)	[-]	Características de los medios
	6. Wells (1969)	[-]	Características de los medios
1988-91	1. Leckenby y Kishi (1982a)	[*]	Uso modelos por agencias
	2. Leckenby y Boyd (1984b)	[*]	Uso modelos por agencias
	3. Leckenby y Kishi (1982b)	[**]	Modelos de selección de medios
	4. Rust y Leone (1984)	[**]	Modelos de selección de medios
	5. Kreshel, Lancaster y Toomey (1985)	[*]	Uso modelos por agencias
	6. Lancaster, Kreshel y Harris (1986)	[*]	Uso modelos por agencias
	7. Leckenby y Kishi (1984)	[**]	Modelos de selección de medios
	8. Metheringham (1964)	[**]	Modelos de selección de medios
[**]	Directamente relacionado con los modelos de selección de medios		
[*]	Indirectamente relacionado con los modelos de selección de medios		
[-]	No guarda relación con los modelos de selección de medios		

Fuente: Elaboración propia a partir de Pasadeos, Barban, Yi y Kim (1997)

En síntesis, y siempre de acuerdo con Pasadeos, Bargan, Yi y Kim (1997), parece evidente que con el tiempo se ha producido una diversificación de los temas abarcados por la literatura sobre planificación de medios. Cuando se trataba de una disciplina joven, allá por los sesenta, la mayoría de los trabajos versaban sobre modelos de selección de medios. Con el tiempo el énfasis parece haberse desplazado de la creación de modelos, a la evaluación e introducción de mejoras marginales en los existentes y a la determinación de los patrones de uso de los mismos por parte de las agencias. Pese a que el número de artículos sobre modelización ha decrecido con los años, las principales influencias sobre la literatura (autores y trabajos más citados) todavía proceden de este área.

1.4.2 La literatura sobre modelización en la planificación de medios publicitarios

La modelización en planificación de medios gira en torno a tres grandes ejes (Wilson y Machleit, 1985): los modelos de exposición, los modelos acerca de la función de respuesta y los modelos de planificación de medios. Antes de entrar en más detalles, debe precisarse que la terminología seguida en esta clasificación no es estrictamente la empleada en el artículo referenciado. Si algo caracteriza el desarrollo tipológico en este campo es, desde nuestro punto de vista, la heterogeneidad en los términos utilizados para hacer referencia a una misma realidad. En aras a racionalizar esta diversidad, se ha optado en esta tesis por utilizar la terminología acuñada en los trabajos que, desde nuestra perspectiva, están más reconocidos. Así, entenderemos por *modelos de exposición* (Chandon, 1985)¹³ aquellos que estiman cuántos individuos (cobertura) resultarán expuestos, cuántas veces (distribución de frecuencias) a los soportes que conforman un plan de medios. Hemos renunciado a

¹³ Traducción dada al término *Exposure Models*, y que en España ha sido ya aplicada en diversos trabajos (León, 1986; 1988). El trabajo de Chandon (1985) es, a todos los efectos, la revisión más profunda realizada de este tipo de modelos como reconocen, entre otros, Leckenby y Ju (1990).

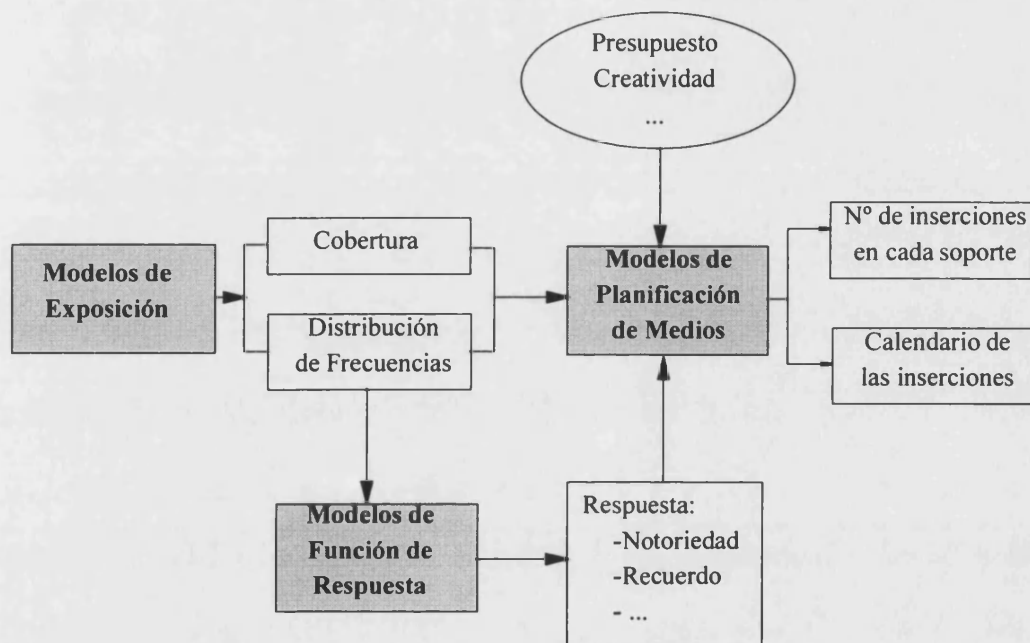
emplear términos como “modelos de alcance y frecuencia (Leckenby y Ju, 1980) o modelos de estimación del alcance (Wilson y Machleit, 1985). Por su parte, un *modelo de planificación de medios* (Chandon, 1985) es aquel que permite seleccionar la combinación adecuada de soportes publicitarios, el número de inserciones a colocar en cada uno de ellos y el momento del tiempo en que deben realizarse las inserciones, todo ello dentro del límite presupuestario y restricciones sobre disponibilidad de medios. Hemos renunciado aquí a términos como modelos de asignación de medios (Wilson y Machleit, 1985: *media allocation models*), modelos de medios (León, 1986; 1988) o a introducir adjetivos que lo precisen: modelos completos de planificación de medios (Leckenby y Ju, 1990: *comprehensive media planning models*).

Resulta evidente que, cualquier modelo de planificación de medios debe llevar implícito un modelo de exposición, a partir del momento en que las variables de cobertura y distribución de frecuencias son un *input* fundamental de aquellos. Por este motivo, aunque nuestra tesis se centre en los modelos de exposición, en la revisión de la literatura que ofrece este apartado, se ha optado por incluir en la misma a ambos tipos de modelos, que habitualmente se suelen agrupar bajo la etiqueta de *modelos de selección de medios* (Pasadeos, Barban, Yi y Kim, 1997).

El tercer foco de investigación en la modelización en planificación de medios, hace referencia a las funciones de respuesta. Esta línea de investigación, de larga tradición en marketing, cobró un nuevo impulso, precisamente, como consecuencia de ser parte fundamental en los modelos de planificación de medios. En la medida en que vinculan la frecuencia de exposición a la publicidad, con algún tipo de respuesta por parte del individuo (en términos de recuerdo, preferencia de marca, intención de compra, etc.) deben formar parte fundamental, como trata de poner de manifiesto la figura 1.3, de los algoritmos que forman parte de los modelos de planificación de medios. Sin embargo la única vinculación con los modelos de exposición, objeto de este trabajo, es que toman como *input* los datos sobre

distribución de frecuencias que éstos proporcionan. Por este motivo hemos optado por dejarlos fuera de la revisión de la literatura que efectuaremos a continuación.

Figura 1.3. La modelización en la planificación de medios



Fuente: Elaboración propia

Sin embargo, pueden consultarse Wilson y Machleit (1985) para una buena revisión sobre esta cuestión. Allí se revisa el trabajo de Zielske (1959) quien analizó la evolución del recuerdo en dos patrones de inserción diferentes: exposiciones semanales durante trece semanas y mensualmente durante una año, comprobando que las curvas de olvido eran distintas. Grass (1968) y Grass y Wallace (1969) quienes en diferentes experimentos constataron el efecto *desgaste*, esto es el decaimiento del recuerdo a niveles elevados de frecuencia, que posteriormente ha sido abordado en trabajos como los de Appel (1971), Calder y Sternthal (1980) y Craig, Sternthal y Leavitt (1976). El nivel óptimo de frecuencia de exposición, o equivalentemente, cuál es el perfil de la curva de respuesta, fue inicialmente

analizado en el trabajo clásico de Krugman (1972). Posteriormente Simon y Arndt (1980) evaluaron empíricamente dos curvas de respuesta alternativas (cóncava decreciente y con forma de S).

Respecto a los esquemas temporales de las inserciones, tras el mencionado trabajo de Zielske (1959), diversos autores han encontrado que el esquema en “pulsaciones” ofrece mejores resultados (Moran, 1976; Rao y Miller, 1975; Strong, 1977) desarrollando Strong (1974; 1977) y Zielske y Henry (1980) modelos para estimar el recuerdo del anuncio según diferentes tipos de pulsaciones temporales.

1.4.2.1 *Los modelos de exposición*

Dado que en el capítulo segundo se procederá al desarrollo de una tipología de este tipo de modelos, prestando especial atención a las hipótesis que los sustentan, nos limitaremos en este apartado a tratar de dar una visión cronológica de su creación, sin entrar en detalles sobre su formulación o limitaciones. Cabe añadir que en esta revisión se incluirán solamente los modelos fundamentales, siendo el capítulo segundo mucho más exhaustivo a este respecto. Por otra parte se hará referencia a diversos modelos de propietarios que, posteriormente no serán desarrollados. El motivo es que, precisamente por su carácter comercial, sus algoritmos no son accesibles para el investigador y, por ello, su evaluación resulta imposible.

El origen de los modelos de exposición, cabe buscarlo en el trabajo seminal de Agostini (1961) quien desarrolla una sencilla fórmula para estimar la cobertura, cuyo parámetro, la constante $k = 1.125$, se estimaba mediante regresión sobre datos de audiencias de revistas francesas. Trabajos posteriores, llevados a cabo en diferentes países, han obtenido otros valores para k , demostrando que no es tal constante. Así, Bower (1963), con datos americanos y estadounidenses, obtuvo que era necesaria una constante mayor, puesto que la curva con la constante de Agostini

(1961) sobreestimaba sistemáticamente estos datos. Kuhn, (1963), con datos alemanes, concluye que la constante k debe depender del número de soportes incluidos en el plan. Marc (1963), con datos franceses, obtiene que la constante debe variar según el tipo de revista que se incorpore al plan de medios (v.g.: 1.6 para decoración, 1.4 para negocios, 1.3 para las revistas de carácter general y 1.5 para las revistas especializadas en el sector de lavanderías y limpieza en seco). Caffyn y Sagovsky (1963), al comparar esta fórmula con los métodos de Sainsbury, afirma que la constante de Agostini (1961), debería hacerse depender del número de soportes incluidos en el plan, conclusión similar a la de Claycamp y McClelland, 1968). En España, Redondo (1994) ha establecido para la k un valor de 1'048. Pese a estas consideraciones, la importancia de este trabajo reside, desde nuestra perspectiva, en haber sido el primero en el intento de modelizar los procedimientos de estimación de la cobertura y la distribución de contactos..

Hofmans (1966), demostró que la curva obtenida empíricamente por Agostini (1961) mediante regresión, era equivalente a la suma de una serie geométrica infinita, en cada uno de cuyos términos se incluía una expresión variable de la k de Agostini, obteniendo así un modelo alternativo.

Hasta este momento, todos los trabajos se basaban en la búsqueda de formas funcionales que se adaptaran a los datos de audiencias disponibles. Este planteamiento tiene la grave limitación de no permitir la obtención de la otra variable fundamental para la planificación: la distribución de frecuencias. Hay que esperar hasta 1964 para encontrarnos con, quizás¹⁴, el trabajo fundamental en el desarrollo de los modelos de exposición: el de Metheringham (1964). La contribución de esta obra cabe buscarla en dos aspectos. Por un lado, abandona el enfoque de ajuste de curvas, para utilizar funciones de probabilidad, hecho que permite la obtención de la

¹⁴ Como se vio en el apartado 1.4.1 (cuadro 1.10), este trabajo es el único que se ha mantenido entre los más citados en los tres periodos en los que Pasadeos, Barban, Yi y Kim (1997) han dividido la treintena de años de literatura revisada sobre planificación de medios.

distribución de frecuencias de exposición. Aunque el autor, tal vez, no era consciente de ello, su planteamiento era equivalente a la utilización de una distribución Beta Binomial (BBD) como demostraron posteriormente Little y Lodish (1969). Por otro lado, según Leckenby y Ju (1990), la mayoría de los modelos de exposición desarrollados con posterioridad en los Estados Unidos, se encuentran, de una u otra manera, apoyados en la distribución BBD introducida en este trabajo, como a continuación se verá.

En primer lugar, nos encontramos con un conjunto de modelos que han optado por estimar por otros procedimientos los parámetros de la distribución BBD. El procedimiento de Metheringham (1964), se denomina de estimación directa, por cuanto que los datos que se requieren pueden obtenerse directamente de estudios de audiencia. Un procedimiento de estimación indirecta mediante regresión lineal ha sido aplicado por Headen, Klompmaker y Teel (1976;1977;1979) y por Rust y Klompmaker (1981) a planes de medios en televisión. Básicamente, consiste en ajustar la expresión teórica de la BBD a un conjunto de distribuciones de frecuencias reales. La varianza de cada ajuste es explicada a continuación con diversas variables independientes que recogen las características del plan de medios: *Gross Rating Points* (GRPs), número de soportes, duplicación total entre los soportes, etc. Para calcular la distribución de frecuencias de un nuevo plan, basta con sustituir sus características en la expresión así deducida.

En segundo lugar, un planteamiento alternativo, ha consistido en tratar de mejorar la BBD. Así, Sabavala y Morrison (1981) introducen en la misma un tercer parámetro que permite que los individuos tengan diferentes probabilidades de exposición en distintos momentos del tiempo, convirtiendo así la BBD en una distribución beta binomial geométrica (BBG). Una segunda modificación ha sido introducida por Leckenby y Boyd (1984a) y consiste en estimar la cobertura por la fórmula de Hofmans (1966) y ajustar la BBD a esta cobertura de una manera secuencial: distribución beta binomial de Hofmans (HBBD).

Un bloque distinto de modelos ha optado por abandonar la BBD como la distribución de probabilidades que sustenta el modelo y sustituirla por distribuciones alternativas. Leckenby y Ju (1989) ofrecen dos de estos métodos, la distribución Poisson-Binomial (PBD), donde los únicos datos necesarios para estimar los parámetros de la distribución son el número total de inserciones y la audiencia de los soportes incluidos en el plan y la distribución Hofmans Poisson-Binomial (HPBD), donde previamente se estima la cobertura mediante la fórmula de Hofmans (1966) y a continuación, los parámetros de la PBD se van ajustando en un proceso iterativo hasta que proporcione la misma cobertura que proporciona la fórmula de Hofmans. Bajo este mismo enfoque de distribuciones de probabilidad alternativas, Leckenby y Kishi (1981) combinan las fórmulas de cobertura de Kwerel (1969) y Hofmans (1966) con la distribución geométrica obteniendo los modelos de Kwerel - geométrico y Hofmans - geométrico (KGD y HGD respectivamente).

Los modelos hasta ahora reseñados, se caracterizan porque sus algoritmos van promediando todos los soportes del plan de medios en un soporte “medio” para, a continuación, plantear un plan de n inserciones en ese soporte. Un enfoque distinto lo plantean los modelos llamados por Leckenby y Ju (1990) *multivariados*, los cuales tratan de mantener tantos parámetros a estimar como soportes están incluidos en el plan de medios. El problema con que se encuentran es que, para que el proceso sea multivariado “puro” sería necesario estimar 2^m parámetros, lo cual, para un plan de medios típico sería inviable, hecho que fuerza a la aplicación de aproximaciones en la estimación. Los modelos más importantes dentro de esta tendencia, son el Dirichlet Multinomial (DMD: Chandon, 1985), el Dirichlet Multinomial Multivariable (DMMD Boyd, 1985), el Dirichlet Multinomial Modificado (MDD: Leckenby y Kim, 1994b), todos ellos aplicados al medio revistas y la modificación del DMD para un plan combinado de varios medios propuesto por Rust y Leone (1984), el MMDMD.

Finalmente, para soslayar el problema que plantea estimar tal cantidad de parámetros en el enfoque multivariado antes expuesto, se ha desarrollado el enfoque de modelos de agregación, aunque el único cuyo algoritmo ha sido explícitamente expuesto es el del modelo de agregación secuencial de Morgensztern (MSAD: Leckenby y Rice, 1986a), en el cual, tras estimar la cobertura mediante la fórmula de Morgensztern (Chandon, 1985) se estima una distribución de probabilidad para cada soporte utilizando una BBD. Estas distribuciones se combinan secuencialmente para formar una única distribución.

El cuadro 1.11 ofrece un resumen de los principales modelos de exposición expuestos en los párrafos anteriores.

Cuadro 1.11. Principales modelos de exposición.

Filosofía	Modelo	Autor
Trabajos seminales:	Agostini	Agostini (1961)
Estimación de la cobertura	Hofmans	Hofmans (1966)
Basados en la BBD	BBD	Metheringman (1964)
	HBBB	Leckenby y Boyd (1984a)
	BBG	Sabavala y Morrison (1981)
	BBD-Est. Ind.	Rust y Klompmaker (1981) Headen Klompmaker y Teel (1976; 1977; 1979)
Basados en otras distribuciones	PBD	Leckenby y Ju (1989)
	HPBD	Leckenby y Ju (1989)
	KGD	Leckenby y Kishi (1981)
	HGD	Leckenby y Kishi (1981)
Modelos multivariados	DMD	Chandon (1985)
	DMMD	Boyd (1985)
	MDD	Leckenby y Kim (1994)
	MMDMD	Rust y Leone (1984)
Modelos de agregación	MSAD	Leckenby y Rice (1986a)

Fuente: Elaboración propia

1.4.2.2 Modelos de planificación de medios publicitarios

Mientras que los modelos de exposición se centran únicamente en la determinación de la cobertura y la distribución de contactos existen, tal como dijimos con anterioridad, otro grupo de modelos, los de planificación de medios, que resuelven el problema completo de la planificación: repartir de una manera eficiente el presupuesto entre un conjunto de soportes. Aunque, existen algunos modelos que abordan el problema de la distribución temporal de las inserciones (Simon, 1982); Strong, 1977; Thompson, 1981, Zielske y Henry, 1980; Zufryden, 1973;1975a) no serán tratados salvo que resuelvan también el problema general de la planificación de medios que acabamos de formular. Seguiremos, fundamentalmente, los trabajos de Bigné (1990b; 1991b) y Leckenby y Ju (1990) en la revisión de las principales contribuciones en este campo, aunque a efectos tipológicos nos centraremos más en las técnicas utilizadas en la resolución de los algoritmos: programación lineal, entera, por objetivos y simulación, que en la filosofía de los mismos: evaluación u optimización.

Los primeros modelos de planificación de medios se desarrollaron tomando como base la utilización de la *programación lineal continua* (Bass y Lonsdale, 1966, 1967; Broadbent, 1962; Brown y Warshaw, 1965; Day, 1962; Engel y Warshaw, 1964; Miller y Starr, 1960; Stasch, 1965). Estos modelos constan, fundamentalmente, de dos partes: una función objetivo y un conjunto de restricciones. La función objetivo habitualmente se plantea como:

$$\text{Max } R = f(X_j)$$

donde R es el rendimiento del plan, $f(X_j)$ es una función de los soportes que recoge los objetivos perseguidos. Esta función objetivo, como indica Gensch (1970b), incluye ponderaciones que intentan recoger factores como la población objetivo, la

diferencia de prestigio entre soportes, diferentes valores de exposición según cual sea el formato del anuncio insertado, etc. Estas consideraciones cualitativas, son las que sirven para distinguir unos modelos de otros.

Las restricciones, por su parte, pueden ser de distintos tipos:

$$C_1X_1 + \dots + C_nX_n \leq B \quad [1.1]$$

$$C_2X_2 + C_5X_5 \geq 0,25B \quad [1.2]$$

$$5 \leq X_j \leq 52 \quad [1.3]$$

$$X_j \geq 0 \quad [1.4]$$

donde el término C_j de las hace referencia al coste por inserción. La restricción [1.1] limita, en función del presupuesto B , la cantidad de inserciones que pueden realizarse. Las restricciones del tipo [1.2] incorporan criterios de experiencia por parte del planificador y fuerzan a que en determinados soportes (v.g. X_2 y X_5) se haga una cantidad importante de inserciones (la cuarta parte del presupuesto), independientemente del criterio meramente optimizador del algoritmo. Restricciones como la [1.3] indican el número mínimo o máximo de inserciones que pueden realizarse en determinados soportes como consecuencia, por ejemplo, de su carácter mensual, semestral, etc. y, finalmente, la restricción [1.4] es la lógica de no negatividad de las inserciones. Cabe, naturalmente, el planteamiento alternativo de minimizar la función de costes, estableciendo como restricción los niveles de audiencia que se pretenden lograr.

Bigné (1990b), señala que los algoritmos de la programación lineal continua provocan una serie de limitaciones en estos modelos que han llevado al desarrollo de procedimientos alternativos de resolución del problema de planificación. La primera

de ellas es que la programación lineal permite que el número de inserciones adopte valores fraccionales, cuando es evidente que deben ser enteros; no permite tratar los problemas de la acumulación y duplicación de las audiencias, no suelen proporcionar la distribución temporal de las inserciones, no tienen en cuenta los descuentos en la compra de medios y, finalmente, las funciones objetivos permiten optimizar un único objetivo ligado, por lo general, al número de exposiciones.

Los trabajos de Steuer y Oliver (1976), Charnes, Cooper, Devoe, Lerner y Reinecke (1968) y Charnes, Cooper, Lerner y Snow (1968), manteniéndose dentro de la programación lineal, incorporan en sus modelos la posibilidad de considerar simultáneamente varios objetivos, a través de la técnica de la *programación por objetivos*. El planteamiento es similar a la programación lineal continua, con la salvedad de que es necesario establecer un valor numérico específico para cada uno de los objetivos y plantear una función para cada uno de ellos. La resolución pasará por buscar una solución que minimice la suma ponderada de las desviaciones de las funciones objetivos respecto a sus objetivos respectivos.

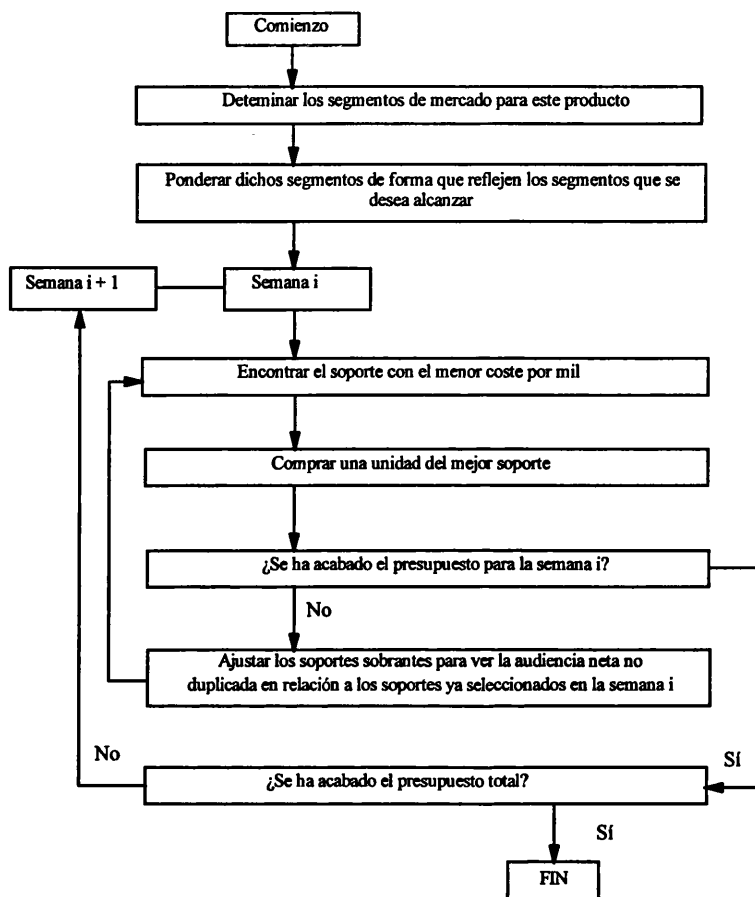
Zangwill (1965), Zufyden (1973;1975b) y Keown y Duncan (1979) abordan el problema de valores fraccionales de las inserciones recurriendo a la *programación entera*, técnica que fuerza a que las variables objetivo sean de este tipo. Como se puso de manifiesto en este mismo capítulo, diversos estudios (Appel, 1971; Calder y Sternthal 1980; Craig, Sternthal y Leavitt, 1976), aún no poniéndose de acuerdo con cual es la forma funcional de la función de respuesta, concluyen que el comportamiento ante exposiciones adicionales no es lineal (efecto *wearout*). Intentando recoger esta realidad, Locander, Scamell, Sparkman y Burton (1978) introducen funciones de respuesta no lineales que reflejan la disminución de la efectividad marginal publicitaria a medida que se incrementa el número de exposiciones. Manteniéndose dentro de la programación entera como técnica de resolución, Deckro y Murdok (1987) recurren a la programación multiobjetivo para permitir que el planificador pueda introducir en el modelo la consideración de varios

objetivos simultáneamente. La programación entera, sin embargo, aún suponiendo un importante avance sobre la programación lineal, no está libre de limitaciones. La primera de ellas son los problemas de tiempo de cálculo derivado de la complejidad de sus algoritmos. Por otra parte, no consideran la distribución de frecuencias de contacto, las duplicaciones ni, por lo general, los descuentos por compra de medios.

Una técnica alternativa de abordar el problema de planificación de medios lo ofrecen los *modelos iterativos*, también conocidos como modelos de *análisis marginal*. Estos modelos han alcanzado mayor popularidad entre las agencias de publicidad que entre los investigadores, debido, probablemente, a la facilidad de entender su proceder (Wilson y Machleit, 1985). Este tipo de modelos, como ilustra la figura 1.4, llevan a cabo el proceso de optimización tomando como restricción el presupuesto. Para ello, seleccionan en primer lugar, el soporte con mayor valor en función del criterio de selección definido, y van repitiendo esta acción: incorporar nuevos soportes, hasta agotar el presupuesto.

Como se ha indicado, estos modelos gozaron de cierta popularidad entre las agencias, las cuales los desarrollaron sin hacer públicos, en algunos casos, sus algoritmos. Entre los más conocidos (Bigné, 1990b), cabe citar el HAMM de la Young & Rubicam, el Mather de la Benson y Mather, el modelo de Benson para la English Electronic - Leo - Marconi, el Mediaplan, de las sociedades francesas CFRO, Publicis, Publi Service y Montford-Bel, y el modelo de Brown (1967) y sus revisiones y el POMSIS de Aaker (1968).

Figura 1.4. Estructura de un modelo iterativo de planificación de medios



Fuente: Gensch (1970b, pág. 201)

Las principales limitaciones de los mismos, cabe buscarlas en que no siempre aportan una solución óptima (Brown, 1967) y, por otra parte, que la elección de un soporte condiciona las siguientes. Si al ejecutarse el plan no puede comprarse un soporte, la optimización se invalida debido a su carácter iterativo (Shocker, 1970). Por otro lado, entre sus ventajas, cabe señalar que permiten la consideración de los descuentos por compra de medios, pueden estudiarse las duplicaciones y acumulaciones de audiencia derivadas de una inserción adicional y pueden recogerse factores cualitativos de los soportes (Bigné, 1990b).

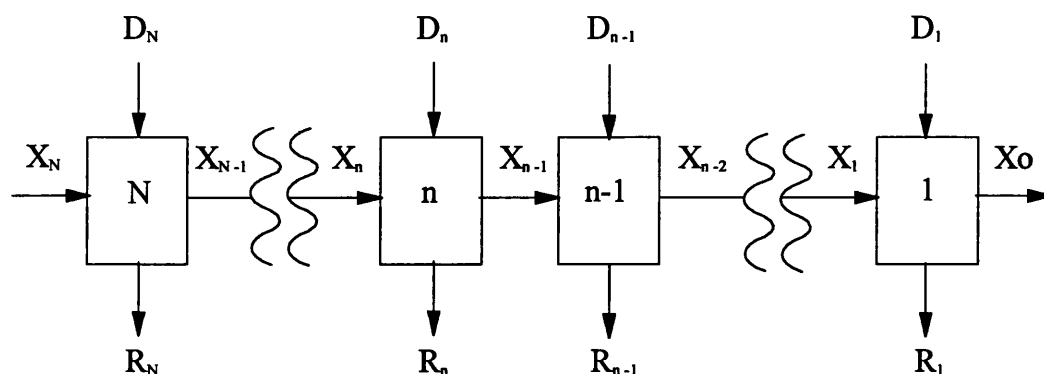
Algunos autores intentan soslayar el problema de la no linealidad de la función de respuesta recurriendo a la *programación dinámica*, técnica que, simultáneamente, considera el efecto acumulativo de la publicidad de unos periodos sobre los subsiguientes (Wilson y Machleit, 1985). Conceptualmente, la programación dinámica (Leckenby y Ju, 1990) aborda un problema complejo dividiéndolo en una secuencia de pequeños problemas. La función objetivo de la que debe ser maximizada mediante programación dinámica es:

$$\sum_{i=1}^n R_i$$

donde $R_i = f(S_i, d_i)$ siendo R_i la utilidad obtenida en la etapa i que es una función de la situación del proceso en la etapa i (S_i) y de la decisión adoptada en esa etapa del proceso (d_i). La programación dinámica selecciona la mejor decisión (con mayor utilidad) en la etapa n -ésima (la última) para luego proceder en sentido inverso con la etapa $n-1$ ésima hasta llegar a la primera etapa del proceso de decisión, tal y como ilustra la figura 1.5. De este modo la programación dinámica es una técnica que, según Leckenby y Ju (1990) que selecciona la mejor alternativa teniendo en consideración todas las posibles combinaciones de decisiones.

El número de aplicaciones de modelos de programación dinámica es bajo debido, fundamentalmente, al hecho de que el problema se vuelve inmanejable a efectos de cálculo aunque el número de periodos de tiempo, soportes o segmentos de mercado no sea demasiado elevado. Entre ellos, el modelo pionero fue el de Maffei (1960), aunque en realidad era más un ejercicio para demostrar cómo se podía repartir un presupuesto determinado entre tres medios en un mercado de prueba.

Figura 1.5. Sistema de programación dinámica



Fuente: Gensch (1970b; pág. 202)

Más importancia ha tenido el modelo presentado por Little y Lodish (1966), el MEDIAC. Este modelo, divide a la población en segmentos de mercado mutuamente exclusivos, cada uno de los cuales con su propio potencial de ventas y hábitos de audiencia. Las exposiciones sirven para aumentar lo que los autores denominan “valor de exposición” de cada segmento que también se halla expuesto a un factor de olvido, de tal manera que las ventas aumentan al aumentar las exposiciones aunque con rendimientos decrecientes. El modelo pretende maximizar el volumen de ventas generadas por las mencionadas inserciones. Como veremos posteriormente, debido a las mencionadas limitaciones de cálculo, los autores sustituyeron este modelo en trabajos ulteriores por otra versión basada en técnicas heurísticas.

En general, puede afirmarse (Leckenby y Ju, 1990) que los métodos basados en algún procedimiento de programación matemática, ya sea esta programación lineal, modelos iterativos o programación dinámica, no han funcionado nunca aceptablemente a un nivel práctico debido, fundamentalmente, a las hipótesis simplificadoras demasiado irreales que es necesario aplicar para adaptar el problema de planificación de medios al formato requerido por los algoritmos de resolución.

Estos mismos autores señalan que, aunque se ha intentado sofisticar los modelos para aproximarlos a la realidad (De Kluiver, 1978; Kaplan y Shecker, 1971; Locander, Scamell, Sparkman y Burton, 1978; Schreiber, 1968), estas modificaciones no han podido aplicarse la mayoría de las veces debido a que los problemas de cálculo las hacían impracticables.

Una forma de intentar que las hipótesis simplificadoras para reducir la realidad a unos pocos parámetros no empobrezcan radicalmente a los modelos, pasa por recurrir a técnicas de *simulación*. Estos modelos utilizan datos individuales de los sujetos con toda su complejidad, para intentar representar cómo un soporte puede alcanzarles. La mayoría de estos modelos han sido desarrollados por las agencias de publicidad y no en el ámbito académico (Wilson y Machleit, 1985; Leckenby y Ju, 1990). Este carácter comercial, hace que sus procedimientos de cálculo no sean conocidos con todo detalle. En todo caso, suponen una aproximación diferente al problema de la planificación de medios, en cuanto que no buscan la generación de un plan óptimo, como los enfoques revisados hasta el momento, sino que comparan distintos planes de medios ofreciendo una herramienta para evaluarlos.

Los más conocidos (puede recurrirse a Gensch, 1970b; 1973; o a Bigné, 1991b para una descripción más detallada de los mismos) son el modelo CAM realizado para la London Press Exchange por Beale, Hugues y Broadbent (1966), el modelo SCAL realizado en Francia por la empresa SEMA, el Media Planex elaborado por Agostini y Steinberg (1965) y el modelo de Friedman (1970). Para ilustrar la forma de actuar de estos modelos, describiremos brevemente el AD-ME-SIM de Gensch (1969) elaborado conjuntamente por este autor y la agencia J. Walter Thompson. Los *inputs* necesarios para este modelo son: el plan de medios que se pretende evaluar (soportes seleccionados, número de inserciones y esquema temporal de las mismas), un listado del coste de los soportes y los descuentos que pueden obtenerse y varias ponderaciones subjetivas del planificador para distintos segmentos, soportes, aspectos creativos y niveles de exposición. A continuación,

basándose en una muestra de individuos proveniente de un estudio de investigación de audiencias, los valores históricos de probabilidad de resultar expuesto de cada individuo se transforman en los patrones de exposición actuales mediante simulación de Montecarlo. Con los niveles de exposición así estimados, se pondera a cada individuo para determinar qué valor tiene para él una exposición, y esta información se combina con el resto de ponderaciones subjetivas: adecuación de cada medio al mensaje publicitario concreto, adecuación del formato publicitario, las ponderaciones de la distribución de frecuencias (valor de cada contacto) y los descuentos por compra de medios. Con estos datos, el modelo proporciona la cobertura y distribuciones de frecuencias semanales y acumuladas, el coste del soporte y un índice del impacto actual que va a tener el esquema temporal de inserciones sobre la población objetivo.

La limitación fundamental de estos modelos, se deriva del hecho de que la muestra simulada debería ser muy amplia para que fuera significativa en planes de medios dirigidos a poblaciones objetivo muy segmentadas (Cannon, 1988). Asimismo, los modelos de simulación son, fundamentalmente, descriptivos e incapaces de encontrar una solución óptima al problema de planificación de medios. Su filosofía se basa, como ya se ha indicado, en introducir el menor número posible de simplificaciones de la realidad, lo cual les lleva a reducir al máximo el uso de fórmulas necesarias en cualquier procedimiento optimizador. Esto hace que el mejor uso de los modelos de simulación sea, más bien, complementar los programas basados en la optimización que sustituirlos (Lilien y Kotler, 1983).

El último enfoque para resolver el problema de planificación de medios, procede de los *modelos heurísticos* que intentan, no imponiendo restricciones tan rígidas como los modelos basados en programación matemática, abordar de una manera más realista la selección de medios. En general (Leckenby y Ju, 1990) estos modelos proceden de una manera secuencial partiendo de un soporte dado, y

añadiendo a continuación aquel soporte que proporciona mayor incremento en la utilidad del dinero invertido¹⁵.

Aunque son muchos los trabajos que se han desarrollado bajo el enfoque heurístico (Brown, 1967; Kotler, 1964, Lee y Burkart, 1960; Lee, 1962; 1963), las principales contribuciones cabe buscarlas en el modelo MEDIAC desarrollado por Little y Lodish (1969), el ADMOD de Aaker (1975) y el VIDEAC de Rust (1985). Otros modelos menos conocidos, son el MISER (Boyd, Claycamp y McClelland, 1970) para anunciantes industriales y el BMH-SOLEM (Bimm y Millman, 1978) desarrollado para la planificación de televisión en Canadá.

Como grandes contribuciones de algunos de estos modelos, puede señalarse que el MEDIAC considera segmentos de mercado, incluyendo ventas potenciales y probabilidades de exposición en cada uno de ellos, costes de los medios (aún suponiendo que son lineales) y los individuos están sujetos a olvido. Además, considera que la respuesta de estos individuos aumenta con las exposiciones pero a tasas decrecientes. En extensiones posteriores (Lodish, 1971), el modelo fue adaptado para que considerara el efecto de los planes de los competidores, siendo el único modelo conocido que admite esta posibilidad (Wilson y Machleit, 1985). Algunos autores (Lilien y Kotler, 1983) señalan que la sofisticación del modelo ha contribuido, probablemente, a lo limitado de su uso.

El ADMOD destaca porque, además de realizar la planificación de medios, permite adoptar la decisión presupuestaria y considerar la creatividad. Aunque normalmente estas decisiones recaen sobre diferentes personas en la agencia, ya se indicó extensamente con anterioridad la profunda interrelación existente entre ellas. Entre las limitaciones del modelo se señalan (Wilson y Machleit, 1985) que es muy

¹⁵ Este procedimiento es tan similar a los modelos iterativos descritos con anterioridad, que Leckenby y Ju (1990) no distinguen entre su trabajo entre unos y otros. Nosotros hemos mantenido la distinción siguiendo el criterio de Wilson y Machleit (1985) y Bigné (1991b).

exigente en cuanto a los requerimientos de datos iniciales. Necesita una estimación de la función de respuesta para diferentes decisiones creativas. El autor propone recurrir a la técnica de laboratorio propuesta por Ray y Sawyer (1971) para esta estimación, pero ello requeriría una investigación para cada segmento de mercado y aspecto creativo considerado lo que, evidentemente, sería muy costoso en términos económicos.

Finalmente el VIDEAC destaca, fundamentalmente, por estar diseñado para la realización de planes de medios en la televisión aunque, conceptualmente, es muy parecido al MEDIAC y puede ser considerado una extensión del mismo (Leckenby y Ju, 1990). Uno de los mayores avances que presenta es que incorpora la distribución beta binomial en su algoritmo y el concepto de frecuencia efectiva en lugar del de exposición al soporte del MEDIAC. Por el contrario, no considera aspectos tales como estacionalidad de la audiencia, descuentos en la compra de medios y el esquema temporal de las inserciones. El cuadro 1.12, resume los principales modelos sobre planificación de medios reseñados en este epígrafe.

A modo de conclusión, Leckenby y Ju (1990) ponen de manifiesto el descenso gradual en el desarrollo de modelos de planificación de medios que se ha venido produciendo desde su inicio en los años sesenta. Atribuyen este hecho a una serie de factores, entre los que cabe destacar el efecto *intercambio*¹⁶ que se produce entre la complejidad creciente de los modelos, sin la cual no se puede lograr resultados ajustados a la realidad y, por otro lado, la necesidad de que los modelos sean comprensibles y fáciles de utilizar para que sean incorporados por las agencias.

Otra de las razones argüidas en esta línea es que muchos de los modelos reseñados, requieren de la atribución de ponderaciones subjetivas por parte de los planificadores, los cuales ni se consideran capaces de establecerlos adecuadamente,

¹⁶ Traducción dada al efecto *trade off*

ni querrían que fuera necesario hacerlo (Simon y Thiel, 1980; Wilson y Machleit, 1985).

Cuadro 1.12. Principales modelos de planificación de medios

Filosofía	Técnica	Modelo
Optimización	Programación lineal	Miller y Starr (1960) Broadbent (1962) Day (1962) Engel y Warshaw (1964) Brown y Warshaw (1965) Stasch (1965) Bass y Lonsdale (1966; 1967) Charnes, Cooper, Lerner y Snow (1968) Charnes, Cooper, Devoe, Lerner y Reinecke (1968) Steuer y Oliver (1976)
	Programación entera	Zangwill (1965) Zufryden (1973; 1975b) Locander, Scamell, Sparkman y Burton (1978) Keown y Duncan (1979)
	Iterativos	HAMM Mather Mediaplan Brown (1967) POMSIS (Aaker, 1968)
	Programación dinámica	Maffei (1960) MEDIAC (Little y Lodish, 1966)
Evaluación	Simulación	SCAL Media Planex (Agostini y Steinberg, 1965) CAM (Beale, Hugues y Broadbent, 1966) ADMESIM (Gensch, 1969) Friedman (1970)
	Heurísticos	MEDIAC (Little y Lodish, 1969) MISER (Boyd, Claycamp, McClelland, 1970) ADMOD (Aaker, 1975) BMH-SOLEM (Bimm y Milman, 1978) VIDEAC (Rust, 1985)

Fuente: Elaboración propia a partir de Leckenby y Ju (1990) y Bigné (1990b;1991b)

En este apartado, se ha ofrecido una visión general de la evolución de la modelización en la literatura sobre planificación de medios. Tras ponerse de manifiesto que el peso tanto cualitativo como cuantitativo de los trabajos sobre modelos es muy importante en la misma, se han revisado las principales aportaciones en las dos grandes áreas más vinculadas con el contenido de esta tesis: los modelos de exposición y los modelos de planificación de medios. No se ha pretendido, en ningún caso, describir con profundidad cada uno de ellos, dado que los que son objeto del análisis específico de este trabajo, los modelos de exposición, lo serán en el capítulo tercero. Respecto a los modelos de planificación de medios se han abordado solamente en cuanto que requieren como *input* información procedente de los modelos de exposición y, por ello, se ha optado por comentar los principales en cada uno de los enfoques empleados en su construcción: programación matemática, simulación y modelos heurísticos.

1.5 LA MODELIZACIÓN EN LA PRÁCTICA DE LA PLANIFICACIÓN DE MEDIOS PUBLICITARIOS

Como se ha indicado con anterioridad, la planificación de medios tienen por misión la asignación eficiente de recursos limitados: los presupuestos publicitarios, entre usos alternativos: los medios y soportes publicitarios. La relevancia de que esta asignación sea efectivamente eficiente, se pone de manifiesto al analizar el volumen de los recursos a los que estamos haciendo referencia. La cifra de 608.686 millones de pesetas de inversión publicitaria en medios convencionales en el ejercicio 1996 en España (Infoadex, 1997) es suficientemente elocuente por sí misma, incluso cuando se la compara con la de los países de nuestro entorno, puesto que está entre las más elevadas de los países de nuestro entorno (cuadro 1.13).

La creación y evaluación de herramientas que permitan la adecuada inversión de los presupuestos publicitarios de las empresas debe convertirse, así, en un objetivo importante en el ámbito académico, hecho que, en cierta forma, justifica la elaboración de trabajos como el presente. Sin embargo, el desarrollo por los investigadores de los modelos expuestos en páginas precedentes, muchos de ellos de elevada complejidad conceptual, carecería de sentido alguno si los individuos que deben proceder a dicha inversión, no hiciesen uso de los mismos. Esta es la cuestión que se pretende acometer en este epígrafe.

No son muchos, sin embargo, los estudios que abordan esta cuestión fundamental. Los más destacados, como se vio en el epígrafe 1.4.1 son los realizados, con una década de diferencia, para el caso estadounidense por el profesor Leckenby (Leckenby y Kishi, 1982a; Leckenby y Kim, 1994a).

Cuadro 1.13. La inversión publicitaria en Europa: 1994.

PAÍS	INVERSIÓN			
	Millones de ECUS	ECUS per capita	% del PIB	Ordenación Europa según inversión total
Alemania	15.474	190,2	0,90	1
Austria	1.249	155,8	0,75	8
Bélgica	1.143	113,2	0,59	9
Dinamarca	995	191,5	0,80	10
España	3.569	91,2	0,88	5
Finlandia	703	138,4	0,85	12
Francia	7.383	127,8	0,66	3
Grecia	815	78,3	1,01	11
Holanda	2.515	163,9	0,90	6
Irlanda	389	109,0	0,88	14
Italia	4.208	73,6	0,49	4
Portugal	594	60,1	0,80	13
Reino Unido	9.693	166,3	1,13	2
Suecia	1.332	152,3	0,80	7
EE.UU.	70.286	272,6	1,24	--
Japón	28.890	231,7	0,80	--
Total Unión Europea	50.062	129,4	0,82	--

Fuente: Eurostat y DGIII (1997) pp. 25-6/7/8

Estos trabajos se basaron en cuestionarios enviados a las doscientas agencias de publicidad norteamericanas más importantes, de acuerdo con sus cifras de facturación. Mientras el primero de los estudios basó sus resultados en 91 cuestionarios válidos, el segundo lo hizo en base a 66. De sus conclusiones, se desprende que la cobertura y la distribución de frecuencias, son efectivamente dos de los factores más utilizados por las agencias de publicidad a la hora de realizar sus planes de medios. Como muestra el cuadro 1.14 aunque ha perdido peso, la cobertura es la variable más empleada (93% de las agencias en 1982, 81% en 1994). Algo similar ocurre con la distribución de frecuencias que pasa de un 84% en 1982 a un 74% en 1994. La relevancia de estas variables para la planificación, lleva a los usuarios a recurrir mayoritariamente a algún modelo para su estimación (cuadro 1.15). Así, en 1982 solamente el 8,8% de las agencias declaraba no emplear modelo

alguno para obtener la cobertura, elevándose esta cifra al 15,4% para la distribución de frecuencias.

Cuadro 1.14. Factores utilizados en la planificación de medios y su importancia.

FACTORES	UTILIZACIÓN % de agencias		IMPORTANCIA Media*	
	1.982	1.994	1.982	1.994
Cobertura	93,4	81,0	4,4	4,5
CPM	89,0	77,8	4,5	4,2
Factores cualitativos	87,9	77,8	4,4	4,1
Coste total	85,7	77,8	4,3	4,3
Distribución de frecuencias	84,6	74,6	4,0	4,2
GRPS	91,2	73,0	4,2	4,1
OTS	90,1	73,0	4,2	4,0
Cobertura efectiva	59,3	68,3	4,4	4,5
Continuidad	70,3	65,1	3,8	3,8
Análisis por quintiles	63,7	58,7	3,0	3,3
Distribución del presupuesto por tipo de medio	53,8	50,8	3,8	3,4
Otros	9,9	11,1	4,0	4,4

* 1 = Nada importante; 5 = Muy importante

Fuente: Leckenby y Kishi (1982a); Leckenby y Kim (1994a)

**Cuadro 1.15. Número de modelos utilizados en la planificación de medios.
Porcentajes verticales sobre las agencias**

NÚMERO	COBERTURA		DISTRIBUCIÓN DE FRECUENCIAS	
	1.982	1.994	1.982	1.994
Ninguno	8,8	17,5	15,4	22,2
Uno	25,3	36,5	35,2	36,5
Dos	36,3	23,8	29,7	23,8
Tres	17,6	15,9	11,0	12,7
Cuatro	8,8	5,0	6,6	3,2
Cinco	2,2	0,0	1,1	0,0
Seis o más	1,1	1,6	1,1	1,6
Total	100	100	100	100

Fuente: Leckenby y Kishi (1982a); Leckenby y Kim (1994a)

Llama poderosamente la atención el hecho de que, en una década, se ha incrementado significativamente el porcentaje de agencias de publicidad que, en Estados Unidos, no recurren a modelo alguno para la estimación de la cobertura (17,5%) y la distribución de frecuencias (22,2%). Dado que, como se ha visto con anterioridad, esta reacción no puede ser debida a que estas variables no se consideren importantes, Leckenby y Kim (1994a) explican este fenómeno basándose en la creencia de que, al haberse incrementado notablemente el número de modelos informatizados existentes y, paralelamente su complejidad, resulta cada vez más difícil para los planificadores comprender su lógica interna, hecho que les lleva a dudar de la precisión de los métodos por ellos empleados. Esta situación lleva a Leckenby y Kim (1994a) a propugnar que la *Advertising Research Foundation* establezca un comité que lleve a cabo auditorías objetivas del funcionamiento de los modelos cuyos algoritmos no sean públicos (modelos de “propietario” o de “caja negra”).

Entrando en qué modelos en concreto son los más utilizados entre las agencias estadounidenses (cuadro 1.16), mientras que en 1982 lo eran aquellos proporcionados por la empresa Telmar (que ofrecía el MEDIAC y otros programas específicos), tanto en la estimación de la cobertura, como de la distribución de contactos, una década después, *Interactive Market Systems* (IMS) se ha convertido en el principal proveedor en ambos campos.

Si se presta atención a los modelos clásicos, el de Metheringham es el más utilizado. En este sentido, si tenemos en cuenta que, como ya se explicó, el modelo de Metheringham es un modelo basado en la distribución beta binomial, al igual que el modelo de la matriz beta, los métodos que toman como punto de partida para su construcción a esta distribución, conservan un peso muy importante en la planificación de medios: 30% en la estimación de la cobertura y 31,5% en la distribución de frecuencias.

Cuadro 1.16. Modelos utilizados en la estimación de la frecuencia y cobertura.
Porcentajes verticales

MODELOS	COBERTURA		DISTRIBUCIÓN DE FRECUENCIAS	
	1.982	1.994	1.982	1.994
IMS (aparte del MODAL)	9,6	26,0	10,5	25,0
Telmar (aparte del MEDIAC)	20,9	18,0	23,0	18,5
Metheringham	16,6	15,0	13,2	13,0
Beta Binomial	7,0	12,0	11,2	15,2
MODAL II	12,3	9,0	11,8	8,7
MEDIAC	7,0	2,0	7,9	2,2
Simulación	3,7	4,0	3,3	3,3
Matriz Beta	3,7	3,0	3,3	3,3
Sainsbury	3,7	0,0	0,7	0,0
Binomial negativa	2,1	1,0	2,6	1,1
Marketmath	1,1	0,0	1,3	0,0
Morgensztern	1,1	0,0	0,0	0,0
Agostini	0,5	0,0	0,7	0,0
Casualidad Constante	0,5	1,0	0,0	0,0
Kwerel	0,0	0,0	0,0	0,0
Hofmans	0,0	0,0	0,0	0,0
Multinomial Dirichlet	0,0	0,0	0,0	0,0
Loglinear	--	0,0	--	0,0
Expansión Canónica	--	0,0	--	0,0
Otros	10,2	9,0	10,5	9,8
Total	100	100	100	100

Fuente: Leckenby y Kishi (1982a); Leckenby y Kim (1994a)

Mucho más complejo resulta afirmar algo acerca del peso que tiene la modelización en la práctica de la planificación de medios en España. Redondo (1995), basándose en cifras generales de inversión y concentración de la misma, hipotetiza que es difícil que las campañas de aquellos anunciantes que en el *ranking* elaborado en base al volumen de inversión, ocupan lugares por encima del 250 puedan estar utilizando estos programas cuando la inversión media que realizan es de 1,75 millones de pesetas en diarios y de 3,19 millones en revistas. De ser así, este hecho habría dejado en 1994 fuera de la asignación mediante programas de planificación de medios a un volumen de inversión superior a los 45.000 millones de pesetas en estos dos medios.

Basándose en razonamientos similares, afirma que si según la AIMC en 1991 había 41 agencias asociadas y, por ello, en condiciones de acceder al fichero individualizado del Estudio General de Medios, mientras que el número de agencias rondaba las mil, es evidente que una gran cantidad de ellas no está recurriendo a los procedimientos aquí descritos¹⁷.

El único estudio *ad hoc* realizado en España sobre el comportamiento de las agencias de publicidad respecto a la planificación de medios del que tenemos conocimiento, es el realizado por Bigné (1989; 1990c). Los resultados del mismo (cuadro 1.17), parecen confirmar empíricamente las impresiones manifestadas por Redondo (1995). Así, se comprueba que el 21,8% de las agencias declaran utilizar siempre o casi siempre la intuición en los planes de medios que realizan. Asimismo, el autor considera que el elevado porcentaje de agencias que declaran utilizar modelos propios (46,5% siempre o casi siempre) se debe a una incorrecta identificación del concepto de modelo, refiriéndose más bien a un método de trabajo o una forma de planificar, con lo que estarían, probablemente, en el grupo anterior. Las agencias que sí utilizan programas comercialmente conocidos recurren fundamentalmente a los programas TOM comercializados por ODEC (55,1% siempre o casi siempre) y, en menor medida, los que en aquel momento proporcionaba ECOTEL (17,9%). Posteriormente, este mismo autor (Bigné, 1990d), estableció una tipología de las agencias de publicidad españolas basándose, precisamente, en el interés de las mismas por diversos temas de investigación relacionados con la planificación de medios.

De sus conclusiones, se desprende que un 11,6% son agencias selectivas e interesadas en la investigación de audiencias y un 28,2% tienen un carácter generalista y también con un gran interés por estos temas. Por contra, establece que

¹⁷ Aunque este razonamiento pone de manifiesto lo que, desde nuestro punto de vista es una realidad, que muchísimas agencias pequeñas no recurren a programa alguno, no deja de ser cierto que no es necesario pertenecer a la AIMC para poder utilizar los programas comerciales más habituales, basta con encargar la compra de medios a una central.

la mayoría de ellas (42,4%) están interesadas sólo de manera muy general por estos temas, no poseyendo intereses muy marcados en ninguno de ellos y, sobre todo, obtiene que un 8,9% de las agencias demostraban escaso interés por la investigación en la mayoría de las cuestiones analizadas.

Cuadro 1.17. Número de modelos utilizados en la planificación de medios
Porcentajes horizontales

MÉTODO	Frecuencia de utilización				
	Siempre	Casi Siempre	A Veces	Nunca	NS/NC
Intuición	7,7	14,1	44,9	21,8	11,5
Modelos propios	16,7	29,5	24,4	16,7	12,8
ODEC-UNITEC	25,6	29,5	21,8	12,8	10,3
ECOTEL	6,4	11,5	19,2	42,3	20,5

Fuente: Bigné (1989)

En este epígrafe hemos tratado de comprobar el uso que se hace en la práctica de los distintos modelos de planificación de medios dado que, como se ha señalado, parece lógico que el elevado volumen de inversión en publicidad que se realiza hoy en día, esté guiado por criterios de eficiencia, muchos de los cuales vendrán proporcionados por estos modelos. Se ha puesto de manifiesto que, aunque el uso de estos métodos es mayoritario, tanto en el caso estadounidense como en el español, todavía están muy lejos de ser empleados por todas las agencias, habiéndose apreciado, incluso, una regresión en su utilización en la última década en los Estados Unidos.

2

Tipología de los modelos de exposición

2.1 INTRODUCCIÓN

El objetivo de este capítulo es doble. En primer lugar, se pretende realizar una tipología detallada de los modelos de exposición. Así, mientras en el capítulo anterior nuestro interés era, únicamente, aportar una visión cronológica de la aparición de los mismos, sin entrar en las hipótesis que los sustentan y sin que por ello, la clasificación resultante tuviera otra misión que la de permitir al lector tener una perspectiva comprensiva de las principales aportaciones que el mundo académico ha ofrecido en este campo y el momento del tiempo en que se han producido.

Esta visión de conjunto, necesaria en cuanto que resulta útil para ofrecer la necesaria perspectiva en unas pocas páginas, deviene en insuficiente cuando se trata de elaborar una clasificación que, además de ser exhaustiva, añada valor al conocimiento de los modelos analizados, al tomar como herramienta fundamental de su construcción, las hipótesis que subyacen en los mismos.

En segundo lugar, el objetivo de este capítulo reside en aproximarnos progresivamente al objeto de esta tesis doctoral. Como se verá inmediatamente, los modelos de exposición responden a dos grandes enfoques (Lamas, 1993). Por un lado están aquellos tradicionalmente llamados modelos “fórmula”, que nosotros denominaremos *modelos de agregación*, caracterizados por estimar la cobertura y la distribución de frecuencias tomando como base datos globales del plan de medios: audiencia de los distintos soportes, duplicaciones entre ellos, acumulación de las audiencias entre distintas inserciones, etc. y derivando por cálculo de probabilidades una expresión, de ahí su denominación de fórmula, para la estimación de la cobertura y distribución de contactos. El término “agregación”, pretende indicar que los datos de audiencia y duplicaciones de soportes con los que se trabaja, proceden de los estudios de audiencia, tras haber agregado éstos la información que les proporcionan los individuos entrevistados. Hemos considerado, con Redondo (1996), que el término “fórmula” es equívoco, en cuanto que parece indicar que los demás no las utilizan. El segundo enfoque da lugar a los *modelos de individuación* que, a diferencia de los anteriores, estiman la cobertura y la distribución de contactos, tratando la información individuo a individuo, mediante las probabilidades de exposición de los mismos proporcionadas por los estudios de audimetría, tal y como se expondrá seguidamente.

La presente tesis doctoral se centra en los modelos del primer enfoque, esto es, los modelos de agregación. Varias son las razones en las que se apoya esta decisión. La primera es de índole práctico. La mayoría de los modelos de individuación, son modelos de propietario utilizados en diversos programas

comerciales de planificación de medios y, por ello, con procedimientos de cálculo no siempre accesibles a la evaluación externa. Si bien es cierto que los más utilizados en España (el TOM de ODEC y el MARKSEL de SOFRES) son transparentes en este sentido¹, no ocurre lo mismo con otros programas de amplia utilización en el extranjero como el MODAL de IMS o el MEDIAC de Telmar. La introducción de estos modelos en nuestro análisis, provocaría una asimetría evidente: mientras los modelos de agregación son exhaustivamente revisados, desde los trabajos seminales, hasta las últimas aportaciones en este campo, los modelos de individuación quedarían reducidos a los disponibles en el caso español, dejando fuera del análisis a la mayoría de los existentes.

La segunda razón es de índole metodológica. Martínez Ramos (1992) señala acertadamente que hay que ser conscientes de que los modelos de agregación son el punto de partida, sino el elemento central, de los modelos de individuación. Como se verá muy pronto, la distribución de contactos de un plan de medios para un individuo, independientemente de cómo se agregue para el conjunto de la población objetivo, debe ser aproximada por una distribución de probabilidad, generalmente la binomial, según un modelo de agregación.

Pese a todo lo expuesto, consideramos que supondría una innecesaria limitación de este trabajo no tratar en modo alguno los modelos de individuación. Por este motivo el presente capítulo se ha estructurado en dos epígrafes. En el primero de ellos se hará una revisión de los principales enfoques en el desarrollo de estos modelos, haciendo énfasis en uno de los elementos fundamentales para su aplicación: el proceso de probabilización de los datos de los paneles de audiencias.

En el segundo de los epígrafes, se llevará a cabo una clasificación de los modelos de agregación que van a ser descritos y evaluados en capítulos posteriores.

¹ Puede verse el artículo ya reseñado de Lamas (1993) para una explicación del cálculo de la cobertura y la distribución de contactos en los mismos.

Se atenderá, para ello, a las hipótesis que subyacen en los mismos, fundamentalmente las referentes a la estacionariedad o no de las probabilidades de exposición respecto a los individuos o a las inserciones, homogeneidad o heterogeneidad de los individuos respecto a las probabilidades de exposición, homogeneidad o heterogeneidad de los soportes respecto a la probabilidad de exposición y, finalmente, las hipótesis concernientes al carácter aleatorio o no de los fenómenos de la duplicación y acumulación de las audiencias.

2.2 MODELOS DE INDIVIDUACIÓN

2.2.1 Procedimientos de probabilización

Como se ha señalado, los modelos de individuación estiman la cobertura y la distribución de contactos de un plan de medios a partir de las probabilidades que tiene cada individuo de la muestra que conforma el estudio de audiencias de resultar expuesto al soporte. Lo que distingue unos modelos de otros, es el procedimiento que emplean para pasar de las probabilidades individuales de exposición, a las probabilidades de la población objetivo de resultar expuesta a ese soporte. Por consiguiente, es imperativo antes de describir los principales modelos, señalar cuáles son los métodos habituales de obtención de esas probabilidades individuales en los diferentes estudios de audiencias, ya que son el *input* fundamental de aquellos.

Los estudios de audiencias pueden ser clasificados atendiendo a diversos criterios (Bigné, 1994): método de investigación empleado, amplitud de la investigación, ámbito espacial del estudio, momento del tiempo en que se realiza la investigación o ámbito subjetivo del estudio. Según el primer criterio, nos encontramos, fundamentalmente, con dos tipos: los estudios continuos de tipo panel y los estudios periódicos basados en la encuesta. Los primeros se han empleado tradicionalmente para medir las audiencias en televisión, apoyándose básicamente en

instrumentos como el audímetro², mientras que los segundos son aplicados principalmente al medio impreso. En España, la empresa SOFRES, A. M. mantiene un panel de audímetros, mientras que la Asociación para la Investigación de Medios de Comunicación (AIMC), realiza cuatrimestralmente el Estudio General de Medios (EGM). A efectos de descripción del procedimiento de probabilización, la estructura de ambos es similar y viene recogida en la figura 2.1. Mientras que en el panel de audímetros, los periodos serían espacios cortos de tiempo³, en el EGM son las distintas oleadas cuatrimestrales. Asimismo, la exposición de un individuo i al soporte j en el periodo k sería un suceso dicotómico en el EGM (1 = hay exposición; 0 = no hay exposición), mientras que en el panel de audímetros vendría indicado el tiempo de contacto.

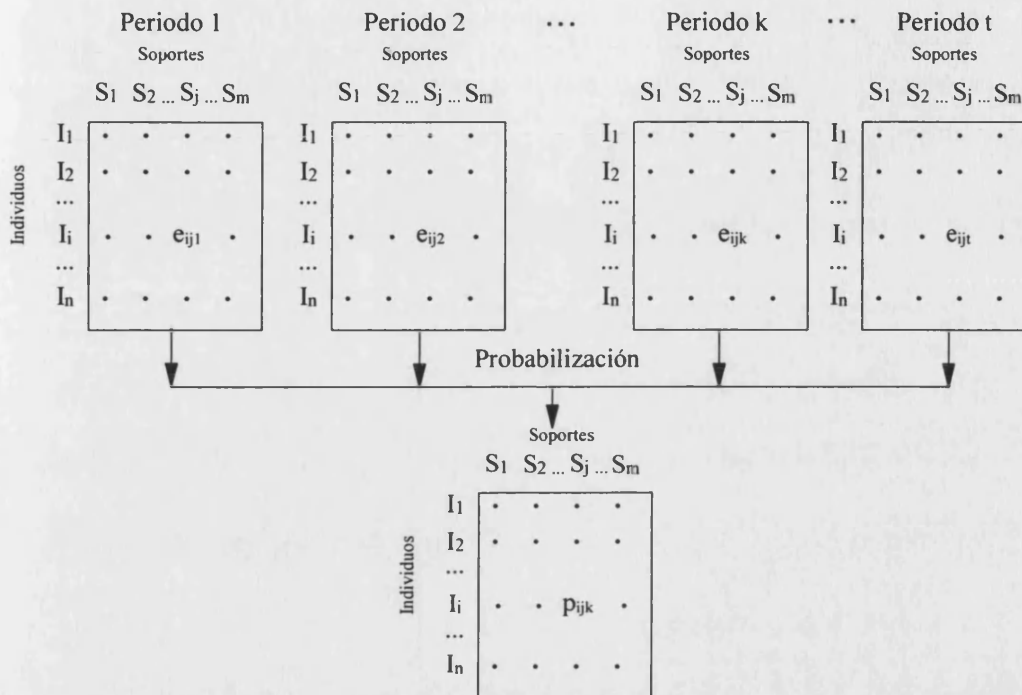
Seguiremos a Lamas (1993) para ilustrar el proceso de probabilización en un estudio continuo. El audímetro proporciona, como se ha indicado, el tiempo que el individuo i ha estado expuesto al soporte j (v. g.: un programa semanal determinado) en cada uno de los periodos considerados. Supóngase que se pretende determinar la probabilidad de exposición de un individuo al primer cuarto de hora del programa (periodo k). Para ello se decide, en primer lugar, la cantidad de datos pasados que quieren utilizarse para su cálculo. Así, en el ejemplo del cuadro 2.1, tomando un mes, la probabilidad de contacto p_{ijk} se obtendrá:

$$P_{ijk} = \frac{\text{Tiempo de contacto real con el soporte}}{\text{Tiempo máximo potencial de contacto con el soporte}}$$

² Si bien es posible su extensión otros medios, configurándose como un instrumento de recogida de audiencias multimedia. Véase Bigné (1991d).

³ El panel de audimetría de Sofres, A. M. proporciona diariamente los resultados de audiencia del día anterior, y éstos se ofrecen: minuto a minuto, por cuarto de hora, por programas, por bloques, por *spot* y las audiencias acumuladas y medias del total de día y por franjas horarias (SOFRES, A.M., 1995).

Figura 2.1. Proceso de probabilización en los estudios de audiencias



Fuente: Redondo (1997)

El cuadro 2.1 ejemplifica el procedimiento de cálculo para dos individuos, uno de los cuales ha formado parte de la muestra útil⁴ durante los cuatro programas que se están utilizando para el cálculo de la probabilidad, mientras que el segundo fue excluido en los dos primeros programas, dado que Sofres A.M. realiza controles periódicos sobre la muestra para asegurarse de la adecuada colaboración de los individuos. Se controlan especialmente los niveles de audiencia extremos, los individuos sin audiencia y, al menos una vez al año, se hace un control coincidental telefónico para verificar la correcta utilización del mando a distancia en la declaración de audiencia individual. El individuo 2 puede haber resultado excluido en esos dos periodos debido, por ejemplo, a un mal funcionamiento del audímetro.

⁴ La audiencia útil es aquella formada por los individuos no excluidos por los diversos controles que realiza la empresa SOFRES A.M.

Cuadro 2.1. Ejemplo de probabilización en estudios continuos

	Individuo 1		Individuo 2	
	En muestra útil	Tiempo de contacto	En muestra útil	Tiempo de contacto
Semana 1	Sí	10 minutos	No	--
Semana 2	Sí	0 minutos	No	--
Semana 3	Sí	2 minutos	Sí	3 minutos
Semana 4	Sí	15 minutos	Sí	0 minutos
Tiempo de contacto	27 minutos		3 minutos	
Tiempo máx. potencial	60 minutos		30 minutos	
Probabilidad de contacto	0,45		0,1	

Fuente: Lamas (1993; pág. 59)

Como toda predicción de un comportamiento futuro en base al pasado, este método está asumiendo una serie de hipótesis. En primer lugar supone que los comportamientos de los individuos son estables en el tiempo, lo cual en el corto plazo no es demasiado restrictivo. Pero en segundo lugar, está considerando que la probabilidad de que un individuo contacte con un soporte no está condicionada a la probabilidad de que contacte con otro. Esta asunción sí es mucho más irreal, dado que las diversas cadenas de televisión tienen programas con contenidos similares y, si a un individuo le resulta atractivo un determinado programa que ve en una de ellas, probablemente busque otro similar en el resto de cadenas. En todo caso, Lamas (1993) considera que, a efectos prácticos, la aproximación a la realidad es suficiente para justificar la aceptación de este procedimiento.

En los estudios periódicos para los medios impresos, el proceso es distinto. En primer lugar debe determinarse el número de lectores medio de un número (AIR: *Average Issue Readership*). Para obtener dicho valor, existen fundamentalmente tres procedimientos para este fin (Lamas, 1997; Martínez Ramos, 1992): el método “*Through The Book*” (TTB), el método de “la primera lectura ayer” (*First Readership Yesterday*: FRY) y el método de la “lectura del último periodo” (*Recent Reading*).

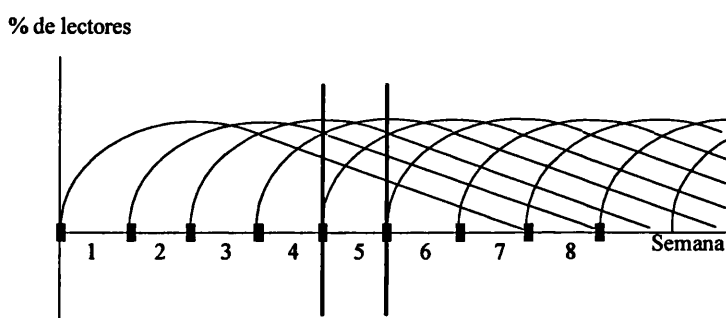
El método TTB está basado en el reconocimiento de un *número específico* de la publicación cuya audiencia se está midiendo, y fue introducido en los años 30 en los Estados Unidos. El entrevistador mostraba diversos números de la revista y preguntaba al entrevistado, ante cada uno de ellos, si lo había leído. Este procedimiento ha sufrido, a lo largo del tiempo, diversas modificaciones. Cuando el número de soportes objeto de análisis fue creciendo, al entrevistador le era imposible llevar números completos de cada uno, por lo que se construían “esqueletos” de los ejemplares con sólo una proporción del total de páginas. Uno de las decisiones críticas de este procedimiento, era fijar la antigüedad de los ejemplares que se muestran, dado que si el número es muy reciente, la medida no recoge la audiencia a lo largo de la vida del mismo (existen muchas posibilidades de que todavía se siga leyendo en el futuro) y si es demasiado antiguo, se corre el riesgo de que el olvido infraestime el reconocimiento del mismo. Lo habitual era enseñar ejemplares de 4 a 5 semanas para las revistas semanales y de 10 a 12 semanas para las mensuales.

En 1995, el Simmons Market Research Bureau (SMRB), que desarrollaba la encuesta en Estados Unidos, renunció a esta metodología adoptando la técnica del *Recent Reading*, debido a los grandes inconvenientes que proporciona este procedimiento: elevados costes, longitud de las entrevistas, cansancio del entrevistado, y dificultad para mantener los niveles de respuesta. En estos momentos su uso está limitado a Canadá (Lamas, 1997).

La idea básica del procedimiento de *Recent Reading* (Martínez Ramos, 1992; pág. 243) es que “el número de personas que entran en contacto con un número cualquiera de una revista, a lo largo de la vida de dicho número (la cual puede durar varias semanas hasta que agota la casi totalidad de sus lectores) es igual al número de personas que en una semana (si estamos hablando de una revista semanal, igual sucedería para las revistas mensuales o quincenales) entran en contacto con cualquier número de dicha revista”. Resulta evidente que, dado que un número cualquiera de una revista (figura 2.2) puede permanecer vivo varias semanas, sería imposible

mediante una encuesta conocer cuántas personas entran en contacto con ese número a lo largo de su vida. Pero si esa cantidad de personas, como se hipotetiza, es la misma que entran en contacto en una semana determinada con un número medio de la revista, ya no es necesario conocer el comportamiento de la población a lo largo de varias semanas, basta conocerlo en una, lo que es accesible mediante encuesta.

Figura 2.2. Supuesto básico del método “lectura del último periodo”



Fuente: Martínez Ramos (1992; pág. 244)

Veamos, siguiendo la descripción realizada por González y Carretero (1997), cómo se calcula la probabilidad de exposición de un individuo determinado, a partir de las preguntas realizadas en la encuesta que, en este ejemplo, es la del Estudio General de Medios. Tomando como ejemplo los periódicos, el EGM realiza a cada entrevistado tres preguntas que recogen sucesivamente el hábito de lectura, la frecuencia y la audiencia del último periodo (Méndez, 1996):

- a. ¿Qué periódicos - diarios ha leído u hojeado en los últimos 30 días?.
- b. *Para cada periódico leído en los últimos 30 días:* De cada siete ejemplares de (el periódico X) que salen en una semana, ¿Cuántos acostumbra Vd. a leer u hojear.

- c. Sin contar el día de hoy ¿Cuándo ha leído por última vez un ejemplar del periódico X?

Mediante la combinación de las preguntas *a* y *b*, se obtiene el número de personas que han sido audiencia *alguna vez* a lo largo de la semana de uno, dos, tres, hasta siete ejemplares: bastará obtener la distribución de frecuencias de la pregunta *b* para cada uno de los soportes enunciados en *a*. Mediante la pregunta *c* se obtiene la audiencia *del último periodo*. La segunda pregunta, como muestra el cuadro 2.2, proporcionaría la probabilidad teórica de exposición a una inserción en el periódico X para cada individuo, según los ejemplares que haya declarado leer. Basta dividir los casos favorables (número de ejemplares que declara leer) por los casos posibles (7 ejemplares semanales).

Cuadro 2.2. Ejemplo de probabilización en los estudios periódicos.

Ejemplares leídos. Frecuencia	Audiencia alguna vez	Audiencia último periodo	Probabilidad teórica	Probabilidad asignada
0	30.000	2.000	0/7 = 0,00	2.000/30.000 = 0,07
1	60.000	6.000	1/7 = 0,14	6.000/60.000 = 0,10
2	78.000	12.480	2/7 = 0,29	12.480/78.000 = 0,16
3	69.000	20.010	3/7 = 0,43	20.010/69.000 = 0,29
4	53.000	22.260	4/7 = 0,57	22.260/53.000 = 0,42
5	84.000	50.400	5/7 = 0,71	50.400/84.000 = 0,60
6	70.000	55.300	6/7 = 0,86	55.300/70.000 = 0,79
7	86.000	81.700	7/7 = 1,00	81.700/86.000 = 0,95

Fuente: González y Carrero (1997; pp. 161-162)

El EGM, sin embargo, según el método que estamos describiendo, corrige la frecuencia de lectura declarada por la del último periodo que recoge la pregunta *c*. La probabilidad asignada a cada individuo, según el número de ejemplares que ha declarado leer, se calcula dividiendo la audiencia del último periodo por la audiencia de alguna vez como muestra el cuadro 2.2.

Con las probabilidades individuales así calculadas, los modelos de individuación se distinguen unos de otros en función del procedimiento que siguen para agregar esas probabilidades ante un plan de medios dado. Según Lamas (1993) nos encontramos ante tres grandes grupos de modelos de individuación que desarrollaremos brevemente a continuación: los modelos de distribución exacta, los modelos basados en simplificaciones del cálculo exacto y los modelos de simulación que se desarrollaran en el epígrafe siguiente.

El método de la “primera lectura ayer” (*First Readership Yesterday: FRY*), es una variante del método de lectura reciente. Con la intención de reducir al mínimo los problemas asociados al recuerdo, se opta por concentrarse en las lecturas que tuvieron lugar el día anterior a la entrevista, pidiendo al entrevistado que indique cuáles de esas lecturas corresponden a ejemplares que se han tenido en las manos por primera vez. Este procedimiento elimina los sesgos procedentes de lecturas replicadas y paralelas (Martínez Ramos, 1990), obteniéndose la estimación de la audiencia multiplicando el valor del “número de lecturas de ayer por primera vez” por el número de días del período de publicación correspondiente. Este método se introdujo en Holanda y Dinamarca en 1982, para ser utilizado en combinación con entrevistas telefónicas asistidas por ordenador. La principal limitación de esta técnica, es que el número de lectores de ayer para un día promedio, es pequeño para muchas publicaciones (Tchaoussoglou, 1995), ha llevado a Holanda a abandonar en 1996 esta técnica.

2.2.2 Modelos de distribución exacta

Estos modelos agregan las probabilidades individuales al total de la población objetivo, mediante la aplicación estricta y estricta del cálculo de probabilidades (Lamas, 1993). Estos modelos presumen que, para cada individuo, la

distribución de contactos se distribuye según una binomial. Suponiendo que el plan estuviera compuesto por un único soporte, esta sería:

$$B(x; p, n) = \frac{n!}{(n-x)! x!} p^x (1-p)^{n-x} \quad [2.1]$$

donde p es la probabilidad de exposición del individuo al soporte en cuestión, obtenida por los procedimientos descritos en el epígrafe anterior, n es el número de inserciones en ese soporte y $B(x)$, por tanto, la probabilidad de tener exactamente x contactos.

Sin embargo, en general el plan estará formado por m soportes, con n_1, n_2, \dots, n_m inserciones en cada uno de ellos y donde, además, las probabilidades de exposición a cada soporte también serán distintas: p_1, p_2, \dots, p_m . Por lo tanto, la probabilidad de que ese individuo aislado resulte expuesto a exactamente x inserciones, se obtendrá por la combinación de las m binomiales resultantes mediante una función multinomial:

$$M(x) = \sum_j \left[\prod_i B_i(x_{ji}; p_i, n_i) \right] \quad [2.2]$$

donde el sumatorio debe extenderse a todas las combinaciones de inserciones en soportes (i inserciones en j soportes = $i \times j$) que sumen la misma cantidad x :

$$\sum_i x_{ji} = x$$

Dicho de otro modo, para cada individuo, es necesario calcular una binomial para cada soporte. A continuación se calcula el producto cartesiano de las mismas previa sustitución en ellas de todas las inserciones que pueden realizarse y, por último, se suman aquellos productos cuyo orden de contactos ($i \times j$) sea el mismo. Veámoslo con el ejemplo ofrecido por Lamas (1993).

Considérese el plan de medios formado por dos soportes. Un individuo concreto tiene una probabilidad de exposición al primero de 0,2 y al segundo de 0,1, según los resultados de un estudio de audiencias. El planificador decide realizar dos inserciones en el primer soporte y una en el segundo. Aplicando la expresión [2.1] a los datos recién expuestos, la distribución simple de las dos binomiales sería:

Cuadro 2.3. Cálculo de las binomiales.

$B_1(x)$	$B_2(x)$
$B_1(0) = 0,64$	$B_2(0) = 0,90$
$B_1(1) = 0,32$	$B_2(1) = 0,10$
$B_1(2) = 0,04$	

Fuente: Lamas (1993; pág. 60)

Las combinaciones posibles (producto cartesiano) de las dos binomiales, vendría dado por:

Cuadro 2.4. Combinaciones de las binomiales.

Contactos 1 ^{er} Soporte	Contactos 2 ^o Soporte	Probabilidad
0	0	$0,64 \times 0,90 = 0,576$
0	1	$0,64 \times 0,10 = 0,064$
1	0	$0,32 \times 0,90 = 0,288$
1	1	$0,32 \times 0,10 = 0,032$
2	0	$0,04 \times 0,90 = 0,036$
2	1	$0,04 \times 0,10 = 0,004$

Fuente: Lamas (1993; pág. 60)

pudiéndose calcular, a partir de [2.2] la distribución de contactos del individuo concreto que estamos analizando, sumando aquellas probabilidades que proceden de un número de contactos en cada soporte que suman, sucesivamente, cero, uno y dos contactos:

Cuadro 2.4. Distribución de contactos.

M(x)	Probabilidad
M(0)	$B_1(0) \times B_2(0) = 0,576$
M(1)	$B_1(0) \times B_2(1) + B_1(1) \times B_2(0) = 0,352$
M(2)	$B_1(1) \times B_2(1) + B_1(2) \times B_2(0) = 0,068$
M(3)	$B_1(2) \times B_2(1) = 0,004$

Fuente: Lamas (1993; pág. 60) y elaboración propia

La probabilidad de que el individuo que estamos analizando, resulte expuesto a al menos, una de las inserciones del plan de medios (probabilidad que derivará posteriormente en la cobertura del plan), vendrá dada por:

$$C = 1 - \prod_{i=1}^m (1 - p_i)^{n_i} \quad [2.3]$$

aunque, evidentemente, puede calcularse también como $1 - M(0) = 0,4240$.

Las probabilidades así calculadas deben, sin embargo, elevarse ahora el total de individuos que conforman la población objetivo del plan de medios. Para ello bastaría la simple adición de los datos individuales ponderados por el factor de elevación del individuo. Este procedimiento implica un número elevadísimo de cálculos que se hace insostenible cuando se incrementa el número de soportes. Lamas (1993) indica que un plan con 20 soportes y 4 inserciones en cada uno de ellos precisaría calcular las probabilidades de 5^{20} (95 billones) de combinaciones posibles para cada uno de los individuos. Esta explosión de cálculo provoca la aparición de simplificaciones del método teórico aquí expuesto.

2.2.3 Simplificaciones del cálculo exacto

2.2.3.1 Reducción a un soporte promedio

Consiste en reducir a un soporte promedio y, por ello, ficticio, todos los soportes del plan. Siguiendo a Lamas (1993), este procedimiento calcula para cada individuo el número de inserciones con probabilidad no nula (n^*) del plan:

$$n^* = \sum_{i=1}^m n_i^* \quad \text{con} \quad \begin{aligned} n_i^* &= n_i \text{ si } p_i \neq 0 \\ n_i^* &= 0_i \text{ si } p_i = 0 \end{aligned}$$

donde n_i es el número de inserciones en cada uno de los m soportes y p_i es la probabilidad individual de exposición del sujeto analizado al soporte i , calculado por los procedimientos anteriormente descritos. De una manera análoga, se calcula la probabilidad promedio de exposición al plan de medios:

$$p = \frac{\sum_{i=1}^m n_i p_i}{n^*} = \frac{\sum_{i=1}^m n_i^* p_i}{n^*}$$

De esta manera, para cada individuo se calcula exclusivamente una binomial $B(x, p, n^*)$ a partir de la cual es inmediato calcular la cobertura y la distribución de contactos como se vio en el procedimiento anterior. A continuación, basta agregar las probabilidades para el conjunto de los individuos ponderados por su factor de elevación.

Este método simplifica de manera espectacular el número de cálculos necesarios. Al reducir el plan de medios a un soporte promedio evita tener que calcular para cada individuo todas las binomiales resultantes del producto cartesiano de las mismas y el número de inserciones en cada soporte, reduciéndolo a una sola

binomial por individuo. Debe notarse, sin embargo, que este procedimiento no proporcionará en modo alguno, los mismos resultados que el anterior. A modo de ejemplo, resulta evidente que, en general:

$$C = 1 - \prod_{i=1}^m (1 - p_i)^{n_i} \neq 1 - (1 - p)^{n^*}$$

Este procedimiento, es “el que utiliza la versión del MARKSEL actualmente en uso en España” (Lamas, 1993; pág. 61; Redondo, 1995; pág. 65).

2.2.3.2 Reducción a un soporte con ajuste de cobertura

Este procedimiento consiste, según Lamas (1993), en ajustar la binomial para que los datos que se obtienen de la cobertura mediante la reducción a un soporte, sean congruentes con los que proporcionaría la distribución exacta. Según lo visto en el apartado 2.2.2, para cada individuo la probabilidad de exposición a alguno de los soportes que componen el plan de medios, se obtendría por [2.3]:

$$C = 1 - \prod_{i=1}^m (1 - p_i)^{n_i}$$

mientras que, si llamamos G a la suma de las probabilidades individuales de exposición a cada soporte, multiplicadas por el número de inserciones que se realiza en cada uno de ellos:

$$G = \sum_{i=1}^m n_i p_i \quad [2.4]$$

entonces resulta claro que ni C ni G se ven afectados si eliminamos, como en el subapartado anterior, los soportes con probabilidades nulas. Calculados C y G para cada individuo, se trata de encontrar una probabilidad promedio p que aplicada a ese

individuo, conserve los valores ya fijados de C y G . En otros términos, como se desprende al analizar el siguiente sistema de ecuaciones, se trata de encontrar una probabilidad p y un factor de corrección L que mantengan los valores precalculados de C y G :

$$\begin{aligned} [1 - (1 - p)^{n^*}] L &= C \\ (n^* p) L &= G \end{aligned}$$

Resolviendo el anterior sistema por reducción, se llega, llamando $q = 1 - p$, a la siguiente ecuación:

$$\frac{n^* C}{G} = 1 + q + q^2 + \dots + q^{n^*-1}$$

que tiene solución única. Calculado así q y por ende p , la binomial se corrige por el factor L , esto es:

$$P(x) = B(x; p, n^*) L$$

siendo el resto del proceso de agregación a la población es análogo a los métodos descritos con anterioridad.

2.2.3.3 Método TOM

Los programas TOM mantienen la expresión [2.3] para el cálculo de la probabilidad de exposición de un individuo a al menos una de las inserciones que conforman el plan de medios. Definimos la variable G como se hizo en la expresión [2.4] del subapartado anterior.

El modelo aplicado por los TOM calcula, para cada individuo j , la *ratio* R_j entre ambas variables:

$$R_j = \frac{G_j}{C_j}$$

ratio que se descompone en su parte entera E y decimal D . Por tanto, obviamente:

$$R_j = E_j + D_j$$

Cada individuo de la muestra del EGM, para el que se ha calculado la *ratio* anterior, tiene un factor de elevación poblacional M_i , esto es, representa a M_i individuos de la población. Esos individuos se dividen en tres grupos, a cada uno de los cuales se les asigna el siguiente número de contactos:

- a. *Primer grupo*. Formado por $(1-C_i) \cdot M_i$ personas, tienen asignados cero contactos.
- b. *Segundo grupo*. Formado por $(1-D_i) \cdot C_i \cdot M_i$ personas, y tienen asignados E_i contactos.
- c. *Tercer grupo*. Formado por $D_i \cdot C_i \cdot M_i$ personas, se les asignan $E_i + 1$ contactos.

La cobertura y la distribución de contactos, se logra sumando los resultados obtenidos para cada individuo. Dada la importancia práctica de este método, puesto que el TOM es el programa empleado por la mayoría de las agencias de publicidad, merece la pena ilustrar el procedimiento con un ejemplo, para lo cual adaptaremos el ofrecido por Redondo (1995).

Considérese una plan de medios formado por tres inserciones en el soporte S_1 , cuatro inserciones en el soporte S_2 , dos inserciones en el soporte S_3 y tres inserciones en el soporte S_4 . Sean, asimismo, los factores de elevación de cada

individuo y sus probabilidades de exposición a cada soporte, los que se muestran en el cuadro 2.5.

Cuadro 2.5. Factores de elevación y probabilidades individuales de exposición.

Individuo	Factor de elevación	Probabilidades de exposición al soporte i			
		S_1	S_2	S_3	S_4
j	M_j				
1	750	0,25	0,00	0,90	0,00
2	700	0,37	0,64	0,00	0,31
3	825	0,00	0,86	0,54	0,00
4	675	0,30	0,76	0,00	0,00

Fuente: Redondo (1995; pág. 46) y elaboración propia

A continuación, mediante las expresiones [2.3] y [2.4] se calculará para cada individuo los valores de C y G , así como su ratio que se descompondrá en sus partes entera E y decimal D , tal y como muestra el cuadro 2.6. Por ejemplo para el individuo 1:

$$C_1 = 1 - (1-0,25)^3 \cdot (1-0,9)^2 = 0,9958$$

$$G_1 = 3 \cdot 0,25 + 2 \cdot 0,9 = 2,55$$

$$G_1 / C_1 = 2,5608 \rightarrow E_1 = 2; D_1 = 0,5608$$

Cuadro 2.6. Cálculo de D_j y E_j .

Individuo	Variables calculadas				
	C_j	G_j	G_j / C_j	D_j	E_j
1	0,9958	2,55	2,5608	0,5608	2
2	0,9986	4,60	4,6064	0,6064	4
3	0,9999	4,52	4,5205	0,5205	4
4	0,9989	3,94	3,9443	0,9442	3

Fuente: Redondo (1995; pp. 47-48) y elaboración propia

Como se ha indicado en la descripción del método, el paso siguiente consiste en calcular el número de individuos que conforman cada grupo de exposición poblacional y contactos que les corresponde a cada uno de ellos, de acuerdo con las anteriores variables calculadas a nivel muestral. El cuadro 2.7 ofrece los resultados en nuestro ejemplo que detallaremos para el individuo 1:

Grupo 1. Individuos⁵ = $(1 - C_j) \cdot M_j = (1 - 0,9958) \cdot 750 \cong 3$; Contactos = 0;

Grupo 2. Individuos = $(1 - D_j) \cdot C_j \cdot M_j = (1 - 0,5608) \cdot 0,9958 \cdot 750 \cong 328$;
Contactos = $E_j = 2$

Grupo 3. Individuos = $D_j \cdot C_j \cdot M_j = 0,5608 \cdot 0,9958 \cdot 750 \cong 419$; Contactos =
 $E_j + 1 = 3$.

Cuadro 2.7. Individuos contactados y número de contactos.

j	Individuos			Contactos		
	Grupo 1	Grupo 2	Grupo 3	Grupo 1	Grupo 2	Grupo 3
	$(1 - C_j) \cdot M_j$	$(1 - D_j) \cdot C_j \cdot M_j$	$D_j \cdot C_j \cdot M_j$	0	E_j	$E_j + 1$
1	3	328	419	0	2	3
2	1	275	424	0	4	5
3	0	396	429	0	4	5
4	1	37	637	0	3	4

Fuente: Redondo (1995; pp. 47-48) y elaboración propia

La obtención de la distribución de contactos es inmediata a partir del cuadro anterior. Basta sumar todos aquellos individuos que han recibido los mismos contactos que, como se observa en el cuadro 2.8, varían entre ninguno y cinco.

⁵ El resultado se ha redondeado hasta el entero más próximo, dado que carecería de sentido ofrecer decimales cuando se está calculando el número de individuos expuestos a un número determinado de contactos.

Cuadro 2.8. Distribución de contactos.

Contactos	Individuos
0	3 + 1 + 0 + 1 = 5
1	0
2	328
3	419 + 37 = 456
4	275 + 396 + 637 = 1.308
5	424 + 429 = 853
Total	2.950

*Fuente: Redondo (1995; pp. 47-48)
y elaboración propia*

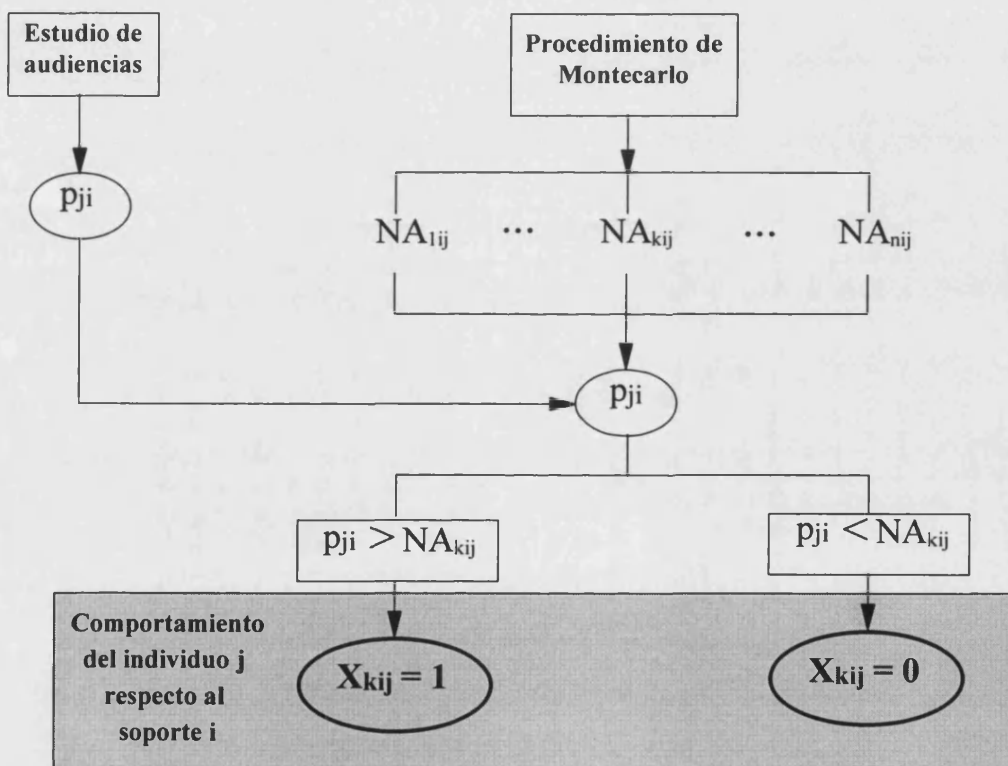
La cobertura del plan de medios puede obtenerse fácilmente a partir del cuadro anterior. En término de individuos será todos aquellos individuos de la población objetivo que han recibido, al menos, una exposición, es decir, todos menos aquellos que no han recibido ninguna: $2.950 - 5 = 2.945$. En términos de probabilidad será el porcentaje que representan esos individuos sobre la población objetivo, es decir, $2.945 / 2.950 = 0,9983$.

2.2.4 Métodos de simulación

La determinación de la cobertura y la distribución de contactos de estos modelos se fundamenta en la obtención de los comportamientos efectivos respecto a los soportes, mediante simulación basada en las probabilidades de exposición a los mismos de cada individuo, información que viene proporcionada por los estudios de audiencias. Aunque existen diferentes procedimientos de simulación, el más utilizado (Martínez Ramos, 1992) para simular el comportamiento de los individuos respecto a los soportes, es el de Montecarlo, y en él centraremos nuestro desarrollo.

Supongamos que tenemos un plan de medios formado por m soportes $S_1 \dots S_m$ en cada uno de los cuales se realizan n inserciones⁶. Del estudio de audiencias que se utilice se dispone, como muestra la figura 2.3, de la probabilidad de exposición p_{ji} de un individuo j al soporte i ($i = 1, 2, \dots, m$).

Figura 2.3: Esquema de funcionamiento de los métodos de simulación



Fuente: Elaboración propia

⁶ A efectos de claridad expositiva supondremos que se realiza el mismo número n de exposiciones en cada soporte aunque, como se desprende claramente del método, el esquema de inserciones puede ser asimétrico.

El procedimiento de simulación debe encargarse de transformar esa probabilidad de exposición a un soporte en un comportamiento, es decir la exposición efectiva o no del individuo j a la inserción k que se realiza en el mencionado soporte i , lo que denominaremos x_{kij} . En otras palabras, debe ofrecer la matriz de exposiciones del individuo j que recoge el cuadro 2.9.

Cuadro 2.9. Matriz de exposiciones del individuo j .

Inserción	Soporte						
	S_1	S_2	...	S_i	...	S_m	
I_1	x_{11j}	x_{12j}	...	x_{1ij}	...	x_{1mj}	x_{1j}
I_2	x_{21j}	x_{22j}	...	x_{2ij}	...	x_{2mj}	x_{2j}
...
I_k	x_{k1j}	x_{k2j}	...	x_{kij}	...	x_{kmj}	x_{kj}
...
I_n	x_{n1j}	x_{n2j}	...	x_{nij}	...	x_{nmj}	x_{nj}
	x_{1j}	x_{2j}	...	x_{ij}	...	x_{mj}	NC_j

Fuente: Elaboración propia

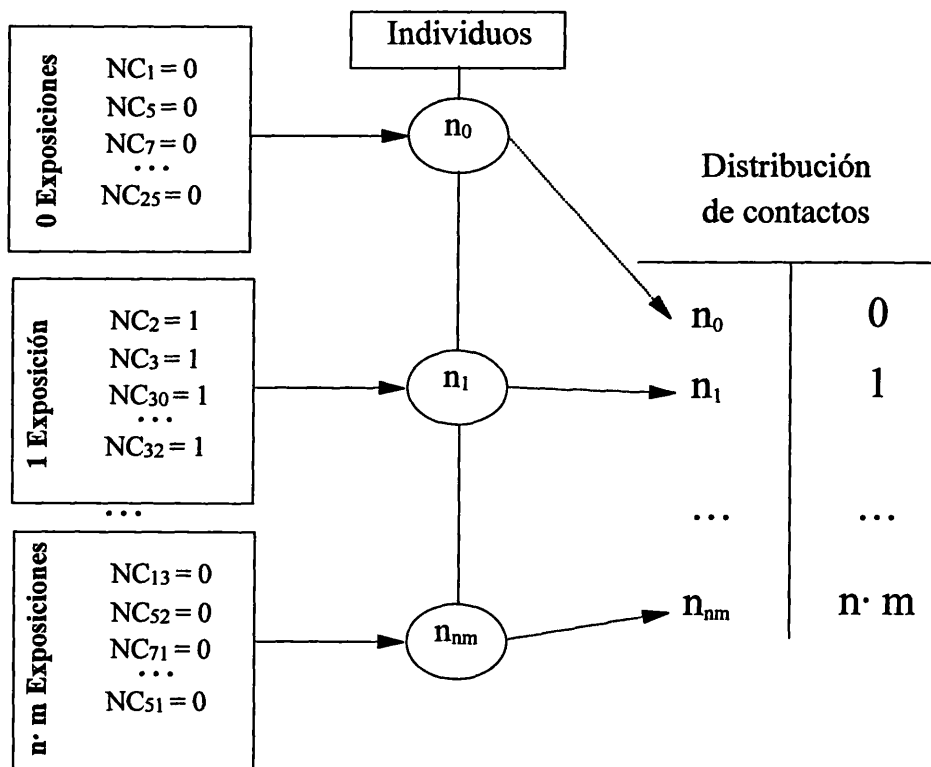
La forma de proceder es la descrita en la figura 2.3. El método de Montecarlo genera un número aleatorio NA_{kij} para la inserción k éxima en el soporte i . Si este número aleatorio, comprendido entre cero y uno, es inferior a la probabilidad de exposición del individuo al soporte p_{ji} , entonces se considera que el individuo ha resultado expuesto a la inserción, recogándose este hecho en la asignación de un 1 en la celda x_{kij} de la anterior matriz. En caso contrario, es decir que el número aleatorio sea superior a la mencionada probabilidad, se considera que el individuo no ha entrado en contacto con la inserción y a la celda se le asigna un cero.

Pasar de las matrices individuales, como la recogida para el sujeto j en el cuadro anterior, a las distribuciones de probabilidad, es inmediato. En cada matriz se suman todas las celdas, con lo que se obtendrá el número de contactos que ha recibido cada individuo NC_j :

$$NC_j = \sum_{i=1}^m x_{i,j} = \sum_{k=1}^n x_{i,j}$$

Para obtener la distribución de contactos, basta sumar el número de individuos que tengan el mismo número de contactos como trata de ilustrar la figura 2.4.

Figura 2.4: Generación de la distribución de contactos en el proceso de simulación



Fuente: Elaboración propia

2.3 MODELOS DE AGREGACIÓN

En el presente epígrafe, se clasificarán los modelos de exposición que van a desarrollarse y evaluarse en capítulos anteriores, atendiendo a dos criterios fundamentales (Chandon, 1985):

- a. En primer lugar, atendiendo a las características del plan de medios al que va a aplicarse el modelo. Si éstos están formados por diversas inserciones en un único soporte, hablaremos de modelos de acumulación de audiencias; si están constituidos por una única inserción en diversos soportes: modelos de duplicación de audiencias; y, finalmente, los planes habituales constituidos por diversas inserciones en distintos soportes: modelos de audiencia neta acumulada.
- b. El segundo criterio hace referencia a cuáles son las hipótesis que subyacen bajo cada uno de los modelos.

El objeto del primer criterio (figura 2.5) es parcelar en tres bloques expositivamente manejables, el gran número de modelos existentes, mientras que el segundo criterio es el tipológicamente relevante puesto que construye la clasificación en base a las características de los modelos que añaden valor a su conocimiento: los supuestos que se han utilizado en su construcción.

La *primera hipótesis* que sirve para clasificar a los modelos pertenecientes a cualquiera de los grupos formados mediante la aplicación del criterio *a*, es el *carácter determinístico o estocástico de la exposición* por parte de un individuo a una inserción cualquiera del plan de medios. En el primer caso, nos encontramos ante los modelos llamados *ad hoc* y, en el segundo, ante los modelos conocidos como estocásticos.

Cuando se representa gráficamente la cobertura de un plan de medios, se observa que ésta crece a una tasa decreciente al aumentar el número de inserciones realizadas. Los modelos *ad hoc*, se limitan a buscar una función matemática que cumpla esta propiedad. Una vez estimados los parámetros de la misma tras ajustarla a los datos existentes mediante la regresión lineal u otro procedimiento de ajuste de datos, se emplea para predecir la cobertura para cualquier nivel de inserciones.

Como indica León (1988), no se realiza ninguna explicación acerca de cuál es la naturaleza de los fenómenos de acumulación o duplicación, con lo que esta es considerada de una manera determinística, limitándose a encontrar la función que mejor reproduzca la evolución de la cobertura al variar el número de inserciones. Dos son, según Martínez Ramos (1992), las principales limitaciones conceptuales⁷ de este enfoque:

- a. Por un lado, no permiten conocer la distribución de contactos: se sabe cuántos individuos contactan, pero no cuántas veces contacta cada uno.
- b. Por otra parte, y derivada de lo anterior, con estos modelos se puede evaluar un conjunto de soportes seleccionados, pero nunca afirmar que una campaña analizada mediante estos modelos sea la mejor posible, dado que la eficacia de una campaña, es un concepto fuertemente ligado, como ya se indicó en el capítulo primero, a la frecuencia de exposición.

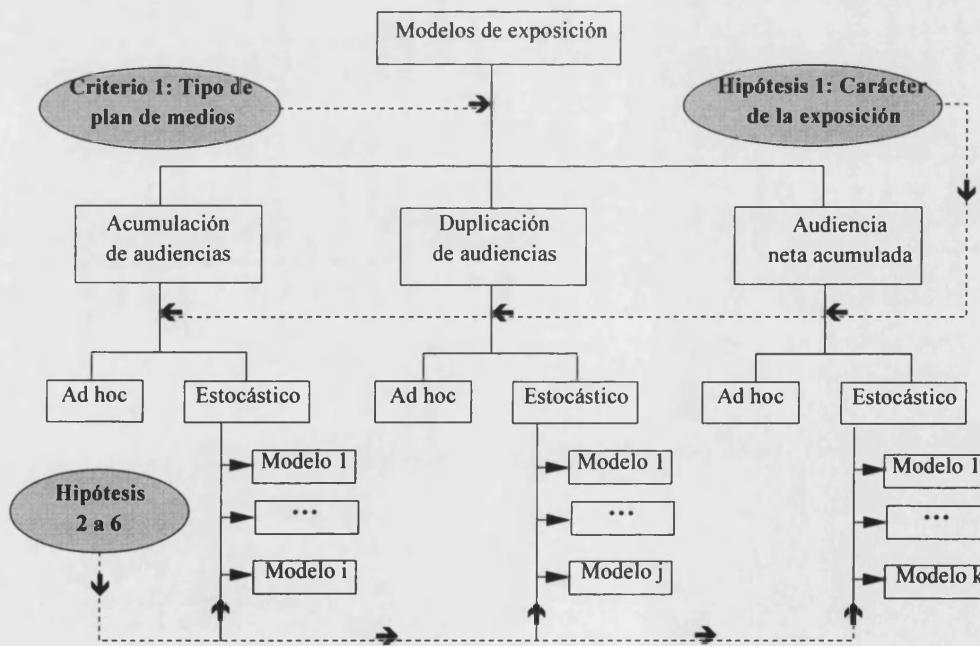
Los modelos estocásticos, por su parte, tratan de recoger los patrones de comportamiento de la audiencia mediante determinadas distribuciones de probabilidad. Los mencionados patrones de comportamiento, son vistos como procesos estocásticos que afectan de un lado a dos entidades: los individuos y las

⁷ Las limitaciones prácticas se derivarán de la mayor o menor precisión en las estimaciones de la cobertura que estos métodos procuren y, por tanto, es objeto de tratamiento en capítulos posteriores.

inserciones y, de otro, a un suceso de carácter dicotómico: la exposición o no al anuncio.

Los modelos estocásticos, se basan en la hipótesis principal de que todos los individuos tienen una cierta probabilidad individual de exposición, que es la que modela sus hábitos de lectura (o de ver televisión u oír radio). Sin embargo, para hacer operativos los modelos, es necesario realizar una serie de hipótesis adicionales acerca de esa probabilidad. Estas hipótesis, que examinaremos a continuación, permiten diferenciar entre unos modelos estocásticos y otros, tal y como se recoge de nuevo en la figura 2.5.

Figura 2.5. Criterios para clasificar los modelos de exposición.



Fuente: elaboración propia

La hipótesis segunda, hace referencia a qué características se atribuyen a las probabilidades de exposición de los individuos. Así, sea p_{ijk} , la probabilidad de

exposición del individuo i a la inserción k -ésima realizada en el soporte j , se dice que en los modelos de exposición, la población es *homogénea* cuando cada uno de los N individuos de la población objetivo, tienen la misma probabilidad de exposición ante un soporte dado. Esto es:

$$p_{1jk} = p_{2jk} = \dots = p_{ijk} = \dots = p_{Njk}$$

Si alguna de estas probabilidades fuera distinta, estaríamos ante modelos donde la población es heterogénea. Asumir esta hipótesis, implicaría aceptar que todos los individuos, independientemente de sus características sociodemográficas, sus temas de interés, etc., tienen la misma probabilidad de exposición ante un soporte. Ello implica asumir que, por ejemplo, que aquellos individuos con interés por el motociclismo o que posean una moto, tienen la misma probabilidad de resultar expuestos a una inserción en una revista especializada en esta materia, que aquellos que carecen de interés por ese tema.

La *tercera hipótesis*, se centra en los soportes. Así, se contraponen aquellos modelos que consideren que todos los soportes tienen la misma probabilidad de ser leídos, vistos u oídos por un individuo determinado (soportes homogéneos), frente a los que asumen que estas probabilidades son distintas en cada uno de los m soportes que conforman el plan de medios. Manteniendo la notación empleada al describir la hipótesis anterior:

$$p_{i1k} = p_{i2k} = \dots = p_{ijk} = \dots = p_{imk}$$

La asunción de esta hipótesis plantea paradojas similares a la expuesta en el caso anterior. Manteniendo el ejemplo, una revista especializada en motociclismo, tendría la misma probabilidad de provocar contactos publicitarios en individuos con interés por el tema que una revista de información económica.

La *cuarta hipótesis*, se basa en el carácter aleatorio o no del fenómeno de la duplicación de las audiencias. Cuando se considera que la probabilidad de que un individuo resulte expuesto a un soporte, no tiene ninguna influencia sobre la probabilidad de que resulte expuesto o no a otro, estamos ante modelos de duplicación aleatoria. En caso contrario la duplicación se considera no aleatoria. Formalizando esta definición con la notación ya empleada, estaremos ante un modelo de duplicación aleatoria, cuando se cumpla que:

$$p_{i1k} \cap p_{i2k} = p_{i1k} \cdot p_{i2k}$$

Intentando ilustrar las consecuencias de esta hipótesis, tal como hemos hecho con las anteriores, asumirla imposibilitaría, por ejemplo, que un individuo que nunca ha tenido interés por los documentales sobre viajes y que, por diversas circunstancias, viese uno que está siendo emitido en La 2, resultara tan entusiasmado por la experiencia que, una vez finalizado, consultase inmediatamente la cartelera y cambiase a Antena 3 donde está programado otro de similares características. La exposición a La 2 ha ejercido una influencia sobre la exposición a Antena 3 que, la hipótesis de aleatoriedad de la duplicación de audiencias no contempla.

La *quinta hipótesis* se formula en términos análogos para el fenómeno de la acumulación de las audiencias. Serán modelos de acumulación aleatoria, aquellos que consideren que la probabilidad de que un individuo resulte expuesto a una inserción en un soporte determinado, no se ve influenciada porque haya o no resultado expuesto a una inserción anterior en ese mismo soporte. Con la notación ya explicitada:

$$p_{ij1} \cap p_{ij2} = p_{ij1} \cdot p_{ij2}$$

Manteniendo el ejemplo que hemos utilizado para ilustrar la hipótesis anterior, el hecho de que ese individuo quedara muy satisfecho con el documental

sobre viajes que vio en La 2, no podría ejercer ninguna influencia sobre la probabilidad de que, a la semana siguiente, encendiese el televisor para, en el mismo programa, ver un nuevo documental. Las suscripciones a revistas o periódicos, tampoco encajan bien con la hipótesis señalada.

Finalmente, la *sexta hipótesis*, introduce la consideración del factor tiempo tanto para los individuos como para los soportes. Estaremos ante modelos estacionarios, cuando las probabilidades de exposición sean estables en el tiempo. Esto implica para los soportes que sus audiencias no varían en el corto plazo de una campaña de publicidad y para los individuos que, en este mismo plazo, no modifican sus hábitos de exposición. Es evidente que esta hipótesis, en el corto plazo, no es demasiado exigente, dado que es constatable una cierta fidelidad de los individuos por los programas de radio que habitualmente escuchan, los periódicos que suelen comprar o determinados programas de televisión. Sin embargo, no es necesario considerar periodos de tiempo demasiado largos para que la hipótesis de estacionariedad choque con el comportamiento real de los individuos. Por ejemplo, la probabilidad de que un sujeto vea el telediario de TVE-1, no puede ser nunca la misma a lo largo del año laboral, que en el mes de agosto mientras pasa sus vacaciones en la playa.

La combinación de estas cinco últimas hipótesis, permite distinguir entre los diferentes modelos estocásticos de exposición. Los cuadros 2.10, 2.11 y 2.12 sistematizan las hipótesis que subyacen bajo los modelos que se desarrollarán en el capítulo tercero, para cada uno de los grupos establecidos por el primer criterio: modelos de acumulación de audiencias, duplicación de audiencias y audiencia neta acumulada. Referimos a ese capítulo para derivar las implicaciones que, en el desarrollo de cada uno de los modelos, tienen las distintas simplificaciones que cada uno de ellos asume.

Cuadro 2.10. Características de los modelos estocásticos de acumulación de audiencias.

Modelo	Población	Soportes	Duplicación	Acumulación	Estacionariedad		Autor
					Individuos	Soportes	
Acumulación Binomial	Homogénea	N. P.	N. P.	Aleatoria	Estacionarios	Estacionarios	Chandon (1985)
Dos segmentos	Heterogénea	N. P.	N. P.	No Aleatoria	Estacionarios	Estacionarios	Chandon (1985)
Tres segmentos	Heterogénea	N. P.	N. P.	No Aleatoria	Estacionarios	Estacionarios	Chandon (1985)
Poltiz	Heterogénea	N. P.	N. P.	No Aleatoria	Estacionarios	Estacionarios	Poltiz (1958)
Beta Binomial	Heterogénea	N. P.	N. P.	No Aleatoria	Estacionarios	Estacionarios	Hyett (1958)
Binomial Negativa	Heterogénea	N. P.	N. P.	No Aleatoria	Estacionarios	Estacionarios	Green y Stock (1971)
Broadbent	Heterogénea	N. P.	N. P.	No Aleatoria	Estacionarios	Estacionarios	Broadbent (1964)
Morgensztern	Heterogénea	N. P.	N. P.	No Aleatoria	Estacionarios	Estacionarios	Morgensztern (1970)
Beta Binomial Modificado	Heterogénea	N. P.	N. P.	No Aleatoria	Estacionarios	Estacionarios	Chandon (1985)
Markov	Homogénea	N. P.	N. P.	No Aleatoria	Estacionarios	No estacionarios	Chandon (1985)

N.P.= Debido a las características de estos modelos la hipótesis no es procedente

Fuente: Elaboración propia

Cuadro 2.11. Características de los modelos estocásticos de duplicación de audiencias.

Modelo	Población	Soportes	Duplicación	Acumulación	Estacionariedad		Autor
					Individuos	Soportes	
Binomial	Homogénea	Homogéneos	Aleatoria	N.P.	N.P.	N.P.	Chandon (1985)
Sainsbury	Homogénea	Heterogéneos	Aleatoria	N.P.	N.P.	N.P.	Caffyn y Sagovsky (1963)
Kwerel	Homogénea	Homogéneos	No Aleatoria	N.P.	N.P.	N.P.	Kwerel (1964)
Sainsbury modificado	Homogénea	Heterogéneos	No Aleatoria	N.P.	N.P.	N.P.	Caffyn y Sagovsky (1963)
Agregación secuencial duplicación cte.	Homogénea	Heterogéneos	Constante	N.P.	N.P.	N.P.	Goodhart y Erhenberg (1969)
Agreg. secuencial duplicación variable	Homogénea	Heterogéneos	No Aleatoria	N.P.	N.P.	N.P.	Chandon (1985)
Independencia condicional	Homogénea	Heterogéneos	No Aleatoria	N.P.	N.P.	N.P.	Chandon (1985)
Hofmans ampliada	Homogénea	Heterogéneos	No Aleatoria	N.P.	N.P.	N.P.	Chandon (1985)
Duplicación de Metheringham	Heterogénea	Homogéneos	No Aleatoria	N.P.	N.P.	N.P.	Metheringham (1964)

N.P.= Debido a las características de estos modelos la hipótesis no es procedente

Fuente: Elaboración propia

Cuadro 2.12: Características de los modelos estocásticos de audiencia neta acumulada

Modelo	Población	Soportes	Duplicación	Acumulación	Estacionariedad		Autor
					Individuos	Soportes	
Binomial	Homogénea	Homogéneos	Aleatoria	Aleatoria	Estacionarios	Estacionarios	Lee y Burkart (1960)
Binomial múltiple	Homogénea	Heterogéneos	Aleatoria	Aleatoria	Estacionarios	Estacionarios	Chandon (1985)
Multinomial	Homogénea	Heterogéneos	No Aleatoria	Aleatoria	Estacionarios	Estacionarios	Chandon (1985)
Markov Bernouilli	Homogénea	Homogéneos	No Aleatoria	No Aleatoria	Estacionarios	No Estacionarios	Chandon (1985)
Metheringham	Heterogénea	Homogéneos	No Aleatoria	No Aleatoria	Estacionarios	Estacionarios	Metheringham (1964)
Dirichlet Multinomial	Heterogénea	Heterogéneos	No Aleatoria	No Aleatoria	Estacionarios	Estacionarios	Chandon (1985)
Poisson Binomial	Homogénea	Homogéneos	Aleatoria	Aleatoria	Estacionarios	Estacionarios	Leckenby y Ju (1989)
Kwerel Geométrico	Homogénea	Homogéneos	No Aleatoria	No Aleatoria	Estacionarios	Estacionarios	Leckenby y Kishi (1981)
Hofmans Geométrico	Homogénea	Heterogéneos	No Aleatoria	No Aleatoria	Estacionarios	Estacionarios	Leckenby y Kishi (1981)
Hofmans Betabinomial	Homogénea	Heterogéneos	No Aleatoria	No Aleatoria	Estacionarios	Estacionarios	Leckenby y Boyd (1984a)
Hofmans Poisson	Homogénea	Heterogéneos	Aleatoria	Aleatoria	Estacionarios	Estacionarios	Leckenby y Ju (1989)

N.P.= Debido a las características de estos modelos la hipótesis no es procedente

Fuente: Elaboración propia

3

Revisión de los modelos de exposición

3.1 MODELOS DE ACUMULACIÓN DE AUDIENCIAS

3.1.1 Modelos ad hoc

Como ya se señaló en el capítulo anterior, si se representa gráficamente la audiencia de un soporte frente al número de inserciones, se observará que la audiencia crece a una tasa decreciente al aumentar el número de las mismas. Los modelos *ad hoc*, se limitan a buscar una forma funcional que cumpla esta propiedad, a estimar los parámetros que la especifican con los datos disponibles, y a utilizarla

para predecir la cobertura, de otros planes de medios. Estos modelos, como ya se ha señalado, no están diseñados para predecir la distribución de contactos. Dependiendo de cuál sea la forma funcional que se ajuste, nos encontraremos ante los modelos de acumulación logarítmica (Chandon, 1985), acumulación exponencial (Chandon, 1985), modelo de acumulación de Agostini (1962) y modelo de acumulación de Hofmans (1966).

3.1.1.1 Modelo de acumulación logarítmica

En este modelo, propuesto por Chandon (1985), la forma funcional que se ajusta a los datos, viene dada por:

$$R_N = a + b \log N \quad [3.1]$$

donde R_N es el alcance acumulado logrado al realizar N inserciones en un soporte dado, y a y b son dos constantes positivas que es necesario estimar para especificar el modelo. Para esta estimación son necesarios, por tanto, dos datos: el alcance acumulado cuando se realizan una (R_1) o dos (R_2) inserciones en el soporte que constituye el plan de medios. Sustituyendo sucesivamente estos dos datos en [3.1] y resolviendo el sistema para a y b , la estimación de los parámetros es:

$$\hat{a} = R_1$$

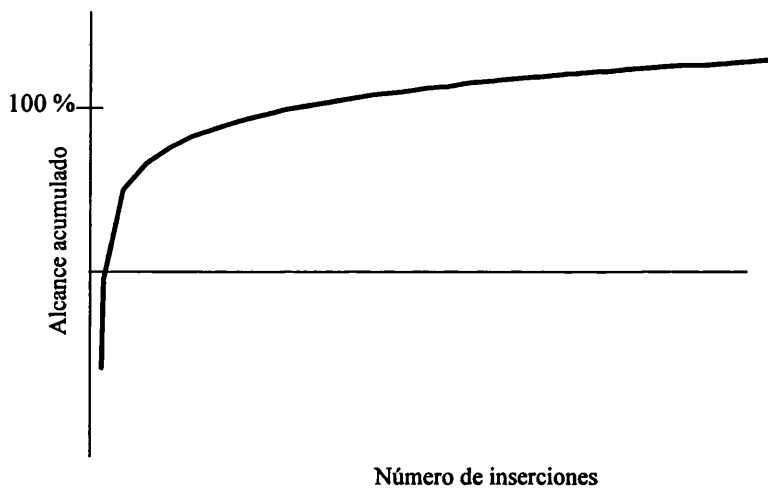
$$\hat{b} = \frac{R_2 - R_1}{\log 2}$$

quedando la expresión que nos permitirá calcular la cobertura para N inserciones en el soporte considerado como:

$$R_N = R_1 + \frac{R_2 - R_1}{\log 2} \log N$$

Independientemente de la mejor o peor capacidad predictiva que demuestre este modelo, lo que será analizado en el capítulo siguiente, la acumulación logarítmica presenta dos importantes limitaciones teóricas que pueden apreciarse gráficamente en la figura 3.1. La función logarítmica no presenta asíntota horizontal que la acote superiormente. Esto implica que, cuando crezca el número de inserciones, el alcance acumulado predicho podrá superar la población total, lo que evidentemente carece de sentido. En segundo lugar, cuando no se realiza ninguna inserción, el alcance acumulado debe ser cero. Pues bien, la función logarítmica tiende a $-\infty$ cuando el número de inserciones tiende a cero.

Figura 3.1. Función logarítmica



3.1.1.2 Modelo de acumulación exponencial

La función exponencial, al igual que la logarítmica, tiene la propiedad de crecer a tasas decrecientes (concavidad), tal y como lo hace el alcance acumulado, por lo que es propuesta por Chandon (1985) para servir de base a un segundo modelo *ad hoc* de acumulación. Su expresión general es la siguiente:

$$R_N = RMAX - ab^N \quad [3.2]$$

donde $a \geq 0$ y $b \leq 1$ son los dos parámetros a estimar que, junto con el tercer parámetro $RMAX$, que representa la cobertura máxima que puede lograr el plan de medios, especifican el modelo. Dado que es necesario estimar tres parámetros, se necesitarán tres datos: el alcance acumulado de una (R_1), dos (R_2) y tres (R_3) inserciones. Sustituyendo sucesivamente estos tres datos en [3.2] y resolviendo el sistema de ecuaciones para a , b y $RMAX$, los parámetros pueden calcularse mediante:

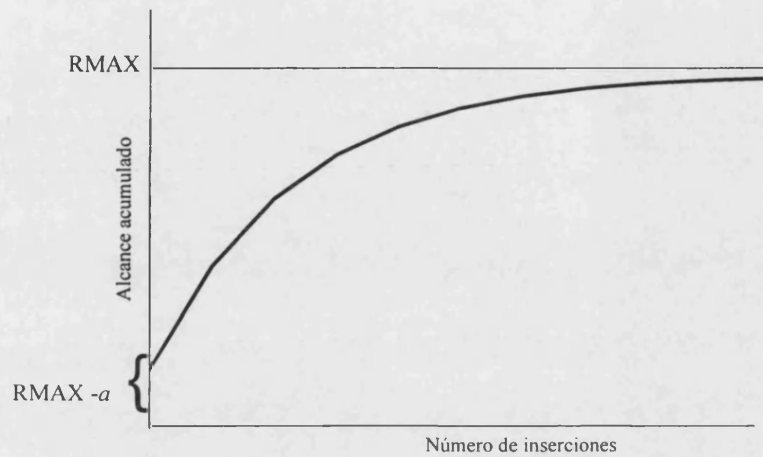
$$\hat{a} = \frac{(R\hat{M}AX - R_1)^2}{R\hat{M}AX - R_2}$$

$$\hat{b} = \frac{R\hat{M}AX - R_2}{R\hat{M}AX - R_1}$$

$$R\hat{M}AX = \frac{R_2^2 - R_1R_3}{2R_2 - R_1 - R_3}$$

Observando la figura 3.2 y analizando la expresión [3.2], se desprende que la función exponencial tiene una propiedad muy deseable: posee una asíntota horizontal que actúa como cota superior, de tal forma que cuando el número de inserciones N tiende a ∞ , la audiencia acumulada tiende a esa cota, $RMAX$, lo que evita que el modelo llegue a proporcionar coberturas superiores a la población objetivo total. Por contra, cuando no se realizan inserciones, el alcance predicho por [3.2] es $RMAX - a$, que no es, como dictaría la lógica, cero.

Figura 3.2. Función exponencial



3.1.1.3 Modelo de acumulación de Agostini

Agostini (1962) propone este modelo para calcular el alcance acumulado, bajo el supuesto de que la acumulación es un proceso aleatorio, es decir, que no hay relación entre resultar expuesto a una inserción en el soporte considerado y resultarlo en inserciones sucesivas. Sea R_N el alcance acumulado tras N inserciones y llamemos $\Delta_N = R_N - R_{N-1}$ al incremento del alcance acumulado logrado al introducir la inserción n -ésima. Llamemos, asimismo,

$$\overline{R_{N-1}} = 1 - R_{N-1}$$

al porcentaje de la población no alcanzada tras la inserción $N-1$ éxima. Pues bien, Agostini (1962), demuestra que la *ratio* y :

$$y = \frac{\Delta_N}{\overline{R_{N-1}}} \quad [3.3]$$

decrecía al aumentar N , lo que convierte a esa *ratio* en una posible aproximación funcional de los alcances acumulados, dado que el cumplimiento de esa propiedad es la que ha llevado, en los dos modelos anteriores, a seleccionar las función logarítmica y exponencial como elementos centrales de los modelos *ad hoc* expuestos. Representando el logaritmo de la *ratio* dada en [3.3] frente al logaritmo del número de inserciones con los datos ofrecidos por Politz (1953; 1958), llega a la conclusión de que parecen seguir una línea recta con pendiente negativa, que presume es:

$$\log y = b - a \log N \quad [3.4]$$

Igualando [3.3] y [3.4] tras despejar y en esta segunda expresión, se llega a:

$$\frac{R_N - R_{N-1}}{1 - R_{N-1}} = \frac{a}{N^b} \quad [3.5]$$

donde toda la notación es conocida salvo a y b que son los dos parámetros que es necesario estimar para especificar el modelo. De nuevo son necesarios para este menester los datos de audiencia acumulada tras las dos primeras inserciones, esto es, R_1 y R_2 . Sustituyendo R_1 para $N = 1$ en [3.5] y teniendo en cuenta que $R_0 = 0$, se llega a que:

$$\hat{a} = R_1$$

y volviendo a sustituir este resultado en [3.5] para $N = 2$, el parámetro b se estima como:

$$\hat{b} = \frac{1}{\log 2} \log \frac{R_1(1-R_1)}{R_2-R_1}$$

con los parámetros a y b así estimados, basta ir sustituyendo sucesivamente $N = 3$, $N = 4$, etc. para obtener el alcance acumulado para el número de inserciones que se desee. En general, se llega a que:

$$1 - R_N = (1 - R_1) \left(1 - \frac{R_1}{2^b}\right) \cdots \left(1 - \frac{R_1}{N^b}\right) \quad [3.6]$$

Analizando la expresión [3.6], vemos que, al igual que el modelo exponencial, tiene una ventaja teórica sobre el logarítmico: el alcance acumulado nunca puede ser superior a la población total, dado que el lado derecho de la expresión [3.6] es un producto de términos menores que la unidad que tiende, por este motivo a cero. Por lo tanto, si $1 - R_N \rightarrow 0$, entonces $R_N \rightarrow 1$, no superando nunca este valor.

3.1.1.4 Modelo de acumulación de Hofmans

Apoyándose en la similitud existente entre considerar N inserciones en un soporte y una inserción en N soportes distintos, Hofmans (1966), tomando como base la fórmula de Agostini (1961)¹ desarrolla dos modelos, uno de acumulación de audiencias, que analizaremos a continuación y un segundo de duplicación de audiencias². Partiendo, como se ha indicado de la expresión de Agostini (1961) para el cálculo de la cobertura:

$$R_m = \frac{A^2}{A+kD} \quad [3.7]$$

siendo:

A_i la audiencia del soporte i

¹ Este modelo de duplicación de audiencias se analiza en el epígrafe 3.2.1.1 de este mismo capítulo, y al mismo debe referirse el lector para la deducción de las expresiones que de él se utilizan.

² Véase epígrafe 3.2.1.2 en este mismo capítulo.

$A = \sum_{i=1}^m A_i$ es la audiencia bruta de los m soportes

$A_{ij} = A_i \cap A_j$ es la audiencia duplicada entre los soportes i y j

$D = \sum_{i=1}^{m-1} \sum_{j=i+1}^m A_{ij}$ es la suma de las audiencias duplicadas

k es una constante que, en ese estudio, toma el valor de 1,125

R_m es la cobertura del plan de medios formado por una inserción en m soportes.

Para adaptar la expresión [3.7] al problema de la acumulación de audiencias, Hofmans (1966) sustituye m , el número de soportes, por N , el número de inserciones, realizando una serie de supuestos adicionales:

- a. Se acepta la simplificación de que la audiencia de un soporte es constante para todos sus números, con lo que la audiencia bruta de Agostini (1961), es simplemente N veces la audiencia de un número: $A = \sum_{i=1}^N R_i = NR_1$.

[3.8]

- b. Se admite también, que la duplicación entre dos inserciones A_{ij} es constante e igual a d , es decir, no depende de qué dos soportes estemos considerando:

$$A_{ij} = A_{i'j'} = d \quad \forall i, i' \in N; \quad \forall j, j' \in N; \quad i \neq i'; \quad j \neq j'$$

Entonces, por paralelismo con la formulación de Agostini se cumplirá que:

$$D = \sum_{i=1}^{N-1} \sum_{j=i+1}^N A_{ij} = d[(N-1) + (N-2) + \dots + 2 + 1] \quad [3.9]$$

dado que la parte entre corchetes de la expresión [3.9], es la suma de los términos de una progresión aritmética, cuyo primer término es 1, el último $N-1$ y tiene, por tanto, $N-1$ términos, podemos escribirla:

$$D = \frac{N(N-1)}{2} \cdot d \quad [3.10]$$

y dado que en las simplificaciones señaladas d había sido considerada constante, el alcance acumulado de dos inserciones, sería igual al doble del de una inserción menos la parte duplicada entre ambas, o sea: $R_2 = 2R_1 - d$. Despejando d , sustituyendo en [3.10] y esta, junto con [3.8] en [3.7], nos permite obtener el alcance acumulado sin más que conocer esta dato para las dos primeras inserciones mediante:

$$R_N = \frac{(NR_1)^2}{NR_1 + k(2R_1 - R_2) \binom{N}{2}} \quad [3.11]$$

El propio Hofmans (1966), demuestra que la principal limitación teórica de la expresión [3.11], es que la supuesta constante k no es tal constante. Hofmans trata de encontrar una función cuya variable independiente sea el número de inserciones, que replique a k y soslayar, así, esta limitación. Encuentra que la forma funcional que mejor ajusta a sus datos³ es:

$$k_N = k(N-1)^{a-1} \quad [3.12]$$

donde la introducción de un parámetro adicional a obliga a disponer de un dato adicional de audiencia acumulada R_3 , estimándose el parámetro mediante la expresión:

$$\hat{a} = \frac{\log \left[\frac{(3R_1 - R_3)R_2}{(2R_1 - R_2)R_3} \right]}{\log 2}$$

De tal forma que el modelo de Hofmans (1966) quedará completamente especificado sin más que sustituir [3.12] en [3.11]:

³ Concretamente, los procedentes del estudio de Politz de 1964 (Politz, 1965).

$$R_N = \frac{(NR_1)^2}{NR_1 + k(N-1)^{\hat{a}} \cdot \frac{N}{2} \cdot d}$$

siendo k , no el valor obtenido por Agostini (1961), sino:

$$k = \frac{2R_1}{R_2}$$

3.1.2 Modelos estocásticos de acumulación de audiencias

Los modelos estocásticos, como se indicó en el capítulo segundo, recogen los patrones de comportamiento de la audiencia mediante determinadas distribuciones de probabilidad. Estos patrones son vistos como procesos estocásticos que afectan de un lado a las inserciones y a los individuos y, de otro, a un suceso de carácter dicotómico: la exposición o no al anuncio. Las hipótesis que se realizan sobre las inserciones y los individuos, determinan el desarrollo teórico de los modelos que expondremos a continuación, refiriéndose fundamentalmente a la:

- a. *Estacionariedad frente a las inserciones.* La probabilidad individual de exposición a determinada inserción, es independiente de las exposiciones pasadas.
- b. *Estacionariedad respecto a los individuos.* La probabilidad de exposición del individuo i a la inserción N , es independiente de la probabilidad de exposición del individuo j a esa misma inserción.
- c. *Homogeneidad de los individuos.* Las probabilidades de exposición son idénticas para todos los individuos.

3.1.2.1 Modelo binomial de acumulación

Asume las hipótesis de estacionariedad de inserciones e individuos y la de homogeneidad de estos últimos. Estas hipótesis determinan que la probabilidad p de que el individuo i -ésimo resulte expuesto a la n -ésima inserción realizada en el soporte analizado, sea constante para todos los individuos e inserciones, pudiendo aproximarse como la relación entre la audiencia promedio del soporte en las N inserciones y el total de la población objetivo POB :

$$p = \frac{\sum_{i=1}^N A_i}{POB}$$

Las hipótesis también determinan que la variable aleatoria “exposición a una inserción”, pueda considerarse que sigue una distribución Bernouilli, dado que se trata de una variable dicotómica que se concreta en dos sucesos mutuamente excluyentes entre sí: resultar o no expuesto. Asimismo, dado que estamos realizando n pruebas o ensayos de Bernouilli independientes en las mismas condiciones, es decir, la probabilidad de éxito p , es la misma en cada prueba, podemos afirmar que la variable aleatoria “número de veces que se presenta el suceso tener éxito (resultar expuesto)”, sigue el modelo binomial⁴. Esto nos permite utilizar las definiciones de la función binomial:

Su *función de cuantía*, en nuestro caso la probabilidad de que un sujeto resulte expuesto a K inserciones de un total de N (veces que realizamos el experimento aleatorio), vendrá dada por la expresión:

⁴ Una explicación mucho más detallada de cómo la distribución de Bernouilli sirve de base a la distribución binomial, puede encontrarse en Cuadras, Echeverría, Mateo y Sánchez (1991), pp. 232-238. En ocasiones no se introduce el concepto de variable Bernouilli, considerándola como un caso particular de la distribución binomial para $n = 1$. Un ejemplo de este planteamiento teórico aparece en López Cachero (1990), pp. 306-309.

$$P(K/N, p) = \binom{N}{K} p^K (1-p)^{N-K} \quad 0 \leq K \leq N \quad [3.13]$$

de modo que, si esa es la probabilidad que afecta a un sujeto, para obtener el número de ellos de una población con POB miembros que resultarán expuestos a K inserciones de las N realizadas (distribución de contactos E_K^N), bastará con multiplicar ambos términos:

$$E_K^N = POB \times P(K/N, p)$$

Como la probabilidad de resultar expuesto a al menos una inserción, es la unidad menos la probabilidad de no resultar expuesto a ninguna, podemos escribir, sustituyendo a la vez [3.13] para $K = 0$:

$$P(K \geq 1/N, p) = 1 - P(K = 0/N, p) = 1 - (1-p)^N \quad [3.14]$$

Dado que hemos asumido la hipótesis de homogeneidad de los individuos, al ser todos idénticos podemos poner que el alcance acumulado de N inserciones medido en individuos, será el resultado de multiplicar la población total POB por la probabilidad individual de resultar alcanzado al menos una vez de las N posibilidades [3.14], es decir:

$$R_N = POB \times [1 - (1-p)^N]$$

3.1.2.2 Modelo de acumulación de dos segmentos

Este modelo, junto con el que le sigue de tres segmentos, es una síntesis que realiza Chandon (1985) de los trabajos de Lee y Burkhart (1960), Lee (1962), Ellis (1966), y Frankel (1962). Estos modelos relajan la hipótesis de homogeneidad de los individuos, es decir, el supuesto de que las probabilidades de exposición son las

mismas para todos ellos. Este modelo, recoge las diferencias entre los individuos considerando que existen dos tipos (segmentos):

- a. Los *lectores*, que representan un porcentaje a de la población y que tienen una probabilidad p de resultar expuestos.
- b. Los *no lectores*, que son el $1 - a$ por ciento de la población y cuya probabilidad de exposición es cero.

A partir de aquí, el desarrollo teórico es paralelo al ya expuesto para el modelo binomial. La probabilidad de que un individuo resulte expuesto a K de las N inserciones, viene dada por [3.13] con la salvedad de que sólo un a por cien de la población es candidata a resultar expuesta:

$$E_K^N = a \cdot POB \cdot \binom{N}{K} p^K (1-p)^{N-K} \quad 0 \leq K \leq N \quad [3.15]$$

expresión que adopta el siguiente enunciado particular cuando $K = 0$:

$$E_K^N = POB \cdot 1 - a + a \cdot (1-p)^N$$

y análogamente la cobertura:

$$R_N = a \cdot POB \cdot [1 - (1-p)^N] \quad [3.16]$$

Como puede comprobarse, es necesario estimar los parámetros para que el modelo quede completamente especificado: a y p , por lo que serán necesarios dos datos de audiencia acumulada del soporte tras una (R_1) y dos (R_2) inserciones. sustituyendo en [3.16] y resolviendo el sistema para los parámetros, sus estimadores quedan:

$$\hat{p} = \frac{2R_1 - R_2}{R_1}$$

$$\hat{a} = \frac{R_1^2}{(2R_1 - R_2) \cdot POB}$$



3.1.2.3 Modelo de acumulación de tres segmentos

La introducción de más segmentos debe, lógicamente, captar mejor la heterogeneidad de la población, redundando en modelos con una capacidad de predecir la cobertura y la distribución de contactos mejorada. El coste de esta mejora radica, sin embargo, en que será necesario conocer el alcance acumulado de más inserciones en el soporte para estimar los parámetros y especificar el modelo. El siguiente modelo, presentado como hemos indicado por Chandon (1985), trata de llegar a un compromiso entre ambos factores introduciendo tres segmentos:

- a. Los *lectores habituales*, que son una proporción a de la población y que tienen una probabilidad p_1 de resultar expuestos.
- b. Los *lectores esporádicos*, que representan un porcentaje b de la población, con una probabilidad p_2 de exposición.
- c. Los *no lectores* que, lógicamente, son un $1 - a - b$ por cien de la población y cuya probabilidad de exposición p_3 es nula.

La probabilidad de que un individuo perteneciente a cada grupo sea alcanzado al menos una vez, sigue viniendo dada por la expresión [3.16], sólo que ahora el término p es distinto en función del segmento al que se pertenezca (p_i), con lo que la cobertura (en términos de probabilidad) quedará:

$$R_N = a [1 - (1 - p_1)^N] + b [1 - (1 - p_2)^N] \quad [3.17]$$

y análogamente, la distribución de contactos se expresará de forma similar a la ecuación [3.15]:

$$E_K^N = \binom{N}{K} [a p_1^K (1 - p_1)^{N-K} + b p_2^K (1 - p_2)^{N-K}]$$

salvo para $K = 0$, que se calculará como:

$$E_0^N = 1 - a - b + a(1 - p_1)^N + b(1 - p_2)^N$$

Dado que para especificar el modelo hemos de conocer los alcances acumulados de las cuatro primeras inserciones ($R_1 \dots R_4$), sustituyéndolos en [3.17], obtendríamos este sistema de cuatro ecuaciones:

$$\begin{aligned} R_1 &= ap_1 + bp_2 \\ R_2 &= a[1 - (1 - p_1)^2] + b[1 - (1 - p_2)^2] \\ R_3 &= a[1 - (1 - p_1)^3] + b[1 - (1 - p_2)^3] \\ R_4 &= a[1 - (1 - p_1)^4] + b[1 - (1 - p_2)^4] \end{aligned}$$

que, además de no ser lineales, no tiene garantizada la existencia de solución. Para intentar resolver este problema, Chandon (1985) apunta el siguiente procedimiento de cálculo. Si resolvemos las dos primeras ecuaciones para a y b , tenemos:

$$\begin{aligned} \hat{a} &= \frac{R_1(2 - p_2) - R_2}{p_1(p_1 - p_2)} \\ \hat{b} &= \frac{R_2 - R_1(2 - p_1)}{p_2(p_1 - p_2)} \end{aligned}$$

donde si fuésemos capaces de calcular p_1 y p_2 por otra vía, el modelo quedaría completamente especificado. Dado que a y b han de ser positivos (son proporciones poblacionales), deberá cumplirse que:

$$\hat{a} = \frac{R_1(2-p_2)-R_2}{p_1(p_1-p_2)} > 0$$

donde el denominador ha de ser necesariamente positivo, ya que son probabilidades y, obligatoriamente, la probabilidad de exposición de los lectores habituales ha de ser mayor que la de los esporádicos, de modo que para que se cumpla la desigualdad:

$$(2-p_2)R_1 > R_2$$

y del mismo modo para el parámetro b se obtiene:

$$R_2 > (2-p_1)R_1$$

uniendo ambas condiciones, llegamos a que debe observarse lo siguiente:

$$(2-p_2)R_1 > R_2 > (2-p_1)R_1$$

Sin embargo, son muchos los valores de p_1 y p_2 que satisfacen la anterior condición. Chandon (1985) propone una regla sencilla de selección: elegirlos de modo que el intervalo $[p_1, p_2]$ esté centrado en $2 - \frac{R_2}{R_1}$ y que p_1 y p_2 estén lo suficientemente distantes para que no puedan confundirse los lectores habituales con los esporádicos.

3.1.2.4 Modelo de acumulación de Politz

Este modelo mantiene las hipótesis de estacionariedad y la de heterogeneidad de la población. Sin embargo, a diferencia de los dos últimos métodos analizados, esta heterogeneidad no se recoge mediante la creación de segmentos, sino asignando a cada individuo su propia probabilidad de exposición, siguiendo el conjunto de las mismas una determinada distribución de probabilidad para la población. Los distintos modelos que expondremos, varían en la forma funcional que consideran más adecuada para esa distribución de probabilidad poblacional. Veremos, por tanto, de forma general las expresiones que permiten obtener la cobertura y la distribución de contactos, particularizándolas posteriormente para cada uno de los modelos desarrollados.

Sea p_i la probabilidad individual de exposición del sujeto i a una inserción en el soporte considerado, y sea $f(p)$ el número de individuos asociados a un valor p_i de la probabilidad individual. En general, el alcance acumulado de N inserciones, puede escribirse como el total de individuos de la población menos aquellos que no han resultado expuestos:

$$R_N = POB - E_0^N = \sum_{K=1}^N E_K^N$$

Por otra parte, la población total puede escribirse como la suma de todos los individuos asociada con todos los posibles valores de p y con todas las posibles frecuencias de exposición, es decir:

$$POB = \int_0^1 f(p) dp = \sum_{K=0}^N E_K^N$$

Sea, asimismo, un individuo i , cuya probabilidad de exposición a una inserción sea p_i , como la exposición-no exposición es una variable Bernouilli, la probabilidad vendrá dada por [3.13] para cada individuo i :

$$P(K/N, p_i) = \binom{N}{K} p_i^K (1 - p_i)^{N-K} \quad 0 \leq K \leq N$$

El número de individuos que tienen esa probabilidad personal es $f(p_i)$, de modo que la cantidad de sujetos expuestos a K inserciones de las N realizadas, de entre aquellos que tienen la probabilidad personal p_i (lo cual denotaremos como iE_K^N), serán el producto de la probabilidad por los individuos con esa probabilidad, es decir:

$$iE_K^N = \binom{N}{K} p_i^K (1 - p_i)^{N-K} f(p_i)$$

Agregando (integrando por ser variable continua) para todas las probabilidades, tendremos la distribución de contactos:

$$E_K^N = \binom{N}{K} p^K (1 - p)^{N-K} f(p) dp \quad [3.18]$$

Entrando ya en el desarrollo propuesto por Politz (1958), éste considera que la función $f(p)$, puede aproximarse satisfactoriamente mediante el polinomio:

$$f(p) = a + bp + cp^2 + dp^3$$

expresión que tiene muchísimas limitaciones. En primer lugar no es una verdadera distribución de probabilidad⁵ y, además, al tener que estimar cuatro parámetros, son

⁵ Entendemos por una función de distribución *verdadera*, aquella que cumple las siguientes propiedades:

Continúa en la página siguiente...

necesarios cuatro datos de alcance. El primer problema será resuelto en modelos sucesivos introduciendo distribuciones reales como la beta binomial o la binomial negativa. Para hacer frente al segundo, Chandon (1985) recomienda aproximar una recta:

$$f(p) = a + bp$$

que sustituida en [3.18] conduce a:

$$E_K^N = \binom{N}{K} \left[a \int_0^1 p^K (1-p)^{N-K} dp + b \int_0^1 p^{K+1} (1-p)^{N-K} dp \right] \quad [3.19]$$

dado que la función beta se define⁶ del siguiente modo:

$$\beta(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx \quad p > 0, q > 0 \quad [3.20]$$

adaptando la notación a la empleada por nosotros, podemos escribir [3.19] como:

$$E_K^N = \binom{N}{K} [a\beta(K+1, N-K+1) + b\beta(K+2, N-K+1)] \quad [3.21]$$

Dado que para poder especificar el modelo es necesario conocer los datos correspondientes a las dos primeras inserciones E_0^2 y E_1^2 , sustituyendo en [3.21] y teniendo en cuenta que al ser K y N enteros, la función beta puede escribirse como:

$$\beta(K+1, N-K+1) = \frac{K!(N-K)!}{(N+1)!}$$

llegamos al sistema:

$$1. f(x) \geq 0$$

$$2. \int_{-\infty}^{+\infty} f(x) dx = 1$$

⁶ Véase al respecto López Cachero (1990), p. 131

$$6E_1^2 = 2a + b$$

$$12E_0^2 = 4a + b$$

de donde pueden estimarse los parámetros del modelo:

$$\hat{a} = 6E_0^2 - 3E_1^2 \quad [3.22]$$

$$\hat{b} = 12(E_1^2 - E_0^2)$$

Para expresar las estimaciones en términos de cobertura, en lugar de distribución de contactos, basta tener en cuenta que las personas no expuestas a ninguna de las dos inserciones realizadas es igual a toda la población menos las que han resultado expuestas a al menos una de esas dos exposiciones, es decir:

$$E_0^2 = 1 - R_1 \quad [3.23]$$

y que, por otra parte, dado que Chandon (1985) demuestra que la relación que existe entre la cobertura obtenida tras realizar N inserciones, y la que se obtuvo en las $N - 1$ anteriores es:

$$R_{N-1} = R_N - \frac{E_1^N}{N}$$

se sigue que:

$$R_1 = R_2 - \frac{E_1^2}{2} \quad [3.24]$$

Sustituyendo [3.24] y [3.23] en [3.22], los parámetros pueden también calcularse en términos de cobertura como:

$$\hat{a} = 6(1 - 2R_2 + R_1)$$

$$\hat{b} = 12(3R_2 - 2R_1 - 1)$$

Con estas estimaciones, teniendo en cuenta que la población alcanzada al menos una vez tras las N inserciones realizadas (R_N) es toda la población menos los que no han sido alcanzados en ninguna ocasión, es decir E_0^N , el alcance acumulado total quedará teniendo en cuenta [3.21] como:

$$R_N = 1 - [\hat{a}\beta(1, N+1) + \hat{b}\beta(2, N+1)]$$

3.1.2.5 Modelo de acumulación beta binomial

Si Politz (1953) elegía una forma funcional polinómica para aproximar la función de probabilidad de los individuos $f(p)$, con las limitaciones que hemos indicado, Hyett (1958) propone una distribución de probabilidad verdadera, la distribución beta. Las propiedades matemáticas ya señaladas de una distribución de probabilidad verdadera, se traducen una serie de ventajas teóricas en los modelos basados en ellas:

- a. La función es siempre positiva, lo que evita el absurdo de poblaciones negativas.
- b. El área bajo la curva suma uno, con lo que alcances superiores a la población total devienen imposibles.

Hyett (1958) opta por la función beta debido a que es muy flexible: según el valor de los parámetros que la especifican puede ser cóncava o convexa, no tener puntos de inflexión o tener dos, o ser simétrica o asimétrica. Asimismo, depende únicamente de dos parámetros, hecho que reduce a dos los datos de audiencia

acumulada necesarios para su estimación. Finalmente, si $f(p)$ se distribuye según una beta, la distribución de contactos, como veremos, sigue también una distribución compuesta conocida: la beta binomial.

Pues bien, admitamos que $f(p)$ sigue una distribución beta de parámetros A y B , que se concreta en la expresión (López Cachero, 1990; p. 331):

$$f(p) = \frac{1}{\beta(A,B)} p^{A-1}(1-p)^{B-1} \quad [3.25]$$

expresión que puede ponerse en función de la integral euleriana de segunda especie del siguiente modo⁷ (López Cachero, 1990; p. 332):

$$\beta(A, B) = \frac{\Gamma(A)\Gamma(B)}{\Gamma(A+B)} \quad [3.26]$$

Sustituyendo [3.26] en [3.25] y ambas en [3.18], llegamos a:

$$E_K^N = \binom{N}{K} \frac{\Gamma(A+B)}{\Gamma(A)\Gamma(B)} \int_0^1 p^{K+A-1}(1-p)^{N+K+B-1} dp$$

expresión que equiparada con la definición de función beta que se ofreció en [3.20], conduce a:

$$E_K^N = \binom{N}{K} \frac{\Gamma(A+B)}{\Gamma(A)\Gamma(B)} \cdot \frac{\Gamma(K+A)\Gamma(N-K+B)}{\Gamma(A+N+B)} \quad [3.27]$$

Para obtener la cobertura, basta tener en cuenta que para N inserciones, ésta será el total de la población menos aquellos individuos que no hayan resultado expuestos a ninguna de esas N inserciones, esto es:

⁷ El utilizar la función gamma, veremos que es especialmente útil cuando obtengamos la expresión definitiva de la distribución beta binomial, dado que aparecen factoriales donde están implicados los términos A y B , que no tienen porqué ser enteros, lo que imposibilitaría su cálculo.

$$R_N = 1 - E_0^N \quad [3.28]$$

Sustituyendo [3.27] para $K = 0$ en [3.28], y teniendo en cuenta que la función gamma cumple la propiedad: $\Gamma(x) = (x-1)\Gamma(x-1)$, llegamos a la expresión que nos permite obtener el alcance acumulado:

$$R_N = 1 - \frac{(N+B-1)(N+B-2)\dots B}{(A+B+N-1)\dots(A+B)}$$

Chandon (1985) obtiene, sin embargo, una fórmula más operativa para obtener la cobertura, si se tiene en cuenta que para especificar el modelo es necesario disponer de los datos de cobertura tras una y dos inserciones:

$$R_N = R_{N-1} + \frac{B+N-2}{A+B+N-1}(R_{N-1} - R_{N-2})$$

Resta, únicamente, obtener las estimaciones de los parámetros A y B para que el modelo quede definitivamente especificado. Puesto que, por definición:

$$\begin{aligned} R_1 &= E_1^1 \\ R_2 &= E_1^2 + E_2^2 \end{aligned}$$

si obtenemos las expresiones que figuran a la derecha del signo igual mediante [3.27] y resolviendo el sistema para A y B , obtendríamos los estimadores de los parámetros:

$$\begin{aligned} \hat{A} &= \frac{R_1(R_2 - R_1)}{2R_1 - R_1^2 - R_2} \\ \hat{B} &= \frac{\hat{A}(1 - R_1)}{R_1} \end{aligned}$$

3.1.2.6 Modelo de acumulación de la binomial negativa

Greene y Stock (1971) demuestran que la distribución beta binomial, converge a una binomial negativa (DBN) cuando el número de inserciones tiende a infinito. El uso de esta distribución como aproximación de $f(p)$, solo tendrá sentido con N muy grandes, lo que la hace poco útil para planes de medios en periódicos o revistas, pero ha sido sugerido su uso en planes para televisión (Friedman, 1971). Precisamente, el modelo que vamos a desarrollar, no centra su interés en el número de exposiciones a una inserción (fenómeno discreto), sino en la exposición durante un periodo continuado de tiempo: no interesa de N programas consecutivos de televisión, cuánta gente ha visto K de ellos, sino durante un periodo dado de tiempo, cuánta gente ha resultado expuesta cuántas veces. No se razona en términos de probabilidad de exposición, sino de tasa de exposición por unidad de tiempo.

La principal variación en las hipótesis que sustentan al modelo es que, aunque para periodos largos de tiempo se considera que la tasa de exposición de un individuo es constante e independiente de la del resto (hipótesis de estacionariedad para individuos e inserciones tal como ha sido definida por nosotros), para periodos cortos de tiempo la tasa de exposición individual variará en torno a su media μ , siguiendo una distribución de Poisson y para toda la población según una distribución gamma⁸.

Para el desarrollo del modelo DBN, adoptaremos la siguiente notación:

μ_i = Tasa personal promedio de exposición del individuo i .

$f(\mu)$ = Distribución de la tasa personal promedio para toda la población

⁸ Un buen análisis de la distribución de Poisson, cuyas características irán especificándose a lo largo del presente modelo, puede encontrarse en López Cachero (1990) pp. 311-313 y en Cuadras, Echeverría, Mateo y Sánchez (1991), pp. 261-276. Del mismo modo la distribución gamma se encuentra en estos mismos tratados: López Cachero (1990) pp. 329-331 y Cuadras, Echeverría, Mateo y Sánchez (1991) p. 352.

$P(K/\mu)$ = Probabilidad de recibir K exposiciones dada una tasa promedio de exposición igual a μ

Como hemos señalado, $P(K/\mu)$ sigue una distribución de Poisson, cuya función de densidad es (Cuadras, Echeverría, Mateo y Sánchez, 1990; p. 261):

$$f(k) = P[X = k] = e^{-\lambda} \frac{\lambda^k}{k!} \quad [3.29]$$

con $k = 0, 1, 2, \dots$ y donde $\lambda > 0$ es, además del parámetro de la distribución, la esperanza matemática de la misma. Adaptando [3.29] a la notación que venimos empleando, tendremos:

$$P(K/\mu) = e^{-\mu} \frac{\mu^K}{K!}$$

El segundo supuesto que se había enunciado, era que $f(\mu)$ seguía una distribución gamma que, también según Cuadras, Echeverría, Mateo y Sánchez (1990; p. 352), tiene una función de densidad dada por:

$$\begin{aligned} & \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1} & \text{si } x > 0 \\ 0 & & \text{si } x < 0 \end{aligned}$$

con $p > 0$ y $\alpha > 0$ parámetros de la distribución. Adaptando de nuevo la notación a la que hemos venido empleando hasta el momento (concretamente $x = \mu$; $\alpha = b$; $h=1/a$), podemos escribir:

$$f(\mu) = \frac{e^{-\mu/a} \mu^{b-1}}{a^b \Gamma(b)}$$

y dado que la distribución gamma es estrictamente continua, integrando para todas las posibles tasas de exposiciones, obtenemos la probabilidad de ocurrencia de K exposiciones a lo largo del tiempo:

$$P(K/a, b) = \int_0^{\infty} P(K/\mu) f(\mu) d\mu = \frac{\Gamma(K+b)}{\Gamma(K+1)\Gamma(b)} (1+a)^{-b} \left(\frac{a}{1+a}\right)^K \quad [3.30]$$

expresión que es equivalente a E_K^N , teniendo en cuenta las precisiones que se hicieron al principio del desarrollo del modelo y que sigue la mencionada distribución binomial negativa.

Para obtener la expresión del alcance acumulado, se partirá, por un lado, de la esperanza matemática m de la distribución DBN (López Cachero, 1990; p. 322):

$$m = a b \quad [3.31]$$

mientras que, por otro lado, Chandon (1985), demuestra que m equivale a:

$$m = N R_t \quad [3.32]$$

despejando a en [3.31] y sustituyendo en [3.30], se llega a:

$$P(K/m, b) = \frac{\Gamma(K+b)}{\Gamma(K+1)\Gamma(b)} \left(1 + \frac{m}{b}\right)^{-b} \left(\frac{m}{m+b}\right)^K \quad [3.33]$$

Como, por otro lado, sabemos que el alcance acumulado de N inserciones es siempre toda la población (uno en términos porcentuales) menos la aquellos individuos que no han visto ninguna de las N inserciones:

$$R_N = 1 - P(0/m, b) \quad [3.34]$$

calculando la probabilidad que aparece en [3.34] mediante [3.33] y sustituyendo m según [3.32], llegamos a la buscada expresión del alcance acumulado para N inserciones:

$$R_N = 1 - \left(\frac{b}{b + NR_1} \right)^b$$

restando únicamente, para que el modelo quede completamente especificado, estimar los parámetros a y b . Chandon (1985) opta para ello por seguir el método de los momentos, consistente en igualar los momentos de la distribución de frecuencias observada con los momentos de la distribución teórica y resolver para los parámetros desconocidos. Como, para el caso del modelo que se está exponiendo, son necesarios sólo dos parámetros, bastará con calcular la media (m) y varianza (σ^2) de la distribución observada e igualar con sus expresiones teóricas, que en el caso que nos ocupa son:

$$\begin{aligned} m &= ab \\ \sigma^2 &= m(a + 1) \end{aligned}$$

y resolviendo el sistema anterior para los parámetros, obtenemos sus estimadores:

$$\begin{aligned} \hat{a} &= \frac{\sigma^2 - m}{m} \\ \hat{b} &= \frac{m^2}{\sigma^2 - m} \end{aligned}$$

3.1.2.7 Modelo de acumulación de Broadbent

Broadbent (1964), plantea un modelo en el cual, los individuos se agrupan en segmentos con hábitos de exposición similares que no se traducen, como ocurría en los modelos de acumulación de dos y tres segmentos, en probabilidades constantes,

sino que cada segmento tiene su propia distribución de probabilidad individual.

Broadbent (1964), propone la existencia de cuatro segmentos, a saber:

- a. Los *no lectores*, cuya probabilidad de exposición es nula, y representan un ω_0 de la población de interés.
- b. Los *lectores totalmente leales*, cuya probabilidad de exposición es la unidad y que representan un ω_1 por cien de la población.
- c. Los *lectores esporádicos*, que son un ω_2 por cien de los individuos y cuyo número decrece la aumentar la probabilidad de exposición.
- d. Los *lectores habituales*, son un ω_3 por cien de la población, aumentando su número al aumentar la probabilidad de exposición.

Este autor hace el supuesto de que los dos primeros segmentos son homogéneos, mientras que los restantes siguen una distribución potencial. Por tanto, la función $f(p)$ que recoge el número de individuos asociados a cada probabilidad de exposición, adopta en el modelo de Broadbent (1964) la expresión:

$$\begin{array}{ll} \omega_0 & \text{si } p = 0 \\ \omega_2(1+r)(1-p)^r + \omega_3(1+r)p^r & \text{si } 0 < p < 1 \\ \omega_1 & \text{si } p = 1 \end{array} \quad [3.35]$$

donde puede observarse que es necesario estimar cuatro parámetros: ω_0 , ω_1 , ω_2 y r , dado que ω_3 se puede obtener teniendo en cuenta que $\sum \omega_i$ ha de ser necesariamente uno. Se necesitarán, por tanto, datos de alcance acumulado para las cuatro primeras inserciones. Sustituyendo en [3.18] la expresión [3.35] para $0 < p < 1$, llegamos a:

$$E_K^N = \binom{N}{K}(1+r) \left[\omega_2 \int_0^1 p^K (1-p)^{N-K+r} dp + \omega_3 \int_0^1 p^{K+r} (1-p)^{N-K} dp \right] \quad [3.36]$$

dado que en [3.20] ya hemos explicitado la definición de la función beta y en [3.26] su relación con la función gamma, [3.36] puede escribirse del siguiente modo:

$$E_K^N = \binom{N}{K}(1+r) \left[\frac{\omega_2 \Gamma(K+1) \Gamma(N-K+r+1) + \omega_3 \Gamma(K+r+1) \Gamma(N-K+1)}{\Gamma(N+r+2)} \right] \quad [3.37]$$

Para obtener la distribución de contactos de toda la población, basta añadir a la expresión anterior ω_0 si $K = 0$ y ω_1 si $K = N$. Para estimar los parámetros, dado que se debe disponer de los datos cobertura y distribución de contactos de las cuatro primeras inserciones, Chandon (1985), demuestra que resolviendo el sistema resultante de sustituir en [3.37] sucesivamente $E_1^4 \cdots E_4^4$, se llega a:

$$\frac{3(r+2)}{r^2+8r+18} = \frac{E_2^4}{E_1^4+E_2^4+E_3^4}$$

donde tras obtener r iterativamente, con la condición de que $r > 0$, y sustituir en [3.37] sucesivamente para E_1^1, E_0^4, E_4^4 se obtendrían el resto de parámetros del modelo. Chandon (1985), sin embargo, simplifica el modelo considerando únicamente dos segmentos al eliminar los lectores totalmente fieles y los no lectores. En este caso, al hacer $\omega_0 = \omega_1 = 0$ y teniendo en cuenta que cambiamos la notación: $\omega_3 \equiv \omega \rightarrow \omega_2 = 1-\omega$ la expresión [3.37] puede escribirse teniendo en cuenta la propiedad [3.26] como:

$$E_K^N = \binom{N}{K}(1+r) [(1-\omega) \beta(K+1, N-K+r+1) + \omega \beta(K+r+1, N-K+1)] \quad [3.38]$$

sustituyendo uno de los dos únicos datos necesarios para estimar los parámetros: E_1^2 , se llega a:

$$2(R_2 - R_1) = \frac{2(r+1)}{(r+2)(r+3)}$$

ecuación que nos permitirá obtener la estimación de r y que sustituida en [3.38] para E_2^2 nos lleva a la expresión del segundo parámetro:

$$\hat{\omega} = \frac{R_1(\hat{r}+2)-1}{\hat{r}}$$

3.1.2.8 Modelo de acumulación de Morgensztern

El modelo desarrollado por Morgensztern (1970), al igual que el de Broadbent (1964), considera la existencia de segmentos, aunque cada uno de ellos con su propia distribución de probabilidad, no con un valor constante de la misma. Parte de la definición de D_N , que es el número de sujetos alcanzados por la primera de N inserciones, que también resultan expuestos a la $N+1$ ésima. Si suponemos que la audiencia del soporte es aproximadamente constante entre dos números o dos programas, y llamamos:

A_N = audiencia del número N

$\Delta_N = R_{N+1} - R_N$ o incremento del alcance acumulado producido por la inserción $N+1$ ésima

entonces es evidente que, teniendo en cuenta la hipótesis de audiencia constante:

$$D_N = A_N - \Delta_N = A_1 - (R_{N+1} - R_N)$$

proponiendo a partir de aquí Morgensztern (1970) la siguiente expresión para la fórmula de la distribución de contactos⁹:

⁹ Se ha de tener presente que la notación ha de sufrir cambios, ya que los hay también conceptualmente: no se refiere el autor al porcentaje de individuos expuestos a K de N inserciones realizadas, sino los individuos expuestos a K inserciones de entre los que al menos han resultado expuestos a 1 de esas N .

$$I_K^N = (R_{K+1} - R_K) + (R_{N-K-1} - R_{N-K}) \quad [3.39]$$

expresión que Chandon (1985) demuestra que:

- a. Es inexacta en el cálculo de E_K^N cuando se aproxima por I_K^N
- b. Es simétrica ($I_K^N = I_{N-K}^N$) lo que no es una propiedad deseable, dado que no hay ninguna razón para creer que el número de personas expuestas a K inserciones de las N realizadas, debe ser el mismo que los expuestos a $N-K$ de las mismas.
- c. La suma de los I_K^N para todos los K no suma toda la población.

Morgensztern (1970) va rectificando el modelo para solventar las anteriores limitaciones. Así, postula que las curvas empíricas de alcance acumulado, pueden aproximarse con una función exponencial idéntica a [3.2] que, sustituida en [3.39] lleva a (siendo D la constante $a(1-b)$):

$$I_K^N = D(b^K + b^{N-K}) \quad [3.40]$$

la simetría puesta de manifiesto por Chandon (1985) es evidente en esta expresión y es corregida por Morgensztern (1970) introduciendo dos coeficientes a y c :

$$I_K^N = D(ab^K + cb^{N-K}) \quad [3.41]$$

y, a continuación, para evitar que la distribución de contactos no sume la población total, fuerza a que D tome aquel valor que lleve a:

$$\sum_{K=1}^N l_K^N = 1$$

quedando la fórmula de Morgensztern (1970) con estos cambios:

$$l_K^N = \frac{b-1}{b(a+1)(b^{N-1})} (ab^K + b^{N-K+1}) \quad [3.42]$$

que no goza de las propiedades indeseables que se pusieron de manifiesto en las expresiones anteriores.

A partir de aquí, Morgensztern (1970) comienza a introducir en su expresión los segmentos para recoger la heterogeneidad de la población, postulando la existencia de dos: los fieles, y los ocasionales. Supone que los que resultan expuestos a una cantidad determinada de números o programas, son proporcionales a los que resultan expuestos a una cantidad mayor de los mismos, expresando este supuesto matemáticamente como sigue:

$$L_K^N = \rho(L_K^N + L_K^{N+1} + \dots + L_N^N) \equiv V_K$$

$$O_K^N = \mu(O_K^N + O_K^{N+1} + \dots + O_N^N) \equiv W_K$$

donde hemos llamado L_K^N al número de sujetos leales expuestos a K de las N inserciones realizadas, y O_K^N al número de lectores, oyentes o televidentes ocasionales en las mismas condiciones. La formulación anterior, equivale a asumir que la distribución de contactos adopta, para los dos segmentos, la siguiente forma:

$$L_K^N = V_1 \left(\frac{\rho}{1-\rho} \right) (1-\rho)^K$$

$$O_K^N = W_1 \left(\frac{\mu}{1-\mu} \right) (1-\mu)^K \quad [3.43]$$

donde $V_1 = \sum_{i=1}^N L_i^N$ y $W_1 = \sum_{i=1}^N O_i^N$ son el total de lectores fieles y esporádicos.

Para facilitar el cálculo de los parámetros que especificarán el modelo, Morgensztern (1970), realiza una serie de simplificaciones adicionales acerca de ellos. Así, hipotetiza que la proporción entre lectores leales y ocasionales, es la constante a independiente de la cantidad de inserciones consideradas:

$$a = \frac{V_1}{W_1}$$

y, por otra parte, sobre las siguientes proporciones obtenidas a partir de [3.43]:

$$\frac{L_{K+1}^N}{L_K^N} = 1 - \rho$$

$$\frac{O_{K+1}^N}{O_K^N} = 1 - \mu$$

realiza el supuesto de que son recíprocas y relacionadas con el parámetro b de este modo:

$$1 - \rho = \frac{1}{1 - \mu} = b$$

de este modo, el modelo de Morgensztern (1970), acaba dependiendo de dos parámetros a y b , cuyo significado es el siguiente:

- a. a representa la proporción entre lectores leales y ocasionales.
- b. b es la *ratio* entre dos términos consecutivos de la distribución de contactos de los lectores fieles, siendo $1 / b$ la misma *ratio* para los lectores que se han denominado ocasionales.

Una vez estimados los parámetros, lo que se verá inmediatamente, las distribuciones de contactos de ambos tipos de lectores y la conjunta, podrá obtenerse del siguiente modo:

$$L_K^N = \frac{a(b-1)b^{K-1}}{(a+1)(b^N-1)}$$

$$O_K^N = \frac{(b-1)b^{K-1}}{(a+1)(b^N-1)}$$

$$I_K^N = L_K^N + O_K^N = \frac{b-1}{b(a+1)(b^N-1)}(ab^K + b^{N-K+1})$$

Sin embargo, el modelo analizado, no puede limitarse al cálculo de I_K^N , dado que como el resto de modelos tratados ha de ser capaz de ofrecer la distribución de contactos que compute también los individuos no alcanzados, lo que hemos venido denotando como E_K^N . La extensión, sin embargo, es sencilla. Anteriormente, obtuvimos un coeficiente normalizador D de tal forma que se cumpliese:

$$\sum_{K=1}^N I_K^N = 1$$

basta, ahora, buscar un nuevo coeficiente D que normalice las expresiones de tal forma que incluya la distribución de frecuencias completa:

$$\sum_{K=0}^N E_K^N = 1 \quad [3.44]$$

Por analogía, con [3.42], E_K^N puede ponerse como:

$$E_K^N = D(ab^K + b^{N-K+1})$$

donde D es la constante que habrá que normalizar para que se cumpla [3.44].

Sustituyendo, pues, en [3.44] y despejando D , se llega a:

$$D = \frac{b-1}{(a+b)(b^{N+1}-1)}$$

de tal forma que la expresión definitiva de la distribución de contactos para toda la población, queda:

$$E_K^N = \frac{b-1}{(a+b)(b^{N+1}-1)}(ab^K + b^{N-K+1}) \quad [3.45]$$

Resta por obtener, únicamente, la expresión del alcance acumulado. En se momento, se definió I_K^N como el porcentaje de individuos expuestos a K inserciones de los que, al menos, habían estado expuestos a 1 de ellas. Formalizando matemáticamente esta expresión, se llegaría a:

$$I_K^N = \frac{E_K^N}{R_N} \rightarrow R_N = \frac{E_K^N}{I_K^N}$$

sustituyendo en esa expresión [3.42] y [3.45], llegamos a la expresión de la cobertura:

$$R_N = \frac{b(a+1)(b^N-1)}{(a+b)(b^{N+1}-1)} \quad [3.46]$$

Para que el modelo quede completamente especificado, basta con que se obtengan los estimadores de los dos parámetros a y b , para lo que son necesarios los datos de alcance acumulado de las dos primeras inserciones: R_1 y R_2 . Sustituyendo esa información sucesivamente en [3.46] y resolviendo el sistema por reducción, llegamos a:

$$b^3(R_2 - R_1) - b^2R_1 + bR_1 + (R_1 - R_2) = 0 \quad [3.47]$$

La estimación \hat{b} así obtenida, sustituida en [3.46] para $N = 1$, nos permite obtener la estimación de a :

$$\hat{a} = \frac{\delta [1 - R_1(1 + \delta)]}{R_1(\delta + 1) - \delta}$$

El problema que aparece en estas estimaciones, es que en la formulación del modelo realizada por Mogensztern, b debe ser mayor que 1, y la expresión [3.47] no garantiza este hecho. Esto provoca que el modelo no sea aplicable a todos los soportes, sino solamente a aquellos en los que el alcance acumulado de las dos primeras inserciones guarden determinada relación que se deduce directamente de [3.47]. Si esta expresión cambiara de signo en el intervalo $(1, +\infty)$ aseguraría un punto de corte con el eje x en ese intervalo, obteniéndose así un valor de b que cumpla que $b > 1$. Para que esto ocurra, la primera derivada de [3.47] debe anularse en $1 < b < +\infty$. Derivando e igualando a cero, obtenemos:

$$b = \frac{-2R_1 \pm \sqrt{(2R_1)^2 - 12(R_1 - R_2)(-R_1)}}{6(R_1 - R_2)}$$

como para que b sea un número real, el radicando debe ser positivo, resolviendo la inequación resultante obtenemos la **condición de aplicabilidad del modelo**:

$$R_2 \leq \frac{4}{3}R_1$$

3.1.2.9 Modelo de acumulación beta binomial modificado

El modelo de acumulación beta binomial propuesto por Hyett (1958) que ya se ha desarrollado en este mismo capítulo, tiene la limitación de que la función beta binomial, al ser una distribución absolutamente continua, da un peso nulo a las probabilidades de exposición $p = 0$ y $p = 1$, lo que significa que ningún individuo de la población puede ser no lector o lector fiel. El modelo beta binomial modificado (Chandon, 1985) trata de corregir el modelo anterior considerando explícitamente a esos dos grupos.

Se considera a la población objeto de estudio formada por dos tipos de individuos: los lectores fieles, que son un porcentaje ω de la población y, por lo tanto, con una probabilidad personal de exposición igual a la unidad y, por otra parte, los lectores probables, en este caso un $1 - \omega$ por cien de la misma. Las probabilidades de exposición de estos últimos se distribuyen según una beta binomial de parámetros A y B , como se resume en el cuadro 3.1.

Cuadro 3.1. Características de la población del modelo beta binomial modificado para dos segmentos

Segmento	p	Distribución de probabilidad poblacional
Fieles	1	$f(1) = \omega$
Probables	$0 \leq p < 1$	$f(p) = (1 - \omega) \frac{\Gamma(A+B)}{\Gamma(A)\Gamma(B)} p^{A-1} (1-p)^{B-1}$

Fuente: Chandon (1985) p. 239

Para obtener la distribución de contactos, bastará aplicar las expresiones obtenidas en su momento para la distribución beta binomial [3.27] a un porcentaje $1 - \omega$ de la población, lo que conducirá a:

$$\begin{aligned}
 E_K^N &= \omega + (1 - \omega) \frac{\Gamma(A+K)}{\Gamma(A+B+N)} \frac{\Gamma(A+B)}{\Gamma(A)} \quad \text{si } K = N \\
 E_K^N &= (1 - \omega) \binom{N}{K} \frac{\Gamma(A+K) \Gamma(B+N-K) \Gamma(A+B)}{\Gamma(A) \Gamma(B) \Gamma(A+B+N)} \quad \text{si } K < N
 \end{aligned}
 \tag{3.48}$$

donde $0 \leq K \leq N$ han de ser enteros y $A, B < 0$ reales. Dado que es conocido que el alcance acumulado tras N inserciones es la unidad menos aquellos que no han resultado expuestos a ninguna de las N inserciones:

$$R_N = 1 - E_0^N$$

sustituyendo en la expresión anterior [3.48] para $K = 0$, llegamos a la obtención de la cobertura como sigue:

$$R_N = 1 + (\omega - 1) \frac{\Gamma(B+N) \Gamma(A+B)}{\Gamma(B) \Gamma(A+B+N)}$$

Para estimar los tres parámetros que especifican el modelo: A , B y ω , son necesarios los datos de tres inserciones, que sustituidos sucesivamente en [3.48] componen un sistema de tres ecuaciones, de los cuales obtenemos las estimaciones:

$$\hat{B} = \frac{2(E_0^2)^2 - E_0^3(E_0^1 + E_0^2)}{E_0^1 E_0^3 - (E_0^2)^2}$$

$$\hat{A} = \frac{(\hat{B} + 1)(E_0^1 - E_0^2)}{E_0^2}$$

$$\hat{\omega} = \frac{(\hat{A} + \hat{B})E_1^1 - \hat{A}}{\hat{B}}$$

expresiones en las que quedan implícitas las condiciones de aplicabilidad del modelo, dado que los tres parámetros del modelo deben ser positivos:

$$E_0^1 > E_0^2 > E_0^3$$

$$\hat{A} < E_1^1(\hat{A} + \hat{B})$$

3.1.2.10 Modelo de acumulación binomial de Markov

Este modelo deja de asumir la hipótesis de estacionariedad. Concretamente, el modelo de acumulación binomial de Markov (Chandon, 1985), elimina la estacionariedad respecto a las inserciones basándose en las siguientes hipótesis:

Hipótesis Primera. La probabilidad de exposición a un número (programa de radio, programa de TV), está condicionada a la exposición o no al número anterior del siguiente modo:

- a. Si un individuo estuvo expuesto al número anterior, tiene la misma probabilidad de resultar expuesto al actual que tuvo de estar expuesto a aquel.
- b. Si no estuvo expuesto al número anterior, tiene la misma probabilidad de resultar expuesto al presente que tenía respecto al anterior.
- c. Las dos probabilidades anteriores son distintas.

Hipótesis Segunda. La probabilidad de resultar expuesto al número actual, es mayor si se estuvo expuesto al anterior que si no se estuvo.

Hipótesis Tercera. La audiencia entre números es constante.

Para el desarrollo del modelo, se utilizará la siguiente notación:

$P(1/1) = d$	Probabilidad de exposición a un número si se estuvo expuesto al anterior
$P(0/1)$	Probabilidad de no exposición a un número si sí se estuvo expuesto al anterior
$P(1/0)$	Probabilidad de exposición a un número cuando no se estuvo expuesto al anterior
$P(0/0)$	Probabilidad de no exposición a un número cuando no se estuvo expuesto al anterior
p	Probabilidad no condicionada de exposición a un número

q Probabilidad no condicionada de no exposición a un número

Por construcción, debe cumplirse que:

$$\begin{aligned} p + q &= 1 \\ P(1/1) + P(0/1) &= 1 \\ P(1/0) + P(0/0) &= 1 \end{aligned} \quad [3.49]$$

además, como por la hipótesis tercera, la audiencia ha de permanecer constante de un número a otro, la probabilidad p deberá ser constante. Utilizando este resultado, podemos escribir:

$$\begin{aligned} p &= p \times P(1/1) + q \times P(1/0) \\ q &= p \times P(0/1) + q \times P(0/0) \end{aligned} \quad [3.50]$$

Puesto que el modelo quedará perfectamente especificado conociendo p y d , es necesario encontrar sus estimadores. Así, Chandon (1985), considera que el estimador de la probabilidad de exposición aun soporte \hat{p} es, simplemente, la *ratio* entre la audiencia de un número y el total de la población, esto es, R_1 en términos porcentuales:

$$\hat{p} = R_1 \quad [3.51]$$

Por otra parte, para este mismo autor, el estimador de la probabilidad de exposición a un soporte, cuando no se estuvo expuesto al número anterior de ese soporte, d , es la *ratio*, entre la audiencia duplicada entre dos números sucesivos y la audiencia promedio:

$$\hat{d} = \frac{2R_1 - R_2}{R_1} \quad [3.52]$$

Sustituyendo los estimadores [3.51] y [3.52] en las expresiones [3.49] y [3.50] se llega a:

$$\begin{aligned}
 P(1/1) &= d \\
 P(0/1) &= 1 - d \\
 P(0/0) &= \frac{1 - 2p + pd}{q} \\
 P(1/0) &= \frac{(1 - d)p}{q}
 \end{aligned}
 \tag{3.53}$$

La probabilidad de que un individuo no resulte expuesto a ninguna de los N anuncios insertados, puede ponerse como la probabilidad de no resultar expuesto al primero por la probabilidad de no resultar expuesto a uno si no se estuvo expuesto al anterior, así $N - 1$ veces:

$$P(0 \text{ de } N) = q P(0/0) P(0/0) \cdots N - 1 \text{ veces}$$

$$P(0 \text{ de } N) = q [P(0/0)]^{N-1}$$

sustituyendo $P(0/0)$ por la expresión [3.53], podemos poner:

$$P(0 \text{ de } N) = q \left[\frac{1 - 2p + dp}{q} \right]^{N-1} = q \left[\frac{1 - 2R_1 + R_1 \frac{2R_1 - R_2}{R_1}}{q} \right]^{N-1} = q \left[\frac{1 - R_2}{1 - R_1} \right]^{N-1}$$

donde, dado que $q = 1 - d = 1 - R_1$, nos queda:

$$P(0 \text{ de } N) = (1 - R_1) \left[\frac{1 - R_2}{1 - R_1} \right]^{N-1}$$

como todos los individuos son iguales, bastará multiplicar la población total por la anterior probabilidad para obtener los individuos que no han resultado expuesto a

ninguna de las N inserciones realizadas y, a partir de aquí, obtener la expresión del alcance acumulado:

$$R_N = 1 - E_0^N = POB(1 - P(0 \text{ de } N)) =$$
$$R_N = POB \left[1 - (1 - R_1) \left[\frac{1 - R_2}{1 - R_1} \right]^{N-1} \right]$$

Chandon (1985) solamente es capaz de obtener la expresión para el alcance acumulado. Por este motivo ha de recurrirse a un procedimiento general para convertir la distribución de contactos en alcances acumulados, el cual hemos desarrollado en el epígrafe 4.2.1.

3.2 MODELOS DE DUPLICACIÓN DE AUDIENCIAS

3.2.1 Modelos ad hoc

Como ya se indicó en el capítulo anterior, los modelos *ad hoc* consideran la exposición como un fenómeno determinístico, por lo que se hace innecesario formular hipótesis sobre si todos los individuos tienen la misma probabilidad de resultar expuestos a un anuncio inserto en un soporte dado o si, por otra parte, los soportes pueden considerarse iguales respecto a las probabilidades de exposición que generan. Estos modelos, mediante técnicas de regresión lineal, se limitan a ajustar una función a los datos conocidos de cobertura de un plan de medios dado, que es utilizada posteriormente para predecir esta misma variable en otros planes. Estos modelos no pueden, por el contrario, estimar la distribución de frecuencias de los planes de medios cuya cobertura calculan.

3.2.1.1 Modelo de duplicación de Agostini

Agostini (1961), presenta este modelo como un procedimiento de estimación de la cobertura de un plan de medios constituido por una inserción en diversos soportes, cuando solamente se dispone de la audiencia individual A_i de cada uno de ellos. Sea A la audiencia bruta (sin tener en cuenta duplicaciones) resultante de realizar una inserción en m soportes distintos:

$$A = \sum_{i=1}^m A_i \quad [3.54]$$

si llamamos A_{ij} a la audiencia duplicada de los soportes i y j :

$$A_{ij} = A_i \cap A_j$$

y llamamos D a la suma de todas las audiencias duplicadas:

$$D = \sum_{i=1}^{m-1} \sum_{j=i+1}^m A_{ij} \quad [3.55]$$

la cobertura R_m del plan de medios deberá ser, debido a las duplicaciones, siempre inferior a la audiencia bruta A . También es evidente que la *ratio* R_m / A deberá ser igual a la unidad cuando no hayan duplicaciones ($R_m = A$) y disminuir cuando las duplicaciones crezcan. Esta serie de propiedades lógicas, son recogidas por Agostini (1961) en la función:

$$\frac{R_m}{A} = \frac{1}{1+k\frac{D}{A}} \quad [3.56]$$

donde el único elemento cuya notación no se ha indicado es la constante k . Sustituyendo [3.54] y [3.55] en [3.56] y despejando R_m , se llega a:

$$R_m = \frac{\left(\sum_{i=1}^m A_i \right)^2}{\sum_{i=1}^m A_i + k \sum_{i=1}^{m-1} \sum_{j=i+1}^m A_{ij}} \quad [3.57]$$

Puesto que Agostini (1961) dispone de los datos de audiencia y de las duplicaciones de 30 revistas gracias a un estudio elaborado en 1957 por el *Centre d'Etude des Supports de Publicité (CESP)*, y aprovechando el hecho de que la cobertura puede ser escrita en función de las duplicaciones, triplicaciones y, en general, las m -plicaciones como sigue¹⁰:

$$R_m = A - \Sigma \text{duplic} + \Sigma \text{triplic} - \dots \pm \Sigma m\text{-plicaciones} \quad [3.58]$$

¹⁰ Puede consultarse Chandon (1985), pág. 317 para una ilustración de esta fórmula.

ajustando la expresión [3.57] a los datos disponibles, estima el parámetro $k = 1,125$. A partir de este momento la cobertura de cualquier plan de medios formado por una inserción en m soportes distintos ya puede ser calculada mediante [3.57] sin más datos que las audiencias individuales de los soportes y las duplicaciones entre los mismos.

El procedimiento de estimación de k mediante el ajuste de [3.57] a una base de datos determinada, ya plantea a Agostini (1961; pág. 14) la cuestión de que “no hay prueba alguna de que la curva estimada sea válida para otros soportes distintos a los que sirvieron para construirla”. Así, Kuhn (1963), Bower (1963), Caffyn y Sagovsky (1963), Marc (1963) y mucho más recientemente Redondo (1994) obtienen otros valores para k utilizando distintas bases de datos. Desde otra perspectiva, Hofmans (1966), basa su modelo, como veremos inmediatamente, en la demostración de que la constante del modelo de Agostini no es tal constante. Todas estas limitaciones, si embargo, no deben ocultarnos que, si alguna virtud tienen los modelos *ad hoc*, no es precisamente la coherencia teórica, dado que ésta resulta imposible cuando no se formulan hipótesis explícitas respecto al comportamiento de los individuos o los soportes frente a las inserciones. Es más, esta limitación resultaría de poca importancia si el modelo fuera capaz de predecir de una manera ajustada la cobertura frente al hecho de que es incapaz de obtener la segunda variable fundamental en la planificación de medios: la distribución de contactos.

3.2.1.2 Modelo de duplicación de Hofmans

Como acabamos de indicar, el modelo de Hofmans (1966) se basa en la demostración de que el parámetro k de la expresión [3.57] no es una constante. Tras deducir una expresión alternativa a la misma (k_{ij}), el autor modifica la fórmula de Agostini con la introducción de ese parámetro dando lugar al modelo que a continuación se analiza.

Aplicando la fórmula de Agostini al caso particular de dos soportes, obtendríamos:

$$R_2 = \frac{(A_1 + A_2)^2}{A_1 + A_2 + kA_{12}} \quad [3.59]$$

mientras que, por otro lado, la cobertura puede calcularse de una manera exacta mediante [3.58], obteniéndose para dos soportes:

$$R_2 = A_1 + A_2 - A_{12} \quad [3.60]$$

igualando [3.59] y [3.60] se llega fácilmente a que:

$$k = \frac{A_1 + A_2}{A_1 + A_2 - A_{12}} \quad [3.61]$$

si el plan de medios que se está analizando consta de m soportes, de la expresión anterior se deduce claramente que k no es una constante, dado que dependerá de con qué par de los m existentes sea calculada. Definiendo k_{ij} como la particularización de [3.61] para cada par de soportes y sustituyendo en [3.57] obtenemos la fórmula de Hofmans (1966):

$$R_m = \frac{\left(\sum_{i=1}^m A_i \right)^2}{\sum_{i=1}^m A_i + \sum_{i=1}^{m-1} \sum_{j=i+1}^m k_{ij} A_{ij}} \quad [3.62]$$

donde k_{ij} se obtiene:

$$k_{ij} = \frac{A_i + A_j}{A_i + A_j - A_{ij}}$$

3.2.1.3 Modelo de duplicación de Little y Lodish

Little y Lodish (1969) desarrollan un modelo completo de planificación de medios, el MEDIAC, en cuyo marco han de resolver el problema de estimar la cobertura de planes formados por una inserción en diversos soportes. Como se verá a continuación, la formulación de hipótesis respecto a las probabilidades de exposición de los individuos frente a los soportes, convertiría a este modelo en estocástico, según lo expuesto en el capítulo segundo de esta tesis. Sin embargo, la determinación mediante procedimientos de regresión del parámetro b que especifica el modelo, que se aplica indistintamente de cuáles sean los soportes concretos que constituyen el plan de medios, hacen que esta fórmula devenga en un modelo *ad hoc*.

Sea p_i la probabilidad de que un individuo resulte expuesto al soporte i y p_{ij} la probabilidad de que resulte expuesto simultáneamente a los soportes i y j . Estas probabilidades se obtendrán empíricamente sin más que dividir los datos de audiencia y duplicación por el total de individuos que forman la población objetivo y que denotaremos como POB . Inicialmente, Little y Lodish (1969), consideran que la duplicación es un fenómeno aleatorio, con lo que, bajo esta hipótesis de independencia estocástica de sucesos¹¹, la probabilidad de resultar expuesto a los dos soportes sería el producto de las probabilidades individuales, esto es:

¹¹ Esta propiedad, que aplicaremos con frecuencia en modelos sucesivos, se basa en la aplicación a la expresión general de probabilidad condicionada:

$$P(A \cap B) = P(A/B) \cdot P(B)$$

de la consecuencia que se deriva de la independencia de sucesos, esto es que la ocurrencia previa de B no afecta a la probabilidad de que acaezca A :

$$P(A/B) = P(A)$$

Continúa en la página siguiente...

$$p_{ij} = p_i \times p_j \quad [3.63]$$

Los autores corrigen, a continuación, esta propiedad considerando que la forma funcional que rige la relación entre las probabilidades no es [3.63] sino [3.64], tomando b , tras ajustar la curva a los datos de que disponen, el valor $b = 0,65$.

$$p_{ij} = (p_i \times p_j)^b \quad [3.64]$$

Es, como hemos comentado, esta determinación del parámetro b , lo que convierte al modelo de Little y Lodish (1969) en *ad hoc*. La principal limitación del modelo comentado, radica en que los autores no fueron capaces de derivar una expresión que pudiera calcular la cobertura de planes de medios que incluyesen más de dos soportes. Para ese caso, la expresión deducida es:

$$R_2 = POB[(p_i + p_j) - (p_i \times p_j)^{0,65}]$$

3.2.1.4 Fórmula de duplicación de Starch

Starch (1967) deduce un procedimiento para el cálculo de la cobertura de un plan de medios que se ha clasificado como *ad hoc* al no explicitarse las hipótesis en las que se basa la fórmula obtenida. Sea $R_{i/n}$ el incremento en la cobertura obtenido al añadir el soporte i a los n que ya formaban el plan de medios analizado. Este incremento puede calcularse, según Starch (1967) mediante la expresión:

con lo que se cumple que:

$$P(A \cap B) = P(A) \cdot P(B)$$

Puede consultarse, al respecto, Cuadras, Echevarría, Mateo y Sánchez (1991), pp. 110-119

$$R_{i/n} = \left(\frac{A_i - A_{in}}{A_i}\right) \left(\frac{A_i - A_{in-1}}{A_i}\right) \dots \left(\frac{A_i - A_{i1}}{A_i}\right) A_i \quad [3.65]$$

donde A_i es la audiencia del soporte i y A_{ij} ($j = n \dots 1$) es la duplicación entre los soportes i y j . Nótese que n varía desde 1 hasta i que es el soporte cuyo aporte en términos de cobertura se está calculando. La cobertura de los m soportes que forman el plan de medios se obtendrá por la simple adición de los aportes marginales, esto es:

$$R_m = \sum_{i=1}^m R_{i/n}$$

Como se observa, estamos ante un procedimiento secuencial que va introduciendo sucesivamente los m soportes que forman el plan de medios. Este hecho plantea a Chandon (1985) la duda de que el dato de cobertura R_m obtenido, llegue a depender del orden en que se introduzcan los soportes. Tras demostrar que esta limitación del modelo efectivamente está presente, concluye que el error se minimiza introduciendo siempre aquel soporte que presente una *ratio* α más pequeña:

$$\alpha = \frac{\sum_{i=1}^{n-1} A_{ij}}{A_j} \quad [3.66]$$

3.2.1.5 Modelo de Lozano

Lozano (1985), basándose en su conocimiento de los datos de audiencia como director-gerente del EGM en España, propone la expresión [3.67] como procedimiento para calcular la cobertura de un plan de medios formado por una

inserción en m soportes diferentes. Aunque el autor trata de revestir a la expresión de una fundamentación teórica consistente, los errores matemáticos cometidos en su deducción, unido a las hipótesis poco verosímiles en que la fundamenta¹² nos ha llevado a clasificar este modelo como *ad hoc*.

La fórmula propuesta es:

$$R_m = \frac{A}{\frac{A}{POB} + \frac{2}{3} - \frac{1}{3} \log \frac{100 \times A}{m \times POB}} \quad [3.67]$$

siendo A la audiencia bruta de los m soportes obtenida como suma simple de las audiencias individuales y POB la población objetivo del plan de medios.

3.2.2 Modelos estocásticos

Como ya se comentó en el capítulo segundo, a diferencia de los modelos *ad hoc*, los modelos estocásticos establecen un sistema de hipótesis relativas a las probabilidades de exposición de los individuos respecto a un conjunto de soportes. La hipótesis fundamental (Chandon, 1985) es que la exposición a un anuncio es un fenómeno aleatorio (no determinístico) y, por lo tanto, sujeto a las leyes de la probabilidad. El resto de hipótesis, que diferencian a unos modelos de otros, hacen referencia a:

- a. Si cada individuo tiene o no la misma probabilidad de resultar expuesto a un soporte (**probabilidades individuales de exposición homogéneas o heterogéneas respectivamente**).

¹² Entre otras que la audiencia de un soporte en un momento dado es igual al número de individuos que en alguna ocasión han resultado expuestos al mismo.

- b. Si todos los soportes tienen la misma probabilidad de generar contactos (**probabilidades de exposición del soporte** homogéneas o heterogéneas caso de no ser así).
- c. Si los soportes pueden o no ser considerados independientes unos de otros, es decir si la **duplicación entre soportes es un fenómeno aleatorio** o si, por el contrario, la exposición a un soporte modifica la probabilidad de exposición a otro.

Puesto que en el capítulo segundo se clasificaron todos los modelos que van a desarrollarse en función de las hipótesis señaladas, a continuación se desarrollaran todos ellos sin solución de continuidad.

3.2.2.1 *Modelo binomial*

Este modelo, desarrollado por Chandon (1985), se basa en las siguientes hipótesis:

- a. Probabilidades de exposición del soporte homogéneas:

$$A_i = A_j = \dots = A_m = A$$

- b. Probabilidades individuales de exposición (individuos a, b, etc.) son homogéneas:

$$P_{a,i} = P_{b,i} = \dots = P_i \quad \forall i \in M \quad \forall a, b \in POB$$

- c. Duplicación aleatoria: la exposición a un soporte no modifica la probabilidad de resultar expuesto a otro:

$$P(A_i/A_j) = P(A_i) \cdot P(A_j) \quad \forall i, j \in M; \quad i \neq j$$

siendo A_i la audiencia del soporte i , $P_{a,i}$ la probabilidad de que el individuo a , forme parte de la audiencia del soporte i , M es el conjunto de los m soportes que forman el plan de medios, POB , la población objetivo del mismo y $P(A_i)$, la probabilidad a priori de resultar expuesto al soporte i .

Uniendo las dos primeras hipótesis, obtenemos que la probabilidad de exposición de cualquier individuo a un soporte determinado es:

$$P = \frac{\bar{A}}{POB}$$

donde el numerador se calcula como la media de las audiencias de cada soporte, esto es:

$$\bar{A} = \frac{1}{m} \sum_{i=1}^m A_i$$

y por la hipótesis c la probabilidad de exposición es una variable bernouilli¹³: cada soporte es un suceso con probabilidad de ocurrencia p , por ello la probabilidad de exposición a k soportes será:

$$P(k/m, p) = \binom{m}{k} p^k (1-p)^{m-k}$$

y la probabilidad de resultar expuesto a k o más exposiciones es:

$$P(x \geq k/m, p) = \sum_{x=k}^m \binom{m}{x} p^x (1-p)^{m-x} \quad [3.68]$$

¹³ Una explicación más detallada de qué es una distribución bernouilli y de cómo ésta sirve de base a la distribución binomial, se expuso en este mismo capítulo al desarrollar el modelo binomial de acumulación de audiencias.

como todos los individuos son idénticos, los que han resultado expuestos a k de los m soportes (definición de distribución de contactos, que denotamos como E_k^m) serán:

$$E_k^m = POB \cdot P(k/m, p) = POB \cdot \binom{m}{k} p^k (1-p)^{m-k}$$

La cobertura, por su parte, vendrá dada por el número de individuos que han resultado expuestos a, al menos, un soporte de los m . , Dado que por definición R_m , en término de individuos, no de probabilidades, puede escribirse como:

$$R_m = POB \cdot P(x \geq 1/m, p)$$

basta sustituir $k = 1$ en [3.68] para llegar a la expresión que nos permite obtener la cobertura del plan de medios:

$$R_m = POB - \frac{(POB - \bar{A})^m}{POB^{m-1}}$$

3.2.2.2 Modelo de Sainsbury

Este método, tanto en la versión original como en la modificada, fue desarrollado por E. J. Sansbury en la *London Press Exchange*, es formalizado y analizado por Caffyn y Sagovsky (1963) quienes lo contraponen al ya analizado modelo *ad hoc* de Agostini (1961).

Las hipótesis en las que se basa, son las mismas que las que subyacían en el modelo binomial, con la salvedad de que se relaja la homogeneidad de las probabilidades de exposición, es decir:

- a. Probabilidades de exposición del soporte heterogéneas:

$$A_i \neq A_j \neq \dots \neq A_m \quad \forall i, j \in M$$

- b. Probabilidades individuales de exposición (individuos a, b, etc.) son homogéneas:

$$P_{a,i} = P_{b,i} = \dots = P_i \quad \forall i \in M \quad \forall a, b \in POB$$

- c. Duplicación aleatoria: la exposición a un soporte no modifica la probabilidad de resultar expuesto a otro:

$$P(A_i/A_j) = P(A_i) \cdot P(A_j) \quad \forall i, j \in M; \quad i \neq j$$

De las dos primeras hipótesis se desprende que la probabilidad de que un individuo resulte expuesto al soporte i , vendrá dado por:

$$P_i = \frac{A_i}{POB}$$

mientras que por la hipótesis c., la probabilidad de exposición continuará siendo una variable bernouilli, pero con diferentes probabilidades de exposición en cada soporte. Si llamamos S_r a la suma de las probabilidades de exposición de los individuos que han resultado expuestos a r soportes:

$$\begin{aligned} S_1 &= \sum_{i=1}^m P_i \\ S_2 &= \sum_{i=1}^{m-1} \sum_{j=i+1}^m P_{ij} \\ &\dots \\ S_r &= \sum_{i=1}^{m-r+1} \sum_{j=i+1}^{m-r+2} \dots \sum_{r=m-1}^m P_{ij\dots r} \end{aligned} \quad [3.69]$$

Chandon (1985) demuestra que por el Teorema de Waring, la probabilidad de que un individuo resulte expuesto a exactamente k soportes es:

$$P(k/m, P_1, P_2 \cdots P_m) = \sum_{i=k}^m (-1)^{i-k} \binom{i}{k} S_i \quad 0 \leq k \leq m \quad S_0 = 1 \quad [3.70]$$

y a al menos k soportes:

$$P(k \geq m, P_1, P_2 \cdots P_m) = \sum_{i=k}^m (-1)^{i-k} \binom{i-1}{k-1} S_i \quad 0 \leq k \leq m \quad S_0 = 1$$

como por la hipótesis b el número de individuos expuestos a k soportes, es decir la distribución de contactos, puede obtenerse multiplicando las probabilidades individuales por el total de la población objetivo, tenemos que:

$$E_k^m = POB \cdot P(k/m, P_1, P_2 \cdots P_m)$$

y análogamente, los que han resultado expuestos a, al menos, k soportes:

$$E_{x \geq k}^m = POB \cdot P(x \geq k/m, P_1, P_2 \cdots P_m) \quad [3.71]$$

como por definición la cobertura de una distribución de contactos es el número de individuos que han resultado expuestos a, al menos, una de las m inserciones realizadas, sustituyendo $k = 1$ en [3.71] y teniendo en cuenta [3.70]:

$$\begin{aligned} R_m &= E_{x \geq 1}^m = POB \cdot P(x \geq 1/m, P_1, P_2 \cdots P_m) \\ &= POB \cdot [1 - P(0/m, P_1, P_2 \cdots P_m)] \\ &= POB \cdot [S_1 - S_2 + S_3 - \cdots \pm S_m] \end{aligned}$$

3.2.2.3 Modelo de Kwerel

Este modelo, desarrollado por Kwerel (1964) en su tesis doctoral y que nos llega en su formulación final en un artículo del *Journal of Advertising Research*

(Kwerel, 1969), mantiene las hipótesis de homogeneidad de las probabilidades individuales y de exposición del soporte, pero considera que la duplicación no es aleatoria, es decir, que la probabilidad de exposición a un soporte es diferente en función de si el individuo resultó o no expuesto a otro de los que conforman el plan de medios, es decir:

$$P(A_i/A_j) \neq P(A_i) \rightarrow P(A_i \cap A_j) \neq P(A_i) \cdot P(A_j) \quad \forall i, j \in M; \quad i \neq j$$

La única información necesaria para especificar el modelo de Kwerel, es el número de soportes m que componen el plan de medios, la suma de las audiencias de los soportes A , que a efectos de cálculo se aproxima por la audiencia promedio de los soportes \bar{A} :

$$\bar{A} = \frac{A}{m} = \frac{\sum_{i=1}^m A_i}{m}$$

y, finalmente, la suma de las audiencias duplicadas D , que son también promediadas por \bar{D} a esos mismos efectos:

$$\bar{D} = \frac{D}{\binom{m}{2}} = \frac{2D}{m(m-1)} = \frac{2 \sum_{i=1}^{m-1} \sum_{j=i+1}^m A_{ij}}{m(m-1)}$$

Kwerel (1969) demuestra que el alcance de un soporte viene acotado por dos límites, uno superior BU y otro inferior BL , que responden a las expresiones:

$$BU = A - \frac{2}{m}D$$

$$BL = (h+1)\bar{A} - \frac{h(h+1)}{2}\bar{D}$$

donde h es el más pequeño de los valores $m - l$ o k , siendo este último el entero más grande contenido en \bar{A}/\bar{D} . Como en nuestros cálculos no se desea acotar la cobertura, sino dar una estimación puntual, el autor deduce la expresión:

$$R_m = \frac{2BU \cdot BL}{BU + BL}$$

Para calcular la distribución de contactos, Kwerel (1969), llega a:

$$E_k^m = \binom{m}{k} \alpha^{k-2} (1 - \alpha)^{m-k} \bar{D}$$

salvo para $k = 1$, donde el número de individuos que han resultado expuestos a uno solo de los m soportes utilizados en el plan de medios se obtiene mediante:

$$E_1^m = A - \frac{2D}{m-1} [1 + (1 - \alpha) + (1 - \alpha)^2 + \dots + (1 - \alpha)^{m-2}]$$

siendo α la raíz positiva resultante de resolver iterativamente la ecuación:

$$A - R_m - \bar{D} \cdot \left[\sum_{i=1}^{m-1} (m-i)(1-\alpha)^{i-1} \right] = 0$$

3.2.2.4 *Fórmula de Sainsbury modificada*

Este método, desarrollado como ya se ha indicado por E. J. Sansbury para la *London Press Exchange* y que se presenta en el desarrollo ofrecido por Caffyn y Sagovsky (1963), se basa en la heterogeneidad de los soportes y la no aleatoriedad de las duplicaciones, aunque mantiene la homogeneidad de las probabilidades individuales de exposición. En síntesis:

- a. Probabilidades de exposición del soporte heterogéneas:

$$A_i \neq A_j \neq \dots \neq A_m \quad \forall i, j \in M$$

- b. Las probabilidades individuales de exposición (individuos a, b, etc.) son homogéneas:

$$P_{a,i} = P_{b,i} = \dots = P_i \quad \forall i \in M \quad \forall a, b \in POB$$

- c. Duplicación aleatoria: la exposición a un soporte no modifica la probabilidad de resultar expuesto a otro:

$$P(A_i/A_j) = P(A_i) \cdot P(A_j) \quad \forall i, j \in M; \quad i \neq j$$

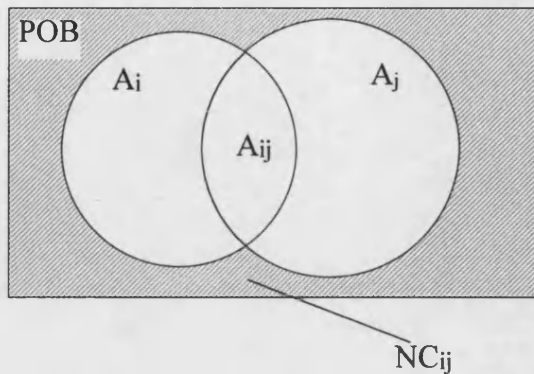
Caffyn y Sagovsky (1963) definen, en primer lugar, el concepto de la *no cobertura* (NC_i) del soporte i como el número de individuos de la población que no forman parte de la audiencia de i , es decir:

$$NC_i = \overline{A}_i = POB - A_i$$

y análogamente, la *no cobertura* de un par de soportes i y j (NC_{ij}), como:

$$NC_{ij} = \overline{A}_{ij} = POB - (A_i \cup A_j) = POB - A_i - A_j + A_{ij}$$

que es la superficie rayada de la figura 3.1:

Figura 3.1. Definición de *no cobertura*

Fuente: Elaboración propia

La probabilidad q_i de que un individuo pertenezca a la *no cobertura* del soporte i , puede definirse como:

$$q_i = \frac{\overline{A_i}}{POB} = \frac{POB - A_i}{POB} = 1 - P_i$$

y análogamente, para el par de soportes:

$$q_{ij} = \frac{\overline{A_{ij}}}{POB} = \frac{POB - A_i - A_j + A_{ij}}{POB} = 1 - P_i - P_j + P_{ij}$$

El modelo de Sainsbury modificado se basa en definir, dado que las probabilidades de exposición no son aleatorias, unas probabilidades de no exposición al soporte *ajustadas* q_i^d que sí tengan esta propiedad deseable, es decir:

$$q_i^d = \frac{q_{ij}}{q_j}$$

de tal forma que la *no cobertura ajustada* del soporte i sería:

$$A_i^d = POB \cdot q_i^d$$

por analogía, Caffyn y Sagovsky (1963), modifican la fórmula de Sainsbury introduciendo esas probabilidades y *no coberturas* ajustadas, de tal forma que la cobertura del plan de medios se calculará como:

$$R_m = POB - \frac{\prod_{i=1}^m A_i^d}{POB^{m-1}}$$

y, aunque estos autores no derivan una expresión para la distribución de contactos, Chandon (1985), utilizando la misma analogía con la fórmula de Sainsbury para el cálculo de esta variable, llega a:

$$E_k^m = \sum_{i=k}^m (-1)^{i-k} \binom{i}{k} S_i$$

3.2.2.5 Modelo de agregación secuencial con duplicación constante

Desarrollado por Goodhart y Erhenberg (1969), se basa en las hipótesis siguientes (cuya notación no reiteraremos):

- a. Las revistas son heterogéneas frente a las probabilidades de generar contactos.
- b. Las probabilidades individuales de exposición son, sin embargo, iguales.
- c. La duplicación no es aleatoria.

La base de este modelo, reside en el supuesto de que, aunque las duplicaciones no son aleatorias, las probabilidades tienen las mismas propiedades que este fenómeno otorga salvo en una constante multiplicativa α , esto es:

$$P_{ij} = \alpha \cdot P_i \cdot P_j \quad \forall i, j; \quad i \neq j; \quad i, j \in M \quad [3.72]$$

Aunque Chandon (1985) demuestra que, con los datos de audiencia de que dispone esta propiedad no se cumple empíricamente y que, por otra parte, a nivel teórico sólo sería cierta bajo ciertas hipótesis poco plausibles, llega a que la cobertura del plan de medios puede obtenerse con la expresión:

$$R_m = A - \alpha S_2 + \alpha^2 S_3 - \dots \pm \alpha^{m-1} S_m \quad [3.73]$$

siendo A la suma de las audiencias individuales de los soportes, y S_i la suma de las audiencias *i-plicadas*, que se obtienen a partir de [3.69] aplicando la propiedad [3.72]:

$$\begin{aligned} S_2 &= \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{A_i A_j}{POB} \\ S_3 &= \sum_{i=1}^{m-2} \sum_{j=i+1}^{m-1} \sum_{k=j+1}^m \frac{A_i A_j A_k}{POB^2} \\ &\dots \\ S_m &= \frac{A_i A_j A_k \dots A_m}{POB^{m-1}} \end{aligned}$$

y calculando la constante α como la media de las α entre todos los pares de revistas implicadas en el análisis, es decir:

$$\alpha = \frac{\sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{P_{ij}}{P_i P_j}}{\binom{m}{2}}$$

Aplicando las expresiones anteriores a la deducción que de la distribución de contactos obtuvimos en el modelo de Sainsbury, llegaríamos a:

$$E_k^m = \sum_{i=k}^m (-1)^{i-k} \binom{i}{k} \alpha^{i-1} S_i \quad [3.74]$$

3.2.2.6 Modelo de agregación secuencial con duplicación variable

Como ya se ha indicado, Chandon (1985) demuestra que la hipótesis en que se basa el modelo anterior acerca de que la ratio entre las probabilidades de resultar expuesto a dos soportes y el producto de las probabilidades de resultar expuesto a cada uno de ellos por separado es una constante, no soporta su contraste frente a los datos reales. Por este motivo, modifica el modelo de agregación secuencial eliminando esta restricción que, por otra parte, es teóricamente innecesaria por cuanto no conduce a una simplificación excesiva del proceso de cálculo.

Identificando los segundos términos (duplicaciones) de la expresión general para el cálculo de la cobertura:

$$R_m = A - \text{Duplic} + \text{Triplíc} - \dots \pm m - \text{plicaciones}$$

con la expresión particular que obtenía el modelo anterior [3.73], y sustituyendo las definiciones de S_2 , D y α se llega a la igualdad

$$\alpha \cdot S_2 = \text{Duplic}$$

$$\alpha \cdot \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{A_i A_j}{POB} = \sum_{i=1}^{m-1} \sum_{j=i+1}^m A_{ij}$$

despejando α , llegamos a su expresión general que, como se observa, no es constante:

$$\alpha = \frac{\sum_{i=1}^{m-1} \sum_{j=i+1}^m A_{ij}}{\sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{A_i A_j}{POB}}$$

Pues bien, basta sustituir el valor α así obtenido en [3.73] y [3.74] para obtener respectivamente la cobertura y la distribución de contactos del plan de medios, sin recurrir a la hipótesis de que la duplicación es constante.

3.2.2.7 Modelo de independencia condicional

Las hipótesis que sustentan a este modelo propuesto por Chandon (1985) son las siguientes.

- a. Las probabilidades de exposición individuales son homogéneas.
- b. Las probabilidades de exposición de los soportes son heterogéneas.
- c. La duplicación es un fenómeno no aleatorio.

La principal aportación de este modelo desarrollado por Chandon (1985), radica en que, aunque como se ha indicado, la duplicación es un fenómeno no aleatorio y, por tanto, las probabilidades de exposición a los soportes no son independientes, sí que lo son las probabilidades condicionadas:

$$P_{ijk} = P_{ilk} \cdot P_{jlk} \quad \forall i, j, k \in M; i \neq j \neq k \quad [3.75]$$

donde la lógica que subyace bajo este supuesto, es que existen grupos de soportes fuera del ámbito de interés de determinados individuos (i y j) y, el que se resulte expuesto a un soporte habitual k , no tiene porqué afectar a la probabilidad de resultar expuesto a esos soportes no habituales.

Por definición de probabilidad condicionada, tenemos que:

$$P_{i/k} = \frac{P_{ik}}{P_k} \quad [3.76]$$

de tal forma que [3.75] queda:

$$P_{ij/k} = \frac{P_{ik}}{P_k} \cdot \frac{P_{jk}}{P_k} \quad [3.77]$$

por analogía con [3.76], la probabilidad de resultar expuesto a dos soportes, condicionada a que se resultó expuesto a un tercero es:

$$P_{ijk} = \frac{P_{ijk}}{P_k} \quad [3.78]$$

y sustituyendo [3.77] en [3.78] y despejando llegamos a:

$$P_{ijk} = \frac{P_{ik} \cdot P_{jk}}{P_k} = \frac{P_{ij} \cdot P_{jk}}{P_j} = \frac{P_{ij} \cdot P_{ik}}{P_i} \quad [3.79]$$

Chandon (1985) considera que no es plausible que se vaya a cumplir la identidad [3.79] y recomienda aproximar la probabilidad de que un individuo resulte expuesto a los soportes i , j y k mediante la media aritmética de los tres términos de la igualdad. En todo caso, a partir de este momento el procedimiento de cálculo se reduce al descrito en el modelo de Sainsbury, sin más que sustituir las probabilidades calculadas mediante [3.79], y expresiones análogas para 4, 5 o m soportes en:

$$\begin{aligned}
 S_1 &= \sum_{i=1}^m P_i \\
 S_2 &= \sum_{i=1}^{m-1} \sum_{j=i+1}^m P_{ij} \\
 &\dots \\
 S_r &= \sum_{i=1}^{m-r+1} \sum_{j=i+1}^{m-r+2} \dots \sum_{r=m-1}^m P_{ij\dots r}
 \end{aligned}$$

y a continuación calcular la cobertura y la distribución de contactos según las expresiones que en su momento obtuvimos, esto es:

$$R_m = POB \cdot [S_1 - S_2 + S_3 - \dots \pm S_m]$$

$$E_k^m = POB \cdot \sum_{i=k}^m (-1)^{i-k} \binom{i}{k} S_i \quad 0 \leq k \leq m \quad S_0 = 1$$

3.2.2.8 Fórmula de Hofmans ampliada

Chandon (1985), plantea una extensión de la fórmula de Hofmans (1966), en base a las hipótesis de:

- a. Homogeneidad de individuos
- b. Heterogeneidad de soportes
- c. Duplicación no aleatoria.

Hofmans (1966), como ya vimos, llegó a que la cobertura puede estimarse como la suma de una serie geométrica infinita de base la audiencia bruta A y razón kD/A :

$$R_m = \frac{A}{1 + \frac{kD}{A}} \quad [3.80]$$

donde kD adopta la expresión:

$$kD = \sum_{i=1}^{m-1} \sum_{j=i+1}^m k_{ij} \cdot A_{ij}$$

donde toda la notación es conocida, salvo k_{ij} que se define en función de las audiencias como la *ratio*:

$$k_{ij} = \frac{A_i + A_j}{A_i + A_j - A_{ij}}$$

Esta aproximación, sin embargo, sólo sería válida cuando $kD < A$, puesto que en caso contrario la serie geométrica no converge. Dado que ningún argumento lógico garantiza que se vaya a cumplir tal condición, el método provoca fuertes errores en la estimación de la cobertura y distribuciones de contactos inconsistentes (v.g. frecuencias negativas). Chandon (1985), sin embargo, modifica la fórmula de Hofmans para salvar esta limitación. Ya hemos visto que la cobertura puede calcularse como:

$$R_m = S_1 - S_2 + S_3 - \dots \pm S_m$$

Denotaremos como $\Sigma R_{v/m}$ a la suma de las coberturas de todas las combinaciones de v soportes de las m posibles. Chandon (1985) pone, por inducción, $\Sigma R_{v/m}$ en función de S_i , llegando a:

$$\Sigma R_{v/m} = \sum_{i=1}^v (-1)^{i+1} \binom{m-i}{v-i} S_i$$

despejando, podemos obtener la suma de las v -plicaciones como:

$$S_v = \sum_{i=1}^{v-1} (-1)^{i+1} \binom{m-i}{v-i} S_i - \Sigma R_{v/m} \quad \text{si } v \text{ es par} \quad [3.81]$$

$$S_v = \sum_{i=1}^{v-1} (-1)^i \binom{m-i}{v-i} S_i + \Sigma R_{v/m} \quad \text{si } v \text{ es impar} \quad [3.82]$$

El procedimiento, a partir de aquí, es sencillo. Basta calcular $\Sigma R_{v/m}$ para todas las combinaciones mediante la expresión [3.80], donde S_i se estimará mediante la fórmula de Hofmans, y tras obtener las S_v mediante [3.81] y [3.82], sustituir en las expresiones generales de la cobertura y la distribución de contactos que obtuvimos para el modelo de Sainsbury, es decir:

$$R_m = POB \cdot [S_1 - S_2 + S_3 - \dots \pm S_m]$$

$$E_k^m = POB \cdot \sum_{v=k}^m (-1)^{v-k} \binom{v}{k} S_v \quad 0 \leq k \leq m \quad S_0 = 1$$

3.2.2.9 *Formula de duplicación de Metheringham*

Las hipótesis que sustentan a este modelo propuesto por Metheringham (1964) son las siguientes:

- a. Las individuos son heterogéneos. Esto implica que tienen su propia probabilidad individual de exposición a un soporte. Cuando este hecho se une a que la probabilidad es independiente de la del resto de sujetos y a si se resultó o no expuesto a la última inserción en un soporte dado, convierte a la exposición en un proceso bernouilli y hace que las probabilidades individuales se distribuyan según una función beta para cualquier segmento de población.

- b. Los soportes son homogéneos, esto es, sus audiencias son todas iguales. Dado que sabemos que esta hipótesis es totalmente irreal, Metheringham (1964) sustituye esa audiencia constante por la media de las audiencias de los soportes, al igual que hace con las duplicaciones.
- c. La duplicación no es un proceso aleatorio.

Al considerar los soportes homogéneos, se hace equivalente el problema de la duplicación entre diferentes soportes, con el de la acumulación de audiencias en un único soporte. Como, por hipótesis, las probabilidades individuales siguen una distribución beta, la acumulación de m inserciones es una beta-binomial. Cuando revisamos los modelos de acumulación de audiencias, vimos que la expresión general de la distribución de contactos bajo este supuesto era:

$$E_k^m = \binom{m}{k} \frac{\Gamma(a+b) \Gamma(a+k) \Gamma(m+b-k)}{\Gamma(a) \Gamma(b) \Gamma(a+b+m)} \quad [3.83]$$

donde a y b , son los dos parámetros que caracterizan a toda distribución beta-binomial y que es necesario estimar a partir de los datos disponibles. La audiencia de cualquiera de los soportes analizados (son homogéneos), puede escribirse como el número de individuos que han resultado expuestos a una inserción cuando sólo se ha realizado una en un único soporte, sustituyendo esta definición en [3.83] y teniendo en cuenta el procedimiento de cálculo de la audiencia constante de los soportes descrita en la hipótesis b :

$$\frac{1}{m} \sum_{i=1}^m A_i = \bar{A} = E_1^1 = \frac{a}{a+b} POB \quad [3.84]$$

Análogamente, la duplicación constante promedio de los soportes considerados puede obtenerse como el número de individuos que han resultado

expuestos a dos inserciones de un plan de medios compuesto por dos inserciones en dos soportes, teniendo en cuenta también el procedimiento descrito en la hipótesis b para el cálculo de la duplicación en el modelo de Metheringham (1964) y sustituyendo en [3.83], llegamos a:

$$\frac{1}{\binom{m}{2}} \sum_{i=1}^{m-1} \sum_{j=i+1}^m A_{ij} = \bar{D} = E_2^2 = \frac{a(a+1)}{(a+b+1)(a+b)} \cdot POB \quad [3.85]$$

si definimos las probabilidades de exposición p y duplicación p_{ij} como:

$$p = \frac{\bar{A}}{POB} \quad [3.86]$$

$$p_{ij} = \frac{\bar{D}}{POB}$$

llegamos a la estimación de los parámetros a y b que especifican el modelo de Metheringham sin más que resolver el sistema formado por [3.84] y [3.85] sustituyendo en ellas las expresiones recogidas en [3.86]:

$$\hat{a} = \frac{p^2 - p \cdot p_{ij}}{p_{ij} - p^2}$$

$$\hat{b} = \frac{\hat{a}(1-p)}{p}$$

y a partir de aquí, siguiendo las expresiones ya obtenidas anteriormente en este capítulo para el modelo beta binomial de acumulación de audiencias, obtenemos la distribución de contactos y la cobertura:

$$E_k^m = \binom{m}{k} \frac{\Gamma(a+b) \Gamma(a+k) \Gamma(m+b-k)}{\Gamma(a) \Gamma(b) \Gamma(a+b+m)} \cdot POB$$

$$R_m = POB \cdot E_0^m$$

3.3 MODELOS DE AUDIENCIA NETA ACUMULADA

Nos encontramos, finalmente, ante los modelos que abordan la planificación de medios de un modo más realista: los planes están formados por n inserciones en m soportes distintos. Estos modelos deben, como los anteriores, responder a una doble pregunta: ¿cuántos individuos han resultado expuestos a, al menos, una de las inserciones del plan? y, por otra parte ¿cuántos han resultado expuestos a una, dos, hasta $n \cdot m$ veces?.

Por mantener la simetría con el resto de apartados, se han dividido en modelos estocásticos y modelos *ad hoc*. Sin embargo debe precisarse que el creador de la tipología que se está siguiendo, Chandon (1985), agrupa a los modelos estrictamente *ad hoc*, según la definición reiteradamente expuesta, junto a los modelos de agregación secuencial bajo el epígrafe de modelos de aproximación. Pero dado que el único modelo de audiencia neta acumulada que es simultáneamente de agregación secuencial, el MODAL, es un modelo de propietario y, por lo tanto, no disponible para su evaluación, se hace innecesario mantener el epígrafe de modelos de acumulación cuando sólo reuniría modelos *ad hoc* enturbiando, por otra parte, la estructura de epígrafes que se está siguiendo. Debe señalarse, finalmente, que se ha invertido el orden de exposición habitual, procediéndose en primer lugar al desarrollo de los modelos estocásticos para pasar, a continuación, a los *ad hoc*. El motivo, como se verá, es que nos encontramos ante un modelo *ad hoc* que se basa precisamente en promediar dos modelos estocásticos que, lógicamente, deben ser previamente expuestos.

3.3.1 Modelos estocásticos

3.3.1.1 Modelo binomial

Este modelo, propuesto y utilizado en los trabajos de Lee y Burkart (1960), se basa en las hipótesis de:

- a. Población homogénea. Cada individuo de la población tiene la misma probabilidad de exposición a un soporte determinado.

$$p_{1i} = p_{2i} = \dots = p_i \quad \forall i \in M$$

- b. Soportes homogéneos. La probabilidad de exposición es la misma para cada uno de ellos.

$$p_i = p_j = \dots = p \quad \forall i, j \in M$$

- c. La duplicación de las audiencias es un suceso aleatorio.

$$p_{ij} = p_i \cdot p_j = p^2 \quad \forall i, j \in M \quad i \neq j$$

- d. La acumulación de las audiencias es, también, un suceso aleatorio.

$$p_{ii} = p_i \cdot p_i = p^2 \quad \forall i \in M$$

- e. Las probabilidades de exposición son, finalmente, estacionarias respecto al tiempo.

Las hipótesis aquí expuestas, llevan a que la probabilidad de exposición a distintas inserciones en diferentes soportes, sea equivalente a la de distintas inserciones en un soporte “promedio” cuya audiencia sea la media simple de todos ellos:

$$p = \frac{\bar{A}}{POB}$$

donde:

$$\bar{A} = \frac{1}{m} \sum_{i=1}^m A_i$$

siendo m el número de soportes implicados en el plan de medios y A_i la audiencia promedio de cada uno de ellos. Cuando el plan consta de n inserciones en cada soporte, la probabilidad de resultar expuesto a K de ellas, cualquiera que sea su reparto entre los m soportes vendrá dada, en este modelo, por la distribución binomial:

$$P(K / p, N) = \binom{N}{K} p^K (1-p)^{N-K}$$

siendo $N = n \cdot m$, de tal forma que, la distribución de contactos se obtendrá como:

$$E_K^N = P(K / p, N) \cdot POB = \binom{N}{K} p^K (1-p)^{N-K} \cdot POB$$

y la cobertura del plan de medios, que venimos reiteradamente definiendo como los individuos que han resultado expuestos a, al menos, una inserción entre las N realizadas, se calcula de manera inmediata como sigue:

$$R_N = p(K \geq 1 / p, N) \cdot POB = 1 - (1-p)^N \cdot POB$$

3.3.1.2 Modelo binomial multivariable con independencia de soportes

Este modelo, propuesto por Chandon (1985), se apoya esencialmente en el anterior, con la salvedad de relajar la hipótesis de homogeneidad de los soportes. En síntesis:

- a. La población es homogénea
- b. Los soportes son heterogéneos.

$$p_i \neq p_j \quad \forall i, j \in M$$
- c. La duplicación y acumulación de audiencias son sucesos aleatorios y,
- d. Las probabilidades de exposición son estacionarias respecto al tiempo.

Bajo estas hipótesis, la exposición al soporte es una variable bernouilli con probabilidades distintas en cada uno de ellos:

$$p_i = \frac{A_i}{POB}$$

Se procederá a continuación a deducir las expresiones de la cobertura y la distribución de contactos para el caso de dos soportes, con probabilidades respectivas de exposición p_1 y p_2 , en los que se realizan n_1 y n_2 inserciones para, posteriormente, generalizar al caso de m soportes. La matriz de exposición, en este caso de orden $n_1 \cdot n_2$, contendrá en su elemento ij el número de individuos que han resultado expuestos

i veces al soporte 1 y j veces al 2. Llamemos $E(X,Y)$ a esa matriz y $e(x,y)$ a cada uno de sus elementos.

Por la hipótesis c , duplicación y acumulación son sucesos aleatorios, la distribución de probabilidad de $E(X,Y)$ vendrá dada, por tanto, por:

$$P(x = k_1, y = k_2) = P(x = k_1) \cdot P(y = k_2)$$

y como X e Y siguen una distribución binomial al proceder de una variable bernouilli¹⁴, tenemos que:

$$P(x = k_1, y = k_2) = \binom{n_1}{k_1} \binom{n_2}{k_2} p_1^{k_1} p_2^{k_2} q_1^{n_1-k_1} q_2^{n_2-k_2} \quad [3.86]$$

donde $0 \leq k_1 \leq n_1$ $0 \leq k_2 \leq n_2$.

A partir de la expresión anterior, es fácil obtener la probabilidad de resultar expuesto a al menos una inserción:

$$P(x \geq 1, y \geq 1) = 1 - P(x = 0, y = 0) = 1 - q_1^{n_1} q_2^{n_2} \quad [3.87]$$

probabilidad que basta multiplicar por la población objetivo del plan de medios para obtener la cobertura. La generalización a n_1, n_2, \dots, n_m inserciones en m soportes diferentes se deriva fácilmente para la cobertura a partir de [3.87]:

$$R_N = \left(1 - \prod_{i=1}^m q_i^{n_i}\right) \cdot POB$$

¹⁴ En el apartado 3.1.2.1 de este mismo capítulo, al desarrollar el modelo binomial de acumulación de audiencias, se dan las referencias para profundizar en cómo una distribución de Bernouilli sirve de base a la distribución binomial.

Para obtener la distribución de contactos es necesario, previamente, generalizar la expresión [3.86] para el caso de m soportes, lo que es inmediato:

$$P(x_1 = k_1, x_2 = k_2, \dots, x_m = k_m) = \prod_{i=1}^m \binom{n_i}{k_i} p_i^{k_i} q_i^{n_i - k_i} \quad [3.88]$$

a continuación, debe calcularse la probabilidad para todas las combinaciones coherentes del vector $\{x_1, x_2, \dots, x_m\}$, agrupando aquellas cuyos $\{k_1, k_2, \dots, k_m\}$ sumen lo mismo. El ejemplo siguiente aclara lo señalado:

$$\begin{aligned} E_2^{n_1+n_2+\dots+n_m} = & P(x_1 = 1, x_2 = 1, x_3 = 0, \dots, x_m = 0) \cdot \text{POB} \\ & + P(x_1 = 0, x_2 = 1, x_3 = 1, \dots, x_m = 0) \cdot \text{POB} \\ & + P(x_1 = 0, x_2 = 0, \dots, x_{m-1} = 1, x_m = 1) \cdot \text{POB} \\ & + P(x_1 = 2, x_2 = 0, x_3 = 0, \dots, x_m = 0) \cdot \text{POB} \\ & + P(x_1 = 0, x_2 = 2, x_3 = 0, \dots, x_m = 0) \cdot \text{POB} \\ & + P(x_1 = 0, x_2 = 0, x_3 = 0, \dots, x_m = 2) \cdot \text{POB} \end{aligned}$$

3.3.1.3 Modelo multinomial

Las hipótesis en las que se basa este modelo desarrollado por Chandon (1985), son las siguientes:

- a. La población es homogénea.
- b. Los soportes son heterogéneos
- c. La duplicación es un suceso no aleatorio, mientras que la acumulación de las audiencias continúa siendo aleatoria.

- d. Las probabilidades de exposición son estacionarias en el tiempo.

El desarrollo requiere de un ligero cambio en la notación que viene recogido en el siguiente cuadro.

Cuadro 3.2. Notación en el modelo multinomial

Notación	Suceso	Prob	Nº exposic
A_0	No exposición a ninguno de los m soportes	p_0	0
A_1	Exposición a uno sólo de los m soportes	p_1	1
A_2	Exposición a cualquier par de soportes	p_2	2
...
A_m	Exposición a todos los soportes	p_m	m

Fuente: Chandon (1985) p.463

Cuando bajo las hipótesis señaladas, se realizan n inserciones en los m soportes habituales, la probabilidad de resultar expuesto K veces viene dada, según Chandon (1985) por la siguiente distribución multinomial:

$$P(K / n, p_1 \cdots p_m) = \sum \frac{n!}{k_0! k_1! \cdots k_m!} p_0^{k_0} p_1^{k_1} \cdots p_m^{k_m} \quad [3.89]$$

donde $k = 0, 1, 2 \dots nm$ y $p_0 = 1 - p_1 - p_2 - \dots - p_m$, extendiéndose el sumatorio a todos los enteros no negativos k_0, k_1, \dots, k_m que cumplan que:

$$\sum_{i=0}^m k_i = n$$

$$\sum_{i=1}^m i \cdot k_i = K$$

Como las probabilidades $p_0, p_1, p_2, \dots, p_m$, que son las probabilidades de resultar expuesto *de manera exclusiva*¹⁵ a exactamente 0, 1, 2, ... m soportes de los m que componen el plan de medios, no son conocidas, se ha de recurrir para su cómputo, al procedimiento que se describió al desarrollar el modelo de Sainsbury¹⁶ :

$$P_i = S_i - \binom{i+1}{i} S_{i+1} + \binom{i+2}{i} S_{i+2} - \dots \pm \binom{m}{i} S_m$$

donde $S_0 = 1$ y S_i es la suma de las probabilidades no exclusivas de resultar expuesto a i soportes. Como los datos de los que se dispone habitualmente son las audiencias de cada soporte y las duplicaciones entre todos los pares, podremos calcular fácilmente S_1 y S_2 . El problema será, pues, derivar S_i para $i > 2$ a partir de S_1 y S_2 . Para ello Chandon (1985) propone seguir el procedimiento ya descrito al exponer la fórmula de Hofmans ampliada¹⁷, esto es, calcular la cobertura con la fórmula de Hofmans¹⁸ para una inserción en cada soporte:

$$R_m = \frac{A^2}{A+kD}$$

siendo A la suma de las audiencias de los soportes y

$$kD = \sum_{i=1}^{m-1} \sum_{j=i+1}^m k_{ij} A_i$$

mientras que k_{ij} responde a la expresión ya conocida:

$$k_{ij} = \frac{A_i + A_j}{A_i + A_j - A_{ij}}$$

¹⁵ Las probabilidades *no exclusivas* hacen referencia a las resultantes de dividir la audiencia de un soporte por la población. Son no exclusivas en cuanto que la probabilidad de estar expuesto a un soporte dependerá también de los valores de las duplicaciones, triplicaciones, etc. Las probabilidades *exclusivas* de resultar expuesto a i soportes cualquiera de un plan de medios sí que tienen en cuenta estos fenómenos.

¹⁶ Véase el apartado 3.2.2.2 de este mismo capítulo.

¹⁷ Véase apartado 3.2.2.8

¹⁸ Véase apartado 3.1.1.4

Así obtenido R_m se obtienen los sucesivos S_i por el procedimiento recurrente ya detallado cuando se desarrolló el modelo de la fórmula de Hofmans ampliada¹⁹, esto es:

$$S_v = \sum_{i=1}^{v-1} (-1)^{i+1} \binom{m-i}{v-i} S_i - \Sigma R_{v/m} \quad \text{si } v \text{ es par}$$

$$S_v = \sum_{i=1}^{v-1} (-1)^i \binom{m-i}{v-i} S_i + \Sigma R_{v/m} \quad \text{si } v \text{ es impar}$$

siendo

$$\Sigma R_{v/m} = \sum_{i=1}^v (-1)^{i+1} \binom{m-i}{v-i} S_i$$

Una vez calculados los sucesivos S_i , y obtenida la distribución de contactos E_K^N para el caso de una inserción en cada soporte por el mencionado procedimiento de la fórmula ampliada de Hofmans, las p_k que sustituidas en [3.89] permiten obtener la distribución final de contactos, se calculan del siguiente modo:

$$p_k = \frac{E_K^m}{POB} = \frac{\sum_{v=K}^m (-1)^{v-K} \binom{v}{K} S_v}{POB}$$

Estas probabilidades deben ser agregadas para aquellos k_i que sumen el mismo número de exposiciones para constituir así, como se ha indicado, la distribución de contactos.

¹⁹ Véase apartado 3.2.2.8. En la expresión siguiente se ha utilizado S_i para hacer referencia a los S_i desconocidos que quedan por calcular, manteniendo S_i para los ya obtenidos. No se olvide que nos encontramos ante un procedimiento recurrente.

3.3.1.4 Modelo de Markov-Bernouilli

Este modelo, creado por Chandon (1985) viene sustentado en las siguientes hipótesis:

- a. La población es homogénea.
- b. Los soportes son homogéneos.
- c. La duplicación y la acumulación no son aleatorios y, finalmente,
- d. Las probabilidades de exposición ya no son estacionarias en el tiempo, sino condicionadas a la exposición-no exposición a la inserción anterior.

Consideremos la matriz de probabilidades de un suceso bernouilli independiente. Sea $X_n = 0$ la no exposición a la inserción n -ésima y $X_n = 1$ la exposición a esa misma inserción. Por definición de independencia, resultar expuesto a la inserción n -ésima no depende de haberlo sido a las $n - 1$ anteriores, luego:

$$P(X_n = 1 / X_{n-1} = 1) = P(X_n = 1 / X_{n-1} = 0) = p$$

$$P(X_n = 0 / X_{n-1} = 0) = P(X_n = 0 / X_{n-1} = 1) = q$$

calculándose la probabilidad de exposición p mediante la media aritmética de las audiencias de los soportes:

$$p = \frac{\bar{A}}{POB} = \frac{\frac{1}{m} \sum_{i=1}^m A_i}{POB}$$

quedando, de esta manera, la matriz de transición como:

	$X_n = 0$	$X_n = 1$
$X_{n-1} = 0$	q	p
$X_{n-1} = 1$	q	p

Sin embargo, como se señala en la hipótesis d las probabilidades no son estacionarias y, por lo tanto, la matriz anterior debe modificarse teniendo en cuenta las consecuencias de este hecho: se ha de hacer que la probabilidad de exposición a la inserción siguiente sea superior a p cuando la exposición anterior ha ocurrido, y menor que p en caso contrario:

$$P(X_n = 1 / X_{n-1} = 1) = \lambda \quad 1 > \lambda > p$$

λ se convierte así en un índice de acumulación cuando n y $n - 1$ se dan en un mismo soporte, y de duplicación en soportes diferentes. Como lo que se pretende con el presente modelo, es construir un soporte promedio del conjunto de los m existentes, λ será la ratio entre el número promedio de dobles exposiciones (acumulaciones o duplicaciones) frente al número de exposiciones. En términos de probabilidad:

$$\lambda = P(X_n = 1 / X_{n-1} = 1) = \frac{P(X_n=1, X_{n-1}=1)}{P(X_n=1)} \quad [3.90]$$

Chandon (1985) profundiza, a continuación, sobre la expresión anterior. Por un lado, el numerador es la probabilidad de ocurrencia de una doble exposición, con lo cual puede ponerse según lo dicho como:

$$P(X_n = 1, X_{n-1} = 1) = \frac{\sum \text{duplicaciones} + \sum \text{acumulaciones}}{\text{Posibilidades de doble exposición}} \quad [3.91]$$

Las posibilidades de doble exposición dependen, lógicamente, del número de inserciones y de soportes de la siguiente forma:

$$\binom{nm}{2} = \frac{nm(nm-1)}{2}$$

Teniendo esto en cuenta, y llamando A_{ij} a la audiencia duplicada entre dos soportes i y j cualesquiera y A_i^{kl} a la audiencia común entre dos números k y l del mismo soporte i , entonces la probabilidad de doble exposición, recogida en la expresión [3.91] queda como:

$$P(X_n = 1 / X_{n-1} = 1) = \frac{\left[\sum_{i=1}^m \sum_{k=1}^{n-1} \sum_{l=k+1}^n A_i^{k,l} + n^2 \sum_{i=1}^{m-1} \sum_{j=i+1}^m A_{ij} \right]}{POB} / \binom{nm}{2}$$

Sustituyendo la anterior expresión en [3.90], y teniendo en cuenta que $P(X_n = 1) = p$, λ puede ponerse como:

$$\lambda = \frac{\left[\sum_{i=1}^m \sum_{k=1}^{n-1} \sum_{l=k+1}^n A_i^{k,l} + n^2 \sum_{i=1}^{m-1} \sum_{j=i+1}^m A_{ij} \right]}{n(nm-1) \sum_{i=1}^m A_i}$$

Con λ ya calculada, resta obtener los otros tres elementos de la nueva matriz de transición que recoge las modificaciones consecuencia de las hipótesis del modelo. Por un lado, como $P(X_n = 1 / X_{n-1} = 1) + P(X_n = 0 / X_{n-1} = 1) = 1$, el elemento (2,1) de la nueva matriz queda: $P(X_n = 0 / X_{n-1} = 1) = 1 - \lambda$. Por otro lado, la hipótesis de audiencia constante entre inserciones en un mismo soporte, implica simultáneamente que:

$$\begin{aligned} P(X_n = 1 / X_{n-1} = 1) \cdot P(X_{n-1} = 1) + P(X_n = 1 / X_{n-1} = 0) \cdot P(X_{n-1} = 1) = \\ = P(X_n = 1) = p \end{aligned}$$

y que

$$P(X_n = 0 / X_{n-1} = 1) \cdot P(X_{n-1} = 1) + P(X_n = 0 / X_{n-1} = 0) \cdot P(X_{n-1} = 1) =$$

$$= P(X_n = 0) = q$$

Resolviendo el sistema formado por las dos expresiones anteriores, la matriz de transición queda definitivamente como sigue:

	$X_n = 0$	$X_n = 1$
$X_{n-1} = 0$	$(1-2p + \lambda p) / q$	$(1 - \lambda) p / q$
$X_{n-1} = 1$	$1 - \lambda$	λ

Chandon (1985) demuestra que, a partir de la anterior matriz, la probabilidad de que un individuo resulte expuesto a un vector de exposiciones $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$ donde cada x_i toma el valor 0 si no se ha resultado expuesto a la inserción i -ésima en el soporte promedio y 1 en el caso contrario, se obtiene mediante la expresión:

$$P(X_1 = x_1, \dots, X_n = x_n) = \lambda^r (1 - \lambda)^{2(s-r)-t} (1 - 2\lambda + \lambda p)^{n-1-2s+t+r} p^{s-r} q^{-(n-2-s+t)} \quad [3.92]$$

con $x_i = 0, 1; i = 1, 2, \dots, n$ y con:

$$r = \sum_{i=2}^n x_{i-1}x_i \quad ; \quad s = \sum_{i=1}^n x_i \quad ; \quad t = x_1 + x_n$$

Para el caso extremo de no exposición a ningún soporte, la expresión [3.92] queda ($r = 0; s = 0; t = 0$):

$$P(X_1 = 0, \dots, X_n = 0) = (1 - 2p + \lambda p)^{n-1} q^{2-n}$$

y como la cobertura del plan de medios se obtiene precisamente, multiplicando la población objetivo por uno menos la probabilidad anterior, ésta queda:

$$R_N = [1 - (1 - 2p + \lambda p)^{N-1} q^{2-N}] \cdot POB$$

Resta en este modelo, determinar la distribución de contactos. Para ello, Chandon (1985), desarrolla un procedimiento para, por recurrencia, obtener las probabilidades de ocurrencia de cada cadena de posibles contactos a las n inserciones. Para facilitar el seguimiento del proceso, consideremos el siguiente cuadro.

Cuadro 3.3: generación de las probabilidades en el modelo de Markov-Bernoulli

Inserción	1	2	3	4	5	6	7	8
1	0 P(1,1)	1 P(1,2)						
2	00 P(2,1)	01 P(2,2)	10 P(2,3)	11 P(2,4)				
3	000 P(3,1)	001 P(3,2)	010 P(3,3)	011 P(3,4)	100 P(3,5)	101 P(3,6)	110 P(3,7)	111 P(3,8)
.
.
n

Fuente: Chandon (1985) p. 477 y elaboración propia

En el cuadro anterior, aparecen en negrita los sucesos que pueden ocurrir cuando se realizan 1, 2 o n inserciones en el soporte promedio. Así, leyendo la segunda fila, puede ocurrir que, cuando el plan de medios está formado por dos inserciones, $POB \cdot P(2,1)$ individuos no vean ninguna de las dos, $POB \cdot P(2,2)$ individuos no vean la primera y sí la segunda, $POB \cdot P(2,3)$ individuos vean la primera y no la segunda, y $POB \cdot P(2,4)$ individuos vean las dos. $P(i, j)$ es, pues, la probabilidad de ocurrencia del suceso que aparece en la fila i y la columna j de la matriz señalada.

Las primeras probabilidades proceden de parámetros ya conocidos o de operaciones entre ellos. Así, si llamamos respectivamente a , b , c y d a los elementos (1,1), (1,2), (2,1) y (2,2) de la matriz de transición ya especificada con anterioridad y p se calcula como se indicó al principio de este apartado, el cuadro 3.4 nos muestra en **negrita** cómo calcular las primeras probabilidades (hasta tres inserciones):

Cuadro 3.4: Cálculo de las primeras probabilidades en el modelo de Markov-Bernouilli

Inserción	1	2	3	4	5	6	7	8
1	q P(1,1)	p P(1,2)						
2	qa P(2,1)	qb P(2,2)	pc P(2,3)	pd P(2,4)				
3	qaa P(3,1)	qab P(3,2)	qbc P(3,3)	qbd P(3,4)	pca P(3,5)	pcb P(3,6)	pdc P(3,7)	pdd P(3,8)

Fuente: Chandon (1985) p. 477 y elaboración propia

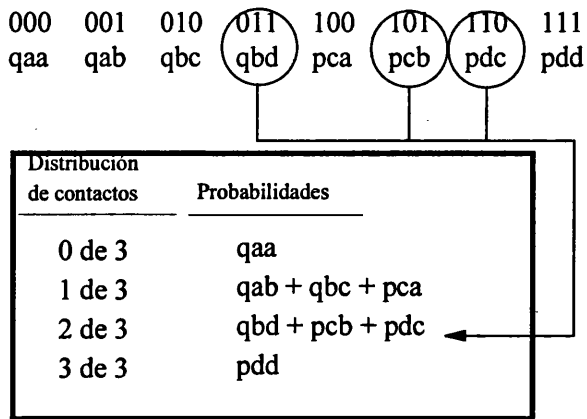
Para planes formados por más de tres inserciones, es necesario utilizar el siguiente procedimiento de recurrencia para el cálculo. Si llamamos $P(n, j)$ a la probabilidad correspondiente a la celda situada en la fila n -ésima y columna j -ésima de la ampliación de la anterior matriz, tendremos:

$$\begin{aligned}
 P(n, 2j - 1) &= P(n - 1, j) \cdot a && \text{si } j \text{ es impar} \\
 P(n, 2j) &= P(n - 1, j) \cdot b
 \end{aligned}$$

$$\begin{aligned}
 P(n, 2j - 1) &= P(n - 1, j) \cdot c && \text{si } j \text{ es par} \\
 P(n, 2j) &= P(n - 1, j) \cdot d
 \end{aligned}$$

Con las probabilidades así calculadas, bastará sumar aquellas cadenas de sucesos que sumen el mismo número de exposiciones, como se observa en la siguiente figura como ejemplo para un plan de tres inserciones:

Figura 3.3: Ejemplo de cálculo de la distribución de contactos



3.3.1.5 Modelo de Metheringham

Este modelo fue desarrollado por Metheringham (1964), quien los sustentó en las siguientes hipótesis:

- La población es heterogénea. Cada individuo tiene su propia probabilidad personal de exposición que, sin embargo, se distribuye como una beta para el conjunto.
- Los soportes son homogéneos. Ello no quiere decir que cada individuo tenga la misma probabilidad de exposición respecto a cada soporte, sino que estas probabilidades individuales se distribuyen de tal forma que todos los soportes acaban con la misma distribución beta de probabilidades de exposición.
- La acumulación y duplicación de las audiencias se promedian entre los soportes para conformar un soporte “tipo”.

- d. Las probabilidades de exposición son estacionarias en el tiempo.

El desarrollo del modelo es paralelo al que se expuso con este mismo nombre entre los de duplicación de audiencias²⁰. Como se recordará, la expresión que permitía obtener la distribución de contactos, era la siguiente:

$$E_K^N = \binom{N}{K} \frac{\Gamma(A+B) \Gamma(A+K) \Gamma(N+B-K)}{\Gamma(A) \Gamma(B) \Gamma(A+B+N)} \quad [3.93]$$

donde A y B son los dos parámetros de la distribución beta que deben estimarse. Por su parte, la cobertura del plan de medios, se obtenía como:

$$R_N = 1 - E_0^N \quad [3.94]$$

Pues bien, en la modificación del modelo para el caso de n inserciones en m soportes distintos, para estimar los parámetros A y B , se calcula en primer lugar la audiencia promedio de los soportes implicados en el plan:

$$\bar{R}_1 = \frac{\sum_{i=1}^m A_i n_i}{\sum_{i=1}^m n_i}$$

donde A_i es la audiencia de cada soporte y n_i el número de inserciones en cada soporte i que constituye el plan de medios. El paso siguiente consiste en obtener la audiencia de dos inserciones en ese soporte “tipo”, promediando las duplicaciones y las acumulaciones. Como ya se ha visto en modelos anteriores, si hay m soportes y el vector de inserciones es el siguiente: $S = \{ n_1, n_2, \dots, n_i, \dots, n_{nm} \}$, existirán:

²⁰ Puede verse el apartado 3.2.2.9

$\sum_{i=1}^m \binom{n_i}{2}$ acumulaciones, con $\binom{n_i}{2} = 0$ si $n_i < 2$ y $\sum_{i=1}^m \sum_{j=i+1}^{m-1} n_i n_j$ duplicaciones

lo que hace un total de dobles exposiciones igual a:

$$\binom{N}{2} \text{ con } N = \sum_{i=1}^m n_i$$

De esta forma, la cobertura promedio de dos inserciones en el soporte “tipo”, viene dada por:

$$\overline{R_2} = \frac{\sum_{i=1}^m \binom{n_i}{2} R_{2i} + \sum_{i=1}^m \sum_{j=i+1}^{m-1} n_i n_j R_{1ij}}{\binom{N}{2}}$$

siendo R_{2i} la audiencia acumulada tras dos inserciones en el soporte i y R_{1ij} el alcance neto del par de soportes i y j cuando se realiza una inserción en cada uno de ellos. Con los datos así obtenidos, los parámetros que especifican el modelo viene dados por:

$$\hat{A} = \frac{\overline{R_1}(\overline{R_2} - \overline{R_1})}{POB(2\overline{R_1} - \overline{R_2}) - \overline{R_1}^2}$$

$$\hat{B} = \frac{\hat{A}(POB - \overline{R_1})}{\overline{R_1}}$$

Basta sustituir estos parámetros en [3.93] y ésta expresión para $K = 0$ en [3.94] para obtener, respectivamente, la distribución de contactos y la cobertura del plan de medios.

3.3.1.6 *Modelo Dirichlet multinomial*

Este modelo fue introducido por primera vez en 1976 por Chandon (1985), aunque ha sufrido diversas mejoras consecuencia, fundamentalmente, debido a que su buena capacidad predictiva (Chandon, 1985; Leckenby, 1981; Leckenby y Kishi 1982b; Rust y Leone 1982) lo ha convertido en uno de los más utilizados. Nosotros vamos a exponerlo en la versión ofrecida por Leckenby y Kishi(1984).

Las hipótesis en que se basa este modelo, como ya se indicó en el capítulo anterior, son las siguientes:

- a. La población es heterogénea.
- b. Los soportes son heterogéneos, viniendo cada uno de ellos caracterizado por su propia distribución beta de probabilidades.
- c. La duplicación no es aleatoria, mientras que la acumulación sigue una distribución beta binomial para cada soporte y, finalmente,
- d. Las probabilidades de exposición son estacionarias entre inserciones sucesivas.

Consideremos un vector de exposiciones $\mathbf{X} = (x_0, x_1, \dots, x_m)$ donde el elemento x_i del vector toma el valor j si se ha resultado expuesto a j inserciones en cualquier combinación de i de los m soportes que conforman el plan de medios. Pues bien la probabilidad de ocurrencia de ese suceso cuando se realizan n inserciones en los mencionados m soportes, viene dada por la composición de las distribuciones multinomial y Dirichlet multivariable:

$$f(\mathbf{X} | \mathbf{A}, n) = \frac{n!}{\prod_{i=0}^m x_i!} \frac{\left[\Gamma\left(\sum_{i=0}^m A_i\right) \right] \left[\prod_{i=0}^m \Gamma(x_i + A_i) \right]}{\left[\prod_{i=0}^m \Gamma(A_i) \right] \left[\Gamma\left(n + \sum_{i=0}^m A_i\right) \right]} \quad [3.95]$$

donde $\mathbf{A} = (A_0, A_1, \dots, A_m)$ es el vector de parámetros de la distribución Dirichlet que sirve de base al modelo, y que deben ser estimados para que éste quede especificado. Esta especificación se basará en la propiedad que cumplen de que:

$$A_i = p_i \cdot S \quad [3.96]$$

donde p_i representa la probabilidad individual de resultar expuesto a ninguno, cualquiera, cualquier dos ... de los m soportes cuando se realiza una única inserción en cada uno de ellos, cumpliéndose por tanto, al ser sucesos excluyentes que:

$$\sum_{i=0}^m p_i = 1 \quad 0 \leq p_i \leq 1$$

y donde, por otro lado, S es la suma de los parámetros i :

$$S = \sum_{i=0}^m A_i \quad [3.97]$$

Si se pudiera estimar el vector \mathbf{P} y el valor de S , a partir de [3.97] se obtendrían los parámetros A_i que especifican el modelo. Para estimar el vector \mathbf{P} , se proponen varios métodos: la distribución geométrica de Hofmans, la distribución beta binomial y el teorema de Waring. Dado que trabajos posteriores (Leckenby y Kim, 1994b) al que está sirviendo de base a esta exposición (Leckenby y Kishi, 1984) demuestran que es la primera de ellas la que aporta mejores resultados, será ésta la utilizada. Se basa en estimar cada elemento p_i de \mathbf{P} de este modo:

$$\hat{p}_i = E_i^m \quad i = 1, 2, \dots, n$$

siendo E_i^m la distribución de contactos de un plan de medios formado por una inserción en m soportes distintos calculada por el método de Hofmans ya descrito en esta tesis en el apartado 3.2.2.8.

Con el vector \mathbf{P} así estimado, es necesario aproximar de algún modo \mathbf{S} . Leckenby y Kim (1994b) aportan tres alternativas. Todas ellas pasan por calcular los parámetros α y β de la distribución beta para cada uno de los soportes por separado según las expresiones que ya dedujimos al desarrollar el modelo beta binomial de acumulación:

$$\hat{\alpha} = \frac{R_1(R_2 - R_1)}{2R_1 - R_1^2 - R_2}$$

$$\hat{\beta} = \frac{\hat{\alpha}(1 - R_1)}{R_1}$$

a continuación \mathbf{S} puede estimarse mediante una media aritmética de los parámetros α y β obtenidos para cada distribución (Chandon, 1985):

$$\hat{S} = \frac{\sum_{i=1}^m R_{1i} (\hat{\alpha}_i + \hat{\beta}_i)}{\sum_{i=1}^m R_{1i}}$$

expresión que, en el caso de no realizarse el mismo número de inserciones en cada soporte debería adaptarse (Leckenby y Kim, 1994b):

$$\hat{S} = \frac{\sum_{i=1}^m n_i R_{1i} (\hat{\alpha}_i + \hat{\beta}_i)}{\sum_{i=1}^m n_i R_{1i}}$$

Danaher (1988a; 1988b) considera, por el contrario, que es recomendable recurrir a una media geométrica para esta estimación:

$$\hat{S} = \left[\prod_{i=1}^m (a_i + \beta_i) \right]^{1/m}$$

siendo, en todos los casos n_i el número de inserciones en el soporte i y R_{ji} la audiencia del soporte i . El resto de la notación es conocida.

Basta, a continuación, obtener mediante la expresión [3.96] el vector A. Sustituyendo esa estimación en la expresión [3.95] para todas las posibles combinaciones coherentes del vector X y agregando aquellos casos en que la suma de los elementos de ese vector obtenga un mismo valor, se obtendrá la distribución de frecuencias. El ejemplo siguiente, para 3 soportes y 2 inserciones, aclara el procedimiento.

Cuadro 3.5. Ejemplo de cálculo de la distribución de contactos por el método Dirichlet multinomial

x_0	x_1	x_2	x_3	$\sum_{i=1}^n x_i$	$f(X A, n)$	E_i^2
2	0	0	0	0	f_0	$E_0^2 = \Sigma f_0.$
1	1	0	0	1	f_{11}	$E_1^2 = \Sigma f_1.$
1	0	1	0	1	f_{12}	
1	0	0	1	1	f_{13}	
0	1	1	0	2	f_{21}	$E_2^2 = \Sigma f_2.$
0	0	1	1	2	f_{22}	
0	1	0	1	2	f_{23}	
0	2	0	0	2	f_{24}	
0	0	2	0	2	f_{25}	
0	0	0	2	2	f_{26}	

Fuente: elaboración propia

3.3.1.7 Modelo de la distribución Poisson-Binomial

Como ya se ha expuesto reiteradamente a lo largo de estas páginas, la tendencia en el desarrollo de los modelos pasa por abandonar distribuciones sencillas pero muy restrictivas en cuanto a las hipótesis que las sustentan, como la distribución binomial, y sustituirlas por otras mucho más exigentes en cuanto a los datos que se requieren para estimarlas, como la distribución beta binomial en la suposición de que hipótesis más realistas deberán traducirse en estimaciones más ajustadas. Algunos trabajos, por contra, han optado por intentar mejorar la distribución binomial para aprovechar su principal ventaja: no requiere de datos de duplicaciones en su especificación y lograr que sea capaz de ofrecer buenas predicciones de la cobertura y la distribución de frecuencias sin, por ello, complicar operativamente el proceso de cálculo.

En esta línea se encuentra el desarrollo del modelo basado en la distribución Poisson Binomial (PBD) por parte de Leckenby y Ju (1989). En el mismo se compone la distribución de Poisson²¹ cuya función de densidad viene dada por la expresión [3.98] con la distribución binomial ya conocida [3.99]:

$$Pr(X = K) = e^{-\lambda} \frac{\lambda^K}{K!} \quad [3.98]$$

$$Pr(X = K) = \binom{N}{K} p^K (1-p)^{N-K} \quad [3.99]$$

donde todos los elementos de la notación empleada ya son conocidos, a excepción de $\lambda > 0$ que es el parámetro de la distribución de Poisson. La función de densidad de la distribución compuesta Poisson Binomial, resulta:

²¹ Puede consultarse Cuadras, Echeverría, Mateo y Sánchez (1991; pp. 261-276) para una descripción detallada de la distribución de Poisson.

$$Pr(X=K) = \phi^K \frac{1}{K!} \sum_{j=1}^N e^{-j\phi} j^K \binom{N}{j} p^j (1-p)^{N-j} \quad [3.100]$$

donde j sigue una distribución binomial de parámetros N y p y ϕ es el parámetro que procede de la distribución de Poisson.

Para la estimación de estos parámetros Leckenby y Ju (1989), proponen el siguiente procedimiento:

- a. N se aproximará por la suma del total de las inserciones que componen el plan de medios (número de inserciones en cada soporte \times número de soportes).
- b. p se estimará mediante la audiencia promedio de los m soportes que constituyen el plan y,
- c. El parámetro ϕ , recurriendo a la expresión:

$$\phi = \frac{\sum_{i=1}^m n_i A_i}{\sum_{i=1}^m n_i \frac{1}{m} \sum_{i=1}^m A_i} = \frac{\sum_{i=1}^m n_i A_i}{\hat{N} \hat{p}}$$

De esta forma, para obtener la cobertura bastará sustituir $K = 0$ en [3.100]:

$$R_N = (1 - Pr(X=0)) \cdot POB$$

y la distribución de contactos sustituyendo sucesivamente $K = 1, 2, \dots, N$ en esa misma expresión.

3.3.1.8 Modelo de la distribución Kwerel-Geométrica

Como se señaló en el capítulo primero de esta tesis, algunos modelos optan por sustituir la función beta binomial que se encuentra implícita en la mayoría de los que hemos descrito, por una distribución de probabilidades alternativas. Leckenby y Kishi (1981) ofrecen dos de estos modelos. El primero de ellos, que desarrollaremos a continuación, calcula la cobertura del plan de medios según la fórmula de Kwerel descrita en el epígrafe 3.2.2.3 de esta tesis, para a continuación calcular la distribución de frecuencias de exposición mediante una distribución de probabilidades geométrica (modelo de la Distribución Kwerel Geométrica: KGD). El segundo de los modelos, que desarrollaremos en el apartado siguiente, aproxima la cobertura mediante el modelo de Hofmans, ya tratado en el apartado 3.1.1.4, obteniendo la distribución de frecuencias, también mediante la distribución geométrica (modelo de la Distribución Hofmans Geométrica: HGD).

En el modelo KGD, Leckenby y Kishi (1981) adaptan las expresiones que Kwerel (1969) dedujo para la duplicación de audiencias, a un esquema de audiencia neta acumulada. Para ello definen la cobertura del plan de medios como:

$$R = \frac{2BU \cdot BL}{BU + BL} \quad [3.101]$$

donde:

$$BU = A - \frac{2}{N} D$$

$$BL = (h + 1) \bar{A} - \frac{h(h + 1)}{2} \bar{D}$$

siendo h el más pequeño de los valores $N - 1$ o k , donde k es el entero más grande contenido en la *ratio* \bar{A}/\bar{D} , N es el número total de inserciones y donde, además:

$$\begin{aligned}\bar{D} &= \frac{2D}{N(N-1)} \\ \bar{A} &= \frac{A}{N} \\ A &= \sum n_i A_i \\ D &= \sum \binom{n_i}{2} d_i + \sum n_i n_j d_{ij}\end{aligned}$$

siendo A_i la audiencia del soporte i , d_i la acumulación resultante al realizar dos inserciones en el soporte i , d_{ij} la audiencia duplicada al realizar una inserción en los soportes i y j y n_i y n_j el número de inserciones que el plan de medios realiza, respectivamente, en los soportes i y j .

Con la cobertura del plan calculada según [3.101], la distribución de frecuencias de exposición se genera utilizando las siguientes expresiones de la distribución de probabilidad geométrica.

Para determinar la probabilidad de recibir una exposición:

$$Pr(1, N) = A - \frac{2D}{N-1} [1 + (1-a) + (1-a)^2 + \dots + (1-a)^{N-2}]$$

y, para el resto de probabilidades:

$$Pr(K, N) = \binom{N}{K} \alpha^{K-2} (1-a)^{N-K} \bar{D}$$

siendo α el resultado de resolver la ecuación de grado $N-2$ ésimo siguiente, conociendo que está acotado entre 0 y 1:

$$A - R - \bar{D} \left[\sum_{i=1}^{N-1} (N-i)(1-a)^{i-1} \right] = 0$$

3.3.1.9 Modelo de la distribución Hofmans Geométrica

Este modelo, propuesto como se ha señalado por Leckenby y Kishi (1981) es idéntico al anterior en el procedimiento de estimación de la distribución de frecuencias de contacto. La única diferencia reside en que estima la cobertura del plan de medios R , no mediante la fórmula de Kwerel (1969), sino recurriendo a la fórmula de Hofmans (1966):

$$R = \frac{A^2}{A + KD + kd}$$

siendo:

$$KD = \sum \sum n_i n_j K_{ij} d_{ij}$$

$$K_{ij} = \frac{A_i + A_j}{A_i + A_j - d_{ij}}$$

$$kd = \sum \frac{n_i(n_i - 1)^a (k_i d_i)}{2}$$

En estas expresiones, todas las variables y sus procedimientos de cálculo ya han sido explicitados en el modelo anterior, salvo el parámetro a que se estima mediante:

$$\hat{a} = \frac{\log \left[\frac{(3R_1 - R_3)R_2}{(2R_1 - R_2)R_3} \right]}{\log 2}$$

y las variables k_i y d_i que adoptan la expresiones:

$$k_i = 2R_{1i} - R_{2i}$$

$$d_i = 2R_{1i} / R_{2i}$$

Con la cobertura así estimada, basta referirse al modelo KGD para la estimación de la distribución de frecuencias de exposición.

3.3.1.10 Método de la distribución Hofmans Beta Binomial

Leckenby y Boyd (1984a) desarrollan un modelo similar al anterior, con la diferencia de que el cálculo de la distribución de frecuencias, no se realiza mediante la distribución geométrica de probabilidades, sino recurriendo a la distribución beta binomial (modelo de la distribución Hofmans Beta Binomial: HBBD).

El procedimiento descrito por estos autores, comienza obteniendo la cobertura del plan de medios mediante la fórmula de Hofmans, tal y como se explica en el modelo anterior. La principal innovación, sin embargo, radica en que la especificación de la distribución beta binomial no se realiza con la estimación directa que se aplicó en el modelo de Metheringham (1964) descrito en el apartado 3.3.1.5, sino que se recurre a un procedimiento iterativo de estimación directa denominado estimación mediante media y ceros que describimos a continuación.

- a. **Paso 1.** Mediante la aplicación de la fórmula de Hofmans descrita en el procedimiento anterior, se estima la cobertura del plan de medios: R_H .
- b. **Paso 2.** Arbitrariamente se da un valor A_0 al parámetro A de la distribución BBD (véase expresión 3.93).
- c. **Paso 3.** Mediante la siguiente expresión, se estima el parámetro B de la distribución BBD.

$$B_0 = \frac{A_0 \left(\sum_{i=1}^m n_i - \sum_{i=1}^m n_i A_i \right)}{\sum_{i=1}^m n_i A_i}$$

donde m es el número de soportes, n_i el número de inserciones realizadas en el soporte i cuya audiencia es A_i .

- d. **Paso 4.** Con los parámetros A_0 y B_0 así estimados, se calcula el número de individuos que no han resultado expuestos a ninguna de las N inserciones realizadas, mediante la expresión [3.39]. Dado que la cobertura es uno menos ese valor, se calcula así la cobertura que daría la aplicación directa de la BBD, y que llamaremos H_{BBD} .
- e. **Paso 5.** Se comparan las dos estimaciones de la cobertura obtenidas. Si la estimación mediante la fórmula de Hofmans es superior a la estimación mediante la BBD, se considera un nuevo valor del parámetro A (A_i) ligeramente inferior al anterior (A_0), procediendo a la inversa en caso contrario. Se repite el proceso hasta que la diferencia entre las dos estimaciones de la cobertura sea inferior al nivel de precisión decidido de antemano. Cuando esto sea así:
- f. **Paso 6.** Con los parámetros A_F y B_F finales, se calcula la distribución de contactos según [3.93].

3.3.1.11 Modelo de la distribución Hofmans Poisson Binomial

La misma filosofía es aplicada por Leckenby y Ju (1989) en la extensión del modelo PBD comentado en el apartado 3.3.1.7, y que da lugar al modelo de la distribución Hofmans Poisson Binomial (HPBD), y cuyas etapas describimos a continuación.

- a. **Paso 1.** Mediante la fórmula de Hofmans descrita, por ejemplo, en 3.3.1.9, se calcula la cobertura del plan de medios que denotaremos de nuevo, R_H .
- b. **Paso 2.** Con este valor, se estiman los parámetros ϕ y p de la distribución PBD (ϕ_0 y p_0) cuya expresión señalamos en [3.100] del siguiente modo:

$$p_0 = \frac{R_H}{m}$$

$$\phi_0 = \frac{\sum_{i=1}^m n_i A_i}{N p_0}$$

- c. **Paso 3.** Mediante la expresión [3.100], se estima la probabilidad de que un individuo no resulte expuesto a ninguna de las inserciones realizadas. Restando a uno, esta probabilidad, se obtendrá la cobertura del plan de medios estimada mediante la PBD, que denominaremos R_{PBD} .
- d. **Paso 4.** Si la cobertura así estimada, es superior a la que obtuvimos aplicando la fórmula de Hofmans, elegiremos un nuevo valor de p (p_1) ligeramente inferior al anterior (p_0) o superior en caso contrario. Repetiremos el proceso hasta que la diferencia entre las dos estimaciones de la cobertura sea inferior al nivel de precisión que se haya establecido de antemano. Una vez alcanzado este límite, con los parámetros p_F y ϕ_F obtenidos en la última iteración, pasamos al paso siguiente.
- e. **Paso 5.** Mediante la expresión [3.100] se estima la distribución de frecuencias del modelo HPBD.

3.3.2 Modelos *ad hoc*

3.3.2.1 Modelo de Hofmans

Hofmans (1966) combina sus dos modelos *ad hoc* de acumulación y duplicación de audiencias, para ofrecer un modelo de audiencia neta acumulada. Antes de comenzar su desarrollo, debe precisarse que, como la mayoría de modelos *ad hoc*, presenta la grave limitación de no permitir la obtención de la distribución de contactos. Inicialmente, el modelo calcula la cobertura de n inserciones realizadas en m soportes distintos, mediante la expresión:

$$R_n = \frac{\left(\sum_{i=1}^m n_i A_i \right)^2}{\sum_{i=1}^m n_i A_i + \sum_{i \neq j} KD + \sum_{i \neq j} kd} \quad [3.102]$$

donde n_i indica el número de inserciones realizadas en el soporte i y A_i es la audiencia del mismo. Por otra parte las expresiones $\sum_{i \neq j} KD$ y $\sum_{i \neq j} kd$ se obtienen:

$$\sum_{i \neq j} KD = \sum_{i=1}^m \sum_{j=i+1}^{m-1} n_i n_j k_{ij} A_{ij} \quad [3.103]$$

$$\sum_{i \neq j} kd = \sum_{i=1}^m \frac{n_i (n_i - 1)}{2} k_i d_i \quad [3.104]$$

La expresión [3.103] es la parte de la fórmula de Hofmans que aborda el cálculo de la duplicación y hace, para ello, uso de ecuaciones ya descritas por nosotros en el apartado 3.2.1.2. Así, k_{ij} se calculará como:

$$k_{ij} = \frac{A_i + A_j}{A_i + A_j - A_{ij}}$$

siendo A_{ij} las audiencias duplicadas resultantes de una inserción en los soportes i y j . Por otra parte, la expresión [3.104] es la componente acumulativa de la fórmula de Hofmans y recurre al modelo de acumulación expuesto en 3.1.1.4. Así, el parámetro a se estima por:

$$\hat{a} = \frac{\log \left[\frac{(3R_1 - R_3)R_2}{(2R_1 - R_2)R_3} \right]}{\log 2} \quad [3.105]$$

y las variables k_i y d_i de [3.104] mediante las expresiones:

$$d_i = 2R_{1i} / R_{2i}$$

$$k_i = 2R_{1i} - R_{2i}$$

donde toda la notación es conocida salvo R_{1i} y R_{2i} que pretende representar el alcance, respectivamente, de una y dos inserciones en el soporte i .

El modelo expuesto, sin embargo, fue modificado con posterioridad por Chandon (1985), partiendo de la base de que la necesidad de conocer la audiencia acumulada de tres inserciones en un soporte, provocada por la expresión [3.105] es un requerimiento de datos demasiado exigente cuando existen otros modelos que, como demuestra en su estudio, conociendo solamente el alcance acumulado de dos inserciones, ofrecen mejores resultados predictivos. Por este motivo, propone un procedimiento que pasa por calcular para cada soporte la audiencia acumulada tras las n inserciones mediante alguno de esos métodos, concretamente los de Agostini o beta binomial²². Con ese dato, bastaría estimar el parámetro a mediante la expresión:

$$\hat{a} = \frac{\log \left[\frac{R_n(nR_1 - R_n)}{R_2(2R_1 - R_2)} \right]}{\log (n-1)}$$

para que el cálculo de la cobertura mediante [3.102] no requiriese, como puede comprobarse, más que los datos de alcance acumulado tras una y dos inserciones en cada uno de los soportes.

3.3.2.2 Modelo de Morgensztern

Morgensztern (1970), propone una modificación de la fórmula de Hofmans que, como consecuencia, tendrá la misma limitación común a aquel y a todos los modelos *ad hoc*: no permitirá el cálculo de la distribución de contactos. Para este autor, la cobertura de un plan formado por n inserciones en m soportes distintos, adopta la expresión:

$$R_N = \frac{\left(\sum_{i=1}^m R_{ni} \right)^2}{\sum_{i=1}^m R_{ni} + \sum_{i=1}^m \sum_{j=i+1}^{m-1} \frac{k_{ij} A_{ij} R_{ni} R_{nj}}{A_i A_j}} \quad [3.106]$$

donde k_{ij} se calcula como en el modelo anterior mediante la expresión:

$$k_{ij} = \frac{A_i + A_j}{A_i + A_j - A_{ij}}$$

siendo toda la notación conocida, salvo R_{ni} (R_{nj}) que es el alcance acumulado tras n inserciones en el soporte i (j). Como se desprende del análisis de la expresión [3.106], la utilización de este alcance acumulado, implica la necesidad de estimarlo mediante algún modelo de acumulación de audiencias, a partir de su valor tras una y dos inserciones. Chandon (1985), al igual que en el modelo anterior, recomienda recurrir a los modelos de Agostini o al beta binomial.

3.3.2.3 *Modelo promedio Metheringham-Dirichlet multinomial*

Chandon (1985), al contrastar empíricamente los modelos, encuentra que el de Metheringham y el Dirichlet multinomial presentan comportamientos especulares. Es decir, cuando el primero subestima la distribución de contactos, el segundo la sobrestima y viceversa. Este hecho le lleva a proponer un modelo cuya distribución de contactos se obtenga de promediar las distribuciones de frecuencias obtenidas tras aplicar los dos modelos anteriores, modelo que no requiere de comentario adicional alguno dada la sencillez de su concepción.

4

Objetivos y diseño metodológico

4.1 OBJETIVOS: FORMULACIÓN Y JUSTIFICACIÓN

En los capítulos precedentes, se ha intentado establecer la evidencia de una serie de hechos que, encadenados, justifican la necesidad y el interés de llevar a cabo la labor que constituyen los objetivos de esta tesis doctoral: estimar la precisión de las estimaciones que los modelos de exposición realizan de la cobertura y la distribución de contactos de un plan de medios y, a continuación, identificar qué factores determinan esa mayor o menor precisión estimativa.

En primer lugar, se ha puesto de manifiesto el papel central que juegan las variables de cobertura y distribución de contactos en una correcta planificación de medios y que se deriva de dos aspectos. Por un lado, como señalan Leckenby y Kishi (1984), estas variables son un *input* primordial de los modelos de planificación de medios, tal y como se definieron en el capítulo 1; es decir, los modelos que realizan la tarea completa de repartir de una manera eficiente el presupuesto entre el conjunto de soportes posibles estableciendo, en la mayoría de los casos, el esquema temporal de las inserciones. Nos estamos refiriendo a modelos ya reseñados como los de Aaker (1975), Little y Lodish (1966; 1969), Zufryden (1975a), Rust (1985) o Rust y Stout (1989), entre otros. Cabe esperar que la utilización de un modelo que ofrezca una estimación más precisa de la cobertura y la distribución de contactos, se traducirá en que los planes de medios obtenidos mediante los modelos de planificación, serán más precisos. Y éste, no es un hecho que carezca de importancia, dados los volúmenes de inversión publicitaria que se dan en nuestro entorno y, también, en nuestro país. Así, si el modelo empleado estuviera subestimando sistemáticamente la cobertura de los planes de medios que ayuda a construir, el anunciante podría alcanzar sus objetivos de cobertura con un presupuesto menor (Leckenby y Kishi, 1984). A título ilustrativo, basta señalar que, para una empresa como el Corte Inglés, una reducción en su inversión publicitaria de tan sólo un 1%, le hubiese supuesto un ahorro, en 1994, de más de cien millones de pesetas (Infoadex, 1995).

Por otro lado, el origen de la importancia de estas variables para una correcta planificación de medios, también se deriva de la influencia que tienen la mismas en el establecimiento de alguna medida de la eficacia de la publicidad. Como se señaló de una manera amplia en el capítulo primero de esta tesis doctoral, el concepto de frecuencia efectiva, esto es, la frecuencia o rango de frecuencias a los cuales la respuesta del mercado al mensaje inserto en un soporte determinado, es la deseada por el planificador (Sissors y Bumba, 1989), es uno de los pilares fundamentales sobre los cuales se ha construido el concepto de eficacia publicitaria. La mayoría de los estudios realizados (Naples, 1979; Sissors, 1982) relacionan la exposición *al*

anuncio con diversas medidas de eficacia y, aunque no hay un indicador generalmente aceptado que permita obtener la exposición al anuncio a partir de la exposición *al soporte*, que es la que nos proporcionan los modelos que estamos analizando, todas las aproximaciones al problema, requieren de una correcta estimación de la distribución de frecuencias o de la frecuencia media de exposición del individuo al mismo. Así, Naples (1979) relaciona la eficacia publicitaria en términos de preferencia de marca con diversos niveles de OTS y Leckenby y Wedding (1984) multiplican cada uno de los valores de la distribución de frecuencias de contacto por unos coeficientes estimados experimentalmente para determinar su aproximación a la eficacia de la publicidad: los valores de respuesta al mensaje/medio (MMR).

También se ha puesto de manifiesto en capítulos precedentes un hecho que, desde nuestra perspectiva, contribuye a justificar la necesidad de estimar la precisión en las estimaciones de los modelos de exposición, como es el uso generalizado que hacen de los mismos, los planificadores de las agencias de publicidad y centrales de medios. Leckenby y Kim (1994a) constatan que más del 80% de las agencias utilizan algún modelo en la estimación de la cobertura y, aproximadamente este porcentaje, los emplean en la estimación de la distribución de contactos. Sin embargo, tanto este trabajo, como los de Leckenby y Kishi (1982a), Kreshel, Lancaster y Toomey (1985) o Lancaster, Kreshel y Harris (1986) se destaca que los usuarios de estos modelos son conscientes de que los errores en las estimaciones que proporcionan (a) existen y (b) deben ser estimados y corregidos.

Una justificación adicional para el desarrollo de este capítulo, procede de los mecanismos disponibles para la corrección del error de estimación. En palabras de Leckenby y Boyd (1984), hay dos tipos de error que afectan a la cobertura y la distribución de contactos. Por un lado, el error de muestreo de los estudios de audiencia que proporcionan la información que les sirve como *input* y, por otro lado, el error de estimación que aporta el propio modelo. Mientras que sería relativamente

caro reducir el error de los estudios, dado que ello implicaría aumentar el tamaño de la muestra de los mismos, ya de por sí elevado, el coste de la reducción del error a través de la mejora de la modelización es mínimo en comparación con el anterior. Uno implica cambios en el proceso de recogida de datos, mientras que el otro se basa en modificaciones conceptuales.

Finalmente, debe destacarse que diversos autores (Leckenby y Ju, 1990) han señalado la conveniencia para el enriquecimiento de la planificación de medios, del desarrollo de estudios de carácter cross-cultural, que permitan evaluar la capacidad de estimación de los modelos frente a bases de datos que reflejen realidades socioculturales diferentes de Estados Unidos país donde, como se verá, se han desarrollado casi en exclusiva, estos trabajos con anterioridad. Esta perspectiva dota, también, de sentido adicional a nuestro trabajo.

Visto lo anterior, la *primera cuestión a investigar* en nuestra tesis doctoral, puede formularse en los siguientes términos:

***CUESTIÓN PRIMERA [CI₁]: EVALUACIÓN Y
JERARQUIZACIÓN DE LOS MODELOS DE
EXPOSICIÓN***

CI_{1A}. Evaluar los modelos de exposición, determinando el error cometido en la determinación de la cobertura y distribución de contactos de los planes de medios, en los medios prensa y diarios.

CI_{1B}. Llevar a cabo la jerarquización de los distintos modelos, tomando como punto de partida los resultados del proceso de evaluación señalado con anterioridad, aportando directrices para la creación futura de nuevos

modelos y para la selección adecuada, entre los actuales, por parte del planificador.

CI_{1c}. Como acción derivada del proceso jerarquizador, analizar la influencia que tienen sobre el error la sucesiva ruptura de las hipótesis simplificadoras en las que se basan los modelos, respecto a las probabilidades de exposición de la población y los soportes.

Algunos autores (Leckenby y Kishi, 1981; 1982b) afirman, sin embargo, que no basta con establecer qué modelos funcionan mejor que otros para un conjunto de planes de medios. En la práctica, estos planes difieren en variables tales como sus niveles de audiencia bruta, de acumulación y duplicación de las mismas, el número de soportes que los integran y, también, la distinta cantidad de inserciones que se realizan en cada uno de ellos. Son evidentes las importantes implicaciones que, para la correcta planificación de medios, se derivarían del hecho de detectar las relaciones entre esas características internas de los planes y la precisión de sus estimaciones: sería posible seleccionar el modelo más adecuado, no por su capacidad predictiva global, sino atendiendo a las peculiaridades del plan de medios al cual va a aplicarse. Este hecho nos lleva a formular nuestra *segunda cuestión a investigar* en los siguientes términos:

CUESTIÓN SEGUNDA [CI₂]: DETERMINACIÓN DE LOS FACTORES DE ERROR

CI_{2a}. Determinar la influencia que ejercen las características estructurales de los planes de medios sobre la precisión de los distintos modelos de exposición.

CI_{2B}. Establecer la adecuación o no de los diferentes modelos al medio diarios cuando, inicialmente, fueron concebidos únicamente para el medio revistas.

CI_{2C}. Tomando como base los resultados anteriores, establecer un patrón de selección del modelo más adecuado en función de esas características internas de los planes de medios, que sirva como criterio adicional para discriminar entre los modelos que, de acuerdo con la jerarquización realizada en Q1, tengan precisiones similares en sus estimaciones.

COROLARIO

Derivar de los resultados anteriores las directrices normativas pertinentes para el diseño futuro de nuevos modelos de exposición. Si ello es posible se creará un nuevo modelo que, si estas directrices son válidas, debería ofrecer mejores estimaciones de la cobertura y la distribución de contactos que los evaluados.

Debe señalarse que en España no se han desarrollado estudios de esta naturaleza lo que, bajo nuestra perspectiva, acentúa su necesidad. Así, existen trabajos que ilustran el funcionamiento de algún modelo concreto como el Dirichlet multinomial (León, 1986) o la fórmula de Hofmans (1966) ampliada (Barreiro, 1991), pero no evalúan su funcionamiento. Más recientemente, en su tesis doctoral, Redondo (1995) realiza el esfuerzo más notable de los llevados a cabo en esta línea, sin embargo, contrasta sólo el funcionamiento de algunos modelos (no evalúa los de acumulación de audiencias y solo da ejemplo de cálculo de la mayoría de los de audiencia neta acumulada) y lo hace, no frente a la realidad medida mediante un

estudio de audiencias, sino frente a la interpretación que de la misma realiza otro modelo, el conjunto de programas TOM.

Dado que, sin embargo, trabajos como el que presentamos sí que han sido habituales en el entorno norteamericano (Chandon, 1985; Leckenby y Kishi, 1981; Kishi, 1982; Leckenby y Boyd, 1985; Leckenby y Kim, 1993; Leckenby y Hong, 1997) y algo menos frecuentes en el resto de países como pueda ser Nueva Zelanda (Danaher, 1988b; 1989; 1991) o Japón (Kishi, 1987), se han seguido las directrices marcadas por los mismos en el diseño metodológico que, inmediatamente, se precisará.

Justificados y formulados los objetivos de esta tesis doctoral, en los epígrafes siguientes se describirá, en primer lugar, la metodología empleada en este proceso: describiendo la fuente de datos utilizada¹, los patrones seguidos en la confección de los planes de medios, la justificación de las definiciones de error empleadas y las técnicas de análisis estadístico utilizadas..

4.2 DISEÑO DE LA INVESTIGACIÓN

En términos generales, la evaluación del funcionamiento de un modelo de exposición es sencilla: comparar la cobertura y distribución reales, medidas mediante un estudio de audiencias, con las estimaciones que proporcionan los diferentes modelos, previa definición de las pertinentes medidas de error. Esta sencillez es, sin embargo aparente: los estudios de audiencias no siempre se ajustan exactamente a las necesidades del investigador, la operativización de la mayoría de los modelos es compleja y no existe un acuerdo generalizado acerca de qué medidas del error son las más adecuadas. En el presente epígrafe se revisará el diseño experimental seguido en esta tesis doctoral, destacando los problemas que han debido abordarse y las

¹ Para el caso español, la fuente de datos que utilizaremos es el Estudio General de Medios (EGM) que lleva a cabo unas 40.000 entrevistas repartidas en tres oleadas (Méndez, 1996).

soluciones que se han adoptado, siempre tomando en consideración las indicaciones de los principales trabajos internacionales en este campo.

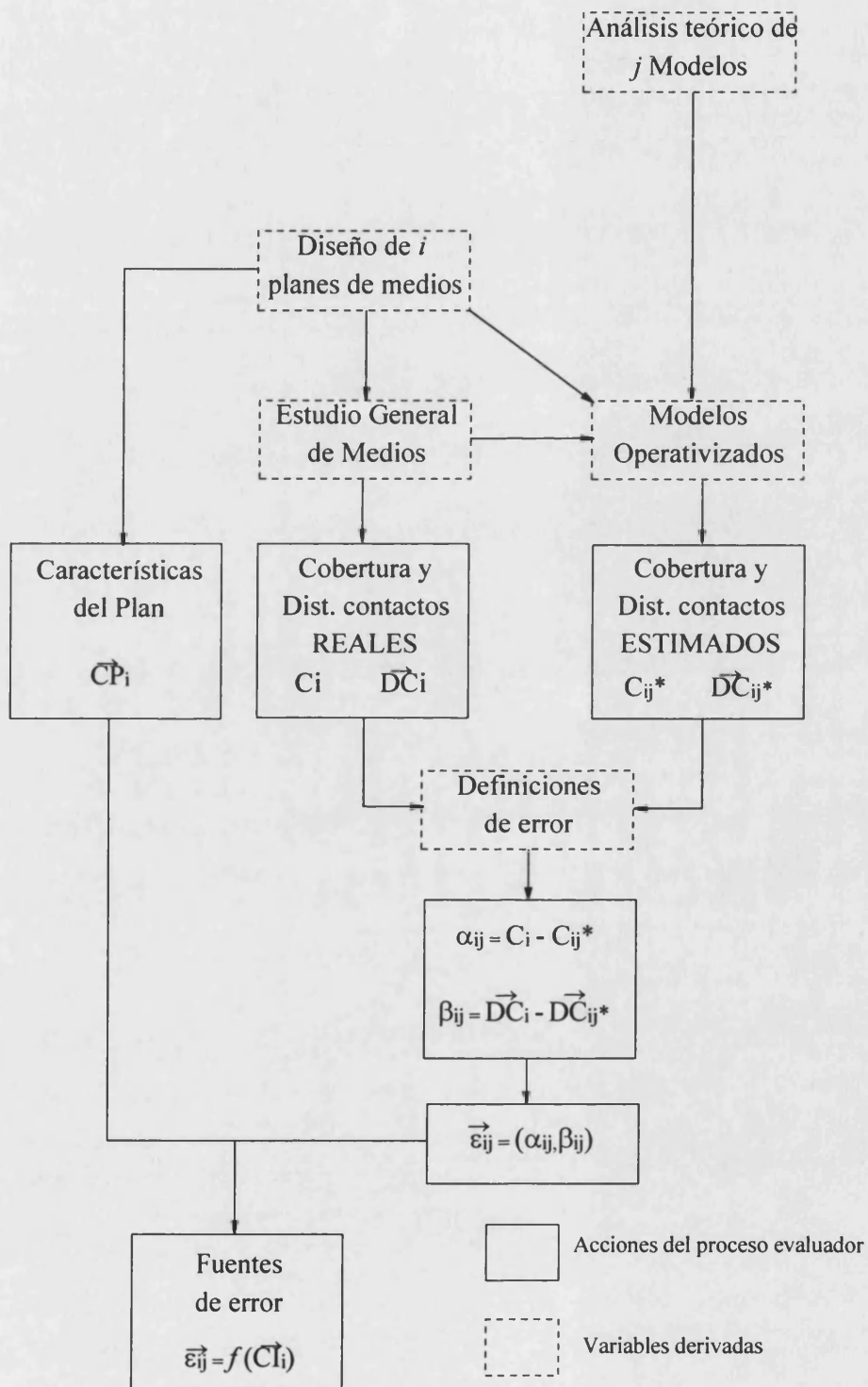
Uno de los principales problemas con que se encuentra el investigador que desee conocer qué modelos de exposición funcionan mejor, es la falta de homogeneidad en los diseños experimentales llevados a cabo para la contrastación de la precisión de las estimaciones que producen (Leckenby y Ju, 1990). Este hecho afecta, por una parte, a elementos que necesaria y, diríamos que de manera deseable por el interés de estudios coss-culturales antes señalado, han de ser diferentes, como las bases de datos empleadas. Nada justifica, sin embargo, que el número de planes analizados sea desproporcionadamente diferente en cada trabajo ni, tampoco, que las definiciones de error empleadas sean tan distintas.

El evitar las innecesarias diferencias metodológicas, ha llevado recientemente a autores como Leckenby y Kim (1992) ha establecer una serie de criterios básicos que deberían seguir todos los trabajos como el presente en sus contrastaciones empíricas. La autoridad del profesor Leckenby en nuestra línea de investigación es tan destacada (véase el cuadro 1.10 como prueba de ello), que se ha decidido seguir en todo momento las indicaciones de sus trabajos, como se podrá comprobar a lo largo del desarrollo de este epígrafe.

La figura 4.1, ilustra el diseño metodológico empleado. En una primera fase, se diseñan un número suficiente de i planes de medios para que los resultados puedan ser generalizables. Cada uno de estos planes, incluye un conjunto de características internas: su audiencia bruta, su nivel de duplicación entre soportes y de acumulación entre dos inserciones, diferente número de soportes, soportes de distinto tamaño en términos de audiencia, de modo tal que permitirán establecer, con posterioridad, si el error cometido por el modelo es función de alguna de ellas. Finalmente, están formados bien por diversas inserciones en un único soporte, bien por una inserción en diversos soportes y el caso más general, varias inserciones en varios soportes.

La base de datos seleccionada para el desarrollo de este trabajo, ha sido el Estudio General de Medios. Las características del mismo, nos permitirán, como se verá, obtener mediante tabulación los valores reales de la cobertura y la distribución de contactos de cada uno de los planes. A la vez, nos ofrecerá los datos necesarios para que los modelos descritos en el capítulo 3, estimen estas mismas variables. De la comparación de los valores reales y estimados se obtendrán diversas medidas del error que cada modelo comete, tanto en la estimación de la cobertura, como de la distribución de contactos, lo que nos permitirá jerarquizar los modelos analizados en base a la precisión de sus estimaciones. En una segunda fase, se intentará poner en relación el error cometido por cada modelo, con las características descritas anteriormente de cada plan. El objeto es establecer si un modelo es más adecuado que otro para planes de determinadas características

Figura 4.1. Diseño experimental de la tesis doctoral



Fuente: Elaboración propia.

Señalados de manera global los principios en los que se basa el diseño experimental propuesto, se pasará a continuación a analizar, de una manera detallada, cada una de las etapas avanzadas.

4.2.1 Fuente de datos: el Estudio General de Medios.

Leckenby y Boyd (1985) señalan que, la contrastación de los términos de cobertura y distribución de contactos estimados por los modelos de exposición, frente a los datos tabulados de un estudio de audiencias, ha demostrado ser la mejor forma de valorar la precisión en las estimaciones de los mismos. El cuadro 4.1 sintetiza los estudios de audiencia empleados en la mayoría de trabajos empíricos con el objetivo mencionado, comprobándose que, todos ellos, han recurrido a esta alternativa metodológica.

Las principales diferencias entre los estudios de audiencia utilizados, proceden, sin embargo, del procedimiento de recogida de información propio de cada uno de los ellos. Como se puede comprobar, en todos los estudios estadounidenses, que son la mayoría, la técnica utilizada ha sido el método *Through The Book* (TTB), cuya fundamentación se discutió suficientemente en el apartado 2.2.1. Este procedimiento provoca, según Leckenby y Boyd (1984a), serias limitaciones en la evaluación de los modelos. El Simmons Market Research Bureau (SMRB), recoge los datos de lectura y visión de programas, únicamente en dos momentos del tiempo. Como se basa en mostrar un número del soporte para preguntar si ha sido visto o leído, este diseño provoca que la exposición más allá de dos inserciones no sea medida, no siendo posible analizar planes que contengan más de dos inserciones en cada soporte.

Cuadro 4.1. Fuente de datos de los principales estudios empíricos sobre evaluación de los modelos de exposición.

Estudio	Fuente de datos	Año de la muestra	País	Método de recogida datos
Leckenby y Kishi (1981)	W.R. Simmons	1.972	EE.UU.	TTB
Kishi (1982)	SMRB	1.976	EE.UU.	TTB
Leckenby y Kishi (1982b)	W.R. Simmons	1.972	EE.UU.	TTB
Leckenby y Kishi (1982c)	SMRB	1.979	EE.UU.	TTB
Leckenby y Boyd (1984a)	SMRB	1.979	EE.UU.	TTB
Leckenby y Kishi (1984)	SMRB	1.979	EE.UU.	TTB
Leckenby y Rice (1985a)	SMRB	1.982	EE.UU.	TTB
Chandon (1985)	W.R. Simmons	1.972	EE.UU.	TTB
Boyd (1985)	SMRB	1.976	EE.UU.	TTB
Leckenby y Boyd (1985)	SMRB	1.979	EE.UU.	TTB
Leckenby y Rice (1985b)	SMRB	1.982	EE.UU.	TTB
Leckenby y Rice (1986a)	SMRB	1.983	EE.UU.	TTB
Leckenby y Rice (1986b)	SMRB	1.982	EE.UU.	TTB
Danaher (1988b)	AGB	1.985	N. Zelanda	RR
Danaher (1989)	AGB	1.985	N. Zelanda	RR
Leckenby y Ju (1989)	SMRB	1.979	EE.UU.	TTB
Danaher (1991)	AGB	1.985	N. Zelanda	RR
Leckenby y Kim (1993)	SMRB	1.979	EE.UU.	TTB
Leckenby y Kim (1994b)	SMRB	1.979	EE.UU.	TTB
Aldás (1995)	EGM	1.995	España	RR
Redondo (1996)	EGM	1.994	España	RR
Leckenby y Hong (1997)	Estudio propio	1.996	EE.UU.	TTB

SMRB = Simmons Market Research Bureau

EGM = Estudio General de Medios

AGB = AGB Research New Zealand Ltd.

TTB = Through The Book

RR = Recent Reading

Fuente: Elaboración propia

Este hecho deriva en que los estudios norteamericanos analicen la capacidad predictiva de los modelos de exposición ante planes tremendamente irreales en término de número de inserciones, pero comporta además, desde nuestra perspectiva, una seria limitación adicional: impide la evaluación de los modelos de acumulación de audiencias. Mientras que nada impide que se analicen los modelos de duplicación de audiencias: una inserción en diversos soportes es compatible perfectamente con los datos suministrados por el SMRB, ni tampoco los modelos de audiencia neta acumulada, aunque sea con planes poco habituales en la práctica de dos inserciones en diversos soportes, la mayoría de los modelos de acumulación necesitan para

calcular los parámetros que los especifican, al menos las coberturas tras las dos primeras inserciones. Si la única información que proporciona el SMRB, ya se utiliza para especificar el modelo, el investigador carece de observaciones “reales” con las cuales contrastar las estimaciones del mismo.

El método *Recent Reading* carece, al menos en parte, de esta limitación. Así lo señala Danaher (1988b; 1989; 1991) cuando describe la metodología seguida por AGB-McNair Surveys New Zealand Ltd. en la realización del estudio de audiencia en que basa su estudio. Según este autor, las preguntas relevantes para su análisis, son las siguiente (ejemplo para revistas semanales):

1. ¿Ha leído u hojeado usted personalmente algún número de [nombre de la revista] en los últimos siete días. No importa el lugar. *Respuesta si / no.*
2. ¿Cuántos números distintos de la revista [nombre de la revista] lee u hojea usted personalmente en un mes típico? *Respuesta: 0, 1, 2, 3, 4 números.*

Estas preguntas son modificadas, según la frecuencia de aparición de la publicación, de acuerdo con el cuadro 4.2:

Cuadro 4.2. Preguntas relevantes del estudio AGB-McNair Surveys New Zealand Ltd.

Revista	Periodo Pregunta 1	Periodo Pregunta 2	Contestación Pregunta 2
Semanal	7 días	Mes típico	0-4 números
Quincenal	2 semanas	Últimos 3 meses	0-6 números
Mensual	1 mes	Últimos 6 meses	0-6 números
Bimensual	2 meses	Últimos 12 meses	0-6 números

Fuente: Danaher (1988b; 1989; 1990)

De esta forma se obtienen, simultáneamente, dos ventajas. Por un lado, los planes de medios pueden ser más realistas, dado que es procedente la realización de entre cuatro y seis inserciones en cada plan de medios y, en segundo lugar, pueden evaluarse los modelos de acumulación de audiencias, puesto que aunque se utilicen las coberturas de las dos primeras inserciones para calcular los parámetros y especificar los modelos, restan los datos de cobertura tras 3 hasta 6 inserciones (cuatro en las revistas semanales), para comparar con las estimaciones de los mismos.

La situación para el caso español, y por ello en nuestro estudio, es bastante similar. La fuente de datos a la que se ha recurrido, ha sido el Estudio General de Medios (EGM) realizado por la Asociación para la Investigación de los Medios de Comunicación (AIMC). Las principales características metodológicas de este estudio de audiencia, son las siguientes (Méndez, 1996):

- a. Es un estudio *poblacional*, en el sentido de que no se trata de representar a los lectores u oyentes o espectadores, sino que busca una representación adecuada de la población española en su conjunto. A esta muestra se le interroga, entre otras cosas², acerca de su comportamiento en relación al consumo de medios, de modo que el que los mencionados lectores, oyentes o espectadores estén representados en la misma, es más una consecuencia que una premisa.
- b. Es un estudio *multimedia*. Es decir, estudia para cada individuo su comportamiento respecto a los distintos medios de comunicación, concretamente se proporcionan datos de las audiencias de: diarios, suplementos, revistas, cine, radio y televisión.

² Además de contestar a diversas variables de clasificación, el entrevistado es cuestionado acerca del equipamiento de su hogar, como indicador de su estatus, sobre el nivel del consumo que realiza en más de cuarenta productos de consumo del hogar y de consumo individual y, finalmente, acerca de sus estilos de vida: temas de banca/seguros, ocio, vacaciones, compras, etc.

- c. Es un estudio *anual*. El diseño muestral, es anual, aunque se divida posteriormente en tres oleadas la realización de las entrevistas a efectos operativos. Este es un hecho muy importante, dado que sería incorrecto utilizar en nuestro análisis los datos correspondientes a una sola ola. En ese caso, podríamos encontrarnos con estratos o segmentos de población, y por ello con datos de audiencia de determinados soportes, infra o sobrerrepresentados, en cuanto que sólo adquieren su verdadero valor de representación en el resultado anual.
- d. El *universo* está constituido por los individuos de 14 o más años residentes en hogares unifamiliares de la España peninsular, Baleares y Canarias. En el caso del ya mencionado estudio neozelandés del AGB, el universo lo componen los individuos de diez o más años.
- e. La *muestra* anual está en torno a los 40.000 individuos, dividida en tres muestras de igual tamaño y diseño. A su vez, cada muestra se subdivide en siete submuestras iguales, una para cada día de la semana. Esto se debe a que se estudian medios de aparición diaria, para los que el comportamiento de la audiencia difiere de unos días a otros. El muestreo es aleatorio, polietápico y estratificado, definiéndose los estratos por el cruce provincia-hábitat. La afijación, es básicamente proporcional, con una ligera sobrerrepresentación de los hábitats superiores a 50.000 habitantes y un mínimo muestral de 300 unidades por provincia y año³. El tamaño muestral es uno de los mayores de Europa (Bigné, 1994), lo que confiere gran credibilidad a sus resultados.
- f. La *recogida de información* se produce mediante entrevista personal en el hogar del entrevistado, por un equipo de entrevistadores entrenados especialmente para la realización del EGM. El cuestionario es estructurado

³ En García Ferrando, Ibáñez y Alvira (1993) puede seguirse el desarrollo, paso a paso, del diseño de una muestra de idénticas características a la descrita.

y cerrado casi en su totalidad: la codificación de las emisoras de radio se produce posteriormente en gabinete.

- g. Finalmente, la *supervisión* de las entrevistas se realiza por el instituto encargado del trabajo de campo, pero además un 10% es supervisado directamente por AIMC, con un equipo de inspección propio e incluso, en ocasiones, los cuestionarios son sometidos a una nueva comprobación por parte de alguno de los asociados que así lo requieran, después de la tabulación de los resultados.

Nuestro estudio se ha centrado, por los motivos que se expondrán con posterioridad, en los medios prensa y revistas. Las preguntas del EGM que se han utilizado para obtener mediante tabulación los valores reales de cobertura y distribución de contactos de los planes de medios diseñados, son las siguientes:

1. **A continuación, le voy a mostrar los títulos de unos determinados periódicos-diarios, le agradecería que me fuera diciendo cuáles de ellos ha leído u hojeado en los últimos 30 días. ¿Qué periódicos de *Información General* [*Información Económica; Información Deportiva*] ha leído u hojeado en los últimos 30 días?**
2. De cada 7 ejemplares de [periódico] que salen en una semana, ¿Cuántos acostumbra a leer u hojear? *Respuesta de 0 a 7.*
3. Sin contar el día de hoy, ¿Cuándo ha leído por última vez un ejemplar de [periódico]?

Para las revistas, las preguntas son esencialmente las mismas, con las adaptaciones que recoge el cuadro 4.3.

Cuadro 4.3. Preguntas relevantes del EGM para el medio revistas.

Revista	Periodo Pregunta 1	Periodo Pregunta 2	Contestación Pregunta 2
Semanal	6 meses	1 mes	Común: 1. Todos/casi todos 2. Aprox. la mitad 3. Pocos números 4. Ninguno
Quincenal	6 meses	3 meses	
Mensual	6 meses	6 meses	

Fuente: AIMC

A diferencia del medio diarios, el medio revistas plantea un problema operativo importante. La respuesta a la pregunta 2 en los periódicos, es directamente una frecuencia de exposición: de siete ejemplares de una semana, se han leído, uno, dos... o siete. Sin embargo la equivalencia entre la contestación a esa pregunta y una frecuencia, no es directa en el medio revistas (cuadro 4.3). Como ilustra la figura 4.2, los niveles de la respuesta, y el número de ejemplares a los que se ha resultado expuesto, no son los mismos, lo que requiere una recodificación de las respuestas para soslayar esta limitación.

Figura 4.2. Discrepancia entre respuestas y número de contactos

Periodicidad	Número de revistas a las que se ha podido resultar expuesto						
	0	1	2	3	4	5	6
Mensual	Ninguno	Pocos números		Aproximadamente la mitad		Todos o casi todos los números	
Quincenal	Ninguno	Pocos números		Aproximadamente la mitad		Todos o casi todos los números	
Semanal	Ninguno	Pocos números		Todos o casi todos los números			
				Aproximadamente la mitad			
	Respuestas posibles a la pregunta 2						

Fuente: Elaboración propia

Con el fin de resolver esta limitación de la base de datos utilizada, se ha empleado el siguiente procedimiento de recodificación de las variables: para las

revistas mensuales y quincenales, se ha generado un número aleatorio comprendido entre cero y uno para cada uno de los individuos. Si el número era superior o igual a 0'5, se le asignaba la respuesta numérica más alta correspondiente con su contestación real. Es decir, si había contestado *pocos números*, se le asignaba dos, si había contestado *aproximadamente la mitad*, cuatro y *todos o casi todos*, seis. Caso de ser el número aleatorio inferior a 0'5, se le asignaba la respuesta numérica inferior. Si el individuo había contestado que no había leído ningún número, el procedimiento no se aplicaba y se le asignaban cero contactos.

Para el caso de las revistas semanales, el procedimiento ha sido muy similar, La única diferencia reside en que, al ser solamente posibles un máximo de cuatro contactos, algunos niveles de frecuencia, reciben individuos provenientes de dos contestaciones distintas a la pregunta dos. Así, se le asignan dos contactos, tanto a los que han contestado *pocos números*, cuando el número aleatorio generado es superior a 0'5, como a los que han contestado *aproximadamente la mitad*, cuando su número aleatorio es inferior a esa cifra, tal y como se observa, de nuevo, en la figura 4.2.

Las características de la base de datos empleada, imponen una limitación adicional al diseño experimental, relacionado con el número de inserciones que pueden realizarse en cada soporte al diseñar los planes de medios que se van a emplear para evaluar los distintos modelos. Aparentemente, y según las preguntas realizadas en el EGM que se han detallado con anterioridad, sería posible diseñar planes de medios en los cuales, en cada soporte, se realizasen inserciones cuyo número podría variar entre uno y cuatro, en las revistas semanales, de uno a seis en las quincenales y de uno a siete en el medio diarios. Sin embargo, para poder obtener mediante tabulación los valores reales de la cobertura y la distribución de contactos, y compararlos así con las estimaciones de los modelos, no es posible, como se verá a continuación, diseñar planes con cualquier número de inserciones dentro del

intervalo permitido, siendo necesario plantearse planes de medios formados por el número máximo de inserciones posibles en cada soporte.

Supóngase que, para evaluar un modelo determinado, se plantea un plan de medios formado por tres inserciones en una revista mensual. Como quiera que el EGM pregunta al entrevistado cuántos de los seis números posibles, ha leído u hojeado, e imaginando que su comportamiento real es el ilustrado en la figura 4.3, éste contestará que tres de ellos, pero ¿qué tres en concreto?, Si el plan de medios que se ha diseñado es el tipo A, los contactos reales habrán sido sólo 2, correspondientes a las inserciones primera y tercera. Si por contra, el plan de medios ha sido del tipo B, el único contacto real, dado el comportamiento del sujeto, habrá sido la segunda. Como los datos proporcionados por el EGM no nos permiten saber exactamente cuáles han sido exactamente los números con los que se ha tenido contacto, la única manera de evitar este inconveniente, es plantear planes de medios formados por el máximo número de inserciones posibles, en este caso seis. Como puede comprobarse en la figura 4.3, el plan de medios que hemos llamado “experimental”, por ser el que hemos planteado en el desarrollo de esta tesis doctoral, al saturar el horizonte temporal con el número máximo de inserciones, hará que el número de contactos tabulados según los datos del EGM (tres), se correspondan exactamente con los recibidos por el individuo en cuestión (tres), haya sido cual haya sido el comportamiento real de éste respecto a las inserciones.

Figura 4.3. Contactos reales y estimados de planes alternativos.

Plan de Medios	Inserciones						Contactos Reales	Contactos Estimados
	1	2	3	4	5	6		
A	1		1		1		2	3
B		1		1		1	1	3
Experimental	1	1	1	1	1	1	3	3
Comportamiento Real	1	1	1					

Fuente: Elaboración propia

Otra limitación que impone las características de la base de datos, y que ha sido necesario resolver en el proceso de evaluación de los modelos, hace referencia a la estimación de la cobertura tras N inserciones en los modelos de acumulación de audiencias. De la lectura del fichero de microdatos del EGM, se puede obtener de manera inmediata la distribución de contactos tras N inserciones (7 para diarios, 4 para revistas semanales y 6 para revistas mensuales, como se ha visto). Sin embargo, como se puso de manifiesto en el capítulo tercero, algunos modelos requieren para su estimación los datos de la cobertura un número determinado de inserciones. En este sentido, Chandon (1985) ya diseñó un procedimiento para obtener esta última variable a partir de la anterior, que es el que nosotros hemos seguido y que, a continuación, expondremos.

Supongamos la siguiente distribución de contactos generada por cinco inserciones en un soporte determinado:

Cuadro 4.4. Ejemplo de distribución de contactos

Contactos	0	1	2	3	4	5	Total
Individuos	50	20	10	5	5	10	100

Fuente: elaboración propia

Supongamos, además, que sea cuál sea la probabilidad individual de exposición de un sujeto determinado, ésta es independiente tanto del número de la revista en el que se esté realizando la inserción como, también, de la cantidad de exposiciones a números anteriores. Pues bien, entre los individuos que han recibido cero contactos, la probabilidad de que hubiese resultado expuesto a un número en concreto, es nula. Entre los individuos que han resultado expuestos a una de las cinco inserciones, la probabilidad de que ese contacto fuera, digamos, con el primer

número es $1/5$, $2/5$, entre los que han recibido dos contactos y, así, sucesivamente. Estos resultados pueden generalizarse mediante la ley hipergeométrica⁴ como se tratará de ilustrar a continuación.

Considérese un fenómeno aleatorio como el que nos ocupa, modelizado mediante la variable aleatoria ξ . Este fenómeno consiste en que, de un colectivo integrado por un total de N elementos, cada uno de los cuales puede tener dos modalidades exclusivas, se consideran n de ellos, de tal manera que x reúnen una de dichas modalidades y $n - x$ de ellos reúnen la otra. De acuerdo con el principio de que una vez que un elemento ha sido seleccionado al azar, para aparecer en el grupo de los n , deja de pertenecer al colectivo formado por los N o, dicho de otra forma, no se produce su sustitución, la composición del colectivo se altera y, con ella, la probabilidad correspondiente a cada elemento, ante cada extracción para constituir el grupo de n elementos.

Veámoslo tomando como ejemplo el caso que nos ocupa y que ilustramos en el cuadro 4.4. En él tenemos cinco inserciones ($N = 5$), de las cuales una es la primera de las realizadas ($N_1 = 1$) y el resto no son la primera (N_2). Considérese que queremos conocer la cobertura tras dos extracciones ($n = 2$). En primer lugar se determinará cuál es la probabilidad entre los individuos que pertenecen a cada uno de los niveles de la distribución de contactos de haber visto la primera ($\xi = 1$). López Cachero (1990; pág. 318), demuestra que esta probabilidad se puede escribir de forma general como:

$$P(\xi = x) = \frac{\binom{N_1}{x} \binom{N_2}{n-x}}{\binom{N}{n}}$$

que en el ejemplo que nos ocupa, adoptaría el valor:

⁴ Se ha seguido para ello López Cahero (1990), pp. 317-322 y Cuadras, Echeverría, Mateo y Sánchez (1991), pp. 277-290.

$$P(\xi = 1) = \frac{\binom{1}{1} \binom{4}{2-1}}{\binom{5}{2}} = \frac{2}{5}$$

Si se repite el proceso para $n = 0, 1, 2, \dots, 5$, obtendremos la probabilidad de que resultaran sometidos a la primera inserción, los que han visto sucesivamente ninguna, una y así hasta cinco inserciones. Multiplicando estas probabilidades por el número de individuos que han visto cero, una, dos, hasta cinco inserciones, se obtendrían los individuos que han visto la primera de las inserciones en cada grupo, como ilustra el cuadro 4.5.

Cuadro 4.5. Probabilidades de exposición a la primera inserción.

Frecuencia	Individuos	Probabilidad exposición a la 1ª inserción	Individuos expuestos a la 1ª inserción
0 de 5	50	0	0
1 de 5	20	1/5	4
2 de 5	10	2/5	4
3 de 5	5	3/5	3
4 de 5	5	4/5	4
5 de 5	10	1	10
Total	100		25

Fuente: elaboración propia

Siendo la última columna la cobertura de la primera inserción. De manera análoga a la explicada anteriormente, podemos obtener la cobertura de las dos primeras inserciones a partir de la ley hipergeométrica. Si un individuo ha resultado expuesto a una de las cinco inserciones, la probabilidad de que esa inserción fuera la primera o la segunda, viene dada por la suma de:

$$\binom{1}{1} \binom{4}{0} / \binom{5}{1} + \binom{1}{1} \binom{4}{1} / \binom{5}{1} = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

donde el primer sumando sería la probabilidad de exposición a la primera inserción, y el segundo a la segunda. De este modo, las probabilidades de exposición a los dos primeros números vendría dado por el cuadro 4.6.

Cuadro 4.6. Probabilidades de exposición a las dos primeras inserciones.

Frecuencia	Individuos	Probabilidad exposición a las dos primeras inserciones	Individuos expuestos a a las dos primeras inserciones
0 de 5	50	0	0
1 de 5	20	2/5	8
2 de 5	10	7/10	7
3 de 5	5	9/10	4,5
4 de 5	5	1	5
5 de 5	10	1	10
Total	100		34,5

Fuente: elaboración propia

El proceso descrito puede generalizarse, desembocando en lo que sería la fórmula general de conversión de la distribución de frecuencias en la cobertura de un plan de medios:

$$R_t = \sum_{K=1}^N E_K^N \left[1 - \frac{\binom{N-t}{K}}{\binom{N}{K}} \right] \quad [4.1]$$

donde la notación empleada corresponde a:

N número de inserciones para las cuales está disponible la distribución de contactos.

K número de inserciones a las cuales se ha resultado expuesto.

t número de inserciones de las que se quiere el alcance acumulado .

R_t , alcance acumulado tras t inserciones en un soporte.

E_K^N que es el número de individuos expuestos a K inserciones de las N realizadas.

Debe cumplirse, además que $N \geq t \geq 0$, tomando K los valores $K = 1, 2, \dots, N$ y adoptando la convención de que, cuando $K \geq N - t$, entonces:

$$\binom{N-t}{K} = 0$$

Chandon (1985), diseñó también el procedimiento inverso para pasar de la cobertura de un plan de medios, a su distribución de contactos. Como se vio al desarrollar el modelo de acumulación binomial de Markov (epígrafe 3.1.2.10) (Chandon, 1985), este procedimiento es necesario ante la incapacidad de algunos modelos de calcular ésta variable. Aunque esta no es una limitación procedente de las características de la base de datos, se ha creído oportuno su desarrollo llegados a este punto, debido a su paralelismo con el procedimiento inverso que acabamos de ilustrar.

Supóngase ahora, que se dispone del alcance acumulado generado por un plan de cinco inserciones, tal y como se recoge en el cuadro 4.7., y asumamos que las cinco inserciones son equiprobables.

Cuadro 4.7. Ejemplo de alcance acumulado

Inserciones	1	2	3	4	5	Total Población
Alcance acumulado	250	345	410	460	500	1.000

Fuente: elaboración propia

En primer lugar, se observa fácilmente que el total de individuos no expuestos (E_0^5), puede obtenerse restando al total de la población (POB), aquellos

sujetos que han sido alcanzados al menos una vez (cobertura tras las cinco inserciones realizadas, R_5):

$$E_0^5 = POB - R_5 = 1000 - 500 = 500$$

y, mediante [4.1], se puede expresar el resto de elementos de la distribución de contactos ($E_1^5 \dots E_5^5$) en función de las audiencias acumuladas:

$$\begin{aligned} R_5 &= E_1^5 + E_2^5 + E_3^5 + E_4^5 + E_5^5 = 500 \\ R_4 &= \frac{4}{5}E_1^5 + E_2^5 + E_3^5 + E_4^5 + E_5^5 = 460 \\ R_3 &= \frac{3}{5}E_1^5 + \frac{9}{10}E_2^5 + E_3^5 + E_4^5 + E_5^5 = 410 \\ R_2 &= \frac{2}{5}E_1^5 + \frac{7}{10}E_2^5 + \frac{9}{10}E_3^5 + E_4^5 + E_5^5 = 345 \\ R_1 &= \frac{1}{5}E_1^5 + \frac{2}{5}E_2^5 + \frac{3}{5}E_3^5 + \frac{4}{5}E_4^5 + E_5^5 = 250 \end{aligned}$$

Las expresiones anteriores, no son sino un sistema de cinco ecuaciones con cinco incógnitas, que puede escribirse matricialmente como sigue:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ \frac{4}{5} & 1 & 1 & 1 & 1 \\ \frac{3}{5} & \frac{9}{10} & 1 & 1 & 1 \\ \frac{2}{5} & \frac{7}{10} & \frac{9}{10} & 1 & 1 \\ \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} & 1 \end{bmatrix} \begin{bmatrix} E_1^5 \\ E_2^5 \\ E_3^5 \\ E_4^5 \\ E_5^5 \end{bmatrix} = \begin{bmatrix} 500 \\ 460 \\ 410 \\ 345 \\ 250 \end{bmatrix}$$

que utilizando notación compacta, puede escribirse como:

$$A \cdot E = R$$

donde A es la matriz de coeficientes, E el vector desconocido de la distribución de contactos y R el vector del alcance acumulado. Premultiplicando por la inversa de la matriz de coeficientes, obtendremos la distribución de contactos buscada:

$$E = A^{-1} \cdot R$$

que en el ejemplo que estamos siguiendo, llevaría a:

$$\begin{bmatrix} 5 & -5 & 0 & 0 & 0 \\ -10 & 20 & -10 & 0 & 0 \\ 10 & -30 & 30 & -10 & 0 \\ -5 & 20 & -30 & 20 & -5 \\ 1 & -5 & 10 & -10 & 5 \end{bmatrix} \begin{bmatrix} 500 \\ 460 \\ 410 \\ 345 \\ 250 \end{bmatrix} = \begin{bmatrix} 200 \\ 100 \\ 50 \\ 50 \\ 100 \end{bmatrix}$$

Chandon (1985; pág. 119), propone que el mismo resultado puede escribirse de manera algebraica de la siguiente forma:

$$E_K^N = (-1)^K \binom{N}{K} \sum_{t=0}^K (-1)^{t+1} \binom{K}{t} R_{N-t}$$

En todo caso, nosotros proponemos una vía alternativa de conversión que funciona adecuadamente en cualquiera de las dos direcciones, tanto desde la distribución de contactos hacia la cobertura, como para derivar la distribución de contactos de esta. Así, manteniendo la notación anterior, llamaremos E al vector que recoge la distribución de contactos y R al vector que contiene la cobertura. Si definimos la matriz de coeficientes hipergeométricos H como aquella cuyos elementos vienen definidos como sigue:

$$h_{ij} = \left[1 - \frac{\binom{i-1}{j}}{\binom{N}{j}} \right] \quad [4.2]$$

siendo \mathbf{H} una matriz $N \times N$, donde N es el número de inserciones que estemos considerando, i es la fila que ocupa el elemento h_{ij} y j la columna. Con la condición adicional de que si $i - 1 > j \rightarrow h_{ij} = 1$. En este caso, hemos comprobado que el producto matricial:

$$\mathbf{H} \cdot \mathbf{E} = \mathbf{R}$$

funciona bien en los dos sentidos, es decir, también para:

$$\mathbf{E} = \mathbf{H}^{-1} \cdot \mathbf{R}$$

como se ilustra en los dos ejemplos planteados en esta exposición. La matriz \mathbf{H} , depende exclusivamente del número de inserciones N que se hayan realizado, que son, en nuestro caso, $N = 5$ para los dos ejemplos. Por tanto, en ambos casos, calculando cada uno de sus elementos según la expresión [4.2], obtenemos:

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0,8 & 1 & 1 & 1 & 1 \\ 0,6 & 0,9 & 1 & 1 & 1 \\ 0,4 & 0,7 & 0,9 & 1 & 1 \\ 0,2 & 0,4 & 0,6 & 0,8 & 1 \end{bmatrix}$$

Aplicándola al primer caso, es decir, a la conversión de la distribución de contactos \mathbf{E} del cuadro 4.4, en cobertura, se llegaría a:

$$\mathbf{R} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0,8 & 1 & 1 & 1 & 1 \\ 0,6 & 0,9 & 1 & 1 & 1 \\ 0,4 & 0,7 & 0,9 & 1 & 1 \\ 0,2 & 0,4 & 0,6 & 0,8 & 1 \end{bmatrix} \begin{bmatrix} 20 \\ 10 \\ 5 \\ 5 \\ 10 \end{bmatrix} = \begin{bmatrix} 50 \\ 46 \\ 41 \\ 34,5 \\ 25 \end{bmatrix}$$

que son los valores R_1, \dots, R_5 del alcance acumulado antes obtenidos.

De la misma manera, para el segundo caso -convertir los alcances acumulados del cuadro 4.7, en la distribución de contactos- basta aplicar:

$$E = H^{-1} \cdot R$$

es decir,

$$\begin{bmatrix} 5 & -5 & 0 & 0 & 0 \\ -10 & 20 & -10 & 0 & 0 \\ 10 & -30 & 30 & -10 & 0 \\ -5 & 20 & -30 & 20 & -5 \\ 1 & -5 & 10 & -10 & 5 \end{bmatrix} \begin{bmatrix} 500 \\ 460 \\ 410 \\ 345 \\ 250 \end{bmatrix} = \begin{bmatrix} 200 \\ 100 \\ 50 \\ 50 \\ 100 \end{bmatrix}$$

que, como puede comprobarse, genera idénticos resultados a los obtenidos por el procedimiento propuesto por Chandon (1985)

El único aspecto importante que nos resta por comentar, referente a la base de datos empleada en esta tesis doctoral, es que se ha empleado el fichero de microdatos del Estudio General de Medios, correspondiente al tercer acumulado móvil de 1996. Aunque es opinión general que el funcionamiento de los modelos no se ve afectado por el hecho de que los datos sean más o menos actuales (Leckenby y Kim, 1993), como demuestra el hecho del desfase entre el año de realización del estudio y la fecha de los datos de la mayoría de trabajos que recoge el cuadro 4.1, se ha preferido que fueran recientes con el único ánimo de que, al analizar los soportes que se incluyen en cada plan, no aparezcan demasiados soportes desaparecidos, objetivo que, debido al entorno turbulento en que se desenvuelven actualmente los medios, no es demasiado fácil de lograr.

4.2.2 Diseño de los planes de medios

En relación con el diseño de los planes de medios que se van a utilizar para evaluar el funcionamiento de los distintos modelos de exposición, la primera decisión importante a adoptar, es en qué medios van a evaluarse los modelos. Como señalan Leckenby y Kim (1992), los trabajos sobre estimación de la capacidad predictiva de los modelos de exposición, nacen del interés de los propietarios de cabeceras de revistas por conocer exactamente cuántos sujetos leían sus publicaciones. Este hecho provocó que la mayoría de los estudios de audiencia se centraran históricamente en este medio, como ocurre inicialmente con los trabajos de Politz (1950; 1953; 1958), o los estudios de la W.R. Simmons and Associates, más tarde la Simmons Market Research Bureau (SMRB).

La disponibilidad de los datos, y el escaso desarrollo inicial de otros medios, provocó que la mayoría de los modelos que se crean inicialmente, se centren en las revistas: Agostini (1961), Kwerel (1964), Metheringham (1964), Chandon (1985). Hay que esperar hasta la década de los setenta para que aparezcan algunos modelos desarrollados específicamente para el medio televisión: Headen, Klompmaker y Rust (1979), Headen, Klompmaker y Teel (1977); Leckenby y Rice (1985); Liebman y Lee (1974), Rice (1985) o Leckenby y Rice (1986). Sin embargo, las características específicas que adopta la medición de las audiencias en televisión, han frenado la creación de modelos de exposición en este campo. Aunque algunos estudios de audiencias basados en la encuesta, como el EGM, incorporan a la televisión en sus cuestionarios, la gran variación que se da en las audiencias día a día, fruto entre otras cosas del fenómeno de la contraprogramación, hace que la planificación en este medio se base fundamentalmente en los datos proporcionados por el audímetro que ofrece su información diariamente y que, por ello, refleja inmediatamente estas variaciones. Dado que esta técnica permite “contar” a los individuos según qué programa estén viendo, así como constatar sus migraciones entre soportes, los

procedimientos simplificadores que suponen en sí mismos los modelos de exposición devienen en innecesarios.

Este hecho ha provocado que el desarrollo de nuevos modelos se haya centrado en las dos últimas décadas, al medio revistas: Leckenby y Kishi (1981), Leckenby y Boyd (1984a), Boyd (1985), Leckenby y Kim (1994) y Danaher (1988b; 1989; 1991). Estas consideraciones nos han llevado a evaluar los modelos de exposición en dos medios. Por un lado, el medio revistas, para el que fueron diseñados la mayoría de ellos y en el cual, se han centrado los trabajos empíricos tendentes a la estimación de su precisión recogidos en el cuadro 4.1. Por otro lado, se ha creído conveniente diseñar también planes de medios en el medio diarios en este proceso de evaluación por diversas razones. La primera de ellas es que este medio, importante en términos de inversión publicitaria⁵, no ha sido utilizado para evaluar los modelos de exposición salvo en trabajos muy puntuales (Kurzbard y Soley, 1985). Y, en principio, creemos que este olvido no se debe tanto a que los creadores de los modelos considerasen que no eran adecuados para este medio, como a que los evaluadores se han dejado llevar por la inercia de aquellos trabajos iniciales muy constreñidos a las revistas debido a la limitación de datos tabulados existentes. Lógicamente, al ser ésta una variación sobre los diseños experimentales de los trabajos realizados hasta el momento, se ha planteado como una cuestión a investigar si el funcionamiento de estos modelos, diseñados en principio para el medio revistas, es o no equiparable a los resultados que se obtienen en el medio periódicos.

La segunda razón que justifica la selección de medios realizada, es que, utilizar medios que no han sido empleados por la literatura sobre evaluación de los modelos de exposición, y que vienen recogidos en el cuadro 4.1, implicaría agravar

⁵ En 1995, según Infoadex, la inversión publicitaria en diarios absorbió el 31,6% de la inversión total en medios convencionales, mientras que la televisión recibió el 37,4% de esta inversión y las revistas el 13,2%.

un problema en el que se entrará con detalle en el epígrafe siguiente: la endémica falta de comparabilidad de resultados propia de esta línea de investigación.

La segunda decisión que debe adoptarse en el diseño de los planes de medios que han de servir para evaluar los modelos, se descompone en dos aspectos: cuántos planes deben construirse y cómo debe repartirse esa cantidad entre todos los planes posibles.

Como se observa en el cuadro 4.8, la diferencia en el número de planes que se utiliza para evaluar los modelos es muy grande. Así, varían desde unas pocas decenas en los estudios de Leckenby y Kishi (1981; 1982b), Leckenby y Rice (1986a) y Leckenby y Hong (1997), hasta cifras superiores al millar en los de Leckenby y Boyd (1984a; 1985). Algunos trabajos que evalúan modelos en el medio televisión han llegado a utilizar más de dos mil planes distintos (Headen, Klompmaker y Teel, 1976). Ante esta diversidad, parece lógico intentar establecer de una manera coherente los límites superior e inferior entre los que debería moverse un tamaño muestral razonable. En este sentido, Iacobucci (1994) recomienda un tamaño mínimo de 30 casos (planes de medios) dado que, de no ser así, los datos tendrían insuficiente poder para detectar diferencias significativas (Cohen, 1977), pudiéndose afirmar que dos modelos gozan de la misma capacidad para estimar la cobertura y la distribución de contactos, cuando puede no ser así. Por el contrario, Greene y Stock (1967), señalan que tamaños muestrales muy elevados, provocan que los *test* basados en el estadístico *F* den siempre como significativas todas las diferencias de error. Por este motivo, Leckenby y Kim (1992; pág. 105), afirman que “un tamaño muestral entre 60 y 100 planes de medios seleccionados aleatoriamente, serán suficientes para una aplicación inteligente de la inferencia estadística”.

Siguiendo estas directrices, en este trabajo se ha optado por ajustarse al límite superior del intervalo considerado razonable por Leckenby y Kim (1992) y se ha trabajado con 100 planes de medios considerando que, de este modo, el compromiso

entre capacidad de inferencia y detección adecuada de las distintas capacidades predictivas de los modelos queda garantizado.

Cuadro 4.8. Tamaños muestrales y procedimiento de selección de los planes, de los principales estudios empíricos sobre evaluación de los modelos de exposición.

Estudio	Número de planes	Método de selección de los planes
Leckenby y Kishi (1981)	45	Muestreo estratificado. No consta si aleatorio.
Kishi (1982)	515	Muestreo estratificado. No consta si aleatorio.
Leckenby y Kishi (1982b)	45	Muestreo estratificado. No consta si aleatorio.
Leckenby y Kishi (1982c)	515	Muestreo aleatorio estratificado
Leckenby y Boyd (1984a)	1.033	Muestreo aleatorio estratificado
Leckenby y Kishi (1984)	515	Muestreo aleatorio estratificado
Leckenby y Rice (1985a)	150	Muestreo aleatorio
Chandon (1985)	57	No consta
Boyd (1985)	660	Muestreo aleatorio estratificado
Leckenby y Boyd (1985)	1.033	Muestreo aleatorio estratificado
Leckenby y Rice (1985b)	200	Muestreo aleatorio estratificado
Leckenby y Rice (1986a)	44	Muestreo aleatorio estratificado
Leckenby y Rice (1986b)	242	Muestreo aleatorio estratificado
Danaher (1988b)	600	Estratificado por audiencia del soporte.
Danaher (1989)	480	Muestreo aleatorio estratificado
Leckenby y Ju (1989)	515	Muestreo aleatorio estratificado
Danaher (1991)	840	Muestreo estratificado. No consta si aleatorio.
Leckenby y Kim (1993)	479	Muestreo aleatorio estratificado
Leckenby y Kim (1994b)	479	Muestreo aleatorio estratificado
Aldás (1995)	10	Muestreo estratificado. No aleatorio.
Redondo (1996)	400	No consta
Leckenby y Hong (1997)	28	No consta

Fuente: Elaboración propia

Una salvedad debe realizarse, sin embargo, a esta decisión. Los modelos de acumulación de audiencias, se caracterizan por evaluar planes que están formados por n inserciones en un único soporte. El diseño de cien planes de medios para este tipo de modelos implicaría, por tanto, la consideración de cien soportes diferentes. Este hecho llevaría a tener que incluir prácticamente todos los diarios para poder alcanzar este número. Como los planes diseñados tienen ámbito nacional, y cada plan está formado por un solo soporte, carecería de sentido incluir a los diarios regionales

para alcanzar este número. Por este motivo para este bloque de modelos, el número de planes diseñados es de 49.

Determinado el tamaño muestral adecuado, resta por establecer el proceso de selección de ese número de planes de medios entre todos los posibles. La revisión de la literatura y, en particular, las aplicaciones empíricas sintetizadas en el cuadro 4.4, nos apuntan la respuesta. Casi todos los trabajos optan por un muestreo aleatorio estratificado, procedimiento que ilustraremos a continuación. La primera etapa, la estratificación, consiste en:

- a. Agrupar a los soportes en grupos homogéneos respecto a su contenido editorial, intentando que estas agrupaciones reflejen algunas de las propiedades que utilizarían en la práctica los planificadores de medios para su selección: un mismo público objetivo (Kishi, 1982)
- b. Por otra parte, se pretende evitar que se llegue a planes de medios carentes de sentido al estar formados por soportes de características totalmente divergentes que, probablemente nunca coincidirían en planes de medios reales (Leckenby y Boyd, 1984a).

La clasificación de los soportes empleada en la mayoría de estos estudios (Leckenby y Kishi, 1982b;1982c;1984; Leckenby y Boyd, 1984a;1985; Leckenby y Kishi, 1984; Leckenby y Rice, 1985a; Leckenby y Ju, 1989) es la formada por revistas del tipo (1) Negocios o finanzas, (2) Noticias, (3) Hombres, (4) Mujeres, (5) Salud y Hogar, (6) Motor, (7) Caza, Pesca y Deportes y (8) Temas específicos.

Dado que no existen trabajos en nuestro país que lleven a cabo una clasificación de los soportes, con un fin análogo: construir planes de medios para la evaluación del funcionamiento de distintos modelos, se ha creído conveniente, con el

fin de soportar en la literatura la clasificación empleada, basar nuestra clasificación en la realizada en los trabajos reseñados, pero adaptándola a las realidades específicas del mercado español de medios, con los criterios que a continuación se expondrán.

- a. Se han fusionado los niveles de *negocios y finanzas* con el de *noticias*, debido a que, en el caso español, el número de soportes del primer tipo es demasiado reducido como para conformar un grupo en sí mismo, al estar formado únicamente por Actualidad Económica y Dinero⁶.
- b. Se ha eliminado la categoría de *hombres* que, para los autores de los trabajos reseñados, incluye en el caso norteamericano a numerosas revistas como Playboy, Oui o Esquire. En España, el EGM sólo incluía a Playboy y Penthouse y no parecía procedente, por las diferencias evidentes del contenido editorial, fusionar este nivel con el de noticias o negocios y finanzas, aunque su público sea también, mayoritariamente hombres.
- c. La categoría de *mujeres* de los trabajos norteamericanos, que incluyen a revistas como Ladies' Home Journal, Woman's Day o Better Home & Gardens, se ha desdoblado en dos para el caso español, atendiendo a la existencia de dos grupos de revistas que, tienen el tamaño suficiente en número de soportes para ser considerados independientemente, y también resultan claros sus contenidos editoriales muy diferenciados. Estos grupos son el que ha sido etiquetado como *mujer y moda*, que sería el equivalente al norteamericano y el de *revistas del corazón*, inexistente en aquella clasificación.
- d. En los trabajos reseñados, el grupo de *noticias*, incluye también a las guías de televisión que, desde nuestra perspectiva, están muy alejados tanto en su

⁶ No se olvide que todos los comentarios a los soportes existentes en el mercado, hacen referencia al año 1996, dado que la base de datos utilizada es la correspondiente al tercer acumulado móvil de este año.

contenido editorial, como en su público objetivo de soportes como Tiempo o Tribuna. Dado que, como puede comprobarse en el cuadro 4.9, aunque reducidas en número, tienen audiencias elevadas, han constituido en nuestra clasificación un grupo propio bajo el nombre de *televisión*.

- e. El grupo de salud y hogar, que incluía en el caso norteamericano a revistas como House Beautiful, Family Health y Apartment Life, se ha mantenido en nuestro trabajo bajo el nombre de *decoración*, aunque trasladando su componente de salud, minoritario en el caso español y, desde luego, poco cercano editorialmente, al grupo de *temas específicos* que, para nosotros, agrupa todas aquellas revistas cuyo único nexo es que su objeto son aficiones concretas, como el motor, la caza, los viajes, etc.
- f. Los *diarios*, cuya inclusión ya se ha justificado anteriormente, forman su propio grupo, y en él se han incorporado las principales cabeceras de ámbito nacional y, las que sin serlo, tienen audiencias elevadas, como La Vanguardia, El Periódico o Correo Español, así como todas las deportivas. El motivo de no considerar diarios regionales, ya se apuntó con anterioridad. Se pretende evitar que en el proceso de selección aleatoria de planes, se consideraran en un mismo plan de medios dos diarios regionales que rara vez coincidirían y, sobre todo, que aportarían muy poco al estudio del fenómeno de la duplicación de audiencias. Finalmente, debe señalarse que ABC no ha sido considerado porque, en 1996, estaba excluido del EGM.

En todos los grupos, se han incluido a todos los soportes existentes a la hora de formar los planes de medios, con la única salvedad del bloque de *temas específicos*, dado que, al contener tal cantidad de revistas, complicaría en exceso el proceso de aleatorización que inmediatamente se describirá. En este grupo se seleccionaron solamente los principales soportes, adoptando para ello el criterio de

incluir a los que contasen con audiencias superiores a los trescientos mil individuos. El cuadro 4.9, sintetiza las categorías construidas y los soportes considerados en cada una de ellas.

Cuadro 4.9. Clasificación de los soportes utilizados.

Categoría	Soporte	Audiencia <i>miles de individuos</i>
Diarios	Marca	2.620,9
	El País	1.422,5
	El Mundo	1.007,3
	El Periódico	959,2
	La Vanguardia	705,5
	Correo Español	635,8
	As	514,1
	Sport	438,3
	Mundo Deportivo	383,9
	Diario 16	260,2
Noticias y Económicas	Interviú	1.202,8
	Tiempo	479,6
	Época	205,9
	Tribuna	201,3
	Actualidad Económica	80,5
Corazón	Dinero	38,4
	Hola	3.179,0
	Pronto	3.084,7
	Lectura	2.141,4
	Diez Minutos	1.982,0
	Semana	1.930,6
	Mía	1.391,1
	Nuevo Vale	773,8
Nuevo Plus	352,5	

Fuente: Elaboración propia en base a EGM: 3º acumulado 1996.

Cuadro 4.9. Clasificación de los soportes utilizados.
Continuación

Categoría	Soporte	Audiencia <i>miles de individuos</i>
Mujer y Moda	Clara	810,3
	Cosmopolitan	746,8
	Burda	720,0
	Ragazza	655,8
	Woman	587,5
	Elle	565,0
	Patrones	464,7
	Prima	355,3
	Dunia	292,9
	Vogue	290,9
	Telva	270,4
Televisión	Teleindiscreta	1.237,7
	Supertele	1.079,4
	TP	1.058,8
	Telenovela	300,6
Decoración	El Mueble	1.085,8
	Nuevo Estilo	814,3
	Mi Casa	734,1
	Labores del Hogar	595,0
	Casa Jardín	301,0
Temas Específicos	Muy Interesante	2.555,2
	Fotogramas	842,8
	Natura	425,1
	Geo	346,6
	Solo Moto	340,7

Fuente: Elaboración propia en base a EGM: 3° acumulado 1996.

Una vez distribuidos los soportes dentro de cada grupo, y decidido el número de planes totales que se van a construir, se reparten proporcionalmente entre cada grupo atendiendo al número de soportes que contiene cada uno de ellos. Sólo resta elegir qué planes concretos se escogerán dentro de todas las combinaciones posibles. Esto se logra listando las combinaciones, asignando un número de orden y seleccionando mediante una tabla de números aleatorios: este procedimiento puede denominarse la fase de aleatorización del procedimiento de muestreo.

En este trabajo se ha mantenido esencialmente este proceder introduciendo una modificación provocada por el tipo de modelos que se evalúan y que ya se apuntó con anterioridad. Así, los estudios que se han venido referenciando, se

centran únicamente en la evaluación de los modelos de audiencia neta acumulada, mientras que en esta tesis doctoral se evalúan también los modelos de acumulación. En ellos, los planes están formados por diversas inserciones en un único soporte, de tal forma que ha sido necesario incluir todos los soportes sin selección aleatoria alguna, para tener el tamaño muestral suficiente para asegurar la potencia de las pruebas estadísticas utilizadas, que se describirán detalladamente en el epígrafe siguiente.

Finalmente, los cuadros 4.10, 4.11 y 4.12, recogen las características fundamentales de los planes de medios que se han construido para evaluar los modelos de exposición, en términos de: número de planes, número de inserciones, cobertura, audiencia bruta, y suma de las audiencias acumuladas y duplicadas.

Cuadro 4.10. Características de los planes de medios.
Modelos de acumulación de audiencias

Características de los planes	Total	Media	Desviación típica	Mínimo	Máximo
Número de planes	49				
Número de inserciones		5,47	1,19	4	7
Cobertura		7,10	7,01	0,80	30,00
GRPs		21,8	19,5	2	84
Suma audiencias acumuladas		14,65	13,58	1,02	67,83

Fuente: elaboración propia

Cuadro 4.11. Características de los planes de medios.
Modelos de duplicación de audiencias

Características de los planes	Total	Media	Desviación típica	Mínimo	Máximo
Número de planes	99				
Número de soportes		3,51	0,68	3	5
Cobertura		15,59	11,01	4,02	45,10
GRPs		21,34	19,11	4,9	83,5
Suma audiencias duplicadas		7,02	13,10	0,00	67,01

Fuente: elaboración propia

Cuadro 4.12. Características de los planes de medios.
Modelos de audiencia neta acumulada

Características de los planes	Total	Media	Desviación típica	Mínimo	Máximo
Número de planes	99				
Número de inserciones por soporte		5,47	1,19	4	7
Número de soportes		3,51	0,68	3	5
Cobertura		22,97	13,12	6,90	52,00
GRPs		101,63	70,82	29,88	334,36
Suma audiencias acumuladas		56,91	39,53	10,01	158,10
Suma audiencias duplicadas		143,32	201,11	7,00	1.080,00

Fuente: elaboración propia

En ellos, puede comprobarse, como se justificó con anterioridad al describir el diseño experimental, que las inserciones están acotadas entre 4 y 7 por soporte. Este hecho se traduce en que, cuando se trata de planes de medios diseñados para evaluar los modelos de acumulación de audiencias, los GRPs alcanzados son muy bajos (una media de 21,8), dado que es difícil incrementar este valor cuando sólo se puede incluir un soporte y el número de inserciones está acotado superiormente en siete. Sin embargo, consideramos que este hecho se traduce más en niveles bajos de frecuencia de contacto que a escasa cobertura, dado que llegar al 7% de la población con este tipo de planes, parecen cifras razonables.

Lo mismo puede concluirse en los planes destinados a evaluar modelos de duplicación de audiencias y que, por ello, sólo pueden incluir dos soportes. De nuevo los GRPs son bajos (media del 21,34%), pero se alcanzan coberturas razonables: 15,59%, lo que nos hace concluir que la principal limitación de los planes de medios que se han construido, radica en el escaso nivel logrado de frecuencia hecho, por otra parte, difícilmente soslayable con las limitaciones que la base de datos imponía a estos tipos de planes y que se han remarcado ampliamente con anterioridad.

Los planes construidos para evaluar los modelos de audiencia neta acumulada parecen, por contra, muy próximos a lo que podrían ser campañas reales. Así, se

consigue una media de más de 100 GRPs, con puntas de 334, llegando la cobertura a niveles superiores al 20% y, por ello, con OTSs en el entorno de cinco. De esta forma, entendemos el diseño que se ha adoptado es adecuado en el sentido en que los planes construidos se aproximan lo bastante a la planificación real de medios y que, cuando esto no es así, la limitación no procede tanto del diseño, como de que se intenta evaluar modelos que, si bien tienen gran interés académico, no se corresponden con procedimientos habituales en la práctica.

4.2.3 Definición del error

Uno de los hechos que, según Leckenby y Boyd (1990) o Leckenby y Kim (1992), más contribuyen a entorpecer la confrontación de los resultados obtenidos por las distintas investigaciones sobre la evaluación de los modelos de exposición, es la gran diversidad de definiciones de error empleadas en la comparación de los datos estimados y reales de la cobertura y la distribución de contactos (véase cuadro 4.13). Resulta evidente que no es operativo utilizar todas las medidas posibles de error, entre otras cosas, porque no todas son acertadas, como posteriormente intentaremos demostrar. Así, Leckenby y Kim, (1992; pág. 105) señalan, en base a los resultados de sus trabajos, que “un estudio nunca deberá utilizar todas las medidas de error posibles para evaluar un modelo, un estudio debe utilizar en torno a cuatro de ellas”, añadiendo a continuación que “sería extremadamente útil que se pudiera llegar a un acuerdo entre los investigadores, para seleccionar dos o tres medidas del error para la distribución de contactos y dos o tres para la cobertura”. Dado que el acuerdo todavía no se ha producido, en este epígrafe explicaremos todas las definiciones de error que se han empleado en la literatura, y que vienen sintetizadas en el cuadro 4.13 para, a continuación, justificar la selección que se ha realizado en este trabajo.

Cuadro 4.13. Definición de error en los principales estudios empíricos sobre evaluación de los modelos de exposición.

Estudio	Definición de error	
	Cobertura	Distribución de contactos
Leckenby y Kishi (1981)	AER	APE; APEj
Kishi (1982)	AER	APE; $\pm 5\%$, χ^2
Leckenby y Kishi (1982b)	--	APE, APEj, $\pm 5\%$
Leckenby y Kishi (1982c)	AER, RER	APE; APEj
Leckenby y Boyd (1984a)	AER	AED; ALE
Leckenby y Kishi (1984)	AER	APE
Leckenby y Rice (1985a)	AER	AED
Chandon (1985)	AER	$\pm 5\%$, χ^2
Boyd (1985)	AER	$\pm 5\%$; AED
Leckenby y Boyd (1985)	AER	--
Leckenby y Rice (1985b)	AER	AED
Leckenby y Rice (1986a)	AER	AED; $\pm 5\%$
Leckenby y Rice (1986b)	AER; $\pm 5\%$	AED
Danaher (1988b)	RER; AEGRP	MSE; EPOR
Danaher (1989)	RER; AEGRP	MSE; EPOR
Leckenby y Ju (1989)	AER; $\pm 5\%$	AED; MSE; $\pm 5\%$
Danaher (1991)	RER; AEGRP	MSE; EPOR
Leckenby y Kim (1993)	AER	APE
Leckenby y Kim (1994b)	AER	APE
Aldás (1995)	AER	MSEC
Redondo (1996)	AER	--
Leckenby y Hong (1997)	AER	APE

Fuente: *Elaboración propia*

4.2.3.1 Medidas de error en la estimación de la cobertura.

Como se recoge en cuadro 4.13, las definiciones del error cometido en la estimación de la cobertura utilizadas con más frecuencia por la literatura, son las correspondientes al *error promedio en la cobertura* (AER⁷) y el *error relativo en la cobertura* (RER):

⁷ Para facilitar la revisión por parte del lector de los trabajos originales, se ha optado por no adaptar al español las siglas de las definiciones de error.

$$AER = \frac{\sum_{j=1}^N |\hat{R}_j - R_j|}{N} \quad [4.3]$$

$$RER = \frac{\sum_{j=1}^N \frac{|\hat{R}_j - R_j|}{R_j}}{N} \quad [4.4]$$

Donde la notación empleada se corresponde con:

- R_j Cobertura real del plan de medios j
- \hat{R}_j Cobertura estimada del plan de medios j
- N Número de planes de medios evaluados

Como se comprueba al analizarla, la expresión [4.3] es, simplemente, la media para todos los planes de medios, de la diferencia en valor absoluto entre los valores reales y estimados de la cobertura. La expresión [4.4] es análoga, con la salvedad de que la suma de las diferencias son relativizadas por la cobertura real de cada plan de medios, antes de obtenerse su media aritmética.

Danaher (1988b; 1989; 1991) introduce lo que llama *error absoluto en la estimación de los GRP's* (AEGRP), que define como:

$$AEGRP = \frac{\sum_{j=1}^N \left| \sum_{i=1}^{k_j} i(O_{ij} - E_{ij}) \right|}{N} \quad [4.5]$$

donde toda la notación ya ha sido empleada, salvo O_{ij} que es el porcentaje real del público objetivo que ha resultado expuesto a i inserciones de las k_j que se han realizado en total en el plan j . E_{ij} es el porcentaje estimado por el modelo de ese mismo concepto. Cuando se analiza detenidamente la expresión [4.5] se entiende el nombre dado por el autor a esa formulación del error. Cuando a cada nivel de

frecuencia, el porcentaje de la población objetivo, ya sea este observado o estimado, se multiplica por el número de contactos que ha recibido, estamos ante la definición *Rating Point*, al sumar este dato para todos los niveles de contacto, se obtienen los *puntos de rating brutos* o GRP's. La expresión [4.5], se limita a calcular la media aritmética de las diferencias entre los valores estimados y observados de este concepto para todos los planes de medios que sirven para evaluar el modelo.

La única definición de error en la estimación de la cobertura que nos resta por definir es la que hemos denotado como $\pm 5\%$ sugerida en el ámbito de la cobertura por Leckany y Rice (1986b) y Leckenby y Ju (1989) y que se corresponde con el *número de estimaciones correctas de la cobertura, que se dan en un intervalo de $\pm 5\%$ del valor real*. Así, cada estimación se clasifica como correcta o incorrecta de acuerdo con el siguiente criterio:

$0,95 \cdot R_j \leq \hat{R}_j \leq 1,05 \cdot R_j$	Estimación correcta
$\hat{R}_j > 1,05 \cdot R_j$	Estimación incorrecta, sobrestimación
$\hat{R}_j < 0,95 \cdot R_j$	Estimación incorrecta, subestimación

y la cantidad de estimaciones correctas calculadas para un modelo en concreto, se convierte en un indicador de su precisión.

De las cuatro medidas señaladas, dos de ellas presentan, desde nuestra perspectiva, serias limitaciones que nos ha llevado a renunciar a su utilización. El número de estimaciones correctas dentro de un 5% del valor real plantea el problema de que ignora qué ocurre con aquellas estimaciones que quedan fuera del intervalo. Así, dos modelos con la misma cantidad de estimaciones correctas, serían siempre considerados como igualmente satisfactorios, cuando cabe la posibilidad de que uno de ellos tenga las estimaciones clasificadas como incorrectas muy concentradas en torno a los límites del intervalo, y el otro modelo no, lo que, evidentemente,

convertiría a la predicción de la cobertura que realice el primero en más acertada que la del segundo.

La limitación del AEGRP es, también, bastante evidente: permite que se compensen, dentro de cada plan de medios, los errores positivos con los negativos a los diferentes niveles de exposición, hecho que redundaría en una subestimación del error cometido por cada modelo.

Es cierto que ambas limitaciones serían fácilmente corregibles modificando las definiciones originales. Sin embargo, hemos creído más conveniente recurrir a definiciones de error que ya hayan sido ampliamente utilizadas y, por ello, contrastadas en este tipo de trabajos, de tal forma que se puedan comparar nuestros resultados, con los de las mencionadas investigaciones, antes que aportar nuestras propias medidas. Por este hecho, en el presente estudio, el error en la estimación de la cobertura se medirá mediante los valores de AER y RER.

4.2.3.2 *Medidas de error en la estimación de la distribución de contactos.*

La variedad de definiciones de error en la estimación de la distribución de contactos, es mucho mayor en la literatura que las referentes a la cobertura. De la revisión de la literatura efectuada, y cuyos resultados se han sintetizado en el cuadro 4.13, nos encontramos hasta con nueve modos distintos de operativizar esta variable. Como hemos hecho con anterioridad, enunciaremos en primer lugar todas y cada una de las expresiones para, a continuación, seleccionar justificadamente las utilizadas en este trabajo.

Entre las más utilizadas, se encuentra el *error porcentual promedio* (APE) y su variante, el *error porcentual promedio para cada nivel de exposición j* (APE_j),

que han sido empleadas simultáneamente en los trabajos de Leckenby y Kishi (1981), Leckenby y Kishi (1982b) y Leckenby y Kishi (1982c) y cuyas expresiones son las siguientes (la notación es conocida del subapartado anterior):

$$APE = \frac{\sum_{j=1}^N \left[\frac{\sum_{i=0}^{k_j} |O_{ij} - E_{ij}|}{\sum_{i=0}^{k_j} O_{ij}} \right]}{N} \quad [4.6]$$

$$APE_j = \frac{\sum_{j=1}^N \sum_{i=0}^{k_j} \frac{|O_{ij} - E_{ij}|}{O_{ij}}}{N} \quad [4.7]$$

Ambas expresiones son muy parecidas. En la [4.6], tras calcular las diferencias en valor absoluto entre los valores observados y estimados de la distribución de contactos a cada nivel de inserción, éstos se suman para todos ellos, relativizándose por la suma, también para todos los niveles de contacto, de los valores observados. De esta manera se tiene el error porcentual relativo para cada plan y, a continuación, se obtiene la media simple de este valor entre el conjunto de planes utilizados para evaluar el modelo. La expresión del APEj [4.7], procede del mismo modo, con la salvedad de que cada diferencia entre el valor observado y estimado de la distribución de contactos, a cada nivel, se relativiza por el valor observado de la misma. A continuación se agrega este valor para el conjunto de las inserciones y se obtiene, después, su media aritmética.

En algunos trabajos (Leckenby y Boyd, 1984a), se trabaja con el error en términos absolutos y no porcentuales. Así nos encontramos con las definiciones del *error absoluto de la distribución* (AED) y del *error promedio a cada nivel de exposición* (ALE):

$$AED = \frac{\sum_{j=1}^N \sum_{i=0}^{k_j} |O_{ij} - E_{ij}|}{N} \quad [4.8]$$

$$ALE = \frac{\sum_{j=1}^N \frac{\sum_{i=0}^{k_j} |O_{ij} - E_{ij}|}{k_j}}{N} \quad [4.9]$$

La primera de estas expresiones recoge, únicamente, la suma para todos los niveles de exposición, de los valores absolutos de las diferencias entre valores observados y estimados de la distribución de contactos. A continuación se obtiene la media simple de este valor, entre todos los planes de medios evaluados.

La segunda expresión [4.9], relativiza la suma de los valores absolutos de las diferencias, según el tamaño en número de inserciones de cada plan de medios, antes de obtener la media del valor para todos ellos.

En un trabajo anterior (Aldás, 1995), ya propusimos una medida similar a ALE, que presentaba la única diferencia de evitar la compensación de errores positivos y negativos mediante el cuadrado del valor, en lugar de mediante el valor absoluto y que, para mantener el orden de magnitud del error, optaba por obtener su raíz cuadrada antes de proceder a calcular su media para el conjunto de planes de medios, tal y como se observa en [4.10]:

$$MSEC = \frac{\sum_{j=1}^N \sqrt{\frac{\sum_{i=0}^{k_j} (O_{ij} - E_{ij})^2}{k_j}}}{N} \quad [4.10]$$

Kishi (1981), Leckenby y Kishi (1982b), Chandon (1985), Boyd (1985), Leckenby y Rice (1986a) y Leckenby y Ju (1989), utilizan para el calcular el error en la estimación de contactos, un procedimiento similar al descrito anteriormente para la

cobertura: el número de estimaciones correctas dentro de un intervalo de $\pm 5\%$ del valor observado. Una estimación se considerará, respectivamente, correcta o incorrecta si cumple:

$$\begin{aligned} 0,95 \cdot O_{ij} \leq E_{ij} \leq 1,05 \cdot O_{ij} & \text{ Estimación correcta} \\ E_{ij} > 1,05 \cdot O_{ij} & \text{ Estimación incorrecta, sobreestimación} \\ E_{ij} < 0,95 \cdot O_{ij} & \text{ Estimación incorrecta, subestimación} \end{aligned}$$

Kishi (1982) y Chandon (1985) emplean un procedimiento de cálculo del error, que coincide con la definición de una función ji-cuadrado:

$$\chi^2 = \frac{\sum_{j=1}^N \sum_{i=0}^{k_j} \frac{(O_{ij}-E_{ij})^2}{E_{ij}}}{N} \quad [4.11]$$

que, puede comprobarse, es una expresión equivalente a APE_j, con dos grandes diferencias: el procedimiento para evitar que se compensen los errores positivos y negativos, consiste en elevar al cuadrado la diferencia, en lugar de tomar su valor absoluto, y el error se relativiza por el valor estimado en lugar de por el observado.

Danaher (1988b; 1989; 1991), utiliza el *error cuadrático medio* (MSE), que es una medida de error habitualmente utilizado en estadística y, por otra parte, el *error en las probabilidades de exposición del plan de medios* (EPOR), ya propuesto por Liebman y Lee (1974). Sus expresiones son las siguientes:

$$MSE = \frac{\sum_{j=1}^N \frac{1}{k_{j+1}} \sum_{i=0}^{k_j} (E_{ij}-O_{ij})^2}{N} \quad [4.12]$$

$$EPOR = \frac{\sum_{j=1}^N \frac{\sum_{i=1}^{k_j} |O_{ij}-E_{ij}|}{1-O_{0j}}}{N} \quad [4.13]$$

Puede comprobarse que MSE es muy similar a ALE, dado que relativizan el error en función del número de inserciones realizadas en cada soporte, antes de obtener la media para el conjunto de los mismos que integran el plan. La única diferencia entre ambos, radica en que el procedimiento para evitar compensaciones de errores positivos y negativos en MSE, consiste en elevar al cuadrado las diferencias, mientras que en ALE se optaba por obtener su valor absoluto.

Sin embargo, EPOR, es una medida bastante distinta a las expresadas hasta ahora, y merece una atención especial. En primer lugar, no calcula las diferencias entre los valores estimados y observados de la distribución de contactos para todos los valores de la misma, elimina el nivel de cero contactos. Este hecho es importante, porque este valor suele ser el más elevado en todos los planes de medios (siempre suele haber una mayor porcentaje de la población objetivo que no ha recibido contactos, que los que han recibido alguno) y, por ello, da más importancia a la hora de determinar si un modelo es más preciso que otro, a los aciertos en la predicción de cada uno de los niveles de contactos producidos efectivamente, no de los individuos no alcanzados. Debe tenerse en cuenta, para valorar adecuadamente este comentario, que la habilidad del modelo para predecir los individuos no contactados, ya se analizó al calcular el error de cada uno de ellos en la estimación de la cobertura del plan que precisamente, en términos porcentuales, es cien menos el porcentaje de individuos que no han recibido ningún contacto.

A la hora de relativizar el error anterior, lo hace dividiendo por el porcentaje de individuos que ha recibido al menos un contacto ($1 - O_{0j}$), para luego obtener la media del valor para el conjunto de los planes de medios considerados.

A modo de síntesis, podemos señalar que las distintas medidas de error difieren unas de otras en función de:

- a. Procedimiento empleado para evitar que los errores negativos se compensen con los positivos. Así, APE, APEj, AED, ALE y EPOR optan por calcular el valor absoluto de la diferencia, mientras que MSEC, ji-cuadrado y MSE optan por elevar al cuadrado esa diferencia. Dado que no encontramos motivo alguno que nos permita afirmar que un procedimiento es más adecuado que otro, no consideraremos esta característica como primordial para elegir la definición de error a utilizar en nuestro estudio.
- b. Mientras algunas medidas acumulan las diferencias entre valores observados y estimados sin más (AED, ALE, MSEC, MSE), otras optan por relativizar previamente esos valores por: la suma de los valores observados (APE), los valores observados a cada nivel de exposición antes de sumar para todos ellos (APEj) o los valores estimados (ji-cuadrado). Dado que los datos de la distribución de contactos, ya sea la real o la estimada, se ofrecen normalmente en términos de porcentaje de la población objetivo, creemos redundante plantear en términos porcentuales errores calculados sobre cifras que ya lo son. Dado que este hecho añade únicamente tiempo de cálculo, se ha optado por centrarse en aquellas definiciones de error que no realizan esta operación.
- c. La única salvedad a la afirmación anterior la constituye, desde nuestro punto de vista, la relativización que lleva a cabo Danaher (1988b; 1989; 1991) en su definición de EPOR, por los motivos ya señalados en su momento.
- d. Sí que consideramos importante que el número de inserciones que se realiza en cada plan de medios, corrija siempre el error calculado para, de este modo, intentar aislar un factor que

potencialmente, pudiera influir en la precisión de los modelos.
En esta línea, están ALE, MSEC, MSE.

Analizadas en su conjunto todas las consideraciones anteriores, este estudio incluirá como método de cálculo del error en la distribución de contactos, en primer lugar, el método de EPOR, en cuanto que es el único que prioriza en la evaluación del modelo a los individuos que sí han recibido contactos frente al resto de fórmulas de error que, en conjunto, incluye también a la población no contactada, ya considerada en la estimación de la cobertura, como anteriormente hemos indicado. Para la segunda expresión del error, existen dos métodos en igualdad de condiciones: ALE y MSE. Hemos optado por MSE únicamente por ser el error cuadrático medio una medida de error habitualmente empleada en estadística.

El cuadro 4.14, resume los procedimientos de cálculo del error que, definitivamente, se emplearán en esta tesis doctoral, habiendo sido toda la notación empleada, ya explicitada con anterioridad.

4.14. Medidas de error utilizadas en esta tesis doctoral.

Estimación de la cobertura

Estimación de la distribución de contactos

$$AER = \frac{\sum_{j=1}^N |\hat{R}_j - R_j|}{N}$$

$$MSE = \frac{\sum_{j=1}^N \frac{1}{k_{j+1}} \sum_{i=0}^{k_j} (E_{ij} - O_{ij})^2}{N}$$

$$RER = \frac{\sum_{j=1}^N \frac{|\hat{R}_j - R_j|}{R_j}}{N}$$

$$EPOR = \frac{\sum_{j=1}^N \frac{\sum_{i=1}^{k_j} |O_{ij} - E_{ij}|}{1 - O_{oj}}}{N}$$

Fuente: *Elaboración propia*

4.2.4 Factores de error.

Como se ha puesto de manifiesto a lo largo de este capítulo, han sido muy abundantes los trabajos encaminados a evaluar la precisión en las estimaciones de los modelos de exposición, partiendo de diversas definiciones de error en el cálculo de la cobertura y la distribución de contactos. Sin embargo, el resultado de este tipo de análisis, siempre es una jerarquización de los modelos atendiendo al error *promedio* que cometen en el conjunto de los planes de medios, formándose grupos de modelos homogéneos en cuanto que las diferencias del error cometido entre ellos no son estadísticamente significativas.

La cuestión natural que se sigue, es si es posible, de alguna manera, discriminar entre los modelos que pertenecen al mismo grupo, estableciendo algún criterio que permita elegir entre ellos atendiendo, no a su comportamiento global, que es idéntico, sino a determinadas características de los planes cuya cobertura y distribución de contactos se va a estimar con ellos. En este sentido, poco o muy poco se ha investigado acerca de si las propiedades estructurales de los planes de medios, puede o no explicar los niveles de error inherentes a los mismos. Parece poco probable, que un modelo consiga el mismo nivel de precisión sean cuales sean las características internas del plan de medios al que se aplica, por lo tanto parece importante establecer un conjunto de factores que ayuden a determinar bajo qué condiciones determinados modelos funcionarán mejor que otros.

Tan solo dos trabajos han afrontado este problema (Leckenby y Kishi, 1981; 1982b) que, realmente, son dos versiones de la misma investigación. En ambos, se sugiere que son doce los posibles factores que pueden explicar el error cometido por los modelos:

1. Audiencia bruta.

2. Suma de las audiencias acumuladas.
3. Suma de las audiencias duplicadas.
4. Número de soportes empleados.
5. Combinación de soportes grandes, medianos y pequeños.
6. Combinación de soportes grandes y medios.
7. Combinación de soportes grandes y pequeños.
8. Combinación de soportes medios y pequeños.
9. Relación entre la suma de las audiencias duplicadas y la audiencia bruta.
10. Relación entre la suma de las audiencias acumuladas y la audiencia bruta.
11. Suma de las audiencias duplicadas y acumuladas.
12. Relación entre la suma de las audiencias duplicadas y acumuladas y la audiencia bruta.

En esta tesis doctoral se han considerado esos mismos factores, siendo definidos y operativizados tal y como a continuación se describe.

Audiencia Bruta. Es la suma de las audiencias de los soportes que conforman el plan de medios, sin tener en cuenta las acumulaciones y duplicaciones de las audiencias que se producen. Si llamamos n_i al número de inserciones que se realizan en el soporte i y A_i a la audiencia de un número del mismo, y N al número de soportes que constituyen el plan de medios, calcularemos la audiencia bruta como sigue:

$$\bar{A} = \sum_{i=1}^N n_i \cdot A_i \quad [4.14]$$

Suma de las audiencias duplicadas. Es la suma de las duplicaciones entre cada par de soportes formado entre los que conforman el plan de medios. Si llamamos d_{ij} a la audiencia duplicada entre dos soportes cualquiera i y j de los N que componen el plan de medios, definiremos este concepto como:

$$D = \sum_{i=1}^N \sum_{j=1}^N n_i \cdot n_j \cdot d_{ij} \quad i \neq j \quad [4.15]$$

Suma de las audiencias acumuladas. Análogamente al anterior, es la acumulación de audiencia que se produce entre dos inserciones consecutivas en un soporte dado, agregada para el conjunto de soportes que conforman el plan de medios. Si se realizan n_i inserciones en el soporte i , denotaremos d_i a la audiencia duplicada entre dos inserciones consecutivas ese soporte que, además, se considerará la misma entre cualquiera otras dos de las posibles combinaciones que son, en total, el número combinatorio que aparece en la expresión [4.16]. Entonces para el conjunto de los N soportes del plan, la suma de las audiencias acumuladas será:

$$AC = \sum_{i=1}^N \binom{n_i}{2} d_i \quad [4.16]$$

Número de soportes. Dado que Leckenby y Kishi (1981, 1982b) utilizan planes de medios formados exclusivamente por 3 y 4 soportes, introducen esta variable como una ficticia en el procedimiento de regresión que se comentará en el apartado siguiente. Así, adopta el valor 0 cuando el plan contempla cuatro soportes y uno si contemplaba 3. Dado que en nuestro caso, el número de soportes es mucho más variado, se ha incluido éste sin categorizar en niveles mediante variables ficticias, construyendo la variable que se expondrá a continuación para recoger la influencia de los tamaños relativos de los soportes que integran el plan.

Combinaciones de tamaño de los soportes. Leckenby y Kishi (1981; 1982b) clasifican a un soporte como pequeño, cuando su audiencia es inferior al 3,83% de la población objetivo, mediano si está entre ese valor y el 15,34% y grande si supera este porcentaje. Dado que los valores de estos niveles no están justificados y, por ello, pueden ser válidos para el caso norteamericano y carecer de sentido para el español, en nuestro trabajo, se ha optado por dividir a los soportes en terciles atendiendo a su audiencia y, mediante ese procedimiento, clasificarlos como

pequeños, medianos o grandes. Cada plan de medios se hace, a continuación, pertenecer a uno de los cuatro tipos recogidos en los factores 5 a 8, mediante cuatro variables ficticias y, por supuesto, en función del tamaño que tienen los soportes que integran cada plan.

Relación entre la suma de las audiencias duplicadas y la audiencia bruta.

Esta relación se calcula a partir de variables ya definidas como sigue:

$$\frac{D}{A} = \frac{\sum_{i=1}^N \sum_{j=1}^N n_i \cdot n_j \cdot d_{ij}}{\sum_{i=1}^N n_i \cdot A_i} \quad i \neq j \quad [4.17]$$

Relación entre la suma de las audiencias acumuladas y la audiencia bruta.

Análogamente a la expresión anterior:

$$\frac{AC}{A} = \frac{\sum_{i=1}^N \binom{n_i}{2} d_i}{\sum_{i=1}^N n_i \cdot A_i} \quad [4.18]$$

Suma de las audiencias duplicadas y acumuladas. Se deriva directamente de las expresiones [4.15] y [4.16]:

$$D + AC = \sum_{i=1}^N \sum_{j=1}^N n_i \cdot n_j \cdot d_{ij} + \sum_{i=1}^N \binom{n_i}{2} d_i \quad [4.19]$$

Relación entre la suma de las audiencias duplicadas y acumuladas y la audiencia bruta. Se obtiene inmediatamente a partir de [4.19] y [4.14]:

$$\frac{D+AC}{\bar{A}} = \frac{\sum_{i=1}^N \sum_{j=1}^N n_i \cdot n_j \cdot d_{ij} + \sum_{i=1}^N \binom{n_i}{2} d_i}{\sum_{i=1}^N n_i \cdot A_i} \quad [4.20]$$

El análisis que proponen Leckenby y Kishi (1981, 1982b), es meramente exploratorio, en la medida en que no hipotetizan qué influencia puede tener cada una de las variables descritas sobre el error cometido por los modelos, ni en qué sentido se producirá la misma (incrementando o disminuyéndolo). El objetivo de estos autores, es sintetizar en una serie de medidas objetivas las variables que pueden caracterizar a un plan de medios y, a continuación, explorar en qué sentido influyen para establecer un criterio adicional de selección de modelos. Puesto que consideran imposible el establecimiento de *a priori* alguno acerca de esta influencia, tratan de ser exhaustivos al recoger las posibles características del plan, de tal forma que no sólo incluyen medidas del tamaño del mismo, como la audiencia bruta o el número de soportes, o de variables tan ligadas al grado de complejidad de los modelos, como los fenómenos de la acumulación y duplicación de las audiencias, sino que plantean todas las relaciones posibles entre ellos. La filosofía de este proceder, radica en dejar que las técnicas de análisis de datos saquen a la luz qué factor influye y cuál no y con qué intensidad lo hace, siendo la única preocupación del investigador asegurarse de que todos los posibles factores de error, se han recogido como variable independiente.

4.2.5 Técnicas de análisis de datos

4.2.5.1 Evaluación y jerarquización de los modelos de exposición.

De lo comentado en el epígrafe 4.2.3, se desprende que serán más precisos aquellos modelos que cometan menor error en sus estimaciones, medido éste por cualquiera de las expresiones elegidas. Sin embargo, analizando con más

detenimiento el problema, nos encontramos con que estas medidas de error, son siempre una media aritmética del error cometido por el modelo en cada uno de los planes de medios. Si pretendemos inferir una conclusión acertada de esa información, no se puede comparar, sin más, las medias obtenidas, es necesario establecer previamente, si esos valores medios son significativamente distintos entre los diferentes modelos: nos encontramos ante el campo clásico de aplicación del análisis de la varianza de un factor. Si la heterogeneidad de los diseños metodológicos era la nota común entre los trabajos analizados: bases de datos diferentes, distinto número de planes de medios evaluados y diversidad entre las definiciones de error manejadas, todos los estudios del cuadro 4.13, coinciden en utilizar esta técnica de inferencia estadística para el análisis de sus datos.

Dado que, ante el acuerdo generalizado, se impone mantener el mismo procedimiento de análisis de la información descrito, so pena de incurrir en el error de contribuir a la falta de comparabilidad de los resultados que se ha venido criticando, justificaremos brevemente la adecuación de la técnica de análisis señalada a las características de los datos, así como los procedimientos seguidos para constatar la procedencia de su utilización en base a las hipótesis subyacentes en dichas técnicas.

El cuadro 4.15, recoge la información de la que se dispone para el análisis una vez que se ha tabulado la información proporcionada por el EGM: cobertura y distribución de contactos observadas de cada uno de los j planes de medios diseñados, cada uno de los cuales está caracterizado por k contactos posibles correspondientes a n inserciones en los m soportes de cada plan. También, tras estimar estas mismas variables con los modelos objeto de evaluación, se dispone de la cobertura y distribución de contactos estimadas.

Cuadro 4.15. Información disponible para el análisis

Plan de Medios	Cobertura		Distribución de contactos							
	Observada	Estimada	Observada				Estimada			
			0	1	...	k=n.m	0	1	...	k=n.m
1	R_1	\hat{R}_1	O_{01}	O_{11}	...	O_{k1}	E_{01}	E_{11}	...	E_{k1}
2	R_2	\hat{R}_2	O_{02}	O_{12}	...	O_{k2}	E_{02}	E_{12}	...	E_{k2}
3	R_3	\hat{R}_3	O_{03}	O_{13}	...	O_{k3}	E_{03}	E_{13}	...	E_{k3}
...
j	R_j	\hat{R}_j	O_{0j}	O_{1j}	...	O_{kj}	E_{0j}	E_{1j}	...	E_{kj}
...
N	R_N	\hat{R}_N	O_{0N}	O_{1N}	...	O_{kN}	E_{0N}	E_{1N}	...	E_{kN}

Fuente: Elaboración propia

Las expresiones del error, descritas con anterioridad en el cuadro 4.14, se pueden descomponer todas ellas en dos partes. La primera es una medida del error de estimación cometido en cada plan de medios, que denotaremos respectivamente AER_j , RER_j , MSE_j y $EPOR_j$ y que se corresponden con las expresiones del cuadro 4.15. La segunda parte es, siempre, el cálculo de la media de los valores anteriores para todos los planes de medios. Como se observa fácilmente al comparar los cuadros 4.15 y 4.16, toda la información necesaria para calcular las expresiones del segundo, se encuentran presentes en el primero.

Cuadro 4.16. Medidas del error cometido en cada plan de medios.

Estimación de la cobertura	Estimación de la distribución de contactos
$AER_j = \hat{R}_j - R_j $	$MSE_j = \frac{1}{k_j+1} \sum_{i=0}^{k_j} (E_{ij} - O_{ij})^2$
$RER_j = \frac{ \hat{R}_j - R_j }{R_j}$	$EPOR_j = \frac{\sum_{i=1}^{k_j} O_{ij} - E_{ij} }{1 - O_{0j}}$

Fuente: Elaboración propia

El análisis de varianza de un factor (ANOVA), nos permite plantear la cuestión de si el valor medio de una variable continua (el error en cualquiera de las cuatro medidas señaladas), es o no significativamente distinto en cada uno de los M grupos que crea una variable categórica (cada uno de los modelos). Es decir, se pretende contrastar las siguientes hipótesis nulas que dotarán de sentido a la ordenación jerárquica de los modelos en función de su precisión:

$$H_0^a : AER^1 = AER^2 = \dots = AER^M$$

$$H_0^b : RER^1 = RER^2 = \dots = RER^M$$

$$H_0^c : MSE^1 = MSE^2 = \dots = MSE^M$$

$$H_0^d : EPOR^1 = EPOR^2 = \dots = EPOR^M$$

Dicho de otra forma, las hipótesis nulas que se plantean parten de la base de que el error que cometen todos los modelos es el mismo. Será necesario poder rechazar estas hipótesis a un nivel de significación razonable, para establecer sobre esta base las clasificaciones oportunas de los modelos.

No creemos que sea adecuado profundizar en los fundamentos de esta técnica de inferencia estadística, dado que es de sobra conocida y no supondría valor añadido alguno a la aportación de este epígrafe. Sin embargo, sí que creemos necesario subrayar que el análisis de la varianza, como la regresión múltiple, son dos de las pruebas estadísticas más potentes y, por tanto, con requisitos muy exigentes en cuanto al cumplimiento de las hipótesis que sustentan las pruebas paramétricas (Manzano, 1995), en consecuencia, se mantendrán las cautelas necesarias para su aplicación.

En particular, y con respecto a la primera de las hipótesis que hace referencia a la exigencia de que los residuos *tengan una distribución normal*. Es decir, en nuestro caso que, dentro de cada modelo, la diferencia entre el valor medio del error, v.g. AER^1 , y los errores calculados para cada plan de medios, AER^1_j , se distribuyan

normalmente. El hecho de que los residuos no se aproximen a una distribución normal, en general, no afecta de forma importante al estadístico F que se construye para rechazar o no la hipótesis nula (Visauta y Batallé, 1986). La razón se debe a que, como se están comparando medias, puede ser válida la aplicación del teorema central del límite a datos procedentes de una distribución no normal (Uriel, 1995). Sin embargo, se ha preferido comprobar la normalidad de la población mediante la aplicación del *test* de Kolmogorov-Smirnov.

La segunda hipótesis que se exige para la aplicación del ANOVA, es que las varianzas de la medida del error en cada uno de los modelos sea la misma, esto es, que las *poblaciones sean homoscedásticas*. De nuevo, Uriel (1995; pág 184), afirma que “el estadístico F no se ve muy afectado por el hecho de que no exista homoscedasticidad, siempre que las muestras de los diferentes grupos sean del mismo o similar tamaño. En este sentido, algunos autores afirman que el no cumplimiento de esta hipótesis puede afectar a los contrastes del análisis de la varianza cuando la razón entre los distintos tamaños muestrales del grupo más grande y más pequeño, es mayor de 2”. En nuestro caso, todos los grupos (los modelos), tienen el mismo número de planes de medios, con lo que al ser el tamaño igual, no sería muy grave según lo anterior, que las poblaciones no fueran homoscedásticas. En todo caso, se ha comprobado esta eventualidad mediante el test de Levene y, caso de ser poblaciones heteroscedásticas, se ha utilizado como variable dependiente el logaritmo del error, siguiendo la recomendación de Danaher (1989).

Finalmente, la *independencia de los errores* viene garantizada por el propio diseño experimental. El valor del dato de error suministrado por un modelo en un momento, no depende en absoluto, del dato suministrado por otro modelo anteriormente (Cuadras, Echeverría, Mateo y Sánchez, 1991).

El análisis de varianza, sin embargo, nos ofrece una información insuficiente para nuestro objetivo. Caso de rechazarse la hipótesis nula de igualdad de las

medidas de error, sabemos que existe al menos una diferencia significativa, es decir, que algún modelo comete un error significativamente mayor o menor que el resto, pero no podemos señalar cuántas diferencias significativas hay en total ni entre qué medias concretamente. Para solventar este problema, diversos trabajos entre los referenciados en el cuadro 4.13, recurren a las llamadas pruebas de comparación múltiple entre medias (Cuadras, Echeverría, Mateo y Sánchez, 1991), análisis *ex-post* (Uriel, 1995) o contrastes múltiples *a posteriori* (Manzano, 1995). Concretamente Danaher (1989) y Leckenby y Kim (1993; 1994) emplean el *test* de Tukey (también llamado HSD: *Honestly Significant Difference*) para este propósito. Dado que no es ésta una prueba con un uso extendido, señalaremos brevemente sus fundamentos según la descripción de Cuadras, Echeverría, Mateo y Sánchez, 1991.

En primer lugar, este *test* calcula todas las diferencias entre las medias del error (v.g. AER) en cada grupo (modelos). A continuación se obtiene la desviación típica de estas diferencias que, de acuerdo con Tukey, se define:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N X_{ij}^2 - \frac{j}{n} \left(\sum_{i=1}^N X_{ij} \right)^2}{n}}$$

cuya notación adaptada a nuestros objetivos implicaría que N es el número total de planes multiplicado por el número de modelos analizados, X_{ij} es el valor del error para el plan de medios i y el modelo j y n , se corresponde con el número total de planes de medios.

A continuación, cada una de las diferencias de medias inicialmente calculadas, se estandariza dividiendo por la desviación típica hallada, comparándose entonces con unos valores tabulados por Tukey en 1953 y que dependen del número de medias que se están comparando, del número de grados de libertad de la media cuadrática intra grupo y del nivel de significación deseado. Los valores de la

diferencia de medias superiores al valor tabulado, serán significativamente distintos. Una reproducción de esta tabla, está disponible en Cuadras, Echeverría, Mateo y Sánchez (1991; pp. 933 y 934) y, evidentemente, este proceso ya está implementado en el software estadístico que hemos empleado.

En síntesis, el análisis de varianza nos permitirá establecer, en primer lugar, si algún modelo funciona significativamente mejor o peor que el resto gracias a que la media del error es significativamente distinta a la de los demás. Caso de ser así, la prueba de Tukey nos dirá exactamente cuáles modelos son los que tienen ese comportamiento diferente y, entonces sí, los valores medios del error permitirán la ordenación de los mismos en base a su precisión en el cálculo de la cobertura y la distribución de contactos.

4.2.5.2 *Determinación de los factores de error*

Para determinar cuáles de los factores de error potenciales, descritos en el epígrafe 4.2.4, explican significativamente el error cometido por los diferentes modelos evaluados, se ha mantenido la técnica de análisis de datos de los trabajos de Leckenby y Kim (1981;1982b). Así, pues, se han construido cuatro funciones de regresión, cada una de las cuales tiene como variables dependientes las cuatro definiciones del error para cada plan de medios que se especificaron en el cuadro 4.16. Por su parte, actúan como variables independientes los doce factores de error $X_1 \cdots X_{12}$, tal y como se definieron y operativizaron en el epígrafe 4.2.4., esto es:

$$\begin{aligned} AER_j &= a_1 + \sum \beta_K X_K + \varepsilon_1 & K = 1, \dots, 12 \\ RER_j &= a_2 + \sum \beta_K X_K + \varepsilon_2 & K = 1, \dots, 12 \\ EPOR_j &= a_3 + \sum \beta_K X_K + \varepsilon_3 & K = 1, \dots, 12 \\ MSE_j &= a_4 + \sum \beta_K X_K + \varepsilon_4 & K = 1, \dots, 12 \end{aligned}$$

Ante lo común de la técnica utilizada, no entraremos en detalles hasta la presentación de los resultados, solamente indicaremos que, tal como se ha señalado en el análisis de la varianza, se han realizado todas las pruebas necesarias para asegurar el cumplimiento de las hipótesis básicas de aplicabilidad de la técnica, esto es: normalidad, independencia, homoscedasticidad y corrección de la multicolinealidad cuando ésta era importante. Consecuencia de este último aspecto, ha sido la supresión de los factores de error números 11 y 12, suma de las audiencias duplicadas y acumuladas y la relación entre esta suma y la audiencia bruta. Como se desprende inmediatamente al analizar su definición en las expresiones [4.19] y [4.20], estas pueden obtenerse sin más que sumar factores de error anteriores, de tal forma que la multicolinealidad sería perfecta. Aunque Leckenby y Kishi (1981, 1982b) mantienen estos factores en los cuadros que sintetizan los resultados de sus regresiones, ningún coeficiente es significativo, lo que nos hace suponer que, en realidad, nunca fueron incorporados en las mismas.

5

Análisis de resultados e implicaciones

5.1 INTRODUCCIÓN.

En el presente capítulo, se presentan los resultados obtenidos en el proceso de evaluación, jerarquización y determinación de los principales factores de error¹ en los modelos de exposición, de acuerdo con los objetivos planteados en el capítulo precedente. Los resultados se han dividido en tres grandes apartados, atendiendo a los tres tipos de modelos analizados: modelos de acumulación de audiencias, modelos de duplicación de audiencias, y modelos de audiencia neta acumulada. Cada

¹ Se ha preferido el término error a otros alternativos como desviación por dos motivos: (a) es el habitualmente utilizado en trabajos análogos al presente y (b) evita la confusión con términos estadísticos también utilizados como la desviación típica.

uno de estos apartados, tendrá una estructura idéntica. Así, en primer lugar, se analizarán los errores cometidos en la estimación de la cobertura de los planes de medios diseñados y, a continuación, en los que se ha incurrido al predecir la distribución de contactos. Atendiendo a esos errores, se clasificarán los modelos en función de su mejor o peor capacidad de estimar las mencionadas variables, agrupando aquellos cuyo comportamiento se haya demostrado similar, identificando al mejor modelo y estableciendo, de este modo, una jerarquía que facilite al planificador la elección del más adecuado.

En una segunda fase se relacionará, dentro de cada apartado, el error cometido con las características estructurales de los planes, lo que hemos venido llamando factores de error. Se pretende, con ello que, ante modelos que demuestren comportamientos similares en el conjunto de los planes analizados, el planificador cuente con un criterio adicional de discriminación entre ellos que posibilite su jerarquización y selección.

5.2 EVALUACIÓN DE LOS MODELOS DE ACUMULACIÓN DE AUDIENCIAS.

5.2.1 Error en la estimación de la cobertura.

Todos los modelos de acumulación de audiencias analizados estiman la cobertura del plan de medios. Algunos de ellos, como se vio en el capítulo tercero, sólo permiten calcular esa variable: modelos *ad hoc*, mientras que los modelos estocásticos ofrecen, además, la distribución de contactos. Sin embargo, antes de comenzar la exposición de las conclusiones del proceso evaluador, debe señalarse un resultado adicional obtenido, referente a las condiciones de aplicabilidad del modelo de Morgensztern (1970) que nos han llevado a tratarlo como un modelo *ad hoc* cuando en el capítulo tercero se clasificó como estocástico.

En el desarrollo teórico del modelo de Morgensztern (1970) expuesto en el apartado 3.1.2.8, se obtuvo una condición de aplicabilidad del modelo: el parámetro b , sólo podía calcularse cuando la cobertura tras dos inserciones fuera inferior o igual a cuatro terceras partes de la audiencia de un número del soporte. De los cuarenta y nueve planes de medios diseñados para evaluar el modelo, treinta y seis de ellos cumplen esta condición y, en principio, permiten obtener la distribución de contactos. Sin embargo, se ha constatado que Morgensztern (1970) primero y Chandon (1985) después, no consideran que la expresión que permite obtener el parámetro a a partir del b , es en si misma una segunda condición de aplicabilidad. Si el mencionado parámetro a debe ser positivo para ofrecer distribuciones de contactos coherentes (porcentajes poblacionales positivos), esto sólo se cumple cuando:

$$\frac{1}{1+b} < R_1 < \frac{b}{1+b}$$

condición que no cumplen ninguno de los planes de medios construidos, lo que provoca que el parámetro a sea negativo y las distribuciones de contactos obtenidas mediante [3.45] totalmente incoherentes.

Sin embargo, en lugar de excluir dicho modelo de la evaluación, dado que la estimación que realiza de la cobertura era razonable, se ha optado por considerar a este modelo como *ad hoc* e incluirlo en el proceso. En estos modelos, al no formularse hipótesis alguna sobre el comportamiento de los individuos, la única característica que le es exigible a la forma funcional empleada para estimar la cobertura, es que refleje bien la realidad de los datos y en este caso, como se verá, así era.

Otra decisión que se ha adoptado en el proceso de análisis, ha sido el tratamiento dado a aquellos modelos con un comportamiento mucho peor que el resto (*outliers*) que, por serlo, impedían la estabilización de la varianza mediante las transformaciones habituales de la variable dependiente: logaritmos neperianos,

arcosenos y raíces cuadradas. Este hecho, imposibilitaba la realización del análisis de la varianza al provocar que los datos violaran la hipótesis de homoscedasticidad, según la cual la varianza de la medida del error debería ser la misma en cada uno de los modelos. Una vez identificados, y demostrado que su capacidad predictiva es significativamente inferior al resto de modelos, se han excluido del análisis para permitir establecer comparaciones adecuadas entre los modelos con mejor ajuste. El cuadro 5.1 ofrece los valores medios para los 49 planes construidos, de los errores promedio (AER) y relativo (RER) cometidos en la estimación de la cobertura de cada uno de los modelos, tal y como se definieron en el capítulo cuarto. En un primer análisis, en el que luego se profundizará, se identifican claramente a tres modelos como *outliers*: Politz, binomial y Broadbent. Esta afirmación es cierta independientemente de la medida de error que se emplee, como se ve gráficamente en la figura 5.1.

Cuadro 5.1. Errores promedio en la estimación de la cobertura. Identificación de *outliers*.

Modelo	AER		RER	
	Media*	D. Típica	Media*	D. Típica
Logarítmico	0,0081	0,0008	0,1097	0,0061
Exponencial	0,0017	0,0003	0,0313	0,0070
Agostini	0,0056	0,0003	0,0661	0,0079
Hofmans	0,0008	0,0000	0,0178	0,0010
Binomial	0,1170	0,0087	2,0404	0,6774
Dos segmentos	0,0068	0,0017	0,1127	0,1064
Tres segmentos	0,0059	0,0000	0,0781	0,0006
Politz	0,2885	0,0070	8,4304	6,1128
Beta binomial	0,0076	0,0001	0,1027	0,0030
Binomial negativa	0,0303	0,0085	0,4812	0,1769
Broadbent	0,0923	0,0053	1,5655	0,4377
Morgensztern	0,0036	0,0008	0,0602	0,0033
Beta binomial modificada	0,0013	0,0002	0,0245	0,0008
Markov	0,0257	0,0006	0,4122	0,1201

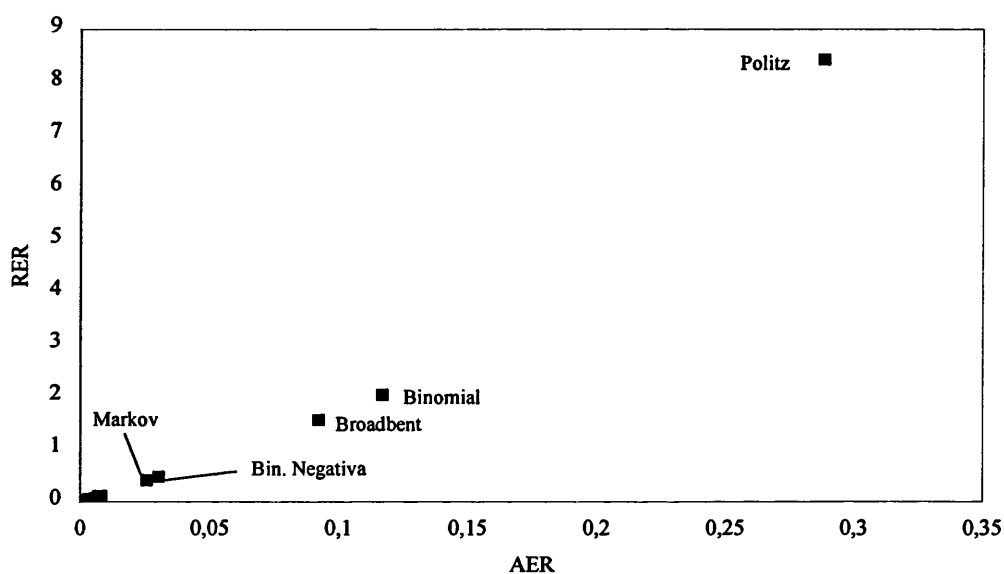
*N = 49, salvo Morgensztern N=36

Fuente: Elaboración propia.

Una vez constatado el hecho de que existe un grupo formado por los tres modelos señalados, cuya capacidad predictiva de la cobertura es muy pobre, y

excluidos del análisis posterior por las razones argumentadas, se han transformado las variables AER y RER tomando logaritmos neperianos, con el fin de estabilizar la varianza y, de este modo, poder determinar si las diferencias de comportamiento del resto de modelos, que aparecen en el cuadro 5.1, son estadísticamente significativas, utilizando para ello una técnica como el análisis de varianza que, por ser muy paramétrica, requiere que las poblaciones sean homoscedásticas.

Figura 5.1. Identificación de *outliers*



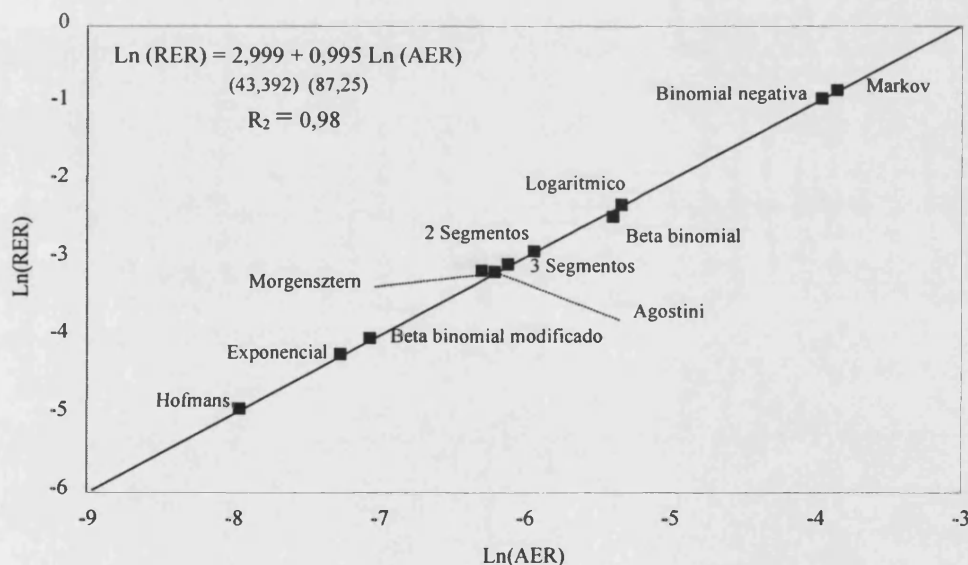
Pero antes de exponer los resultados a los que se ha llegado, se ha creído conveniente determinar si las dos medidas de error en la estimación de la cobertura utilizadas, son coherentes entre sí, es decir, si utilizadas por separado permitirían llegar a las mismas conclusiones. Una forma de comprobarlo consiste en regresar los errores medios cometidos por los modelos que permanecen en el análisis y medidos mediante una de ellas, sobre los resultantes de utilizar la otra medida. Si ambas son coherentes, el ajuste de la recta de regresión debería ser alto, dado que ello indicaría que una variable no es sino una transformación lineal de la otra.

Como puede comprobarse en la figura 5.2, este ajuste es prácticamente perfecto ($R^2 = 0,9$), lo que nos permitiría jerarquizar a los modelos de acumulación de audiencias en su capacidad predictiva de la cobertura, a partir de una sola de las definiciones de error, dado que una puede ponerse, con absoluta precisión, como una combinación lineal de la otra del siguiente modo:

$$\ln RER = -2,999 + 0,995 \ln AER$$

Se ha preferido, sin embargo, mantener ambas variables en previsión de que en futuros contrastes: de la capacidad predictiva de la distribución de contactos, o en los siguientes grupos de modelos analizados, afloraran diferencias.

Figura 5.2. Coherencia de los errores tipo AER y RER



El análisis de varianza resumido en el cuadro 5.2 confirma que el error es una variable dependiente del tipo de modelo ($p \leq 0,01$), en otras palabras, que no todos los modelos gozan de la misma capacidad predictiva de la cobertura, y esto es cierto independientemente de la medida del error que se haya empleado. Sin embargo, para

lograr el objetivo de ordenar a estos modelos en función de su capacidad predictiva, es necesario determinar cuáles provocan esta significatividad conjunta de las diferencias, para lo cual, como se justificó en el capítulo cuarto, se ha realizado un análisis *post hoc* basado en la prueba de Tukey, sintetizado en los cuadros 5.3 y 5.4.

Cuadro 5.2. Análisis de varianza para las definiciones de error

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
ln AER	Entre modelos	10	79,34	43,17*
	Error	515	1,84	
ln RER	Entre modelos	13	78,65	60,41*
	Error	515	1,30	

* Significativo para $p \leq 0,01$

Fuente: elaboración propia

Esta prueba nos permite obtener las conclusiones que a continuación expondremos, y que nos permitirán contestar a la cuestión a investigar CI_{IB} . En primer lugar, debe recordarse que existe un grupo de modelos: Politz, binomial y Broadbent, que demuestran una capacidad predictiva muy inferior al resto. Tanto es así que, como se ha señalado, se les ha considerado como *outliers* y excluido del análisis para facilitar la extracción de conclusiones respecto a los demás, pero no debe olvidarse que, en el proceso de clasificación que estamos desarrollando, ocupan el último lugar en cuanto a la capacidad de estimar la cobertura. El resto de modelos, atendiendo simultáneamente a las dos medidas del error, puede agruparse en tres grupos² como se señala a continuación y se ilustra en la figura 5.3.

² Tres era el número de grupos homogéneos con medias no significativamente distintas que se derivaban de la realización de las pruebas *post hoc* en la mayoría de los análisis realizados. Cuando no ha sido así, se ha optado, en aras a la simetría de los resultados, por mantener los tres grupos de error, indicándose en todo momento qué criterio se ha adoptado para asignar a un grupo determinado a un modelo que debería formar su propio grupo.

Cuadro 5.3. Matriz de diferencias pareadas de lnAER. Test de Tukey.

	Hofmans	BBMOD	EXP	MORG	AGOS	3SEG	2SEG	BBIN	LOG	MKV
Beta binomial modificado	-0,8599									
Exponencial	-0,6940	0,2049								
Morgensztern	-1,6643**	-0,7653	-0,9702*							
Agostini	-1,7588**	-0,7599	-1,0648**	-0,0946						
3 segmentos	-1,8487**	-0,9497*	-1,1547**	-0,1844	-0,0899					
2 segmentos	-2,0256**	-1,1267*	-1,3316**	-0,3613	-0,2667	-0,1769				
Beta binomial	-2,5667**	-1,6678**	-1,8727**	-0,9024	-0,8079	-0,7180	-0,5411			
Logarítmico	-2,6237**	-1,7248**	-1,9297**	-0,9594*	-0,8649	-0,7750	-0,5981	-0,0570		
Markov	-3,9959**	-3,0970**	-3,3019**	-2,3317**	-2,2371**	-2,1472**	-1,9703**	-1,4292**	-1,3722**	
Binomial negativa	-4,1028**	-3,2039**	-3,4087**	-2,4385**	-2,3439**	-2,2541**	-2,0772**	-1,5361**	-1,4790**	-0,1068

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ Fuente: *Elaboración propia*

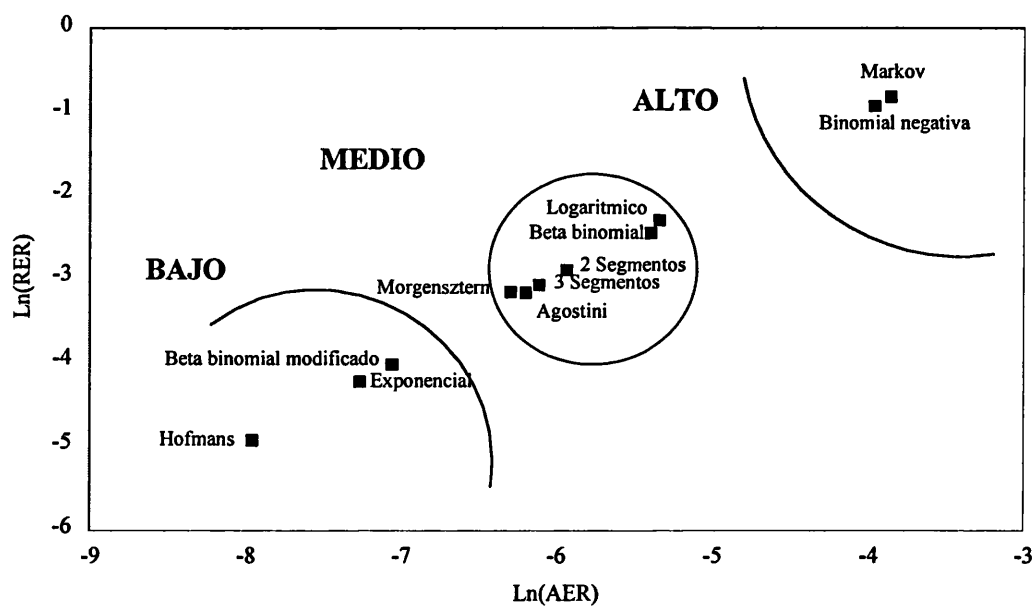
Cuadro 5.4. Matriz de diferencias pareadas de lnRER. Test de Tukey.

	Hofmans	EXP	BBMOD	MORG	AGOS	3SEG	2SEG	BBIN	LOG	MKV
Beta binomial modificado	-0,6940									
Exponencial	-0,8989**	-0,2049								
Morgensztern	-1,7653**	-1,0713**	-0,8664*							
Agostini	-1,7588**	-1,0648**	-0,8599**	0,0064						
3 segmentos	-1,8487**	-1,1547**	-0,9497**	-0,0834	-0,0899					
2 segmentos	-2,0256**	-1,3316**	-1,5716**	-0,2603	-0,2668	-0,1769				
Beta binomial	-2,4706**	-1,7765**	-1,5716**	-0,7052	-0,7117	-0,6219	-0,4450			
Logarítmico	-2,6237**	-1,9297**	-1,7248**	-0,8584*	-0,8649**	-0,7750*	-0,5981	-0,1532		
Markov	-3,9959**	-3,3019**	-3,0970**	-2,2306**	-2,2370**	-2,1472**	-1,9703**	-1,5254**	-1,3722**	
Binomial negativa	-4,1028**	-3,4087**	-3,2039**	-2,3374**	-2,3439**	-2,2541**	-2,0442**	-1,6322**	-1,4790**	0,1685

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: Elaboración propia*

- Grupo de *error bajo*. Formado por los modelos de Hofmans, exponencial y beta binomial modificado.
- Grupo de *error medio*. Integrado por los modelos de Morgensztern, Agostini, tres y dos segmentos, beta binomial y logarítmico.
- Grupo de *error alto*. Estaría compuesto por los modelos que, sin ser *outliers*, muestran un peor comportamiento, esto es, los modelos de Markov y de la binomial negativa.

**Figura 5.3. Jerarquización de los modelos de acumulación de audiencias.
Estimación de la cobertura.**



Fuente: Elaboración propia

Como se vio con detalle en el capítulo tercero, cada modelo está caracterizado por la asunción de una serie de hipótesis que pretenden simplificar la realidad para facilitar la modelización. El coste de oportunidad de estas simplificaciones radica, precisamente, en que cabe esperar que la capacidad

predictiva de aquellos modelos que realicen las asunciones más restrictivas, sea peor que el de aquellos otros que intentan simplificar en menor medida la complejidad del comportamiento de los individuos frente a los medios. Con el análisis realizado, se dispone de elementos de juicio para contrastar este *a priori* que habíamos formulado como cuestión a investigar CI_{IC} , en el ámbito de la estimación de la cobertura, pudiendo llegarse a las siguientes conclusiones:

- a. Puede comprobarse que, entre los tres modelos que forman el grupo de los que cometen menor error, se encuentran dos modelos *ad hoc* de construcción muy sencilla, Hofmans y exponencial, frente a un sólo modelo estocástico. Como se indicó, estos modelos no consideran función de probabilidad alguna para recoger los patrones de comportamiento de la audiencia, tal y como realizan los modelos estocásticos, por lo que puede concluirse que la introducción de estas funciones, no redundan en mejoras significativas en la predicción de la cobertura.
- b. Centrándonos en las hipótesis que sustentan a los modelos estocásticos, parece que sí es importante considerar la heterogeneidad de la población respecto a las probabilidades de exposición. Los únicos modelos que no la consideran, o bien se encuentran entre los dos con peor comportamiento (Markov), o bien son *outliers* (binomial).
- c. No es relevante, sin embargo, la forma en que se recoja esa heterogeneidad. Así nos encontramos con que poseen la misma capacidad predictiva, tanto modelos que la introducen mediante la consideración de segmentos de población (modelos de dos y tres segmentos), como los que recurren a asignar probabilidades individuales que luego se distribuyen poblacionalmente según distintas funciones: modelos beta binomial, binomial negativa, etc.

- d. Por el contrario, sí que parece importante, que cuando se emplean funciones de distribución continuas, se corrijan para evitar que den un peso nulo a los extremos (lectores fieles y no lectores). La corrección del modelo beta binomial, considerando este hecho, logra convertir al modelo beta binomial modificado en uno de los mejores predictores de la cobertura.
- e. Todos los modelos que demuestran mejores resultados, consideran el carácter no aleatorio de la acumulación de audiencias.
- f. Finalmente, no parece que la ruptura de la hipótesis de estacionariedad respecto a las inserciones, realizada por el modelo de acumulación binomial de Markov, compense con su funcionamiento la complejidad que introduce.

En el capítulo cuarto, se justificó la consideración del medio diarios en los planes de medios diseñados, argumentando que recogían una parte importante de la inversión publicitaria realizada en España y que, por otro lado, no eran utilizados para contrastar la capacidad predictiva de los modelos de exposición porque se consideraba que la mayoría de éstos se habían diseñado exclusivamente para el medio revistas. Nos planteamos, llegados a este punto, comprobar si el funcionamiento de los modelos en su conjunto, difiere según el tipo de medio que se emplea en el plan: diarios, revistas mensuales y revistas semanales, tal y como señalamos en la cuestión a investigar CI_{2B} . El cuadro 5.5, muestra las medias de error en la estimación de la cobertura cometidos por el conjunto de los modelos según el tipo de soporte. El cuadro 5.6 resume el análisis de varianza que ha empleado como factor el tipo de medio y como variable dependiente las dos definiciones de error empleadas y, finalmente, el cuadro 5.7 recoge el análisis *post hoc* que nos indica la fuente de las diferencias.

Cuadro 5.5. Medias de lnAER y lnRER según el tipo de soporte.

Medios	lnAER		lnRER	
	Media	D. típica	Media	D. típica
Diarios	-5,1664	1,3855	-2,3054	1,1983
R. Semanales	-6,0365	2,0898	-3,3206	1,8569
R. Mensuales	-6,1868	1,6581	-2,8227	1,5881

Fuente: elaboración propia

Cuadro 5.6. Análisis de varianza para las definiciones de error tomando como factor el tipo de soporte

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
lnAER	Entre tipo soporte	2	38,73	0,98*
	Error	523	3,18	
lnRER	Entre tipo soporte	2	35,88	2,86*
	Error	523	2,65	

* Diferencias significativas para $p \leq 0,01$

Fuente: elaboración propia

Cuadro 5.7. Matriz de diferencias pareadas de lnAER y lnRER Test de Tukey

medios	lnAER		lnRER	
	Mensuales	Semanales	Mensuales	Semanales
Semanales	-0,1502		0,4979**	
Diarios	-1,0203**	-0,8701**	-0,5173*	-1,0152**

** diferencias significativas para $p \leq 0,01$

* diferencias significativas para $p \leq 0,05$

Fuente: elaboración propia

El cuadro 5.5 muestra que los errores más elevados son cometidos por los modelos cuando estiman los planes formados por periódicos. Este hecho parece confirmar que estos modelos no han sido diseñados para este medio, al menos en la estimación de la cobertura. Adicionalmente, se ha realizado la prueba de Tukey, para el lnAER, y esta concluye en la misma línea: todas las diferencias entre los grupos son significativas, es decir, que la ordenación del error revistas mensuales, semanales

y periódicos parece consistente. Sin embargo, el lnRER obtiene resultados diferentes. Así, el comportamiento de los modelos sería el mismo cuando evalúan revistas mensuales y periódicos (para $p \leq 0.01$), y significativamente peor que cuando el plan de medios está formado por revistas semanales. Este hecho, nos hace pensar si, más que depender del tipo de soporte, el error cometido por los modelos no dependerá del número de inserciones que constituyen el plan. Recordemos que, para periódicos y revistas mensuales, éste número era mayor (siete y seis respectivamente) que para las revistas semanales (cuatro). Esta cuestión se dirimirá cuando se analicen los factores de error, es decir, la influencia de las características internas de los planes sobre el error cometido por los modelos y, en ese momento, se podrá dar una respuesta mucho más categórica a la cuestión planteada.

Evaluados y jerarquizados los modelos de acumulación de audiencias en función de su capacidad predictiva de la cobertura de los planes de medios, se pasa a continuación a desarrollar esta misma tarea para la estimación de las distribuciones de contactos.

5.2.2 Error en la estimación de la distribución de contactos

Como se indicó reiteradamente al construir la tipología de los modelos de exposición en el capítulo segundo, la principal limitación de los modelos *ad hoc* es que son incapaces de deducir la distribución de contactos de un plan de medios. Por este motivo; en esta parte del análisis, el conjunto de modelos objeto del mismo, queda reducido a los modelos estocásticos de acumulación de audiencias, esto es, el modelo binomial, los de dos y tres segmentos, el modelo de Politz, el beta binomial y su versión modificada, la binomial negativa, Broadbent y Markov. Como se justificó en el epígrafe anterior, se ha eliminado de esta parte del análisis al modelo de Morgensztern, al concluir en la fase de operativización del mismo, que las condiciones de aplicabilidad lo convertían realmente en un modelo *ad hoc*.

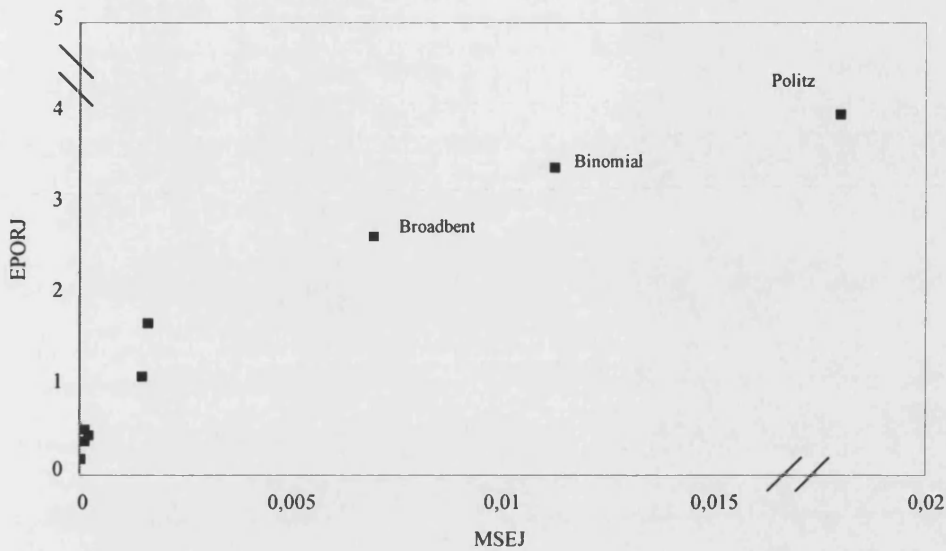
El cuadro 5.8, resume, en respuesta a la cuestión a investigar CI_{1A} el error cometido por los mencionados modelos en la estimación de la distribución de contactos, para el conjunto de los planes de medios, medido éste mediante las expresiones ya analizadas en el capítulo anterior: error cuadrático medio (MSEJ) y error en las probabilidades de exposición al plan (EPORJ).

Cuadro 5.8. Errores promedio en la estimación de la distribución de contactos. Identificación de *outliers* (N=49).

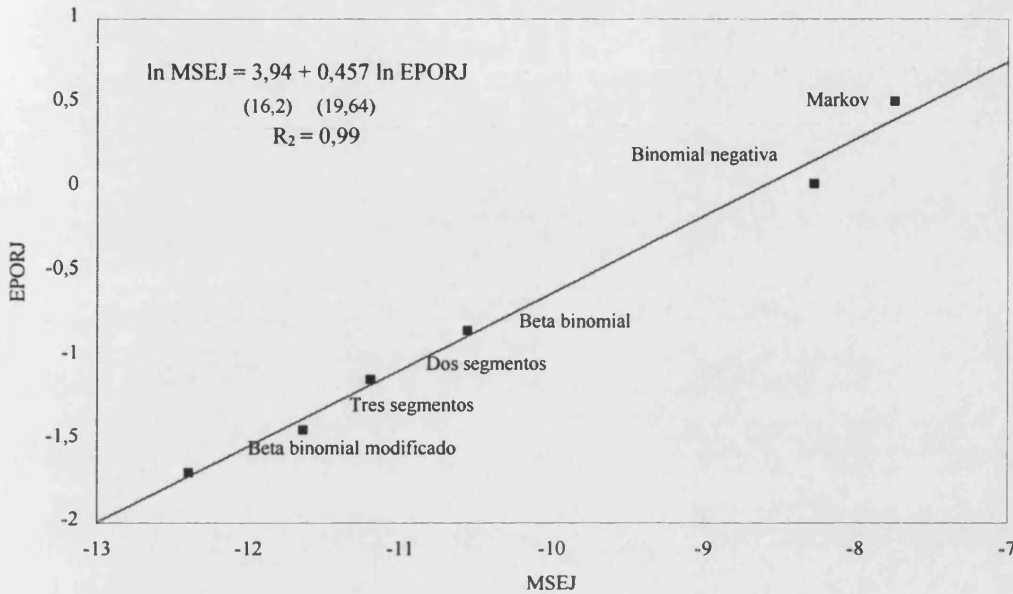
Modelo	MSEJ		EPORJ	
	Media	D. Típica	Media	D. Típica
Binomial	0,0112	0,0068	3,4076	0,8698
Dos segmentos	0,0001	0,0001	0,5141	0,5039
Tres segmentos	0,0001	0,0001	0,3788	0,4643
Politz	0,1611	0,0002	50,2323	40,1941
Beta binomial	0,0002	0,0001	0,4408	0,1024
Binomial negativa	0,0014	0,0001	1,1011	0,3458
Broadbent	0,0069	0,0001	2,6467	0,5608
Beta binomial modificada	0,0001	0,0009	0,1931	0,0096
Markov	0,0016	0,0001	1,6873	0,2419

Fuente: Elaboración propia.

Analizando conjuntamente el cuadro 5.8 y su traducción gráfica, la figura 5.4, antes incluso de comprobar si las diferencias son estadísticamente significativas o no, ya se aprecia que los resultados son muy similares a los obtenidos en el apartado anterior, es decir, que aquellos modelos de acumulación que demuestran una peor capacidad predictiva de la cobertura de los planes (Politz, binomial y Broadbent), también muestran sus deficiencias en la estimación de la distribución de contactos. Como hicimos y justificamos con anterioridad, eliminaremos a estos modelos del análisis, que no de la clasificación jerárquica, para centrarnos en las diferencias que se observan en el resto de modelos. Mantendremos, para ello, la estructura expositiva del epígrafe anterior: análisis de la coherencia de las dos medidas de error, análisis de la significatividad de las diferencias de error cometidas por los distintos modelos de acumulación analizados, implicaciones de los resultados para la introducción de hipótesis simplificadoras en los modelos y, finalmente, análisis de las diferencias por tipo de soporte.

Figura 5.4. Identificación de *outliers*.

En relación con la distribución de contactos, se observa que, al igual que ocurría en el análisis de la cobertura, las dos medidas de error empleadas para determinar la precisión en las estimaciones de los diferentes modelos, ofrecen resultados totalmente paralelos. De nuevo se ha regresado el logaritmo natural de EPORJ sobre esta misma transformación de MSEJ, y el ajuste perfecto de la recta de regresión ($R^2 = 0.99$), demuestra, como se aprecia en la figura 5.5, que una variable es tan equivalente a la otra que, en realidad, se presenta como una transformación lineal de aquella. Pese a ello, como hicimos con anterioridad, mantendremos por paralelismo a ambas en el desarrollo subsiguiente.

Figura 5.5. Coherencia de los errores tipo EPORJ y MSEJ.

Fuente: Elaboración propia

Respecto al comportamiento de los modelos en la estimación de la distribución de contactos, de nuevo existen diferencias significativas entre el conjunto de los mismos, como pone de manifiesto el cuadro 5.9. Cuando a través del análisis *post hoc*, (cuadros 5.10 y 5.11) se profundiza en qué modelos son los causantes de esas diferencias, nos encontramos ahora con más dificultades a la hora de establecer grupos homogéneos, dado que alguno de ellos ocupa posiciones fronterizas lo que, unido a las ligeras divergencias entre los resultados que ofrece la prueba de Tukey según la medida de error que se emplee, contribuye a dificultar la clasificación

Cuadro 5.9. Análisis de varianza para las definiciones de error

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
ln MSEJ	Entre modelos	5	173,45	42,62*
	Error	288	4,07	
ln EPORJ	Entre modelos	5	36,52	85,49*
	Error	288	0,43	

* Significativo para $p \leq 0,01$ **Cuadro 5.10 Matriz de diferencias pareadas de lnMSEJ.
Test de Tukey.**

	BBMD	3SEG	2SEG	BBIN	BNEG
3 segmentos	-0,7568				
2 segmentos	-1,2005*	-0,4436			
Beta binomial	-1,8381**	-1,0813	-1,0813		
Binomial negativa	-4,1247**	-3,3678**	-3,3678**	-2,2866**	
Markov	-4,6511**	-3,8943**	-3,8943**	-2,8130**	-0,5265

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: Elaboración propia***Cuadro 5.11 Matriz de diferencias pareadas de lnEPORJ.
Test de Tukey.**

	BBMD	3SEG	2SEG	BBIN	BNEG
3 segmentos	-0,2552				
2 segmentos	-0,5564**	-0,3012			
Beta binomial	-0,8484**	-0,5933**	-0,2921		
Binomial negativa	-1,7206**	-1,4654**	-1,1642**	-0,8722**	
Markov	-2,2087**	-1,9535**	-1,6524**	-1,3603**	-0,4881**

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: Elaboración propia*

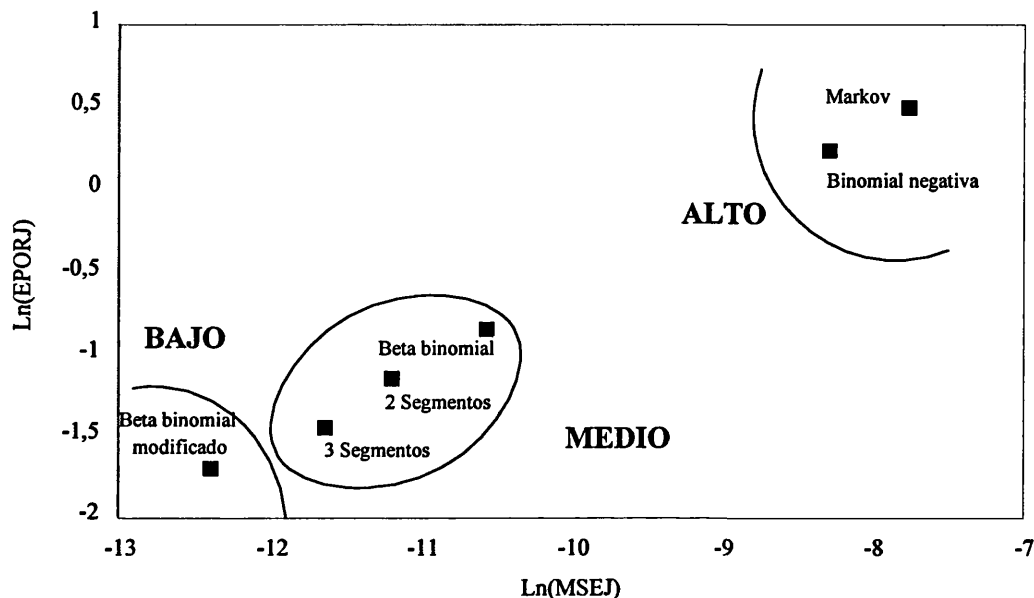
Cuando un modelo no se puede incluir en un grupo porque tiene una media de error significativamente distinta de alguno de sus miembros, se ha optado por incluirlo en aquel en el que esté presente el modelo con la media no significativamente distinta más cercana, de esta manera, con la salvedad señalada, podemos diferenciar de nuevo tres grandes grupos en respuesta a la cuestión a

investigar CI_{1B} , que etiquetaremos de manera análoga a como lo hicimos en la estimación de la cobertura y cuya ilustración se observa en la figura 5.6:

- a. Grupo de *error bajo*. Formado exclusivamente por el modelo beta binomial modificado.
- b. Grupo de *error medio*. Integrado por los modelos beta binomial y de dos y tres segmentos. El modelo de tres segmentos, atendiendo al $\ln MSEJ$, no tiene una media de error significativamente distinta del beta binomial modificado pero, como se ha comentado, se ha optado por incluirlo en este segundo grupo al estar su media más cercana al error cometido por el modelo de dos segmentos que por el beta binomial modificado.
- c. Grupo de *error alto*. Estaría compuesto, de nuevo, por los modelos que, sin ser *outliers*, muestran un peor comportamiento, esto es, los modelos de Markov y de la binomial negativa.

Al abordar la cuestión CI_{1C} , referente a la conveniencia de introducir hipótesis simplificadoras en el diseño de los modelos, dado que la clasificación de los modelos estocásticos de acumulación de audiencias, atendiendo a la estimación que realizan de la distribución de contactos, es la misma que se derivó en el cálculo de la cobertura, las implicaciones sobre la conveniencia de imponer determinadas hipótesis son idénticas a las que expusimos entonces y no se reiterarán con detalle. En síntesis se destaca la importancia de que los modelos incorporen la heterogeneidad de la población, independientemente de que lo hagan mediante la consideración de segmentos o asignando probabilidades individuales que se distribuyan según distintas funciones que, por otra parte, deberán corregirse para no dar un peso nulo a los extremos. Finalmente no parece mejorar la capacidad predictiva de los modelos el incorporar la no estacionariedad de las inserciones.

**Figura 5.6. Jerarquización de los modelos de acumulación de audiencias.
Estimación de la distribución de contactos.**



El último aspecto que debe ser comentado respecto al funcionamiento de los modelos en su estimación de la distribución de contactos se centra, como vimos, en la detección de diferencias de comportamiento en función de si el plan de medios está formado sólo por diarios o por distintos tipos de revistas, cuestión que, al enunciar los objetivos, denominamos CI_{2B} .

Cuadro 5.12. Medias de lnMSEJ y lnEPORJ según el tipo de soporte.

Soporte	lnMSEJ		lnEPORJ	
	Media	D. típica	Media	D. típica
Diarios	-9,8814	2,4274	-0,5509	0,9505
R. Semanales	-9,9035	3,2007	-1,0654	1,1888
R. Mensuales	-10,8494	2,0588	-0,8268	0,8268

Fuente: elaboración propia

Cuadro 5.13. Análisis de varianza para las definiciones de error tomando como factor el tipo de soporte

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
lnMSEJ	Entre tipo soporte	2	32,76	4,82*
	Error	291	6,78	
lnEPORJ	Entre tipo soporte	2	7,82	7,85*
	Error	291	1,00	

* Diferencias significativas para $p \leq 0.01$ *Fuente: elaboración propia***Cuadro 5.14. Matriz de diferencias pareadas de lnMSEJ y lnEPORJ. Test de Tukey.**

Soporte	lnMSEJ		lnEPORJ	
	Mensuales	Semanales	Mensuales	Semanales
Semanales	-0,9459*		0,4586**	
Diarios	-0,9679*	-0,0220	-0,0560	-0,5146**

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: elaboración propia*

Los resultados obtenidos cuando los modelos estiman la distribución de contactos, confirman de una manera clara los que ya se obtuvieron en el apartado anterior. Aunque existen diferencias entre los errores cometidos cuando los planes de medios están formados por distinto tipo de soportes, estas diferencias no vienen ocasionadas por los periódicos. Como muestra el cuadro 5.14, el comportamiento de éstos es idéntico al de las revistas semanales cuando se utiliza como medida del error MSEJ, y análogo al de las revistas mensuales cuando se utiliza EPORJ. Luego no puede concluirse que los modelos funcionen peor cuando deben estimar cobertura y distribución de contactos de planes formados por periódicos, la explicación a estas diferencias, habrá que buscarla no en el tipo de soporte, sino en las características del plan de medios. Este es el objeto del epígrafe siguiente, la determinación de los factores de error de los modelos de acumulación de audiencias en su predicción de la

cobertura y la distribución de contactos, y que constituyen, como enunciamos en el capítulo cuarto, la cuestión a investigar CI_{2B} .

5.2.3 Factores de error de los modelos de acumulación de audiencias.

En epígrafes anteriores, se ha determinado qué modelos de acumulación de audiencias, funcionan mejor que otros en la estimación de la cobertura y la distribución de contactos, sin indagar en las razones que explican esas diferencias de comportamiento. El objetivo de este epígrafe, es determinar qué características de los planes de medios, entre las especificadas en el capítulo cuarto, determinan que cada modelo por separado, sea capaz de predecir mejor o peor las variables reseñadas. La información que se obtenga, unida a la jerarquización lograda en el epígrafe anterior, permitirá al planificador tener un criterio claro de selección del modelo más adecuado, atendiendo a su conocimiento general del plan de medios que pretenda evaluar. Pero, asimismo, nos permitirá conocer mejor el funcionamiento de los modelos, aportando directrices para la creación futura de nuevos modelos.

En una primera fase, se determinarán qué factores tienen una mayor influencia en el error que comete cada modelo³. Se analizarán el número de inserciones (IN), la audiencia bruta del plan de medios (AB), el tamaño relativo del soporte utilizado en el plan, respecto a todos los posibles (TG: tamaño grande; TM: tamaño mediano⁴), suma de las audiencias acumuladas (SAC) y, finalmente, la relación entre las audiencias acumuladas y la audiencia bruta (RACAB). Las conclusiones de este análisis, se aplicarán a continuación a aquellos modelos que en

³ La definición detallada de estas variables, ya se ofreció en el epígrafe 4.2.4, remitiendo a las expresiones [4.12] a [4.18] para obtener más información sobre su construcción.

⁴ Como se indicó en el capítulo cuarto, esta variable se ha introducido como una *dummy*, por lo cual es necesario eliminar uno de los niveles, el tamaño pequeño, para evitar la multicolinealidad perfecta y, a continuación, interpretar los coeficientes de las variables que se mantienen en el análisis en términos relativos al nivel suprimido.

el análisis realizado en los epígrafes precedentes, demostraron un mejor capacidad predictiva para establecer los criterios de selección que pueda utilizar el planificador y cumplimentar así, la cuestión a investigar CI_{2c} .

Los cuadros 5.15 y 5.16 ofrecen los coeficientes estandarizados que resultan significativos, de las variables señaladas, obtenidas al regresar el error en la estimación de la cobertura ($\ln AER$ y $\ln RER$) utilizándolas a ellas como regresores, de acuerdo con las expresiones:

$$\ln AER = a_1 + \beta_1 IN + \beta_2 AB + \beta_3 TG + \beta_4 TM + \beta_5 SAC + \beta_6 RACAB$$

$$\ln RER = a_1 + \beta_1 IN + \beta_2 AB + \beta_3 TG + \beta_4 TM + \beta_5 SAC + \beta_6 RACAB$$

Como puede comprobarse, para la mayoría de los modelos, la capacidad explicativa del error atribuible a estas variables es elevado y, en todos los casos, significativa la relación.

La principal conclusión a la que se puede llegar, es que el error depende fundamentalmente del tamaño de los planes de medios en términos de audiencia. En la medida en que este aumenta, bien porque el soporte utilizado tenga una audiencia elevada, o porque se realice un mayor número de inserciones en el mismo, el error cometido por la mayoría de los modelos se aumenta necesariamente. Como se observa en el cuadro 5.15, los coeficientes más elevados corresponden siempre a la audiencia bruta (logarítmico, Agostini, beta binomial, binomial negativa, beta binomial modificado y Markov) o al número de inserciones (exponencial y dos segmentos). A resultados muy similares se llega mediante la medida del error a través de RER, dado que la mayoría de variables distintas de estas dos, no sólo son menos importantes cuantitativamente que ellas, sino que además dejan de ser significativas.

Cuadro 5.15. Factores de error en la estimación de la cobertura: ln AER.
Coefficientes estandarizados.

	IN	AB	TM	TG	SAC	RACAB	R ²	F
Logarítmico		0,8230**	0,232*	0,519**			0,744	20,358**
Exponencial	0,736**			0,406*			0,605	10,707*
Agostini		0,6770*				0,586*	0,533	8,0010**
Hofmans							0,280	2,7160*
2 segmentos	0,431**		0,1840			0,409*	0,803	28,572**
3 segmentos		0,595*	0,268*	0,488*		0,681**	0,646	12,775**
Beta binomial		0,770**	0,287**	0,357**			0,743	20,222**
Binomial negativa	0,420**	0,898**	0,296**	0,500**	-0,405**	-0,394**	0,897	60,642**
Morgensztern			0,449**				0,596	7,1250**
Beta binomial modificado		0,684*					0,429	5,2560**
Markov	0,350**	0,919**	0,311**	0,538**	-0,381*		0,879	50,619**

** Significativo para $p \leq 0,01$

* Significativos para $p \leq 0,05$

Fuente: Elaboración propia

Cuadro 5.16. Factores de error en la estimación de la cobertura: ln RER.
 Coeficientes estandarizados.

	IN	AB	TM	TG	SAC	RACAB	R ²	F
Logarítmico								1,2650
Exponencial	0,770*						0,587	9,930**
Agostini								1,5120
Hofmans								1,7760
2 segmentos	0,474**	-0,492*					0,806	29,107**
3 segmentos						0,656*	0,331	3,4645**
Beta binomial								1,2400
Binomial negativa	0,992**					-1,362**	0,857	41,858**
Morgensztern								1,1780
Beta binomial modificado								2,1170
Markov	1,224**					-0,676**	0,626	11,725**

** Significativo para $p \leq 0,01$

* Significativos para $p \leq 0,05$

Fuente: Elaboración propia

Debe llamarse la atención sobre el hecho de que, en dos de los modelos que mejor comportamiento demostraron en análisis precedentes, el modelo de Hofmans y el beta binomial modificado, su mejor capacidad predictiva deriva, precisamente, de que o no son sensibles a la influencia del tamaño del plan (modelo de Hofmans), o bien esta influencia es relativamente baja en comparación con la que ejercen sobre el resto de modelos (modelo beta binomial modificado). Obsérvese en el cuadro 5.15 que ningún coeficiente, salvo la constante, resulta significativo en el modelo de Hofmans y que el coeficiente de la audiencia bruta en el beta binomial modificado, además de sólo ser significativo para $p \leq 0.05$, es más pequeño que en el resto de modelos.

Un comportamiento ligeramente diferente al de los dos grupos especificados, lo constituye el modelo de tres segmentos. Aunque la audiencia bruta tiene una influencia importante sobre el error, es más relevante la ejercida por la *ratio* entre ésta y la audiencia acumulada, es decir que el error crece cuando una mayor proporción de la audiencia es fiel al soporte y se acumula a mayor velocidad fruto de las inserciones sucesivas.

Los resultados que se obtienen al analizar los factores de error que afectan a la estimación de la distribución de contactos, son muy similares. Si el error se mide a través de MSEJ (cuadro 5.16), es la audiencia bruta quien ocasiona un mayor crecimiento del mismo, mientras que si la medida es EPORJ (cuadro 5.17), es el número de inserciones. En todo caso, la estimación de la cobertura de los modelos, es muy sensible al tamaño de los planes.

El modelo que mejores resultados daba en análisis precedentes, el modelo beta binomial modificado, destaca de nuevo por su comportamiento particular. Su error no depende de ninguna de estas variables cuando se mide mediante EPORJ y, cuando se recurre a MSEJ, aumenta cuando lo hace la audiencia bruta, pero disminuye al aumentar el número de inserciones. Esto lo hace mucho más adecuado

para aquellos planes de medios caracterizados por muchas inserciones en soportes pequeños que cuando éstos están formados por soportes que proporcionan audiencias elevadas con pocas inserciones.

La principal implicación para el planificador que, en conclusión, puede derivarse de los análisis realizados, es la siguiente: tanto cuando se pretenda calcular la cobertura del plan de medios que se quiera evaluar, como su distribución de contactos, el modelo recomendado es el beta binomial modificado. Dos son las razones que justifican esta afirmación. Por un lado, garantiza que el error promedio en ambas variables, será menor que si se recurre a cualquier otro modelo y, además, es insensible a los factores que provocan incrementos relevantes del error en otros modelos, fundamentalmente el tamaño del plan de medios, ya sea a través del número de inserciones, como de la audiencia bruta del mismo.

Cuadro 5.17. Factores de error en la estimación de la distribución de contactos: lnMSEJ.
Coeficientes estandarizados.

	IN	AB	TM	TG	SAC	RACAB	R ²	F
2 segmentos	0,364**	0,511**	0,150*	0,274*		0,565**	0,857	41,880**
3 segmentos		0,869**	0,217*			0,603**	0,821	32,088**
Beta binomial	-0,363**	0,668**	0,233**	0,459**		0,234*	0,923	83,995**
Binomial negativa		0,794**	0,274**	0,458**	-0,329**	-0,276**	0,933	96,885**
Beta binomial modificado	-0,322**	0,813**	0,278**	0,428**	-0,353*		0,904	66,213**
Markov	-0,161**	0,870**	0,290**	0,511**	-0,340*	-0,436**	0,925	86,041**

** Significativo para $p \leq 0,01$ * Significativos para $p \leq 0,05$ *Fuente: Elaboración propia*

Cuadro 5.18. Factores de error en la estimación de la distribución de contactos: lnEPORJ.
Coeficientes estandarizados.

	IN	AB	TM	TG	SAC	RACAB	R ²	F
2 segmentos	0,741**						0,880	51,460**
3 segmentos	0,637**						0,736	19,546**
Beta binomial	-0,381**						0,572	9,3470**
Binomial negativa	0,681**					-1,363**	0,923	96,369**
Beta binomial modificado							0,489	6,6910**
Markov	0,573**						0,433	5,3480**

** Significativo para $p \leq 0,01$ * Significativos para $p \leq 0,05$ *Fuente: Elaboración propia*

5.3 EVALUACIÓN DE LOS MODELOS DE DUPLICACIÓN DE AUDIENCIAS.

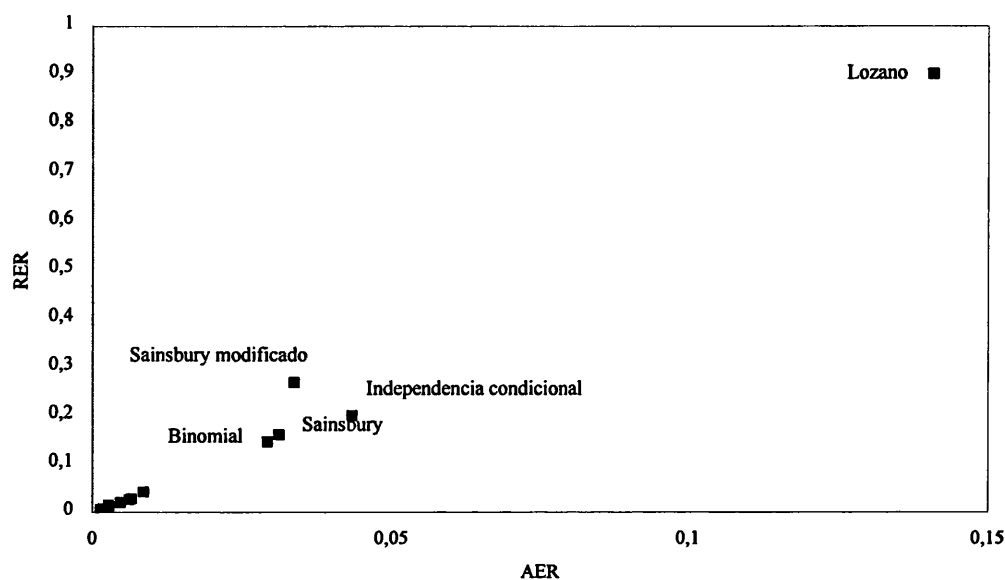
5.3.1 Error en la estimación de la cobertura.

La primera diferencia que se aprecia en el resultado de la evaluación de los modelos de duplicación, frente a los que se han presentado de los modelos de acumulación de audiencias, al proceder al desarrollo de la cuestión CI_{1A} es que gozan de un comportamiento mucho más homogéneo. Como se aprecia en el cuadro 5.19, y gráficamente en la figura 5.7, no existe un claro grupo de *outliers* cuyo comportamiento es radicalmente peor que el resto. El único modelo que destaca negativamente en este sentido, es el de Lozano pero, dado que no ha dificultado el proceso de estabilización de la varianza, puesto que una vez tomados los logaritmos naturales, los datos verificaban la hipótesis de homoscedasticidad, no se ha excluido del análisis. En todo caso, se ha optado por suprimirlo de las representaciones gráficas para facilitar, de este modo, que se ganase claridad en las mismas pero, reiteramos, está presente en todos los tratamientos realizados.

Cuadro 5.19. Errores promedio en la estimación de la cobertura. Identificación de *outliers*. N = 99.

Modelo	AER		RER	
	Media	D. Típica	Media	D. Típica
Agostini	0,0028	0,0037	0,0137	0,0098
Hofmans	0,0019	0,0046	0,0085	0,0115
Starch	0,0061	0,0126	0,0269	0,0532
Lozano	0,1407	0,0975	0,9025	0,1394
Binomial	0,0293	0,0397	0,1470	0,1011
Sainsbury	0,0312	0,0399	0,1613	0,1127
Kwerel	0,0047	0,0088	0,0212	0,0196
Sainsbury modificado	0,0337	0,0423	0,2695	0,3319
Agregación sec. dup. constante	0,0085	0,0138	0,0425	0,0558
Agregación sec. dup. variable	0,0065	0,0141	0,0287	0,0553
Independencia condicional	0,0434	0,0694	0,2006	0,1724
Hofmans ampliado	0,0027	0,0067	0,0156	0,0494
Metheringham	0,0014	0,0032	0,0063	0,0087

Fuente: Elaboración propia.

Figura 5.7. Identificación de *outliers*

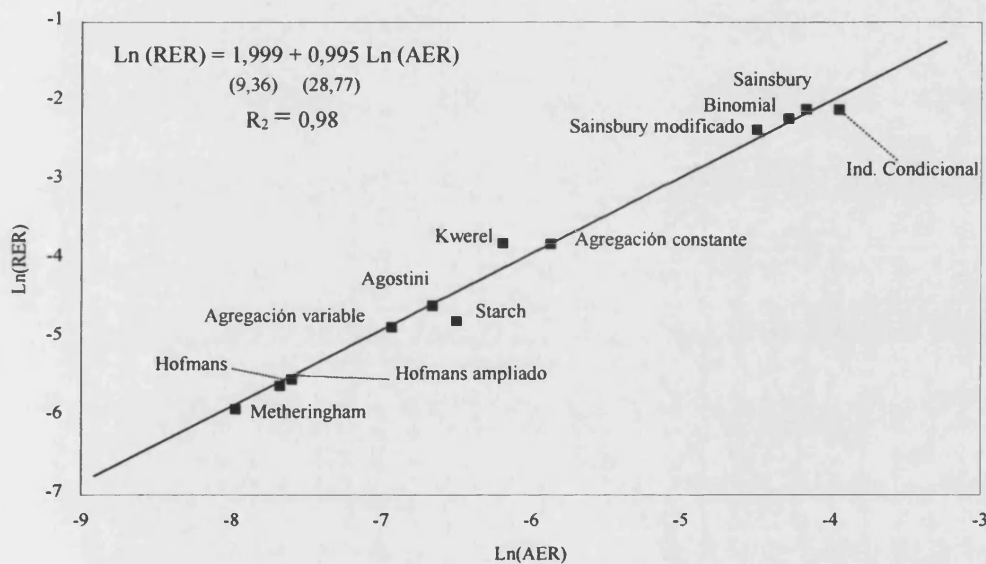
Fuente: Elaboración propia

Debemos señalar que, por motivos distintos, se ha eliminado del análisis a uno de los modelos *ad hoc* que se desarrollaron en el capítulo tercero, el modelo de Little y Lodish (1969). Como pudo comprobarse en el epígrafe 3.2.1.3, los autores no derivaban una expresión que permitiese calcular la cobertura de un plan de medios que incluyera más de dos soportes, mientras que todos los planes que se han diseñado para evaluar al resto, contaban entre tres y cinco. No se ha creído ortodoxo modificar el carácter aleatorio del proceso de selección incorporando *ad hoc* planes de dos soportes que permitiera evaluar un modelo, puesto que este hecho podría producir resultados engañosos, al no estar todos los modelos comparados sobre la misma base común de planes.

Al igual que ocurría con los modelos de acumulación de audiencias, las dos medidas de error en la estimación de la cobertura empleadas, ofrecen resultados simétricos. Como muestra la figura 5.8, ambos tipo de error se ajustan de manera casi perfecta a una recta, con lo que uno de ellos sería, prácticamente, una

combinación lineal del otro. Aunque, por ello, la utilización de las dos medidas de error no va a aportar más información que si recurriésemos sólo a una de ellas, se mantendrán ambas por mantener la simetría con el epígrafe anterior y, también, seguir la recomendación de Leckenby y Kim (1992) ya señalada en el capítulo cuatro de emplear, cuando menos, dos medidas del error de la cobertura y dos de la distribución de contactos.

Figura 5.8. Coherencia de los errores tipo AER y RER



Fuente: Elaboración propia

Entrando ya en los resultados del proceso de evaluación, el análisis de varianza realizado utilizando como variable dependiente los dos tipos de error, y como factor cada uno de los modelos, demuestran que existen diferencias significativas entre los errores que cometen cada uno de ellos (Cuadro 5.20). De nuevo, lo más relevante, es conocer qué modelos provocan esas diferencias y, sobre todo, qué grupos de modelos tienen un comportamiento homogéneo entre sí y distinto del resto para, de este modo, facilitar su jerarquización (cuestión a investigar CI_{1B}). Como ya hicimos en el epígrafe anterior, se ha recurrido al análisis *post hoc* de

Tukey para indagar qué diferencias de medias son significativas. Los resultados, que comentaremos a continuación, se sintetizan en los cuadros 5.21 y 5.22 para las medidas de AER y RER respectivamente.

Cuadro 5.20. Análisis de varianza para las definiciones de error

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
ln AER	Entre modelos	12	312,12	120,51*
	Error	1.274	2,59	
ln RER	Entre modelos	12	312,13	185,90*
	Error	1.274	1,68	

* Significativo para $p \leq 0,01$

Fuente: elaboración propia

Estas pruebas nos permiten identificar tres grandes grupos en función del error cometido, además del modelo de Lozano, de comportamiento mucho peor que el resto:

- a. Grupo de *error bajo*, formado por los modelos de Hofmans, Hofmans ampliado y, sobre todo, el modelo de Metheringham.
- b. Grupo de *error medio*, en el que se incorporaría a los modelos de agregación secuencial con duplicaciones constante y variable y a los modelos de Agostini, Kwerel y Starch.
- c. Grupo de *error alto*, formado por los modelos de Sainsbury y Sainsbury modificado, el modelo binomial y el modelo de independencia condicional.

Cuadro 5.21. Matriz de diferencias pareadas de LNAER. Test de Tukey.

	METH	HOF	HOFAM	DUPV	STAR	AGOS	KWE	DUPC	SAIMD	BIN	SAI	INDCD
Hofmans	-0,2971											
Hofmans ampliada	-0,3764	-0,0792										
Duplicación variable	-1,0427**	-0,7455	-0,6663									
Starch	-1,1212**	-0,8240*	-0,7448	0,0785								
Agostini	-1,3163**	-1,0191**	-0,9399**	-0,2736	-0,1951							
Kwerel	-1,7843**	-1,4872**	-1,4080**	-0,7417	-0,6631	-0,4681						
Duplicación constante	-2,1014**	-1,8043**	-1,7250**	-1,0587**	-0,9802**	-0,7851*	-0,3171					
Sainsbury modificado	-3,5508**	-3,2536**	-3,1725**	-2,5081**	-2,4295**	-2,2345**	-1,7664**	-1,4493**				
Binomial	-3,6907**	-3,3936**	-3,3143**	-2,6480**	-2,5695**	-2,3744**	-1,9064**	-1,5893**	-0,1400			
Sainsbury	-3,8094**	-3,5122**	-3,4330**	-2,7667**	-2,6882**	-2,4931**	-2,0250**	-1,7080**	-0,2586	-0,1187		
Indep. condicional	-3,9377**	-3,6405**	-3,5613**	-2,8950**	-2,8165**	-2,6214**	-2,1533**	-1,8362**	-0,3869	-0,2470	-0,1283	
Lozano	-5,7955**	-5,4983**	-5,4191**	-4,7528**	-4,6743**	-4,4792**	-4,0111**	-3,6941**	-2,2447**	-2,1048**	-1,9861**	-1,8578**

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$

Fuente: Elaboración propia

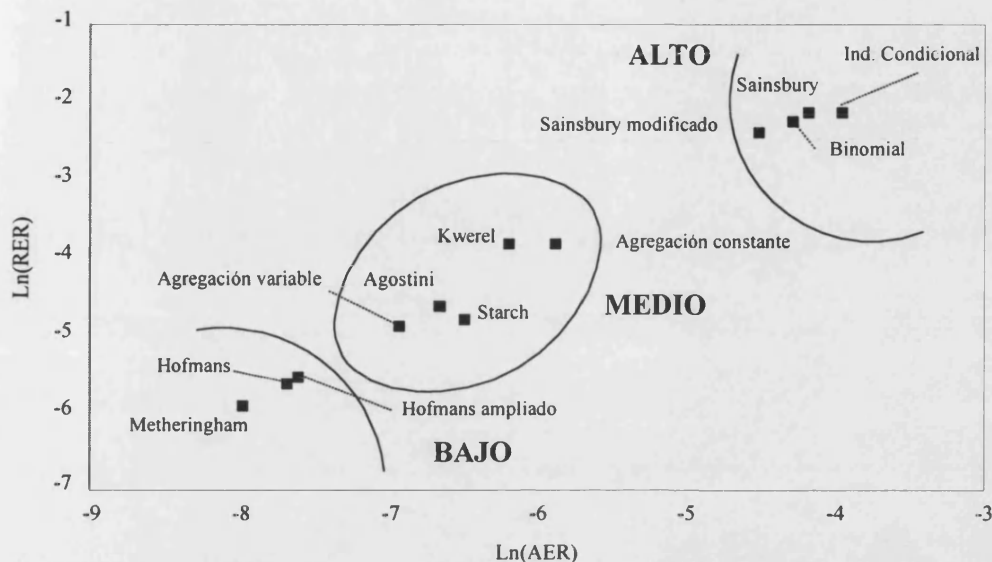
Cuadro 5.22. Matriz de diferencias pareadas de LNRER. Test de Tukey.

	METH	HOF	HOFAM	DUPV	STAR	AGOS	KWE	DUPC	SAIMD	BIN	SAI	INDCD
Hofmans	-0,2971											
Hofmans ampliada	-0,3763	-0,0792										
Duplicación variable	-1,0426**	-0,7455	-0,6663									
Starch	-1,1212**	-0,8241**	-0,7448	0,0785								
Agostini	-1,3163**	-1,0191**	-0,9399**	-0,2736	-0,1950							
Kwerel	-1,7843**	-1,4872**	-1,4080**	-0,7416	-0,6631	-0,4680						
Duplicación constante	-2,1014**	-1,8043**	-1,7250**	-1,0587**	-0,9802**	-0,7851*	-0,3170					
Sainsbury modificado	-3,5508**	-3,2536**	-3,1725**	-2,5081**	-2,4295**	-2,2345**	-1,7664**	-1,4493**				
Binomial	-3,6907**	-3,3936**	-3,3143**	-2,6480**	-2,5695**	-2,3744**	-1,9064**	-1,5893**	-0,1399			
Sainsbury	-3,8094**	-3,5122**	-3,4330**	-2,7667**	-2,6882**	-2,4931**	-2,0250**	-1,7080**	-0,2586	-0,1186		
Indep. condicional	-3,9377**	-3,6405**	-3,5613**	-2,8950**	-2,8165**	-2,6214**	-2,1533**	-1,8362**	-0,3869	-0,2469	-0,1282	
Lozano	-5,7955**	-5,4983**	-5,4191**	-4,7528**	-4,6743**	-4,4792**	-4,0111**	-3,6941**	-2,2447**	-2,1048**	-1,9861**	-1,8578**

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$

Fuente: Elaboración propia

**Figura 5.9. Jerarquización de los modelos de duplicación de audiencias.
Estimación de la cobertura.**



Fuente: Elaboración propia

En estos grupos, tal como ocurría con los modelos de acumulación de audiencias, existen modelos que presentan diferencias significativas con los que se han incluido en su mismo grupo. Como se indicó anteriormente, se ha buscado mantener una clasificación operativa a los efectos de selección de modelos y, la de tres grupos lo es. Cuando un modelo podía estar en dos grupos pero presentaba diferencias significativas con alguno de los presentes en ambos, se ha clasificado en aquel en el que la diferencia, aunque significativa, era más baja.

La clasificación obtenida, nos permite realizar algunas consideraciones acerca de la conveniencia en términos de capacidad predictiva de la cobertura, de introducir o no algunas de las hipótesis simplificadoras que subyacen bajo estos modelos y que, con detalle, se expusieron en el capítulo tercero (cuestión a investigar CI_{1c}). En síntesis:

- a. Al igual que ocurría con los modelos de audiencia acumulada, los modelos *ad hoc*, pese a su sencilla construcción, demuestran una elevada precisión en sus estimaciones de la cobertura. El modelo de Hofmans se encuentra entre los de error bajo y los de Starch y Agostini cometen un error medio. No hay modelos *ad hoc* que incurran en grandes errores al calcular la cobertura. Esta conclusión, explica, como se vio en el capítulo tercero, que la mayoría de los últimos modelos que se han desarrollado de audiencia neta acumulada, estimen la cobertura mediante fórmulas *ad hoc* como la de Hofmans y recurran a métodos estocásticos para la predicción de la distribución de contactos.
- b. No asumir la hipótesis de homogeneidad de la población, provoca mejoras sustantivas en las estimaciones de la cobertura. Aunque solamente un modelo introduce la heterogeneidad, el de Metheringham, es el que demuestra unos mejores resultados.
- c. Finalmente, el considerar el carácter aleatorio del fenómeno de la duplicación de las audiencias se demuestra como una hipótesis demasiado restrictiva, dado que los dos modelos que la incorporan están entre los que cometen un error elevado: Binomial y Sainsbury.

Como hicimos en el epígrafe anterior para contestar a la cuestión a investigar CI_{2B} , resta por analizar si los resultados confirman que, aunque inicialmente todos estos modelos se evaluaron con planes de medios formados únicamente por revistas, lo que ha llevado a afirmar que están diseñados exclusivamente para este medio, el comportamiento es o no peor cuando se trata de estimar la cobertura de planes en los que se incluyen diarios.

Cuadro 5.23. Medias de lnAER y lnRER según el tipo de soporte.

Soporte	lnAER		lnRER	
	Media	D. típica	Media	D. típica
Diarios	-6,5841	2,2188	-4,3895	2,1701
R. Semanales	-4,7751	2,1103	-3,2314	1,9022
R. Mensuales	-5,7599	2,3406	-3,7054	2,2321

Fuente: elaboración propia

Cuadro 5.24. Análisis de varianza para las definiciones de error tomando como factor el tipo de soporte

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
lnAER	Entre tipo soporte	2	415,12	85,77*
	Error	1.284	4,84	
lnRER	Entre tipo soporte	2	132,87	30,363*
	Error	1.284	4,38	

* Diferencias significativas para $p \leq 0.01$

Fuente: elaboración propia

Cuadro 5.25. Matriz de diferencias pareadas de lnAER y lnRER. Test de Tukey.

Soporte	lnAER		lnRER	
	Mensuales	Semanales	Mensuales	Semanales
Semanales	-1,4934**		-0,4739**	
Diarios	0,3157	1,4934**	0,6840**	1,1581**

** diferencias significativas para $p \leq 0,01$

* diferencias significativas para $p \leq 0,05$

Fuente: elaboración propia

Los resultados sintetizados en los cuadros anteriores demuestran, de una manera incluso más clara que en el análisis realizado de los modelos de acumulación, que la adecuación de los modelos descritos al medio diarios es completa. Así se observa que la media del error cometido en este medio, es inferior que en las revistas mensuales y semanales (cuadro 5.23). Asimismo, siendo las diferencias estadísticamente significativas (cuadro 5.24), lo son entre todos los tipos de soportes

(cuadro 5.25), no encontrándose los periódicos enfrentados a un grupo común formado por las revistas. La única diferencia no significativa aparece al considerar el error tipo AER y agruparía, además, a los diarios y revistas mensuales frente a las revistas semanales, no a estas últimas frente a los diarios. Este último resultado nos hace suponer, de nuevo, que las diferencias no aflorarán del medio utilizado en los planes sino, más bien, de otras características internas de los planes (número de soportes, nivel de duplicación, audiencia bruta, etc.) que serán analizados con posterioridad.

5.3.2 Error en la estimación de la distribución de contactos

Antes de comentar, para dar respuesta a las cuestiones a investigar CI_{IA} y CI_{IB} los resultados obtenidos en el proceso de evaluación y jerarquización de los modelos de duplicación de audiencias, respecto a la precisión que demuestran en el cálculo de la distribución de contactos, debe realizarse solamente una precisión. Como se detalló suficientemente en el capítulo tercero, los modelos de Sainsbury y Sainsbury modificado, difieren únicamente en el procedimiento empleado para estimar la cobertura. La forma en que operan para calcular la distribución de contactos es idéntica. Por ello, carecería de sentido mantenerlos como dos modelos distintos cuando, para la variable analizada, no lo son. Sin embargo, se ha decidido, en aras a conservar la simetría con el apartado anterior, mantener a ambos modelos en los cuadros y tablas que vendrán a continuación, aunque no se hará referencia a ellos como dos modelos distintos en los comentarios que procedan.

El primer hecho que merece ser destacado, y que se deriva directamente de la comparación de los cuadros 5.19 y 5.26, es que, haciendo abstracción de los modelos *ad hoc*, la ordenación de los modelos estocásticos de duplicación de audiencias es, prácticamente, la misma cuando se considera su capacidad para estimar la distribución de contactos, que lo era en la estimación de la cobertura. La única

diferencia, aunque mínima, aparece por el hecho de que el modelo de Sainsbury - Sainsbury modificado, funciona peor que el binomial en la estimación de la distribución de contactos cuando, en la estimación de la cobertura, el de Sainsbury modificado daba mejores resultados que el binomial, aunque el de Sainsbury los daba, también, peores.

Cuadro 5.26. Errores promedio en la estimación de la distribución de contactos. Identificación de *outliers*. N = 99.

Modelo	MSEJ		EPORJ	
	Media	D. Típica	Media	D. Típica
Binomial	0,0016	0,0032	0,3423	0,2062
Sainsbury	0,0017	0,0031	0,3675	0,2062
Kwerel	0,0006	0,0030	0,1545	0,1746
Sainsbury modificado	0,0017	0,0031	0,3675	0,2062
Agregación sec. dup. constante	0,0009	0,0035	0,1960	0,2399
Agregación sec. dup. variable	0,0009	0,0369	0,1613	0,2411
Independencia condicional	0,0062	0,0154	0,5392	0,4416
Hofmans ampliado	0,0007	0,0033	0,1827	0,6935
Metheringham	0,0003	0,0018	0,0726	0,1547

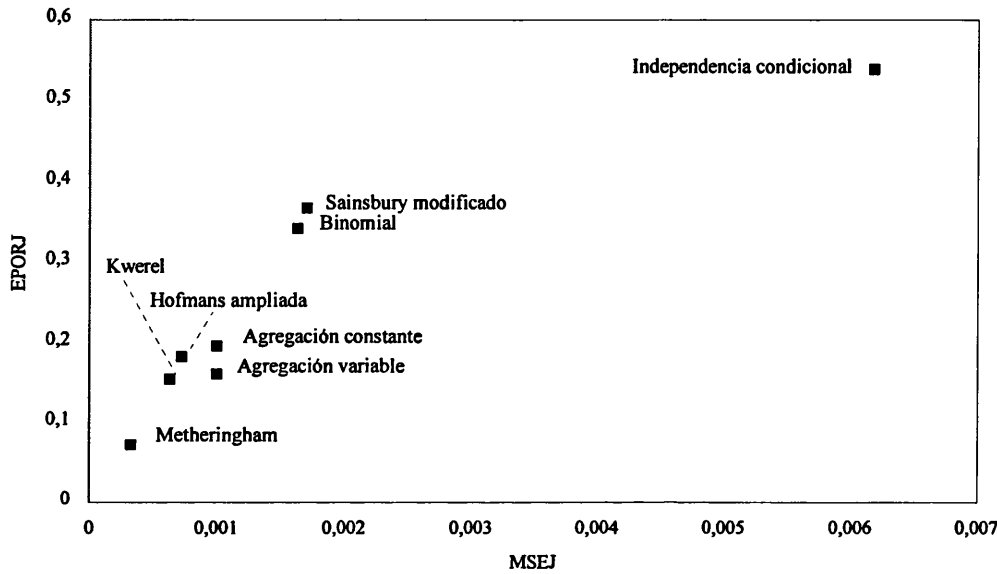
Fuente: Elaboración propia.

Siguiendo el esquema que, en el comentario de los resultados, iniciamos al comentar los modelos de acumulación, respecto a la identificación de los *outliers* nos encontramos ante el mismo caso que en la estimación de la cobertura. Existe un modelo, el de independencia condicional, que parece funcionar peor que el resto (véase la figura 5.10), pero dado que su existencia no afecta a la estabilización de la varianza, una vez se toman los logaritmos neperianos de MSEJ y EPORJ, se mantendrá en el análisis para retener la información que proporciona.

Respecto a los resultados que proporcionan las dos medidas de error utilizadas, nos encontramos con que, de nuevo, estas son totalmente paralelas, pudiendo considerarse una de ellas, sin que ello provocara diferencia alguna en el análisis, como transformación lineal de la otra. La figura 5.11 ilustra este hecho, que viene ratificado por un ajuste casi perfecto ($R^2 = 0.98$) de la recta de regresión entre

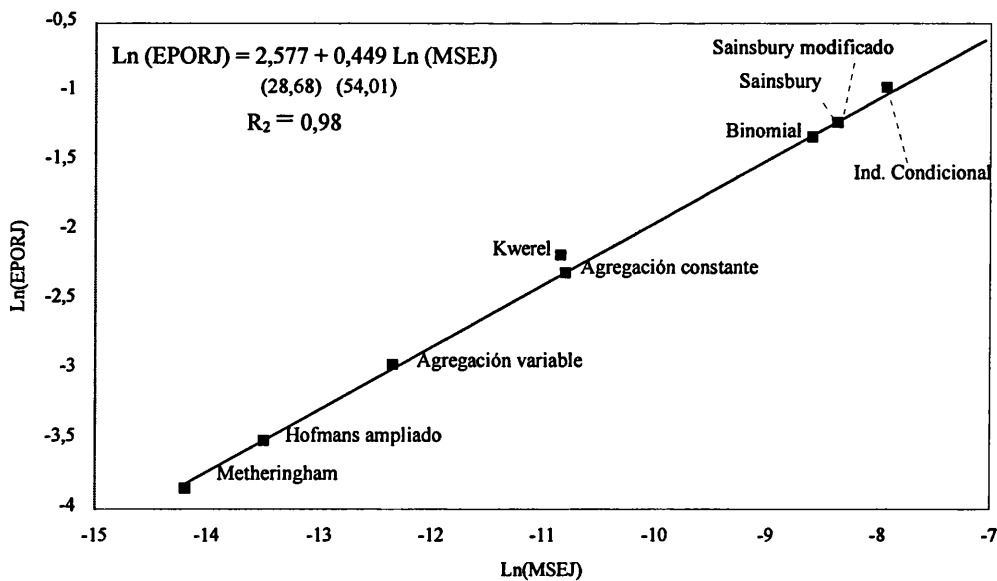
ambas variables. Pese a ello, y por las razones ya argumentadas con anterioridad, se presentarán por separado los resultados obtenidos por cada una de ellas.

Figura 5.10. Identificación de outliers



Fuente: Elaboración propia

Figura 5.11. Coherencia de los errores tipo MSEJ y EPORJ



Entrando ya en la evaluación y jerarquización de los modelos, el cuadro 5.27 demuestra la existencia de diferencias significativas entre el error cometido por los distintos modelos, mientras que en los cuadros 5.28 y 5.29, se pone de manifiesto la presencia de tres grupos en función del error que comenten en la estimación de la distribución de contactos, como se ilustra también en la figura 5.12:

- a. Grupo de *error bajo* formado, al igual que en la estimación de la cobertura, por los modelos de Metheringham y Hofmans ampliado al que se le añadiría el modelo de agregación secuencial con duplicación variable, que pasaría de tener un desempeño medio en la estimación de la cobertura, a ser de los más precisos en la estimación de la distribución de contactos.
- b. Grupo de *error medio*, en el que se incluirían los modelos de Kwerel y agregación secuencial con duplicación constante.
- c. Grupo de *error alto*, formado por el binomio de Sainsbury-Sainsbury modificado, el binomial y el de independencia condicional.

Cuadro 5.27. Análisis de varianza para las definiciones de error

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
ln MSEJ	Entre modelos	8	564,22	57,28*
	Error	882	9,85	
ln EPORJ	Entre modelos	8	114,34	74,09*
	Error	882	1,54	

* Significativo para $p \leq 0,01$

Fuente: elaboración propia

Cuadro 5.28. Matriz de diferencias pareadas de lnMSEJ. Test de Tukey.

	METH	HOFAM	DUPV	KWE	DUPC	BIN	SAI	SAIMD
Hofmans ampliada	-0,6960							
Duplicación variable	-1,8438**	-1,1478						
Kwerel	-3,3509**	-2,6548**	-1,5071**					
Duplicación constante	-3,3929**	-2,6969**	-1,5491**	-0,0420				
Binomial	-5,6031**	-4,9071**	-3,7593**	-2,2522**	-2,2102**			
Sanisbury	-5,8266**	-5,1306**	-3,9828**	-2,4747**	-2,4337**	0,2235		
Sainsbury modificada	-5,8266**	-5,1306**	-3,9828**	-2,4747**	-2,4337**	0,2235	0,0000	
Independencia condicional	-6,2649**	-5,5689**	-4,4211**	-2,9140**	-2,8720**	-0,6618	-0,4383	-0,4383

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: Elaboración propia*

Cuadro 5.29. Matriz de diferencias pareadas de InEPORJ. Test de Tukey.

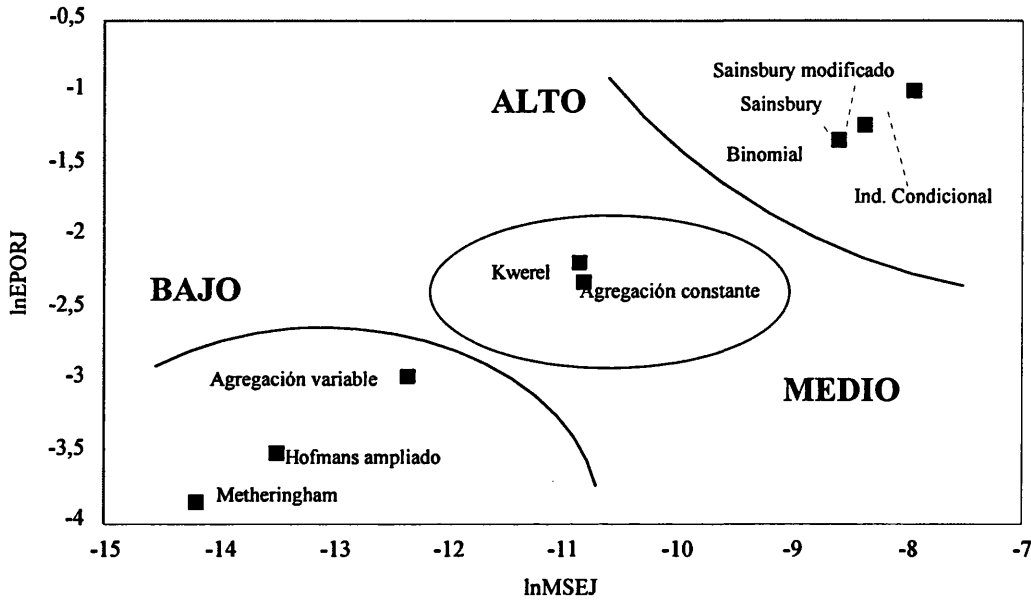
	METH	HOFAM	DUPV	KWE	DUPC	BIN	SAI	SAIMD
Hofmans ampliada	0,3423							
Duplicación variable	-0,8867**	-0,5444						
Kwerel	-1,6747**	-1,3323**	-0,7879**					
Duplicación constante	-1,5472**	-1,2049**	-0,6605**	0,1275				
Binomial	-2,5270**	-2,1847**	-1,6403**	-0,8523**	-0,9798**			
Sanisbury	-2,6318**	-2,2894**	-1,7451**	-0,9570**	-1,0845**	-0,1048		
Sainsbury modificada	-2,6318**	-2,2894**	-1,7451**	-0,9570**	-1,0845**	-0,1047	0,0000	
Independencia condicional	-2,8849**	-2,5426**	-1,9982**	-1,2102**	-1,3377**	-0,3579	-0,2531	-0,2531

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: Elaboración propia*

La formación de los grupos, en este caso, ha sido derivada con facilidad a partir de las diferencias de medias significativas, dado que los grupos homogéneos eran totalmente excluyentes. La única salvedad la constituye el modelo de agregación secuencial con duplicación variable cuyo error, no siendo significativamente distinto del modelo ampliado de Hofmans, sí que lo es del de Metheringham.

Dado que la clasificación en la estimación de la distribución de contactos, es la misma que se obtuvo respecto a la cobertura, las implicaciones respecto a la restrictividad de las hipótesis también son las mismas, pudiendo resumirse en que: la no asunción de la hipótesis de homogeneidad de la población deviene en mejoras sustantivas en la estimación de la distribución de contactos, dado que el único modelo que considera que ésta es heterogénea, el de Metheringham, es el que mejores resultados obtiene. Asimismo, se reitera la conclusión de que considerar el carácter aleatorio del fenómeno de la duplicación de las audiencias, provoca errores elevados, como lo demuestra el hecho de que los dos modelos que así lo consideran, binomial y Sainsbury, están entre los de peor desempeño.

**Figura 5.12. Jerarquización de los modelos de duplicación de audiencias.
Estimación de la distribución de contactos.**



Fuente: Elaboración propia

Respecto a las diferencias en el funcionamiento de los modelos atendiendo al tipo de medio que constituye el plan (cuestión CI_{2B}), los resultados de los cuadros 5.30, 5.31 y 5.32, nos confirman que, lejos de funcionar peor para los diarios, para los cuales no estaban diseñados en principio, es en este medio donde se dan los mejores resultados de los modelos considerados en su conjunto. De nuevo habrá que buscar en las características internas de los planes, los factores determinantes del error.

Cuadro 5.30. Medias de lnMSEJ y lnEPORJ según el tipo de soporte.

Soporte	lnMSEJ		lnEPORJ	
	Media	D. típica	Media	D. típica
Diarios	-12,4640	3,0797	-3,0164	0,0959
R. Semanales	-8,3574	3,4455	-1,5853	0,0736
R. Mensuales	-11,6366	3,5864	-2,2110	0,0934

Fuente: elaboración propia

Cuadro 5.31. Análisis de varianza para las definiciones de error tomando como factor el tipo de soporte

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
lnMSEJ	Entre tipo soporte	2	1.441,85	124,09*
	Error	888	11,62	
lnEPORJ	Entre tipo soporte	2	140,85	62,73*
	Error	888	2,25	

* Diferencias significativas para $p \leq 0.01$ *Fuente: elaboración propia***Cuadro 5.32. Matriz de diferencias pareadas de lnMSEJ y lnEPORJ. Test de Tukey.**

Soporte	lnMSEJ		lnEPORJ	
	Mensuales	Semanales	Mensuales	Semanales
Semanales	-3,2793**		-0,6257**	
Diarios	0,8273*	4,1067**	0,8053**	1,4311**

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: elaboración propia*

5.3.3 Factores de error en los modelos de duplicación de audiencias.

Al igual que se hizo en el epígrafe anterior, las páginas precedentes, han evaluado y ordenado a los modelos de duplicación de audiencias, atendiendo al error promedio que cometían en la estimación de la cobertura y la distribución de contactos de diferentes planes de medios. La hipótesis implícita en el diseño experimental anterior, es que todo el error cometido viene explicado por las características del modelo y que, en consecuencia, los resultados serían los mismos fuesen cuales fuesen las características de los planes utilizados en su evaluación. El objeto del presente epígrafe, es tratar de comprobar si el funcionamiento de los modelos depende, por el contrario, de cuáles sean las características internas del plan

de medios que tienen que evaluar en términos de cobertura o frecuencia y, de esta forma, establecer si existen unos modelos más adecuados que otros, no globalmente, sino para tipos de planes específicos. En este sentido, se formuló la cuestión a investigar CI_{2A} .

Como se detalló en el capítulo cuarto, siguiendo los trabajos de Leckenby y Kishi (1981; 1982b) se identificaron doce factores de error, de entre los cuales, dado que se están evaluando planes de medios formados por 1 inserción en n soportes, son relevantes los siguientes: audiencia bruta del plan de medios (AB), suma de las audiencias duplicadas (AD), ratio entre ambas variables (RADAB), número de soportes que conforman el plan de medios (SOP) y, finalmente, tamaño relativo de los soportes que forman parte del plan. Dado que en su momento se definieron detalladamente todos ellos, sólo se indicará que ésta última se ha construido dividiendo a los soportes en tres grupos según su audiencia por terciles y construyendo cinco variables ficticias según el plan de medios estuviese formado por soportes de los grupos de audiencia baja, media y alta (BMA), alta y media (AM), alta y baja (AB), media y baja (MB) o que todos perteneciesen al mismo grupo. Para evitar la multicolinealidad perfecta, ha sido este último el grupo suprimido, lo que debe tenerse en cuenta a la hora de interpretar los coeficientes.

Cuando se analizan los resultados de los factores determinantes del error en la estimación de la cobertura (Cuadros 5.33 y 5.34), se observan una serie de hechos destacados. Así, en primer lugar, se constata que la composición del plan de medios, en términos de tamaño relativo de los distintos soportes integrados, ejerce poca o ninguna influencia sobre el error. Son muy pocos los modelos en que sus coeficientes son significativos y, donde lo son, demuestran ejercer una influencia mucho menor que el resto de variables consideradas.

En segundo lugar, la *ratio* entre la audiencia duplicada y la audiencia bruta, demuestra ser el factor que más tiene que ver con el error cometido por el modelo en

la estimación de la cobertura. En la medida en que la audiencia duplicada representa un porcentaje mayor de la audiencia bruta del plan de medios, el error cometido crece.

A diferencia de lo que ocurría en los modelos de acumulación, la influencia de los distintos factores de error, es muy parecida en todos los modelos, lo que nos hace concluir que, en este caso, el que los modelos de Metheringham, Hofmans y Hofmans ampliado, demuestren un desempeño mejor, radica exclusivamente en la forma en que están contruidos, y que las características del plan de medios, al influir de manera paralela sobre todos los modelos, no pueden provocar que la clasificación que en su momento establecimos varíe sustancialmente. En todo caso, debe destacarse que, si es necesario elegir entre los tres modelos señalados que demostraron funcionar mejor (cuestión a investigar CI_{2c}), cuando se considere que los soportes que integran el plan de medios que se va a evaluar, van a traducirse en elevadas duplicaciones, deberá optarse siempre por el de Metheringham, dado que es el modelo en el que el error tiene una tendencia más acentuada a disminuir cuando aumenta el valor de esta variable.

Cuadro 5.33. Factores de error en la estimación de la cobertura: ln AER.
 Coeficientes estandarizados.

	SOP	AB	AD	RADAB	BMA	AM	AB	MB	R ²	F
Agostini		1,772**	-1,219**						0,751	33,96**
Hofmans	0,291**		-0,661**	0,903**					0,742	32,39**
Starch	0,254**	0,535**	-0,954**	1,030**					0,840	58,98**
Lozano	0,062*	1,535**	-1,167**	0,390**				-0,195**	0,958	254,44**
Binomial	-0,067*	0,709**	-1,246**	1,486**					0,954	233,63**
Sainsbury		0,749**	-1,272**	1,465**			0,126**		0,947	199,34**
Kwerel	0,2310**	1,207**							0,839	58,542**
Sainsbury modificado				0,659*			0,357*		0,235	3,45**
Duplicación constante	0,202*	1,032**	-1,144**	0,678**					0,619	18,26**
Duplicación variable	0,274**	0,461*	-0,912**	1,093**					0,774	38,61**
Independencia condicional		0,633**	-1,041**	1,359**			0,107**		0,958	259,29**
Hofmans ampliada	0,333**		-0,655**	0,912**					0,705	26,83**
Metheringham		1,115**	-1,193**	0,957**	0,397**	0,236**	0,255**		0,678	23,681**

** Significativo para $p \leq 0,01$

* Significativos para $p \leq 0,05$

Fuente: Elaboración propia

Cuadro 5.34. Factores de error en la estimación de la cobertura: ln RER.
 Coeficientes estandarizados.

	SOP	AB	AD	RADAB	BMA	AM	AB	MB	R ²	F
Agostini		1,468**	-1,298**						0,435	8,66**
Hofmans	0,290**			1,170**					0,639	19,94**
Starch	0,257**		-0,874**	1,336**					0,760	35,63**
Lozano	-0,353**	-0,483*	-1,942**	2,645**					0,739	31,87**
Binomial	-2,201**		-1,202**	2,257**					0,888	89,32**
Sainsbury	-0,165**		-1,269**	2,317**					0,866	72,92**
Kwerel	0,257*								0,538	13,13**
Sainsbury modificado		-0,9900		0,713*			0,333*		0,276	4,29**
Duplicación constante			-1,105**	1,000**					0,371	6,63**
Duplicación variable	0,272**		-0,804**	1,346**					0,693	25,42**
Independencia condicional	-0,114*	-0,282*	-0,961**	2,102**					0,902	103,23**
Hofmans ampliada	0,335**			1,143**					0,602	17,03**
Metheringham		0,670*	-1,164**	1,281**	0,500**	0,284**	0,299*		0,500	11,24**

** Significativo para $p \leq 0,01$

* Significativos para $p \leq 0,05$

Fuente: *Elaboración propia*

Cuando se trata de analizar el error cometido al estimar la distribución de contactos, los resultados son muy similares (cuadros 5.35 y 5.36), pero requieren de alguna matización. También es cierto que la composición de los planes de medios en relación al tamaño relativo de los soportes que lo integran demuestra ejercer poca o nula influencia sobre el error cometido por cada modelo. Asimismo, también es cierto que la *ratio* entre la audiencia duplicada y la audiencia bruta del plan de medios, es la variable determinante del error. Sin embargo, sí que puede encontrarse una explicación a que tres modelos: duplicación variable, Hofmans ampliado y Metheringham tengan una mejor capacidad predictiva que el resto: son los modelos en los cuales la influencia de esta variable, medida por el coeficiente estandarizado, es sustancialmente menor que en el resto. Así, tomando como ejemplo el error medido mediante el \ln MSEJ, el coeficiente es de 0.814, 0.872 y 1.065, respectivamente para Metheringham, Hofmans ampliado y agregación secuencial con duplicación variable, cuando el resto de modelos toma valores de 1.592 para el binomial, 1.582 para el de Sainsbury o 1.442 para el de independencia condicional. Los resultados son análogos cuando se toma como medida del error, el \ln EPORJ.

En síntesis, podemos concluir que, mientras que ninguna variable inherente a las características de cada plan de medios, explica que unos modelos estimen mejor que otros la cobertura del mismo, sí que se demuestra que aquellos modelos con menor sensibilidad a la *ratio* entre la audiencia duplicada y la bruta del plan, son capaces de estimar mejor la distribución de contactos. El hecho de que el modelo de Metheringham haya demostrado cometer un error que no es significativamente distinto del de Hofmans ampliado, se explica por la similitud de los coeficientes de RADAB (0.814 vs. 0.872) para el \ln MSEJ. Por contra, en la medida en que este coeficiente es claramente menor que el del modelo de agregación secuencial con duplicación variable (1.442), la diferencia entre los errores, devenía significativa, como se pudo comprobar en el cuadro 5.28.

Cuadro 5.35. Factores de error en la estimación de la distribución de contactos: In MSEJ.
Coefficientes estandarizados.

	SOP	AB	AD	RADAB	BMA	AM	AB	MB	R ²	F
Binomial	-0,254**	0,727**	-1,316**	1,592**					0,937	167,01**
Sainsbury	-0,230**	0,778**	-1,362**	1,582**					0,929	146,52**
Kwerel	0,222**	1,221**					0,164*		0,845	61,30**
Sainsbury modificado	-0,230**	0,778**	-1,362**	1,582**			0,149**		0,929	146,52**
Duplicación constante	0,254**	0,882**	-0,941**	0,780**		0,138*			0,783	40,55**
Duplicación variable	0,299**	0,528**	-0,930**	1,065**					0,812	48,59**
Independencia condicional	-0,150**	0,639**	-1,066**	1,442**			0,122**		0,957	248,25**
Hofmans ampliada	0,403**		-0,596**	0,872**					0,735	31,16**
Metheringham	0,186*	0,912**	-0,916**	0,814**	0,375**	0,232**	0,238*		0,678	23,71**

** Significativo para $p \leq 0,01$

* Significativos para $p \leq 0,05$

Fuente: Elaboración propia

Cuadro 5.36. Factores de error en la estimación de la distribución de contactos: lnEPORJ.
Coefficientes estandarizados.

	SOP	AB	AD	RADAB	BMA	AM	AB	MB	R ²	F
Binomial		-0,451*	-1,155**	2,423**					0,827	53,77**
Sainsbury	-0,475**	-0,428*	-1,255**	2,523**			0,183*		0,805	46,47**
Kwerel	0,296**								0,600	16,91**
Sainsbury modificado	-0,475**	-0,428*	-1,255**	2,523**			0,183*		0,805	46,47**
Duplicación constante	0,273**		-0,829**	1,156**					0,656	21,46**
Duplicación variable	0,298**		-0,812**	1,302**					0,740	31,99**
Independencia condicional	-0,285**	-0,337*	-0,936**	2,223**			0,153*		0,903	104,61**
Hofmans ampliada	0,419**			1,064**					0,656	21,44**
Metheringham			-0,806**	1,081**	0,447**	0,263**	0,265*		0,546	13,54**

** Significativo para $p \leq 0,01$

* Significativos para $p \leq 0,05$

Fuente: Elaboración propia

5.4 EVALUACIÓN DE LOS MODELOS DE AUDIENCIA NETA ACUMULADA

En la evaluación y jerarquización de los modelos de exposición, se mantendrá también el esquema expositivo que se ha venido desarrollando hasta el momento. Así, en primer lugar se identificarán a aquellos modelos que, por tener ofrecer unas predicciones mucho peor que las del resto, dificultan el proceso de jerarquización al impedir que la varianza pueda estabilizarse mediante las transformaciones habituales: logaritmos neperianos, raíces cuadradas, arcosenos, etc. y provocando, con ello, que no pueda recurrirse a una prueba tan paramétrica, y por ello exigente en términos de homoscedasticidad de la población, como es el análisis de la varianza. Debe reiterarse, sin embargo, que la exclusión de los *outliers* no implica que estos modelos no resulten evaluados, en cumplimiento del objetivo planteado en la cuestión a investigar CI_{1A} , sin que el resultado de esa evaluación, es que su comportamiento, siempre es peor que la del resto de modelos.

A continuación, se comprobará la coherencia entre las distintas medidas del error empleadas. El objetivo de esta prueba, como ya se indicó, no es otro que la economía en los comentarios. Si ambas medidas ofrecen resultados paralelos, deviene innecesario distinguir las conclusiones obtenidas aplicando cada una de ellas. Esto sería inevitable de no existir esta coherencia, como ocurre, adelantamos, al analizar el error cometido en la estimación de la distribución de contactos.

Como se ha venido haciendo hasta el momento, mediante el análisis de la varianza y la prueba *post hoc* de Tukey, se determinará la significatividad de las diferencias de error cometido por cada modelo y se establecerán grupos que tengan un comportamiento homogéneo, de tal forma que esta jerarquía (cuestión CI_{1B}) deje patente qué modelos son susceptibles de ser utilizados en la práctica de la planificación, asegurando resultados ajustados. Asimismo, de acuerdo con la

cuestión CI_{1C} , y tomando como punto de partida los resultados anteriores, se analizará la conveniencia o no, para mejorar la capacidad estimativa de los distintos modelos, de levantar algunas de las hipótesis simplificadoras que los sustentan.

En último lugar, tras determinar si estos resultados son o no sensibles al tipo de medio que constituyen el plan: diarios, revistas semanales o mensuales (cuestión CI_{2B}), se establecerá la influencia que ejerce cada uno de los factores de error que se han identificado sobre el funcionamiento de los modelos analizados (cuestión CI_{2A}) y se establecerán los criterios para la selección del modelo más adecuado (CI_{2C}).

5.4.1 Error en la estimación de la cobertura

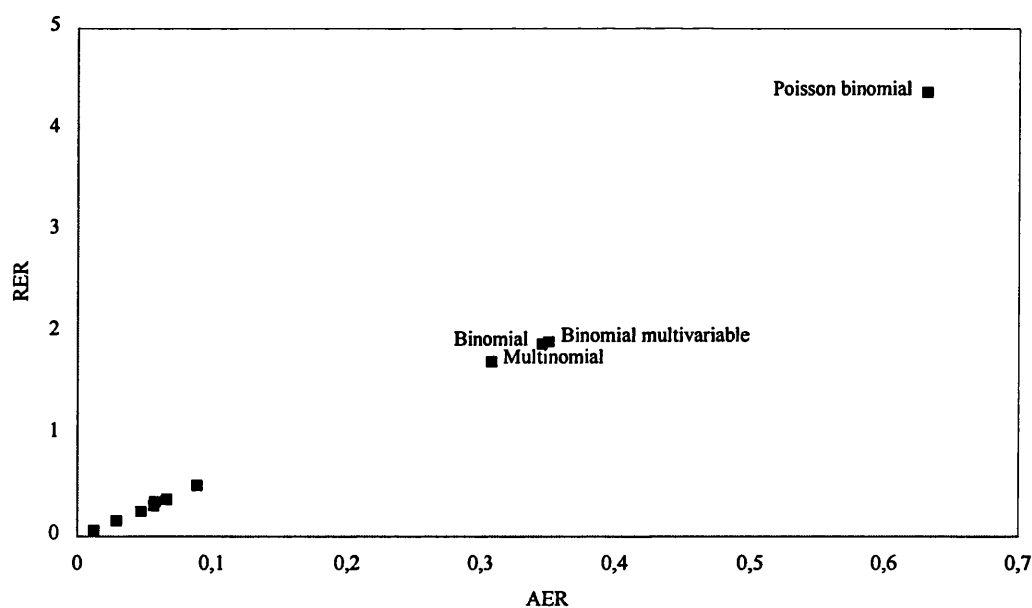
Como muestra el cuadro 5.37 e ilustra la figura 5.13, existe un grupo de modelos que ofrecen un resultado muy pobre en la estimación de la cobertura y que, constatado este hecho y siguiendo con la práctica habitual en este trabajo, no se incluirán en el análisis de la varianza. Son el modelo Poisson binomial, claramente el peor de ellos, y los modelos binomial, binomial multivariable y multinomial. Aunque luego se resaltaré este hecho con más profundidad, debe destacarse que son todos los modelos que utilizan la distribución binomial como base de su construcción, bien considerando soportes y población homogéneos, como se venía haciendo hasta el momento (modelo binomial), bien introduciendo la heterogeneidad de soportes (binomial multivariable y multinomial) e incluso aunque la distribución binomial sea utilizada únicamente para componerla con otra distribución de probabilidad (modelo Poisson binomial).

Cuadro 5.37. Errores promedio en la estimación de la cobertura. Identificación de *outliers*. N = 99.

Modelo	AER		RER	
	Media	D. Típica	Media	D. Típica
Binomial	0,3456	0,0841	1,8837	0,7716
Binomial multivariable	0,3504	0,0860	1,9066	0,7708
Multinomial	0,3078	0,0890	1,7104	1,0215
Markov-Bernouilli	0,0292	0,0241	0,1562	0,1293
Metherringham	0,0478	0,0439	0,2460	0,2098
Dirichlet multinomial	0,0664	0,0406	0,3624	0,2071
Poisson Binomial	0,6313	0,1428	4,3672	3,2295
Kwerel geométrica	0,0582	0,0355	0,3406	0,2586
Hofmans geométrica	0,0887	0,0307	0,5025	0,2754
Hofmans beta binomial	0,0887	0,0307	0,5025	0,2754
Hofmans Poisson binomial	0,0887	0,0307	0,5025	0,2754
Promedio Metherringham Dirichlet	0,0571	0,0388	0,3042	0,1865
Hofmans	0,0887	0,0307	0,5025	0,2754
Morgensztern	0,0123	0,0083	0,0629	0,0508

Fuente: Elaboración propia.

Figura 5.13. Identificación de *outliers*.

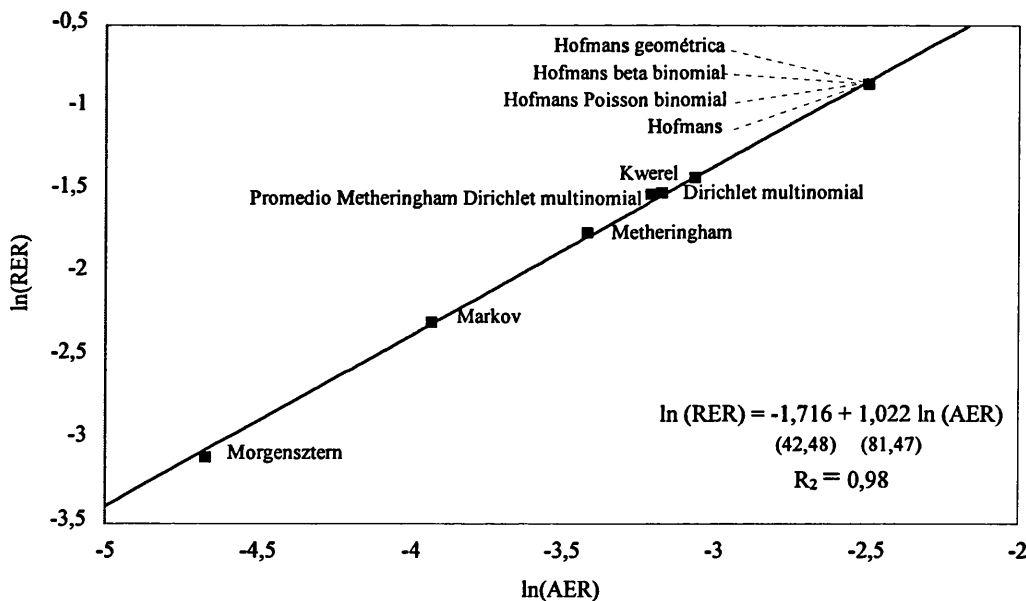


Fuente: Elaboración propia

Antes de pasar a comentar el resto de resultados, hay un hecho que debe destacarse por ejercer una notable influencia sobre los mismos. Entre los modelos de audiencia neta acumulada, como se detalló en el capítulo tercero, existe un grupo que

opta por estimar la cobertura del plan de medios mediante un modelo *ad hoc* para luego determinar la distribución de contactos mediante otro tipo de procedimientos, normalmente recurriendo a distribuciones de probabilidad como la beta binomial. En concreto, cuatro de los modelos analizados por nosotros recurren a la fórmula de Hofmans para estimar la cobertura: además del propio modelo de Hofmans, el Hofmans geométrico, el Hofmans beta binomial y el Hofmans Poisson binomial. Esto provoca que, a la hora de evaluar el error cometido en la estimación de la cobertura, tengamos cuatro modelos que ofrecen los mismos resultados lo que, entre otras implicaciones, favorece coherencia casi perfecta entre los errores tipo AER y RER, dado que diez medidas, cuatro están sobre la misma recta de regresión, como se aprecia en la figura 5.14.

Figura 5.14. Coherencia de los errores tipo AER y RER.



Fuente: Elaboración propia

De nuevo, como refleja la mencionada figura, los resultados que ofrecen las dos medidas del error en la estimación de la cobertura empleadas, son idénticos, pudiendo entenderse una como una transformación lineal de la otra. Este hecho,

como se ha indicado, hará que en la presentación de los resultados que siguen, se haga referencia indistintamente a una u otra.

Como demuestra el análisis de varianza cuyos resultados se sintetizan en el cuadro 5.38, puede afirmarse, sin lugar a dudas, que existen diferencias significativas entre el error cometido en la estimación de la cobertura por los distintos modelos, siendo procedente establecer una ordenación de los mismos en función de esta variable y cumplir así, los propósitos de la cuestión a investigar CI_{1B} .

5.38. Análisis de varianza para las definiciones de error

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
ln AER	Entre modelos	9	51,95	75,56*
	Error	980	0,69	
ln RER	Entre modelos	9	61,95	59,49*
	Error	980	0,87	

* Significativo para $p \leq 0,01$

Fuente: elaboración propia

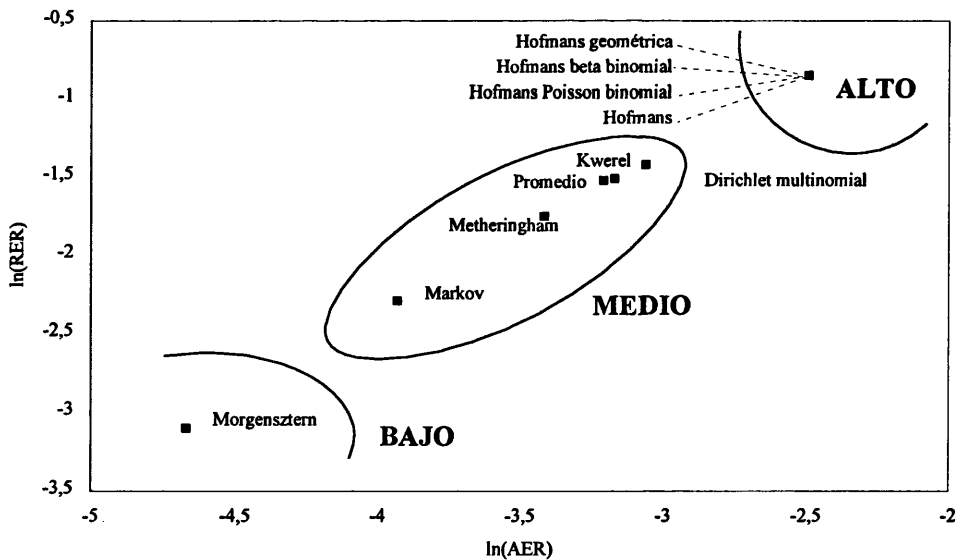
La prueba *post hoc* de Tukey, recogida en los cuadros 5.39 y 5.40, permite la identificación de tres grupos de modelos con un comportamiento homogéneo, dado que las medias de los errores que cometen no son significativamente distintas, tal y como se ilustra también en la figura 5.15:

- Grupo de *error bajo*, formado exclusivamente por el modelo de Morgensztern
- Grupo de *error medio*, en el que se incluirían a los modelos de Kwerel, Markov, Dirichlet multinomial, Metheringham y el modelo obtenido como promedio de estos dos últimos.
- Grupo de *error alto*, formado por todos los modelos que utilizan la fórmula de Hofmans para estimar la cobertura, esto es, el propio

modelo de Hofmans, el Hofmans beta binomial, el Hofmans geométrica y el Hofmans Poisson binomial.

Esta clasificación es muy estable, dado que la única aproximación que se ha realizado para mantener la estructura de tres grupos, ha sido la de incorporar el modelo de Markov al de error medio cuando, como puede comprobarse en los cuadros 5.39 y 5.40, la media de error cometido es significativamente distinta de la del resto de integrantes. Como también lo era de la cometida por el modelo de Morgensztern, se ha incluido en el grupo que incorporaba al modelo con la media más cercana a él.

Figura 5.15. Jerarquización de los modelos de audiencia neta acumulada. Estimación de la cobertura.



Fuente: Elaboración propia

Cuadro 5.39. Matriz de diferencias pareadas de LNAER. Test de Tukey.

	MOR	MKV	MET	PRO	DIR	KWE	HBBD	HPB	HGD
Markov Bernouilli	-0,7381**								
Metheringham	-1,2534**	-0,5153**							
Promedio Meth- Dirichlet	-1,4609**	-0,7723**	-0,2075						
Dirichlet multinomial	-1,4987**	-0,7606**	-0,2454	0,0379					
Kwerel geométrica	-1,6068**	-0,8687**	-0,3534	-0,1459	-0,1081				
Hofmans beta binomial	-2,1756**	-1,4375**	-0,9222**	-0,7147**	-0,6769**	-0,5688**			
Hofmans Poisson binomial	-2,1756**	-1,4375**	-0,9222**	-0,7147**	-0,6769**	-0,5688**	0,0000		
Hofmans geométrica	-2,1756**	-1,4375**	-0,9222**	-0,7147**	-0,6769**	-0,5688**	0,0000	0,0000	
Hofmans	-2,1756**	-1,4375**	-0,9222**	-0,7147**	-0,6769**	-0,5688**	0,0000	0,0000	0,0000

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: Elaboración propia*

Cuadro 5.40 Matriz de diferencias pareadas de LNRER. Test de Tukey.

	MOR	MKV	MET	PRO	DIR	KWE	HBBD	HPB	HGD
Markov Bernouilli	-0,7381**								
Metheringham	-1,2534**	-0,5153**							
Promedio Meth- Dirichlet	-1,4609**	-0,7723**	-0,2075						
Dirichlet multinomial	-1,4987**	-0,7606**	-0,2454	0,0379					
Kwerel geométrica	-1,6068**	-0,8687**	-0,3534	-0,1459	-0,1081				
Hofmans beta binomial	-2,1756**	-1,4375**	-0,9222**	-0,7147**	-0,6769**	-0,5688**			
Hofmans Poisson binomial	-2,1756**	-1,4375**	-0,9222**	-0,7147**	-0,6769**	-0,5688**	0,0000		
Hofmans geométrica	-2,1756**	-1,4375**	-0,9222**	-0,7147**	-0,6769**	-0,5688**	0,0000	0,0000	
Hofmans	-2,1756**	-1,4375**	-0,9222**	-0,7147**	-0,6769**	-0,5688**	0,0000	0,0000	0,0000

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: Elaboración propia*

De la clasificación obtenida, se pueden derivar una serie de conclusiones acerca de la rentabilidad en términos de mejora de la predicción, de complejizar la construcción de los modelos levantando algunas de las hipótesis simplificadoras. Queremos destacar las siguientes:

- a. De nuevo un modelo *ad hoc* es el que ofrece mejor estimación de la cobertura del plan de medios. Este hecho se ha venido repitiendo a lo largo de todos los bloques de modelos analizados, siempre era un modelo *ad hoc* el que mejor comportamiento demostraba, o había uno o varios modelos *ad hoc* en ese grupo de error bajo. Bajo esta perspectiva se explica que las pautas en la construcción de los modelos de exposición más recientes, pasen por recurrir a una fórmula *ad hoc* para la estimación de la cobertura y buscar una distribución de probabilidad que replique de la manera más adecuada posible la distribución de contactos. En esta línea están los modelos que se analizan en este trabajo de Leckenby y Ju (1989): Hofmans Poisson Binomial, Leckenby y Kishi (1981): Kwerel geométrico y Hofmans geométrico y Leckenby y Boyd (1984a): Hofmans beta binomial. También bajo esta filosofía de construcción, se inscribe el modelo que, como corolario a este trabajo, propondremos al final del capítulo: el modelo Morgensztern beta binomial.
- b. Romper la hipótesis de homogeneidad de los individuos, no supone mejoras sustanciales en la estimación de la cobertura. Si bien es cierto que dos de los modelos que conforman el grupo de error medio⁵, Metheringham y Dirichlet multinomial consideran la heterogeneidad de las poblaciones, no es menos cierto que los

⁵ El grupo de error bajo está formado, como se ha visto, exclusivamente por un modelo *ad hoc* y, por ello, no sujeto a hipótesis alguna. Esto hace que todos los comentarios se refieran a comparar las hipótesis en las que se basan los modelos de los grupos de error medio y alto.

modelos de Markov Bernouilli y el Kwerel geométrico, también en este grupo, mantienen la hipótesis de homogeneidad.

- c. Levantar la hipótesis de homogeneidad de los soportes parece contraproducente en términos de precisión en las estimaciones. Aunque el modelo Dirichlet multinomial considera la heterogeneidad de los mismos y está en el grupo de error medio, el resto de modelos de este grupo mantienen el supuesto de homogeneidad. Además, puede comprobarse que todos los modelos del grupo de error alto, también consideran que los soportes son heterogéneos.
- d. Hay dos hipótesis que, sin embargo, sí que demuestran ser demasiado restrictivas y que, de mantenerse, provocan que las estimaciones de los modelos sean bastante pobres. Nos referimos a suponer el carácter aleatorio de los fenómenos de duplicación y acumulación de las audiencias. Aunque no es condición suficiente considerar la no aleatoriedad de los mismos para que el modelo funcione adecuadamente (como lo demuestra el mal comportamiento de los modelos Hofmans beta binomial y Hofmans geométrico), sí que parece ser una condición necesaria: todos los modelos del grupo de error medio la consideran: Markov Bernouilli, Metheringham, Dirichlet multinomial y Kwerel geométrico.
- e. Finalmente el único modelo que considera el carácter no estacionario respecto a las inserciones en los soportes, el de Markov Bernouilli, es el que mejores estimaciones de la cobertura ofrece de entre los modelos estocásticos.

A lo largo de este capítulo, se ha planteado en cada bloque de modelos la cuestión de si el comportamiento de los mismos, es distinto según el tipo de soporte incluido en el plan de medios: diarios, revistas semanales y mensuales, respondiendo

así a la cuestión a investigar CI_{2B} . El origen de esta pregunta está, como se indicó, en el hecho de que la mayoría de los modelos fueron diseñados inicialmente para el medio revistas y, siempre, evaluados en base a planes formados únicamente por soportes de este medio. En epígrafes anteriores ya se ha demostrado que los resultados, o bien no son diferentes entre diarios y revistas o que, cuando lo son, también lo son entre revistas semanales y mensuales, lo que nos ha hecho concluir hasta el momento que los modelos, y la clasificación que derivamos, también son válidos para planes formados por el medio diarios.

La misma conclusión se obtiene en la estimación de la cobertura de los modelos de audiencia neta acumulada y, quizás, de una manera incluso más evidente. Como se observa en el cuadro 5.41, en una primera aproximación, el conjunto de los modelos evaluados comete un error menor, precisamente cuando el plan de medios está formado por diarios, lo que parece confirmar la mencionada adecuación. Si se trata de comprobar si esta diferencia aparente es estadísticamente significativa, el análisis de varianza realizado y resumido en el cuadro 5.42, confirma este hecho, constatando la prueba *post hoc* del cuadro 5.43 que estas diferencias son significativas entre todos los tipos de planes. Luego, al menos en la estimación de la cobertura, podemos concluir que la clasificación que se ha obtenido es también válida para el medio diarios y, además, que las estimaciones serán, probablemente, incluso más fiables cuando se considere planes formados por soportes de este medio.

Cuadro 5.41. Medias de lnAER y lnRER según el tipo de soporte.

Soporte	lnAER		lnRER	
	Media	D. típica	Media	D. típica
Periódicos	-3,7288	1,2474	-2,2998	1,2811
R. Semanales	-2,8546	0,9627	-1,5441	0,9604
R. Mensuales	-3,0500	0,8831	-0,8786	0,8635

Fuente: elaboración propia

Cuadro 5.42. Análisis de varianza para las definiciones de error tomando como factor el tipo de soporte

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
lnAER	Entre tipo soporte	2	60,62	58,65*
	Error	987	1,03	
lnRER	Entre tipo soporte	2	147,05	142,05*
	Error	987	1,04	

* Diferencias significativas para $p \leq 0.01$ *Fuente: elaboración propia***Cuadro 5.43. Matriz de diferencias pareadas de lnAER y lnRER. Test de Tukey.**

Soporte	lnAER		lnRER	
	Mensuales	Semanales	Mensuales	Semanales
Semanales	-1,9551*		0,6655**	
Periódicos	0,6787**	0,8742**	1,4212**	0,7557**

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: elaboración propia*

5.4.2 Error en la estimación de la distribución de contactos

Manteniendo el esquema seguido hasta el momento en el desarrollo de este capítulo, debe determinarse en primer lugar qué modelos tienen un comportamiento tal que, en su calidad de *outliers*, dificulten el análisis del resto de modelos. Nos encontramos, llegados a este punto, con la primera diferencia importante que surge al analizar las estimaciones de la distribución de contactos de los modelos de audiencia neta acumulada: esta identificación no es tan obvia como en casos anteriores. Analizando el cuadro 5.44 y más claramente su traducción gráfica en la figura 5.16, no se observa un grupo de modelos cerca del origen, y por ello con un buen comportamiento general, y otros alejados del mismo, los *outliers*, como ocurría en las figuras 5.4 y 5.10. La causa es que, como se analizará con profundidad posteriormente, las dos medidas de error en la estimación de la cobertura no ofrecen

resultados simétricos, lo que provoca que modelos que funcionan bien atendiendo a una de ellas, funcionen peor cuando se adopta el otro criterio.

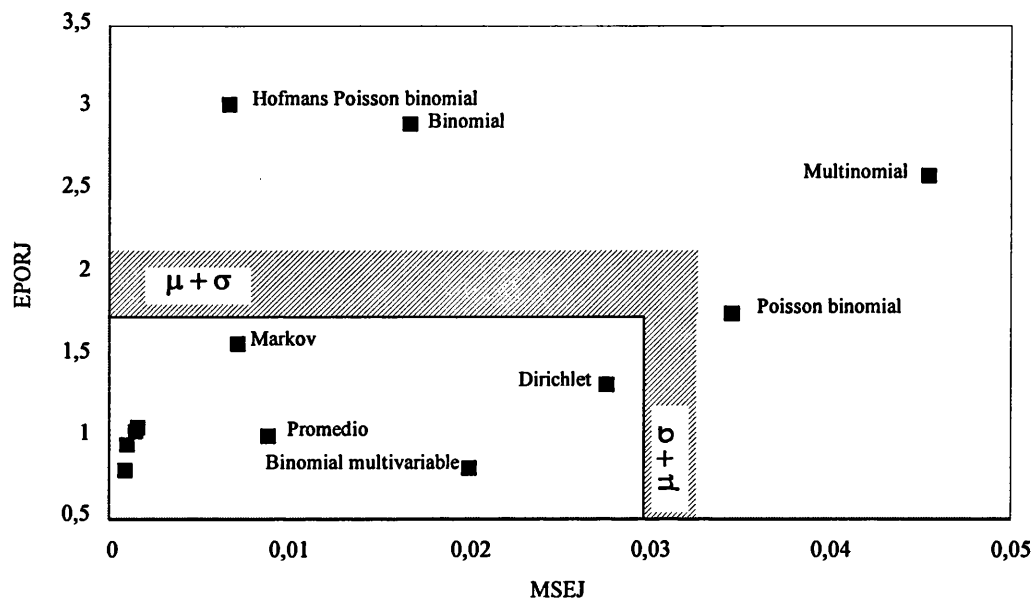
Este hecho tendrá consecuencias en el proceso de jerarquización, pero atendiendo a la identificación de los *outliers*, nos obliga a adoptar un criterio objetivo de selección cuando, con anterioridad, podíamos recurrir sin cometer prácticamente imprecisión alguna, a seleccionarlos gráficamente. Se ha decidido, en consecuencia, considerar como *outliers*, a aquellos modelos que comenten un error que supera a la media en más de la desviación típica. Como se observa en la figura 5.16, esto nos lleva a clasificar como tales a los modelos multinomial, binomial, Hofmans Poisson binomial y Poisson binomial.

De nuevo, por primera vez en todo nuestro análisis, nos encontramos ante el hecho de que, modelos que no eran *outliers* en la estimación de la cobertura (Hofmans Poisson binomial), lo son en la estimación de la distribución de contactos, dándose también el caso contrario: modelos que tenían un comportamiento paupérrimo al predecir la cobertura (binomial multivariable) no pueden considerarse *outliers* cuando estiman la distribución de contactos.

Cuadro 5.44. Errores promedio en la estimación de la distribución de contactos. Identificación de *outliers*. N = 99.

Modelo	EPORJ		MSEJ	
	Media	D. Típica	Media	D. Típica
Binomial	2,9124	1,0702	0,0166	0,0057
Binomial multivariable	0,8239	0,0813	0,0199	0,0094
Multinomial	2,5975	0,8267	0,0454	0,0195
Markov-Bernouilli	1,5805	0,5134	0,0072	0,0074
Metheringham	0,8078	0,3795	0,0008	0,0014
Dirichlet multinomial	1,3312	1,4641	0,0277	0,0679
Poisson Binomial	1,7581	0,6449	0,0346	0,0127
Kwerel geométrica	0,9666	0,3170	0,0010	0,0009
Hofmans geométrica	1,0745	0,3454	0,0016	0,0010
Hofmans beta binomial	1,0374	0,4103	0,0014	0,0009
Hofmans Poisson binomial	3,0409	1,8049	0,0067	0,0033
Promedio Metheringham Dirichlet	1,0166	0,7228	0,0089	0,0183

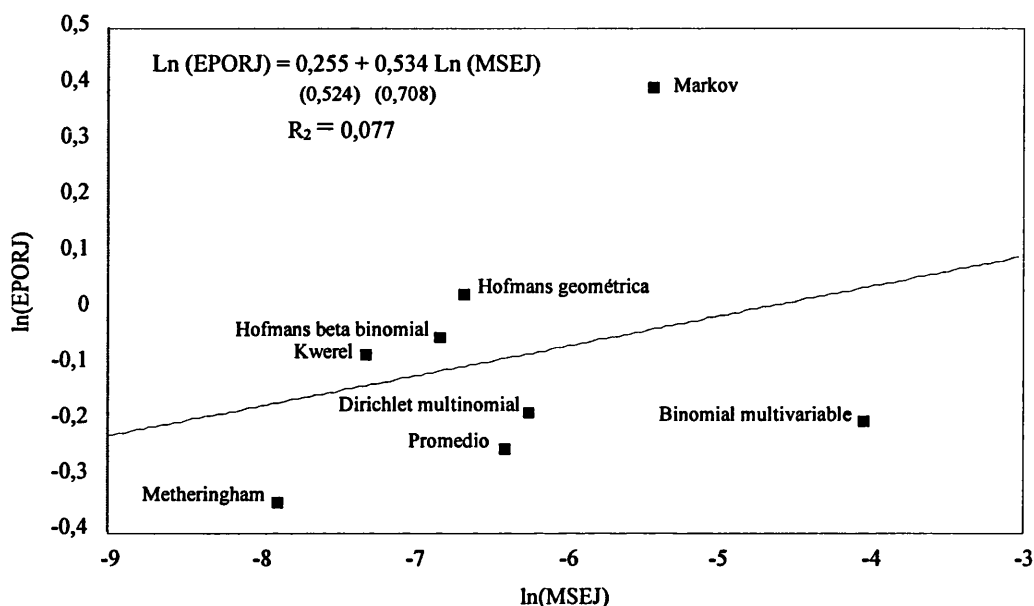
Fuente: Elaboración propia.

Figura 5.16. Identificación de *outliers*.

Fuente: elaboración propia

Ya se ha adelantado que una de las principales diferencias en el comportamiento de los modelos de audiencia neta acumulada cuando estiman la distribución de contactos, es que las dos medidas de error empleadas, MSEJ y EPORJ, no conducen a los mismos resultados, como se aprecia en la figura 5.17. Se impone, por ello, analizar las causas de este hecho y, sobre todo, derivar sus consecuencias.

Figura 5.17. Coherencia de los errores tipo MSEJ y EPORJ



Fuente: Elaboración propia

En el capítulo cuarto, concretamente en el epígrafe 4.2.3.2 se definieron las medidas de error empleadas para valorar la precisión de las estimaciones de la distribución de contactos. Si atendemos a las expresiones [4.12] y [4.13], correspondientes respectivamente a MSEJ y EPORJ, se observa que la principal diferencia entre ellas radica en el procedimiento que emplean para relativizar la suma de los errores cometidos en cada nivel de exposición de la distribución de contactos. Así, mientras MSEJ relativiza dividiendo por el número de niveles de exposición (en un plan de medios formado por n inserciones en m soportes estos niveles serían $n \cdot m + 1$), EPORJ relativiza por uno menos el porcentaje observado de la población objetivo que no ha resultado expuesto ninguna vez al plan de medios. Esa parte de la distribución de contactos es, precisamente, el nexo entre ella y la cobertura del plan de medios: la cobertura se define, en términos porcentuales, como uno menos la población no expuesta. Pues bien, de ahí surgen las discrepancias entre ambas medidas de error. Los perfiles de las distribuciones de contactos observados y estimados pueden ser paralelos, pero aquellos modelos que cometieron un fuerte

error sobrestimando el valor de la cobertura, están subestimando necesariamente el nivel cero de la distribución de contactos, dado que este forma parte del denominador de la expresión [4.13], y EPORJ es anómalamente reducido. Como, por otra parte, MSEJ, sí que incorpora en el cálculo del error la diferencia entre el nivel cero observado y estimado de la distribución de contactos y, además, considerando el cuadrado de la diferencia, ambos tipos de error conducen a resultados contradictorios en estos modelos particulares.

Este hecho se produce fundamentalmente en la estimación realizada por el modelo binomial multivariable (era *outlier* en la estimación de la cobertura) y en el Dirichlet multinomial (y consecuentemente en el modelo que promedia a éste con el de Metheringham). De no ser por ellos, la congruencia entre las dos medidas de error, sería de nuevo elevada pero, dado que existen, atenderemos principalmente al error tipo MSEJ, no sujeto al problema señalado, para la jerarquización de todos ellos.

Como se demuestra en el cuadro 5.45, el error cometido por los distintos modelos, es significativamente distinto. Como venimos realizando hasta el momento, debemos establecer qué modelos causan esas diferencias para, con ello, proceder a su jerarquización, de acuerdo con la cuestión a investigar CI_{1B} . Los cuadros 5.46 y 5.47, sintetizan los resultados de la prueba *post hoc* de Tukey en la que fundamentaremos nuestros comentarios.

5.45. Análisis de varianza para las definiciones de error

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
ln MSEJ	Entre modelos	7	138,40	78,86*
	Error	784	1,75	
ln EPORJ	Entre modelos	7	5,13	16,91*
	Error	784	0,30	

* Significativo para $p \leq 0,01$

Fuente: elaboración propia

Como se ha señalado, las clasificaciones a la que conducen las dos variables de error utilizadas, son diferentes. Así, MSEJ identifica tres grupos, el primero de ellos, que venimos llamando como grupo de *error bajo*, estaría formado por los modelos de Metheringham y Kwerel (las diferencias entre ellos son significativas para $p \leq 0.05$ pero no para $p \leq 0.01$). El segundo grupo, el de *error medio*, quedaría formado por los modelos de Hofmans beta binomial, Hofmans geométrica, Dirichlet multinomial y el promedio entre este y el modelo de Metheringham. Finalmente, el grupo de *error alto*, quedaría integrado por los modelos de Markov y binomial multivariable dado que, aunque las diferencias del error cometido en la estimación de la distribución de contactos, son significativamente distintas entre ellos, carece de sentido formar dos grupos con un solo modelo cuando, además, son los de peor comportamiento.

Si recurrimos a EPORJ como medida del error, el grupo de *error bajo* estaría formado por los mismos que incluía MSEJ, esto es, Metheringham y Kwerel, a los que se unirían el binomial multivariable, Dirichlet y el promedio entre éste y el modelo de Metheringham. El grupo de *error medio*, quedaría reducido a los modelos Hofmans beta binomial y Hofmans geométrica, mientras que el de Markov formaría el grupo de *error alto*.

Cuadro 5.46. Matriz de diferencias pareadas de lnMSEJ. Test de Tukey.

	MET	KWE	HBBD	HGD	PRO	DIR	MKV
Kwerel geométrica	-0,5801*						
Hofmans beta binomial	-1,0670**	-0,4870					
Hofmans geométrica	-1,2175**	-0,6375*	-0,1505				
Promedio Meth- Dirichlet	-1,4807**	-0,9006**	-0,4136	-0,2631			
Dirichlet multinomial	-1,6355**	-1,0554**	-0,5684	-0,4179	-0,1548		
Markov Bernouilli	-2,4609**	-1,8808**	-1,3938**	-1,2433**	-0,9802**	-0,8254**	
Binomial multivariable	-3,8432**	-3,2631**	-2,7762**	-2,6257**	-2,3625**	-2,2077**	-1,3824**

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: Elaboración propia*

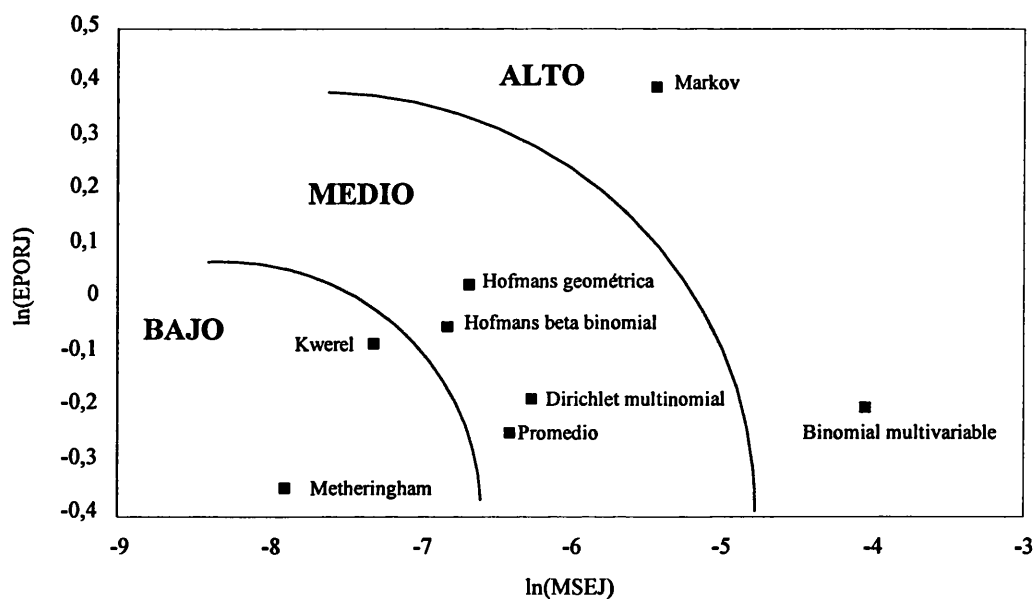
Cuadro 5.47. Matriz de diferencias pareadas de lnEPORJ. Test de Tukey.

	MET	PRO	BMUL	DIR	KWE	HBBD	HGD
Promedio Meth- Dirichlet	-0,0992						
Binomial multivariable	-0,1468	-0,0476					
Dirichlet multinomial	-0,1623	-0,0631	-0,0155				
Kwerel geométrica	-0,2665*	-0,1673	-0,1197	-0,1043			
Hofmans beta binomial	-0,2960**	-0,1968	-0,1491	-1,3367	-0,0294		
Hofmans geométrica	-0,3755**	-0,2763*	-0,2287	-0,2132	-0,1090	-0,0796	
Markov Bernouilli	-0,7417**	-0,6425**	-0,5949**	-0,5794**	-0,4752**	-0,4458**	-0,3662**

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: Elaboración propia*

Se hace necesario, por mantener la simetría con los epígrafes anteriores, ofrecer una clasificación única. Dado que, como se ha justificado, el error EPORJ tiene serias limitaciones cuando las estimaciones de la cobertura que ofrecen los modelos son poco acertadas, y este hecho se demuestra por el bajo error cometido por un modelo que era *outlier* en el cálculo de esta variable, el modelo binomial multivariable, mantendremos los resultados de MSEJ, que vienen ilustrados en la figura 5.18.

Figura 5.18. Jerarquización de los modelos de audiencia neta acumulada. Estimación de la distribución de contactos.



Fuente: Elaboración propia

Las implicaciones para la introducción de hipótesis simplificadoras que se derivan de esta clasificación son muy parecidas a las que se expusieron al valorar los resultados de la estimación de la cobertura. Así, podemos resumirlas en las siguientes:

- a. Poca relevancia de considerar el carácter homogéneo o heterogéneo de la población, dado que, entre los modelos de mejor funcionamiento, tenemos un que asume la homogeneidad (Kwerel geométrico) y otro que no lo hace (Metheringham).
- b. Respecto a la introducción de la heterogeneidad de los soportes, esta complejización de los modelos parece innecesaria, por cuanto los dos modelos señalados, no asumiéndola, demuestran funcionar mejor que el resto.
- c. Al igual que ocurría con la cobertura se demuestra como importante no considerar que los fenómenos de acumulación y duplicación son aleatorios, en cuanto que perjudica a la capacidad de estimación de los modelos.
- d. Finalmente, de manera contraria a lo que ocurría con la cobertura, parece innecesario añadir complejidad mediante la asunción del carácter no estacionario de los soportes, dado el mal comportamiento del modelo de Markov.

Desde nuestro punto de vista, más que atender a las hipótesis en que se basan los modelos y que explicarían el mejor o peor funcionamiento de los mismos, los resultados obtenidos parecen demostrar que, lo realmente determinante, es la elección de la función de probabilidad en que el modelo va a basarse. En los tres bloques de modelos analizados: acumulación de audiencias, duplicación de audiencias y modelos de audiencia neta acumulada, hemos encontrado resultados recurrentes: la función binomial, aunque facilita por la sencillez la construcción de los modelos, da unos pobre resultados. Esta conclusión no varía cuando la distribución binomial se ha compuesto en trabajos recientes con una Poisson para mejorar su capacidad de estimación. Por contra, aquellos modelos que se han construido bajo la distribución beta binomial, principalmente el modelo de

Metheringham, han demostrado una buena capacidad predictiva, sobre todo en el cálculo de la distribución de contactos, aunque algunos modelos *ad hoc*, la superaban en el cálculo de la cobertura. Cuando trabajos recientes han intentado introducir otras distribuciones, como la geométrica, los resultados obtenidos por nosotros, demuestran que han mejorado a la distribución binomial, pero nunca al mejor modelo basado en la beta binomial.

Retomando la línea argumental mantenida a lo largo de este capítulo, llega el momento de comprobar si, también para la estimación de la distribución de contactos, los modelos evaluados son susceptibles de ser utilizados en planes de medios formados por diarios, aunque originalmente no fueran diseñados para ello. El cuadro 5.48, como primera aproximación, confirma los resultados que ya se habían obtenido con anterioridad, dado que, en conjunto, el error cometido cuando el plan de medios está formado por periódicos es inferior al resto. Como se observa en el cuadro 5.49, estas diferencias aparentes, son efectivamente significativas, estadísticamente hablando.

Cuadro 5.48. Medias de lnMSEJ y lnEPORJ según el tipo de soporte.

Soporte	lnMSEJ		lnEPORJ	
	Media	D. típica	Media	D. típica
Periódicos	-7,3458	1,7247	-0,6949	0,6380
R. Semanales	-5,4377	1,5853	0,0982	0,4613
R. Mensuales	-6,6635	1,3193	0,1479	0,3073

Fuente: elaboración propia

Cuadro 5.49. Análisis de varianza para las definiciones de error tomando como factor el tipo de soporte

Tipo de Error	Fuente de Variación	Grados de Libertad	Media Cuadrática	F
lnMSEJ	Entre tipo soporte	2	242,76	103,02*
	Error	789	2,36	
lnEPORJ	Entre tipo soporte	2	50,04	227,51*
	Error	789	0,22	

* Diferencias significativas para $p \leq 0.01$ *Fuente: elaboración propia***Cuadro 5.50. Matriz de diferencias pareadas de lnMSEJ y lnEPORJ. Test de Tukey.**

Soporte	lnMSEJ		lnEPORJ	
	Mensuales	Semanales	Mensuales	Semanales
Semanales	-1,2258**		-0,0495	
Periódicos	0,6823**	1,9081**	0,8428**	0,7932**

** diferencias significativas para $p \leq 0,01$ * diferencias significativas para $p \leq 0,05$ *Fuente: elaboración propia*

5.4.3 Factores de error en los modelos de audiencia neta acumulada

Como se ha venido realizando con los bloques de modelos anteriormente analizados para dar respuesta a la cuestión a investigar CI_{2A} , en este epígrafe se tratará de establecer qué características internas de los planes explican el mejor o peor funcionamiento de los distintos modelos de audiencia neta acumulada. Para este fin, se han regresado las distintas medidas de error que se han venido utilizando sobre los factores de error que se definieron oportunamente el en capítulo cuarto: audiencia bruta del plan de medios (AB), suma de las audiencias acumuladas (SAC) y duplicadas (SAD), número de soportes que integran el plan de medios (SOP), tamaño relativo de los soportes que se incluyen en el plan (BMA, AM, AB, MB) y, finalmente, la *ratio* entre la suma de las audiencias duplicadas y la audiencia bruta

(RADAB) y entre la suma de audiencias acumuladas y esa misma variable (RACAB). El objetivo de este proceder es, como ya se indicó, doble. Por un lado se pretende identificar las variables que explican que unos modelos funcionen mejor que otros, es decir, que sean *outliers* o pertenezcan a los grupos de error alto, medio y bajo. Por otro lado, se busca dar un criterio adicional para, si llegara el caso, elegir entre aquellos modelos que por pertenecer al mismo grupo no cometen un error significativamente distinto. Si podemos identificar ante qué factores internos del plan de medios, es más sensible este error, analizadas las características del plan que se pretende evaluar, podrá predecirse qué modelo proporcionará mejores estimaciones de la cobertura y la distribución de contactos.

Los cuadros 5.51 y 5.52, ofrecen los coeficientes estandarizados de las variables independientes (factores de error) que han actuado como regresores de, respectivamente, el error promedio (lnAER) y el error relativo (lnRER). De su análisis se desprenden las siguientes conclusiones.

La audiencia bruta se muestra como el factor determinante del error que cometen los modelos. Cuanto mayor es ésta, mayor es el error promedio en que se incurre. Así, los *outliers*, tienen coeficientes estandarizados superiores a dos (2,251 el modelo binomial, 2,362 el modelo binomial multivariable y 2,022 el modelo multinomial), y superiores a 1,6 los modelos pertenecientes al grupo de error alto. Por su parte, los modelos de mejor comportamiento, Markov y Morgensztern, lo son, precisamente, por no ser sensibles a esta variable, como demuestra el hecho de que sus coeficientes estandarizados no son significativos. Las *ratios* RACAB y RADAB, también contribuyen de una manera notable en el comportamiento de los modelos, como se desprende de que el error promedio aumenta en los *ouliers* cuando esta *ratio* crece, mientras que disminuye en el resto.

Cuadro 5.51. Factores de error en la estimación de la cobertura: ln AER.
 Coeficientes estandarizados.

	AB	SAC	SAD	SOP	BMA	AM	AB	MB	RADAB	RACAB	R ²	F
Binomial	2,251**	-0,621**	-1,711**	0,217**					0,657**	0,416**	0,888	69,89**
Binomial multivariable	2,362**	-0,798**	-1,743**	0,248**					0,694**	0,527**	0,882	65,84**
Multinomial	2,022**		-1,302**							0,517*	0,594	12,90**
Markov Bernouilli											0,179	1,92
Metheringham	1,501**	-0,965**	-0,703**	0,313**							0,678	18,52**
Dirichlet multinomial	1,592**	-1,579**									0,724	23,06**
Poisson binomial	-1,666**		0,906**	-0,40*							0,992	1.058,*
Kwerel geométrica			-0,631*						-0,529**	-0,783**	0,551	10,78**
Hofmans geométrica	1,695**		-1,128**	0,303**					-0,662**	-0,554*	0,602	13,32**
Hofmans beta binomial	1,695**		-1,128**	0,303**					-0,662**	-0,554*	0,602	13,31**
Hofmans Poisson binomial	1,695**		-1,128**	0,303**					-0,662**	-0,554*	0,602	13,31**
Promedio Meth-Dirichlet	1505**	-1,120**	-0,665**	0,173*							0,793	33,66**
Hofmans	1,695**		-1,128**	0,303**					-0,662**	-0,554*	0,602	13,31**
Morgensztern				0,299*	-0,364*						0,450	7,19**

** Significativo para $p \leq 0,01$

* Significativos para $p \leq 0,05$

Fuente: Elaboración propia

Cuadro 5.52. Factores de error en la estimación de la cobertura: lnRER.
 Coeficientes estandarizados.

	AB	SAC	SAD	SOP	BMA	AM	AB	MB	RADAB	RACAB	R ²	F
Binomial	-1,499**			-0,086*					0,611**	-0,299**	0,943	145,81**
Binomial multivariable	-1,406**								0,642**	-0,223*	0,945	152,08**
Multinomial	-1,451**								0,386**		0,815	38,86**
Markov Bernouilli	-1,273*	1,193*								-0,826*	0,223	2,53*
Metheringham		-0,661*		0,205*							0,694	19,96**
Dirichlet multinomial	0,685*	-1,269**							0,285*		0,767	28,94**
Poisson binomial	-1,904**	0,280**	0,880**	-0,121**				0,099**	0,094*	-0,281**	0,984	549,61**
Kwerel geométrica									-0,343*	-0,889**	0,693	19,91**
Hofmans geométrica	-0,806**								-0,328*	-0,761**	0,809	37,22**
Hofmans beta binomial	-0,806**								-0,328*	-0,761**	0,809	37,17**
Hofmans Poisson binomial	-0,806**								-0,328*	-0,761**	0,809	37,17**
Promedio Meth-Dirichlet		-0,775**							0,297*	-0,328*	0,821	40,43**
Hofmans	-0,806**								-0,328*	-0,761**	0,809	37,22**
Morgensztern	-1,855**	1,124*	0,780*								0,364	5,04**

** Significativo para $p \leq 0,01$

* Significativos para $p \leq 0,05$

Fuente: *Elaboración propia*

La conclusión es distinta cuando se analiza el error relativo medido mediante $\ln RER$. Dado que en este caso la audiencia bruta actúa como reductor del error, las diferencias que determinan la clasificación surgen del hecho de que en los *outliers* las *ratios* RADAB y RACAB tienden a que éste aumente, mientras que en el resto de modelos fuerzan su disminución. Este hecho explica la distinción entre *outliers* y demás modelos, pero cabe preguntarse qué provoca la jerarquización obtenida entre estos últimos. Analizando el cuadro 5.52, se observa que la influencia reductora del error de la audiencia bruta, es inferior entre los modelos del grupo de error alto (coeficientes estandarizados entorno a -0,8) que entre los modelos de mejor funcionamiento (Morgensztern: -1,855; Markov: -1,273).

Merece destacarse por su carácter particular, los resultados del modelo Poisson binomial. Este modelo es el único cuyo error tipo $\ln AER$ se reduce al aumentar la audiencia bruta y, también, cae el $\ln RER$ en mayor proporción que en el resto de modelos cuando ésta aumenta. Sin embargo, su comportamiento global es muy pobre, tanto que, recordemos, oportunamente lo clasificamos como un *outlier*. Este hecho provoca dos consideraciones. Por un lado, que el potencial de este modelo para evaluar planes de medios que deriven en elevadas coberturas, es alto y que su clasificación debe haber venido condicionada porque los planes utilizados en su evaluación eran poco adecuados a sus características⁶. Por otro lado, es evidente que nos encontramos ante un modelo de estructura distinta a la del resto dado que, también, es el único en el que la suma de las audiencias duplicadas ejerce una influencia negativa sobre el error relativo y de los pocos en los que provoca el aumento del error promedio. De nuevo, si los planes que aleatoriamente se han construido para su evaluación, tienen elevadas duplicaciones, puede explicarse el mal comportamiento del único *outlier* no basado en la distribución binomial y del que cabría esperar mejores resultados al ser un modelo de reciente desarrollo.

⁶ Recordemos que, como se indicó en el cuadro 4.12 el número medio de GRPs de los planes construidos es de 101, cifra superada en la práctica totalidad de los planes de medios reales.

De los cuadros anteriores, también se deriva que, por lo general, el error en la estimación de la cobertura no es más sensible a que el plan de medios esté formado por una combinación determinada de tamaño de soportes que otra, como demuestra el hecho de que, la mayoría de coeficientes de estas variable no sean significativos, ni cuando la variable dependiente es el error promedio, como cuando lo es el relativo.

Si analizamos la influencia de los distintos factores de error en la estimación de la distribución de contactos (cuadros 5.53 y 5.54), encontramos resultados muy similares a los ya reseñados. De nuevo la audiencia bruta es el elemento que mayor influencia ejerce sobre el error que cometen los modelos. Así en el error tipo $\ln\text{MSEJ}$, se comprueba que los *outliers* y modelos de peor funcionamiento, tienen los coeficientes estandarizados asociados a esta variable muy superiores al resto (Binomial: 2,556; Binomial multivariable: 2,352 vs. Metheringham: 1,960 y Kwerel geométrica 0,843).

También se confirma que la suma de las audiencias duplicadas y acumuladas moderan la influencia de la audiencia bruta, al disminuir el error tipo $\ln\text{AER}$ cuanto más elevados son estos fenómenos en el plan de medios, mientras que otro elemento determinante en la jerarquización de los modelos en base al tipo de error señalado, aunque con una influencia menor a las anteriores, proviene de las *ratios* RADAB y RACAB. Se observa que mientras que en los modelos de peor comportamiento, el error crece al crecer estas variables (Binomial: RADAB=0,681 RACAB=0,487; Binomial multivariable RADAB=0,621 RACAB=0,495) los modelos que mejor funcionan, ven como el $\ln\text{AER}$ disminuye al aumentar las *ratios* señaladas (v.g. Kwerel geométrica: RADAB= -1,198 RACAB=-0,461).

Cuadro 5.53 Factores de error en la estimación de la distribución de contactos: lnMSEJ.
Coeeficientes estandarizados.

	AB	SAC	SAD	SOP	BMA	AM	AB	MB	RADAB	RACAB	R ²	F
Binomial	2,556**	-0,945**	-1,953**	0,281**					0,681**	0,487**	0,838	45,42**
Binomial multivariable	2,352**	-0,732**	-1,714**	0,244**					0,621**	0,495**	0,898	77,76**
Multinomial	2,558**	-0,982**	-1,435**	0,295**					0,220*	0,563**	0,858	52,56**
Markov Bernoulli	1,827**	-0,464**	-0,855**	-0,119**		0,088*	0,126**	-0,169**		-0,230*	0,943	146,32**
Metheringham	1,960**	-1,067**	-0,888**	0,281**							0,748	26,13**
Dirichlet multinomial	2,141**	-0,832**	-1,372**	-0,119*					0,447**		0,889	70,78**
Poisson binomial	-1,497**	-0,221**	0,834**	-0,060**							0,988	702,06**
Kwerel geométrica	0,843**								-1,198**	-0,461**	0,846	48,36**
Hofmans geométrica	1,326**			0,252**					-0,760**		0,719	22,52**
Hofmans beta binomial	1,903**		-0,864**	0,285**					-0,714**	-0,473*	0,755	27,09**
Hofmans Poisson binomial		0,685*		-0,341**	0,473**			0,657**	0,364*	-1,105**	0,737	24,69**
Promedio Meth-Dirichlet	2,221**	-0,840**	-1,443**						0,403**		0,839	46,02**

** Significativo para $p \leq 0,01$

* Significativos para $p \leq 0,05$

Fuente: *Elaboración propia*

Cuadro 5.54 Factores de error en la estimación de la distribución de contactos: InEPORJ.
 Coeficientes estandarizados.

	AB	SAC	SAD	SOP	BMA	AM	AB	MB	RADAB	RACAB	R ²	F
Binomial	-1,578**		0,323**						0,529**	-0,438**	0,950	166,49**
Binomial multivariable									-0,581**	-0,586**	0,796	34,25**
Multinomial	-1,121**		0,352**						0,274**	-0,420**	0,954	181,06**
Markov Bernouilli	-0,653**		0,256*	-0,285**					0,228**	-0,743**	0,927	111,16**
Metheringham		-0,659*								-0,425*	0,772	29,78**
Dirichlet multinomial	1,382**	-0,610**	-1,148**	-0,252**					0,617**	-0,433	0,849	49,41**
Poisson binomial	-1,350**								0,692**	-0,471**	0,909	88,43**
Kwerel geométrica	-2,120**	1,255**	2,345**						-1,418**	-0,900**	0,686	19,19**
Hofmans geométrica	-2,061**	0,860*	1,694**		-0,343*	-0,237*			-0,697**	-0,954**	0,617	14,15**
Hofmans beta binomial	-0,625*		0,367*						-0,283**	-0,833**	0,874	60,98**
Hofmans Poisson binomial	-1,700**	0,897**	0,554**	-0,360**	0,262**			0,482**	0,441**	-1,007**	0,850	49,68**
Promedio Meth-Dirichlet	1,243**	-0,659*	-1,137**	-0,179*					0,591**	-0,429*	0,787	32,49**

** Significativo para $p \leq 0,01$ * Significativos para $p \leq 0,05$ *Fuente: Elaboración propia*

Cuando se toma como referencia el error tipo lnEPORJ, las conclusiones son muy similares a las obtenidas respecto a lnMSEJ. Así, se observa que el buen funcionamiento del modelo de Kwerel deriva tanto del efecto negativo sobre el error de la audiencia bruta, como de la influencia en este mismo sentido de RADAB. Por contra en los modelos de peor funcionamiento, aunque la audiencia bruta tiende a disminuir el error, la *ratio* entre la audiencia duplicada y la audiencia bruta lo incrementa.

Si fuera necesario elegir entre los dos modelos de mejor comportamiento, Kwerel geométrica y Metheringham para evaluar determinado plan de medios, a la vista de los resultados anteriores, debe tenerse en cuenta que el primero cometerá menor error que el segundo en la medida en que la audiencia bruta de los planes sea elevada, sin embargo, es mucho más sensible que el modelo de Metheringham a los fenómenos de acumulación y duplicación de las audiencias.

5.4.4 Corolario: propuesta de un nuevo modelo de audiencia neta acumulada

A lo largo de la revisión de los modelos de audiencia neta acumulada, se ha puesto de manifiesto, y demostrado a la vez en el proceso evaluador, que la creación de nuevos modelos debe intentar aprovechar el buen funcionamiento de los modelos *ad hoc* en la estimación de la cobertura y, paralelamente, recurrir a la función de probabilidad que mejor replique el perfil de las distribuciones de contactos. En esta línea están modelos como los Hofmans geométrica y Kwerel geométrica que utilizan respectivamente las fórmulas *ad hoc* de Hofmans y Kwerel para estimar la cobertura y la distribución geométrica para la distribución de contactos o el Hofmans beta binomial que cambia esta función por la beta binomial.

Este procedimiento de construcción de modelos tiene un problema fundamental. Si se calcula separadamente la cobertura de la distribución de contactos, se puede incurrir en la incongruencia de que la estimación de la primera difiera del valor que adoptaría el nivel cero de la segunda⁷. Como se vio en el capítulo tercero, este inconveniente se salva ajustando iterativamente los parámetros de la distribución de contactos, hasta que el porcentaje de la población que no ha resultado expuesto a ninguna de las inserciones sea coherente con el valor de cobertura estimado.

Todos los modelos que se han venido creando bajo esta perspectiva han utilizado como estimador de la cobertura, ora la fórmula de Hofmans, ora la de Kwerel. Sin embargo, se ha demostrado en esta tesis doctoral, que para el caso español, el que mejores estimaciones aporta es el modelo de Morgensztern. Este resultado nos ha llevado a proponer como corolario a nuestro trabajo el modelo que a continuación describimos y evaluamos: el Morgensztern beta binomial (MBBD).

Este modelo se basa en el procedimiento seguido por Leckenby y Boyd (1984a) en el desarrollo del modelo Hofmans beta binomial, con la salvedad ya señalada de que la cobertura se estimaría mediante la fórmula propuesta por Morgensztern (1970). Para este autor, la cobertura de un plan de medios formado por n inserciones en m soportes distintos, vendría dada por:

$$R_n = \frac{\left(\sum_{i=1}^m R_{ni} \right)^2}{\sum_{i=1}^m R_{ni} + \sum_{i=1}^m \sum_{j=i+1}^{m-1} \frac{k_{ij} A_{ij} R_{ni} R_{nj}}{A_i A_j}} \quad [5.1]$$

donde k_{ij} se calcula como en la fórmula de Hofmans (1966) mediante la expresión:

⁷ Más concretamente uno menos ese nivel cero. Reiteradamente se ha venido manifestando que la cobertura puede calcularse como uno menos el porcentaje de la población objetivo que no ha resultado expuesto a ninguna de las N inserciones realizadas en el plan de medios.

$$k_{ij} = \frac{A_i + A_j}{A_i + A_j - A_{ij}}$$

siendo toda la notación conocida, salvo R_{ni} (R_{nj}) que es el alcance acumulado tras n inserciones en el soporte i (j). Evidentemente esta cobertura acumulada debe ser estimada de algún modo, puesto que habitualmente sólo se dispondrá de la audiencia del soporte. Como se demostró en este mismo capítulo en el epígrafe 5.2.1, el mejor modelo de acumulación de audiencias para ello es el de Hofmans (1966), remitiendo para su descripción al epígrafe 3.1.1.4 de esta misma tesis doctoral.

A partir de aquí, se recurrirá al procedimiento de estimación de ajuste de los parámetros de la distribución beta binomial denominado de medias y ceros que se describe seguidamente:

- a. **Paso 1.** Mediante la aplicación de la fórmula de Morgensztern descrita en el procedimiento anterior, se estima la cobertura del plan de medios: R_M .
- b. **Paso 2.** Arbitrariamente se da un valor A_0 al parámetro A de la distribución BBD.
- c. **Paso 3.** Mediante la siguiente expresión, se estima el parámetro B de la distribución BBD.

$$B_0 = \frac{A_0 \left(\sum_{i=1}^m n_i - \sum_{i=1}^m n_i A_i \right)}{\sum_{i=1}^m n_i A_i}$$

donde m es el número de soportes, n_i el número de inserciones realizadas en el soporte i cuya audiencia es A_i .

- d. **Paso 4.** Con los parámetros A_0 y B_0 así estimados, se calcula el número de individuos que no han resultado expuestos a ninguna de las N inserciones realizadas ($K = 0$), mediante la expresión deducida

en su momento para la BBD:

$$E_K^N = \binom{N}{K} \frac{\Gamma(A+B) \Gamma(A+K) \Gamma(N+B-K)}{\Gamma(A) \Gamma(B) \Gamma(A+B+N)} \quad [5.2]$$

- e. **Paso 5.** Dado que la cobertura es uno menos ese valor, se calcula así la cobertura que daría la aplicación directa de la BBD, y que llamaremos H_{BBD} .
- f. **Paso 6.** Se comparan las dos estimaciones de la cobertura obtenidas. Si la estimación mediante el método de Morgensztern es superior a la estimación mediante la BBD, se considera un nuevo valor del parámetro A (A_1) ligeramente inferior al anterior (A_0), procediendo a la inversa en caso contrario. Se repite el proceso hasta que la diferencia entre las dos estimaciones de la cobertura sea inferior al nivel de precisión decidido de antemano. Cuando esto sea así:
- g. **Paso 6.** Con los parámetros A_F y B_F finales, se calcula la distribución de contactos según [5.2].

La aplicación de este método conduce a los resultados que se sintetizan en el cuadro 5.55.

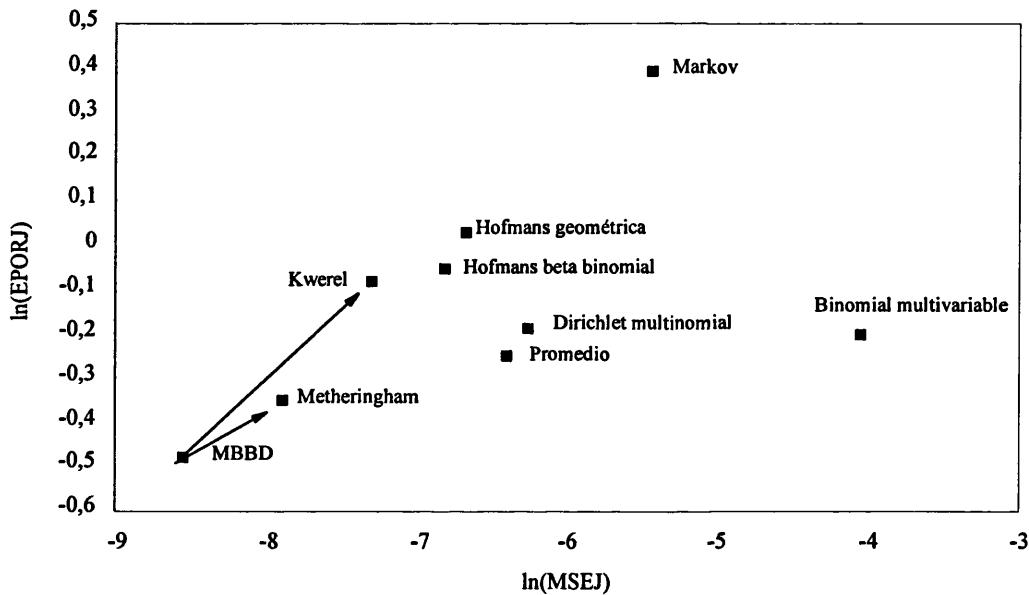
Cuadro 5.55. Funcionamiento del modelo MBBD.

Tipo de Error	Media	D. Típica
AER	0,0123	0,0083
RER	0,0629	0,0508
MSEJ	0,0003	0,0003
EPORJ	0,6788	0,2388
lnAER	-4,6693	0,8298
lnRER	-3,0245	0,7513
lnMSEJ	-8,5467	1,1006
lnEPORJ	-0,4768	0,4664

Fuente: elaboración propia

Aunque es evidente que por su construcción, la estimación de la cobertura tiene la misma precisión que el modelo de Morgensztern, basta comparar los resultados con el cuadro 5.44 para darse cuenta de que el nuevo modelo por nosotros propuesto, mejora sustancialmente el error cometido en la estimación de la distribución de contactos por el que mejor funcionamiento demostraba, el modelo de Metheringham, tal y como se ilustra en la figura 5.19.

Figura 5.19. Funcionamiento comparado del nuevo modelo MBBD.



Fuente: Elaboración propia

Se finaliza, con esta aportación, la presentación de los resultados obtenidos esta tesis doctoral, en los cuales se han evaluado y jerarquizado más de cuarenta modelos de exposición, determinando simultáneamente los factores determinantes del error cometido. Se cubre, de esta forma, la necesidad existente de en el ámbito académico, puesta de manifiesto por Leckenby y Ju (1990), del desarrollo de estudios de carácter cross-cultural, que permitieran evaluar la capacidad predictiva de los modelos frente a bases de datos que reflejen realidades socioculturales diferentes

a Estados Unidos país en el que, como se vio con anterioridad, habían tenido lugar hasta el momento, la mayoría de estudios similares a éste.

Además de clasificar a los modelos atendiendo al error global cometido, se ha explicado éste en base a los factores de error derivados de las características estructurales de los planes de medios. Se ofrece, de este modo, un criterio adicional para seleccionar el mejor modelo cuando sea necesario elegir entre aquellos que, globalmente, demuestran un funcionamiento que no es significativamente diferente.

Finalmente, tras el análisis de los resultados obtenidos, se ha propuesto un nuevo modelo de audiencia neta acumulada que mejora notablemente las estimaciones de los modelos tradicionales.

Conclusiones

A lo largo de la presente tesis doctoral, se ha pretendido profundizar en uno de los elementos más destacados en el ámbito de las decisiones publicitarias que deben abordar las organizaciones, la elección de los medios y soportes que han de servir de base al desarrollo de las campañas de publicidad, esto es, la planificación de medios y, dentro de ella, uno de sus instrumentos analíticos centrales: los modelos de determinación de la cobertura y la distribución de contactos.

La finalidad última de este trabajo, presidido siempre por el principio de conexión entre los ámbitos académico y gerencial, ha sido la de contribuir desde el primero, a la mejora de las herramientas de las que se dispone en las organizaciones para la gestión de una tarea compleja, pero tremendamente importante para la correcta asignación de los recursos publicitarios, la planificación de medios.

Con esta finalidad, el trabajo se ha estructurado en cinco capítulos, el primero de los cuales ha situado la planificación de medios en el ámbito de la toma de decisiones de comunicación de la empresa, destacando la influencia que, sobre la realización del plan de medios y soportes, ejercen las decisiones adoptadas en el resto de etapas del proceso. Seguidamente, en base a la importancia creciente que, derivada del incremento de la turbulencia del entorno mediático y del fuerte crecimiento de las inversiones publicitarias, está adquiriendo la tarea del planificador, se ha hecho hincapié en la necesidad de formalizar este proceso mediante modelización. El capítulo ha finalizado con una revisión del peso cuantitativo y cualitativo de los trabajos sobre modelización en la planificación de medios respecto a la modelización en comunicación.

El capítulo segundo ha ofrecido una tipología de los modelos de exposición basada en las hipótesis que los sustentan y que ha servido de estructura para el desarrollo de los modelos llevado a cabo en el capítulo tercero. En él se han derivado las expresiones para el cálculo de la cobertura y la distribución de contactos en cada uno de los modelos de exposición, expresiones que han servido de base para la posterior informatización de los mismos.

En el capítulo cuarto se han planteado los objetivos de la tesis doctoral, diseñando a continuación una metodología para la evaluación de los modelos, buscando que ésta asegurara la comparabilidad de los resultados, a la vez que estuviera adaptada a las particularidades propias del caso español.

Finalmente, en el capítulo quinto, se presentan los resultados obtenidos, cuyas conclusiones fundamentales serán expuestas a continuación.

La primera cuestión que se ha abordado, hace referencia al análisis del interés que pueda tener, tanto en el entorno académico como en el gerencial, el tema objeto de esta tesis doctoral. Respondiendo a la misma, se ha puesto de manifiesto (§ 1.2) el

importante crecimiento de la inversión publicitaria en España, desde la década de los ochenta hasta nuestros días, alcanzando en este momento cifras superiores a los seiscientos mil millones de pesetas (Infoadex, 1997), solamente en medios convencionales: diarios, suplementos y dominicales, revistas, radio, cine, televisión y publicidad exterior. Ante tales cifras, crece la necesidad manifestada tanto en el ámbito académico (Leckenby y Kim, 1994a; Leckenby y Kishi, 1982a; Kreshel Lancaster y Toomey, 1985; Lancaster Kreshel y Harris, 1986), como en el empresarial (Sperling, 1996; Sylvester, 1991; Abrams, 1991), de garantizar de alguna manera que la planificación de medios asegure la asignación eficiente de esos recursos.

La turbulencia del entorno mediático está, sin embargo, añadiendo complejidad a la labor del planificador de medios. Solamente en nuestro país, en la última década, se ha producido un fuerte incremento de la oferta de medios y de soportes dentro de cada medio, hecho que ha fragmentado la audiencia y reducido la fidelidad de ésta, provocando, simultáneamente, la aparición de prácticas comerciales muy agresivas que han redundado en un fuerte abaratamiento de los espacios publicitarios, la saturación de los mismos y, de manera derivada, una previsible reducción de la eficacia de la publicidad en ellos inserta.

Estos hechos, encadenados, se han traducido en la necesidad de crear nuevos procedimientos analíticos, o mejorar los existentes, de tal forma que su uso permita abordar con garantías la tarea de planificar los medios eficaz y eficientemente, y esta tarea ha de desarrollarse, fundamentalmente, en el ámbito de la universidad (Departamento de Publicidad de la Universidad de Texas en Austin, 1995).

Un elemento clave, sino el más importante, de estos procedimientos analíticos, lo constituyen la determinación de la cobertura y la distribución de contactos logrados por una campaña de publicidad (§ 1.3), a partir del momento en que sirven para discriminar entre soportes y determinar si uno es más adecuado que

otro para formar parte del plan de medios (Martínez Ramos, 1992). Este hecho deriva de que, cualquier medida de eficacia publicitaria que se incorpore al proceso de selección de los soportes, pasa por estimar de alguna manera, la exposición al anuncio a partir de la exposición al soporte obtenida mediante el cálculo de mencionadas variables.

Este papel central de la estimación de la cobertura y la distribución de contactos, se ha manifestado en el gran peso, tanto cuantitativo (Stautamoyer y Lindquist, 1994) como cualitativo (Pasadeos, Barban, Yi y Kim, 1997), de los modelos de exposición dentro del ámbito de la modelización en planificación de medios (§ 1.4). De la revisión de la literatura, se puede concluir que, aunque la etapa más prolífica corresponde a la creación de nuevos modelos llevada a cabo en los años sesenta, en la actualidad el peso del esfuerzo académico, se ha desplazado hacia el análisis del uso de los mismos por parte de las empresas y, sobre todo, a la evaluación del funcionamiento de los distintos modelos, para determinar cuáles son los más precisos en las estimaciones, para dar así respuesta a las exigencias procedentes del ámbito empresarial ya señaladas.

En nuestro país, se ha avanzado lentamente en este tema. Algunos trabajos se han llevado a cabo en el análisis de los modelos completos de planificación de medios (Bigné 1990b; 1991b) o en la creación de nuevos modelos de este tipo (Díez de Castro y Martín Armario, 1994). Respecto a los modelos de exposición, son muy meritorios los trabajos pioneros de León (1986) o Barreiro (1991) que describen el funcionamiento de, respectivamente, los modelos Dirichlet multinomial (Chandon, 1985) y Hofmans (1966). Recientemente, en su tesis doctoral, Redondo (1995) ha abordado la necesaria tarea de contrastar la precisión de los modelos de exposición, utilizando como base de comparación el conjunto de programas TOM.

Bajo esta doble perspectiva, elevado interés académico y práctico de su objeto de estudio, a la vez que un limitado desarrollo del tema en nuestro país, se

inscribe la realización de esta tesis doctoral, en la cual se han clasificado más de cuarenta modelos de exposición (§ 2), atendiendo a las hipótesis en que se sustentaban para, a continuación (§ 3), describir en profundidad la construcción y características de cada uno de ellos. A partir de este análisis, se formularon dos *cuestiones a investigar*:

- a. La primera de ellas (CI_1) se centraba en evaluar cada uno de los modelos, atendiendo al error que cometían en la estimación de la cobertura y la distribución de contactos (CI_{1A}) para, con esta información, proceder a su jerarquización (CI_{1B}), ofreciendo así un criterio para la elección del modelo más preciso y, también, las directrices para la creación de nuevos modelos. Estas directrices, se fundamentan en el análisis de la influencia que, la ruptura sucesiva de las hipótesis simplificadoras en las que se basan los distintos modelos, tiene sobre la precisión en las estimaciones de los mismos (CI_{1C}).
- b. La segunda cuestión a investigar (CI_2), indaga sobre la influencia que ejercen sobre la precisión de los modelos, las características internas de los planes de medios (CI_{2A}), medidas en términos de audiencia bruta, nivel de los fenómenos de acumulación y duplicación y tamaño relativo de los soportes. Dado que la mayoría de ellos fueron diseñados inicialmente para el medio revistas, se analiza su adecuación a un medio distinto, como son los diarios (CI_{2B}). A partir de ambos resultados, se establece un criterio adicional para discriminar entre aquellos modelos que, de acuerdo con la jerarquización inicial, demostraban un funcionamiento idéntico (CI_{2C}).

Para dar respuestas a estas cuestiones a investigar, ha sido necesario llevar a cabo un diseño metodológico (§ 4) que cumpliera dos condiciones fundamentales. Por un lado, que se adaptase a las particularidades de la información disponible en

nuestro país, fundamentalmente el Estudio General de Medios y, con ello, se asegurase su aplicabilidad. Por otro lado que garantizase la comparabilidad de los resultados con los obtenidos en los principales estudios llevados a cabo sobre esta materia. Para ello, de la revisión de la literatura efectuada, se han extraído los requisitos mínimos que debía cumplirse en materia de: tamaño muestral, técnicas de análisis de datos y medidas de error a emplear.

En este sentido, aunque en términos generales el procedimiento de contraste es sencillo: comparar la cobertura y distribución reales, medidas mediante un estudio de audiencias, con las estimaciones que proporcionan los diferentes modelos, ha sido necesario recodificar determinadas preguntas del cuestionario del EGM para convertir en cifras objetivas de número de contactos, respuestas que, de otro modo, no podrían asignarse a un nivel de contactos concreto. Las preguntas realizadas por el EGM permiten, con la recodificación señalada, obtener el número de contactos que ha recibido un individuo. Sin embargo, cuando el plan de medios está formado por un número de inserciones inferior al máximo posible (siete en una semana en diarios, cuatro en un mes en revistas semanales, seis en dos meses en revistas mensuales), no se puede saber si los contactos recibidos por el individuo, se corresponden con las inserciones realizadas por el planificador. Por este motivo, una segunda innovación metodológica propuesta por nosotros, ha consistido en saturar el espacio temporal con el máximo número de inserciones posibles para asegurar, de esta manera, que todos los contactos recibidos por el individuo, se corresponden con inserciones llevadas a cabo en el plan de medios.

Creemos que este diseño, con las mejoras que sean procedentes, constituye un punto de partida para los trabajos futuros sobre la evaluación de los modelos de exposición, debiendo abandonarse la práctica de recurrir a tomar como base de referencia para comparar las estimaciones de los modelos de exposición, los datos de algún modelo comercial, normalmente los programas TOM. Estos programas, asumen sus propias hipótesis simplificadoras e incorporan a otros modelos en sus

algoritmos, pudiendo provocar que la jerarquización obtenida derive más de similitudes en los procesos de cálculo, que de las diferencias de las estimaciones con la realidad.

Llevadas a cabo adecuaciones metodológicas provocadas por las peculiaridades de la base de datos empleada que se han señalado, se ha seguido en todo momento las líneas que, sobre el diseño experimental en este tipo de trabajos, venían marcadas por la revisión de la literatura efectuada. Este proceder viene impuesto porque, según Leckenby y Boyd (1990) o Leckenby y Kim (1992), la divergencias de los diseños experimentales en lo que hace referencia, además de la base de datos empleada, al número de planes de medios construidos para la evaluación de los modelos y la elección de la medida de error a utilizar, suponen la mayor dificultad para la comparabilidad de los resultados obtenidos por los diferentes estudios. Por esta razón, la construcción de 49 planes de medios para evaluar los modelos de acumulación de audiencias y de 99 para los modelos de audiencia neta y audiencia neta acumulada, así como la elección de los errores promedio (AER), relativo (RER), cuadrático medio (MSE) y error en las probabilidades de exposición (EPOR), viene fundamentado en las conclusiones obtenidas en la revisión de la literatura existente.

Respecto a las técnicas de análisis de datos utilizadas, en esta tesis doctoral nos hemos ceñido, de nuevo en aras a la comparabilidad de los resultados, a emplear aquellas que, en la revisión de la literatura llevada a cabo, demostraron ser las herramientas consideradas como más adecuadas por la mayoría de los autores. En este sentido, para establecer si las diferencias de las medias del error cometido por los diferentes modelos eran significativamente distintas, se ha recurrido al análisis de varianza de un factor. Para establecer grupos homogéneos de modelos, en base a que el error cometido por ellos era estadísticamente el mismo, se ha empleado la prueba *post hoc* de Tukey. Finalmente, para analizar la influencia que las características

estructurales de los planes de medios sobre el error, la herramienta empleada ha sido la regresión lineal.

En cuanto a los resultados obtenidos (§ 5), en el *proceso de evaluación y jerarquización* (CI₁), se ha llegado a las conclusiones que, a continuación, se exponen.

Respecto a la *estimación de la cobertura*, analizados los modelos que han demostrado un mejor comportamiento, puede establecerse que:

- a. Los modelos *ad hoc*, pese a ser modelos de construcción muy sencilla, y no realizar hipótesis alguna respecto a las probabilidades de exposición de los individuos y soportes, demuestran una gran precisión en sus predicciones de la cobertura. Así, de los siete modelos que han demostrado mayor precisión, cuatro de ellos, Hofmans (1966) aplicado a la acumulación de audiencias, exponencial (Chandon, 1985), Hofmans (1966) aplicado a la duplicación de audiencias y Morgensztern (1970), son modelos *ad hoc*.
- b. De los modelos estocásticos, los que aportan mayor fiabilidad en la estimación de la cobertura son, fundamentalmente, aquellos basados en la distribución beta binomial: Metheringham (1964) y beta binomial modificado (Chandon, 1985).

Cuando la variable a estimar es la *distribución de contactos*, del análisis de los modelos que demuestran un mejor comportamiento, puede afirmarse que:

- a. Basar la construcción del modelo en la distribución beta binomial, parece ser una garantía de precisión en las estimaciones, como lo demuestra el hecho de que la mitad de los modelos más precisos están basados en esta distribución

- b. Si se abandona esta distribución de probabilidad, no debe recurrirse en ningún caso a la utilización de la distribución binomial, o derivadas compuestas de la misma como la Poisson binomial (Leckenby y Ju, 1989), dado que ninguno de los modelos fundamentados en ella obtienen buenos resultados. Sólo la distribución geométrica, propuesta por Leckenby y Kishi (1981) ofrece en alguno de sus modelos, como el Kwerel geométrico, resultados similares a la beta binomial.

De la síntesis de los anteriores resultados, se ha inducido en esta tesis doctoral un hecho que, bajo nuestra perspectiva, debe fundamentar el desarrollo de futuros modelos de exposición, y que proponemos a continuación:

En la creación de nuevos modelos de exposición, deberá aprovecharse, simultáneamente, la buena capacidad en la estimación de la cobertura demostrada por los modelos ad hoc, y la adecuación de la distribución beta binomial para replicar los perfiles de la distribución de contactos.

En relación con la cuestión a investigar CI_1 , se pretendía establecer la si la eliminación de hipótesis simplificadoras en los modelos estocásticos, se traducían en mejora de la precisión de las estimaciones. De los resultados obtenidos, puede concluirse que, las únicas hipótesis que, claramente, se demuestran demasiado restrictivas y derivan en mal comportamiento de los modelos que las sustentan, son las que suponen la aleatoriedad de los fenómenos de duplicación y acumulación de las audiencias. Su no consideración, se traduce en mejoras sustantivas en la precisión de los modelos que así actúan. Si se analizan los cuadros 2.10, 2.11 y 2.12, en los cuales se resumían las hipótesis en las que se basaban cada uno de los modelos estocásticos, se observa que son, precisamente, los modelos basados en la distribución binomial, los únicos que consideran el carácter aleatorio de los

fenómenos de acumulación y duplicación. Este hecho nos reafirma en la conclusión a la que ya llegamos anteriormente, de que la creación de nuevos modelos, nunca debe pasar por la sustitución de la distribución beta binomial por la binomial o cualquiera de sus variantes.

También se ha obtenido que, el asumir el carácter heterogéneo de la población sólo redundará en mejoras en las estimaciones en los modelos de acumulación y duplicación de audiencias, mientras que no ejerce influencia alguna en los modelos de audiencia neta acumulada. Puesto que los planes que habitualmente se construirán en la práctica, responderán a este último tipo, no parece aventurado concluir que:

Es prácticamente nula la mejora de la precisión de las estimaciones de la cobertura y distribución de contactos, que se deriva de añadir complejidad en la creación de modelos de exposición, al suprimir las hipótesis de homogeneidad de la población y los soportes y no estacionariedad de las inserciones. Basta que los modelos, no estén basados en la distribución binomial o en otras distribuciones que asuman el carácter no aleatorio de los fenómenos de duplicación y acumulación de las audiencias.

La cuestión a investigar CI₂, pretendía establecer la influencia que ejercen las características estructurales de los planes de medios, en términos de audiencia bruta, niveles de acumulación y duplicación y tamaño relativo de los soportes, sobre la precisión de las estimaciones. El objetivo último de este análisis, era ser capaz de ofrecer un criterio adicional para discriminar entre aquellos modelos que, de acuerdo con la jerarquización obtenida en los análisis anteriores, no cometieran errores significativamente distintos. Para este fin, se regresaron los errores cometidos en la estimación de la cobertura y la distribución de contactos, tomando como variables independientes las características estructurales señaladas con anterioridad. De los

resultados que se obtuvieron en el capítulo cinco, pueden extraerse las siguientes conclusiones.

Respecto a la precisión en la estimación de la cobertura:

- a. De los tres modelos de acumulación de audiencias que demuestran una mayor precisión: Hofmans, exponencial y beta binomial modificado, sólo el error cometido por el modelo de Hofmans se muestra insensible a variaciones en las características estructurales de los planes de medios. Dado que los coeficientes estandarizados del resto de modelos son positivos, el error cometido por éstos se incrementará al aumentar el número de inserciones (modelo exponencial) o la audiencia bruta del plan de medios (beta binomial modificado). Por ello, cuando se trate de evaluar la cobertura de planes de medios formados por varias inserciones en un sólo soporte, se recomienda la utilización del modelo de Hofmans.
- b. Los modelos de duplicación de audiencias que cometen errores análogos en la estimación de la cobertura: Metheringham, Hofmans y Hofmans ampliado. Del análisis de los factores de error, se concluye que el error cometido por el modelo de Metheringham crece más rápidamente que el del resto de modelos cuando se incrementa la audiencia bruta, pero que es insensible al número de soportes y, sobre todo, disminuye al aumentar la audiencia duplicada en mayor proporción que aquéllos. Por todo ello, será siempre preferible el modelo de Metheringham, salvo cuando los planes de medios se construyan utilizando soportes que provean duplicaciones bajas entre ellos debido, por ejemplo, a un contenido editorial muy distinto, pues ello hará incrementar la audiencia bruta más rápidamente que la suma de las audiencias duplicadas. En ese caso se recurriría indistintamente a cualquiera de las dos fórmulas de Hofmans.

- c. En cuanto a los modelos de audiencia neta acumulada, no procede al análisis precedente, dado que un único modelo, el de Morgensztern, proporciona las mejores estimaciones de la cobertura con lo que, indistintamente de las características del plan de medios, será éste el más adecuado.

Cuando se analiza la influencia de las características internas de los planes de medios sobre el error cometido por los diferentes modelos en la estimación de la distribución de contactos, de los resultados obtenidos en esta tesis doctoral, podemos llegar a las siguientes conclusiones:

- a. En los modelos de acumulación de audiencias, sólo existe un modelo que proporciona los mejores resultados, el modelo beta binomial modificado, por lo que es irrelevante el resultado del análisis de los factores de error.
- b. En los modelos de duplicación de audiencias, es necesario discriminar entre tres modelos: el de Hofmans ampliado, el modelo de Metheringham y el modelo de agregación secuencial con duplicación variable. Si la variable que se analiza es el error en las probabilidades de exposición, el modelo de Metheringham siempre será preferible a los otros dos, dado que los coeficientes estandarizados de las variables significativas son muy parecidos. Sin embargo, este modelo no es sensible a variaciones en el número de soportes, lo que no ocurre con el modelo de Hofmans ampliado y el de agregación secuencial. Cuando se analiza el error cuadrático medio la conclusión a la que se llega es la misma, con la única salvedad de que el error cometido por Metheringham crece más que el resto cuando se incrementa la audiencia bruta. Dado que el error disminuye al hacerlo la audiencia duplicada, concluiremos lo mismo que en la estimación de la cobertura: el modelo de Metheringham será elegido, salvo cuando el plan

de medios está formado por soportes entre los que se prevea duplicaciones bajas.

- c. Finalmente, dos modelos de audiencia neta acumulada ofrecen resultados equivalentes: el de Metheringham y el Kwerel geométrico, siendo las conclusiones muy similares a las ya expuestas. Tomando como medida del error, el cometido en las probabilidades de exposición, el modelo de Metheringham siempre será el preferido, dado que la única variable a la que es sensible, la suma de las audiencias acumuladas, hace disminuir el error cuando ésta aumenta. Cuando se recurre al error cuadrático medio, éste disminuye cuando se incrementa la suma de las audiencias acumuladas y duplicadas, pero aumenta cuando lo hace la audiencia bruta en mayor medida que el modelo de Kwerel. Luego recurriremos, de nuevo, a Metheringham, siempre que se trabaje con soportes cuyo contenido editorial sea similar y, por ello, tenga una mayor probabilidad de derivar en duplicaciones elevadas.

En relación con los resultados del análisis de los factores de error, nos reafirman en las conclusiones expuestas anteriormente. Vemos que, globalmente, la cobertura es siempre mejor estimada por modelos *ad hoc* y, cuando no es así, se recurre a un modelo basado en la distribución beta binomial. Asimismo, la estimación de las distribuciones de contactos es una labor donde los modelos basados en la distribución beta binomial, ya sea el beta binomial modificado o el de Metheringham, demuestran una mayor precisión.

El último objetivo que pretendíamos alcanzar en la segunda cuestión a investigar CI_2 , era la de establecer si, dado que la mayoría de los modelos expuestos, habían sido desarrollados pensando en su aplicación al medio revistas, cometerían un error significativamente mayor al estimar la cobertura y distribución de contactos de

planes de medios formados por diarios (CI_{2B}). Para verificar esta posibilidad, se llevó a cabo un análisis de varianza de las distintas medidas del error, tomando como factor el tipo de medio que formaba cada plan, esto es, diarios, revistas mensuales y semanales.

En todos los casos analizados, las diferencias de medias del error han sido significativamente distintas, sin embargo, el análisis *post hoc* realizado, ha puesto de manifiesto qué tipo de medio era el que provocaba esas diferencias. En este sentido, se han obtenido las siguientes conclusiones:

- a. En los modelos de acumulación de audiencias, a la hora de estimar la cobertura, cuando el plan de medios está formado por diarios, el error es siempre mayor. Sin embargo, también lo es cuando está formado por revistas mensuales frente a las semanales, por lo que no puede atribuirse a un comportamiento distinto entre las revistas y los diarios. En la estimación de la distribución de contactos, el resultado es más claro dado que el comportamiento de los diarios es el mismo que el de las revistas mensuales cuando se considera MSEJ como medida de error, y que el de las semanales cuando se considera EPORJ.
- b. Si se consideran los resultados obtenidos con los modelos de duplicación de audiencias y los de audiencia neta acumulada, el error cometido en planes formados exclusivamente por periódicos, es siempre menor que el que se comete cuando los planes están formados por soportes del medio revistas. Esta afirmación es válida, tanto para la estimación de la cobertura como para la distribución de contactos.

Por todo lo expuesto, podemos proponer que:

La evaluación y jerarquización realizadas, así como los criterios derivados para la selección y creación de nuevos modelos de exposición, son también válidos para planes de medios formados por periódicos aunque, originalmente, estos modelos fueran concebidos para el medio revistas.

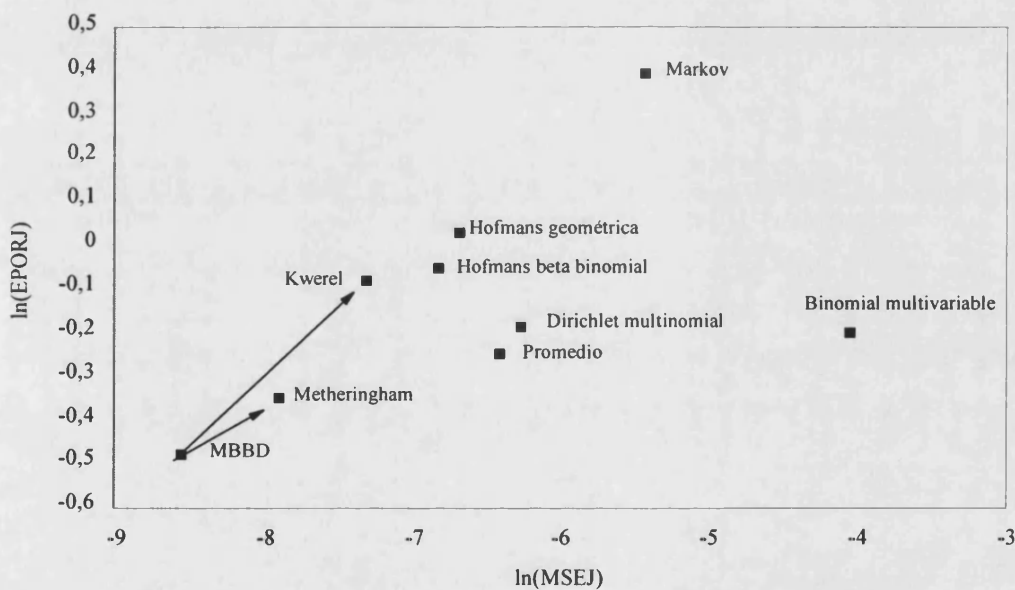
El último objetivo que se propuso en esta tesis doctoral, era verificar la fiabilidad de las conclusiones obtenidas. Para ello se ha construido un nuevo modelo de audiencia neta acumulada que cumplía las recomendaciones que acabamos de exponer. Si éstas son válidas, este modelo debiera ser más preciso en las estimaciones de la cobertura y la distribución de contactos que los que han sido evaluados. Bajo esta perspectiva, el modelo construido, está basado en la estimación de la cobertura mediante la fórmula *ad hoc* que mejores resultados había ofrecido, la de Morgensztern para, a continuación, derivar la distribución de contactos a través de una distribución beta binomial que, reiteradamente, había demostrado ser la más adecuada para este fin.

Cuando se calcula por separado la cobertura de un plan de medios y su distribución de contactos, puede incurrirse, como se señaló, en que los resultados no sean coherentes, dado que el nivel cero de la distribución de contactos (los individuos que no han recibido ninguna exposición) ha de coincidir con el resultado de restar de uno, la cobertura del plan de medios. Para evitar esta posible incoherencia, se recurrió al procedimiento propuesto por Leckenby y Boyd (1984a) en el desarrollo del modelo Hofmans beta binomial, conocido como de medias y ceros, consistente en ir ajustando iterativamente los parámetros de la distribución beta binomial, hasta que el nivel cero de la distribución resulte coherente con la cobertura estimada mediante la fórmula de Morgensztern. Este modelo, que hemos denominado *Modelo Morgensztern beta binomial (MBBD)*, cumple, por su construcción, todas las recomendaciones que se han ido señalando en estas conclusiones y, como se demostró en el capítulo 5, y se ilustra en la figura C.1,

mejora notablemente al mejor de los modelos tradicionales evaluados. Por ello, concluimos finalmente que:

Las recomendaciones para la creación de nuevos modelos, derivadas del proceso de evaluación y jerarquización de los modelos de exposición tradicionales, demuestran su validez a partir del momento en que un nuevo modelo construido bajo sus directrices (MBBD) consigue una mayor precisión en la estimación de la cobertura y la distribución de contactos que los modelos tradicionales.

Figura C.1 Funcionamiento comparado del nuevo modelo MBBD.



Fuente: Elaboración propia

Se finaliza, con esta aportación, el resumen de los resultados obtenidos esta tesis doctoral, en los cuales se han evaluado y jerarquizado más de cuarenta modelos de exposición, determinando simultáneamente los factores determinantes del error cometido. Se cubre, de esta forma, la necesidad existente en el ámbito académico, puesta de manifiesto por Leckenby y Ju (1990), del desarrollo de estudios, que

permitieran evaluar la capacidad predictiva de los modelos frente a bases de datos que reflejen realidades socioculturales diferentes a Estados Unidos, país en el que, como se vio en el desarrollo de este trabajo, habían tenido lugar la mayoría de estudios similares a éste.

En cuanto a las limitaciones de esta tesis doctoral, deben destacarse la siguientes. El diseño experimental propuesto permite la evaluación de los modelos de exposición mediante la utilización de planes de medios más reales que en los estudios estadounidenses señalados, dado que no tiene la limitación provocada por los estudios de audiencias basados en la metodología *Through The Book* de utilizar como máximo dos inserciones por soporte. Sin embargo, aun siendo más amplias las posibilidades, no es posible construir planes de medios que tengan más de siete inserciones en diarios, seis en revistas mensuales y cuatro en semanales. Creemos, pese a ello, que los niveles de cobertura y frecuencia que pueden alcanzarse con los planes así diseñados, están lo suficientemente cercanos a los logrados en campañas reales, como para que los resultados de la evaluación sean adecuados.

La segunda limitación que debe señalarse, es que la evaluación se ha limitado a los modelos creados para los medios impresos. La principal razón, como se señaló oportunamente, es la necesidad de parcelar el tema analizado para que este tuviera una extensión abordable con la suficiente profundidad. Sin embargo, es intención nuestra extender en futuros trabajos, los mecanismos de análisis desarrollados, a los modelos diseñados específicamente para los medios radio y televisión.

A modo de conclusión, podemos afirmar que la planificación de medios es una tarea cuya complejidad creciente marcha paralela a su relevancia para la correcta asignación de los recursos invertidos por las empresas en publicidad. Creemos que, desde el ámbito académico, es importante que se contribuya a dotarla de nuevos instrumentos analíticos que faciliten la labor del planificador y, también, que se trabaje en la determinación de la precisión de los existentes. Bajo este principio

inspirador, e intentando conectar siempre los ámbitos académico y gerencial, se ha inscrito el desarrollo de esta tesis doctoral que ha intentado, por un lado, establecer la precisión de las estimaciones de la cobertura y la distribución de contacto de los modelos de exposición existentes, para facilitar su elección por parte del planificador. Por otro lado, de los resultados obtenidos en esa labor jerarquizadora de los modelos, se ha buscado el establecimiento de una serie de directrices que ayuden a otros investigadores, al desarrollo de nuevos modelos de exposición.