

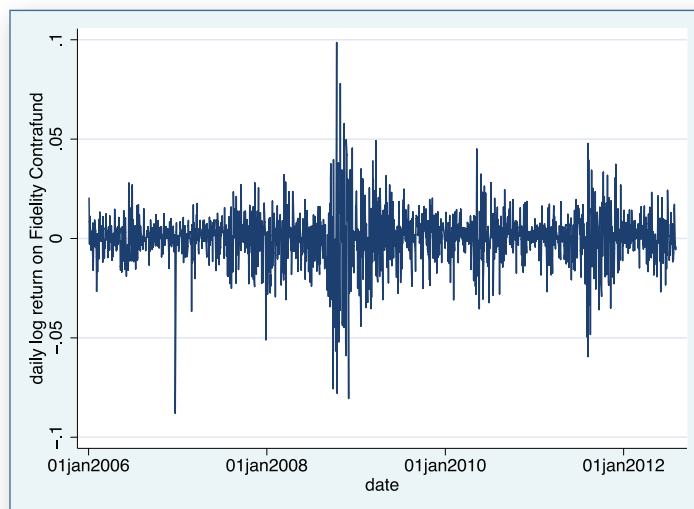
Las Noticias de STATA

Statistics Graphics Data Management & Analysis

En el centro de atención: mgarch

Es posible que usted esté familiarizado con el comando de Stata **arch**, que ajusta modelos de volatilidad univariantes, también conocidos como modelos autorregresivos con heteroscedasticidad condicional generalizados (GARCH). En el marco de los modelos GARCH, se supone que la varianza condicional de una serie es una función de la volatilidad observada hasta el momento. Los modelos GARCH se utilizan comúnmente para modelar los dividendos de las acciones en la bolsa de valores.

A la derecha se muestra un gráfico de los logaritmos de los dividendos de un fondo mutuo. Podemos ver que los dividendos presentan una volatilidad significativa y que los períodos de alta y baja volatilidad tienden a estar conglomerados. Las técnicas para modelar la volatilidad de una serie son ampliamente utilizadas para modelar las variables asociadas a la compra y venta de acciones, el manejo de riesgos y la asignación de activos.



Los modelos GARCH multivariantes (MGARCH) generalizan los modelos GARCH univariantes y permiten incorporar relaciones entre los procesos de volatilidad de varias series. Queremos saber cómo los cambios en la volatilidad de una acción afectan a la volatilidad de otra acción. Estas relaciones se pueden parametrizar de varias maneras. El comando de Stata **mgarch** implementa cuatro parametrizaciones comúnmente utilizadas: el modelo vech diagonal (**mgarch dvech**), el modelo de correlación condicional constante (**mgarch ccc**), el modelo de correlación condicional dinámica (**mgarch dcc**), y el modelo de correlación condicional variable (**mgarch vcc**).

Ejemplo

Utilizaré los datos de logaritmos de dividendos diarios de tres fondos mutuales Fidelity: Intermediate Bond Fund (**bond**), Contrafund (**contra**), y Blue Chip Growth Fund (**bchip**). Usted puede utilizar el comando de Stata **arch** para analizar cada serie por separado. Por ejemplo, para ajustar un modelo GARCH (1,1) a la serie de bonos, escriba:

```
. arch bond, noconstant arch(1) garch(1)
```

Como alternativa, puede usar cualquiera de los modelos MGARCH de correlaciones condicionales. Aquí, he utilizado **mgarch dcc**:

```
. mgarch dcc (bond =, noconstant arch(1) garch(1))
```

Observe lo fácil que es convertir la sintaxis **arch** a la sintaxis **mgarch**. He incluido toda la ecuación entre paréntesis y añadido el signo = después de la variable dependiente. Ahora usted puede adivinar cómo ajustar un modelo GARCH multivariante sin siquiera mirar el archivo de ayuda de **mgarch dcc**:

```
. mgarch dcc (bond =, noconstant arch(1) garch(1))
              (contra =, noconstant arch(1) garch(1))
              (bchip =, noconstant arch(1) garch(1))
```

Novedades de Stata Press4

Novedades en la Librería de Stata.....5

Capacitación.....8

Conferencias y Reuniones de Grupos de Usuarios.....9

En el centro de atención: Curvas de Característica Operativa del Receptor 10

Nuevo curso abierto de capacitación: Structural Equation Modeling Using Stata..... 12

Las Noticias de Stata

Editora ejecutiva:..... Karen Strope

Supervisora de Producción:..... Annette Fett

Editora de la versión en español:.... Isabel Cañete



Sin embargo, todas las ecuaciones tienen los mismos términos ARCH y GARCH, por lo que puedo reescribir la sintaxis sacando hacia afuera los términos comunes:

```
. mgarch dcc (bond =, noconstant)
      (contra =, noconstant)
      (bchip =, noconstant) , arch(1) garch(1)
```

También puedo combinar los términos de las ecuaciones dentro de los mismos paréntesis:

```
. mgarch dcc (bond contra bchip =, noconstant) , arch(1) garch(1)
```

Ahora vemos por qué se requiere el signo = : separa las variables dependientes de las variables independientes. La última sintaxis es la más corta, pero la primera es la más flexible, ya que permite incorporar especificaciones diferentes para las medias y las varianzas de las tres ecuaciones.

Cualquiera de las tres opciones producirá las siguientes estimaciones:

Dynamic conditional correlation MGARCH model

Sample: 1 - 1668
Distribution: Gaussian
Log likelihood = 19677.94

Number of obs = 1668
Wald chi2(.) = .
Prob > chi2 = .

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ARCH_bond						
arch L1.	.0455378	.0085325	5.34	0.000	.0288144	.0622612
garch L1.	.9553776	.008344	114.50	0.000	.9390237	.9717315
_cons	1.40e-08	1.04e-08	1.35	0.176	-6.28e-09	3.43e-08
ARCH_contra						
arch L1.	.069363	.0083893	8.27	0.000	.0529202	.0858057
garch L1.	.8582225	.0148355	57.85	0.000	.8291454	.8872996
_cons	9.00e-06	1.13e-06	7.93	0.000	6.77e-06	.0000112
ARCH_bchip						
arch L1.	.069969	.0081638	8.57	0.000	.0539682	.0859697
garch L1.	.8836878	.011061	79.89	0.000	.8620087	.9053669
_cons	7.51e-06	1.00e-06	7.49	0.000	5.54e-06	9.47e-06
Correlation						
bond						
contra	-.149995	.0498875	-3.01	0.003	-.2477727	-.0522174
bchip	-.180069	.0504977	-3.57	0.000	-.2790426	-.0810954
contra						
bchip	.8697618	.0117909	73.77	0.000	.846652	.8928716
Adjustment						
lambda1	.1878626	.0146178	12.85	0.000	.1592121	.216513
lambda2	.7011739	.0237479	29.53	0.000	.6546289	.7477189

Ready CAP NUM OVR

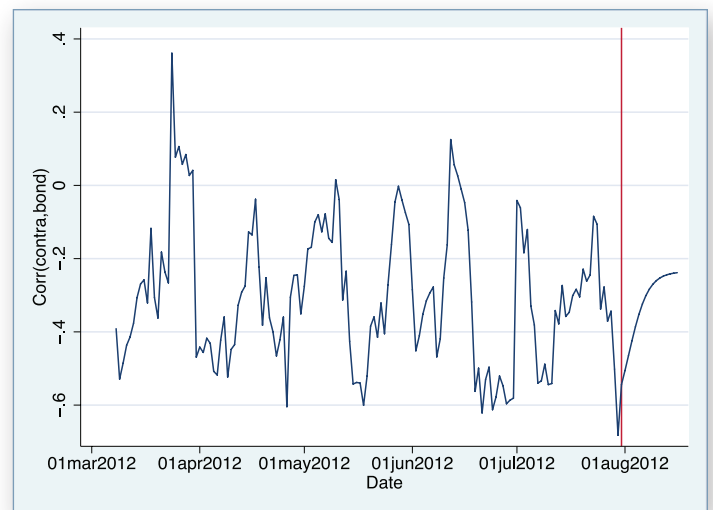
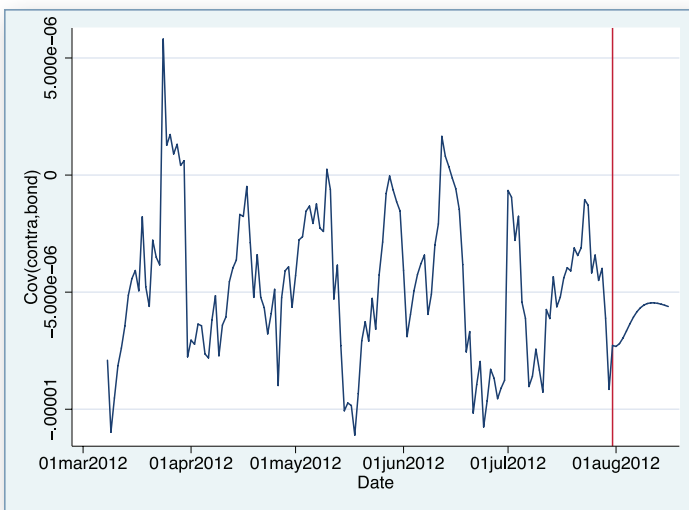
La interpretación de los parámetros de varianza es la mismo que para un modelo GARCH univariante. La parte nueva son los parámetros estimados que miden la dinámica entre las volatilidades. Las quasicorrelaciones condicionales estimadas entre las volatilidades de los dos fondos mutuales de acciones, **contra** y **bchip**, es alta y positiva. Esto significa que una volatilidad alta en Contrafund se asocia a una volatilidad alta en Blue Chip y vice versa.

Las quasicorrelaciones condicionales estimadas entre la volatilidad del fondo mutual de bonos y las volatilidades de los fondos mutuales de acciones son bajas y negativas. Esto significa que un aumento en la volatilidad en el fondo mutual de bonos se asocia con cierta disminución en la volatilidad de los fondos mutuales de acciones. Esta información será de utilidad si usted desea minimizar la volatilidad de un portafolio.

Una vez estimado el modelo, puede estar interesado en proyectar la serie o sus volatilidades. A continuación mostraré cómo obtener predicciones de las varianzas y covarianzas condicionales para toda la serie. En primer lugar, utilizaré **tsappend** para extender los datos, y luego usaré **predict** para obtener las predicciones:

```
. tsappend, add(14)
. predict H*, variance dynamic(td(31jul2012))
```

Las predicciones condicionales hasta el 31 de julio de 2012 son efectuadas dentro de la muestra (in-sample) y un paso adelante. Luego de esa fecha, las predicciones serán fuera de la muestra y dinámicas. A continuación, en el lado izquierdo se presenta un gráfico que muestra la predicción de la covarianza condicional entre Contrafund y el fondo de bonos. Las proyecciones dinámicas se representan a la derecha de la línea vertical. Puede ser más fácil interpretar la relación entre las volatilidades en una escala de correlación; ésta se muestra en el gráfico de la derecha.



Las gráficas muestran que la correlación condicional entre los dos volatilidades varía con el tiempo. Este comportamiento es una propiedad de los modelos MGARCH con correlaciones condicionales dinámicas y variables en el tiempo. En un modelo MGARCH con correlación condicional constante, las correlaciones no variarían en el tiempo y la curva sería una línea horizontal.

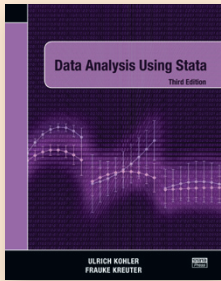
Resumen

Los modelos GARCH multivariantes se utilizan para analizar las relaciones dinámicas entre los procesos de volatilidad de varias series. El comando de Stata **mgarch** proporciona un fácil acceso a algunas de las parametrizaciones de uso común.

— Rafal Raciborski
Econometrista

Novedades de Stata Press

Data Analysis Using Stata, Third Edition



Autores: Ulrich Kohler y Frauke Kreuter
 Editorial: Stata Press
 Derechos de autor: 2012
 ISBN-13: 978-1-59718-110-5
 Páginas: 497; cubierta suave
 Precio: USD 56,00

Data Analysis Using Stata, Third Edition, ha sido completamente renovado para reflejar las capacidades de Stata 12. Este libro será de utilidad tanto a aquellos que están aprendiendo estadística como a los usuarios de otros paquetes estadísticos que desean comenzar a usar Stata. A lo largo del libro, Kohler y Kreuter muestran ejemplos utilizando datos del German Socioeconomic Panel, una encuesta de hogares que contiene información demográfica, de ingresos, de empleo y otra información clave.

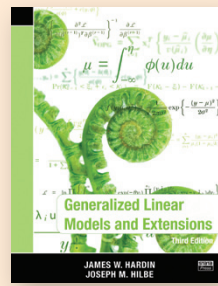
Kohler y Kreuter adoptan un enfoque aplicado, mostrando en primer lugar cómo utilizar la Interfaz Gráfica de Stata y luego describiendo la sintaxis de Stata. El núcleo del libro cubre diversos aspectos de la investigación en ciencias sociales, incluyendo el manejo de los datos, la elaboración de tablas y gráficas, el análisis de regresión lineal, y los modelos logísticos. Los autores describen el manejo de covariables categóricas en Stata y muestran cómo los nuevos comandos **margins** y **marginsplot** simplifican enormemente la interpretación de los resultados de la regresión lineal y logística. Un nuevo capítulo discute aspectos de la inferencia estadística, incluyendo las muestras aleatorias, las muestras de encuestas complejas, la no respuesta y la inferencia causal.

El resto del libro incluye capítulos sobre la lectura de archivos de texto, escritura de programas y de archivos .ado, y el uso de recursos a través de Internet, tales como el comando **search** y el archivo SSC.

Data Analysis Using Stata, Third Edition se ha estructurado de modo que pueda ser utilizado como un curso de estudios individuales o como un libro de texto en un curso introductorio de análisis datos o de estadística. Será de interés para los estudiantes y los investigadores académicos en todas las ciencias sociales.

Para ver tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite stata-press.com/books/data-analysis-using-stata.

Generalized Linear Models and Extensions, Third Edition



Autores: James W. Hardin y Joseph M. Hilbe
 Editorial: Stata Press
 Derechos de Autor: 2012
 ISBN-13: 978-1-59718-105-1
 Páginas: 455; cubierta suave
 Precio: USD 58,00

Los modelos lineales generalizados (MLG) extienden los modelos de regresión lineal a modelos con variables dependientes no gaussianas, e incluso discretas. La teoría de MLG se basa en la familia exponencial de distribuciones, una clase que incluye los conocidos modelos logit, probit y de Poisson. Si bien se puede ajustar estos modelos en Stata utilizando comandos especializados (por ejemplo, **logit** para modelos logit), utilizar **glm**, el comando de Stata para MLG, ofrece algunas ventajas. Por ejemplo, los diagnósticos de los modelos se pueden calcular e interpretar de manera similar, independientemente de la distribución supuesta.

Este texto cubre con detalle los distintos aspectos de los MLG, tanto desde el punto de vista teórico como computacional, con énfasis en Stata. La teoría consiste en mostrar cómo los diferentes MLG son casos particulares de la familia exponencial, mostrando propiedades generales de esta familia de distribuciones, así como la derivación de los estimadores por máxima verosimilitud (MV) y de los estimadores de los errores estándar. Hardin y Hilbe muestran cómo el método de los mínimos cuadrados ponderados iterativamente, es una consecuencia de la estimación MV mediante el uso del algoritmo “scoring” de Fisher.

Los autores también discuten los diferentes métodos de estimación de errores estándar, incluyendo métodos robustos, métodos robustos con conglomerado, Newey–West, producto exterior del gradiente, bootstrap, y jackknife. La cobertura completa que ofrecen sobre diagnósticos del modelo incluye medidas de influencia tales como la distancia de Cook, varias formas de residuos, los criterios de información de Akaike y Bayesiano, y varias medidas de la variabilidad explicada, del tipo del R^2 .

Después de presentar la teoría general, Hardin y Hilbe se detienen en cada una de las distribuciones. Cada distribución tiene su propio capítulo que explica los detalles de cálculo de la aplicación de la teoría general a esa distribución particular. Aquí el pseudocódigo juega un papel importante, ya que permite a los autores describir algoritmos computacionales de una forma relativamente

sencilla. Dedicar un capítulo a cada distribución (o familia, en términos de MLG) permite también incluir ejemplos con datos reales que muestran cómo se ajustan dichos modelos en Stata, y también presentar ciertos diagnósticos y estrategias analíticas que son únicos para esa familia. Los capítulos sobre datos binarios y datos de conteo (Poisson) son excelentes en este sentido. Hardin y Hilbe prestan gran atención a los problemas de sobredispersión e inflación en cero en los modelos de datos de conteo.

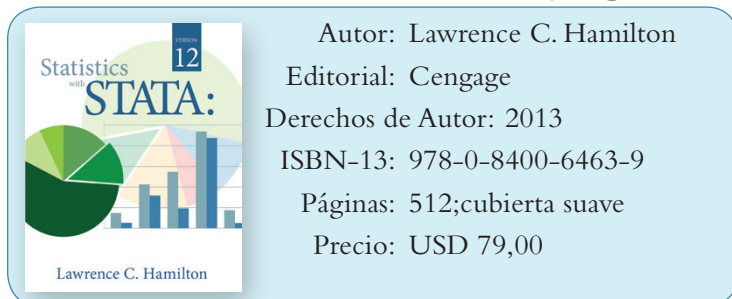
La parte final del texto se refiere a tres tipos de extensiones de los MLG. En primer lugar, los autores cubren las respuestas multinomiales, ambas ordinales y nominales. Aunque las respuestas multinomiales no son estrictamente parte de los MLG, la teoría es similar, ya que se puede pensar en una respuesta multinomial como una extensión de una respuesta binaria. Los ejemplos que se presentan en estos capítulos suelen utilizar programas escritos por los propios autores en Stata, aumentando la capacidad oficial de Stata. En segundo lugar, los MLG se pueden extender a los datos agrupados a través de ecuaciones

de estimación generalizadas (EEG); un capítulo cubre la teoría de EEG y ejemplos. Finalmente, los MLG pueden ser extendidos por los usuarios, quienes pueden programar su propia familia de funciones de enlace para usar con el comando oficial **glm** de Stata, y los autores detallan este proceso. Además de otras adiciones, (por ejemplo, una nueva sección sobre efectos marginales), la tercera edición contiene varios nuevos modelos lineales generalizados extendidos, proporcionando a los usuarios de Stata nuevas formas de captar la complejidad de los datos de conteo. Los nuevos modelos incluyen un modelo binomial negativo de tres parámetros conocido como NB-P, el modelo de distribución de Poisson inversa gaussiana (PIG), el modelo cero inflado generalizado de Poisson (ZIGP), una versión reescrita del modelo generalizado de Poisson, modelos de mezcla de dos y de tres componentes finitas, y un modelo generalizado censurado de Poisson y binomial negativo. Esta edición tiene un nuevo capítulo sobre la síntesis y simulación de datos, pero también muestra cómo construir una amplia variedad de modelos sintéticos y de Monte Carlo a lo largo del libro.

Para ver la tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite stata-press.com/books/generalized-linear-models-and-extensions.

Novedades de La Librería de Stata

Statistics with Stata: Version 12, Eighth Edition



Autor: Lawrence C. Hamilton
 Editorial: Cengage
 Derechos de Autor: 2013
 ISBN-13: 978-0-8400-6463-9
 Páginas: 512; cubierta suave
 Precio: USD 79,00

Statistics with Stata: Version 12 es la edición más reciente de la popular serie *Statistics with Stata*, escrita por el Profesor Lawrence C. Hamilton. Creada con el propósito de llenar el hueco entre los textos de estadística y la documentación de Stata, *Statistics with Stata* muestra cómo usar Stata para realizar una variedad de tareas.

Los tres primeros capítulos cubren cómo iniciarse en el uso de Stata, la manipulación de datos y los gráficos. Hamilton luego introduce una variedad de procedimientos estadísticos disponibles dentro Stata. Éstos incluyen estadísticas de resumen y tablas, ANOVA, regresión lineal (y diagnóstico), métodos robustos de regresión lineal, modelos de regresión para variables dependientes limitadas,

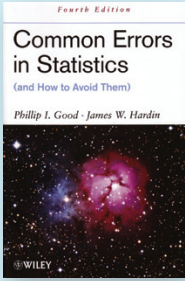
datos de encuestas complejas, análisis de supervivencia, análisis factorial, análisis de conglomerado, modelos de ecuaciones estructurales, imputación múltiple, series de tiempo, y modelos multinivel de efectos mixtos. El capítulo final proporciona una introducción a la programación.

La organización de este libro lo hace ideal para aquellos que se inician en la estadística, para aquellos que desean migrar de otros paquetes estadísticos a Stata, y para los usuarios experimentados en Stata que desean explorar las funcionalidades de Stata con que aún no están familiarizados.

Al comienzo de cada capítulo, Hamilton incluye una sección con ejemplos, donde muestra la sintaxis de Stata para los temas tratados en ese capítulo. Para aquellos que ya están familiarizados con las técnicas estadísticas, pero no con los correspondientes comandos de Stata, esta sección de ejemplos puede ser suficiente para comenzar un análisis utilizando Stata. Luego de las secciones de ejemplos, Hamilton aborda cada tema con más detalle, incluyendo la descripción de los procedimientos estadísticos, ejemplos con datos reales, y la interpretación de la salida de Stata.

Para ver la tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite stata.com/bookstore/statistics-with-stata.

Common Errors in Statistics (and How to Avoid Them), Fourth Edition



Autores: Phillip I. Good y James W. Hardin
 Editorial: Wiley
 Derechos de Autor: 2012
 ISBN-13: 978-1-118-29439-0
 Páginas: 336; cubierta suave
 Precio: USD 49,75

Common Errors in Statistics (and How to Avoid Them), Fourth Edition, escrito por Phillip I. Good y James W. Hardin, contiene una gran cantidad de consejos sobre cómo mejorar el diseño experimental, elaborar cuadros y gráficos informativos, y analizar datos en forma eficiente. Este libro no es un tratado sobre la teoría estadística. Más bien, proporciona información sobre cómo aplicar mejor esa teoría a problemas reales y obtener resultados informativos. Como el título lo indica, el libro ofrece muchos ejemplos de análisis mal ejecutados y luego explica detalladamente cómo estos ejemplos se pueden mejorar.

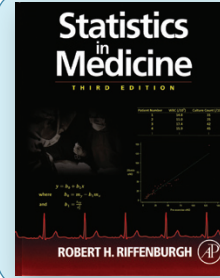
Los autores comienzan por discutir algunos temas fundamentales del análisis estadístico, como las fuentes de error, la recopilación de datos y la formulación de hipótesis. El capítulo 2, sobre las hipótesis, ha sido completamente reescrito y ahora hace hincapié en la importancia de la formulación de una hipótesis nula y todas las alternativas, incluyendo una discusión de las posibles conclusiones. Este capítulo también discute las tradicionales pruebas de Neyman–Pearson, así como la toma de decisiones.

La segunda parte del libro se centra en las pruebas de hipótesis y la estimación de parámetros. Aquí, los autores examinan la evaluación estadística de los datos, así como las fortalezas y limitaciones de varios procedimientos estadísticos. El capítulo 8, sobre cómo informar de los resultados, se ha actualizado para reflejar las fortalezas y debilidades de los p -valores y los intervalos de confianza, así como para mostrar la importante distinción entre la significación estadística y la importancia práctica de los resultados. En el capítulo 10 se explica cómo hacer gráficos explicativos y se incluye una lista de 11 reglas útiles a seguir.

La última parte del libro muestra cómo construir un modelo, incluyendo regresión lineal y no lineal, regresión de cuantiles, modelos de conteo, y modelos para datos de panel (longitudinales). El capítulo final discute la validación del modelo. Esta exposición aplicada sobre errores comunes en materia de estadísticas (y cómo evitarlos), será de utilidad para los profesionales experimentados, y sus muchos ejemplos y explicaciones cuidadosas hacen del libro un texto complementario útil para los estudiantes de estadística.

Para ver la tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite stata.com/bookstore/common-errors-statistics.

Statistics in Medicine, Third Edition



Autor: Robert H. Riffenburgh
 Editorial: Academic Press
 Derechos de Autor: 2012
 ISBN-13: 978-0-12-384864-2
 Páginas: 690; cubierta rígida
 Precio: SD 64,25

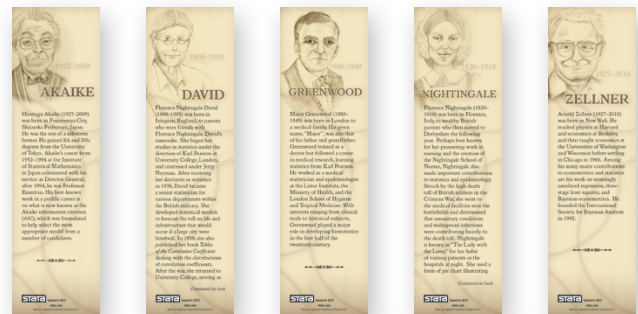
Statistics in Medicine, Third Edition, por Robert H. Riffenburgh, es un libro excelente, útil como referencia para los investigadores de las ciencias médicas y también como libro de texto. Se centra principalmente en la comprensión de los conceptos estadísticos y no en los fundamentos matemáticos y teóricos. Riffenburgh cubre tanto técnicas introductorias como métodos estadísticos avanzados que aparecen comúnmente en las publicaciones médicas.

Riffenburgh comienza con una discusión acerca de la planificación de los estudios y la redacción de artículos para informar de los resultados. A continuación, introduce los métodos estadísticos que normalmente se cubrirían en un curso de bioestadística. Entre éstos, se incluye estadísticas de resumen, distribuciones, tablas de dos vías, intervalos de confianza y pruebas de hipótesis. Además, ofrece una visión general de una variedad de técnicas estadísticas más sofisticadas como regresiones avanzadas, análisis de supervivencia, pruebas de equivalencia, análisis Bayesiano y análisis de series de tiempo.

Para ver tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite stata.com/bookstore/statistics-medicine.

Novedades de la Tienda de Regalos de Stata

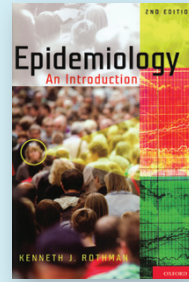
Marcadores: Serie 4



USD 2.25 en Norte América (el precio incluye envío)
 USD 3.50 en el resto del mundo (el precio incluye envío)

Epidemiology: An Introduction, Second Edition

El estudio de la epidemiología requiere un conocimiento profundo de la estadística. Muchos textos se centran en fórmulas, cálculos y software más que en los conceptos que usted necesita saber. Es fácil perderse en la jerga. Este texto es ideal para que los principiantes en la estadística epidemiológica puedan aprender la terminología y entender cómo y cuándo utilizar las herramientas estadísticas. El texto se centra en conceptos, no en fórmulas matemáticas, y examina las técnicas estadísticas en el contexto de los problemas reales que éstas pueden resolver. Este texto cierra la brecha entre lo que se enseña en un texto de introducción a la estadística y lo que se necesita para ser un investigador y analista eficaz.



Autor: Kenneth J. Rothman
 Editorial: Oxford University Press
 Derechos de Autor: 2012
 ISBN-13: 978-0-19-975455-7
 Páginas: 268; cubierta suave
 Precio: USD 33,50

La segunda edición cuenta con dos nuevos capítulos, el capítulo 2 y el capítulo 6. El capítulo 2 ofrece una visión histórica de las primeras contribuciones de los pioneros de la epidemiología y la salud pública. El capítulo 6 presenta un resumen de los conceptos clave utilizados en la epidemiología de las enfermedades infecciosas. Este texto se mantiene al día con la constante evolución de los conceptos epidemiológicos, y ahora viene acompañado por un sitio web (oup.com/us/epi), donde los lectores pueden participar en las discusiones sobre los conceptos presentados en el libro.

Para ver la tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite stata.com/bookstore/epidemiology-introduction.

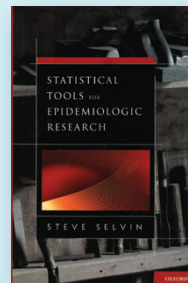
Statistical Tools for Epidemiologic Research

Se trata de un nuevo libro del autor del aclamado *Statistical Analysis of Epidemiologic Data*.

Está pensado como un segundo curso sobre métodos estadísticos. El cálculo se utiliza de vez en cuando, pero la mayor parte del texto requiere sólo un conocimiento de las matemáticas elementales. Los conjuntos de datos en Stata y las salidas están disponibles en el sitio web del libro.

El libro está dedicado principalmente a modelos de regresión, con capítulos separados sobre la regresión logística, la regresión de Poisson, y la regresión logística condicional. Los capítulos enfatizan los conceptos de interacción y de confusión, y las conexiones con las tablas de dos por dos. También contiene algunos capítulos sobre temas a menudo no cubiertos en este nivel, incluyendo errores de clasificación, análisis de datos longitudinales, y suavizado.

Para ver la tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite stata.com/bookstore/statistical-tools-for-epidemiologic-research.



Autor: Steve Selvin
 Editorial: Oxford University Press
 Derechos de Autor: 2011
 ISBN-13: 978-0-19-975596-7
 Páginas: 494; cubierta rígida
 Precio: USD 56,00

Visítenos en APHA 2012

San Francisco, California, 27 al 31 de octubre

La reunión anual 2012 de la American Public Health Association (APHA) se llevará a cabo en San Francisco, California, del 27 al 31 de octubre. Para obtener más información, visite apha.org/meetings/AnnualMeeting.

Deténgase en el stand No 2000 para hablar con las personas que desarrollan y apoyan el software Stata.



Cursos abiertos de Capacitación

Curso	Fechas	Lugar	Costo
Multilevel/Mixed Models Using Stata	4 y 5 de octubre de 2012	Washington, DC	USD 1.295
Panel-Data Analysis Using Stata	16 y 17 de octubre de 2012	Washington, DC	USD 1.295
Structural Equation Modeling Using Stata NUEVO	25 y 26 de octubre de 2012	Washington, DC	USD 1.295
Survey Data Analysis Using Stata	23 y 24 de octubre de 2012	Washington, DC	USD 1.295
Using Stata Effectively: Data Management, Analysis, and Graphics Fundamentals	2 y 3 de octubre de 2012 1 y 2 de noviembre de 2012	Washington, DC San Francisco, CA	USD 950

Multilevel/Mixed Models Using Stata

Este curso es una introducción al uso de Stata para ajustar modelos mixtos/multinivel. Los modelos mixtos pueden contener más de un nivel de efectos aleatorios anidados, por lo que también se los conoce como modelos multinivel o jerárquicos, en particular en las ciencias sociales. El curso es interactivo, con datos reales, y ofrece muchas oportunidades para hacer preguntas específicas de investigación y provee ejercicios donde se aplican los conceptos del curso.

Panel-Data Analysis Using Stata

Este curso ofrece una introducción a la teoría y práctica del análisis de datos de panel. Luego de la introducción de los enfoques de efectos fijos y efectos aleatorios para modelar heterogeneidad no observada a nivel de individuos, el curso abarca modelos lineales con covariables exógenas, modelos lineales con covariables endógenas, modelos lineales dinámicos y algunos modelos no lineales. Incluye también una introducción al método generalizado de momentos. Las clases se complementan con ejercicios y ejemplos en Stata.

Structural Equation Modeling Using Stata

NUEVO

Este curso cubre el uso de Stata para ajustar modelos de ecuaciones estructurales (SEM). El curso presenta una gran variedad de modelos, incluyendo el análisis de trayectoria, el análisis factorial confirmatorio, modelos completos de ecuaciones estructurales, las curvas de crecimiento latente, y más. Los ejemplos muestran el uso del comando **sem** así como el SEM-Builder: una interfaz gráfica para generar diagramas que se pueden utilizar para ilustrar, especificar, y estimar los modelos SEM.

Survey Data Analysis Using Stata

Este curso cubre el uso de Stata para el análisis de datos provenientes de encuestas, asumiendo una población fija. El curso cubre los métodos de muestreo utilizados para recoger datos de encuestas y cómo afectan la estimación de los totales, las razones, y los coeficientes de regresión, así como varios estimadores de varianza implementados en Stata para la estimación con datos de encuestas, incluyendo linealización, replicaciones repetidas balanceadas y jackknife. Cada tema se ilustra con uno o más ejemplos en Stata.

Using Stata Effectively: Data Management, Analysis, and Graphics Fundamentals

Este curso de dos días, dictado por Bill Rising (Director de Servicios Educativos de StataCorp), está diseñado para familiarizarlo con los tres componentes de Stata: manejo de datos, análisis, y gráficos. El curso está pensado tanto para usuarios nuevos de Stata, como para aquellos usuarios que, si bien ya conocen Stata, quieren aumentar su eficiencia en el uso de Stata. Al terminar el curso, usted estará preparado para utilizar eficientemente Stata y de manera reproducible, lo que simplificará significativamente la realización de trabajos en colaboración, o modificaciones posteriores de sus propios análisis.

Ofrecemos un descuento del 15% en la matrícula para grupos de tres o más participantes. Por detalles, escríbanos a la dirección training@stata.com. Para obtener más información sobre los cursos o para inscribirse, visite stata.com/public-training.

Próximas fechas de Netcourse®

Para inscribirse, visite stata.com/netcourse.

NetCourse 101, Introduction to Stata

Un curso introductorio de 6 semanas que enseña a usar Stata en forma interactiva. Cubre varias técnicas detalladas y trucos para convertirlo en un usuario eficiente de Stata.

Fechas: 19 de octubre al 30 de noviembre de 2012

Fecha límite para inscripciones: 18 de octubre de 2012

Precio: USD 95

stata.com/netcourse/intro-nc101

NetCourse 151, Introduction to Stata Programming

Este curso introduce al usuario en lo que la mayoría de quienes usan paquetes estadísticos llaman programación, es decir, la cuidadosa realización de análisis reproducibles.

Fechas: 19 de octubre al 30 de noviembre de 2012

Fecha límite para inscripciones: 18 de octubre de 2012

Precio: USD 125

stata.com/netcourse/programming-intro-nc151

NetCourse 152, Advanced Stata Programming

Este curso le enseñará principalmente a crear nuevos comandos, que resulten indistinguibles de los comandos oficiales de Stata. El curso asume que usted ya entiende en qué casos es necesario programar y por qué, y, hasta cierto punto, cómo hacerlo. En este curso aprenderá a utilizar el comando **syntax** para analizar expresiones (“parsing”), tanto para las expresiones que siguen las reglas de sintaxis estándar en los comandos de Stata, como para aquellas que no las siguen. Aprenderá a manejar y procesar resultados, procesar datos por grupo, y más.

Fechas: 12 de octubre al 30 de noviembre de 2012

Fecha límite para inscripciones: 11 de octubre de 2012

Precio: USD 150

stata.com/netcourse/programming-advanced-nc152

NetCourse 461, Introduction to Univariate Time Series with Stata

Este curso introduce el análisis de datos de series de tiempo univariantes, con énfasis en los aspectos prácticos más requeridos por profesionales e investigadores aplicados. El curso será de interés para una amplia gama de usuarios, incluyendo economistas, analistas financieros, gerentes, y todo aquel que necesite trabajar con series de tiempo.

Fechas: 12 de octubre al 30 de noviembre de 2012

Fecha límite para inscripciones: 11 de octubre de 2012

Precio: USD 295

stata.com/netcourse/univariate-time-series-intro-nc461

NetCourseNow

¿Prefiere elegir el momento y establecer el ritmo de un NetCourse? ¿Quiere tener un instructor personal de NetCourse? Este es el curso para usted.

stata.com/netcourse/ncnow.

Reserve la fecha: Conferencia de Stata



NUEVA ORLEANS
18 AL 19 DE JULIO DE 2013

Lugar: Hyatt French Quarter New Orleans
800 Iberville Street
New Orleans, Louisiana 70112
frenchquarter.hyatt.com

Presidente R. Carter Hill
del comite: Louisiana State University



stata.com/new-orleans13

Reunión 2012 del Grupo de Usuarios de Stata en Polonia

Fecha: 19 de octubre de 2012

Lugar: Faculty of Economic Sciences
Warsaw University
ul. Długa 44/50
PL-00241 Warsaw

Costo: 60 PLN profesionales; 30 PLN estudiantes

Detalles: stata.com/meeting/poland12

En la reunión 2012 del Grupo de usuarios de Stata en Polonia, los usuarios tendrán la oportunidad de compartir sus logros científicos y los resultados de su investigación empírica usando Stata. Además, habrá una conferencia magistral dirigida tanto a usuarios experimentados de Stata como a quienes no utilizan Stata.

Organizadores

Organizadores científicos

- Leszek Morawski (*Katedra Statystyki i Ekonometrii*)
- Jerzy Mycielski (*Katedra Statystyki i Ekonometrii*)
- Paweł Strawiński (*Katedra Statystyki i Ekonometrii*)

Organizadores de logística

- Timberlake Consultants Ltd.
biuro@timberlake.pl

Por información adicional, visite stata.com/meeting/poland12.

En el centro de atención: Curvas de característica operativa del receptor

Supongamos que un médico está tratando de decidir si se debe diagnosticar a una persona con cáncer de páncreas. Existen varias pruebas de diagnóstico que clasifican a un individuo como que padece cáncer (caso) o no (control) sobre la base de uno de los atributos biológicos del individuo. El atributo medido es la variable de clasificación para la prueba. Los niveles individuales del marcador tumoral CA 19-9 y la proteína CA-125 son ejemplos de variables de clasificación que pueden ser utilizadas para diagnosticar el cáncer de páncreas.

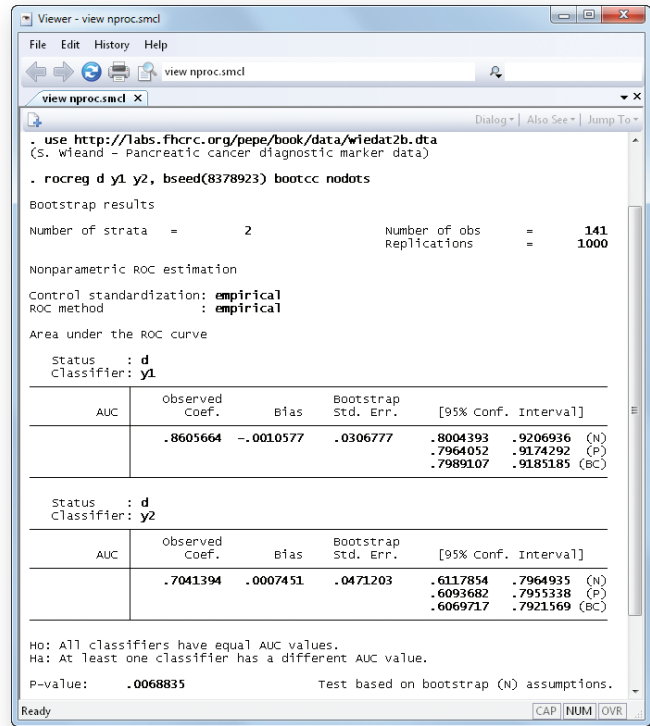
En este artículo, veremos cómo las curvas de característica operativa del receptor (ROC) pueden ser utilizadas para evaluar las pruebas de diagnóstico utilizando Stata. En nuestro primer ejemplo, utilizaremos el análisis de la curva ROC para evaluar las pruebas de diagnóstico para el cáncer de páncreas basadas en los clasificadores CA 19-9 y CA-125.

El usuario de la prueba de diagnóstico puede elegir diferentes valores del clasificador como “puntos de corte” para diagnosticar un individuo. Aquellos que caen por debajo del punto de corte se clasifican como controles, mientras que los que están por encima del punto de corte se clasifican como casos. Cada valor de corte del clasificador corresponde a una proporción de sujetos con la enfermedad clasificados correctamente, (tasa de verdaderos positivos) y a una proporción de sujetos sin la enfermedad clasificados incorrectamente (falsos positivos). Un test ideal tiene puntos de corte con altas tasas de verdaderos positivos y bajas tasas de falsos positivos.

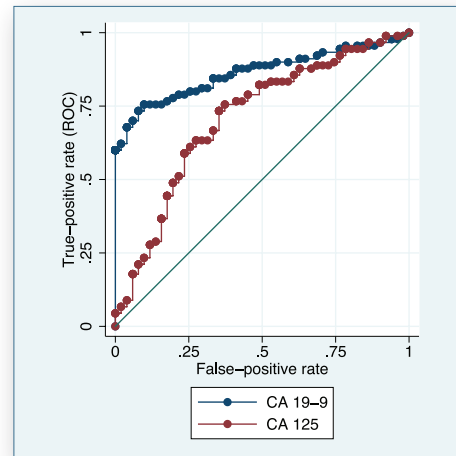
Cuando las condiciones reales de los individuos están determinadas, las tasas de verdaderos positivos y de falsos positivos pueden ser estimadas. La curva ROC es una gráfica de la tasa estimada de verdaderos positivos (valor ROC) en función de la tasa estimada de falsos positivos. Las curvas ROC se pueden usar para evaluar y comparar las pruebas de diagnóstico. Además de las tasas de verdaderos y falsos positivos, también se utiliza el área bajo la curva ROC (AUC) para comparar las pruebas de diagnóstico.

Estimación no paramétrica de la curva ROC

Examinamos los datos de un estudio sobre el cáncer de páncreas con dos clasificadores continuos, **y1** (CA 19-9) y **y2** (CA-125). La variable **d** indica si la persona realmente tiene cáncer. Utilizamos **rocreg** para realizar una estimación no paramétrica de las tasas de verdaderos positivos y falsos positivos para los valores únicos de cada clasificador. **rocreg** también realiza una estimación no paramétrica de la AUC y utiliza el método de bootstrap para obtener el error estándar. Usamos la opción **bootstrap** para especificar que se realice el remuestreo estratificado por casos y controles.



La estimación de la AUC es mayor en el caso de CA 19-9, y rechazamos la hipótesis de que las dos pruebas tienen la misma AUC al nivel 0,01. Para poner estas cifras en perspectiva, utilizamos **rocgplot** para dibujar la curva ROC implícita en los resultados **rocreg**:



El test que usa el clasificador CA 19-9 es claramente preferible al que usa el clasificador CA-125. Para todas las tasas de falsos positivos (excepto las más altas), se observa una mayor tasa de verdaderos positivos con CA 19-9 que con CA-125.

Estimación de la curva ROC bajo hipótesis de normalidad

rocreg también puede estimar curvas ROC paramétricamente mediante el uso de ecuaciones de estimación o de máxima verosimilitud (MV). Con MV, modelamos el clasificador **y** utilizando las covariables de control **z**, las covariables de caso **x**, y un indicador **d** para la presencia de la condición, suponiendo que **y** se distribuye normalmente condicionada a las covariables y al indicador de la condición **d**.

Las covariables de control afectan al clasificador tanto para el grupo de casos como para el grupo de control, mientras que las covariables de casos tienen un efecto adicional solamente sobre la población de casos.

Cuando se utiliza una variable de clasificación que corresponde a este modelo, hay una curva ROC diferente para cada combinación de valores de las covariables de caso. Esencialmente, se utiliza un clasificador diferente $y | x$ para cada combinación de valores de las covariables de caso x . Tanto las covariables de caso como las de control de cada individuo se utilizan para estimar el modelo subjacente; luego se evalúa la prueba de diagnóstico que utiliza el clasificador y bajo las covariables de caso, basándose en la curva ROC específica para las covariables.

Para ilustrar este modelo, examinamos los datos de un estudio audiológico neonatal. Nuestra variable de clasificación es **y2** (TEOAE 80 a 2 kHz). La variable binaria **d** indica si un individuo tiene problemas auditivos. La variable **currage** contiene la edad de cada infante (en meses), y la variable **male** indica el género. Dado que no esperamos que el género tenga un efecto directo sobre la curva ROC, la incluimos en la especificación como covariable de control.

Usaremos **rocreg** con la opción **probit ml** para estimar los parámetros de la curva ROC. Más del 90% de los recién nacidos fueron evaluados en ambos oídos, por lo que los agrupamos usando el identificador de individuo (**id**).

Viewer - view proc.smcl

```

File Edit History Help
view proc.smcl
view proc.smcl x
Dialog | Also See | Jump To
. use http://www.stata-press.com/data/r12/nhns
(Norton - neonatal audiology data)
. rocreg d y2, probit ml ctrlcov(currage male) roccov(currage) cluster(id) nolog
Parametric ROC estimation
Covariate control : linear regression
Control variables : currage male
Control standardization: normal
ROC method : parametric Link: probit

Status : d
Classifiers: y2
Classifier : y2
Covariate control adjustment model:
(Std. Err. adjusted for 2741 clusters in id)

```

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
casecov					
currage	.5729861	.2251764	2.54	0.011	.1316485 1.014324
_cons	-18.2597	8.615108	-2.12	0.034	-35.14501 -1.374403
casesd					
_cons	9.723858	.7540084	12.90	0.000	8.246029 11.20169
ctrlcov					
currage	-.1694575	.0330224	-5.13	0.000	-.2341803 -.1047348
male	-.7122587	.2298645	3.10	0.002	-.2617326 1.162785
_cons	-5.651728	1.284617	-4.40	0.000	-8.169532 -3.133925
ctrlsd					
_cons	6.986167	.098408	70.99	0.000	6.793291 7.179044
Status : d					
ROC Model :					(Std. Err. adjusted for 2741 clusters in id)
	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
y2					
i_cons	-1.877825	.905174	-2.07	0.038	-3.651933 -.1037167
currage	-.0589258	.0235849	2.50	0.012	-.0127002 -.1051514
s_cons	.7184563	.0565517	12.70	0.000	.607617 .8292957

Ready CAP NUM OVR

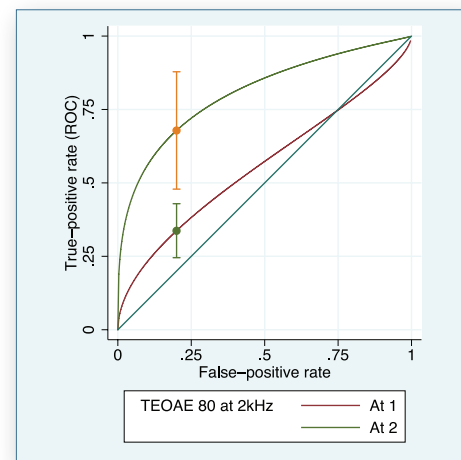
Todos los coeficientes son significativos al nivel 0,05. La tabla “Covariate control adjustment model” indica cómo las covariables afectan directamente a la variable de clasificación. La edad tiene un efecto negativo en el clasificador dentro de la población de control, pero este efecto está moderado por la población de casos.

Ser varón tiene un efecto positivo en el clasificador; no afecta a la curva ROC, porque suponemos que el género no tiene un efecto adicional en la población de casos. Sin embargo, su inclusión en el modelo mejora la precisión global del modelo.

La tabla “ROC Model” muestra cómo las covariables de caso afectan directamente la curva ROC. Aquí vemos que la edad tiene una influencia positiva en la curva. También vemos que el coeficiente para el cuantil normal estándar para la tasa de falsos positivos, referido como **s_cons**, es significativamente menor que 1. Este parámetro puede ser visto como una forma de pendiente para la tasa de falsos positivos, y corresponde a la relación entre la desviación estándar de la población de control y la de la población de casos. En este ejemplo, se estima que la población de control es menos variable que la de casos. Por lo tanto, la tasa de falsos positivos tiene un efecto menor en la curva ROC.

Podemos visualizar las curvas ROC para diferentes edades utilizando **rocregplot**. Aquí contrastamos las curvas correspondientes a los infantes de 35 meses con las de los de 50 meses utilizando la opción **at()**. También especificamos la opción **roc()** para estimar la tasa de verdaderos positivos correspondiente a una tasa de falsos positivos de 0,2.

. rocregplot, at1(currage=35) at2(currage=50) roc(.2)



El clasificador TEOAE 80 a 2 kHz claramente tiene un mejor desempeño a la edad de 50 meses que a la edad de 35 meses.

RESUMEN

Las curvas ROC son una herramienta útil para evaluar y comparar tests de diagnóstico. Los comandos de Stata **rocreg** y **rocregplot** hacen de el trazado y la evaluación de las curvas ROC una tarea simple.



StataCorp
4905 Lakeway Drive
College Station, TX 77845-4512
USA

Cómo contactarnos

979-696-4600 979-696-4601 (fax)
service@stata.com stata.com

Por favor, incluya su número de serie de Stata en toda su correspondencia.



facebook.com/StataCorp



twitter.com/Stata



blog.stata.com

Encuentre un distribuidor cerca de usted
stata.com/worldwide

 Derechos de Autor 2012 por StataCorp LP.

Serious software for serious researchers. Stata es una marca registrada de StataCorp LP. Serious software for serious researchers es una marca registrada de StataCorp LP.

Nuevo curso abierto de capacitación: Structural Equation Models Using Stata

25 al 26 de octubre de 2012

Este curso cubre el uso de Stata para el ajuste de modelos de ecuaciones estructurales (SEM). SEM es una clase de técnicas estadísticas para modelar las relaciones entre las variables, tanto observadas como no observadas. SEM abarca algunos modelos conocidos como la regresión lineal, la regresión multivariante, y el análisis factorial, y se extiende a una variedad de modelos más complicados.

El curso proporcionará una introducción a los SEM lineales. Además, se discutirá una variedad de modelos que caen dentro del marco de SEM, con énfasis en el uso de Stata para ajustar cada uno de los mismos. Éstos incluyen análisis factorial confirmatorio (CFA), análisis de caminos, modelos no recursivos (simultáneos), modelos de crecimiento latente, modelos de mediación, modelos con múltiples grupos, y modelos de ecuaciones estructurales completos. Stata permite ajustar los modelos de ecuaciones estructurales de dos formas: mediante el uso de la sintaxis del comando **sem** o mediante la interfaz gráfica de usuario para dibujar diagramas de caminos. Los ejemplos mostrarán ambos enfoques.

Temas del curso

- Visión general de SEM
 - › Descripción del modelo
 - › Proceso de ajuste y evaluación de modelos de ecuaciones estructurales
 - › Descripción de los diagramas de camino
- Herramientas de Stata para SEM
 - › Ajuste de modelos con el comando **sem**
 - › Construcción de modelos mediante la GUI para SEM
 - › Uso de los comandos **ssd** para trabajar con estadísticos de resumen
- Comprobación e interpretación de los resultados de SEM
 - › Resultados estandarizados
 - › Efectos directos, indirectos y totales
 - › Estadísticas de bondad de ajuste
 - › Índices de modificación
 - › Prueba de Wald y pruebas de scores
 - › Pruebas para múltiples grupos

El costo de inscripción para este curso es de 1.295 dólares por persona, y el espacio está limitado a 24 plazas. Por más información o para inscribirse, visite stata.com/training/structural-equation-modeling-using-stata.

Para ver la lista de todos los cursos abiertos de capacitación, consulte la página 8.