Proves t de Student

Dra. Pilar Serra

Universitat de València

Prova t de Student

- Es tracta d’una prova de contrast d'hipòtesis sobre valors de diverses mitjanes:
  - H_0 (hipòtesi nul·la o d’igualtat): les mitjanes de dues mostres són iguals.
  - H_1 (hipòtesi alternativa o de diferència): les mitjanes de les mostres són significativament distintes.

Prova t de Student

- Prova t de Student per a mostres independents.

- Prova t de Student per a mostres relacionades.

- Prova t de Student per a una mostra.

Prova t de Student per a mostres independents

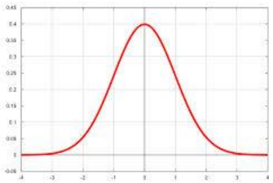
Introducció
Execució de la prova
Resultats
Escriptura dels resultats

VNIVERSITAT DE VALÈNCIA

Introducció

- Hipòtesi sobre els valors de dues mostres obtingudes de dues poblacions INDEPENDENTS.
- Exemple: valoració de l'amplitud de moviment anterior del centre de pressions en la sedestació entre persones amb paraplegia i sense.

Introducció



- Compliment de supòsits
  - NORMALITAT
    - Es comprova amb les proves de Kolmogorov-Smirnov o Shapiro-Wilk.
  - HOMOGENEÏTAT DE VARIÀNCIES
    - La variància d'una mostra ha de ser similar a la variància de l'altra mostra.
    - Es comprova amb la prova de Levene.
  - INDEPENDÈNCIA: puntuacions distintes perquè venen de persones distintes.

Introducció

- Que ocorre si no es compleixen...?
  - El supòsit de normalitat: la prova t de Student és robusta.
    - Alternatives
      - Eliminar valors extrems
      - Transformar dades
      - Usar la prova U de Mann-Whitney
  - El supòsit d'homoscedasticitat
    - Alternatives
      - Transformar dades
      - Usar una prova no paramètrica
  - El supòsit d'independència
    - Alternatives
      - Prova t per a mostres relacionades

Execució de la prova

- Estudi sobre el desplaçament anterior del centre de pressions en sedestació en persones sense paraplegia.

Resultats

- En la taula cal fixar-se en:
  - Comprovació del supòsit d'homoscedasticitat (prova de Levene).
  - La significació.
  - 95% d’interval de confiança per a la diferència.
    - Si entre els límits inferior i superior hi ha el 0, no hi haurà diferències estadísticament significatives.
- La potència s’ha obtenir:

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

Escriptura dels resultats

- Els participants sense paraplegia són capaços de fer un desplaçament anterior significativament més gran (M = 110,14, DE =) que les persones amb paraplegia (M = 35,51, DE =, $t(46)$= 12,50, p<0,05, r = 0,88)

  M = mitjana

  DE = desviació estàndard

  t = prova t

  () = graus de llibertat

  p = nivell de significació

  r = potència

***r* = 0,10 (petita)**: l'efecte explica l'1% de la variància total
***r* = 0,30 (mitjana)**: l'efecte explica el 9% de la variància total
***r* = 0,50 (gran)**: l'efecte explica el 25% de la variància total

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

Prova t de Student per a mostres relacionades

Introducció
Execució de la prova
Resultats
Escriptura dels resultats

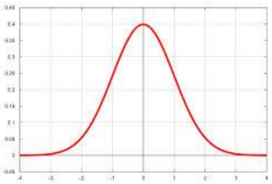
VNIVERSITAT DE VALÈNCIA

Introducció

- Hipòtesi sobre els valors de dues mostres RELACIONADES.
- Exemples
  - 1. Valorar l'amplitud de moviment del centre de pressions en la sedestació en persones amb paraplegia abans d'una intervenció terapèutica i després.
  - 2. Comparar la força de flexió del colze amb el braç dominant i el no dominant.
  - 3. Comparar la capacitat per a mantenir l'equilibri en condicions diferents (ulls oberts i ulls tancats).

Només hi ha un grup de persones, però les mesurem dues vegades.

Introducció



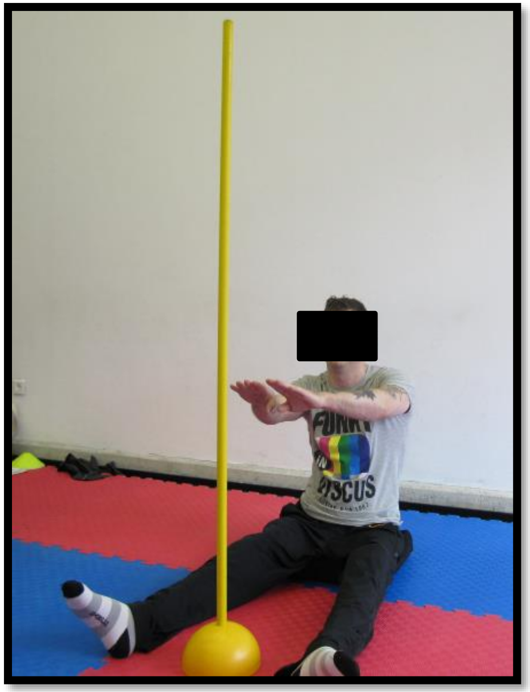
- Compliment dels supòsits
  - NORMALITAT
    - Es comprova amb Kolmogorov-Smirnov o Shapiro-Wilk.

- Què ocorre si no es compleix...?
  - El supòsit de normalitat: la prova t de Student és robusta.
    - Alternatives
      - Eliminar valors extrems
      - Transformar dades
      - Usar la prova de Wilcoxon

Universitat de València

Execució de la prova

- Estudi sobre la repercussió d'un programa terapèutic de control postural en l'equilibri de persones amb paraplegia.



Resultats

- En la taula cal fixar-se en:
  - La significació.
  - 95% d'interval de confiança per a la diferència
    - Si entre els límits inferior i superior hi ha el 0, no hi hauria diferències estadísticament significatives.
- La potència s'ha d'obtenir:

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

Escriptura dels resultats

- Els participants experimentaven un desplaçament anterior significativament més petit abans del tractament (*M* = 35,51 *SE* = 3,33) que després (*M*= 39,32, *SE*= 3,68, *t*(23) = -10,23, *p*<0,05, r = 0,90)

M = mitjana

DE = desviació estàndard

t = prova t

() = graus de llibertat

p = nivell de significació

r = potència

_r_ = 0,10 (petita): l'efecte explica l'1% de la variància total.
_r_ = 0,30 (mitjana): l'efecte explica el 9% de la variància total.
_r_ = 0,50 (gran): l'efecte explica el 25% de la variància total.

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

Prova t de Student per a una mostra

Introducció
Execució de la prova
Resultats
Escriptura dels resultats

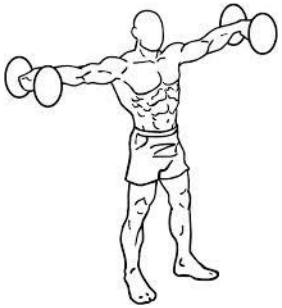
Universitat de València

Introducció

- Hipòtesi sobre els valors de les mitjanes poblacionals.

- Exemple: comparar els resultats obtinguts de la mostra d'un estudi propi amb el valor existent en la bibliografia.

Introducció



Valor = 65 Nm

Execució de la prova

- Estudi sobre desplaçament anterior del centre de pressions en sedestació en persones amb paraplegia comparat amb un valor acceptat per a la població, segons la bibliografia.

Resultats

- En la taula cal fixar-se en:
  - La significació
  - 95% interval de confiança per a la diferència
    - Si entre els límits inferior i superior hi ha el 0, no hi hauria diferències estadísticament significatives.
- La potència s'ha d'obtenir:

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

Escriptura dels resultats

- El desplaçament màxim en sedestació de persones amb paraplegia acceptat segons la bibliografia (*M* = 41,21) no difereix significativament del realitzat per la mostra de persones amb paraplegia del meu estudi (*M* = 35,51 *SE* = 3,31, *t*(23) = -1,71, *p*>0,05, *r* = 1)

M = mitjana

DE = desviació estàndard

t = prova t

() = graus de llibertat

p = nivell de significació

r = potència

r = 0,10 (petita): l'efecte explica l'1% de la variància total.
r = 0,30 (mitjana): l'efecte explica el 9% de la variància total.
r = 0,50 (gran): l'efecte explica el 25% de la variància total.

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

http://statisti.blogs.uv.es/

VNIVERSITAT DE VALÈNCIA

ANÀLISI DE VARIÀNCIA (ANOVA) D'UN FACTOR

Prof. Dra. Pilar Serra

VNIVERSITAT DE VALÈNCIA

ANOVA d'un factor

- Eina estadística que analitza la variància d'un factor per a una variable dependent quantitativa.

- Objectiu: contrastar la hipòtesi que unes quantes mitjanes (més de dues) són iguals (H_0).

- Aquesta tècnica és una extensió de la prova t.
  - Corregeix l'error que es deriva de les comparacions de les mitjanes de *k* grups o *k* repeticions a fi d'evitar augmentar la probabilitat d'error de tipus I.

ANOVA d’un factor

- Tipus

  – ANOVA d’un factor ENTRE SUBJECTES

  – ANOVA d’un factor DE MESURES REPETIDES

ANOVA D'UN FACTOR ENTRE SUBJECTES

Introducció
Execució de la prova
Resultats
Escriptura dels resultats

Vniversitat de València

Introducció

- **Suma de quadrats**
  - Sumatori de la diferència entre les dades observades i la mitjana elevada al quadrat.

$$SS = \sum (x_i - \bar{x})^2$$

- **Suma de quadrats total**
  - Es calcula la diferència entre les dades observades i la mitjana total (la mitjana de tots els grups que es vulga comparar).

$$SS_T = \sum (x_i - \bar{x}_{total})^2$$

Introducció

- **Graus de llibertat totals**: nombre d’observacions que són lliures de variar.

- Exemple
  - Si hi ha 39 subjectes i la mitjana total sabem quina és (73,45), només 38 poden prendre valors de forma lliure; el valor del 39 està condicionat segons els valors que han pres els altres.

Introducció

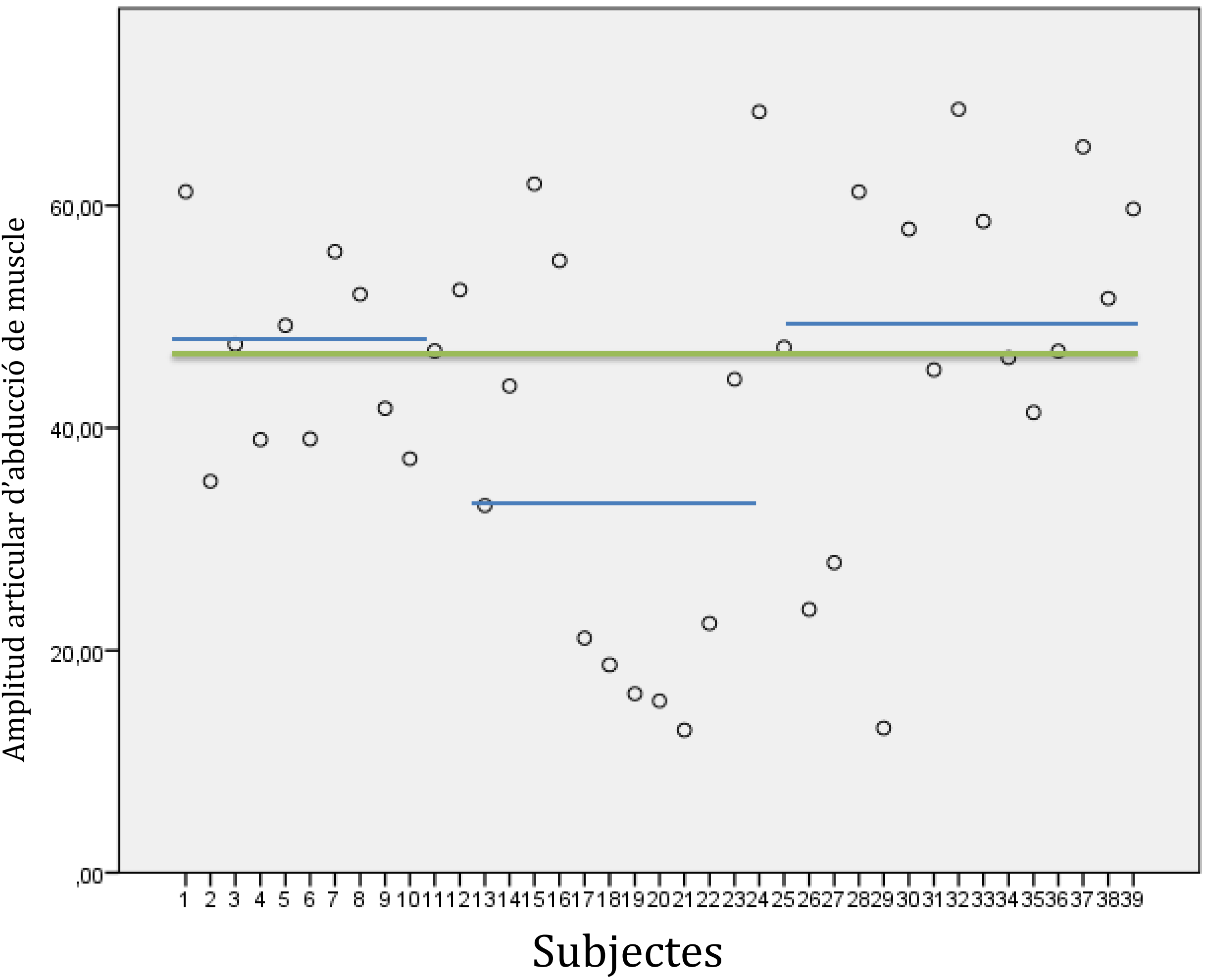
- La **variància** és la suma de quadrats dividida pel nombre d'observacions menys 1.

$$s^2 = SS/(N - 1)$$

- La **variància total** és la variació entre totes les puntuacions independentment de la condició experimental de la qual vinguen.

Introducció

Estudi de l'amplitud articular d'abducció de muscle



N. tendinitis del supraespinós = 11

N. ruptura parcial del supraespinós = 13

N. bursitis de muscle = 15

	Amplitud tendinitis	Amplitud ruptura	Amplitud bursitis
$\overline{X}$	46,42	33,43	50,42
S	9,09	17,45	15,18
S^2	82,69	370,34	230,37

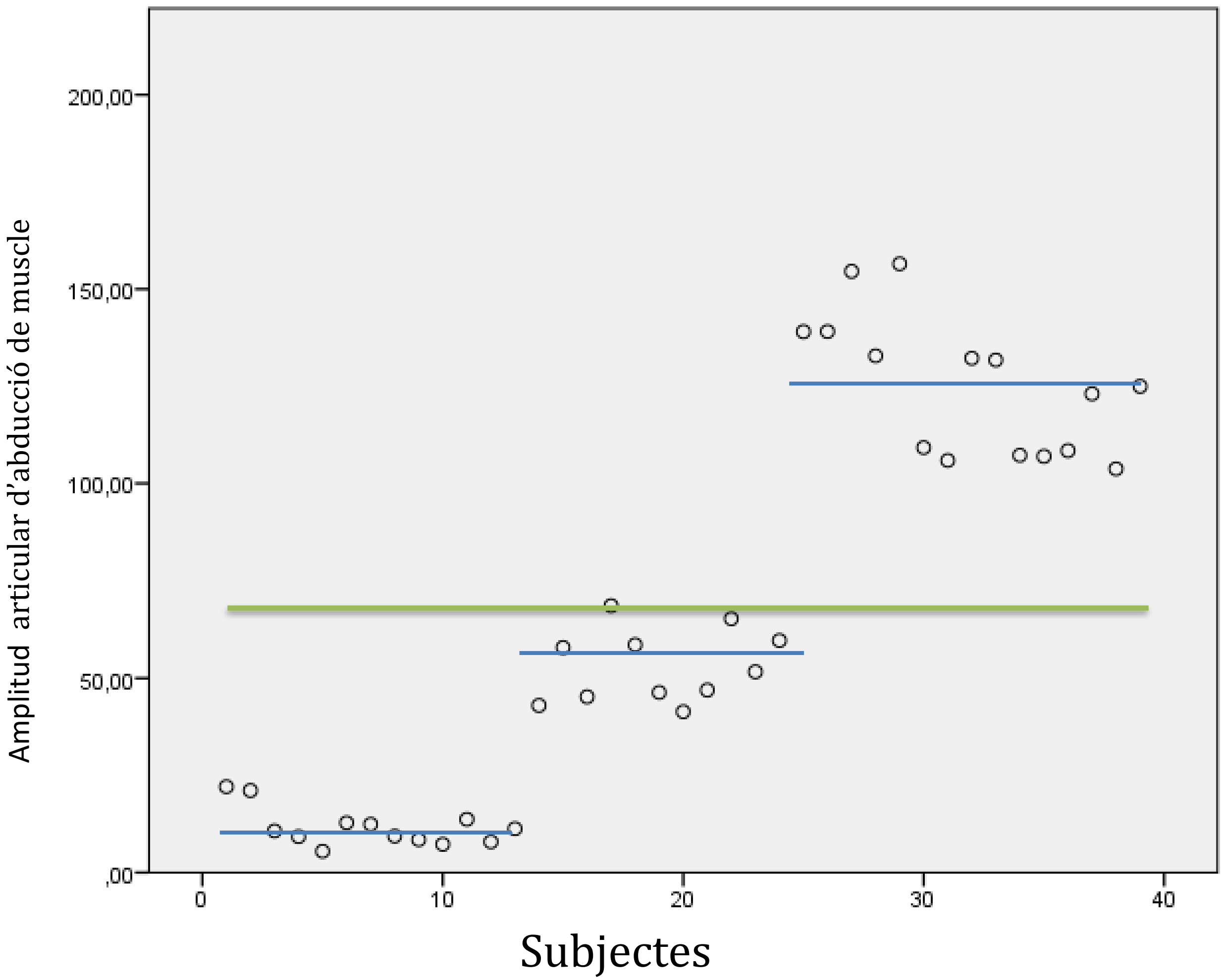
Mitjana total = 43,22

Desviació total = 16,12

Variància total = 259,87

Introducció

Estudi de l'amplitud articular de l'abducció de muscle



N. muscle congelat= 13

N. ruptura parcial supraespinós = 11

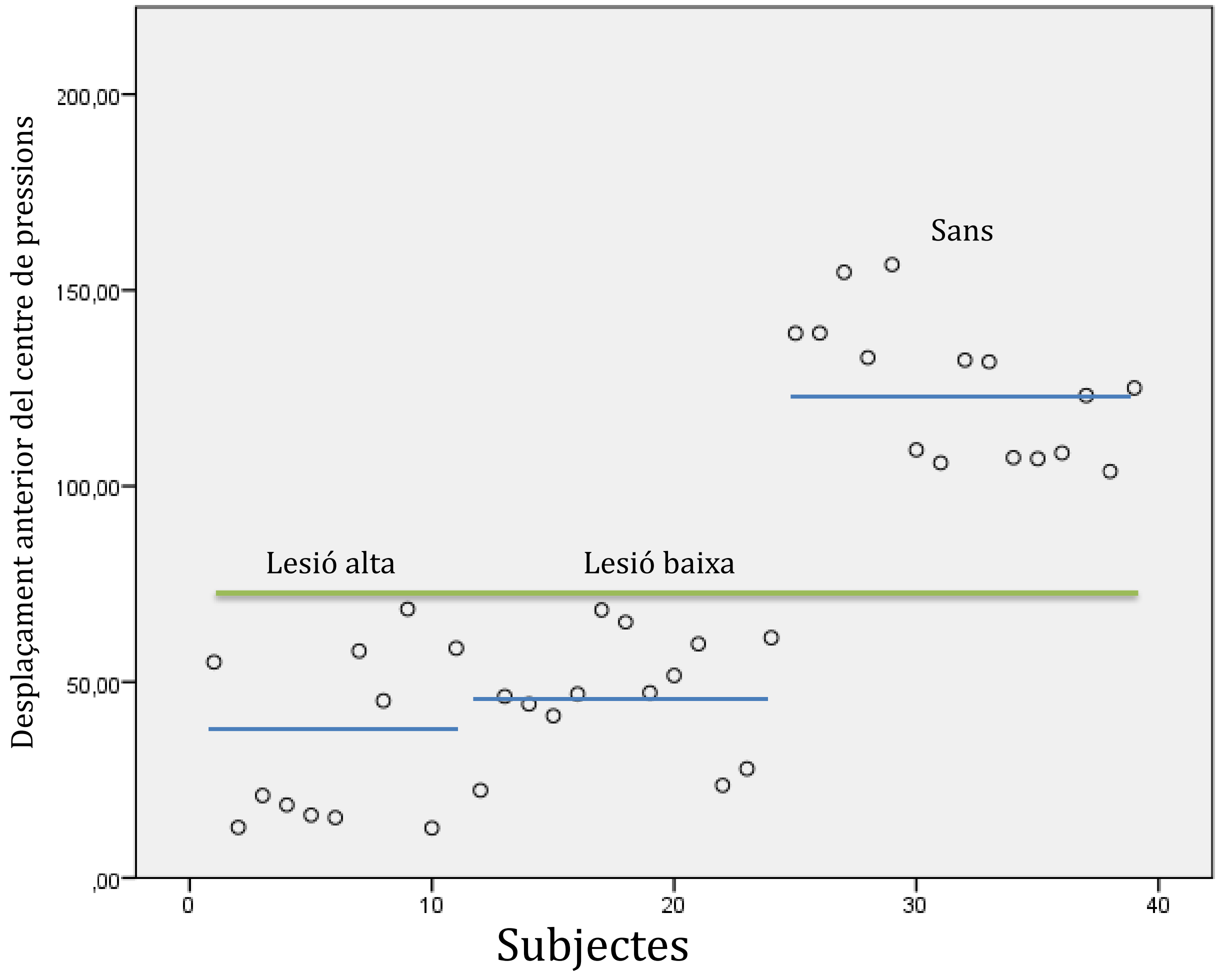
N. contractura trapezi superior = 15

	Amplitud muscle congelat	Amplitud ruptura	Amplitud contractura
$\overline{X}$	11,67	53,15	125,03
S	9,26	9,36	17,74
S^2	24,81	87,59	314,67

Mitjana total = 66,97
Desviació total = 50,78
Variància total = 2.578,84

Introducció

Estudi del desplaçament anterior del centre de pressions



N. lesió alta = 11
N. lesió baixa = 13
N. sans = 15

$$SS_T = \sum (x_i - \bar{x}_{total})^2$$

- Mitjana total
- Mitjana per a cada grup

	Desplaçament lesió baixa	Desplaçament lesió alta	Desplaçament sans
	22,42	55,07	139,03
	46,33	12,99	139,08
	44,37	21,09	154,56
	41,38	18,71	132,83
	46,93	16,11	156,49
	68,45	15,46	109,20
	65,31	57,89	105,89
	47,27	45,23	132,20
	51,64	68,68	131,75
	59,70	12,81	107,23
	23,69	58,58	106,95
	27,91		108,42
	61,26		123,06
			103,76
			124,98
$\bar{X}$	46,67	34,78	125,03
S	15,09	22,13	17,74
S^2	227,59	489,56	314,67
	Mitjana total = 73,45 Desviació total = 45,22 Variància total = 2.044,79		

Introducció

- L'anàlisi de variància permet comprovar:
  - **La variabilitat imputable a les condicions dels grups**.
  - **La variabilitat deguda a l'atzar.**

Variabilitat total = Σ (efectes atribuïbles) + Σ (efectes no atribuïbles/residuals)

$$SS_M = \sum n_k (\bar{x}_k - \bar{x}_{total})^2 \qquad SS_R = \sum (x_{ik} - \bar{x}_k)^2$$

Introducció

- **Suma de quadrats del model (*SS_M*)**
- Objectiu: saber quanta de la variància total es pot explicar amb el model de regressió.
  - Amb ANOVA: quin % de la variància total s'explica per les diferències entre els grups de comparació.
  - Els valors previstos pel model són la mitjana de cada grup.
    - Per cada participant, el valor previst pel model és la mitjana a la qual pertany aquest participant.

- Com es calcula
  - Es calcula la diferència entre les mitjanes de cada grup i la mitjana total.
  - S'eleven al quadrat aquestes diferències.
  - Cada resultat es multiplica pel nombre de participants de cada grup.
  - Se sumen els valors de cada grup.

$$SS_M = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$$

Introducció

$$SS_M = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$$

$$SS_M = 13(46{,}67 - 73{,}45)^2 + 11(34{,}67 - 73{,}45)^2 + 15(125{,}03 - 73{,}45)^2 = 66.557{,}77$$

	Desplaçament lesió baixa	Desplaçament lesió alta	Desplaçament sans
$\bar{X}$	46,67	34,78	125,03
S	15,09	22,13	17,74
S^2	227,59	489,56	314,67
	Mitjana total = 73,45 Desviació total = 45,22 Variància total = 2.044,79		

Introducció

- Els graus de llibertat de SS_M = nombre de grups de comparació menys 1.

  - $SS_M = k - 1$

    - Exemple: en un estudi amb tres grups de comparació (i. e. sense lesió, lesió toràcica baixa i lesió toràcica alta), el nombre de graus de llibertat és 2.

Introducció

- **Suma de quadrats residual (*SS_R*)**
- Objectiu: saber la quantitat de variància que no s'explica pel model.
  - Causada per factors estranys, diferències individuals, etc.

- Com es calcula
  - Es calcula la diferència entre la puntuació de cada subjecte amb la mesura del seu grup.
  - S'eleven al quadrat aquestes diferències.
  - Se sumen els valors de cada grup.

$$SS_R = \sum (x_{ik} - \bar{x}_k)^2$$

Introducció

- Els graus de llibertat de SS_R = graus de llibertat totals – graus de llibertat del model.

$$gl_R = gl_T - gl_M$$

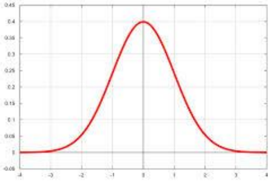
- gl_T = nombre de subjectes menys 1
- gl_M = nombre de grups menys 1

Introducció

- Les hipòtesis que es formulen són:
  - H_0 (hipòtesi nul·la o d'igualtat): les mitjanes de les mostres són iguals.
  - H_1 (hipòtesi alternativa o de diferència): les mitjanes de les mostres són significativament distintes.

Introducció

- Compliment dels supòsits



  - NORMALITAT
    - Es comprova amb Kolmogorov-Smirnov o Shapiro-Wilk.
  - HOMOGENEÏTAT DE LES VARIÀNCIES
    - La variància d'una mostra ha de ser similar a la variància de l'altra mostra.
    - Es comprova amb el test de Levene.
  - INDEPENDÈNCIA: puntuacions distintes perquè venen de persones distintes.

Introducció

- Què ocorre si no es compleix…?
  - El supòsit de normalitat: l'ANOVA és robusta
    - Alternatives
      - Eliminar valors extrems.
      - Transformar dades.
      - Usar la prova Kruskal-Wallis amb les *post-hoc* U de Mann-Whitney i ajust d'errors (taxa d'error de 0,05/nombre total de contrastos).
  - El supòsit d'independència
    - Alternatives
      - ANOVA de mesures repetides.

Introducció

- Què ocorre si no es compleix...?
  - El supòsit d'HOMOSCEDASTICITAT

Alternatives a la ràtio *F*:

$$SS_R = \sum (x_{ik} - \bar{x}_k)^2$$

$$SS_R = \sum s^2{}_k (n_k - 1)$$

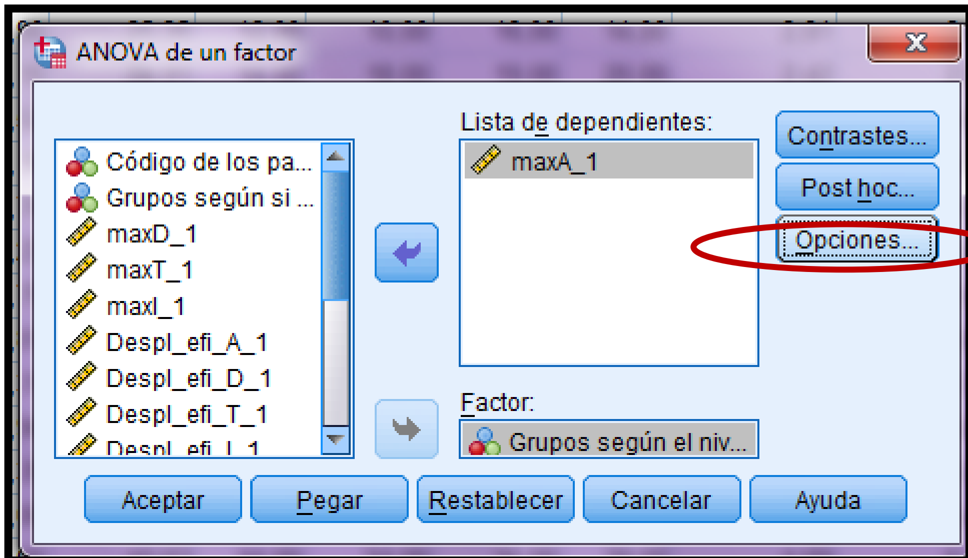
- Brown-Forsythe F.
- Welch F.

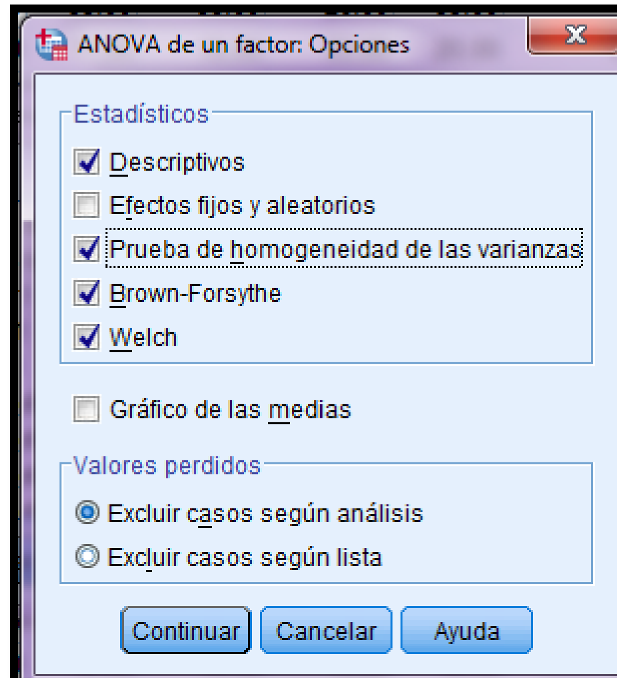
» Ambdós ajusten *F* tenint en compte els graus de llibertat, cosa que elimina els problemes d'incompliment d'homogeneïtat de variàncies.

» Ambdós controlen l'error de tipus I.

» Welch té més potència (i. e. detecta millor l'efecte quan aquest no existeix) llevat que la mesura siga molt gran i tinga una variància molt gran.

Introducció





El compliment del supòsit d'homoscedasticitat determina les proves *post hoc*

Introducció

- Ràtio F

$$F = \frac{MS_M}{MS_R}$$

$$MS_M = \frac{SS_M}{gl_M}$$

$$MS_R = \frac{SS_R}{gl_R}$$

VNIVERSITAT DE VALÈNCIA

Introducció

Proves *post hoc*

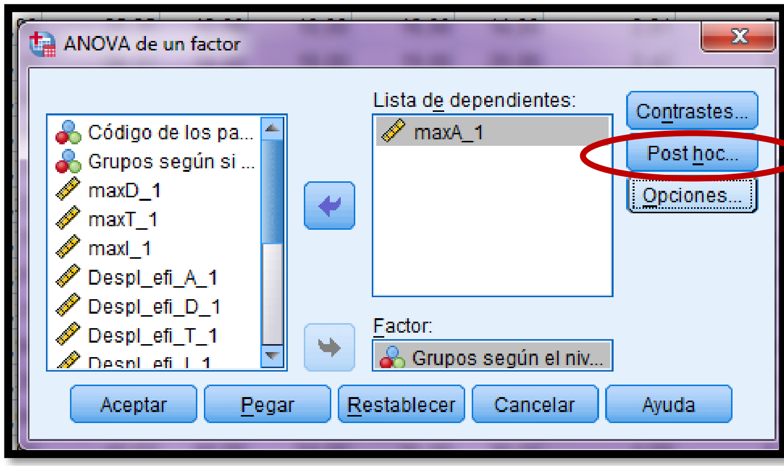
- Han de controlar l’error de tipus I quan es facen múltiples comparacions.

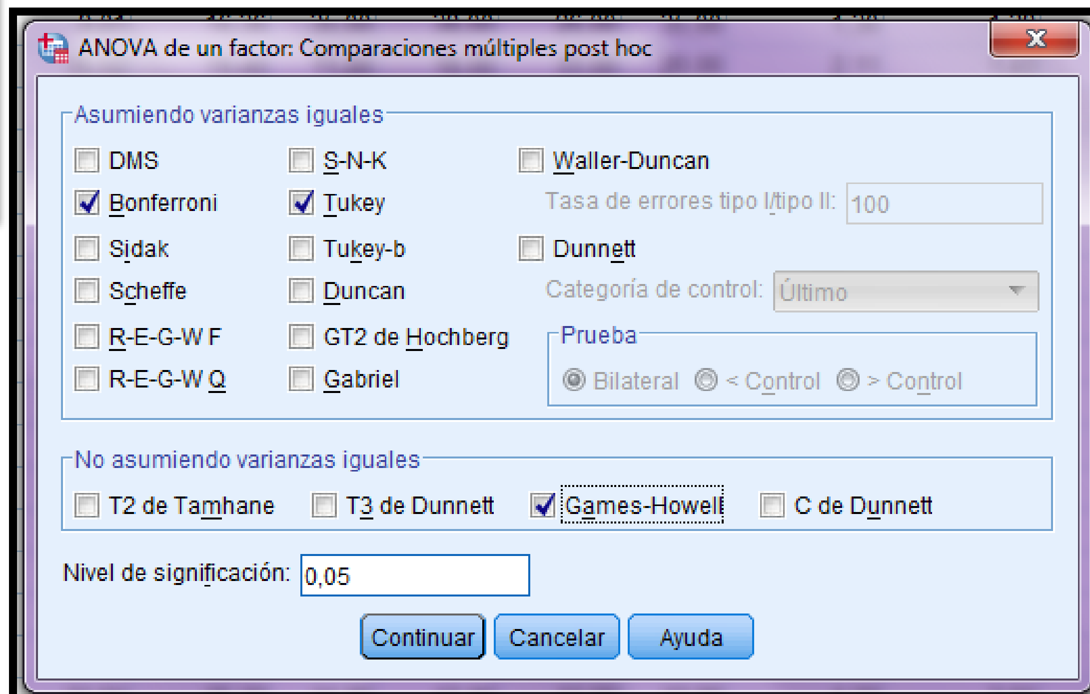
- L’estadístic diferència menys significativa (DMS) no controla l’error de tipus I.

- **Bonferroni** i **Tukey** són les més usades:
  - Bonferroni: més potent quan el nombre de comparacions és baix.
  - Tukey: més potent quan el nombre de comparacions és alt. Millor quan la mida de les mostres dels diversos grups és similar i hi ha homogeneïtat de variàncies entre grups.
- Gabriel : quan les mides mostrals són distintes en els diversos grups.
- **Games-Howell**: quan les variàncies no són iguals i les mides mostrals són distintes.

Introducció

- Proves *post hoc*





Execució de la prova

- Estudi de la capacitat de desplaçament màxim anterior del centre de pressions durant la sedestació en persones amb paraplegia (lesió alta i lesió baixa) i persones sense paraplegia.

Resultats

- Descriptius

Descriptivos

Desplazamiento anterior máximo del cdp

	N	Media	Desviación típica	Error típico	Intervalo de confianza para la media al 95% Límite inferior	Límite superior	Mínimo	Máximo
Sin lesión	15	125,0300	17,73761	4,57983	115,2072	134,8528	103,76	156,49
Lesión alta	13	33,4323	19,24287	5,33701	21,8040	45,0607	12,81	68,45
Lesión baja	11	50,4240	15,17751	4,57619	40,2277	60,6204	12,99	68,68
Total	39	73,4547	45,22159	7,24125	58,7956	88,1139	12,81	156,49

Resultats

Prueba de homogeneidad de varianzas

Desplazamiento anterior máximo del cdp

Estadístico de Levene	gl1	gl2	Sig.
1,361	2	36	,269

$P > 0,05 \rightarrow H_0$

Resultats

ANOVA

Desplazamiento anterior máximo del cdp

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	66557,961	2	33278,980	107,431	,000
Intra-grupos	11151,744	36	309,771		
Total	77709,705	38			

Model → Inter-grupos

Residus → Intra-grupos

$$F = \frac{MS_M}{MS_R}$$

Resultats

- En cas de no complir-se el supòsit d'homoscedasticitat...

Ajust dels graus de llibertat residuals

Pruebas robustas de igualdad de las medias

Desplazamiento anterior máximo del cdp

	Estadístico[a]	gl1	gl2	Sig.
Welch	101,308	2	23,487	,000
Brown-Forsythe	109,858	2	34,988	,000

a. Distribuidos en F asintóticamente.

$$SS_R = \sum (x_{ik} - \bar{x}_k)^2$$

$$SS_R = \sum s^2{}_k (n_k - 1)$$

Resultats

Comparaciones múltiples

Variable dependiente:Desplazamiento anterior máximo del cdp

	(I) Grupos según el nivel de la lesión medular	(J) Grupos según el nivel de la lesión medular	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95% Límite inferior	Límite superior
HSD de Tukey	Sin lesión	Lesión alta	91,59767*	6,66932	,000	75,2959	107,8995
		Lesión baja	74,60594*	6,98658	,000	57,5287	91,6832
	Lesión alta	Sin lesión	-91,59767*	6,66932	,000	-107,8995	-75,2959
		Lesión baja	-16,99173	7,21037	,061	-34,6160	,6326
	Lesión baja	Sin lesión	-74,60594*	6,98658	,000	-91,6832	-57,5287
		Lesión alta	16,99173	7,21037	,061	-,6326	34,6160
Bonferroni	Sin lesión	Lesión alta	91,59767*	6,66932	,000	74,8507	108,3446
		Lesión baja	74,60594*	6,98658	,000	57,0623	92,1495
	Lesión alta	Sin lesión	-91,59767*	6,66932	,000	-108,3446	-74,8507
		Lesión baja	-16,99173	7,21037	,072	-35,0973	1,1138
	Lesión baja	Sin lesión	-74,60594*	6,98658	,000	-92,1495	-57,0623
		Lesión alta	16,99173	7,21037	,072	-1,1138	35,0973
Games-Howell	Sin lesión	Lesión alta	91,59767*	7,03268	,000	74,0672	109,1281
		Lesión baja	74,60594*	6,47429	,000	58,4081	90,8038
	Lesión alta	Sin lesión	-91,59767*	7,03268	,000	-109,1281	-74,0672
		Lesión baja	-16,99173	7,03031	,061	-34,6571	,6737
	Lesión baja	Sin lesión	-74,60594*	6,47429	,000	-90,8038	-58,4081
		Lesión alta	16,99173	7,03031	,061	-,6737	34,6571

*. La diferencia de medias es significativa al nivel 0.05.

Introducció

- Com es calcula la mida de l'efecte?

  – El valor d'eta al quadrat (η^2):

$$\eta^2 = \frac{SS_M}{SS_T}$$

ANOVA

Desplazamiento anterior máximo del cdp

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	66557,961	2	33278,980	107,431	,000
Intra-grupos	11151,744	36	309,771		
Total	77709,705	38			

Escriptura dels resultats

- Hi ha un efecte estadísticament significatiu del fet de patir una lesió medul·lar toràcica en el desplaçament màxim anterior del centre de pressions (DAcdp). $F(2, 36) = 107{,}43$, $p < 0{,}05$, $\eta^2 = 0{,}09$.

Petit: $\eta^2 = 0{,}01$

Mitjà: $\eta^2 = 0{,}06$

Gran: $\eta^2 = 0{,}14$

- Les comparacions múltiples indiquen que hi ha diferències estadísticament significatives en el DAcdp entre els grups de persones sanes i els grups de persones amb paraplegia baixa $t(36) = 13{,}73$, $p<0{,}05$; de persones sanes i persones amb paraplegia alta $t(36) = 10{,}67$, $p><0{,}05$; i que no hi ha diferències en aquesta variable entre els dos grups de paraplègics $p>0{,}05$.

http://statistics.blogs.uv.es

VNIVERSITAT DE VALÈNCIA

ANÀLISI DE VARIÀNCIA (ANOVA) D'UN FACTOR

Prof. Pilar Serra

VNIVERSITAT DE VALÈNCIA

ANOVA d’un factor

- Eina estadística que analitza la variància d’un factor per a una variable dependent quantitativa.

- Objectiu: contrastar la hipòtesi que diverses mitjanes (més de dues) són iguals (H_0).

- Aquesta tècnica és una extensió de la prova t.
  - Corregeix l’error que es deriva de les comparacions de les mitjanes de *k* grups o *k* repeticions a fi d’evitar augmentar la probabilitat d’error de tipus I.

ANOVA d'un factor

- Tipus
  - ANOVA d'un factor ENTRE SUBJECTES
  - ANOVA d'un factor DE MESURES REPETIDES

ANOVA D'UN FACTOR DE MESURES REPETIDES

Introducció
Execució de la prova
Resultats
Escriptura dels resultats

VNIVERSITAT DE VALÈNCIA

Introducció

- Prova que serveix per a comparar les mesures que s'obtenen de diverses valoracions (més de dues) d'un únic grup.

- La variable independent constitueix un factor intrasubjecte.

Introducció

- Tipus d'estudi d'investigació
  - Temps: quan es vol avaluar l'efectivitat d'un tractament.
    - Es pot mesurar abans, durant i després de la intervenció.
  - Condició: quan es vol mesurar la mateixa prova amb estímuls o contextos diferents.
    - Equilibri bipodal estàtic, dinàmic, amb goma escuma, ulls tancats, estimulació optocinètica, etc.
  - Extremitats: quan es vol comprovar una variable amb les quatre extremitats del cos.
    - En cas d'una lesió tendinosa unilateral, podem voler comprovar el costat no lesionat.

Introducció

- Forma de presentar les dades
  - Cada fila correspon a un subjecte.
  - Cada columna correspon a cada condició o cada moment del temps.

Subjectes	Sense pes	1 kg	2 kg	5 kg
1	180	178	130	81
2	171	170	99	65
3	180	180	105	72
4	179	180	102	75
5	177	173	110	83
6	179	179	170	98

Introducció

La ràtio *F* compara la mida de la variància de la nostra manipulació amb l'atzar.

SS_T

$$SS_T = \sum (x_i - \bar{x}_{total})^2$$

- ### *SS* entre grups

  Només si hi ha més d'un grup

- ### *SS* intrasubjecte

  $$SS_W = \sum S^2{}_n (n_n - 1)$$

  - #### SS_M

    Efecte de l'experiment

    $$SS_M = \sum n_k (\bar{x}_k - x_{total})^2$$

  - #### SS_R

    Error

    $$SS_R = SS_W - SS_M$$

En mesures repetides queda dins de la variància intrasubjectes

Introducció

- **Suma de quadrats total**
  - Variància de totes les puntuacions respecte a la mitjana total (ignorant la condició experimental).

Subjectes	Sense pes	1 kg	2 kg	5 kg
1	180	178	130	81
2	171	170	99	65
3	180	180	105	72
4	179	180	102	75
5	177	173	110	83
6	179	179	170	98

$$SS_T = \sum (x_i - \bar{x}_{total})^2$$

$$SS_T = S^2{}_{total}(n-1)$$

X total	138,17
S total	44,65
S^2 total	1.994,06

$$S^2 = \frac{SS_T}{gl_T}$$

Introducció

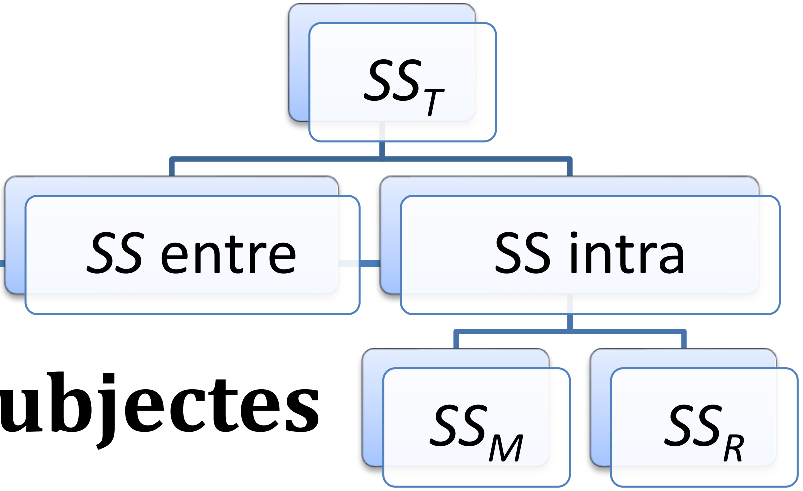
- **Graus de llibertat totals**

$gl_T = N\text{-}1$

N= 6 subjectes × 4 condicions = 24

gl_T = 24-1 = 23

Introducció



- **Suma de quadrats intrasubjectes**
  - És la dispersió dins de la mateixa persona.
  - Com es calcula
    - Adaptació de l'equació dels residus (ANOVA intergrups).

S^2= variància entre cadascuna de les puntuacions del subjecte en cadascuna de les condicions respecte a la mitjana de totes les condicions.

$$SS_R = \sum S^2{}_k (n_k - 1)$$

$$SS_W = \sum S^2{}_n (n_n - 1)$$

n_n = nombre de condicions
En el nostre exemple= 4 (pesos diferents)

$$SS_W = s^2{}_{\text{persona 1}}(n_1 - 1) + s^2{}_{\text{persona 2}}(n_1 - 1) \ldots + s^2{}_{\text{persona n}}(n_n - 1)$$

Introducció

- **Graus de llibertat intrasubjecte**

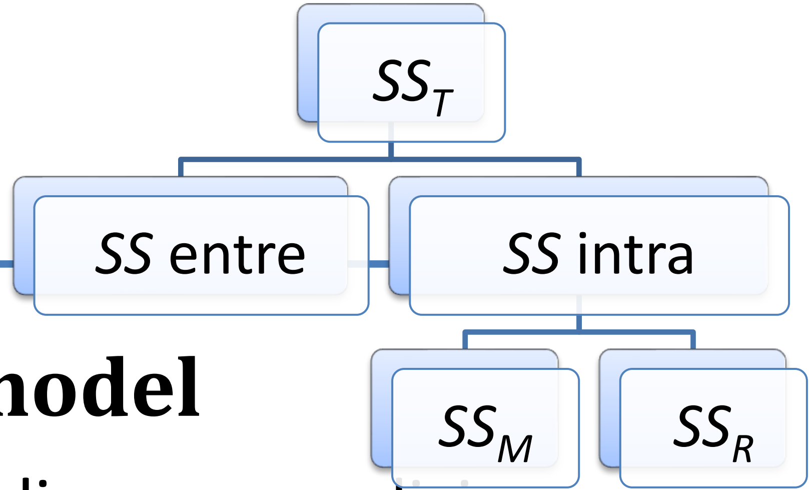
  gl de cada persona = nombre de condicions -1

  Ex. 4 - 1 = 3

  Multiplicació pel nombre de persones (*N*)

  Exemple: (4 - 1) × 6 = 18

Introducció



- **Suma de quadrats del model**
  - Variància que s'obté de les diverses condicions experimentals.
  - Com es calcula
    - S'obté la diferència de la mitjana de cadascuna de les condicions respecte a la mitjana total.
    - La diferència s'eleva al quadrat.
    - Cada resultat de la diferència es multiplica pel nombre de subjectes que hi ha en cadascuna de les condicions.
    - Se sumen totes les diferències.

$$SS_M = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$$

Introducció

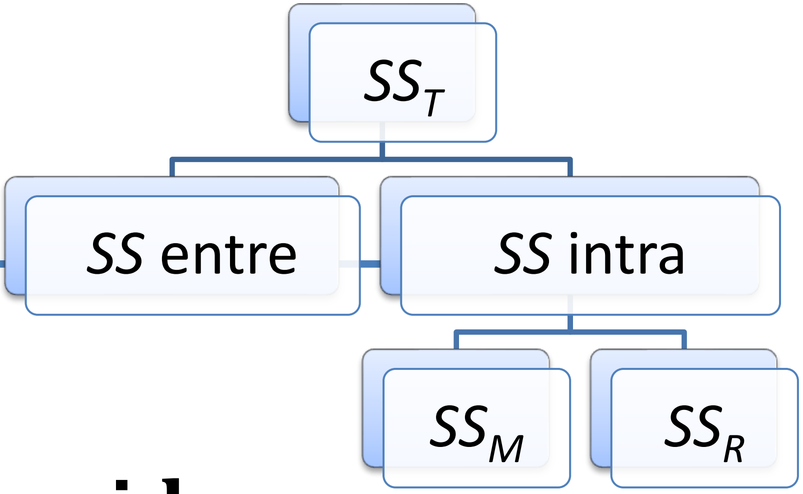
- **Graus de llibertat del model**

$gl_M = k - 1$

K = nombre de condicions

Exemple: 4 - 1 = 3

Introducció



- **Suma de quadrats dels residus**
  - Variació que no s'explica pel model, sinó per les característiques individuals. Com es calcula:

$$SS_R = SS_W - SS_M$$

Introducció

- **Graus de llibertat dels residus**

$$gl_R = gl_W - gl_M$$

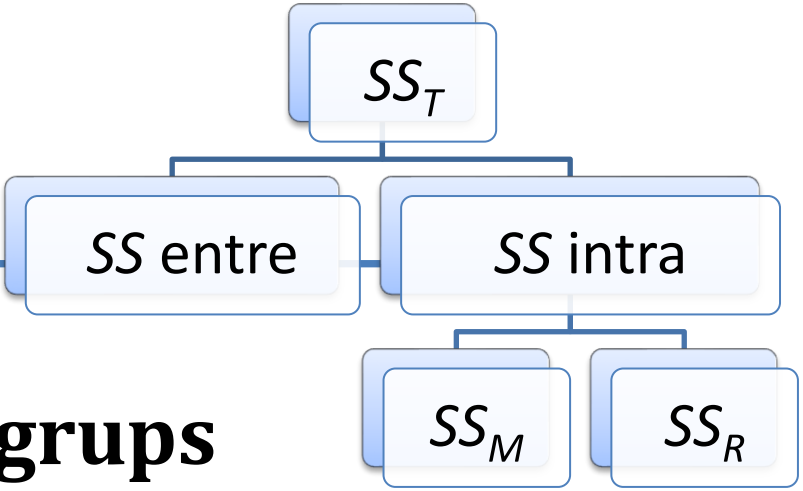
Exemple

$gl_W = (k - 1) \times n = (4 - 1) \times 6 = 18$

$gl_M = \text{k} - 1 = 4 - 1 = 3$

$gl_R = gl_W - gl_M = 15$

Introducció



- **Suma de quadrats entre grups**

  Dispersió deguda a les diferències entre grups de persones.

$$SS_T = SS_{BG} + SS_M + SS_R$$

$$SS_{BG} = SS_T - SS_M - SS_R$$

Introducció

- Ràtio F

$$F = \frac{MS_M}{MS_R}$$

$$MS_M = \frac{SS_M}{gl_M}$$

$$MS_R = \frac{SS_R}{gl_R}$$

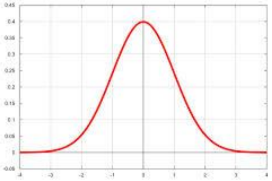
Introducció

- **Proves *post hoc***
  - Bonferroni: realitza les proves t de mostres aparellades i divideix l'error entre el nombre de contrastos (0,05/nre. de contrastos).
  - Sidak: menys conservadora.
  - LSD: no ajusta l'error de tipus I.

Introducció

- Les hipòtesis que es formulen són:
  - H_0 (hipòtesi nul·la o d'igualtat): les mitjanes de les diverses repeticions són iguals.
  - H_1 (hipòtesi alternativa o de diferència): les mesures de les diverses repeticions són significativament distintes.

Introducció



- Compliment de supòsits
  - NORMALITAT
    - Es comprova amb Kolmogorov-Smirnov o Shapiro-Wilk.
  - ESFERICITAT
    - Las variàncies de les diferències entre parells de mesures han de ser similars.
    - Es comprova amb la prova d'esfericitat de Mauchly.
      - Si es compleix, l'estadístic *F* es calcula amb una aproximació univariada.

Introducció

- Què ocorre si no es compleix...?
  - El supòsit de normalitat: L'ANOVA és robusta
    - Alternatives
      - Eliminar valors extrems.
      - Transformar dades.
      - Usar la prova Friedman amb les proves *post hoc* de Wilcoxon i ajust d'errors (taxa d'error de 0,05/nombre total de contrastos).

Introducció

- Què ocorre si no es compleix...?
  - El supòsit d'esfericitat
    - L'estadístic *F* no pot calcular-se amb l'aproximació univariada.
    - Alternatives
      - Aproximació univariada amb un ajust dels graus de llibertat:
        - » Límit inferior
        - » Greenhouse-Geisser $\hat{\varepsilon}$
        - » Huynh-Feldt $\tilde{\varepsilon}$

Multiplica els graus de llibertat (del model i dels residus) per l'esfericitat estimada.

      - Aproximació multivariada

Introducció

– Alternatives

- Aproximació univariada
  - Límit inferior $1/K-1$ → K és el nombre de condicions o moments.
  - Greenhouse-Geisser: pren valors entre el límit inferior i 1.
  - Huynh-Feldt: menys conservador.

- Aproximació multivariada
  - En lloc d'analitzar les diverses condicions o moments de mesura com parelles d'observacions relacionades, ho fa com si foren variables dependents diferents.

Se selecciona la més potent→ en SPSS es demana la potència observada.

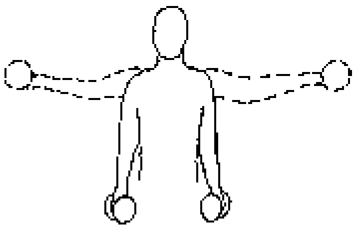
Quina d'aquestes opcions cal triar?

Quan Greenhouse-Geisser és > 0,75 no detecta si hi ha diferències E. S.

Execució de la prova

- Valoració de l'amplitud articular d'abducció de muscle quan s'alcen quatre pesos distints:

Subjectes	Sense pes	1 kg	2 kg	5 kg
1	180	178	130	81
2	171	170	99	65
3	180	180	105	72
4	179	180	102	75
5	177	173	110	83
6	179	179	170	98



Dispersió pel model

Hi ha diferències per la **condició** experimental:

< amplitud a > pes

Dispersió individual /error

Hi ha diferències **individuals**:

Al subjecte 6, el pes l'influeix molt poc, i al 2 molt.

Resultats

Estadísticos descriptivos

	Media	Desviación típica	N
Sin peso	177,6667	3,44480	6
1 kg	176,6667	4,17931	6
2 kg	119,3333	27,15634	6
5 kg	79,0000	11,33137	6

Resultats

Prueba de esfericidad de Mauchly[b]

Medida:Rango_articular_abducción

Efecto intra-sujetos	W de Mauchly	Chi-cuadrado aprox.	gl	Sig.	Epsilon[a]		
					Greenhouse-Geisser	Huynh-Feldt	Límite-inferior
Peso	,002	23,502	5	,000	,359	,379	,333

Contrasta la hipótesis nula de que la matriz de covarianza error de las variables dependientes transformadas es proporcional a una matriz identidad.

a. Puede usarse para corregir los grados de libertad en las pruebas de significación promediadas. Las pruebas corregidas se muestran en la tabla Pruebas de los efectos inter-sujetos.
b. Diseño: Intersección
Diseño intra-sujetos: Peso

P<0,05→ $H1$→ diferència estadísticament significativa en les variàncies de les parelles de mesures. Incompliment del supòsit d'esfericitat.

Resultats

Epsilon[a]		
Greenhouse-Geisser	Huynh-Feldt	Límite-inferior
,359	,379	,333

Pruebas de efectos intra-sujetos.

Medida:Rango_articular_abducción

Origen		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta al cuadrado parcial	Parámetro de no centralidad Parámetro	Potencia observada[a]
Peso	Esfericidad asumida	41387,333	3	13795,778	88,334	,000	,946	265,001	1,000
	Greenhouse-Geisser	41387,333	1,077	38425,001	88,334	,000	,946	95,144	1,000
	Huynh-Feldt	41387,333	1,138	36382,369	88,334	,000	,946	100,486	1,000
	Límite-inferior	41387,333	1,000	41387,333	88,334	,000	,946	88,334	1,000
Error(Peso)	Esfericidad asumida	2342,667	15	156,178					
	Greenhouse-Geisser	2342,667	5,385	434,998					
	Huynh-Feldt	2342,667	5,688	411,874					
	Límite-inferior	2342,667	5,000	468,533					

a. Calculado con alfa = ,05

- Graus de llibertat del model

  $k - 1 = 3$

- Graus de llibertat de l'error

  $(k - 1) \times (n - 1) = 3 \times 5 = 15$

Resultats

- Aproximació multivariada

Contrastes multivariados[c]

Efecto		Valor	F	Gl de la hipótesis	Gl del error	Sig.	Eta al cuadrado parcial	Parámetro de no centralidad Parámetro	Potencia observada[b]
Peso	Traza de Pillai	,998	598,127[a]	3,000	3,000	,000	,998	1794,380	1,000
	Lambda de Wilks	,002	598,127[a]	3,000	3,000	,000	,998	1794,380	1,000
	Traza de Hotelling	598,127	598,127[a]	3,000	3,000	,000	,998	1794,380	1,000
	Raíz mayor de Roy	598,127	598,127[a]	3,000	3,000	,000	,998	1794,380	1,000

a. Estadístico exacto
b. Calculado con alfa = ,05
c. Diseño: Intersección
Diseño intra-sujetos: Peso

Resultats

- No sempre la potència és la mateixa per a tots els ajustaments

Pruebas de efectos intra-sujetos.

Medida:Rango_articular_abducción

Origen		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta al cuadrado parcial	Parámetro de no centralidad Parámetro	Potencia observada[a]
Peso	Esfericidad asumida	2756,125	3	918,708	14,768	,000	,747	44,305	,999
	Greenhouse-Geisser	2756,125	1,094	2519,251	14,768	,010	,747	16,157	,890
	Huynh-Feldt	2756,125	1,169	2358,675	14,768	,008	,747	17,257	,908
	Límite-inferior	2756,125	1,000	2756,125	14,768	,012	,747	14,768	,863
Error(Peso)	Esfericidad asumida	933,125	15	62,208					
	Greenhouse-Geisser	933,125	5,470	170,586					
	Huynh-Feldt	933,125	5,843	159,713					
	Límite-inferior	933,125	5,000	186,625					

a. Calculado con alfa = ,05

VNIVERSITAT DE VALÈNCIA

Resultats

Epsilon[a]		
Greenhouse-Geisser	Huynh-Feldt	Límite-inferior
,359	,379	,333

$$\eta^2_p = \frac{SS_M}{SS_M + SS_R}$$

Pruebas de efectos intra-sujetos.

Medida:Rango_articular_abducción

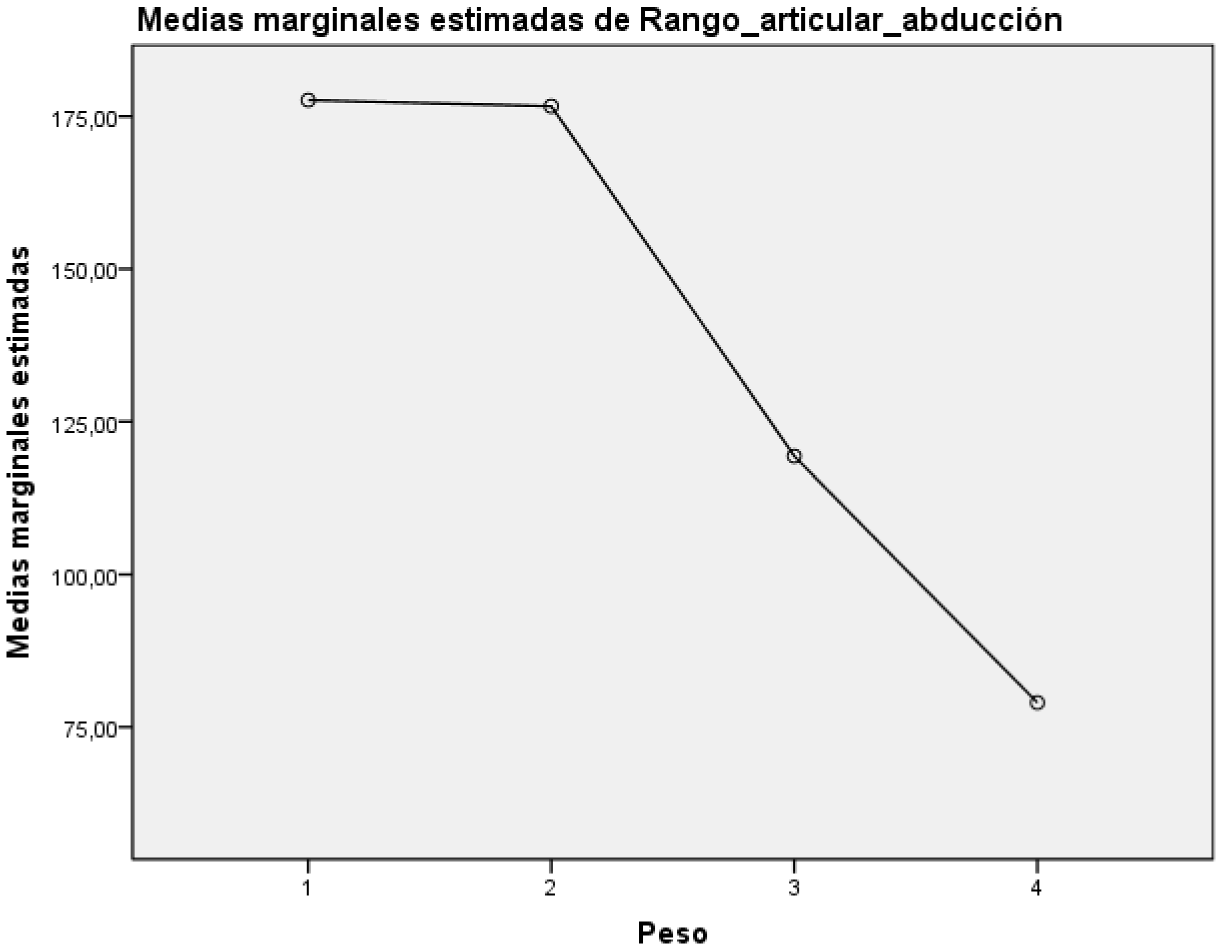
Origen		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta al cuadrado parcial	Parámetro de no centralidad Parámetro	Potencia observada[a]
Peso	Esfericidad asumida	41387,333	3	13795,778	88,334	,000	,946	265,001	1,000
	Greenhouse-Geisser	41387,333	1,077	38425,001	88,334	,000	,946	95,144	1,000
	Huynh-Feldt	41387,333	1,138	36382,369	88,334	,000	,946	100,486	1,000
	Límite-inferior	41387,333	1,000	41387,333	88,334	,000	,946	88,334	1,000
Error(Peso)	Esfericidad asumida	2342,667	15	156,178					
	Greenhouse-Geisser	2342,667	5,385	434,998					
	Huynh-Feldt	2342,667	5,688	411,874					
	Límite-inferior	2342,667	5,000	468,533					

a. Calculado con alfa = ,05

- Graus de llibertat del model:
  $k - 1 = 3$
- Graus de llibertat de l'error.
  $(k - 1) \times (n - 1) = 3 \times 5 = 15$

El 94,6% és per l'efecte de les condicions de valoració

Resultats



Medias marginales estimadas de Rango_articular_abducción
Medias marginales estimadas
175,00
150,00
125,00
100,00
75,00
1
2
3
4
Peso

Resultats

Comparaciones por pares

Medida:Rango_articular_abducción

(I)Peso	(J)Peso	Diferencia de medias (I-J)	Error típ.	Sig.[a]	Intervalo de confianza al 95 % para la diferencia[a]	
					Límite inferior	Límite superior
1	2	1,000	,730	1,000	-2,081	4,081
	3	58,333*	10,632	,016	13,473	103,194
	4	98,667*	4,096	,000	81,384	115,949
2	1	-1,000	,730	1,000	-4,081	2,081
	3	57,333*	10,617	,018	12,539	102,128
	4	97,667*	4,287	,000	79,579	115,755
3	1	-58,333*	10,632	,016	-103,194	-13,473
	2	-57,333*	10,617	,018	-102,128	-12,539
	4	40,333*	7,135	,014	10,228	70,439
4	1	-98,667*	4,096	,000	-115,949	-81,384
	2	-97,667*	4,287	,000	-115,755	-79,579
	3	-40,333*	7,135	,014	-70,439	-10,228

Basadas en las medias marginales estimadas.

a. Ajuste para comparaciones múltiples: Bonferroni.
*. La diferencia de medias es significativa al nivel ,05.

Escriptura dels resultats

- L'amplitud articular d'abducció de muscle és afectada pel pes elevat, $F(1{,}08,\ 5{,}40) = 88{,}33$, $p<0{,}05$, $\eta^2_p=0{,}946$.

- Abans d'especificar el resultat de l'ANOVA, es pot reportar l'incompliment del supòsit d'esfericitat:
  - El test d'esfericitat de Mauchly indica que el supòsit d'esfericitat no es compleix per a l'efecte de la condició en l'amplitud articular d'abducció de muscle ($\chi2(5)=23{,}50$, $p<0{,}05$); per tant, els graus de llibertat s'han corregit amb l'estimació d'esfericitat de Greenhouse-Geisser ($\varepsilon = 0{,}36$).

Escriptura dels resultats

- Pel que fa a les comparacions per parelles, no hi ha diferències estadísticament significatives en l'amplitud articular d'abducció de muscle sense pes i amb un quilogram ($t(5)=1{,}37$, $p>0{,}05$). No obstant això, sí que n'hi ha entre l'abducció sense pes i amb dos quilograms ($t(5)= 5{,}49$, $p<0{,}05$), entre l'abducció sense pes i amb cinc quilograms ($t(5)= 24{,}09$, $p<0{,}05$), entre un quilogram i dos quilograms ($t(5)= 5{,}40$, $p<0{,}05$), entre un i cinc ($t(5)=22{,}78$, $p<0{,}05$) i entre dos i cinc ($t(5)= 5{,}65$, $p<0{,}05$).

Escriptura dels resultats

- Altres opcions
  - Especificació de les diferències i indicació del nivell de significació (i. e. $p < 0,05$).
  - El nivell de significació es pot indicar al principi de la manera següent: "Tots els efectes es reporten amb una significació $p < 0,05$". A continuació, especificació de les diferències.

http://statistics.blogs.uv.es

VNIVERSITAT DE VALÈNCIA

ANÀLISI DE VARIÀNCIA (ANOVA) FACTORIAL

Dra. Pilar Serra Añó

VNIVERSITAT DE VALÈNCIA

ANOVA factorial

- Eina estadística que analitza la variància de dos o més factors per a una variable dependent quantitativa.

- Objectiu: contrastar la hipòtesi que diverses mitjanes són iguals (H_0).

ANOVA factorial

- Tipus
  - ANOVA factorial ENTRE SUBJECTES
  - ANOVA factorial DE MESURES REPETIDES

ANOVA factorial ENTRE SUBJECTES

- Exemple: valoració del dolor (mesurat amb EVA) durant l’abducció de muscle (dependent quantitativa) segons:
  - Factor 1: patologia (tres grups) → Factor entre subjectes
  - Factor 2: sexe (dos grups) → Factor entre subjectes

Grups	Sexe	Dolor
Muscle congelat	Dones	
	Homes	
Ruptura parcial del supraespinós	Dones	
	Homes	
Contractura del trapezi superior	Dones	
	Homes	

ANOVA factorial MESURES REPETIDES

- Exemple: valoració del dolor (mesurat amb EVA) durant l'abducció de muscle (dependent quantitativa) segons:
  - Factor 1: pes (quatre nivells) → Factor intrasubjecte
  - Factor 2: tècnica (dos grups)→ Factor intrasubjecte

Patologia	Dolor							
	0 kg		1 kg		2 kg		5 kg	
	D	No D	D	No D	D	No D	D	No D
Tendinitis del supraespinós								

D = descoaptació; No D = sense descoaptació

ANOVA disseny mixt

- Exemple: valoració del dolor (mesurat amb EVA) durant l'abducció de muscle (dependent quantitativa) segons:
  - Factor 1: patologia (tres grups) → Factor entre subjecte
  - Factor 2: sexe (dos grups) → Factor entre subjecte
  - Factor 3: pes (quatre nivells) → Factor entre subjecte

Patologia	Sexe	Dolor			
		0 kg	1 kg	2 kg	5 kg
Muscle congelat	Dones				
	Homes				
Ruptura parcial del supraespinós	Dones				
	Homes				
Contractura del trapezi superior	Dones				
	Homes				

VNIVERSITAT DE VALÈNCIA

ANOVA disseny mixt

Patologia	**Sexe**	**Dolor**							
		0 kg		1 kg		2 kg		5 kg	
		D	No D	D	No D	D	No D	D	No D
Muscle congelat	Dones								
	Homes								
Ruptura parcial del supraespinós	Dones								
	Homes								
Contractura trapezi superior	Dones								
	Homes								

D = descoaptació; No D= sense descoaptació

ANOVA FACTORIAL ENTRE SUBJECTES

Introducció
Execució de la prova
Resultats
Escriptura dels resultats

VNIVERSITAT DE VALÈNCIA

Introducció

- Prova que serveix per a comparar les mitjanes que s'obtenen de la valoració de més de dos grups.

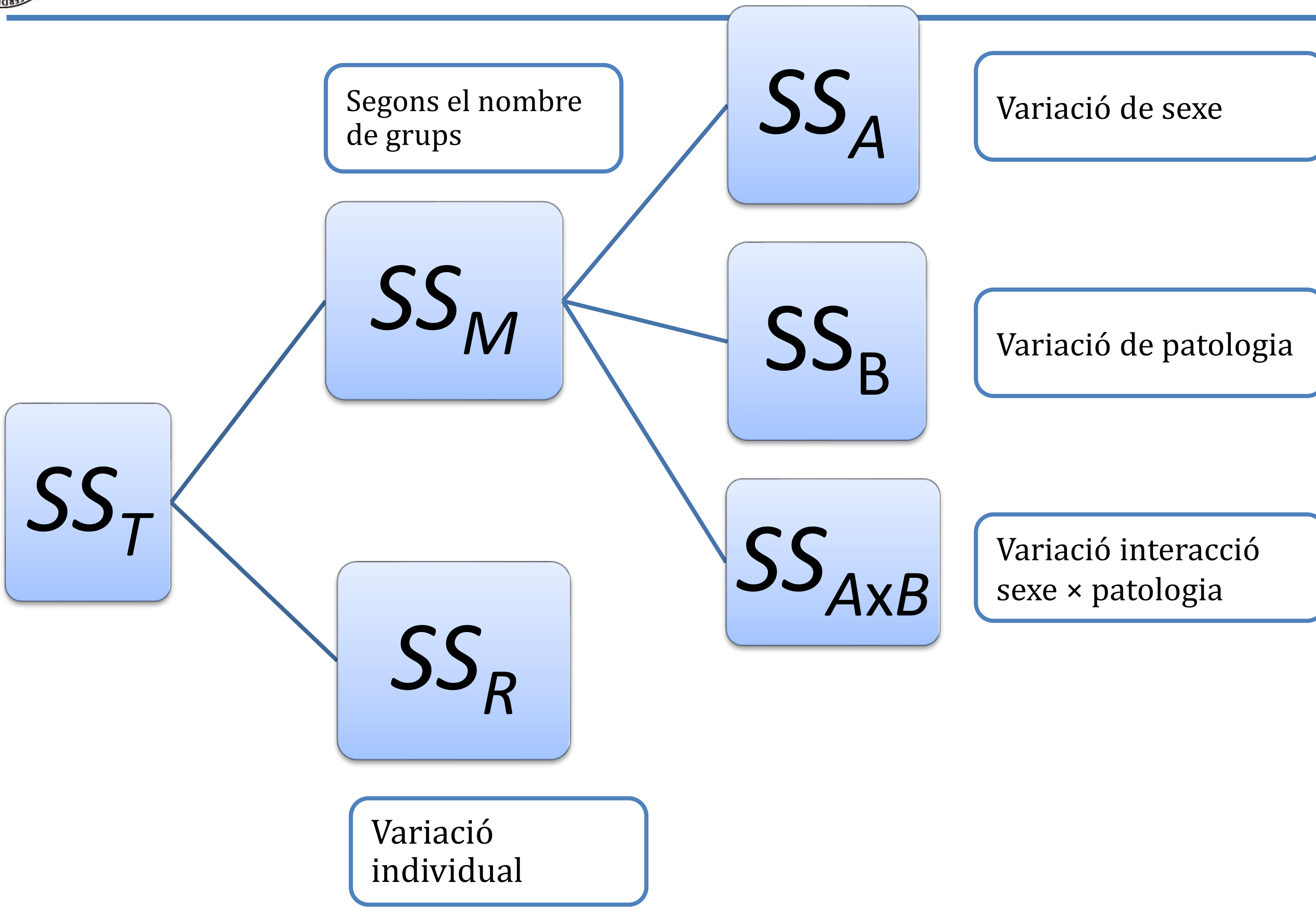
- Dos o més factors entre subjectes: les variables independents sobre les quals es fonamenta l'estudi.

Introducció

- Exemple: valoració de l'amplitud articular (mesurada amb goniòmetre electrònic) durant l'abducció de muscle (dependent quantitativa) segons:
  - Factor 1: patologia (3 grups)
  - Factor 2: sexe (2 grups)

Grups	Sexe	Amplitud articular
Muscle congelat	Dones	
	Homes	
Ruptura parcial del supraespinós	Dones	
	Homes	
Contractura del trapezi superior	Dones	
	Homes	

Introducció



Introducció

- SS_T

  $SS_T = \sum (x_i - \bar{x}_{total})^2$

  - SS_M

    $SS_M = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$

    - SS_A

      $SS_A = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$

    - SS_B

      $SS_B = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$

    - SS_{AXB}

      $SS_{AxB} = SS_M - SS_A - SS_B$

  - SS_R

    $SS_R = \sum (x_{ik} - \bar{x}_k)^2$

Introducció

Significat de la interacció

- Existeix una interacció significativa entre dos factors quan l'efecte d'un d'ells sobre la variable dependent no és el mateix en tots els nivells de l'altre factor.

Introducció

El comportament d'amplitud articular en les diverses patologies no és el mateix en homes i dones.

Patologia	Sexe	Amplitud articular
Muscle congelat	Dones	
	Homes	
Ruptura parcial del suprespinós	Dones	
	Homes	
Contractura de trapezi superior	Dones	
	Homes	

Introducció

- **Suma de quadrats total**
  - Es calcula la diferència entre les dades observades i la mitjana total (la mitjana de tots els grups que s'hagen de comparar).

$$SS_T = \sum \left(x_i - \bar{x}_{total} \right)^2$$

- **Graus de llibertat totals**: nombre d'observacions que són lliures de variar.

  $gl_T = N$-1

Introducció

- ***SS$_M$***: dispersió de les dades que s'explica pel model (en aquest cas, el nostre model té sis grups).

$$SS_M = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$$

- ***gl$_M$***

  K – 1 → K = nombre de grups

  Exemple: 6 – 1 = 5

VNIVERSITAT DE VALÈNCIA

Introducció

- ***SS$_A$***: dispersió de les dades segons el sexe (sense tenir en compte la condició patològica)

$$SS_A = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$$

$$SS_{sexe} = n_{\text{dones}}(\bar{x} - \bar{x}_{total})^2 + n_{\text{homes}}(\bar{x} - \bar{x}_{total})^2$$

- ***gl$_A$***

  K – 1 → K = nombre de grups

  Exemple: 2 – 1 = 1

Introducció

- $\boldsymbol{SS_B}$: dispersió de les dades segons la condició patològica (sense tenir en compte el sexe).

$$SS_B = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$$

$$SS_{\text{patologia}} = n_{\text{muscle cong.}}(\bar{x} - \bar{x}_{total})^2 + n_{\text{ruptura}}(\bar{x} - \bar{x}_{total})^2 + n_{\text{contractura}}(\bar{x} - \bar{x}_{total})^2$$

- $\boldsymbol{gl_B}$

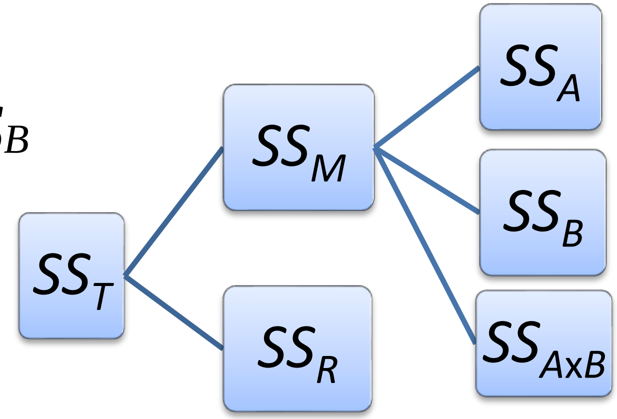
  K – 1 → K = nombre de grups

  Exemple: 3 – 1 = 2

Introducció

- $\boldsymbol{SS_{AxB}}$: dispersió de les dades tenint en compte la condició patològica i el sexe.

$$SS_{AxB} = SS_M - SS_A - SS_B$$



- $\boldsymbol{gl_{AxB}}$

$$gl_M - gl_A - gl_B$$

Exemple: 5 – 1 – 2 = 2

Introducció

- $\boldsymbol{SS_R}$: dispersió individual de les dades

$$SS_R = \sum (x_{ik} - \bar{x}_k)^2$$

Estadísticos descriptivos

Variable dependiente:Rango articular de abducción sin peso

Patología	Género	Media	Desviación típica	N
Contractura del trapecio superior	Mujeres	176,80	2,775	5
	Hombres	177,67	3,445	6
	Total	177,27	3,036	11
Tendinitis del supraespinoso	Mujeres	170,00	5,500	9
	Hombres	178,17	1,941	6
	Total	173,27	5,982	15
Rotura parcial del supraespinoso	Mujeres	136,00	4,000	6
	Hombres	126,13	4,155	8
	Total	130,36	6,416	14
Total	Mujeres	161,50	17,881	20
	Hombres	157,20	26,233	20
	Total	159,35	22,266	40

- $\boldsymbol{gl_R}$

  Nombre de subjectes menys un de cada grup

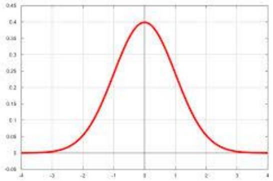
  6 grups = 34

Introducció

- Les hipòtesis que es formulen són:
  - H_0 (hipòtesi nul·la o d'igualtat): les mitjanes de les mostres són iguals.
  - H_1 (hipòtesi alternativa o de diferència): les mitjanes de les mostres són significativament distintes.

Introducció

- Compliment dels supòsits



  - NORMALITAT
    - La distribució de les dades ha de ser similar a la distribució normal.
    - Es comprova amb Kolmogorov-Smirnov o Shapiro-Wilk.
  - HOMOGENEÏTAT DE VARIÀNCIES
    - La variància d'una mostra ha de ser similar a la variància d'una altra mostra.
    - Es comprova amb el test de Levene.
  - INDEPENDÈNCIA
    - Puntuacions distintes perquè venen de persones distintes.

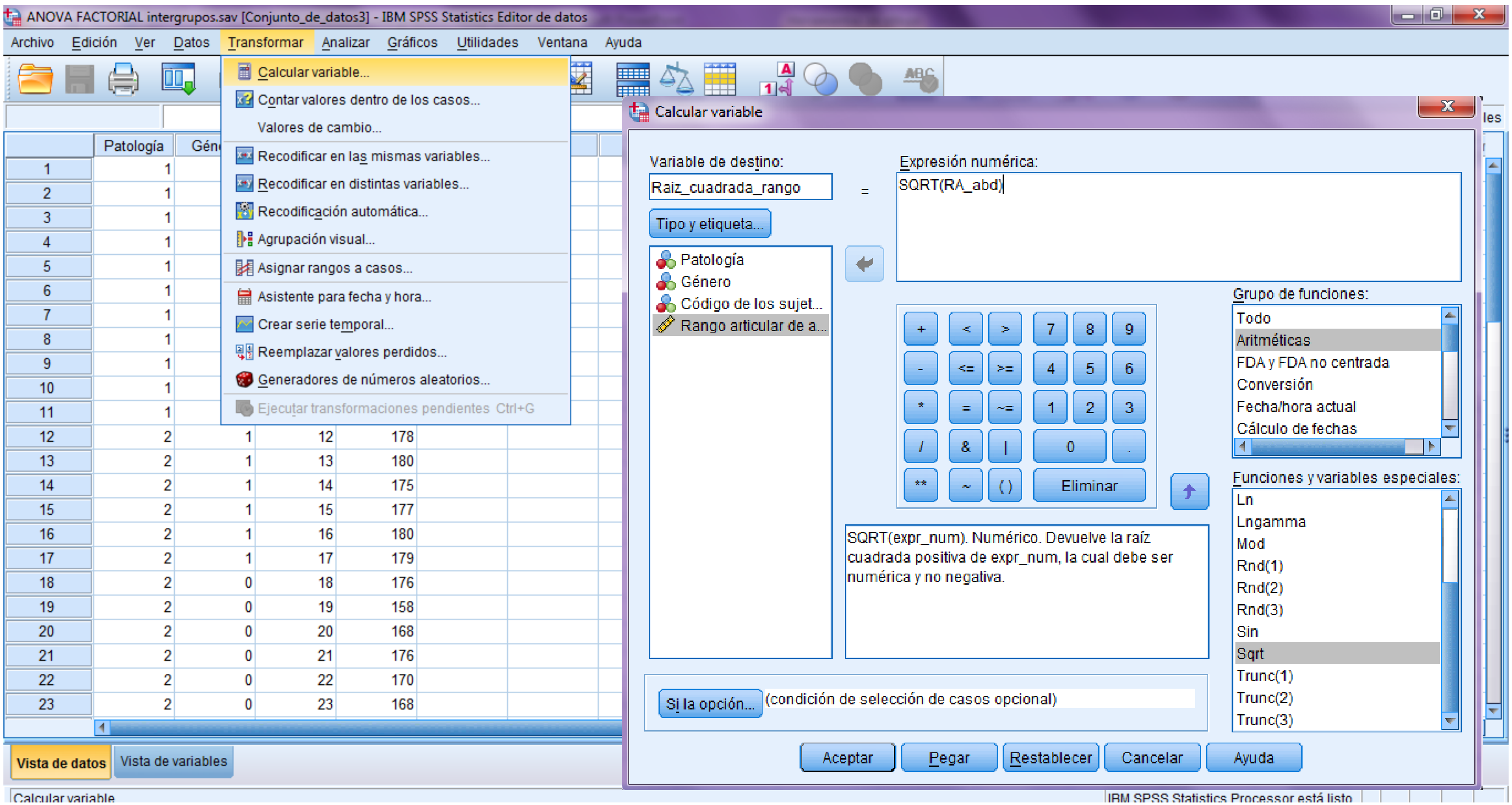
Introducció

- Què ocorre si no es compleix…?
  - El supòsit de normalitat: L'ANOVA és robusta.
    - Alternatives
      - Eliminar valors extrems
      - Transformar dades

Introducció

- Què ocorre si no es compleix…?
  - El supòsit d'homoscedasticitat
    - Alternativa: transformar dades

Introducció



Introducció

*ANOVA FACTORIAL intergrupos.sav [Conjunto_de_datos3] - IBM SPSS Statistics Editor de datos

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda

	Patología	Género	Sujetos	RA_abd	Raiz_cuadrada_rango	var	var	var	var	var	var	var
1	1	1	1	180	13,42							
2	1	1	2	171	13,08							
3	1	1	3	180	13,42							
4	1	1	4	179	13,38							
5	1	1	5	177	13,30							
6	1	1	6	179	13,38							
7	1	0	7	178	13,34							
8	1	0	8	173	13,15							
9	1	0	9	175	13,23							
10	1	0	10	180	13,42							
11	1	0	11	178	13,34							
12	2	1	12	178	13,34							
13	2	1	13	180	13,42							
14	2	1	14	175	13,23							
15	2	1	15	177	13,30							
16	2	1	16	180	13,42							
17	2	1	17	179	13,38							
18	2	0	18	176	13,27							
19	2	0	19	158	12,57							
20	2	0	20	168	12,96							
21	2	0	21	176	13,27							
22	2	0	22	170	13,04							
23	2	0	23	168	12,96							

Vista de datos Vista de variables

IBM SPSS Statistics P

Universitat de València

Introducció

- Tres ràtios *F*

Les mitjanes quadràtiques són el resultat de la divisió de la suma de quadrats corresponent entre els seus graus de llibertat

$$F_A = \frac{MS_A}{MS_R}$$

Efecte principal del tipus d'intervenció

$$F_B = \frac{MS_B}{MS_R}$$

Efecte principal del pes

$$F_{AxB} = \frac{MS_{AxB}}{MS_R}$$

Efecte principal de la interacció

Execució de la prova

- Estudi de l'amplitud articular d'abducció de muscle en homes i dones amb distintes lesions.

Resultats

- Descriptius

Estadísticos descriptivos

Variable dependiente:Rango articular de abducción sin peso

Patología	Género	Media	Desviación típica	N
Contractura del trapecio superior	Mujeres	176,80	2,775	5
	Hombres	177,67	3,445	6
	Total	177,27	3,036	11
Tendinitis del supraespinoso	Mujeres	170,00	5,500	9
	Hombres	178,17	1,941	6
	Total	173,27	5,982	15
Rotura parcial del supraespinoso	Mujeres	136,00	4,000	6
	Hombres	126,13	4,155	8
	Total	130,36	6,416	14
Total	Mujeres	161,50	17,881	20
	Hombres	157,20	26,233	20
	Total	159,35	22,266	40

Resultats

Contraste de Levene sobre la igualdad de las varianzas error[a]

Variable dependiente:Rango articular de abducción sin peso

F	gl1	gl2	Sig.
,768	5	34	,579

Contrasta la hipótesis nula de que la varianza error de la variable dependiente es igual a lo largo de todos los grupos.

a. Diseño: Intersección + Patología + Género + Patología * Género

$P > 0{,}05 \rightarrow H_0$

Resultats

$$F = \frac{MS_M}{MS_R}$$

Pruebas de los efectos inter-sujetos

Variable dependiente:Rango articular de abducción sin peso

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta al cuadrado parcial	Parámetro de no centralidad Parámetro	Potencia observada[b]
Modelo corregido	18783,258[a]	5	3756,652	231,454	,000	,971	1157,272	1,000
Intersección	994282,229	1	994282,229	61259,593	,000	,999	61259,593	1,000
Patología	17579,080	2	8789,540	541,540	,000	,970	1083,080	1,000
Género	,757	1	,757	,047	,830	,001	,047	,055
Patología * Género	576,039	2	288,019	17,745	,000	,511	35,491	1,000
Error	551,842	34	16,231					
Total	1035032,000	40						
Total corregida	19335,100	39						

a. R cuadrado = ,971 (R cuadrado corregida = ,967)
b. Calculado con alfa = ,05

$$F_A = \frac{MS_A}{MS_R} \qquad F_B = \frac{MS_B}{MS_R} \qquad F_{AxB} = \frac{MS_{AxB}}{MS_R}$$

Resultats

Comparaciones por pares

Variable dependiente:Rango articular de abducción sin peso

Género	(I)Patología	(J)Patología	Diferencia de medias (I-J)	Error típ.	Sig.[a]	Intervalo de confianza al 95 % para la diferencia[a] Límite inferior	Límite superior
Mujeres	Contractura del trapecio superior	Tendinitis del supraespinoso	6,800*	2,247	,014	1,141	12,459
		Rotura parcial del supraespinoso	40,800*	2,440	,000	34,657	46,943
	Tendinitis del supraespinoso	Contractura del trapecio superior	-6,800*	2,247	,014	-12,459	-1,141
		Rotura parcial del supraespinoso	34,000*	2,123	,000	28,653	39,347
	Rotura parcial del supraespinoso	Contractura del trapecio superior	-40,800*	2,440	,000	-46,943	-34,657
		Tendinitis del supraespinoso	-34,000*	2,123	,000	-39,347	-28,653
Hombres	Contractura del trapecio superior	Tendinitis del supraespinoso	-,500	2,326	1,000	-6,357	5,357
		Rotura parcial del supraespinoso	51,542*	2,176	,000	46,063	57,021
	Tendinitis del supraespinoso	Contractura del trapecio superior	,500	2,326	1,000	-5,357	6,357
		Rotura parcial del supraespinoso	52,042*	2,176	,000	46,563	57,521
	Rotura parcial del supraespinoso	Contractura del trapecio superior	-51,542*	2,176	,000	-57,021	-46,063
		Tendinitis del supraespinoso	-52,042*	2,176	,000	-57,521	-46,563

Basadas en las medias marginales estimadas.

*. La diferencia de medias es significativa al nivel ,05.

a. Ajuste para comparaciones múltiples: Bonferroni.

VNIVERSITAT DE VALÈNCIA

Resultats

Comparaciones por pares

Variable dependiente:Rango articular de abducción sin peso

Patología	(I)Género	(J)Género	Diferencia de medias (I-J)	Error típ.	Sig.[a]	Intervalo de confianza al 95 % para la diferencia[a]	
						Límite inferior	Límite superior
Contractura del trapecio superior	Mujeres	Hombres	-,867	2,440	,725	-5,824	4,091
	Hombres	Mujeres	,867	2,440	,725	-4,091	5,824
Tendinitis del supraespinoso	Mujeres	Hombres	-8,167*	2,123	,001	-12,482	-3,852
	Hombres	Mujeres	8,167*	2,123	,001	3,852	12,482
Rotura parcial del supraespinoso	Mujeres	Hombres	9,875*	2,176	,000	5,453	14,297
	Hombres	Mujeres	-9,875*	2,176	,000	-14,297	-5,453

Basadas en las medias marginales estimadas.

a. Ajuste para comparaciones múltiples: Bonferroni.
*. La diferencia de medias es significativa al nivel ,05.

Escriptura dels resultats

- Esquema
  - Existeix/No existeix una interacció significativa entre els factors patologia x sexe.
  - Existeix/No existeix un efecte principal significatiu de la patologia.
  - Existeix/No existeix un efecte principal significatiu del sexe.
  - Comparacions per parelles.

Escriptura dels resultats

- Existeix una interacció significativa entre els factors patologia i sexe $F(2, 34)= 17{,}75$, $p<0{,}05$, $\eta^2 =0{,}51$ en l'amplitud articular durant l'abducció de muscle.
- Existeix un efecte principal significatiu de la patologia $F(2, 34)=541{,}54$, $p<0{,}05$, $\eta^2=0{,}97$ en l'amplitud articular durant l'abducció de muscle.
- No existeix un efecte principal significatiu del sexe $F(1, 34)=0{,}05$, $p>0{,}05$, $\eta^2 = 0{,}01$ en l'amplitud articular durant l'abducció de muscle.

Pruebas de los efectos inter-sujetos

Variable dependiente:Rango articular de abducción sin peso

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta al cuadrado parcial	Parámetro de no centralidad Parámetro	Potencia observada[b]
Modelo corregido	18783,258[a]	5	3756,652	231,454	,000	,971	1157,272	1,000
Intersección	994282,229	1	994282,229	61259,593	,000	,999	61259,593	1,000
Patología	17579,080	2	8789,540	541,540	,000	,970	1083,080	1,000
Género	,757	1	,757	,047	,830	,001	,047	,055
Patología * Género	576,039	2	288,019	17,745	,000	,511	35,491	1,000
Error	551,842	34	16,231					
Total	1035032,000	40						
Total corregida	19335,100	39						

Escriptura dels resultats

- Les comparacions múltiples indiquen que hi ha diferències estadísticament significatives entre *x* i *y* ($p < 0,05$).

  – Exemple: en les dones hi ha diferències estadísticament significatives en l'amplitud articular de muscle ($p<0,05$), segons si pateixen una contractura de trapezi superior ($M=176,80$; $DT=2,78$) o una tendinitis del supraespinós ($M=170,00$; $DT= 5,50$).

http://statistics.blogs.uv.es

VNIVERSITAT DE VALÈNCIA

Introducció

- Valoració del tipus de tractament (GE i GC) en l'amplitud articular

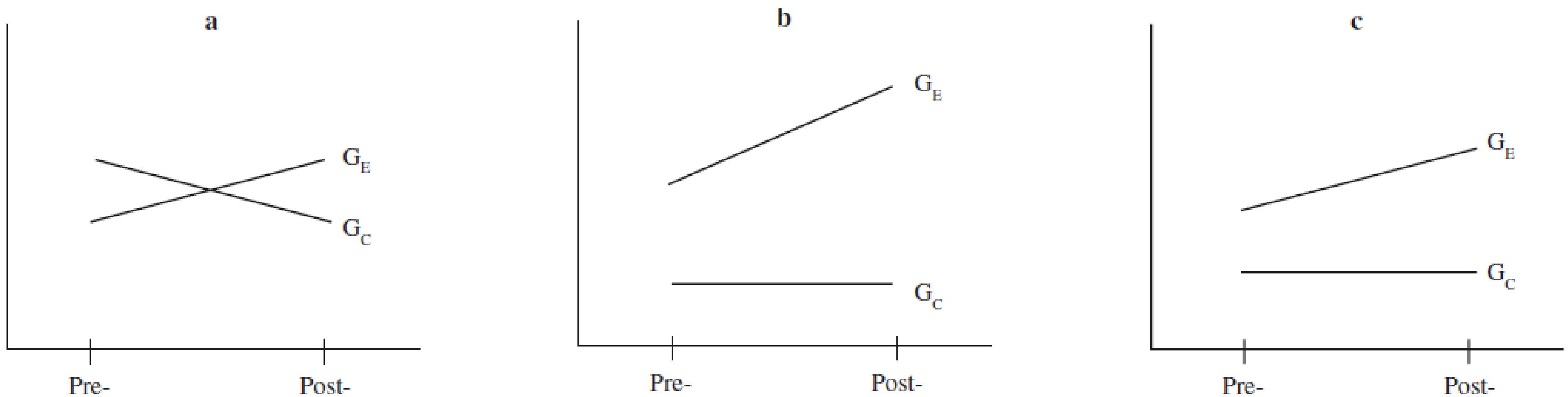


Figura 1. Distintas pautas de interacción en un diseño 2×2

Pardo, A., Garrido, J., Ángel Ruiz, M. i San Martín, R. (2007). "La interacción entre factores en el análisis de varianza: errores de interpretación". *Psicothema*, *19*(2).

ANOVA factorial
MESURES REPETIDES

Introducció
Execució de la prova
Resultats
Escriptura dels resultats

VNIVERSITAT DE VALÈNCIA

Introducció

- Prova que serveix per a comparar mitjanes que s'obtenen de diverses valoracions (més de dues) d'un únic grup.
- Factors intrasubjectes: les variables independents que constitueixen l'estudi.

Introducció

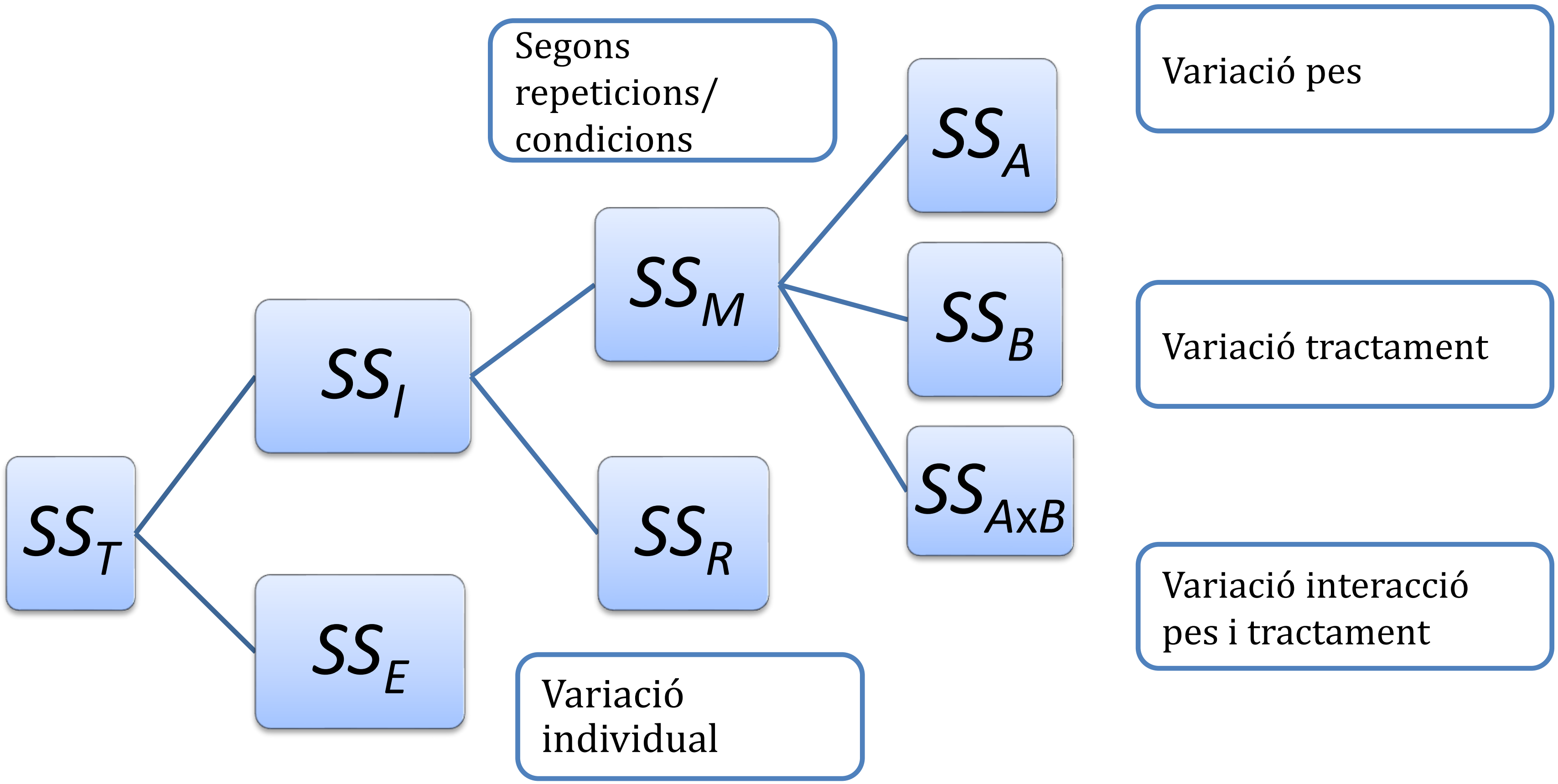
- Exemple: anàlisi de l'efecte del pes en el dolor segons els hagen aplicat una intervenció d'entrenament de força o de flexibilitat (amb un període de depuració).

ANOVA factorial MESURES REPETIDES

- Exemple: valoració del dolor (mesurat amb EVA) durant l'abducció de muscle (dependent quantitativa) segons:
  - Factor 1: pes (quatre nivells) → Factor intrasubjecte
  - Factor 2: intervenció rebuda (dos grups)→ Factor intrasubjecte

Patologia	**Dolor**					
	0 kg		2 kg		5 kg	
	Força	Flexib.	Força	Flexib.	Força	Flexib.
Tendinitis del supraespinós						

Introducció



VNIVERSITAT DE VALÈNCIA

Introducció

- SS_T
  $$SS_T = \sum (x_i - \bar{x}_{total})^2$$
  - *SS* entre grups
    Si hi ha més d'un grup
  - *SS* intrasubjecte
    $$SS_W = \sum S^2{}_n (n_n - 1)$$
    - SS_M
      Efecte de l'experiment
      $$SS_M = \sum n_k (\bar{x}_k - x_{total})^2$$
      - SS_A
        $$SS_A = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$$
      - SS_B
        $$SS_B = \sum n_k (\bar{x}_k - \bar{x}_{total})^2$$
      - SS_{AxB}
        $$SS_{AxB} = SS_M - SS_A - SS_B$$
    - SS_R
      Error
      $$SS_R = SS_W - SS_M$$

Introducció

- Tres ràtios *F*

Les mitjanes quadràtiques són el resultat de la divisió de la suma de quadrats corresponent entre els seus graus de llibertat.

$$F_A = \frac{MS_A}{MS_R}$$ → Efecte principal del pes

$$F_B = \frac{MS_B}{MS_R}$$ → Efecte principal del tipus d’intervenció

$$F_{AxB} = \frac{MS_{AxB}}{MS_R}$$ → Efecte principal de la interacció

Introducció

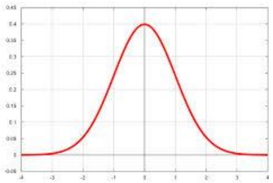
- Per a l'anàlisi de les mitjanes de les diverses condicions, s'**ajusten als efectes principals** mitjançant els ajustos *post hoc*:
  - Bonferroni: fa proves t de mostres emparellades i divideix l'error entre el nombre de contrastos (0,05/nre. de contrastos).
  - Sidak: menys conservador.
  - LSD: no ajusta l'error de tipus I.

Introducció

- Les hipòtesis que es formulen són:
  - H_0 (hipòtesi nul·la o d'igualtat): les mitjanes de les diverses repeticions són iguals.
  - H_1 (hipòtesi alternativa o de diferència): les mitjanes de les diverses repeticions són significativament distintes.

Introducció

- Compliment de supòsits
  - NORMALITAT
    - La distribució de les dades ha de ser similar a la distribució normal.
    - Es comprova amb Kolmogorov-Smirnov o Shapiro-Wilk.
  - ESFERICITAT
    - Les variàncies de les diferències entre parelles de mesures han de ser similars.
    - Es comprova la prova d’esfericitat de Mauchly.
      - Si es compleix, l’estadístic *F* es calcula amb una aproximació univariada.



Introducció

- Què ocorre si no es compleix?
  - El supòsit de normalitat: L’ANOVA és robusta.
    - Alternatives
      - Eliminar valors extrems
      - Transformar dades

Introducció

- Què ocorre si no es compleix?
  - El supòsit d'esfericitat
    - L'estadístic *F* no pot calcular-se amb l'aproximació univariada.
    - Alternatives
      - Aproximació univariada amb <u>ajustament</u> dels graus de llibertat:
        - » Límit inferior
        - » *Greenhouse-Geisser* $\hat{\varepsilon}$
        - » *Huynh-Feldt* $\tilde{\varepsilon}$
      - Aproximació multivariada

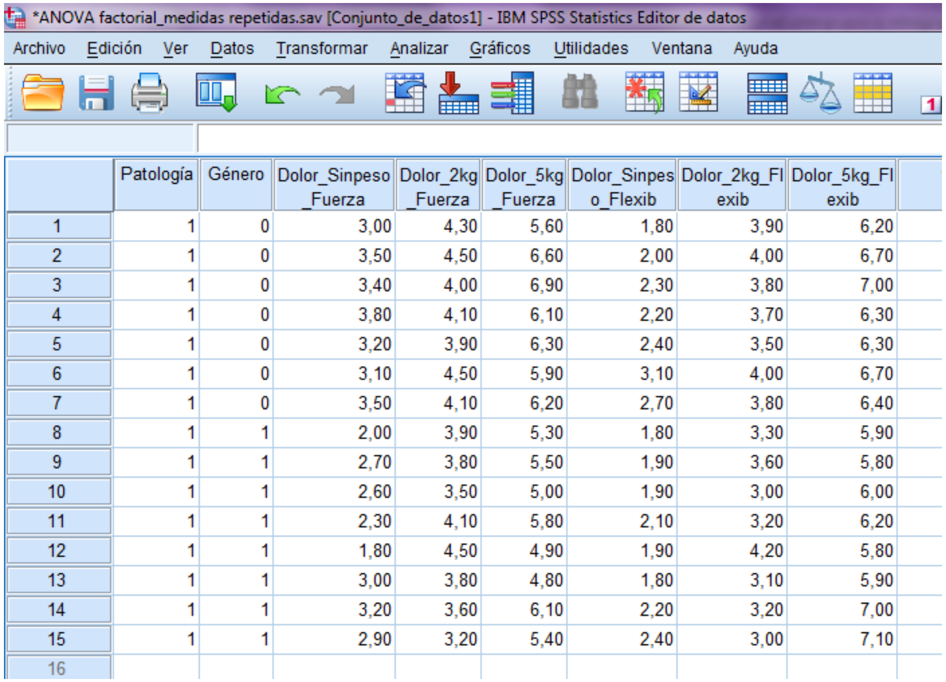
Multiplica els graus de llibertat (del model i dels residus) per l'esfericitat estimada.

Se selecciona la més potent → En SPSS es demana la potència observada.

Quan Greenhouse-Geisser és > 0,75, no detecta si hi ha diferències E. S.

Execució de la prova

- Valoració del dolor durant l'abducció de muscle en persones sanes quan alcen tres pesos distints (0,2 i 5 kg) després d'haver fet dos tipus d'intervenció (entrenament de força i entrenament de flexibilitat).



	Patología	Género	Dolor_Sinpeso_Fuerza	Dolor_2kg_Fuerza	Dolor_5kg_Fuerza	Dolor_Sinpeso_Flexib	Dolor_2kg_Flexib	Dolor_5kg_Flexib
1	1	0	3,00	4,30	5,60	1,80	3,90	6,20
2	1	0	3,50	4,50	6,60	2,00	4,00	6,70
3	1	0	3,40	4,00	6,90	2,30	3,80	7,00
4	1	0	3,80	4,10	6,10	2,20	3,70	6,30
5	1	0	3,20	3,90	6,30	2,40	3,50	6,30
6	1	0	3,10	4,50	5,90	3,10	4,00	6,70
7	1	0	3,50	4,10	6,20	2,70	3,80	6,40
8	1	1	2,00	3,90	5,30	1,80	3,30	5,90
9	1	1	2,70	3,80	5,50	1,90	3,60	5,80
10	1	1	2,60	3,50	5,00	1,90	3,00	6,00
11	1	1	2,30	4,10	5,80	2,10	3,20	6,20
12	1	1	1,80	4,50	4,90	1,90	4,20	5,80
13	1	1	3,00	3,80	4,80	1,80	3,10	5,90
14	1	1	3,20	3,60	6,10	2,20	3,20	7,00
15	1	1	2,90	3,20	5,40	2,40	3,00	7,10
16								

VNIVERSITAT DE VALÈNCIA

Resultats

Estadísticos descriptivos

	Media	Desviación típica	N
Dolor sin peso tras intervención de fuerza	2,9333	,56779	15
Dolor sin peso tras intervención de flexibilidad	2,1667	,37161	15
Dolor con 2kg tras intervención de fuerza	3,9867	,37960	15
Dolor con 2kg tras intervención de flexibilidad	3,5533	,39797	15
Dolor con 5kg tras intervención de fuerza	5,7600	,62198	15
Dolor con 5kg tras intervención de flexibilidad	6,3533	,45019	15

Resultats

Prueba de esfericidad de Mauchly[b]

Medida:Dolor

Efecto intra-sujetos	W de Mauchly	Chi-cuadrado aprox.	gl	Sig.	Epsilon[a]		
					Greenhouse-Geisser	Huynh-Feldt	Límite-inferior
Peso	,675	5,114	2	,078	,755	,826	,500
Tratamiento	1,000	,000	0	.	1,000	1,000	1,000
Peso * Tratamiento	,978	,293	2	,864	,978	1,000	,500

Contrasta la hipótesis nula de que la matriz de covarianza error de las variables dependientes transformadas es proporcional a una matriz identidad.

a. Puede usarse para corregir los grados de libertad en las pruebas de significación promediadas. Las pruebas corregidas se muestran en la tabla Pruebas de los efectos inter-sujetos.

b. Diseño: Intersección
 Diseño intra-sujetos: Peso + Tratamiento + Peso * Tratamiento

$$F = \frac{MS_M}{MS_R}$$

Pruebas de efectos intra-sujetos.

Medida:Dolor

Origen		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta al cuadrado parcial	Parámetro de no centralidad Parámetro	Potencia observada[a]
Peso	Esfericidad asumida	190,140	2	95,070	420,058	,000	,968	840,116	1,000
	Greenhouse-Geisser	190,140	1,509	125,989	420,058	,000	,968	633,939	1,000
	Huynh-Feldt	190,140	1,652	115,081	420,058	,000	,968	694,027	1,000
	Límite-inferior	190,140	1,000	190,140	420,058	,000	,968	420,058	1,000
Error(Peso)	Esfericidad asumida	6,337	28	,226					
	Greenhouse-Geisser	6,337	21,128	,300					
	Huynh-Feldt	6,337	23,131	,274					
	Límite-inferior	6,337	14,000	,453					
Tratamiento	Esfericidad asumida	,920	1	,920	8,094	,013	,366	8,094	,754
	Greenhouse-Geisser	,920	1,000	,920	8,094	,013	,366	8,094	,754
	Huynh-Feldt	,920	1,000	,920	8,094	,013	,366	8,094	,754
	Límite-inferior	,920	1,000	,920	8,094	,013	,366	8,094	,754
Error(Tratamiento)	Esfericidad asumida	1,592	14	,114					
	Greenhouse-Geisser	1,592	14,000	,114					
	Huynh-Feldt	1,592	14,000	,114					
	Límite-inferior	1,592	14,000	,114					
Peso * Tratamiento	Esfericidad asumida	7,537	2	3,768	48,705	,000	,777	97,410	1,000
	Greenhouse-Geisser	7,537	1,956	3,852	48,705	,000	,777	95,287	1,000
	Huynh-Feldt	7,537	2,000	3,768	48,705	,000	,777	97,410	1,000
	Límite-inferior	7,537	1,000	7,537	48,705	,000	,777	48,705	1,000
Error(Peso*Tratamiento)	Esfericidad asumida	2,166	28	,077					
	Greenhouse-Geisser	2,166	27,390	,079					
	Huynh-Feldt	2,166	28,000	,077					
	Límite-inferior	2,166	14,000	,155					

a. Calculado con alfa = ,05

Resultats

Comparaciones por pares

Medida:Dolor

Tratamiento	(I)Peso	(J)Peso	Diferencia de medias (I-J)	Error típ.	Sig.[a]	Intervalo de confianza al 95 % para la diferencia[a]	
						Límite inferior	Límite superior
1	1	2	-1,053*	,174	,000	-1,526	-,580
		3	-2,827*	,121	,000	-3,154	-2,499
	2	1	1,053*	,174	,000	,580	1,526
		3	-1,773*	,165	,000	-2,221	-1,325
	3	1	2,827*	,121	,000	2,499	3,154
		2	1,773*	,165	,000	1,325	2,221
2	1	2	-1,387*	,123	,000	-1,722	-1,051
		3	-4,187*	,099	,000	-4,457	-3,916
	2	1	1,387*	,123	,000	1,051	1,722
		3	-2,800*	,156	,000	-3,224	-2,376
	3	1	4,187*	,099	,000	3,916	4,457
		2	2,800*	,156	,000	2,376	3,224

Basadas en las medias marginales estimadas.

*. La diferencia de medias es significativa al nivel ,05.

a. Ajuste para comparaciones múltiples: Bonferroni.

Resultats

Comparaciones por pares

Medida:Dolor

Peso	(I)Tratamiento	(J)Tratamiento	Diferencia de medias (I-J)	Error típ.	Sig.[a]	Intervalo de confianza al 95 % para la diferencia[a]	
						Límite inferior	Límite superior
1	1	2	,767*	,135	,000	,477	1,056
	2	1	-,767*	,135	,000	-1,056	-,477
2	1	2	,433*	,050	,000	,325	,541
	2	1	-,433*	,050	,000	-,541	-,325
3	1	2	-,593*	,122	,000	-,856	-,331
	2	1	,593*	,122	,000	,331	,856

Basadas en las medias marginales estimadas.

*. La diferencia de medias es significativa al nivel ,05.
a. Ajuste para comparaciones múltiples: Bonferroni.

Escriptura dels resultats

- Hi ha una interacció significativa entre els factors tipus d'intervenció i pes alçat $F(2, 28) = 48{,}71$, $p < 0{,}05$, $\eta^2 = 0{,}77$ en el dolor experimentat durant l’abducció.

- Hi ha un efecte principal del pes $F(2, 28) = 420{,}06$, $p < 0{,}05$, $\eta^2 = 0{,}97$ en el dolor durant l’abducció.

- Hi ha un efecte principal del tipus de tractament $F(1, 14) = 8{,}09$, $p < 0{,}05$, $\eta^2 = 0{,}37$ en el dolor durant l’abducció.

Escriptura dels resultats

- Les comparacions múltiples indiquen que hi ha diferències estadísticament significatives entre *x* i *y* (p < 0,05).
  - Exemple: quan s'efectua l'abducció amb 2 kg, hi ha diferències estadísticament significatives en el dolor (p < 0,05) segons si han rebut un tractament de força (M = 3,99, DT = 0,38) o de flexibilitat (M = 3,56, DT = 0,40).

http://statistics.blogs.uv.es

VNIVERSITAT DE VALÈNCIA

Associació entre variables

Dra. Pilar Serra Añó

VNIVERSITAT DE VALÈNCIA

Associació entre variables

- Proves per a observar si hi ha relació estadísticament significativa entre variables:
  - Categòriques (nominals o ordinals)
  - Quantitatives

Relacions entre variables categòriques. Taules de contingència i prova de χ^2

Introducció
Execució de la prova
Resultats
Escriptura dels resultats

Universitat de València

Introducció

- **<u>Taula de contingència</u>**: taules de doble entrada, a l'entrada de les quals se situen les diverses categories de la variable.
- En les caselles de la taula hi ha la freqüència o nombre de casos.
- Freqüències
  - Observades: nombre real de casos en el nostre estudi.
  - Esperades: nombre de casos que hauria d'haver-hi a la casella si les categories foren independents ($H0$ / $p > 0{,}05$).

Introducció

			Grau de la lesió			TOTAL
			Lleu	Moderat	Greu	
Tipus d'esportista	Esportista ocasional	Recompte	15	8	6	29
		Freqüència esperada	11	8,3	9,7	29
	Esportista de competició	Recompte	9	10	15	34
		Freqüència esperada	13	9,7	11,3	34
	Total	Recompte	24	18	21	63
		Freqüència esperada	24	18	21	63

Introducció

- Per a estimar el grau d'associació entre variables: **khi quadrat** (o una raó de versemblança).

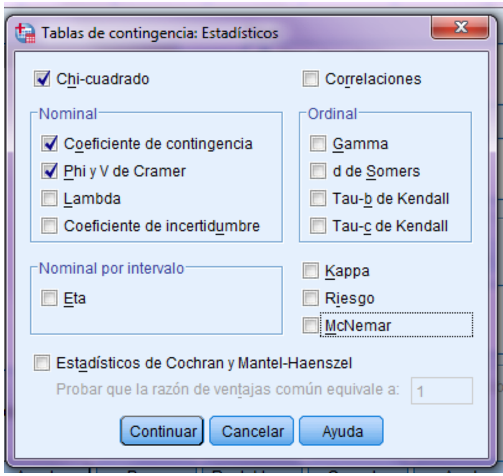
- El khi quadrat parteix de la hipòtesi que les variables són independents (i. e. no hi ha relació entre si i cap variable no exerceix influència sobre l'altra.
  - Si $p > 0,05$, s'accepta la hipòtesi nul·la (H_0)

Ho decideix el SPSS i comunica en els resultats si ocorre qualsevol d'aquests supòsits.

- El nivell de significació del χ^2 es calcula mitjançant tres mètodes:
  - Asimptòtic → Si la mostra té més de 20 subjectes; menys del 20% de les caselles tenen freqüències esperades < 5; no hi ha caselles amb freqüència observada < 1.
  - Montecarlo → No cal complir els supòsits anteriors. Mostres > 30 subjectes.
  - Exacte: no cal complir els supòsits anterior. No cal tenir en compte la mida de la mostra.

Introducció

- Com que χ^2 no indica la relació ni la magnitud de l'associació de les variables, cal usar estadístics segons el tipus de taula que presenta el nostre estudi:
  - Taules 2 x 2: coeficient phi $r(\varphi)$ i risc
  - Taules J x K: V de Cramer i coeficient de contingència.



  - En cas de ser variables relacionades (disseny anterior i posterior, s'usaria l'estadístic McNemar).

Introducció

- Coeficient phi $r(\varphi)$
- Valors de phi: entre 0 i $\sqrt{q} - 1$
  - En què q és el nombre de modalitats de la variable que en tinga menys.
  - En taules de 2 x 2, sempre va entre 0 i 1.

Introducció

- El risc (oportunitat relativa o *odds ratio*) també indica la mida de l'efecte encara que només es pot usar en taules de 2 x 2.
  - Exemple: si es compara el tipus d'esportista (de competició/ocasional) pel grau de lesió (greu/lleu) → Diríem que els esportistes de competició tenen 5,33 vegades més risc que la lesió siga greu que en el cas dels esportistes ocasionals.

Tabla de contingencia Tipo_deportista * Grado de lesión

			Grado de lesión		Total
			1	3	
Tipo_deportista	Deportista ocasional	Recuento	20	9	29
		Frecuencia esperada	13,8	15,2	29,0
	Deportista de competición	Recuento	10	24	34
		Frecuencia esperada	16,2	17,8	34,0
Total		Recuento	30	33	63
		Frecuencia esperada	30,0	33,0	63,0

Estimación de riesgo

	Valor	Intervalo de confianza al 95%	
		Inferior	Superior
Razón de las ventajas para Tipo_deportista (Deportista ocasional / Deportista de competición)	5,333	1,814	15,681
Para la cohorte Grado de lesión = 1	2,345	1,319	4,168
Para la cohorte Grado de lesión = 3	,440	,245	,789
N de casos válidos	63		

Introducció

- V de Cramer $0 < V < 1$
  - 0 = independència absoluta
  - 1 = dependència perfecta

V de Cramer = 0: cap relació
V de Cramer = 0,50: relació moderada
V de Cramer = 0,70: relació moderada-alta
V de Cramer = 1: relació perfecta

- Coeficient de contingència
  - 0 = independència absoluta
  - Valor màxim:

$$Max\ \{C\} = \sqrt{\frac{Min\{r-1, c-1\}}{1 + Min\{r-1, c-1\}}}$$

Segons el valor màxim que pot prendre, es calcula si l'associació és baixa, moderada o alta.

$$C = \sqrt{\chi^2 / (\chi^2 + n)}$$

Introducció

- En una mostra de persones majors, existeix alguna relació entre com se senten durant la marxa i els intents d'alçar-se?

- Variables
  - Marxa (insegur, amb por de caure / segur).
  - Intents (impossible sense ajuda / més d'un intent /
  amb un intent n'hi ha prou).

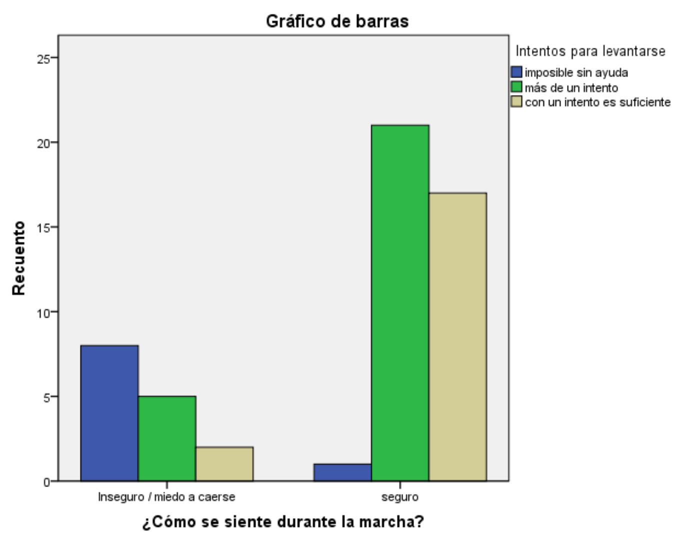
Introducció

- En la taula, cal fixar-se en:
  - Taula de contingència.
  - Proves de khi quadrat (valor de l'estadística i nivell de significació).
  - Mesures simètriques: per a observar la potència o força d'associació.
  - Risc (oportunitat relativa o *odds ratio*).

Resultats

Tabla de contingencia ¿Cómo se siente durante la marcha? * Intentos para levantarse

			Intentos para levantarse			Total
			imposible sin ayuda	más de un intento	con un intento es suficiente	
¿Cómo se siente durante la marcha?	Inseguro / miedo a caerse	Recuento	8	5	2	15
		Frecuencia esperada	2,5	7,2	5,3	15,0
	seguro	Recuento	1	21	17	39
		Frecuencia esperada	6,5	18,8	13,7	39,0
Total		Recuento	9	26	19	54
		Frecuencia esperada	9,0	26,0	19,0	54,0



Resultats

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	20,519[a]	2	,000
Razón de verosimilitudes	19,288	2	,000
Asociación lineal por lineal	14,416	1	,000
N de casos válidos	54		

a. 1 casillas (16,7%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 2,50.

Resultats

Medidas simétricas

		Valor	Sig. aproximada
Nominal por nominal	Phi	,616	,000
	V de Cramer	,616	,000
	Coeficiente de contingencia	,525	,000
N de casos válidos		54	

Escriptura dels resultats

- Existeix una associació estadísticament significativa entre com se sent durant la marxa i el nombre d'intents per a alçar-se χ^2 (2)= 20,519, $p < 0,05$.

- Si es vol reportar l'associació entre variables:
  - S'ha trobat una relació estadísticament significativa i alta (directament proporcional) (c. contingència = 0,525, p < 0,05)

  - X^2 = estadístic khi quadrat
  - () = graus de llibertat
  - *p* = nivell de significació

El valor màxim
del c. contingència és 0,71

http://statistics.blogs.uv.es/

VNIVERSITAT DE VALÈNCIA

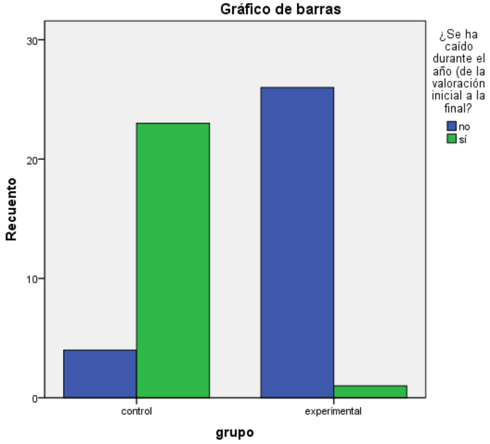
Execució de la prova (2)

- Hi ha alguna relació entre haver fet o no un programa d'exercicis terapèutics d'equilibri i enfortiment per a persones majors, i haver caigut en aquest període de temps (des de la valoració inicial fins a la final)?
- Variables
  - Grup (control/experimental)
  - Caigudes posteriors (sí/no)

Resultats

Tabla de contingencia grupo * ¿Se ha caído durante el año (de la valoración inicial a la final?

			¿Se ha caído durante el año (de la valoración inicial a la final?		Total
			no	sí	
grupo	control	Recuento	4	23	27
		% dentro de grupo	14,8%	85,2%	100,0%
	experimental	Recuento	26	1	27
		% dentro de grupo	96,3%	3,7%	100,0%
Total		Recuento	30	24	54
		% dentro de grupo	55,6%	44,4%	100,0%



Resultats

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	36,300[a]	1	,000		
Corrección por continuidad[b]	33,075	1	,000		
Razón de verosimilitudes	42,986	1	,000		
Estadístico exacto de Fisher				,000	,000
Asociación lineal por lineal	35,628	1	,000		
N de casos válidos	54				

a. 0 casillas (,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 12,00.
b. Calculado sólo para una tabla de 2x2.

Resultats

Medidas simétricas

		Valor	Sig. aproximada
Nominal por nominal	Phi	-,820	,000
	V de Cramer	,820	,000
N de casos válidos		54	

Escriptura dels resultats

- Els participants del grup experimental han caigut menys que els participants del grup de control (3,7% el grup experimental enfront del 85,2% en el grup de control). La relació és estadísticament significativa, ja que khi^2 (1) = 36,300, $p < 0{,}05$, Phi = 0,820).
  - Xi^2 = estadístic Xi quadrat
  - () = graus de llibertat
  - *P* = nivell de significació
  - V de Cramer = potència

Phi o V de Cramer = 0: cap relació
Phi o V de Cramer= 0,50: relació moderada
Phi o V de Cramer = 1: relación perfecta

Correlació

Dra. Pilar Serra Añó

VNIVERSITAT DE VALÈNCIA

Correlació

- Eina estadística per a comprovar si variables DEPENDENTS QUANTITATIVES estan associades.
- La hipòtesi nul·la (H_0) indica que no hi ha associació o relació estadísticament significativa entre les variables en qüestió (p>0,05).

Correlació

- Tipus
  - Correlació simple: s'estudia la dependència únicament entre dues variables.
  - Correlació múltiple: s'estudia la dependència entre més de dues variables.
  - Correlació parcial: quan s'inclou la influència de variables exògenes no considerades en el càlcul dels coeficients.

Correlació lineal simple. Coeficient de correlació de Pearson

Introducció
Execució de la prova
Resultats
Escriptura dels resultats

VNIVERSITAT DE VALÈNCIA

Introducció

- La **correlació** lineal simple tracta d'establir la relació o dependència que hi ha entre les dues variables que intervenen en una **distribució bidimensional**.

- La forma més simple d'analitzar-la és mirar si les variables tenen COVARIÀNCIA.

Introducció

- **Variància d'una única variable**
  – Dispersió de les dades respecte a la mitjana

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{N - 1}$$

Introducció

Covariància

- Per a comprovar si dues variables estan associades, cal veure si ambdues tenen covariància.
- Quan una variable es desvia respecte a la mitjana, l'altra variable es desvia de forma similar.

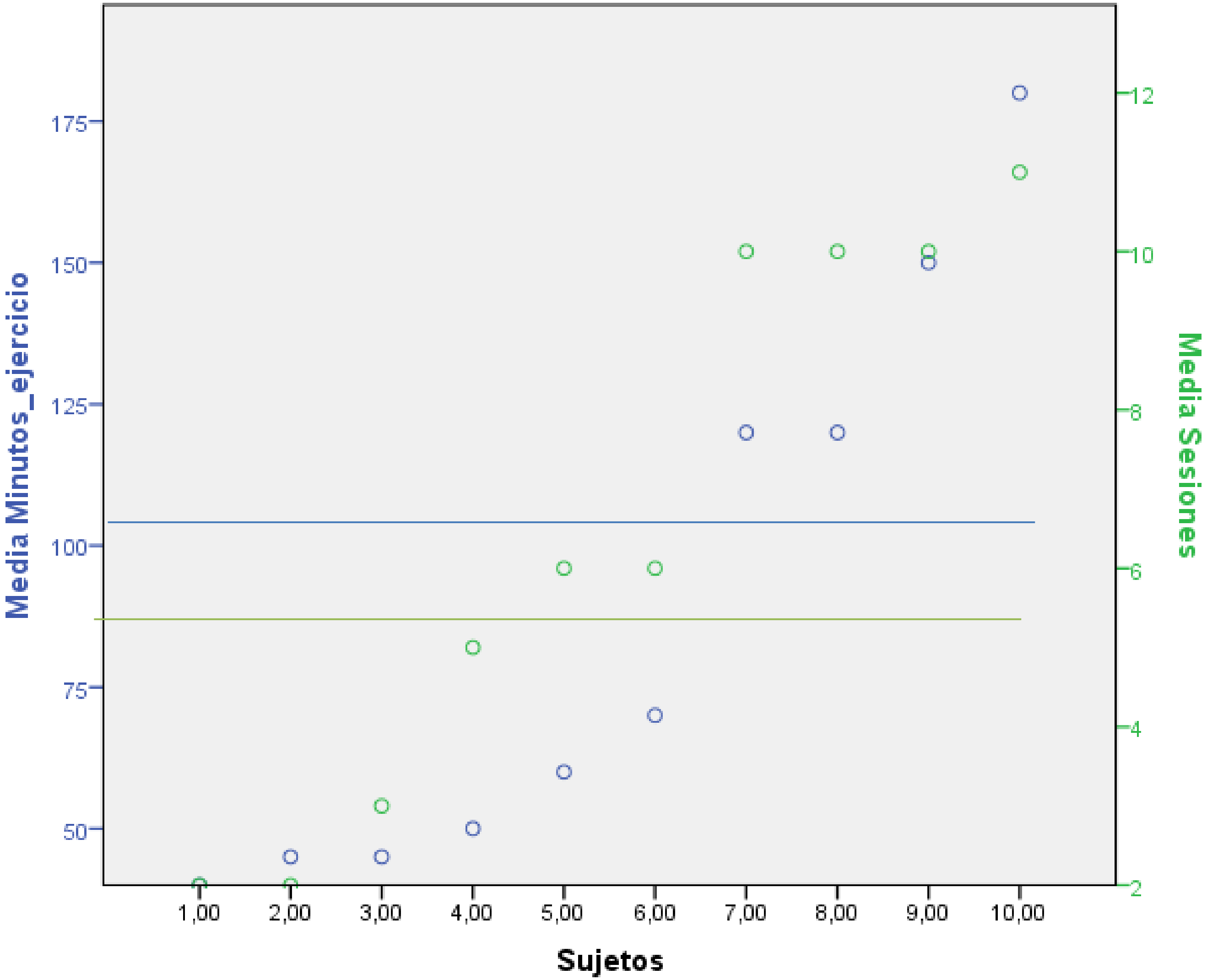
$$s^2 = \frac{\sum (x_i - \bar{x})^2}{N-1} \qquad s_{(x,y)} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N-1}$$

Introducció

Exemple: quantitat de minuts d'exercici físic segons la quantitat de sessions d'escola d'esquena rebudes.

Subjectes	1	2	3	4	5	6	7	8	9	10	$\overline{X}$	DT
Nombre sessions	2	2	3	6	10	10	10	5	6	11	6,5	3,54
Minuts/ setmana	40	45	45	60	120	150	120	50	70	180	88	50,45

Introducció



Introducció

Quin és el problema de la covariància?

- El valor depèn de l'escala escollida per als eixos
  - Varia si expressem l'alçada en metres o en centímetres.

Quina és la solució?

- Covariància estandarditzada = coeficient de correlació

Coeficient de correlació (lineal) de Pearson

- El coeficient de correlació de Pearson parteix de la covariància:

$$s^2_{(x,\,y)} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N - 1}$$

- És el quocient entre la **covariància** i el producte de les **desviacions típiques** d'ambdues variables.

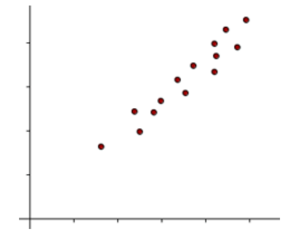
$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{(N - 1) s_x s_y}$$

Introducció

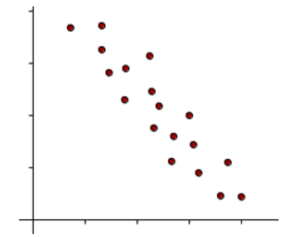
- **Propietats de l'índex de correlació de Pearson**

L'índex de correlació de Pearson és un nombre real comprès entre -1 i +1.

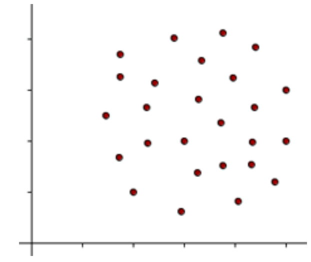
+1 indica una relació lineal positiva perfecta.



-1 indica una relació lineal negativa perfecta.



0 indica absència de relació lineal.



Introducció

- **<u>Propietats de l'índex de correlació de Pearson</u>**
- No té unitats.
- No varia quan les variables es transformen linealment.
- El signe del coeficient de correlació és el mateix que el de la covariància.

  Si la covariància és positiva → correlació directa (positiva).

  Si la covariància és negativa → correlació inversa (negativa).

  Si la covariància és nul·la → no hi ha correlació.

Introducció

- <u>**Propietats de l'índex de correlació de Pearson**</u>

  La magnitud de la correlació es mesura en funció de quant s'aproxima als extrems.

  – Valors pròxims a −1: correlació **forta i inversa/negativa**

  – Valors pròxims a 1: correlació **forta i directa/positiva**

  – Valors pròxims a 0: correlació **dèbil**

- Valors

  – Sense correlació: 0,0

  – Correlació baixa: 0,2 - 0,4

  – Correlació moderada: 0,4 - 0,6

  – Correlació bona: 0,6 - 0,8

  – Correlació molt bona: 0,8 - 1,0

  – Correlació perfecta: 1,0

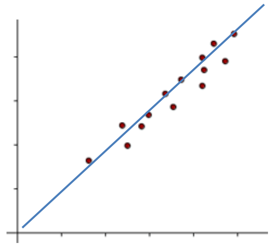
Introducció

- **Propietats**

  Amb el coeficient de correlació de Pearson no es pot establir que hi haja causalitat ni la direcció de la causalitat.

Introducció

- **Coeficient de determinació: R^2_{xy}** → Coeficient de correlació elevat al quadrat.
- Una vegada ajustada la recta de regressió al núvol d'observacions, R^2 mesura la bondat d'ajustament.



- Informa de la proporció de la variància d'una variable que es pot explicar a partir de l'altra proporció de variància.
- Criteri segons Cohen per a valorar si aquesta capacitat explicativa és:
  - Baixa = 0,1
  - Mitjana = 0,3
  - Alta = 0,5

Introducció

- Què es pot fer per a establir la relació de variables quan NO són quantitatives?
  - Variables ordinals

    Coeficient rho de Spearman

    Coeficient de Kendall

    Coeficient de Goodman-Kruskal

  - Variables dicotòmiques

    Coeficient de correlació biserial-puntual

    Coeficient de correlació biserial

    Coeficient de correlació tetracòrica

Execució de la prova

Hi ha correlació lineal entre l'edat i el risc de patir caigudes? (Aquest últim mesurat amb l'escala de Tinetti).

- Variables

  Edat (quantitativa)

  Tinetti (quantitativa)

Escriptura dels resultats

- L'edat i els resultats de l'escala de Tinetti estan correlacionats entre si, $r = -0{,}518$, $p < 0{,}05$.

- De forma opcional podem dir que:
  - La correlació és moderada i negativa.
  - La variable risc de patir caigudes s'explica en un 26,8% per l'edat $R^2 = 0{,}268$).

http://statistics.blogs.uv.es/

VNIVERSITAT DE VALÈNCIA

REGRESSIÓ LINEAL SIMPLE I MÚLTIPLE

Dra. Pilar Serra

VNIVERSITAT DE VALÈNCIA

REGRESSIÓ LINEAL

Introducció
Execució de la prova
Resultats
Escriptura dels resultats

VNIVERSITAT DE VALÈNCIA

Introducció

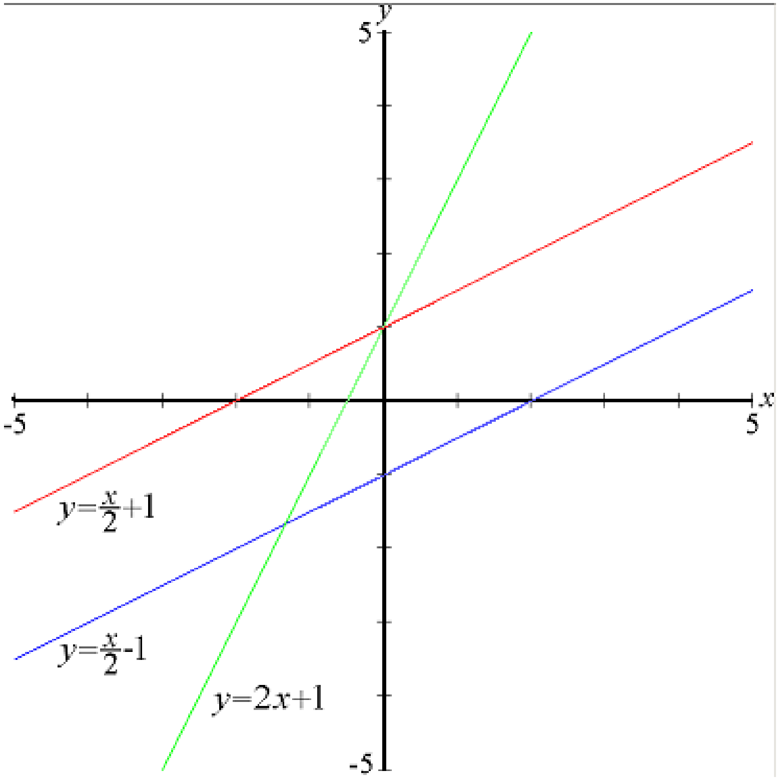
SUBJECTE	FUNCIONALITAT	FORÇA	MOBILITAT	RESISTÈNCIA
1	66	14,5	49,2	29,3
2	56	21,5	35,2	20,3
3	64	24,5	59,8	42,7
4	80	29,5		
5	52	18,5		
6	72	18,5		
7	62	19,5		
8	70	28,5		
9	57	23		
10	61	10		
11	66	21,5		
12	50	19,5		
13	63	28		
14	73	19		
15	69	15		
16	69	22		
17	56	22,5		
18	74	32,5		
19	57	25		
20	58	12,5		
21	80	31,5		
22	74	29,5	60,8	36,7
23	76	19,5	63,2	25,3
24	58	18,5	43,6	25,9
25	80	23,5	66	39,5
26	70	17,5	60	40,5
27	52	7,5	46,4	23,1
28	79	34	61,8	31,2
29	76	34,5	71,2	31,3
		13,5	39	37,5
		26,5	61,6	38,9
		25	44	36
		19,5	31,2	16,3
		16,5	56,8	28,7
		24	63,8	45,2
		9,5	53,6	25,9
		17	33,6	14,4
		26	40,4	23,6
		30,5	77	30,5
		20,5	26	30,5
		3,5	40,4	15,1
		13,5	28,8	14,7
		22,5	38,8	25,7
		23	47,2	26,8
		26	61,8	42,2
		4,5	33	16,5
		22,5	38	33,5
48	64	27,5	50,8	22,7
49	70	25,5	63	28,5
50	69	16	55,8	17,2

La funcionalitat de les persones amb lesió medul·lar es pot estimar mitjançant les qualitats físiques?

Introducció

REGRESSIÓ LINEAL SIMPLE

$$Y_i = (b_0 + b_1 X_i) + \varepsilon_i$$

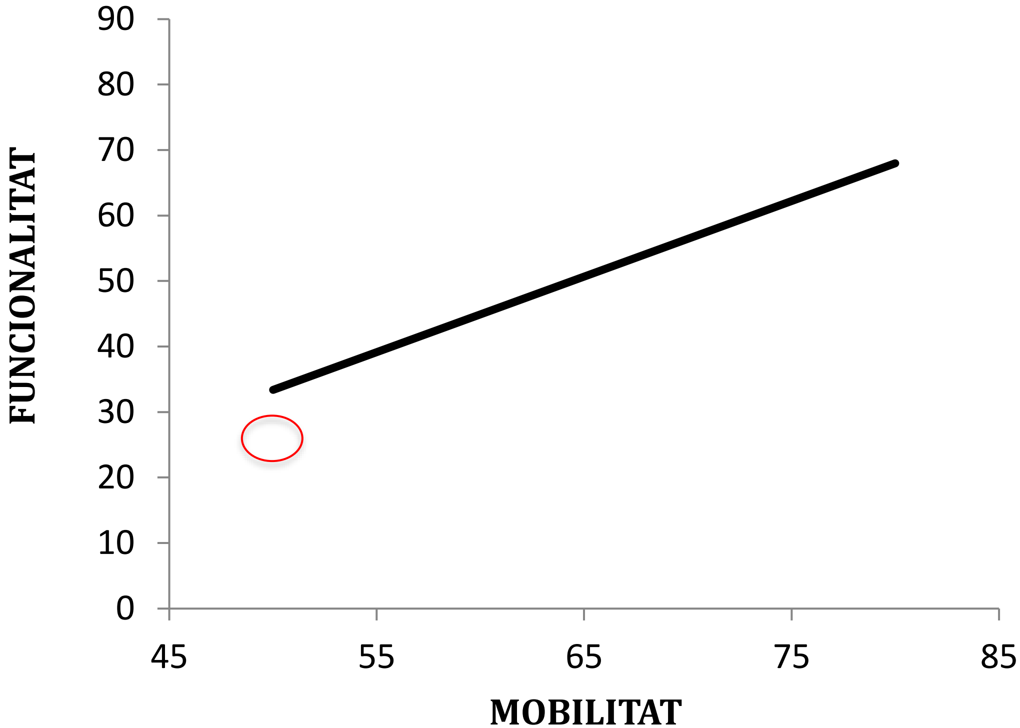


- Permet determinar quins són els coeficients b_0 i b_1 que relacionen linealment la variable d'entrada (*X*) amb la variable d'eixida (*Y*).
- El coeficient b_0 o constant és el valor que pren la variable d'eixida (*Y*) quan la variable d'entrada (*X*) val 0.
- El coeficient b_1 multiplica la variable d'entrada (*X*) i, per tant, determina la inclinació de la recta. Com més gran és b_1, la recta té més inclinació i, per tant, petits canvis en la variable d'entrada (*X*) generen canvis grans en la variable d'eixida (*Y*).

Introducció

REGRESSIÓ LINEAL SIMPLE

$$Y_i = (b_0 + b_1 X_i) + \varepsilon_i$$

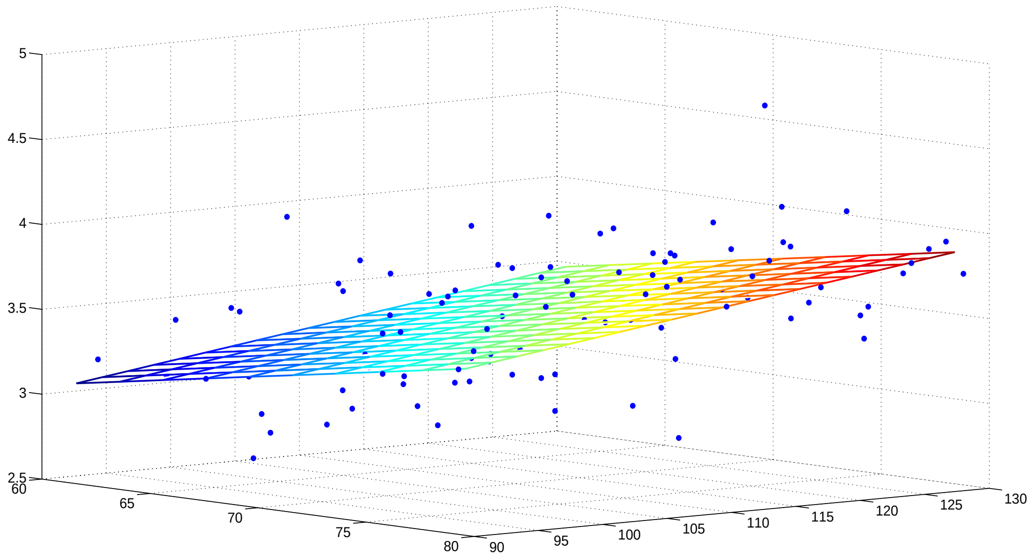


- Hi ha infinites possibles rectes.
- Necessitem trobar la que minimitze les distàncies entre els valors observats i els predits per l'equació.
- Aquesta recta s'obté mitjançant un procés matemàtic denominat mínims quadrats.

Introducció

REGRESSIÓ LINEAL MÚLTIPLE

$$Y_i = (b_0 + b_1 X_{i1} + b_2 X_{i2} + \ldots + b_n X_n) + \varepsilon_i$$



- Hi ha més d'una variable d'entrada (X_1, X_2, ..., X_n) per a predir la variable d'eixida (Y).

Introducció

PER A GENERALITZAR EL MODEL A LA POBLACIÓ

1. COMPLIMENT DE SUPÒSITS

- Tipus de variables
- Independència
- Sense multicol·linealitat perfecta
  - Màxim factor d'inflació de variància (FIV) < 10
  - Mitjana dels FIV ≈ 1
- Independència dels residus
  - Test de Durbin-Watson ≈ 2

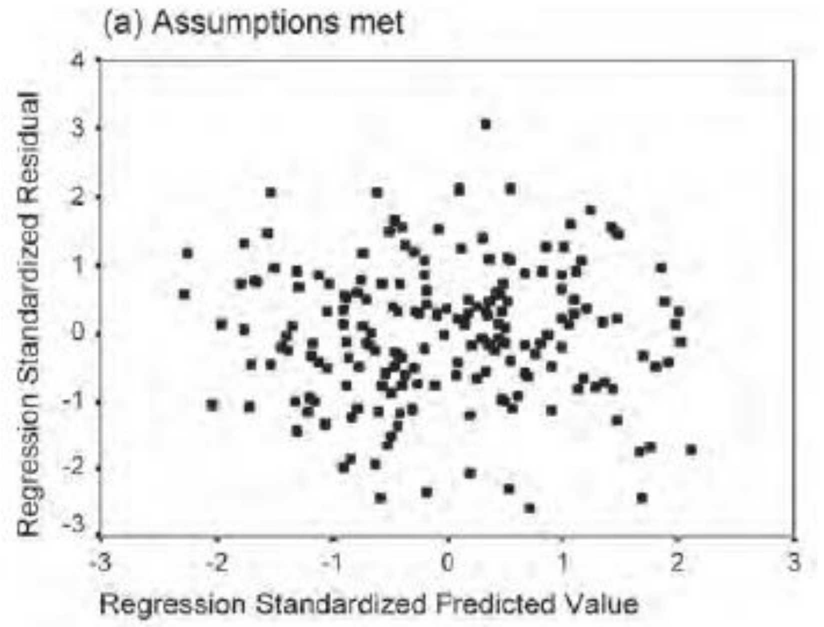
Introducció

PER A GENERALITZAR EL MODEL A LA POBLACIÓ

1. COMPLIMENT DE SUPÒSITS

- Homoscedasticitat dels residus
- Linealitat

Field A.: *Discovering statistics using SPSS.* Londres, Sage, 2009.



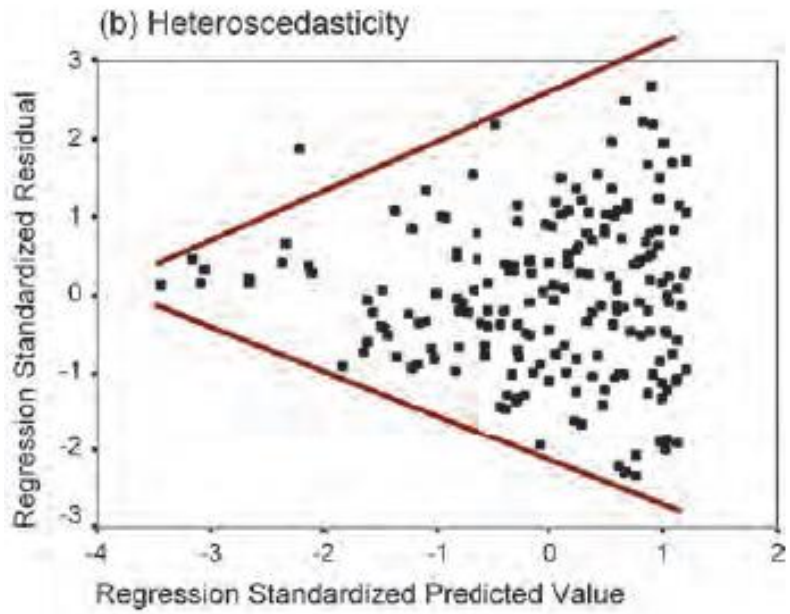
Introducció

PER A GENERALITZAR EL MODEL A LA POBLACIÓ

1. COMPLIMENT DE SUPÒSITS

- Homoscedasticitat dels residus
- Linealitat

Field A.: *Discovering statistics using SPSS.* Londres, Sage, 2009.



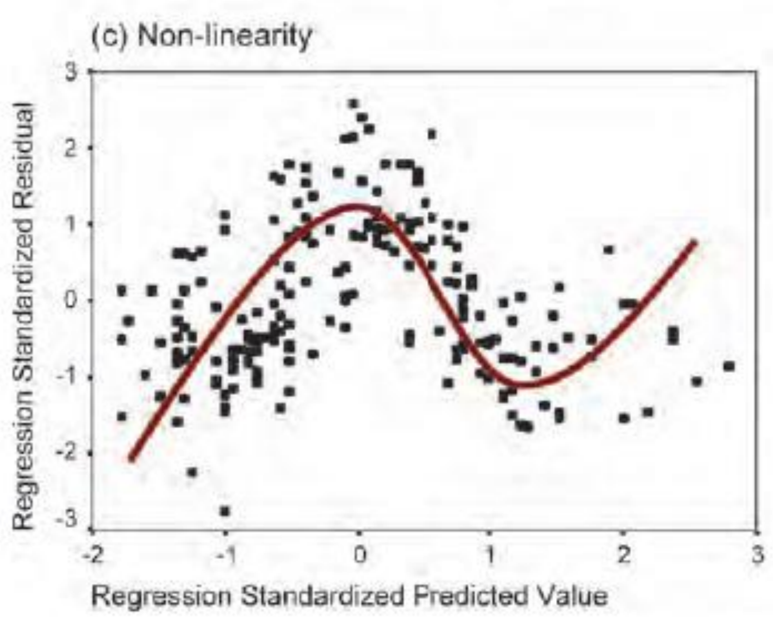
Introducció

PER A GENERALITZAR EL MODEL A LA POBLACIÓ

1. COMPLIMENT DE SUPÒSITS

- Homoscedasticitat dels residus
- Linealitat

Field A.: *Discovering statistics using SPSS.* Londres, Sage, 2009.



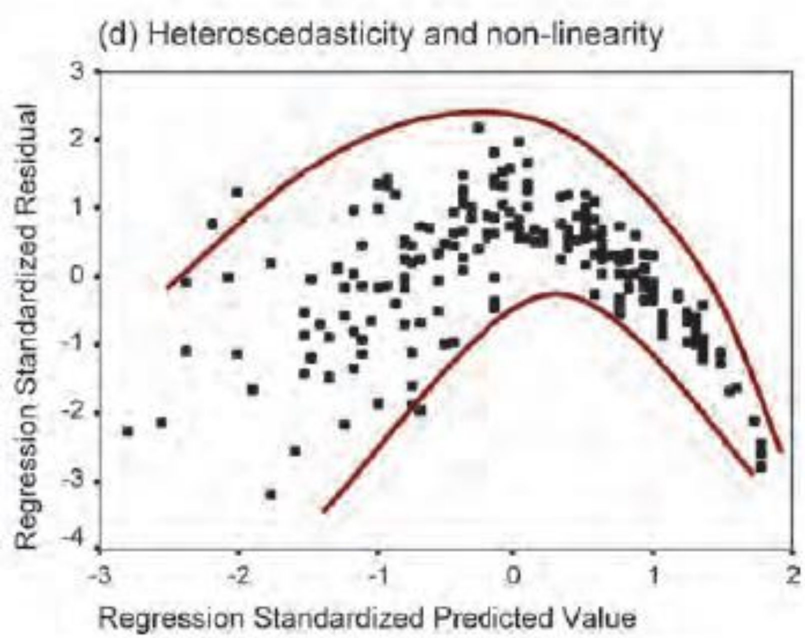
Introducció

PER A GENERALITZAR EL MODEL A LA POBLACIÓ

1. COMPLIMENT DE SUPÒSITS

- Homoscedasticitat dels residus
- Linealitat

Field A.: *Discovering statistics using SPSS.* Londres, Sage, 2009.

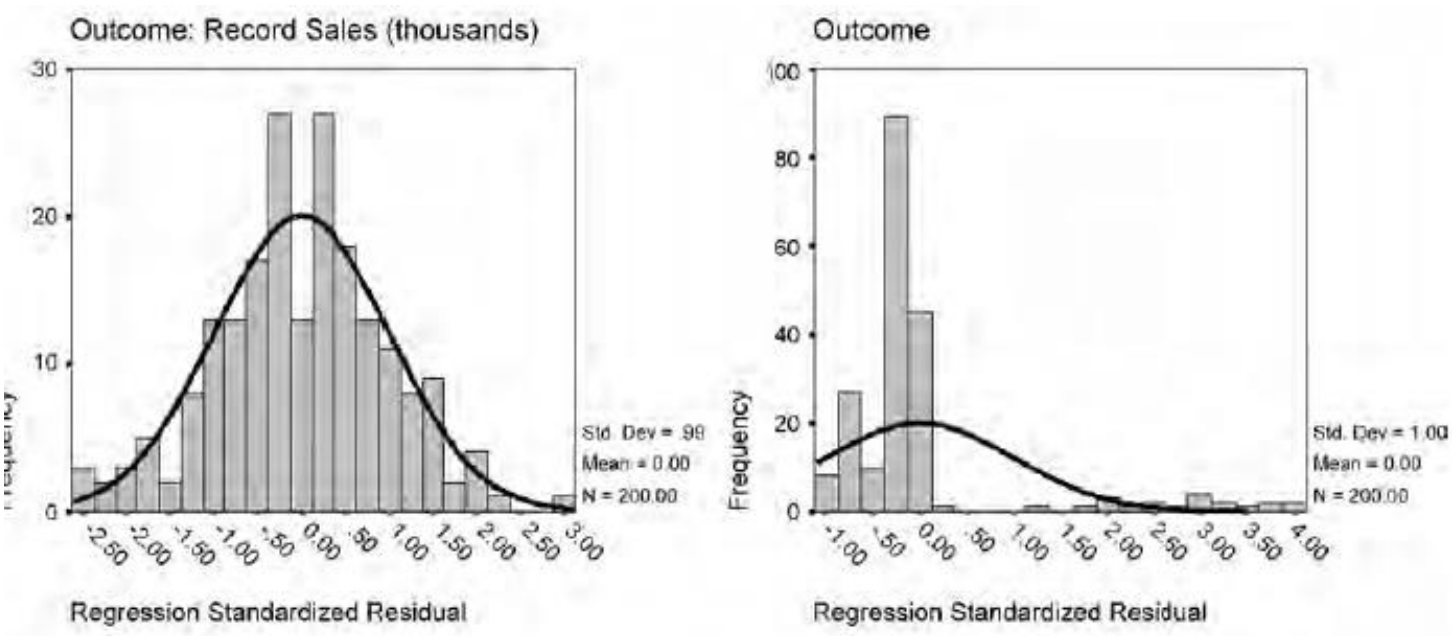


Introducció

PER A GENERALITZAR EL MODEL A LA POBLACIÓ

1. COMPLIMENT DELS SUPÒSITS

- Distribució normal dels residus (test K-S o de S-W)



Field A.: *Discovering statistics using SPSS.*
Londres, Sage, 2009.

Introducció

PER A GENERALITZAR EL MODEL A LA POBLACIÓ

2. VALIDACIÓ CREUADA

- Obtenció del model amb una part de la mostra (e. g., 80%)
- Avaluació del model sobre la mostra restant (e. g., 20%)

$$r = \frac{S_{XY}}{S_X\, S_Y}$$

$$MAD = \frac{1}{n}\sum_{i=1}^{n} | D_t - F_t |$$

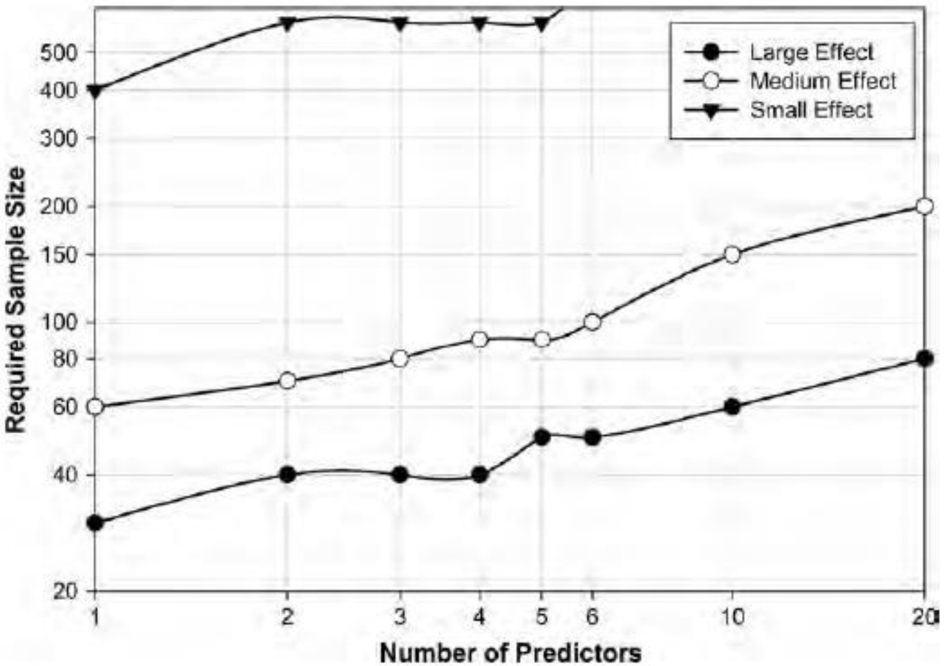
$$MSE = \frac{1}{n}\sum_{i=1}^{n} (D_t - F_t)^2$$

Introducció

PER A GENERALITZAR EL MODEL A LA POBLACIÓ

3. MIDA DE LA MOSTRA

- n = 10 * k
- n = 50 + 8 * k



Escriptura dels resultats

L'anàlisi de regressió lineal mostra l'existència d'una relació entre les variables que és explicada per l'equació següent:

$$Y = 28{,}33 + 0{,}298\, X_1 + 0{,}598\, X_2$$

En què Y es la funcionalitat, X_1 la força i X_2 la mobilitat.

El coeficient de determinació ha sigut de 0,82 i l'error quadràtic mitjà de 15,99. La taula 1 mostra els coeficients tipificats i els valors de probabilitat.

	B	*SE B*	*B* estandarditzat
Constant	28,33	2,56	
Força	0,298	0,09	0,225 *
Mobilitat	0,598	0,05	0,78 *

* Indica que el valor dels coeficients és significativament diferent de 0 ($p < 0{,}005$).

http://statistics.blogs.uv.es

VNIVERSITAT DE VALÈNCIA