

LAS NOTICIAS DE STATA

abril/mayo/junio 2010
Volumen 25, Número 2

Stata 11.1 ya está disponible

Conozca algunas de las nuevas funcionalidades disponibles en la actualización gratuita más reciente de Stata.

p. 1

Stata marca la diferencia

Vea cómo Stata es utilizado en el Banco Mundial.

p. 2

Manejo de datos usando Stata: una guía práctica

Lea acerca del nuevo libro de Michael N. Mitchell sobre manejo de datos.

p. 3

Temas destacados

Conozca acerca de variables factoriales y el método de los momentos generalizado.

p. 4

Actualización del reporte de Stata/MP

Descubra dónde se exploran todos los detalles acerca de Stata/MP.

p. 7

Conferencia de Stata Boston 2010

Haga planes para asistir a la conferencia anual de Stata en Boston.

p. 12

También en este número

Cursos Abiertos de Capacitación.....	8
Novedades de la Librería de Stata	9
Encuentros de usuarios.....	11
Próximas fechas de NetCourse.....	11

Las Noticias de Stata

Editor ejecutivo: Brian Poi
Supervisora de Producción: Annette Fett
Editora de la versión en español: Isabel Cañette

Stata 11.1 ya está disponible.

Stata 11.1 ya está disponible como actualización gratuita de Stata 11. Esta actualización agrega nuevas funcionalidades y extensiones a funcionalidades ya existentes. Si usted tiene Stata11, escriba **update all** en Stata, y luego escriba **update swap**. Alternativamente, seleccione **Help** en el menú principal de Stata, y luego seleccione **Official Updates** y siga las instrucciones desplegadas.

Las siguientes son algunas de las nuevas funcionalidades incorporadas en Stata 11.1:

Imputación múltiple. **mi** ahora soporta oficialmente el ajuste de modelos de datos de panel y de modelos multinivel.

Modelos para datos de conteo truncados. Los comandos **tpoisson** y **tnbreg** ajustan modelos para variable dependiente de conteo con distintas formas de truncamientos a la izquierda, incluyendo truncamientos que varían con las observaciones.

Modelos mixtos. Los modelos lineales mixtos (multinivel) ahora se pueden ajustar con nuevas estructuras de covarianza con los siguientes residuos: exponencial, Toeplitz y por bandas.

Predicción de probabilidades. **predict**, invocado luego de modelos de conteo como **poisson** y **nbreg**, ahora permite predecir la probabilidad de cualquier valor o cualquier intervalo.

Bootstrap para datos de encuestas. Los comandos de estimación ahora pueden estimar errores estándar bootstrap, utilizando pesos de replicación provistos por el usuario.

Pesos de replicaciones de diferencias sucesivas. Ahora las estimaciones para datos de encuestas (survey) aceptan pesos de replicaciones de diferencias sucesivas. Estos pesos son suministrados con varios conjuntos de datos, como los de la Oficina de Censos de los Estados Unidos.

Concordancia. **estat concordance** agrega una nueva medida de concordancia, el coeficiente *K* de Gönen y Heller, que es robusto en presencia de censura.

Bondad de ajuste para datos de encuestas.

Se han incorporado tests de bondad de ajuste para **probit** y **logit** con datos de encuestas.

Errores estándar robustos. Errores estándar robustos o por conglomerado (cluster) han sido agregados a **xtpoisson**, **fe**.

CV para datos de encuestas. El coeficiente de variación (CV) ahora está disponible para datos de encuestas.

Formateo de estimaciones. Ahora es posible personalizar los formatos numéricos en los resultados de las regresiones. Usted puede establecer el número de lugares decimales para los coeficientes, errores estándar, valores *p*, e intervalos de confianza utilizando argumentos en la línea de comando o el comando **set**.

También se puede controlar el despliegue de variables factoriales en tablas de estimación. Estos controles incluyen extra espacio alrededor de las variables factoriales, despliegue de celdas vacías, despliegue de los niveles de base, y la omisión de variables que han sido excluidas debido a colinealidad.

Cualidades técnicas de Stata/MP. Stata/MP, la versión paralela de Stata, ha incrementado su rendimiento en varias áreas, incluyendo aún más comandos de panel paralelizados, mayor paralelización en estimaciones con más de 200 variables independientes, y un mejor ajuste de de MP para números elevados de procesadores o núcleos.

Continúa en p. 2

Adiciones al portapapeles. El soporte del portapapeles en el Editor de Datos ha tenido adiciones. Ahora, las copias en el portapapeles retienen los formatos de las variables y otras características de los datos, cuando son transferidos desde Stata.

Windows XP. La cantidad de memoria disponible para usar en Stata en Windows XP 32 bits ha sido incrementada.

Editor de Archivos Do. Resaltador de sintaxis, marcadores y otras funcionalidades del Editor de archivos Do han sido incorporadas a Stata para Mac.

ODBC. Se ha adicionado soporte ODBC para Sun Solaris (Oracle Solaris).

Ventanas de Diálogo. Las ventanas de diálogo en Unix ahora tienen controles que permiten seleccionar variables de una lista de variables en su conjunto de datos.

Por más información, escriba

help whatsnew luego de actualizar Stata 11 y use las ligas a los distintos comandos.

Si aún no tiene Stata 11, usted se está perdiendo una serie de funcionalidades, incluyendo:

- Imputaciones múltiples (con su propio manual)
- Estimación por el método de los momentos generalizado (GMM)
- Regresión para riesgos en competencia
- Variables factoriales
- Modelos de espacio y estado
- Medias marginales y márgenes predictivos
- Efectos marginales promedio
- Letras griegas, itálicas y otras fuentes en gráficos
- Modelos de factor dinámico
- Modelos GARCH multivariados
- Estructuras de covarianza en los errores para modelos lineales mixtos
- Pruebas de raíz unitaria para datos de panel
- Manejador de Variables
- Nuevo Editor de Datos
- Nuevo Editor de Archivos Do
- Programación orientada a objetos en Mata
- Más comandos paralelizados en Stata/MP
- Documentación en .pdf

Por más información acerca de Stata 11, visite www.stata.com/stata11/.

.....

Stata marca la diferencia en el Banco Mundial: Análisis de pobreza automatizado

El Banco Mundial apoya los Objetivos de Desarrollo del Milenio de las Naciones Unidas de eliminar la pobreza y proveer desarrollo continuo. Para mejorar la situación en un área, primero es necesario entender quiénes son más afectados por la pobreza, y cómo la pobreza se distribuye entre los miembros de la sociedad.

La evaluación de la pobreza es un elemento clave en las estrategias de reducción de la pobreza del Banco Mundial. Estos reportes son producidos periódicamente para prácticamente todos los países estudiados por el banco. Cada evaluación de pobreza incluye varias estadísticas sobre pobreza e ingresos y reportes sobre el progreso de cada país en el alcance de sus objetivos de reducción de pobreza. Históricamente, la producción de una evaluación de pobreza para un país determinado implicaba el contratar un consultor, frecuentemente un estudiante de posgrado o un doctor graduado recientemente. El consultor aprendía sobre los principios de análisis de pobreza y escribía código en Stata para producir las tablas y gráficos requeridos. Este enfoque estaba expuesto a errores porque no había lineamientos precisos para estos reportes. Los programas y la documentación eran producidos por personas con distintas formaciones. Las metodologías y las hipótesis eran muchas veces imprecisas, y los resultados eran difíciles de reproducir. El mantenimiento del código y la preparación de los datos eran procedimientos costosos. Para realizar un análisis similar a otro ya existente, el investigador muchas veces necesitaba volver a hacer todo desde el principio, en lugar de utilizar el código ya existente.

Conjunto de Herramientas Para el Análisis de la Pobreza

Para rectificar y agilizar el proceso de la producción de evaluaciones de pobreza, Michael Lokshin, el economista principal en el Grupo de Investigación de Desarrollo del Banco Mundial, y su equipo, incluyendo Sergiy Radyakin y Zurab Sajaia, escribieron un conjunto de archivos .ado, a los que con el tiempo se empezó a referir como Conjunto de Herramientas Para el Análisis de Pobreza (CHAP), y fueron ampliamente utilizados en el Banco Mundial. El comando escrito por usuarios `xml_tab`, muy popular entre usuarios de Stata, también es producto del trabajo de este equipo. Dicho comando, disponible a través de los archivos SSC (Statistical Software Component), permite a los usuarios grabar resultados de Stata en un formato que se puede incorporar fácilmente en planillas de texto Microsoft Excel.

El CHAP incluye varios programas para análisis dinámico de políticas, entre ellos, comandos para graficar curvas de incidencia de crecimiento, incidencia de pobreza, déficit, y curvas de severidad; también para analizar cambios en la pobreza a través del tiempo debidos a cambios sectoriales y de población, y crecimiento y redistribución.

El CHAP simplificó significativamente el estudio de la pobreza en el Banco Mundial poniendo a disposición de los investigadores un conjunto estándar de comandos de Stata, que aquellos pueden utilizar sin necesidad de reinventar la rueda. Sin embargo el tener una colección de programas en lugar de una interfaz única requería un cierto período de entrenamiento para nuevos investigadores, y limitaba la producción de resultados estandarizados que pudieran ser incluidos fácilmente en los reportes.

ADePT

Para que un mayor número de usuarios se interesara en el CHAP, Lokshin y su equipo decidieron combinar las distintas rutinas y proveer una única interfaz de uso sencillo. El CHAP recibió entonces el nuevo nombre de ADePT (Automated DEC Poverty Tables) y fue adoptado rápidamente por investigadores alrededor del mundo. A diferencia del CHAP, ADePT, disponible en www.worldbank.org/adept/, no es un conjunto de componentes aislados, sino una plataforma integrada. El tener una plataforma integrada permite que los componentes funcionen conjuntamente y facilita la creación de nuevos módulos.

ADePT fue desarrollado utilizando una combinación del lenguaje de programación ado de Stata, Mata, y lenguajes de programación de diálogos, constituyendo más de 150.000 líneas de código. Algunas rutinas fueron desarrolladas en C++ y en lenguaje assembly para maximizar la eficiencia, e incorporadas por medio de las utilidades "plug-in" de Stata. Un ejemplo de este tipo de rutina es el comando `usespss`, que está disponible en los archivos SSC. Este comando permite a los usuarios de Stata leer datos en formato SPSS.

Distintos módulos de ADePT efectúan una variedad de análisis estadísticos, desde una simple tabla de dos entradas hasta estimaciones de ecuaciones simultáneas vía máxima verosimilitud. Otras rutinas permiten calcular errores estándar para distintas medidas de desigualdad y pobreza. ADePT incluso permite realizar simulaciones complejas en forma sencilla.

Para utilizar las primeras versiones de ADePT, los investigadores necesitaban tener Stata instalado en sus computadores. Este requerimiento constituía un impedimento en la distribución de ADePT entre usuarios que no tenían Stata, particularmente en países en vía de desarrollo con bajos ingresos, donde la investigación en temas de pobreza es más crítica.

Numerics by Stata

En 2009, StataCorp anunció **Numerics by Stata**, una versión de Stata que permite a quienes desarrollan software crear aplicaciones en el lenguaje de su preferencia, e invocar Numerics by Stata para crear aplicaciones con interfaces de usuarios de acuerdo a sus necesidades particulares, y utilizando la potencia analítica de Stata. Dichas aplicaciones no requieren que los usuarios finales tengan Stata instalado en sus máquinas.

A fines de 2009, el Banco Mundial modificó ADePT para trabajar con Numerics by Stata. Como resultado, investigadores alrededor del mundo que no tenían Stata y por lo tanto no podían usar versiones previas de ADePT, pudieron bajar la nueva versión de ADePT de la página web del Banco Mundial y comenzar a utilizarla inmediatamente.

Las funcionalidades de automatización de Stata permitieron a los investigadores desarrollar una interfaz en C# utilizando componentes estándar de Windows, y continuar efectuando análisis estadísticos utilizando el lenguaje de programación

de Stata. Esta separación de la interfaz para el usuario de la parte que realiza los cálculos permitió desarrollar versiones localizadas de ADePT. Además de la versión en inglés, también están disponibles versiones en indonesio y en ruso. Se está trabajando para producir versiones en español, búlgaro, rumano, georgiano y portugués, entre otros idiomas. El contar con versiones de ADePT en el lenguaje local facilita la adopción de ADePT por parte de los investigadores.

La localización no se limita al cambio del lenguaje utilizado en la interfaz. Por ejemplo, en noviembre de 2009, el Equipo de Pobreza de Indonesia colaboró con el equipo de desarrollo de ADePT para adaptar el software para poder realizar análisis al nivel de kabupaten (distrito). Esta adaptación permitió a los investigadores utilizar datos de encuestas de hogares para responder preguntas específicas de políticas en Indonesia.

Lokshin estima que actualmente existen más de 1.500 usuarios registrados de ADePT. A medida que el número de usuarios sigue creciendo, se realizan más pruebas y perfeccionamientos del software. Por ejemplo, cuando los usuarios encuentran casos difíciles y excepciones, el equipo de ADePT puede utilizar esa información para modificar el código y hacerlo más robusto a esas situaciones. El equipo de desarrollo utiliza procedimientos automatizados de testeo para asegurar que los resultados de nuevas versiones coinciden con los de las versiones anteriores.

ADePT reduce el tiempo necesario para realizar análisis de datos, dando a los usuarios más tiempo para concentrarse en los resultados. Los usuarios también pueden hacer simulaciones para explorar los posibles resultados de políticas alternativas como transferencias de efectivo o subsidios. Más importante, ADePT es fácil de usar y, con la preparación de datos apropiada, mejora la fidelidad y la coherencia de los resultados

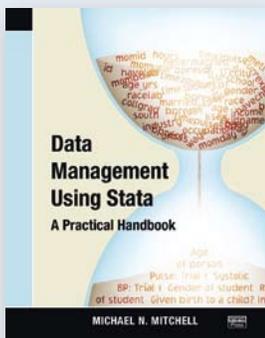
Resumen

El Banco Mundial se ha propuesto reducir la pobreza en el mundo, y Stata ha jugado un papel crucial en este emprendimiento durante muchos años. La flexibilidad de Stata ha contribuido a hacer posible el Conjunto de Herramientas para el Análisis de la Pobreza, que mejoró la eficiencia de investigadores alrededor del mundo. El equipo de desarrollo del Banco Mundial utilizó las ventanas de diálogo programables de Stata, Mata, y un sistema de clases para crear ADePT, lo que mejoró sustancialmente el análisis de la pobreza. El software anunciado recientemente Numerics by Stata permite que investigadores de pobreza alrededor del mundo lleven a cabo sus análisis en forma eficaz y económica.

— Brian Poi, Editor Ejecutivo
y Economista Sénior

Novedades de Stata Press

Data Management Using Stata: A Practical Handbook



Autor: Michael N. Mitchell
Editorial: Stata Press
Derechos de autor: 2010
Páginas: 387; cubierta suave
ISBN-10: 1-59718-076-9
ISBN-13: 978-1-59718-076-4
Precio: USD 48,00

El libro de Michael N. Mitchell *Data Management Using Stata*, se ocupa en forma detallada de tareas de manejo de datos, desde aquellas que podría necesitar quien se inicia en la estadística, hasta aquellas tareas complicadas que podrían confundir a un usuario experimentado.

Mitchell realiza esta tarea utilizando un lenguaje simple y ejemplos ilustrativos.

El libro está estructurado en módulos, y cada uno de éstos se basa en conjuntos de tareas de manejo de datos, no en conjuntos de comandos. Este formato es útil porque permite a los lectores encontrar fácilmente las herramientas que necesitan para resolver problemas específicos. Complementando este formato, el libro está escrito en un estilo que transmitirá a los lectores buenos hábitos en el manejo de datos, aún si los lectores recurren al texto esporádicamente o si leen los capítulos en distinto orden.

A lo largo del libro, Mitchell enfatiza sutilmente la necesidad imperiosa de generar análisis reproducibles y de mantener registros de los mismos.

El libro no se centra en elementos complicados de programación, sino que consolida hábitos simples y señala los beneficios de trabajar en forma cuidadosa y ordenada. Muchos de estos consejos están basados en la experiencia de Mitchell en los Servicios de Tecnología Académica de la UCLA.

Mitchell incluye consejos para aquellos que desean aprender a escribir sus propios comandos para el manejo de datos. Aún usuarios experimentados aprenderán trucos útiles y nuevas formas de abordar sus problemas de manejo de datos.

Este es un libro muy útil, ampliamente beneficioso para quienes estén interesados en el manejo de datos usando Stata.

Para ver tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite: www.stata-press.com/books/dmus.html.

Temas destacados: variables factoriales

Las variables factoriales fueron introducidas en Stata 11. A mi me gustan por muchas razones, entre ellas:

1. Ahorran espacio.
2. Facilitan la sintaxis para especificaciones factoriales completas.
3. Funcionan prácticamente en todas partes, lo que resulta muy cómodo.
4. Simplifican los tests de cambios estructurales.

En primer lugar, en caso de que usted aún no esté usando variables factoriales, voy a introducirlos. Cualquier variable categórica en Stata puede ser tratada como una variable factorial (bueno, no exactamente *cualquier* variable; la variable debe tomar únicamente valores enteros no negativos). A menudo no queremos tratar una variable categórica como cardinal u ordinal, sino más bien como un conjunto de variables indicadoras, una para cada uno de los niveles (valores) que toma la variable. Esto lo hacemos utilizando los operadores de variables factoriales de Stata.

El operador más utilizado es `i.`. También existen otros operadores, y veremos algunos de ellos en este artículo. Usando el venerable `auto.dta`, cuando escribimos `i.rep78`, en realidad nos estamos refiriendo a cinco variables indicadoras, una para cada nivel de `rep78`. (Sí, uno de ellos es un nivel base, que es especial, pero no voy a entrar en esos detalles). Podemos ver las variables indicadoras escribiendo

```
. sysuse auto
. list rep78 i.rep78
```

	rep78	1b. rep78	2. rep78	3. rep78	4. rep78	5. rep78
1.	3	0	0	1	0	0
2.	3	0	0	1	0	0
3.
4.	3	0	0	1	0	0
5.	4	0	0	0	1	0
6.	3	0	0	1	0	0
7.
8.	3	0	0	1	0	0
9.	3	0	0	1	0	0
10.	3	0	0	1	0	0
11.	3	0	0	1	0	0
12.	2	0	1	0	0	0
13.	3	0	0	1	0	0
14.	3	0	0	1	0	0
15.	4	0	0	0	1	0

Ahora, volvamos a mi lista parcial de las razones por las cuáles me gustan las variables factoriales.

1. Ahorran espacio

Uno de los conjunto de datos provenientes de la Encuesta Longitudinal Nacional de la Mujer en los Estados Unidos se puede cargar en Stata escribiendo `sysuse nlsw88`. Una de las variables en ese conjunto de datos es la variable `industry` que tiene 12 niveles, los que representan la industria en la que una mujer estaba trabajando, por ejemplo la manufactura o los bienes inmuebles. Para efectuar una regresión de los salarios (`wage`) en las ramas de actividad, puedo escribir

```
. regress wage i.industry
```

Source	SS	df	MS			
Model	4743.21081	11	431.200983	Number of obs =	2232	
Residual	69374.068	2220	31.2495802	F(11, 2220) =	13.80	
Total	74117.2788	2231	33.2215503	Prob > F =	0.0000	
				R-squared =	0.0640	
				Adj R-squared =	0.0594	
				Root MSE =	5.5901	

wage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
industry 2	9.728472	3.106542	3.13	0.002	3.636439	15.8205
3	1.943813	1.707566	1.14	0.255	-1.404781	5.292407
4	1.880457	1.386852	1.36	0.175	-.8392066	4.60012
5	5.822232	1.478319	3.94	0.000	2.923198	8.721265
6	.5047755	1.389983	0.36	0.717	-2.221028	3.230579
7	4.222053	1.414556	2.98	0.003	1.448061	6.996044
8	1.894669	1.483771	1.28	0.202	-1.015056	4.804394
9	-1.220028	1.46982	-0.83	0.407	-4.102394	1.662338
10	1.103288	1.9174	0.58	0.565	-2.656797	4.863372
11	2.250065	1.369721	1.64	0.101	-.436003	4.936132
12	3.527286	1.419776	2.48	0.013	.7430572	6.311514
_cons	5.621121	1.355806	4.15	0.000	2.96234	8.279902

Aquí utilicé la notación factorial `i.industry` para efectuar la la regresión de la variable `wage` sobre variables indicadoras para los distintos niveles de la industria. ¿Cuál es la diferencia entre esta forma de efectuar la regresión y otras formas que habrían dado el mismo resultado (como la creación de los indicadores con `tabulate`, `generate()` o mediante el uso de `xi`)? La diferencia es que Stata no necesita crear las 12 variables indicadoras para ejecutar la regresión. Los valores de los indicadores fueron creados "al vuelo", usando los valores de la variable `industry`.

2. Facilitan la sintaxis para especificaciones factoriales completas.

Continuando con la regresión del salario de la Encuesta Longitudinal Nacional, yo también podría estar interesado en el efecto de las categorías ocupacionales de la mujer (ventas, profesional, gestión, etc). Podría agregar indicadores para cada nivel de ocupación a la regresión de la misma manera que he incluido indicadores para la industria. ¿Qué pasa si, sin embargo, también se sospecha que los efectos de ocupación difieren entre las industrias (es decir, que existe un efecto de interacción)? Existe una notación de variable factorial para ingresar en el modelo los 12 indicadores para la industria, los 13 indicadores de ocupación, y todos los indicadores para cada combinación de la industria y la ocupación. Simplemente escribimos

```
. regress wage industry##occupation
```

El primer operador, `#`, implica que queremos incluir la interacción, y el segundo operador, `##`, dice que queremos incluir todas las interacciones de bajo nivel, hasta los niveles de las variables categóricas originales.

Para aquellos versados en el diseño experimental, todo esto sonará mundano. Ellos se refieren a esto como una especificación factorial completa, y el comando `anova` lo hace como una especificación estándar. Lo que no es mundano es que con las variables factoriales, esta notación está disponible para casi todos los comandos de estimación, y, de hecho, para casi todos los comandos de Stata.

Lo que también es agradable es que Stata no necesitó crear las $12+13+12 \times 13 = 181$ variables indicadoras. Éstas fueron creadas "al vuelo", a medida que el comando `regress` las fue necesitando para efectuar los cálculos.

Las variables factoriales realmente ahorran espacio.

3. Funcionan prácticamente en todas partes, lo que resulta muy cómodo.

Sí, yo sé que **xi** haría algo similar para las especificaciones factoriales, pero la notación de variables factoriales es comprendida por Stata a un nivel mucho más profundo. Intente obtener predicciones en un conjunto de datos que no sea el conjunto de datos de estimación con **xi**.

- . webuse margex, clear
- . xi: logistic outcome i.treatment*i.group age
- . webuse hstandard
- . predict prbhat

El comando **predict** fallará, porque cuando se utilizó **xi**, Stata simplemente creó las variables indicadoras, y, a pesar de que **hstandard.dta** contiene las mismas variables categóricas, Stata sabe que necesita volver a crear indicadores para hacer la predicción.

Intente hacer lo mismo con las variables factoriales:

- . webuse margex
- . logistic outcome treatment##group age
- . webuse hstandard
- . predict prbhat

¡Funciona! Funciona porque Stata sabe que, por ejemplo, **group==2** es un indicador del segundo grupo, y esta información no se pierde cuando se cambia el conjunto de datos. Por supuesto, usted debe asegurarse de que 2 signifique lo mismo en sus dos conjuntos de datos.

4. Simplifican los tests de cambios estructurales.

A veces creemos que tenemos una buena especificación del modelo, pero no estamos seguros de si los parámetros para un grupo son aplicables a otro grupo. Podemos aplicar lo que economistas llaman a menudo un test de Chow (Chow (1960)), (aunque este tipo de test se realiza en muchas disciplinas).

Imagine un modelo de duración de hospitalización antes y después de un cambio en la política de admisión, un modelo de crecimiento económico antes y durante una guerra, o un modelo de poblaciones de peces en un arrecife antes y después de un huracán. Supongamos que queremos realizar una prueba conjunta para ver si los parámetros coinciden en los dos grupos. Nuevamente, usando **auto.dta**, si suponemos que el precio está determinado por el rendimiento en millas por galón (**mpg**) y el peso (**weight**) (como indicador de tamaño), podríamos ajustar un modelo para esa relación escribiendo

- . sysuse auto, clear
- . regress price mpg weight

¿Qué ocurre si, además, creemos que los precios de los coches de alta y de baja calidad tienen una relación básicamente diferente con el rendimiento y el peso? Tal vez el efecto de la calidad sea reducir la importancia de ambos, el rendimiento y el tamaño. Podemos agrupar los automóviles en función de si sus registros de reparación son mejores que "adecuado" (representado por el número 3)

- . gen highq = rep78 > 3

Ahora podemos estimar ambos, el efecto total de rendimiento y el de peso en el precio, y también podemos estimar las diferencias de esos efectos en los coches de alta calidad. Estos son simplemente los resultados por defecto de una especificación factorial completa que incluye las interacciones con las

variables independientes continuas.

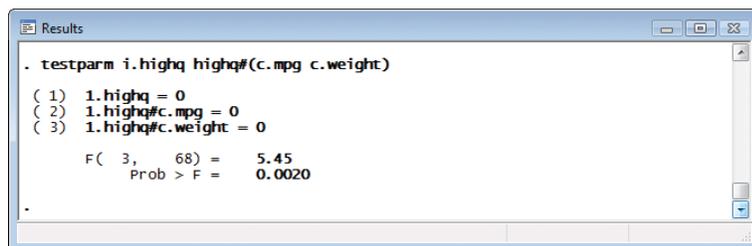
- . regress price highq##(c.mpg c.weight)

Podemos verificar si el indicador de alta calidad cambia los coeficientes de rendimiento y peso escribiendo

- . testparm highq#(c.mpg c.weight)

Esto es casi lo mismo que escribimos para efectuar la regresión, sólo el operador es diferente. Queremos poner a prueba sólo los efectos de interacción, no el efecto general, por lo que escribimos **#** en lugar de **##**. La prueba clásica de Chow incluye el test de si la constante varía entre los dos grupos, por lo que se puede añadir **i.highq** de nuestra prueba:

- . testparm i.highq highq#(c.mpg c.weight)



```

Results
. testparm i.highq highq#(c.mpg c.weight)
( 1) 1.highq = 0
( 2) 1.highq#c.mpg = 0
( 3) 1.highq#c.weight = 0

      F( 3, 68) = 5.45
      Prob > F = 0.0020

```

La prueba confirma nuestra sospecha de que existe una diferencia estructural en la determinación de precios entre los coches de alta calidad y los coches de baja calidad.

Las variables factoriales también nos permiten calcular medias marginales, efectos marginales discretos, etc, con el comando **margins**, lo que es un tema para otro artículo. Si usted no ha utilizado las variables factoriales de Stata 11, le recomiendo que explore las posibilidades.

Le sugiero comenzar leyendo [\[U\] 11.4.3 Factor variables](#) y, a continuación pasar a [\[U\] 25 Working with categorical data and factor variables](#). Los manuales de referencia [\[R\]](#) también utilizan las variables factoriales en muchos ejemplos.

Hay muchas otras razones para que a uno le gusten las variables factoriales. Comience a utilizarlas, y encontrará algunas razones propias.

Referencia

Chow, G. C. 1960. Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28: 591–605.

— Vince Wiggins, Vicepresidente,
Desarrollo Científico

Temas destacados: El método de momentos generalizado

El método de momentos generalizado (MMG) es un marco de estimación muy flexible que se ha convertido en un caballo de batalla en los análisis econométricos modernos. A diferencia de la estimación por máxima verosimilitud, el MMG no requiere que el usuario haga fuertes supuestos de distribución, proporcionando así una estimación más robusta. Por otra parte, el MMG tiene un contexto amplio: otros estimadores de uso común como los de mínimos cuadrados y máxima verosimilitud pueden ser vistos como casos particulares del MMG. El MMG es popular en el área de la economía, no sólo por sus propiedades estadísticas favorables, sino también porque muchos modelos teóricos, como los que involucran esperanzas racionales, se basan naturalmente en los supuestos subyacentes al MMG.

El modelo de regresión lineal ordinaria es un estimador por el MMG. En el modelo $y_i = \mathbf{x}_i' \boldsymbol{\beta} + e_i$, asumimos que el término de error e_i (condicionado a \mathbf{x}_i) tiene un valor esperado de cero: $E[e_i | \mathbf{x}_i] = 0$. Por la ley de valores esperados iterados, esto implica que la condición de momento incondicional $E[\mathbf{x}_i' e_i] = E[\mathbf{x}_i'(y_i - \mathbf{x}_i' \boldsymbol{\beta})] = 0$. El MMG elige $\boldsymbol{\beta}$ tal que la versión muestral de esta condición de momentos esté lo más cerca posible de cero. En el caso de regresión por mínimos cuadrados, dado que el número de parámetros es igual al número de elementos de \mathbf{x}_i , la condición de momentos será exactamente igual a cero en el valor elegido de $\boldsymbol{\beta}$ y la estimación por MMG de $\boldsymbol{\beta}$, será igual a la estimación por mínimos cuadrados.

En el caso más general, tenemos condiciones de momentos como $E[\mathbf{z}_i(y_i - \mathbf{x}_i' \boldsymbol{\beta})]$, donde hay más variables en \mathbf{z}_i que parámetros en $\boldsymbol{\beta}$, por lo que la condición de momentos no será igual a cero. En el MMG, elegimos $\boldsymbol{\beta}$ para que la distancia euclídea (ponderada por una matriz de pesos) de las condiciones de momentos al valor cero sea mínima. Por lo general elegimos la matriz de pesos basándonos en las propiedades del término de error (por ejemplo, heteroscedasticidad o autocorrelación). Una característica clave del MMG es que si seleccionamos la matriz de pesos apropiada, el estimador será eficiente, lo que significa que tendrá menor varianza que cualquier otro estimador basado en esas condiciones de momentos.

El comando de Stata `gmm`, nuevo en Stata 11, permite el ajuste de modelos por medio del MMG en un instante. Usted sólo necesita ingresar sus condiciones de momentos por medio de una sintaxis sencilla, especificar sus instrumentos, y seleccionar el tipo de matriz de pesos que desee. Para los modelos más complicados, puede escribir un programa que calcule la condiciones de momentos, de forma similar a los programas que se escriben para utilizar con `ml`.

Tenemos datos sobre el número de visitas al médico que hace una persona, y queremos modelar esta variable basándonos en el género de la persona y sus ingresos, así como si la persona tiene una enfermedad crónica o un seguro privado. Dado que el número de visitas al médico es una variable de conteo, queremos utilizar la regresión de Poisson. Tenemos la sospecha de que los factores no observados que determinan el número de veces que una persona visita el médico también afectan los ingresos de la persona. Por lo tanto, los ingresos constituyen un regresor endógeno. Programas estándar de estimación del modelo de Poisson (como el comando de Stata `poisson`) no pueden tratar la endogeneidad, pero eso no es un problema para `gmm`. El modelo de Poisson conduce a las condiciones de momentos $E[\mathbf{z}_i \{y_i - \exp(\mathbf{x}_i' \boldsymbol{\beta})\}] = 0$. Como instrumentos para los ingresos, vamos a utilizar la edad de cada persona y la raza.

En Stata, escribimos

```
. webuse docvisits, clear
. gmm (docvis - exp({xb:private chronic female income}+{b0})),
    instruments(private chronic female age black hispanic)
```

En el comando `gmm`, hemos especificado nuestra ecuación de residuos $y_i - \exp(\mathbf{x}_i' \boldsymbol{\beta})$ utilizando una expresión sustituible, como las utilizadas por los comandos

Stata/SE 11.1 - http://www.stata-press.com/data/r11/docvisits.dta - [Results]

Review

Command	_rc
1 webuse docvisits, clear	
2 gmm (docvis - exp({xb:private ...	

Step 1
Iteration 0: GMM criterion q(b) = 16.910173
Iteration 1: GMM criterion q(b) = .82276104
Iteration 2: GMM criterion q(b) = .21832032
Iteration 3: GMM criterion q(b) = .12685935
Iteration 4: GMM criterion q(b) = .12672369
Iteration 5: GMM criterion q(b) = .12672365

Step 2
Iteration 0: GMM criterion q(b) = .00234641
Iteration 1: GMM criterion q(b) = .00215957
Iteration 2: GMM criterion q(b) = .00215911
Iteration 3: GMM criterion q(b) = .00215911

GMM estimation
Number of parameters = 5
Number of moments = 7
Initial weight matrix: unadjusted
GMM weight matrix: robust
Number of obs = 4412

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
/xb_private	.535335	.1599039	3.35	0.001	.2219291 .8487409
/xb_chronic	1.090126	.0617659	17.65	0.000	.9690668 1.211185
/xb_female	.6636579	.0959884	6.91	0.000	.4755241 .8517918
/xb_income	.0142855	.0027162	5.26	0.000	.0089618 .0196092
/b0	-.5983477	.138433	-4.32	0.000	-.8696713 -.327024

Instruments for equation 1: private chronic female age black hispanic _cons

Command
gmm (docvis - exp({xb:private chronic female income}+{b0})), instruments(private chronic female age black hispanic)

C:\Program Files\Stata11

de Stata `nl` y `nlshr` para resolver problemas no lineales por mínimos cuadrados. A continuación se especifican los instrumentos (z) en la opción `instruments()`. Por defecto, `gmm` utiliza el estimador en dos etapas con una matriz de pesos que soporta la heteroscedasticidad de los residuos. `gmm` es capaz de ajustar una amplia variedad de modelos, por lo que a diferencia de comandos como `ivregress`, `ivprobit`, o el aclamado comando escrito por usuarios `ivreg2`, requiere que se incluyan todas las variables exógenas en la opción `instruments()`, no sólo las variables excluidas de la ecuación que se está estimando.

En nuestro ejemplo, se utilizaron datos de corte transversal. Sin embargo, `gmm` también proporciona matrices de peso que son adecuadas para su uso con series de tiempo, datos de panel y datos agrupados en conglomerados (clusters). MMG es como una navaja suiza: se puede utilizar directamente para estimar modelos sencillos, y a medida que lo utiliza, usted descubrirá que también puede realizar muchas otras tareas.

— Brian Poi, Editor Ejecutivo y Economista Sénior

Manténgase informado

Para recibir noticias de Stata actualizadas al minuto, visite nuestra página web:

www.stata.com

Allí podrá encontrar anuncios relativos a las actualizaciones de Stata, próximos cursos abiertos de capacitación, conferencias y reuniones de grupos usuarios de Stata, libros de Stata Press, recursos en Internet para el aprendizaje de Stata, archivos de Statalist, más de 400 preguntas frecuentes acerca de Stata y de estadística, y mucho más. También puede suscribirse a una fuente web RSS para tener los titulares de las noticias directamente en su navegador.

¿Prefiere recibir una alerta de correo electrónico con noticias de su interés? Usted puede suscribirse a nuestro servicio de alerta por correo electrónico en:

www.stata.com/alerts/

Stata/MP Actualización del informe de rendimiento

Stata/MP es la implementación de Stata que se ha programado para aprovechar al máximo los ordenadores multiprocesador y multinúcleo. Es exactamente como Stata/SE en todos los aspectos, excepto que distribuye muchas de las tareas exigentes desde el punto de vista computacional entre todos los núcleos de su computadora, y por lo tanto la ejecución de las tareas es más rápida; mucho más rápida.

Si usted tiene una computadora de doble núcleo, usted puede esperar que Stata/MP ejecute regresiones lineales, regresiones logísticas, y muchos otros comandos de estimación en aproximadamente la mitad del tiempo requerido por Stata/SE. Si tiene un equipo de cuatro núcleos, dichos comandos se ejecutarán en aproximadamente una cuarta parte el tiempo, y si tiene un equipo de 8 núcleos, se ejecutarán en aproximadamente un octavo el tiempo. Usted no tiene que cambiar nada para obtener estos incrementos en las velocidades; Stata/MP simplemente correrá más rápido en ordenadores de varios núcleos.

Teniendo en cuenta todos los comandos, Stata/MP corre 1,6 veces más rápido en los equipos de doble núcleo, 2,1 veces más rápido en los ordenadores de cuatro núcleos, y 2,7 veces más rápido en ordenadores de 8 núcleos. Estas son las medianas de los incrementos de velocidad; la mitad de los comandos corren aún más rápido. Algunos de los comandos que tardan más en ejecutarse han sido aún más paralelizados. Teniendo en cuenta todos los comandos de estimación, Stata/MP corre 1,8 veces más rápido en los equipos de doble núcleo, 2,8 veces más rápido en los ordenadores de cuatro núcleos, y 4,1 veces más rápido en los ordenadores de 8 núcleos.

La Figura 1 resume los aumentos de rendimiento de Stata/MP.

Stata/MP es más rápido porque hemos paralelizado a mano 250 secciones cruciales (Más de 10.000 líneas) del código en C de Stata. Estas secciones ahora dividen sus cálculos a través de los núcleos de su computadora mientras se optimiza cuidadosamente la forma en que se realiza esta división, teniendo en cuenta el tamaño del conjunto de datos, el número de variables, y la estructura de cada cálculo. Debido a estos cuidados, Stata/MP es altamente escalable

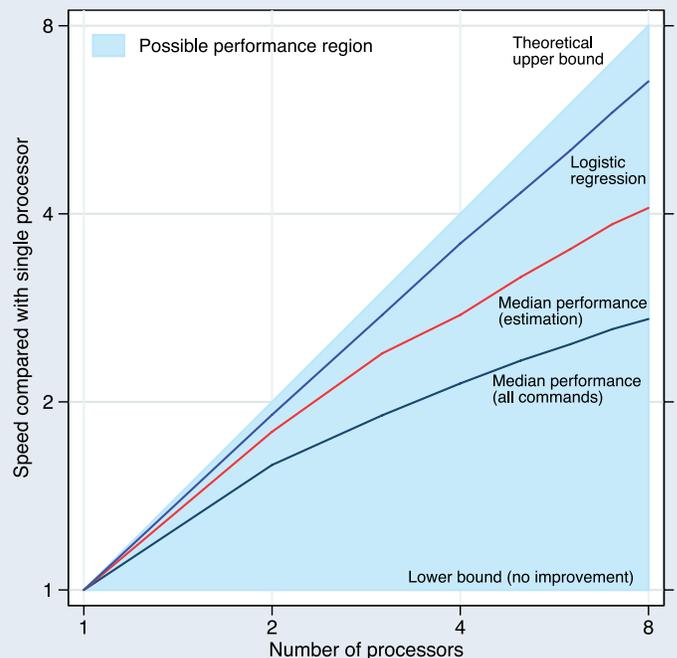


Figura 1

(algunos comandos están casi 100% paralelizados y se ejecutarán más de 40 veces más rápido en 64 núcleos) y también eficiente en el uso de recursos computacionales para comandos menos paralelizados. No se puede obtener este tipo de escalabilidad con técnicas automatizadas para la paralelización de código.

El informe de rendimiento de Stata/MP ofrece una evaluación completa de Stata/MP, incluyendo la evaluación comando por comando. El informe se ha actualizado para Stata 11 y ahora incluye cobertura para todos los comandos de Stata, junto con una sección de discusión sobre Mata. Contiene los gráficos de rendimiento para 488 comandos de Stata.

Lea más acerca de Stata/MP en www.stata.com/statamp/, o vea directamente el informe actualizado en www.stata.com/statamp/report.pdf.

Cursos abiertos de capacitación

Cursos	Fechas	Lugar	Costo
Uso Eficiente de Stata: Fundamentos en Manejo de Datos, Análisis y Gráficos	27 y 28 julio	San Francisco, CA	USD 950
	25 y 26 de agosto	Washington, D.C.	USD 950
	5 y 6 de octubre	Nueva York, NY	USD 950
Modelos Mixtos/Multinivel Usando Stata	16 y 17 de septiembre	San Francisco, CA	USD 1.295
	7 y 8 de octubre	Nueva York, NY	USD 1.295
Imputación Múltiple Usando Stata	22 y 23 de septiembre	Washington, D.C.	USD 1.295

Imputación Múltiple Usando Stata NUEVO

Instructor: Yulia Marchenko, Estadística Sénior de StataCorp y principal responsable del desarrollo de los comandos oficiales de Stata para imputación múltiple.

Este curso de dos días cubre el uso de Stata para realizar análisis de imputación múltiple. La imputación múltiple (IM) es una técnica basada en la simulación para trabajar con datos incompletos. El curso ofrecerá una breve introducción a la imputación múltiple y se centrará en cómo llevar a cabo IM en Stata con el comando **mi**. Las tres etapas de IM (imputación, análisis de datos completos e integración) serán discutidas en detalle, con ejemplos en Stata. Se discutirán varias técnicas de imputación, y la discusión se centrará principalmente en la imputación normal multivariada. Además, se presentarán una serie de ejemplos para explicar cómo manejar datos provenientes de imputación múltiple en forma eficiente. Se presentarán también el análisis lineal y logístico de este tipo de datos, así como varios comandos de postestimación.

Temas del Curso

- Panorama general de la imputación múltiple
 - › IM como un procedimiento estadístico
 - › Etapas de la IM: Imputación, análisis e integración
 - › IM en Stata: el conjunto de comandos **mi**
- Imputación
 - › Técnicas de imputación
 - › Imputación univariada
 - › Imputación multivariada
 - › Verificación de la idoneidad de las imputaciones
- Manejo de datos
 - › Almacenamiento de los datos imputados por IM
 - › Importación de datos imputados por IM ya existentes
 - › Verificación de los datos imputados por IM
 - › Manejo de variables (variables pasivas)
 - › La fusión (merge), añadido (append), y la reorganización de datos de IM
 - › Exportación los datos de IM a una aplicación externa
- Estimación
 - › Uso del comando **mi estimate** para realizar las etapas de análisis e integración de la IM en un solo paso
 - › Estimación de funciones lineales y no lineales de los coeficientes
 - › Pruebas de hipótesis lineales y no lineales

Para obtener más información o para inscribirse, visite www.stata.com/training/mi.html.

Modelos Mixtos/Multinivel Usando Stata

Instructor: Roberto G. Gutierrez, Director de Estadística de StataCorp, y principal responsable del desarrollo de los comandos oficiales de Stata para modelos mixtos.

Este curso de dos días es una introducción al uso de Stata para ajustar modelos mixtos/multinivel. Los modelos mixtos contienen ambos, efectos fijos (análogos a los coeficientes en una regresión lineal estándar) y efectos aleatorios, que no son directamente estimados, sino modelados a través de elementos de la matriz de covariancia. Los modelos mixtos pueden contener más de un nivel de efectos aleatorios anidados, y por eso estos modelos también se llaman modelos multinivel o jerárquicos, particularmente en las ciencias sociales. El enfoque de Stata a los modelos lineales mixtos es asignar efectos aleatorios a los paneles independientes, y dentro de éstos pueden definirse paneles anidados para incluir efectos aleatorios en varios niveles.

Temas del curso

- Introducción a los modelos lineales mixtos
- Coeficientes aleatorios y modelos jerárquicos
- Análisis de postestimación
- Modelos no lineales
- Temas avanzados

Por más información o para inscribirse, visite www.stata.com/training/mixed.html.

Uso Eficiente de Stata: Fundamentos en Manejo de Datos, Análisis y Gráficos

Instructor: Bill Rising, Director de Servicios de Educación de StataCorp.

Este curso de dos días 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 mejorar su fluidez, y acceder a trucos y consejos para aumentar su eficiencia en el uso de Stata. Al terminar el curso, usted estará preparado para utilizar eficientemente Stata para efectuar análisis básicos y gráficos. Usted estará capacitado para hacerlo de manera reproducible, lo que simplificará significativamente la realización de trabajos en colaboración, o modificaciones posteriores de su propio análisis. Por último, usted será capaz de documentar su propio conjunto de datos, para sus colaboradores o para usted mismo.

Si usted tiene actualmente acceso a Stata 11, o si está considerando una actualización o la adquisición de esta versión, este curso, sin lugar a duda, lo volverá más diestro en el amplio rango de funcionalidades de Stata.

Temas del curso

- Conceptos básicos de Stata
- Manejo de datos
- Fluidez en el trabajo
- Análisis
- Gráficos

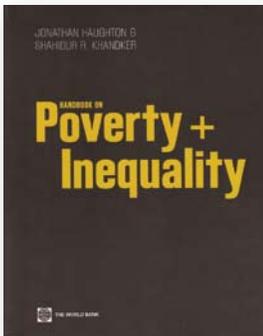
Por más información o para inscribirse, visite www.stata.com/training/eff_stata.html.

Las inscripciones para los cursos de Capacitación son limitadas. Todas las sesiones contarán con computadoras con Stata 11 para uso de los participantes. Todos los cursos de Capacitación se desarrollarán cada uno de los días estipulados de 8:30 a 16:30, y serán dictados en inglés. El costo incluye desayuno continental, almuerzo y refrigerio en la tarde. Se invita a los participantes a traer sus unidades de almacenamiento USB a todas las sesiones. Ésta es la forma más simple y segura de guardar el trabajo de cada sesión.

Para ver las fechas de los próximos cursos, visite www.stata.com/training/public.html.

Novedades de La Librería de Stata

Handbook on Poverty + Inequality



Autores: Jonathan Haughton y Shahidur R. Khandker
 Editorial: World Bank
 Derechos de autor: 2009
 Páginas: 444; cubierta suave
 ISBN-10: 0-8213-7613-6
 ISBN-13: 978-0-8213-7613-3
 Precio: USD 36,00

En el libro *Handbook on Poverty + Inequality*, Jonathan Haughton y Shahidur Khandker proporcionan una introducción a los problemas de definición, medición y análisis de la pobreza y la desigualdad. Si bien este libro se centra más en los temas de pobreza y desigualdad que en los métodos de análisis de datos, algunos de los capítulos describen aspectos importantes de cómo interpretar los resultados de diferentes tipos de análisis de datos. Los autores también proporcionan introducciones breves a los fundamentos de análisis de datos, modelos de regresión, y al uso de datos provenientes de encuestas complejas; además proporcionan un apéndice que presenta de manera sucinta una introducción al uso de Stata orientada a los métodos discutidos en el libro.

Este libro será de utilidad para estudiantes e investigadores que necesiten una introducción a la literatura sobre la pobreza y la desigualdad y que precisen una rápida introducción al análisis de datos. Los investigadores y estudiantes que trabajan en el análisis de datos, probablemente necesiten un libro adicional sobre estadística o econometría, como *Data Analysis Using Stata, Second*

Edition escrito por Ulrich Kohler y Frauke Kreuter, o *Introductory Econometrics: A Modern Approach, Fourth Edition*, escrito por Jeffrey M. Wooldridge.

Para ver la tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite: www.stata.com/bookstore/hopi.html.

Handbook on Impact Evaluation: Quantitative Methods and Practices



Autores: Shahidur R. Khandker, Gayatri B. Koolwal, y Hussain A. Samad
 Editorial: World Bank
 Derechos de autor: 2010
 Páginas: 280; cubierta suave
 ISBN-10: 0-8213-8028-1
 ISBN-13: 978-0-8213-8028-4
 Precio: USD 37,75

En el libro *Handbook on Impact Evaluation: Quantitative Methods and Practices*, Shahidur Khandker, Koolwal Gayatri, y Samad Hussain proporcionan una introducción excelente y no muy técnica a la estimación y la interpretación de los efectos de tratamiento. Si bien este libro está dirigido a profesionales en la economía del desarrollo, también será útil para estudiantes de posgrado e investigadores que necesiten conocer las bases y los conocimientos intuitivos para entender la literatura moderna más técnica. Los autores también proporcionan una breve introducción a Stata orientada a la estimación de los efectos de tratamiento.

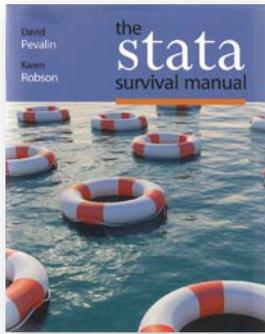
Para ilustrar los aspectos técnicos, Khandker, Koolwal y Samad usan muchos casos de estudio y ejemplos intuitivos tomados tanto de su propio trabajo como de la economía del desarrollo. La capacidad de los autores para presentar información técnica de forma fácil de entender es, definitivamente, uno de los puntos fuertes de este libro.

La amplitud de la cobertura es extraordinaria: luego de distinguir la evaluación cuantitativa del impacto de otras formas de evaluación de programas, los autores proporcionar introducciones al modelo contrafactual, el modelo de asignación al azar, nivelación del grado de propensión (propensity score matching), y varios otros métodos para la selección en no observables.

Los autores no tratan muchos detalles técnicos importantes, pero cada capítulo contiene una buena bibliografía para los lectores que necesiten más información. Los autores proporcionan una introducción de alto nivel que servirá como excelente panorama general o punto de partida para el lector que necesita entender estos temas.

Para ver la tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite: www.stata.com/bookstore/hoie.html.

The Stata Survival Manual



Autores: David Pevalin y
Karen Robson
Editorial: McGraw-Hill
Derechos de autor: 2009
Páginas: 373; cubierta suave
ISBN-10: 0-335-22388-5
ISBN-13: 978-0-335-22388-6
Precio: USD 54,75

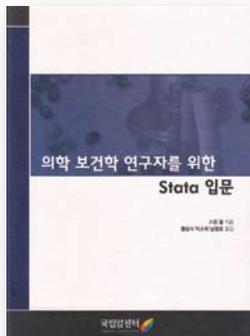
The Stata Survival Manual, escrito por David Pevalin y Robson Karen, es una introducción bien escrita al uso de Stata 10. El estilo es agradable y fluido, y los autores no asumen conocimientos previos de Stata o sofisticación estadística por parte del lector. Tanto el uso de Stata como de las herramientas estadísticas se explican en el libro.

Los autores recorren los aspectos básicos del uso de Stata, comenzando con el uso básico de Stata y continuando con las técnicas más comunes de manejo de datos utilizadas para el análisis, la creación de tablas y gráficos, y la presentación de resultados. Se presta especial atención al trabajo con variables categóricas y a la construcción de escalas a partir de los instrumentos. Las secciones de análisis explican con detalle cómo ajustar interacciones y utilizan gráficos para volverlas accesibles a los lectores sin formación estadística. Cada capítulo comienza con una presentación de nuevas herramientas de Stata y ejemplos simples de su uso. Las herramientas se aplican luego a través de un "Ejercicio Demostración", un ejemplo que continúa a lo largo de todo el libro. Así, el lector puede aprender nuevas herramientas en un entorno sencillo y ver su uso en el análisis de un conjunto de datos de la vida real de principio a fin.

En varios puntos del libro, especialmente en los capítulos centrados en el manejo de datos, los autores señalan diferencias entre Stata e IBM SPSS Statistics para aquellos que están haciendo la transición de IBM SPSS Statistics a Stata. Si bien los autores se centran en hacer uso de los archivos .do (por la reproducibilidad), también muestran cómo utilizar los menús y cuadros de diálogo para los lectores acostumbrados a trabajar de esta forma.

Para ver la tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite: www.stata.com/bookstore/ssm.html.

An Introduction to Stata for Health Researchers, 2nd Edition (Coreano)



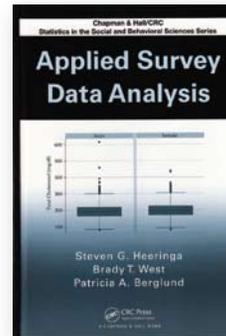
Autor: Svend Juul
Editorial: JasonTG Co., Seoul
Derechos de autor: 2009
Páginas: 450; cubierta suave
ISBN-10: 89-92864-03-5
ISBN-13: 978-89-92864-03-9
Precio: USD 52,00

Traducido por JasonTG, distribuidor de Stata en Corea del Sur, *An Introduction to Stata for Health Researchers, 2nd Edition* ofrece a los lectores de habla

Coreana una visión general de las capacidades de Stata en materia de manejo de datos, gráficos y estadísticas. Este libro se distingue por su cuidadosa atención a los detalles, y no sólo enseña a usar Stata, sino también consolida los hábitos que permiten hacer análisis reproducibles, necesarios en esta área de investigación. El libro se basa en la suposición de que el lector tiene algún conocimiento básico de estadística, pero ningún conocimiento de Stata. Juul edifica las habilidades del lector de la misma forma que un constructor lo haría con un edificio, comenzando con una base sólida en Stata, enmarcando una estructura general que permita llevar a cabo trabajos productivos, y finalmente introduciendo los detalles propios de varios tipos de análisis estadísticos.

Para ver la tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite: www.stata.com/bookstore/ishr2_korean.html.

Applied Survey Data Analysis



Autores: Steven G. Heeringa, Brady T. West, and Patricia A. Berglund
Editorial: CRC Press/Taylor & Francis
Derechos de autor: 2010
Páginas: 462; cubierta rígida
ISBN-10: 1-4200-8066-0
ISBN-13: 978-1-4200-8066-7
Precio: USD 64,00

Applied Survey Data Analysis es un texto de nivel intermedio, basado en ejemplos, sobre los métodos actuales para datos provenientes de encuestas complejas. Será de interés para investigadores de todas las disciplinas que trabajen con datos de encuestas y tengan conocimientos básicos de la metodología estadística para datos estándar (no de encuestas).

Los autores comienzan con un poco de historia y discutiendo algunos conjuntos de datos de encuesta ampliamente utilizados, como los datos de la Encuesta Nacional de Salud y Nutrición (NHANES) de los EE.UU.. A continuación, tratan los conceptos básicos de datos de encuestas: planes de muestreo, pesos, agrupación por conglomerado, pre estratificación y post estratificación, efectos de diseño, y muestras de varias etapas. Luego pasan a discutir los tipos de estimadores de varianzas: estimador por linealización de Taylor, de "jackknife", y de repeticiones repetidas balanceadas.

Las secciones centrales del texto proporcionan una cobertura en profundidad de los tipos de análisis que se pueden realizar con datos de encuestas, incluyendo medias y proporciones, correlaciones, tablas, regresión lineal, regresiones con variables dependiente limitadas (incluyendo logit y de Poisson), y análisis de supervivencia (Incluyendo regresión de Cox). Dos capítulos finales están dedicados a temas avanzados, como la imputación múltiple, el análisis bayesiano, y los modelos multinivel. En el apéndice se ofrece una visión general de algunos de los paquetes de software estadístico más utilizados, incluyendo Stata.

Para ver la tabla de contenidos, u obtener la información necesaria para ordenar este libro, visite: www.stata.com/bookstore/asda.html.

Reuniones de Grupos de Usuarios Reserve la fecha

Encuentro del Grupo de Usuarios de Stata de Alemania

Sophia Rabe-Hereth estará a cargo de la Conferencia Magistral.

Fecha: 25 de junio de 2010
Lugar: Berlin Graduate School of Social Sciences
Luisenstraße 56, House 1
10117 Berlin-Mitte, Germany
Costo: profesionales €35; estudiantes €15
Detalles: www.stata.com/meeting/germany10/

Encuentro del Grupo de Usuarios de Stata del Reino Unido

Fechas: 9 y 10 de septiembre de 2010
Lugar: London School of Hygiene and Tropical Medicine
Keppel Street, London WC1E 7HT, UK
Costo: Dos días: profesionales £94,00; estudiantes £64,63
Un día: profesionales £64,63; estudiantes £47,00
Detalles: www.stata.com/meeting/uk10/

Encuentro del Grupo de Usuarios de Stata de España

Fecha: 14 de septiembre de 2010
Lugar: Universidad Carlos III de Madrid
C/ Madrid 126, 28903 Getafe, Madrid, España
Costo: profesionales €30; estudiantes €30
Fecha límite para resúmenes: 14 de junio de 2010
Detalles: www.stata.com/meeting/spain10/

Encuentro del Grupo de Usuarios de Stata de Portugal

Fecha: 17 de septiembre de 2010
Lugar: University of Minho
Gualtar University Campus, 4710-057 Braga, Portugal
Costo: profesionales €45; estudiantes €20
Fecha límite para resúmenes: 14 de junio de 2010
Detalles: www.stata.com/meeting/portugal10/

Encuentro del Grupo de Usuarios de Stata de Italia

Dates: 11 y 12 de noviembre de 2010
Lugar: Grand Hotel Baglioni
Via Indipendenza, 8
40121 Bologna, Italia
Costo: €90 día 1 solamente; €375 día 1 + un curso
Fecha límite para resúmenes: 30 de agosto de 2010
Detalles: www.stata.com/meeting/italy10/

Próximas fechas de NetCourse®

Para inscribirse, visite www.stata.com/netcourse/.

NC101: Introducción a Stata

Introducción al uso interactivo de Stata.

Fechas: 9 de julio al 20 de agosto de 2010
Fecha límite de inscripción: 8 de julio de 2010
Costo: USD 95
Detalles: www.stata.com/netcourse/nc101.html

NC151: Introducción a la Programación en Stata

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: 9 de julio al 20 de agosto de 2010
Fecha límite de inscripción: 8 de julio de 2010
Costo: USD 125
Detalles: www.stata.com/netcourse/nc151.html

NC152: Programación Avanzada en Stata

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 realizar el análisis de 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 usar el comando `syntax`, a manejar y procesar resultados, procesar datos por grupo, y más.

Fechas: 8 de octubre al 26 de noviembre de 2010
Fecha límite de inscripción: 7 de octubre de 2010
Costo: USD 150
Detalles: www.stata.com/netcourse/nc152.html

NC461: Introducción a las Series de Tiempo Univariadas en Stata

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

Fechas: 8 de octubre al 26 de noviembre de 2010
Fecha límite de inscripción: 7 de octubre de 2010
Costo: USD 295
Detalles: www.stata.com/netcourse/nc461.html

Conferencia de Stata en Boston 2010

Fechas: 15 y 16 de julio de 2010

Lugar: Omni Parker House, Boston
60 School Street
Boston, MA 02108



Costo: Un día USD 125, estudiantes USD 50
Dos días USD 195, estudiantes USD 75

Inscripciones: www.stata.com/meeting/boston10/

La Conferencia de Stata en Boston 2010 se celebrará los días 15 y 16 de julio en el Hotel Omni Parker House, ubicado en el centro de Boston cerca de Boston Common y Park Street T.

Programa

Jueves 15 de julio

Regression for nonnegative skewed dependent variables

Austin Nichols, Urban Institute

Margins and the Tao of interaction

Phil Ender, UCLA Statistical Consulting Group

To the vector belong the spoils: Circular statistics in Stata

Nicholas J. Cox, Durham University

System for formatting tables

John Gallup, Portland State University

Hunting for genes with longitudinal phenotype data using Stata

Chuck Huber, Texas A&M Health Science Center School of Rural Public Health

Bayesian bivariate diagnostic meta-analysis via R-INLA

Ben Adarkwa Dwamena, University of Michigan and VA Ann Arbor Health Systems

Storing, analyzing, and presenting Stata output

Julian Reif, University of Chicago

An efficient data envelopment analysis with a large dataset in Stata

Choonjoo Lee, Korea National Defense University

Competing-risks regression in Stata 11

Roberto G. Gutierrez, StataCorp

Structural equation models with latent variables

Stas Kolenikov

Viernes 16 de julio

Multiple imputation using Stata's mi command

Yulia Marchenko, StataCorp

CEM: Coarsened exact matching in Stata

Matthew Blackwell, Harvard University

Evaluating one-way and two-way cluster-robust covariance matrix estimates

Christopher F. Baum, Boston College

Bootstrap LM test for the Box-Cox tobit model

David Vincent, Hewlett Packard

Teaching a statistical program in emergency medicine research rotations: Command-driven or click-driven?

Muhammad Waseem, Lincoln Medical and Mental Health Center

Report to users

William Gould, StataCorp

Wishes and grumbles: User feedback and Q&A

Organizadores científicos

Elizabeth Allred, *Harvard School of Public Health*
Christopher F. Baum (chair), *Boston College*
Amresh Hanchate, *Boston University*
Marcello Pagano, *Harvard School of Public Health*

Organizadores de logística

Chris Farrar, *StataCorp*
Gretchen Farrar, *StataCorp*
Sarah Marrs, *StataCorp*

Cómo contactarnos

StataCorp
4905 Lakeway Dr.
College Station, TX 77845
USA

Tel 979-696-4600
Fax 979-696-4601
Email service@stata.com
Web www.stata.com

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



Derechos de Autor 2010 por StataCorp LP.

Para localizar un distribuidor de Stata cerca de usted, visite www.stata.com/worldwide/.