

Análisis de datos

en ciencias sociales y de la salud II

Antonio Pardo • Rafael San Martín

Datos disponibles en: www.sintesis.com

MACS

2.ª Edición



EDITORIAL
SÍNTESIS

Análisis de datos

en ciencias sociales y de la salud II

2.ª Edición

PROYECTO EDITORIAL:
Metodología de las Ciencias del Comportamiento y de la Salud

Directores:
Antonio Pardo Merino
Miguel Ángel Ruiz Díaz



Queda prohibida, salvo excepción prevista en la ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra sin contar con autorización de los titulares de la propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual (arts. 270 y sigs. Código Penal). El Centro Español de Derechos Reprográficos (www.cedro.org) vela por el respeto de los citados derechos.

Análisis de datos

en ciencias sociales y de la salud II

Antonio Pardo • Rafael San Martín

2.ª Edición



EDITORIAL
SINTESIS

Consulte nuestra página web: www.sintesis.com
En ella encontrará el catálogo completo y comentado

Reservados todos los derechos. Está prohibido, bajo las sanciones penales y el resarcimiento civil previstos en las leyes, reproducir, registrar o transmitir esta publicación, íntegra o parcialmente, por cualquier sistema de recuperación y por cualquier medio, sea mecánico, electrónico, magnético, electroóptico, por fotocopia o por cualquier otro, sin la autorización previa por escrito de Editorial Síntesis, S. A.

© Antonio Pardo y Rafael San Martín

© EDITORIAL SÍNTESIS, S. A.
Vallehermoso, 34. 28015 Madrid
Teléfono 91 593 20 98
<http://www.sintesis.com>

ISBN:978-84-975670-4-6
Depósito Legal: M. 14.400-2015

Impreso en España - Printed in Spain

Índice de contenidos

Presentación	15
--------------------	----

1. La inferencia estadística

La inferencia estadística	17
El contraste de hipótesis	18
Un ejemplo	21
Qué significa rechazar y no rechazar la hipótesis nula	24
Contrastes bilaterales y unilaterales	27
Errores Tipo I y II, y potencia de un contraste	29
Factores de los que depende la potencia	32
Tamaño del efecto	35
Clasificación de los contrastes de hipótesis	38
Programas informáticos para el análisis de datos	40
Apéndice 1	
Cálculo de la potencia	41
Curva de potencias	43
Ejercicios	46

2. Inferencia con una variable

Contrastes sobre el centro de una distribución	52
La prueba de Wilcoxon para una muestra	52
La prueba de los signos para una muestra	55
La prueba T , la de Wilcoxon y la de los signos	60
Los contrastes sobre el centro de una distribución con SPSS	61
Contrastes sobre la dispersión de una distribución	64
Contrastes sobre la forma de una distribución	66
La prueba de Kolmogorov-Smirnov sobre bondad de ajuste	67
Los contrastes sobre la forma de una distribución con SPSS	68
Apéndice 2	
La prueba de las rachas	73
La prueba de los cuantiles	76
Ejercicios	77

3. Inferencia con dos variables categóricas

Proporciones independientes y relacionadas	80
Homogeneidad marginal y simetría	83
Simetría con variables dicotómicas: la prueba de McNemar	84
Simetría con variables politómicas: la prueba de Bowker	88
Homogeneidad marginal y simetría con SPSS	89
Índices de riesgo	94
Riesgo relativo	95
<i>Odds ratio</i>	97
Relación entre el riesgo relativo y la <i>odds ratio</i>	100
Consideraciones sobre la interpretación de los índices de riesgo	101
Los índices de riesgo con SPSS	104
Índices de acuerdo	106
Acuerdo con variables nominales: <i>kappa</i>	107
El índice de acuerdo <i>kappa</i> con SPSS	112
Asociación entre variables categóricas ordinales	114
Medidas de concordancia-discordancia	116
Medidas de concordancia-discordancia con SPSS	119
Apendice 3	
Simetría relativa	120
Combinación de tablas 2 · 2 (Cochran y Mantel-Haenszel)	121
La paradoja de Simpson	125
Medidas de asociación basadas en la reducción proporcional del error	126
Coeficiente lambda	127
Coeficiente tau	128
Coeficiente de incertidumbre	129
Medidas de reducción proporcional del error con SPSS	129
Muestras pequeñas: la prueba exacta de Fisher	130
Concordancias y discordancias en una tabla de contingencias	132
Cómo reproducir una tabla de contingencias en SPSS	133
Ejercicios	135

4. Inferencia con dos variables cuantitativas

Como comparar dos variables cuantitativas	140
La prueba <i>t</i> de Student para muestras relacionadas	140
La prueba de Wilcoxon para dos muestras	142
La prueba de Wilcoxon para dos muestras con SPSS	145
La prueba de los signos para dos muestras	148
La prueba de los signos para dos muestras con SPSS	149
Como relacionar dos variables cuantitativas	151
El coeficiente de correlación de Pearson	151
El coeficiente de correlación de Spearman	153
El coeficiente de correlación de Spearman con SPSS	157

Medidas del tamaño del efecto	158
Cálculo de la potencia y del tamaño muestral	159
Apéndice 4	
Correlaciones parciales	160
Ejercicios	162
5. Inferencia con una variable categórica y una cuantitativa	
La prueba <i>T</i> de Student para muestras independientes	166
Asumiendo varianzas poblacionales iguales	168
Independencia, normalidad e igualdad de varianzas	172
No asumiendo varianzas poblacionales iguales	173
La prueba <i>T</i> para muestras independientes con SPSS	176
La prueba <i>U</i> de Mann-Whitney	179
La prueba <i>U</i> de Mann-Whitney con SPSS	183
Medidas del tamaño del efecto	184
Medidas del tamaño del efecto con SPSS	189
Cálculo de la potencia y del tamaño muestral	190
Contrastes de equivalencia y no-inferioridad	191
Métodos para demostrar equivalencia	192
Doble contraste unilateral de Schuirmann	192
Intervalo de confianza de Westlake	193
Métodos para demostrar no-inferioridad	195
Limitaciones de los contrastes de equivalencia y no-inferioridad	196
Apéndice 5	
La distribución muestral del estadístico <i>T</i> asumiendo $\sigma_1 = \sigma_2$	197
La prueba de Kolmogorov-Smirnov para dos muestras independientes	199
La prueba de las rachas de Wald-Wolfowitz	199
La prueba de reacciones extremas de Moses	200
Ejercicios	201
6. Análisis de varianza (I). Un factor completamente aleatorizado	
Modelos de análisis de varianza	206
Número de factores	206
Tipo de asignación de las unidades de análisis a las condiciones del estudio	207
Forma de establecer los niveles del factor	208
Clasificación de los modelos de análisis de varianza	209
La lógica del análisis de varianza	210
ANOVA de un factor completamente aleatorizado	213
Supuestos del ANOVA de un factor	217
Independencia	218
Normalidad	219
Igualdad de varianzas (homocedasticidad)	219

Transformación de las puntuaciones	220
Efectos fijos y aleatorios	222
Medidas del tamaño del efecto	222
Cálculo de la potencia y del tamaño muestral	226
Comparaciones múltiples entre medias	228
Comparaciones lineales	228
Tasa de error en las comparaciones múltiples	231
Comparaciones planeadas o <i>a priori</i>	233
Prueba de Dunn-Bonferroni	234
Comparaciones de tendencia	238
Prueba de Dunnett	241
Comparaciones <i>post hoc</i> o <i>a posteriori</i>	243
Prueba de Tukey	243
Prueba de Scheffé	246
Comparaciones múltiples: qué procedimiento elegir	248
ANOVA de un factor completamente aleatorizado con SPSS	249
Supuestos	250
Hipótesis de igualdad de medias	252
Tamaño del efecto y potencia observada	254
Comparaciones <i>post hoc</i>	255
Comparaciones planeadas y de tendencia	257
La prueba de Kruskal-Wallis	260
Apéndice 6	
Distribución muestral del estadístico F	264
Tabla de la distribución F	266
La distribución F con SPSS	266
Estadísticos F robustos: Welch y Brown-Forsythe	267
Comparaciones <i>post hoc</i> : procedimientos alternativos	267
Efectos aleatorios	270
Ejercicios	272

7. Análisis de varianza (II). Dos factores completamente aleatorizados

Estructura de los datos y notación	278
La interacción entre factores	280
ANOVA de dos factores completamente aleatorizados	284
Supuestos del ANOVA de dos factores	291
Efectos fijos y aleatorios	291
Medidas del tamaño del efecto	292
Cálculo de la potencia y del tamaño muestral	294
Comparaciones múltiples	296
Efectos principales	296
Efectos simples	299
Efecto de la interacción	302

ANOVA de dos factores completamente aleatorizados con SPSS	306
Hipótesis globales (efecto de A , de B y de AB)	307
Tamaño del efecto y potencia observada	308
Comparaciones múltiples: efectos principales	309
Comparaciones múltiples: efectos simples	311
Comparaciones múltiples: efecto de la interacción	313
Apéndice 7	
Casillas con tamaños muestrales distintos	315
Más sobre los efectos simples y el efecto de la interacción	317
La sentencia <i>LMATRIX</i>	319
Análisis de los efectos simples	320
Comparaciones entre los niveles de un mismo efecto simple	321
Análisis del efecto de la interacción	323
Modelos jerárquicos o anidados	324
Ejercicios	325
8. Análisis de varianza (III). Un factor con medidas repetidas	
Características de los diseños con medidas repetidas	329
ANOVA de un factor con medidas repetidas	332
Supuestos del modelo	338
Alternativas al estadístico F	339
Estadístico F con los grados de libertad modificados	339
Aproximación multivariada	340
Qué solución elegir	340
Medidas del tamaño del efecto	341
Cálculo de la potencia y del tamaño muestral	342
Comparaciones múltiples	343
ANOVA de un factor con medidas repetidas con SPSS	344
Esfericidad	346
Igualdad de medias	346
Tamaño del efecto y potencia observada	348
Comparaciones planeadas	348
Comparaciones <i>post hoc</i>	350
La prueba de Friedman	351
Apéndice 8	
Cómo estimar ϵ	354
Prueba de Cochran	356
Coeficiente de concordancia W de Kendall	358
Prueba de no-aditividad de Tukey	360
Ejercicios	362

9. Análisis de varianza (IV). Dos factores con medidas repetidas

ANOVA de dos factores con medidas repetidas en ambos	367
Supuestos del modelo	371
Medidas del tamaño del efecto	372
Cálculo de la potencia y del tamaño muestral	372
Comparaciones múltiples	373
ANOVA de dos factores con medidas repetidas en ambos con SPSS	373
Esfericidad	377
Hipótesis globales (efecto de <i>A</i> , de <i>B</i> y de <i>AB</i>)	378
Tamaño del efecto y potencia observada	380
Comparaciones <i>post hoc</i> : efectos principales	380
Comparaciones múltiples: efectos simples	381
Comparaciones múltiples: efecto de la interacción	382
ANOVA de dos factores con medidas repetidas en uno	385
Supuestos del modelo	389
Medidas del tamaño del efecto	390
Cálculo de la potencia y del tamaño muestral	391
Comparaciones múltiples	391
ANOVA de dos factores con medidas repetidas en uno con SPSS	392
Esfericidad multi-muestra e igualdad de varianzas	396
Hipótesis globales (efecto de <i>A</i> , de <i>B</i> y de <i>AB</i>)	397
Comparaciones <i>post hoc</i> : efectos principales	399
Comparaciones múltiples: efectos simples	400
Comparaciones múltiples: efecto de la interacción	401
Apéndice 9	
La sentencia <i>MMATRIX</i>	403
Análisis de los efectos simples	404
Comparaciones entre los niveles de un mismo efecto simple	405
Análisis del efecto de la interacción	407
Ejercicios	409

10. Análisis de regresión lineal

Regresión lineal simple	414
Diagramas de dispersión	416
La recta de regresión	417
Mínimos cuadrados	418
Coeficientes de regresión	419
Coeficientes de regresión tipificados	421
Bondad de ajuste	422
Significación de los coeficientes de regresión	425
Intervalos de confianza	427
Cómo efectuar pronósticos	428

Regresión lineal múltiple	430
La ecuación de regresión	430
Coeficientes de regresión	431
Coeficientes de regresión tipificados	432
Bondad de ajuste	433
Significación de los coeficientes de regresión	434
Importancia relativa de las variables	437
Variables independientes categóricas	440
Regresión jerárquica o por pasos	441
Criterios para seleccionar variables	442
Métodos para seleccionar variables	443
Supuestos del modelo de regresión lineal	445
Linealidad	445
No colinealidad	447
Independencia	449
Normalidad	451
Homocedasticidad	452
Casos atípicos e influyentes	453
Casos atípicos	453
Casos atípicos en Y	453
Casos atípicos en X_j	455
Casos influyentes	455
Cambio en los coeficientes de regresión	457
Cambio en los pronósticos	458
Cambio en los residuos	458
Qué hacer con los casos atípicos e influyentes	458
Regresión lineal con SPSS	459
Regresión múltiple	459
Bondad de ajuste	460
Ecuación de regresión	461
Significación de los coeficientes de regresión	462
Pronósticos	462
Importancia relativa de las variables	462
Chequeo de los supuestos	464
Casos atípicos e influyentes	468
Regresión jerárquica o por pasos	469
Regresión jerárquica o por pasos con variables categóricas	472
Comentarios finales	473
Validación de una ecuación de regresión	473
Consideraciones sobre el tamaño muestral	473
Apéndice 10	
Regresión curvilínea	474
Mediación estadística	477
Ejercicios	480

Apéndice final. Tablas estadísticas	485
Glosario de símbolos	515
Referencias	519
Índice de materias	535

Los contenidos y la exposición del presente manual están orientados a resaltar la utilidad de los procedimientos estadísticos disponibles y las claves para saber elegirlos, aplicarlos e interpretarlos correctamente, sin prescindir de los detalles necesarios para entenderlos. En esta segunda edición revisada se han reorganizado algunos contenidos y se han aligerado otros, así como incluido algunas modificaciones con el propósito de mejorar las explicaciones.

Esta obra va dirigida a estudiantes, profesores e investigadores que se ven en la necesidad de aplicar herramientas estadísticas. Al diseñarla se ha asumido que los profesionales de las ciencias sociales y de la salud no necesitan ser expertos en los fundamentos matemáticos de las herramientas estadísticas que aplican. Además, gracias a los programas informáticos de análisis estadístico, ya no es necesario invertir tiempo en complejos cálculos. Aunque todos los procedimientos se presentan con suficiente detalle para poder aplicarlos con una calculadora de bolsillo, se explica cómo aplicarlos con el programa *SPSS Statistics* y cómo interpretar correctamente los resultados que se obtienen.

Antonio Pardo es profesor titular del departamento de Psicología Social y Metodología de la Universidad Autónoma de Madrid.

Rafael San Martín es catedrático en la Universidad Autónoma de Madrid.

ISBN 978-84-975670-4-6



55 103



9 788497 567046

Datos disponibles en:
www.sintesis.com



EDITORIAL
SÍNTESIS