

BIOESTADÍSTICA AMIGABLE



BIOESTADÍSTICA AMIGABLE

3.^a edición

EDITORES

Director

Miguel Ángel Martínez-González
Catedrático de Medicina Preventiva y Salud Pública,
Universidad de Navarra

Codirectores

Almudena Sánchez-Villegas
Profesora Titular de Medicina Preventiva y Salud Pública,
Universidad de Las Palmas de Gran Canaria

Estefanía Toledo Atucha

Profesora Contratada Doctora de Medicina Preventiva y
Salud Pública, Universidad de Navarra

Javier Faulín Fajardo

Catedrático de Estadística e Investigación Operativa,
Universidad Pública de Navarra

Ámsterdam Barcelona Beijing Boston Filadelfia Londres Madrid
México Milán Múnich Orlando París Roma Sidney Tokio Toronto



ELSEVIER



ELSEVIER

© 2014 Elsevier España, S.L.
Travessera de Gràcia, 17-21
08021 Barcelona, España

Fotocopiar es un delito (Art. 270 C.P.)

Para que existan libros es necesario el trabajo de un importante colectivo (autores, traductores, dibujantes, correctores, impresores, editores...). El principal beneficiario de ese esfuerzo es el lector que aprovecha su contenido.

Quien fotocopia un libro, en las circunstancias previstas por la ley, delinque y contribuye a la «no» existencia de nuevas ediciones. Además, a corto plazo, encarece el precio de las ya existentes.

Este libro está legalmente protegido por los derechos de propiedad intelectual. Cualquier uso fuera de los límites establecidos por la legislación vigente, sin el consentimiento del editor, es ilegal. Esto se aplica en particular a la reproducción, fotocopia, traducción, grabación o cualquier otro sistema de recuperación y almacenaje de información.

ISBN (versión impresa): 978-84-9022-500-4
ISBN (versión electrónica): 978-84-9022-651-3

Depósito legal (versión impresa): xxxxxxxxxxxx
Depósito legal (versión electrónica): xxxxxxxxxxxx

Coordinación y producción editorial: **GEA CONSULTORÍA EDITORIAL, S. L.**

Advertencia

La medicina es un área en constante evolución. Aunque deben seguirse unas precauciones de seguridad estándar, a medida que aumenten nuestros conocimientos gracias a la investigación básica y clínica habrá que introducir cambios en los tratamientos y en los fármacos. En consecuencia, se recomienda a los lectores que analicen los últimos datos aportados por los fabricantes sobre cada fármaco para comprobar las dosis recomendadas, la vía y duración de la administración y las contraindicaciones. Es responsabilidad ineludible del médico determinar las dosis y el tratamiento más indicados para cada paciente, en función de su experiencia y del conocimiento de cada caso concreto. Ni los editores ni los directores asumen responsabilidad alguna por los daños que pudieran generarse a personas o propiedades como consecuencia del contenido de esta obra.

El Editor

PRESENTACIÓN DE LA TERCERA EDICIÓN

p0010 *Bioestadística amigable* es un texto que ha ido ganando popularidad durante la última década. Quizás esto se deba a su orientación eminentemente práctica, con un contenido matemático mínimo. Son muchas las facultades que lo han recomendado para el aprendizaje de la Bioestadística. Esto es un motivo de alegría y de gratitud hacia quienes han confiado en nosotros. También esta grata experiencia, junto con otras razones, nos ha supuesto un buen incentivo para acometer una nueva edición, con la idea de poder mejorar el servicio a nuestros lectores. Hemos tenido la fortuna de contar con la editorial Elsevier para esta nueva edición.

p0015 En las dos ediciones previas, realizadas magníficamente por Díaz de Santos, la intención del libro fue doble: por una parte, se buscaba enseñar al lector a identificar los procedimientos estadísticos que están indicados para afrontar cada diseño de investigación, y, por otra, se orientó el texto hacia el aprendizaje de la interpretación clínica o práctica de los resultados de un programa informático estándar. Siempre hemos procurado alcanzar una cobertura suficientemente amplia y poder abarcar todos los procedimientos estadísticos que están realmente presentes de modo habitual en la investigación publicada en revistas biomédicas.

p0020 Las dos ediciones anteriores, de 2001 y 2006, estaban muy centradas en el programa SPSS. Esto se debía a que SPSS era el *software* más utilizado tanto en hospitales como en investigaciones de laboratorio en el mundo de habla hispana. No obstante, también se incluyeron en la segunda edición aplicaciones con otros programas (Excel, STATA, SAS y Splus o su versión gratuita, R). Pero ya hace tiempo que el software STATA se está demostrando superior en muchos aspectos a SPSS, sin dejar de ser igualmente *amigable*. Además una licencia de STATA es siempre más accesible desde el punto de vista económico que una licencia de SPSS.

p0025 Por otra parte, especialmente en el ámbito de la epidemiología, casi se ha abandonado ya SPSS y se usa preferentemente STATA (en España), o bien SAS (en EE. UU. o Europa). Pero SAS es mucho más caro que SPSS y, por supuesto, que STATA. Otro programa emergente y muy completo es R, que tiene la gran ventaja de que es gratuito. Pero R es menos amigable, y está más pensado para profesionales que ya tienen buenos conocimientos de estadística matemática y programación. Todo esto hace que la balanza se incline por dar preferencia a STATA. No obstante la preferencia por STATA, hemos procurado presentar siempre en este libro el modo de resolver cada procedimiento también con SPSS y con R. También se presentan posibles aplicaciones realizadas con Excel, porque pueden ser muy didácticas. En definitiva, no es imprescindible contar con STATA para que el texto cumpla su utilidad.

p0030 Pero nuestra experiencia nos dice que STATA es ideal para quien empieza desde cero. Por eso hemos dado preferencia a STATA en esta nueva edición, que escribimos con la pretensión de que sea todavía más *amigable*. También hemos puesto los medios para que esta edición sea más breve que la anterior en aras de acercarla aún más al lector interesado en la aplicación práctica de los resultados estadísticos en las ciencias de la vida.

Miguel A. Martínez-González
Houston, Texas, Junio 2013



AUTORES

Francisco Javier Basterra-Gortari

Profesor Asociado de Medicina Preventiva y Salud Pública, Universidad de Navarra.
Especialista MIR en Endocrinología.

Maira Bes-Rastrollo

Profesora Titular de Medicina Preventiva y Salud Pública, Universidad de Navarra.

Jokin de Irala Estévez

Catedrático de Medicina Preventiva y Salud Pública, Universidad de Navarra.

Pedro A. de la Rosa Fernández-Pacheco

Residente de Medicina Preventiva y Salud Pública, Servicio Navarro de Salud-Osasunbidea.

Martín García-López

Especialista MIR en Cardiología. Clínica Universidad de Navarra.

Alfredo Gea Sánchez

Profesor Ayudante Doctor de Medicina Preventiva y Salud Pública, Universidad de Navarra.

Francisco Guillén-Grima

Catedrático de Medicina Preventiva y Salud Pública, Universidad Pública de Navarra.

Cristina López del Burgo

Profesora Contratada Doctora de Medicina Preventiva y Salud Pública, Universidad de Navarra.

Jesús López-Fidalgo

Catedrático de Estadística e Investigación Operativa, Universidad de Castilla-La Mancha.

Nerea Martín-Calvo

Becaria Río Hortega. Medicina Preventiva y Salud Pública, Universidad de Navarra.
Doctora en Medicina. Especialista MIR en Pediatría.

Jorge María Núñez-Córdoba

Especialista en Medicina Preventiva y Salud Pública, Doctor en Medicina.
Clínica Universidad de Navarra.

Miguel Ruiz-Canela

Profesor Titular de Medicina Preventiva y Salud Pública, Universidad de Navarra.

Carmen Sayón-Orea

Profesora Ayudante Doctora de Medicina Preventiva y Salud Pública, Universidad de Navarra.

Juan B. Toledo Atucha

Investigador posdoctoral en el Centro de Investigación de Enfermedades Neurodegenerativas,
Facultad de Medicina, Universidad de Pensilvania, Filadelfia, EE. UU.
Especialista MIR en Neurología.



ÍNDICE DE CAPÍTULOS

Presentación de la tercera edición	v
Autores	vii
1 Introducción a los métodos de la epidemiología y la bioestadística..... 1 <i>M.A. Martínez-González, A. Sánchez-Villegas, J. de Irala</i>	1
1.1. Estadística, estadística aplicada y bioestadística..... 1	1
1.2. Bioestadística descriptiva y bioestadística analítica o inferencial..... 1	1
1.3. Poblaciones y muestras..... 2	2
1.4. Etapas de la investigación científica: relaciones entre la bioestadística y la epidemiología..... 2	2
2 Procedimientos descriptivos..... 13 <i>M.A. Martínez-González, A. Gea, C. Sayón-Orea</i>	13
2.1. Tipos de variables..... 13	13
2.2. Transformación de una variable: categorización y recodificación..... 17	17
2.3. Consejos prácticos sobre categorización de variables cuantitativas..... 26	26
2.4. Representaciones gráficas..... 29	29
2.5. Medidas de tendencia central..... 43	43
2.6. Medidas de dispersión..... 46	46
2.7. Medidas de forma: asimetría y curtosis..... 51	51
2.8. Medidas de posición: cuantiles, percentiles..... 51	51
2.9. Ponderación, medias ponderadas..... 54	54
2.10. Valores extremos (<i>outliers</i>) y consejos sobre su uso..... 56	56
2.11. Preferencia de números..... 56	56
2.12. Índices estadísticos descriptivos con STATA..... 57	57
2.13. Procedimientos descriptivos con Excel..... 58	58
2.14. Procedimientos descriptivos con otros programas..... 59	59
2.15. Datos truncados o censurados..... 63	63
2.16. Resumen de las instrucciones en STATA y SPSS..... 63	63
3 Probabilidad. Distribuciones de probabilidad..... 65 <i>E. Toledo, A. Sánchez-Villegas, M.A. Martínez-González</i>	65
3.1. Introducción..... 65	65
3.2. Conceptos de probabilidad..... 65	65
3.3. Axiomas y propiedades de la probabilidad..... 66	66
3.4. Concepto de independencia..... 69	69
3.5. Probabilidad condicionada..... 69	69
3.6. Inversión de las condiciones: teorema de Bayes..... 72	72
3.7. Factor Bayes para relacionar la odds pretest con la odds postest..... 74	74
3.8. Planteamiento bayesiano, inferencia bayesiana..... 75	75
3.9. Distribuciones de probabilidad discretas..... 77	77

3.10.	Distribuciones de probabilidad continuas: distribución normal.....	80
3.11.	Teorema del límite central	87
3.12.	Condiciones, pruebas y gráficos de normalidad	88
3.13.	Las distribuciones binomial, de Poisson y normal en STATA.....	93
3.14.	Las distribuciones binomial, de Poisson y normal en el programa Excel	95
3.15.	Las distribuciones binomial, de Poisson y normal en otros programas	97
3.16.	Aproximación a las distribuciones binomial y de Poisson con la distribución normal	98
3.17.	Media y desviación estándar de una proporción	99
4	Intervalos de confianza y contraste de hipótesis.....	101
	<i>M.A. Martínez-González, J.B. Toledo, J. López-Fidalgo</i>	
4.1.	Error sistemático y error aleatorio	101
4.2.	Muestreo aleatorio o selección aleatoria	102
4.3.	Diferencia entre selección aleatoria y asignación aleatoria (aleatoriza- ción)	109
4.4.	Conceptos generales sobre estimación de parámetros	111
4.5.	Estimación de una proporción.....	112
4.6.	Estimación de una media.....	118
4.7.	Intervalos de confianza con STATA	120
4.8.	Intervalos de confianza con otros programas	121
4.9.	La distribución <i>t</i> de Student en Excel, STATA y R/Spplus	124
4.10.	Estimación de una mediana.....	126
4.11.	Intervalos de confianza bayesianos.....	128
4.12.	Contraste de hipótesis	130
4.13.	Hipótesis nula e hipótesis alternativa.....	130
4.14.	Errores en el contraste de hipótesis: error tipo 1 y error tipo 2.....	134
4.15.	Interpretación de valores <i>p</i> : significación estadística.....	136
4.16.	Significación estadística frente a significación práctica	138
4.17.	Pruebas a una cola y pruebas a dos colas.....	138
4.18.	Pruebas de contraste de hipótesis frente a intervalos de confianza.....	140
4.19.	Potencia estadística.....	141
4.20.	Estudios que demuestran equivalencia y estudios de no inferioridad	142
4.21.	Pruebas paramétricas y no paramétricas.....	143
4.22.	Resumen de las instrucciones en STATA, R, SPSS y Excel.....	143
5	Datos categóricos y porcentajes: comparación de proporciones	147
	<i>E. Toledo, J.M. Núñez-Córdoba, M.A. Martínez-González</i>	
5.1.	Test de χ^2 de Pearson para dos proporciones	147
5.2.	Test <i>z</i> para comparar dos proporciones	149
5.3.	Intervalo de confianza de la diferencia de dos proporciones	150
5.4.	Relación entre el intervalo de confianza y el valor <i>p</i>	151
5.5.	Ji cuadrado para comparar una proporción con una referencia externa (esperada): cálculo y su relación con la distribución binomial y sus aproximaciones.....	152
5.6.	Test exacto de Fisher.....	154

5.7. Test de McNemar para datos emparejados	158
5.8. Test de tendencia lineal para categorías ordenables lógicamente (variables ordinales)	161
5.9. <i>Odds ratio</i> en tablas 2×2	162
5.10. Error estándar e intervalo de confianza de la <i>odds ratio</i>	163
5.11. Otras medidas en tablas categóricas (tau de Kendall, gamma de Goodman y Kruskal)	164
5.12. Test para proporciones y tablas categóricas con STATA	165
5.13. Representación de intervalos de confianza para proporciones y <i>odds ratio</i> con STATA	167
5.14. Test para proporciones con otros programas	169
5.15. Resumen de las instrucciones en STATA y SPSS	173
6 Comparación de medias entre dos grupos	175
<i>E. Toledo, C. López del Burgo, C. Sayón-Orea, M.A. Martínez-González</i>	
6.1. Test de la <i>t</i> de Student para dos muestras independientes.....	175
6.2. Test para comparar varianzas	178
6.3. Test <i>t</i> para dos medias independientes con varianzas heterogéneas (test de Welch)	179
6.4. Intervalo de confianza para la diferencia de medias	181
6.5. Transformación logarítmica de la variable dependiente en un test <i>t</i> ... 182	
6.6. Test de la <i>t</i> de Student para comparar una media con un valor de referencia.....	185
6.7. Test de la <i>U</i> de Mann-Whitney.....	187
6.8. Test de la <i>t</i> de Student para datos emparejados (muestras relacionadas)	190
6.9. Test de Wilcoxon para datos emparejados.....	192
6.10. Test para medias con STATA.....	195
6.11. Test para medias con otros programas	197
7 Estimación del tamaño muestral	201
<i>M.A. Martínez-González, M. Ruiz-Canela, F. Guillén-Grima</i>	
7.1. Introducción.....	201
7.2. Margen de error.....	201
7.3. Estimación de una proporción.....	201
7.4. Estimación de una media.....	202
7.5. Comparación de dos proporciones.....	203
7.6. Comparación de dos medias	205
7.7. Cálculo de la potencia estadística.....	206
7.8. Curvas de potencia	206
7.9. Uso de STATA para estimar el tamaño muestral y la potencia.....	208
7.10. Programación de Excel para tamaño muestral y potencia.....	209
7.11. Otros programas disponibles para el cálculo del tamaño muestral... 210	
7.12. Resumen de las fórmulas del tamaño muestral.....	210
7.13. Resumen de las instrucciones en STATA	211
8 Comparaciones de <i>k</i> medias (tres o más grupos)	213
<i>M.A. Martínez-González, N. Martín-Calvo, J.B. Toledo</i>	
8.1. Introducción al ANOVA de una vía.....	213
8.2. Relación entre el ANOVA y la <i>t</i> de Student	215

8.3.	ANOVA de una vía con STATA	216
8.4.	Requisitos del ANOVA.....	217
8.5.	ANOVA de una vía con otros programas	218
8.6.	El ANOVA en manos de un lector sagaz de artículos científicos.....	221
8.7.	Test no paramétrico alternativo al ANOVA: Kruskal-Wallis	224
8.8.	Comparaciones múltiples: contrastes <i>a priori</i>	228
8.9.	Contrastes <i>a posteriori</i> (<i>post hoc</i>): correcciones por comparaciones múltiples	231
8.10.	Método de Benjamini-Hochberg basado en ordenar los valores <i>p</i> ...	236
8.11.	Resumen de las instrucciones en STATA y SPSS	238
9	ANOVA factorial, modelos lineales generalizados y ANCOVA	241
	<i>C. Sayón-Orea, E. Toledo, J.M. Núñez-Córdoba, M.A. Martínez-González</i>	
9.1.	Introducción.....	241
9.2.	ANOVA factorial (de dos vías o dos criterios).....	241
9.3.	ANOVA con medidas repetidas (comparación de <i>k</i> medias relacionadas).....	247
9.4.	Equivalente no paramétrico del ANOVA con medidas repetidas: test de Friedman.....	252
9.5.	Ajuste del ANOVA por variables continuas: ANCOVA	254
9.6.	Comparaciones intragrupo e intergrupos con medidas repetidas.....	254
9.7.	Análisis estadístico de ensayos cross-over.....	259
9.8.	ANOVA factorial y ANCOVA: estimación de medias ajustadas en STATA	262
9.9.	ANOVA factorial en SPSS.....	265
9.10.	Test de Friedman con SPSS y con STATA.....	266
9.11.	Resumen de las instrucciones en STATA y SPSS	267
10	Correlación y regresión lineal simple.....	269
	<i>A. Sánchez-Villegas, N. Martín-Calvo, M.A. Martínez-González</i>	
10.1.	Introducción.....	269
10.2.	Correlación.....	269
10.3.	Coefficiente de correlación con STATA.....	279
10.4.	Coefficiente de correlación con otros programas	281
10.5.	Regresión lineal simple	285
10.6.	Usos e interpretación de una regresión lineal.....	298
10.7.	Supuestos del modelo de regresión.....	300
10.8.	Representación gráfica de los residuales en una regresión lineal.....	301
10.9.	Construcción de un modelo de regresión lineal con STATA e instrucciones postestimación	303
10.10.	Regresión lineal con otros programas informáticos	312
10.11.	Relación entre regresión, ANOVA y <i>t</i> de Student	316
10.12.	Uso de la regresión para sustituir al ANOVA factorial.....	319
10.13.	Resumen de correlación y regresión lineal simple.....	323
10.14.	Resumen de las instrucciones en STATA y SPSS.....	324
11	Introducción al análisis de supervivencia.....	327
	<i>E. Toledo, F.J. Basterra-Gortari, M. García-López, M.A. Martínez-González</i>	
11.1.	Introducción.....	327
11.2.	Descripción de la supervivencia: método de Kaplan-Meier	329

11.3.	Pasos para realizar curvas de supervivencia de Kaplan-Meier	331
11.4.	Representación gráfica del estimador de Kaplan-Meier	332
11.5.	Intervalos de confianza para la estimación de supervivencia acumulada	334
11.6.	Análisis de supervivencia con STATA	335
11.7.	Análisis de supervivencia con otros programas	336
11.8.	Curvas de incidencia de Nelson-Aalen	338
11.9.	Comparación de curvas de supervivencia: test del <i>log-rank</i>	338
11.10.	Resumen de las instrucciones en STATA y SPSS	341
12	Introducción a los modelos multivariantes. Regresión lineal múltiple	343
	<i>A. Sánchez-Villegas, C. López del Burgo, M.A. Martínez-González</i>	
12.1.	Introducción	343
12.2.	Primera aproximación al modelo de regresión lineal múltiple	345
12.3.	Primera aproximación al modelo de regresión logística	347
12.4.	Primera aproximación al modelo de regresión de Cox	349
12.5.	Aspectos comunes y diferenciales de los modelos lineal, logístico y de Cox	352
12.6.	Regresión de Poisson	353
12.7.	Otros métodos multivariantes	356
12.8.	Hipótesis nulas en una regresión múltiple	357
12.9.	Interpretación condicional de los valores <i>p</i>	359
12.10.	Intervalos de confianza en la regresión múltiple	359
12.11.	Coefficiente de determinación R^2 y su versión ajustada	360
12.12.	Condiciones de aplicación del modelo de regresión múltiple. Análisis de residuales y verificación de supuestos	361
12.13.	Tolerancia, factor de inflación de varianza y multicolinealidad	364
12.14.	Variables que deben registrarse en una investigación	366
12.15.	Variables categóricas y variables indicadoras (<i>dummy</i>)	367
12.16.	Factores de confusión en la regresión múltiple	369
12.17.	Interacción (o modificación del efecto) en la regresión múltiple ...	373
12.18.	Relaciones no lineales, modelos polinómicos	377
12.19.	Construcción de un modelo de regresión múltiple	380
12.20.	Elección del mejor modelo	386
12.21.	Los métodos automáticos por pasos (stepwise) casi nunca están indicados	386
12.22.	Regresión lineal múltiple en otros programas distintos de STATA	386
12.23.	Resumen de las instrucciones en STATA y SPSS	390
12.23.	Resumen de las instrucciones en STATA y SPSS (<i>cont.</i>)	391
12.23.	Resumen de las instrucciones en STATA y SPSS (<i>cont.</i>)	392
12.23.	Resumen de las instrucciones en STATA y SPSS (<i>cont.</i>)	393
13	Regresión logística	397
	<i>A. Sánchez-Villegas, M. Bes-Rastrollo, M.A. Martínez-González</i>	
13.1.	Introducción	397
13.2.	Conceptos de <i>odds</i> , <i>odds ratio</i> y riesgo relativo	398
13.3.	Ejemplo ridículamente sencillo de regresión logística binaria univariante: interpretación	402

13.4.	Regresión logística binaria con variable independiente cuantitativa: interpretación	406
13.5.	Regresión logística binaria con una variable independiente con > 2 categorías: interpretación	409
13.6.	Regresión logística con múltiples variables independientes.....	412
13.7.	Control de la confusión en regresión logística. La confusión no tiene nada que ver con valores p	412
13.8.	Identificación de la interacción en regresión logística: test de razón de verosimilitud	415
13.9.	Supuesto de linealidad en el logit y uso de términos polinómicos ...	416
13.10.	Ajuste de un modelo multivariable de regresión logística	419
13.11.	Significación estadística en la regresión logística.....	422
13.12.	Test de Hosmer-Lemeshow: bondad de ajuste o calibración.....	424
13.13.	Curvas ROC: discriminación.....	426
13.14.	Criterios de construcción de modelos en regresión logística	428
13.15.	Regresión logística condicional	430
13.16.	Regresión logística en SPSS	431
13.17.	Resumen de las instrucciones en STATA y SPSS.....	433
14	Aspectos avanzados de regresión de Cox	437
	<i>M. Ruiz-Canela, E. Toledo, J. López-Fidalgo, M.A. Martínez-González</i>	
14.1.	Introducción: contexto y utilidad	437
14.2.	Conceptos de <i>hazard</i> y <i>hazard ratio</i>	437
14.3.	Ejemplo ridículamente sencillo de regresión de Cox univariante.....	438
14.4.	La ecuación de la regresión de Cox	439
14.5.	Interpretación de los coeficientes de la regresión de Cox univariante	440
14.6.	Comparación de curvas de supervivencia con la regresión de Cox...	441
14.7.	Regresión de Cox con variable independiente cuantitativa.....	441
14.8.	Interpretación de los coeficientes de variables independientes cuantitativas	443
14.9.	Regresión de Cox con una variable independiente con >2 categorías.....	443
14.10.	Interpretación de coeficientes de variables <i>dummy</i>	443
14.11.	Regresión de Cox con múltiples variables independientes	443
14.12.	Control de la confusión en la regresión de Cox.....	444
14.13.	Intervalos de confianza para la <i>hazard ratio</i> en el modelo de regresión de Cox	445
14.14.	Interacción (modificación del efecto) en regresión de Cox y test de razón de verosimilitud.....	446
14.15.	Interpretación del riesgo basal (<i>baseline hazard</i>)	446
14.16.	Regresión de Cox estratificada	446
14.17.	Tiempo de seguimiento en la regresión de Cox.....	448
14.18.	Regresión de Cox con covariables dependientes del tiempo.....	449
14.19.	Modelos de tiempos de fallo acelerados.....	449
14.20.	Relación entre <i>hazard ratio</i> y razón de densidades de incidencia....	450
14.21.	Similitudes y diferencias entre regresión de Cox y regresión logística.....	450

	14.22. Posibilidades y opciones de la regresión de Cox con STATA	451
	14.23. Resumen de las instrucciones en STATA, SPSS y R	453
15	Análisis de concordancia, validez y pronóstico	455
	<i>M.A. Martínez-González, E. Toledo, A. Sánchez-Villegas</i>	
	15.1. Conceptos y definiciones	455
	15.2. Consideraciones generales sobre estudios de validación de pruebas diagnósticas	455
	15.3. Consistencia interna en escalas cuantitativas: alfa de Cronbach.....	457
	15.4. Reproducibilidad: índice kappa de concordancia en variables cualitativas.....	459
	15.5. Coeficiente de corr elación intraclase: concordancia en variables cuantitativas.....	461
	15.6. Gráficos de Bland-Altman para acuerdo en variables cuantitativas...	464
	15.7. Coeficiente de correlación de concordancia de Lin.....	465
	15.8. Regresión de Passing-Bablok y regresión de Deming.....	466
	15.9. Gráficos de acuerdo-supervivencia	467
	15.10. Validez diagnóstica: sensibilidad, especificidad, valores predictivos, razones de verosimilitud	468
	15.11. Discriminación diagnóstica y pronóstica: curvas ROC.....	475
	15.12. Comparación de curvas ROC.....	478
	15.13. Índice C de Harrell para predicciones en análisis de supervivencia ...	478
	15.14. Índice neto de reclasificación, capacidad de estratificación y otros índices de discriminación.....	481
	15.15. Resumen de las instrucciones en STATA y SPSS.....	484
16	Análisis factorial.....	487
	<i>A. Sánchez-Villegas, M. Bes-Rastrollo, M.A. Martínez-González</i>	
	16.1. Introducción al análisis factorial	487
	16.2. Número de factores para extraer: gráfico de sedimentación	495
	16.3. Cálculos numéricos	497
	16.4. Sinonimias y equivalencias.....	499
	16.5. Condiciones de aplicación del análisis factorial de componentes principales (AFCP)	499
	16.6. Consideraciones sobre el tamaño muestral.....	502
	16.7. Rotación de los factores	502
	16.8. Refinamiento del análisis: eliminación de variables	504
	16.9. Análisis factorial común frente a análisis factorial de componentes principales.....	505
	16.10. Análisis factorial confirmatorio frente al exploratorio.....	507
	16.11. Diferente aproximación en STATA para realizar un análisis factorial de componentes principales	507
	16.12. Análisis factorial de componentes principales con SPSS.....	508
	16.13. Resumen de las instrucciones en STATA y SPSS.....	510
17	Análisis de clústeres o conglomerados	513
	<i>A. Sánchez-Villegas, F. Guillén-Grima, M.A. Martínez-González</i>	
	17.1. Introducción y concepto.....	513
	17.2. Tipos de análisis de clúster.....	513

17.3.	Método para la formación de conglomerados	514
17.4.	Gráficos del análisis de clúster: dendrogramas	520
17.5.	Estandarización y transformación de variables	521
17.6.	Requisitos para la aplicación de los métodos de análisis de clúster	523
17.7.	Clústeres de variables.....	523
17.8.	Ejemplo de análisis de clúster con STATA	523
17.9.	Análisis de clúster con SPSS.....	529
17.10.	Resumen de las instrucciones en STATA y SPSS.....	531
18	Métodos estadísticos en metaanálisis.....	533
	<i>M.A. Martínez-González, P.A. de la Rosa, A. Gea</i>	
18.1.	Revisiones sistemáticas y metaanálisis	533
18.2.	Tareas previas al análisis estadístico	533
18.3.	Escala aditiva o multiplicativa.....	533
18.4.	Efectos estandarizados: d de Cohen	533
18.5.	Método del inverso de la varianza: efectos fijos	534
18.6.	Gráficos de bosque (<i>forest plot</i>).....	538
18.7.	Test de heterogeneidad: estadístico Q.....	540
18.8.	Tau cuadrado: varianza entre estudios.....	541
18.9.	Índice I cuadrado.....	541
18.10.	Gráfico de L'Abbé para heterogeneidad.....	543
18.11.	Metaanálisis de efectos aleatorios: método de DerSimonian-Laird.....	544
18.12.	Análisis de subgrupos	545
18.13.	Metarregresión.....	545
18.14.	Sesgo de publicación: gráfico de embudo (<i>funnel plot</i>)	545
18.15.	Sesgo de publicación: test de Egger.....	547
18.16.	Sesgo de publicación: métodos de MacAskil y de Peters.....	547
18.17.	Sesgo de publicación: otros métodos.....	548
18.18.	Metaanálisis acumulado.....	549
18.19.	Uso de STATA para el metaanálisis.....	549
19	Otros métodos bioestadísticos.....	553
	<i>M.A. Martínez-González, P.A. de la Rosa, A. Gea</i>	
19.1.	Métodos de remuestreo: <i>bootstrap</i> , <i>jackknife</i>	553
19.2.	Método de captura-recaptura para indagar el tamaño de una población	555
19.3.	Análisis de decisiones.....	557
19.4.	Modelos flexibles de regresión con intervalos de confianza (splines).....	561
19.5.	Valores perdidos (missing) y métodos de imputación.....	565
19.6.	Ponderación por el inverso de la varianza y modelos estructurales marginales	570
19.7.	Índices de propensión (<i>propensity scores</i>).....	575
19.8.	Ecuaciones de estimación generalizadas (<i>generalized estimating equations, GEE</i>)	576