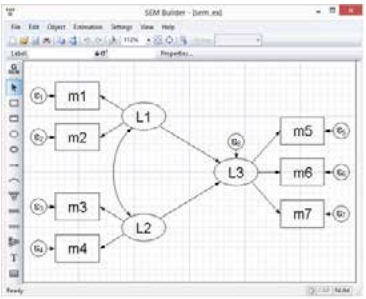
Software Shop®

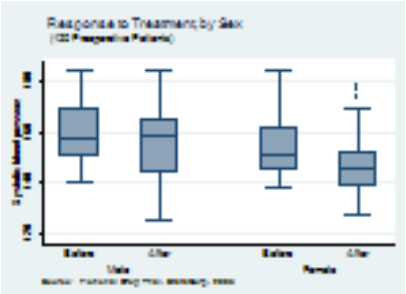
STATA® release 15

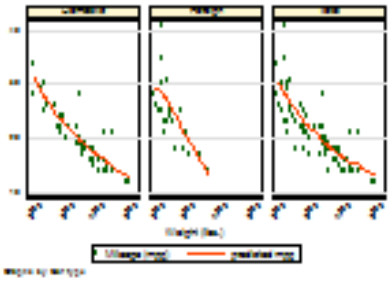
Stata es un software de Estadística completo e integrado que provee todo lo que necesita para el Análisis de Datos, Gestión de Datos y Gráficos.

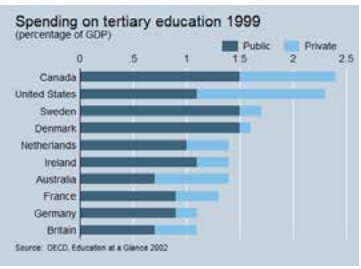
¿Por qué utilizar Stata?

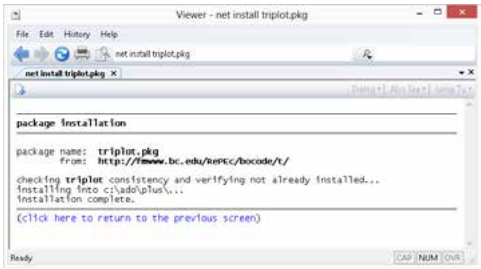
- No se vende por módulos lo que significa que puede tener acceso a todo lo que necesita en un solo paquete.
- Puede adquirirse a través de licencias perpetuas o licencias anuales.
- Es rápido, preciso y de fácil uso.
- Interfaz gráfica de point-and-click y sintaxis de comandos intuitiva.
- Cuenta con una suite amplia de características y cientos de herramientas estadísticas.
- Completa gama de herramientas para la Gestión de Datos.
- Gráficos con Calidad de Publicación.
- Programación de Matrices.
- Multiplataforma: Windows, Mac, Linux/Unix.
- Soporte técnico y recursos para el aprendizaje.
- Ampliamente utilizado. Stata se distribuye en más de 200 países y es utilizado por cientos de miles de profesionales en muchos campos de investigación.

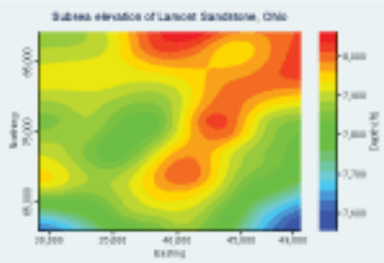












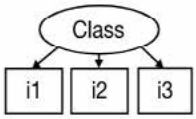
Novedades en Stata 15

ERM=Endogeneity +Selection +Treatment

ERM = Endogeneidad + Selección + Tratamiento

Combina covariables endógenas, selección de muestras, y tratamientos endógenos en modelos con respuestas continuas, binarias, ordenadas y censuradas.

Latent class analysis (LCA)



Análisis de Clase Latente (LCA)

Descubra y entienda las agrupaciones no observables en sus datos. Utilice los modelos de clasificación LCA para encontrar:

- Cuántos grupos tiene
- Quién se encuentra en los grupos
- Qué hace que los grupos sean distintos

bayes: logistic ... and 44 more

Bayes…

Escriba bayes en frente de cualquiera de los 45 comandos de estimación de Stata para ajustar modelos de regresión Bayesianos

Markdown & dynamic documents

Documentos dinámicos

- Cree páginas web desde Stata
- Intercale texto, regresiones, resultados, gráficos, etc.
- Observe los cambios en los datos o comandos realizados automáticamente en su página web.

Linearized DSGEs

Modelos Linealizados Dinámicos Estocásticos de Equilibrio General (DSGEs)

Escriba su modelo en forma algebraica simple. Stata hace el resto: resuelve el modelo, estima los parámetros, estima matrices de transición (con intervalos de confianza), estima y grafica funciones impulso-respuesta, y realiza pronósticos.

Threshold regression

Regresión segmentada (Threshold regression)

Encuentre umbrales y estime parámetros.

Panel-data tobit with random coefficients

Datos Panel Tobit con coeficientes aleatorios

Stata tiene estimadores para efectos aleatorios en datos panel. Ahora puede tener coeficientes aleatorios también.

Search, browse, and import FRED data

Busque, navegue e importe datos FRED (Datos Económicos de la Reserva Federal)

La Reserva Federal de St. Louis tiene disponibles más de 470,000 series de tiempo económicas y financieras internacionales y de Estados Unidos. Ahora puede buscar e importar fácilmente esta información.

Software Shop®

Multilevel regression for interval-measured outcomes

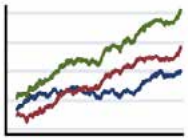
Regresiones Multinivel para resultados medidos por intervalos

Multilevel tobit regression for censored outcomes

Regresión Tobit Multinivel para respuestas censuradas

- Censura izquierda, derecha o conjunta.
- Censura que varía por observación
- Realice inferencia sobre las salidas censuradas o no censuradas
- Errores estándar robustos y agrupados
- Soporte de datos de encuesta

Panel-data cointegration tests



Pruebas de cointegración para datos panel

Pruebas:

- Kao
- Pedroni
- Westerlund

Un total de 9 variantes para las pruebas.

Tests for multiple breaks in time series

Pruebas para múltiples puntos de quiebre en series de tiempo

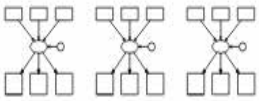
- Prueba de suma acumulada (CUSUM) para estabilidad de los parámetros
- Gráficos con intervalos de confianza

Tests for multiple breaks in time series

Pruebas para múltiples puntos de quiebre en series de tiempo

- Prueba de suma acumulada (CUSUM) para estabilidad de los parámetros
- Gráficos con intervalos de confianza

Multiple-group generalized SEM



Modelos de Ecuaciones Estructurales Generalizados (GSEM) de grupo múltiple

Los Modelos de Ecuaciones Estructurales Generalizados ahora soportan análisis de múltiples grupos. Fácilmente especifique grupos y pruebas de invariancia sobre los parámetros entre los grupos. GSEM incluye:

- Respuestas continuas, binarias, ordinales, de conteo, categóricas y de supervivencia o delos multinivel

ICD-10-CM/PCS

ICD-10-MC/PCS (sistema internacional de clasificación médica)

STATA®

Software Shop®

Power for cluster randomized designs

Potencia para diseños aleatorios agrupados

Análisis de potencia para comparar:

- Una o dos medias muestrales
- Una o dos proporciones muestrales
- Curvas de supervivencia para dos muestras

Cuando aleatoriza conglomerados en lugar de individuos

Power for linear regression models

Potencia para modelos de regresión lineales

- Resolución para:
  - Potencia
  - Tamaño de muestra
  - Tamaño del efecto
- Listas de especificación para:
  - Valores alfa
  - Niveles de potencia
  - Valores beta
  - Tamaños muestrales
  - Y más
- Tablas y gráficos automáticos

Poisson models with sample selection

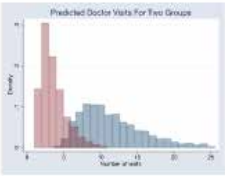
Modelos de Poisson con selección de muestra

More in panel data

Más en datos panel

- Modelos no lineales con efectos aleatorios, incluyendo coeficientes aleatorios
- Modelos Bayesianos de datos panel
- Regresión por intervalos con interceptos aleatorios y coeficientes aleatorios

More in graphics



Más en gráficos

- Transparencias en gráficos
- Exportación SVG

More in statistics

Más en estadística

- Modelos de supervivencia bayesianos
- Probit ordenado sin inflación
- Agregue sus propios métodos de potencia y tamaño de muestra
- Modelos bayesianos de selección de muestra
- Y mucho más

Stata

Software Shop®

More in the interface

Más en la interfase

- Mejoras en el editor de do-files
- Interfaz

And, even more

Y, mucho más

- Generador de números aleatorios
- Mejoras para los plugins de Java

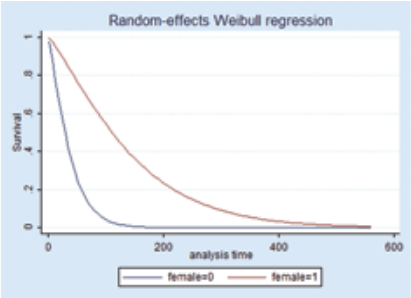
STATA®

CARACTERÍSTICAS Y FUNCIONALIDADES DE STATA

A continuación se referencian de manera detallada más no exhaustiva las características y funcionalidades de Stata.

• Interfaz de usuario gráfica

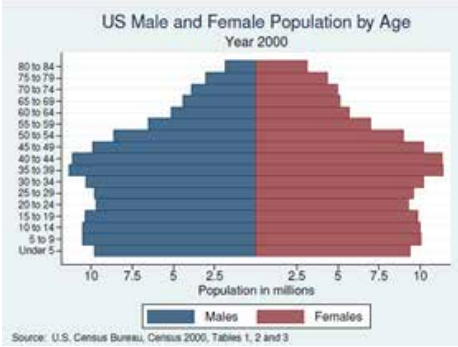
>Menús y diálogos para todas las características
>Editor de datos
>Administrador de variables
>Editor de gráficos
>Project Manager
>Editor Do-file
>Vista previa (clipboard)
>Múltiples ajustes de preferencia



• Gráficos

>Líneas
>Dispersión
>Barras
>Pie
>Máximo-mínimo
>Contorno
>GUI editor
>Gráficos de diagnóstico de regresión
>Gráficos de supervivencia
>Suavizadores no paramétricos
>Distribuciones Q-Q
>Geo-referenciación*

* Requiere de mapas con formatos ESRI shapefiles o MapInfo Interchange Format. Disponible a través de del comando spmap disponible en un download gratuito



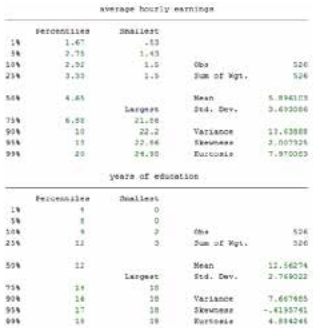
STATA®

• Funciones

- >Estadísticas
- >Números aleatorios
- >Matemáticas
- >Texto
- >Fecha y tiempo

• Estadísticas básicas

- >Resúmenes
- >Tabulaciones cruzadas
- >Correlaciones
- >Pruebas t
- >Pruebas de igualdad de varianza
- >Pruebas de proporciones
- >Intervalos de confianza
- >Variables de factor



average hourly earnings

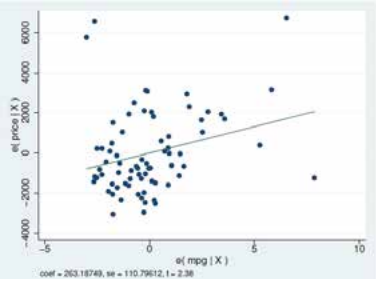
	Percentiles	Smallest		
1%	1.67	.53		
5%	2.75	1.43		
10%	2.92	1.5	Obs	526
25%	3.33	1.5	Sum of Wgt.	526
50%	4.65		Mean	5.896103
		Largest	Std. Dev.	3.693086
75%	6.88	21.86		
90%	10	22.2	Variance	13.63888
95%	13	22.86	Skewness	2.007926
99%	20	24.98	Kurtosis	7.970063

years of education

	Percentiles	Smallest		
1%	4	0		
5%	8	0		
10%	9	2	Obs	526
25%	12	3	Sum of Wgt.	526
50%	12		Mean	12.56274
		Largest	Std. Dev.	2.769022
75%	14	16		
90%	16	18	Variance	7.667985
95%	17	18	Skewness	-.4195741
99%	18	18	Kurtosis	4.894246

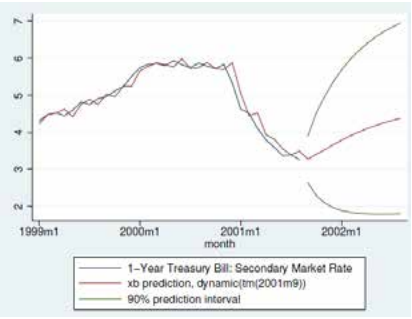
• Modelos lineales

- >Regresión
- >Respuestas censuradas
- >Regresoras endógenas
- >Bootstrap, jackknife
- >Regresiones robustas y robustas en cluster
- >Variables instrumentales
- >Mínimos cuadrados tri-etápicos
- >Restricciones
- >Regresión cuantílica
- >Mínimos cuadrados generalizados



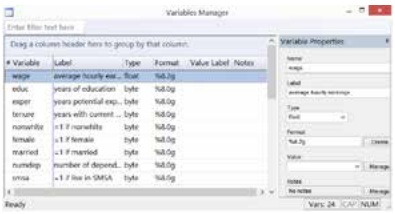
• Series de tiempo

- >ARIMA
- >ARFIMA
- >ARCH/GARCH
- >VAR
- >VECM
- >GARCH multivariado
- >Modelos de componentes no observables
- >Factores dinámicos
- >Modelos de espacio-estado (space-state models)
- >Calendario de negocios
- >Correlogramas
- >Periodogramas
- >Pronósticos
- >Funciones de impulso-respuesta
- >Pruebas de raíz unitaria
- >Filtros y suavizadores
- >Estimaciones "rolling" y recursivas



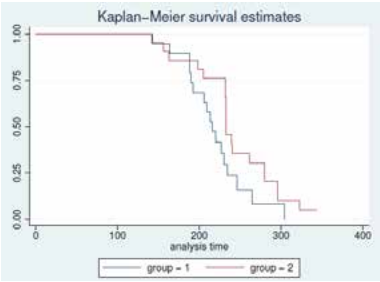
• Gestión de datos

>Transformaciones de datos
>Match-merge-append
>Importación y exportación data
>ODBC
-Importar de Oracle, SQL, Access, Excel, MySQL y DB2
-Exportar datos a tablas ODBC nuevas o existentes
-Ejecución personalizada de comandos SQL de manera individual o en batches
-Personalización de strings de conexión ODBC
-Soporte para ODBC
-Soporte para VARCHARs/CLOBs and BLOBs
>SQL
>XML
>Ordenamiento
>Transposición
>Rotulación
>Haver Analytics database



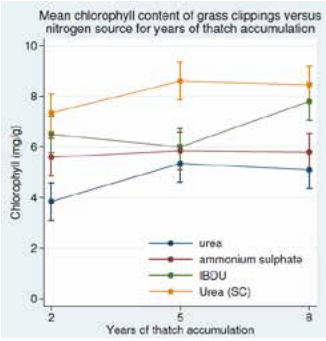
• Datos longitudinales y de panel

>Efectos fijos y aleatorios con errores estándar robustos
>Modelos mixtos lineales
>Probit de efectos aleatorios
>GEE (estimación de ecuaciones generalizada)
>Poisson de efectos fijos y aleatorios
>Modelos dinámicos de datos panel
>Variables instrumentales
>Pruebas de raíz unitaria para paneles



• Análisis de supervivencia

>Estimadores de Kaplan-Meier y Nelson-Aalen
>Regresión Cox
>Modelos paramétricos
>Riesgo competitivo (competing risk)
>Hazards
>Covariables variantes en el tiempo
>Censura izquierda y derecha
>Modelos Weibull, Exponencial y Gompertz

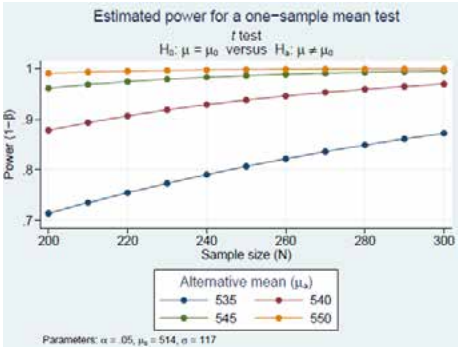


• Modelos multinivel y de efectos mixtos

>Continuos, binarios y de conteo
>Modelos de 2, 3 y múltiples niveles
>Modelos lineales generalizados
>Interceptos aleatorios
>Pendientes aleatorias
>Efectos aleatorios cruzados
>BLUPs de los efectos y los valores estimados
>Modelos jerárquicos
>Estructuras residual-error
>Soporte de información de encuestas en modelos lineales

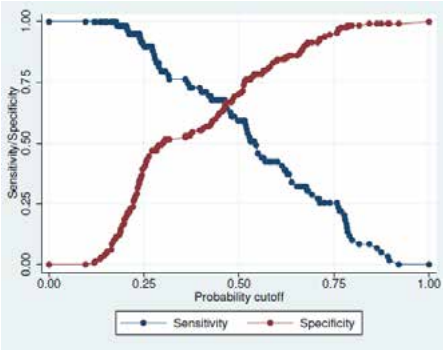
• Potencia y tamaño de la muestra

- >Potencia
- >Tamaño de la muestra
- >Efecto del tamaño
- >Efecto mínimo detectable
- >Medias
- >Proporciones
- >Varianzas
- >Correlaciones
- >ANOVA
- >Estudio caso-control
- >Estudios de cortes
- >Análisis de supervivencia
- >Diseños balanceados y no balanceados
- >Resultados en tablas o gráficos



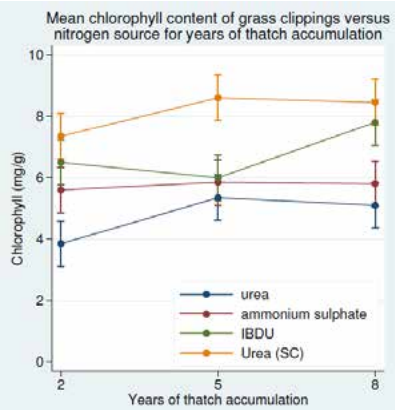
• Modelos de respuesta binaria, limitada y de conteo

- >Logit
- >Probit
- >Tobit
- >Poisson y binomial negativa
- >Logístico condicional, multinomial, anidado, ordenada, ordenado y estereotípico
- >Probit multinomial
- >Modelos de conteo truncados
- >Modelos de selección
- >Efectos marginales



• Efectos de tratamiento

- >IPW (ponderación de probabilidad inversa)
- >Probit
- >Métodos robustos dobles
- >Propensity score matching
- >Ajustes de regresión
- >Matching de covariables
- >Tratamientos multinivel
- >ATE (efectos promedio del tratamiento)
- >ATETs (efectos promedio del tratamiento en el tratado)
- >POMs (potential-outcome means)

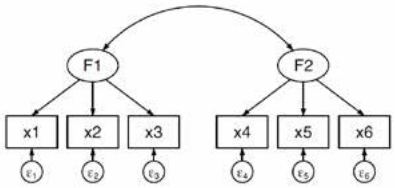


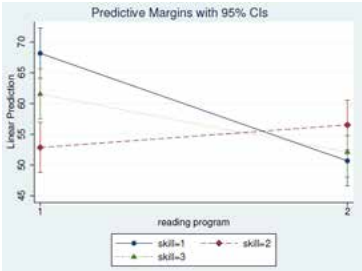
• Modelos lineales generalizados (GLMs)

- >10 funciones de vinculación
- >Vínculos definidos por el usuario
- >Estimación ML e IRLS
- >9 estimadores de varianza
- >7 residuos

• Modelación de ecuaciones estructurales (SEM)

- >Constructor gráfico de diagramas
- >Estimaciones estandarizadas y no estandarizadas
- >Índices de modificación
- >Efectos directos e indirectos
- >Respuesta continuas, binarias, de conteo y ordinales
- >Modelos multinivel
- >Interceptos y pendientes aleatorias
- >Factors scores
- >Predictores empíricos de Bayes y otras
- >Grupos y pruebas de invariancia
- >Bondad de ajuste
- >Datos MAR
- >Datos correlacionados





• Otros métodos estadísticos

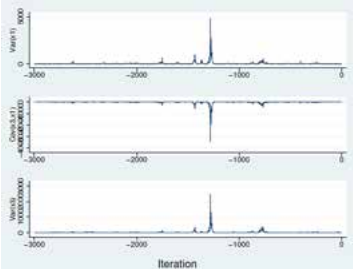
- >Medida de Kappa para acuerdos
- >Alfa de Cronbach
- >Regresión stepwise
- >Pruebas de normalidad

• ANOVA/MANOVA

- >Diseños balanceados y no balanceados
- >Diseños factoriales, anidados y mixtos
- >Medidas repetidas
- >Medias marginales
- >Contrastes

• Imputación múltiple

- >9 métodos univariados de imputación
- >Imputación normal multivariada
- >Ecuaciones encadenadas
- >Exploración de patrones de pérdida
- >Gestión de bases de datos imputadas
- >Ajuste de modelos y resultados agrupados
- >Parámetros de transformación
- >Pruebas conjuntas de las estimaciones de los parámetros
- >Predicciones



• Estadísticas exactas

- >Regresiones Logística y de Poisson exactas
- >Estadísticas exactas para caso-control
- >Pruebas binomiales
- >Prueba excita de Fisher para tablas r x c

• Pruebas, predicciones y efectos

- >Tests de Wald
- >LR tests
- >Combinaciones lineales y no lineales
- >Predicciones y predicciones generalizadas
- >Medias marginales
- >Medias ajustadas
- >Efectos marginales y parciales
- >Modelos de pronóstico
- >Test de Hausman

```
. svydescribe
Survey: Describing stage 1 sampling units
      pweight: finalwgt
          VCE: linearized
  Single unit: missing
     Strata 1: stratid
         SU 1: psuid
        FPC 1: <zero>
```

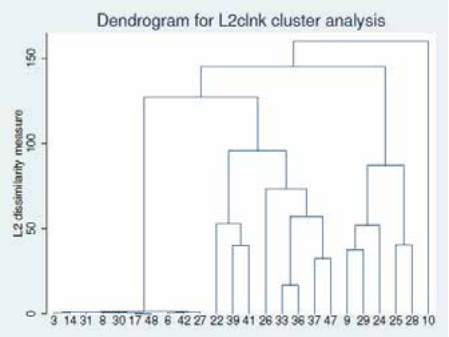
			#Obs per Unit		
Stratum	#Units	#Obs	min	mean	max
1	2	380	165	190.0	215
2	2	185	67	92.5	118
3	2	348	149	174.0	199
(output omitted)					
17	2	393	180	196.5	213
18	2	359	144	179.5	215
20	2	285	125	142.5	160
21	2	214	102	107.0	112
(output omitted)					
31	2	308	143	154.0	165
32	2	450	211	225.0	239
31	62	10351	67	167.0	288

• Métodos de encuesta

- >Diseños multietápicos
- >Estimación de varianza Bootstrap, BRR, jacknife, linealizada y SDR
- >Post-estratificación
- >DEFF
- >Márgenes predictivos
- >Medias, proporciones, razones y totales
- >Tablas de resumen
- >Regresión de variables instrumentales, Probit y Cox

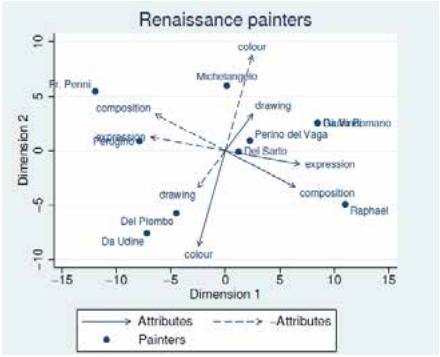
• Contrastes, comparaciones pareadas y márgenes

- >Comparación de medias, interceptos o pendientes
- >Comparación respecto una categoría de referencia
- >Categoría adyacente, gran media, etc.
- >Polinomios ortogonales
- >Ajustes de comparación múltiple
- >Gráficos de medias estimadas y contrastes
- >Gráficos de interacción



• Análisis de conglomerados (cluster analysis)

- >Conglomerados jerárquicos
- >Agrupamiento no jerárquico para Kmeans y kmedian
- >Dendogramas
- >Reglas de parada
- >Análisis extensibles al usuario

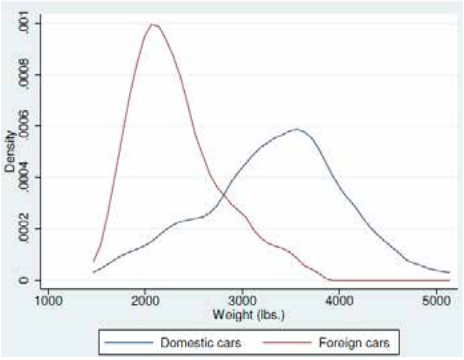


• Métodos multivariados

- >Análisis factorial
- >Componentes principales
- >Análisis discriminante
- >Rotación
- >Escalamiento multidimensional
- >Biplots
- >Dendogramas
- >Análisis extensibles al usuario

• GMM y regresión no lineal

>Análisis factorial
>GMM (método generalizado de momentos)
>Regresión no lineal

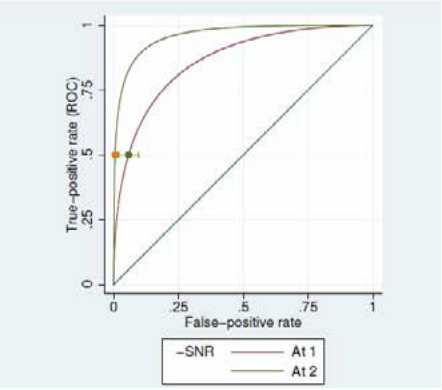


• Máxima verosimilitud simple

>Uso de expresiones simples para especificar verosimilitud
>No se requiere programación
>Datos de encuestas
>Estimadores matriciales

• Métodos no paramétricos

>Wilcoxon-Mann-Whitney
>Rango de signos de Wilcoxon
>Kruskal-Wallis test
>Correlaciones de Spearman y Kendall
>Pruebas de Kolmogorov-Smirnov
>Intervalos de confianza binomiales exactos
>Datos de supervivencia
>Análisis ROC
>Suavizamiento
>Bootstrapping



• Remuestreo y métodos de simulación

>Bootstrap
>Jacknife
>Simulación de Monte Carlo
>Pruebas de permutación

• Epidemiología

>Estandarización de tasas
>Caso-control
>Cohortes
>Matched caso-control
>Mantel-Haenszel
>Farmacocinética
>Análisis ROC
>ICD-9C-M

• Característica de programación

>Adición de nuevos comandos
>Command scripting
>Programación orientada a objetos
>Menú y cuadros de diálogo para programación
>Project Manager
>Plugins

```
. forvalues x = 1(1)4 {
  2.        if mod(`x',2) {
  3.                display "`x' is odd"
  4.                continue
  5.        }
  6.        display "`x' is even"
  7. }
1 is odd
2 is even
3 is odd
4 is even
```

• Programación de matrices – Mata

>Sesiones interactivas
>Desarrollo de proyectos de gran escala
>Optimización
>Inversión de matrices
>Descomposiciones
>Eigenvalores y eigenvectores
>Motor LAPACK
>Números reales y complejos
>Matrices de texto
>Interface a bases de datos de Stata y matrices
>Derivaciones numéricas
>Programación orientada a objetos

• Capacidades de internet

>Habilidad para la instalación de nuevos comandos
>Actualización web
>Noticias de Stata recientes
>Compartir bases de datos y programas
>Ayuda

• Accesibilidad

>Acceso a personas con discapacidades

• Documentación

>27 manuales (Más de 14,000 páginas, todos en PDF e interactivos)
>Navegación por hipervínculos
>Miles de ejemplos trabajados
>Métodos y fórmulas
>Referencias

¿Quiénes utilizan Stata?

Por más de 30 años, Stata ha proveído a cientos de miles de investigadores con todo lo necesario para la estadística, gráficos y gestión de datos. Algunas áreas con un uso muy importante de Stata son:

• Ciencias comportamentales

Tendencias sociales, neurociencia cognitiva, aprendizaje, métodos psicométricos.

- **Educación**
  Desarrollo de nuevas pruebas, investigación en aprendizaje y desarrollo, efectividad de los profesores.

- **Medicina y bioestadística**
  Investigación médica básica, ensayos clínicos, salud pública, cáncer, VIH, estudios clínicos observacionales,

- **Políticas públicas**
  Investigación en comercio, desarrollo internacional, defensa, salud, energía/ambiente, educación, leyes, planeación urbana.

- **Epidemiología**
  Investigación en enfermedades infecciosas, exposición a patógenos, estudios de asociación genómica.

- **Ciencias políticas**
  Investigación en descentralización, comportamiento electoral, distritos legislativos, conflicto interno.

- **Sociología Epidemiología**
  Investigación en salud, raza/etnicidad, familia, género, inequidad, demografía

- **Economía**
  Investigación en selección educativa, salario mínimo, PIB, tendencia en cotizaciones.

- **Finanzas, negocios y mercadeo**
  Investigación en valoración de activos, dinámica del mercado de capitales, gestión de valor al cliente, comportamiento del consumidor y de la firma, branding.

Especificaciones

	Stata/MP	Stata/SE	Stata/IC
# observaciones	1.099.511.627.775	2.147.483.647	2.147.483.647
# de variables	120.000	32.767	2.048
Máximo # de variables en el lado derecho (modelos)	10.998	10.998	798
# de caracteres en un comando	15.480.216	4.227.159	264.408

Velocidad.

Información detallada sobre el rendimiento de cada una de las versiones puede encontrarse en: http://www.stata.com/statamp/report.pdf
Sin embargo, en resumen puede mencionarse que:

* En computadores con procesadores de doble núcleo, Stata/MP corre los comandos en un 71% del tiempo requerido por Stata/SE y los comandos estadísticos en un 59% del tiempo aprox.
* En computadores de procesadores de cuatro núcleos, Stata/MP corre los comandos en un 50% del tiempo requerido por Stata/SE y los comandos estadísticos en un 35% del tiempo aprox.
* Stata MP soporta hasta 64 núcleos.
* Stata/IC es más lento que Stata/SE, pero las diferencias sólo son notables cuando se procesan bases de datos que pongan en el límite a la versión IC (ver recuadro sobre capacidades en cuanto a número de variables permitidas)

Sistemas operativos compatibles

Plataformas

Stata para Windows

Windows 10
Windows 8
Windows 7
Windows Vista
Windows Server 2016, 2012, 2008, 2003
Stata para MacOS
Stata para Mac requiere de procesadores de 64-bit Intel® (Core™2 Duo o mejores) con Mac OS X 10.9 o más nuevo
Stata para Unix
Cualquiera de 64-bit (x86-64 o compatible) o 32-bit (x86 o compatible) con Linux

Requerimientos de hardware

Stata/MP requiere 4GB de memoria
Stata/SE requiere 2GB de memoria
Stata/IC requiere 1GB de memoria
Mínimo 1GB de espacio en disco
Stata para Unix requiere de una tarjeta de video que pueda visualizar miles de colores o más (16-bit o 24-bit color)

¿Qué equipo debo usar para ejecutar Stata / MP?

Stata / MP es el Stata que puede realizar el multiprocesamiento simétrico (SMP) en un equipo con varios procesadores o núcleos. Aunque licencias Stata / MP no son específicas de la plataforma, los sistemas operativos siguientes son compatibles:

Windows (64 bits x86-64)
Windows (32-bit)
Mac (64-bit Intel)
Linux (64 bits x86-64)
Linux (32 bits)
Solaris (64 bits SPARC)
Solaris (64 bits x86-64)
El procesador de una computadora a menudo se llama la unidad central de procesamiento, o CPU. Las computadoras pueden tener múltiples CPUs, y Stata / MP se aprovecha de este diseño para procesar datos más rápido.

También están disponibles los equipos con procesadores que contienen múltiples núcleos. Un núcleo es el motor de cálculo de una CPU, que consta de unidades de procesamiento de enteros y de coma flotante. Así, un ordenador con un procesador de doble núcleo es similar a un ordenador con dos procesadores. De hecho, Stata / MP no distingue entre CPUs y núcleos; que está diseñado para utilizar sistemas con hasta 64 núcleos, ya sean de un solo núcleo (uno por CPU), de doble núcleo (dos por CPU) o varios núcleos (muchos por CPU).

La mayoría de los fabricantes venden equipos multiprocesador como servidores. Así, su elección es comprar ya sea una computadora de escritorio con una CPU de varios núcleos o un equipo servidor con varios procesadores, cada uno de los cuales podría o no podría ser de varios núcleos. Tener más procesadores / núcleos significa que Stata / MP se ejecutará más rápido.

Sea consciente del término "hyperthreaded", sin embargo. Stata / MP corre más rápido en procesadores con hyperthreading, pero no tan rápido como lo haría si tuviera núcleos completos en lugar de hyperthreads. Las computadoras con procesadores múltiples hyperthreaded son adecuados para Stata / MP. El número de procesadores reales es el factor crítico.

Ambos chips apropiados para Stata / MP Intel y AMD de fabricación. Cualquiera de ejecutar un sistema operativo Windows moderno (Vista, 7, o más reciente) o ejecutar Linux.

Ordenadores de doble núcleo basados en Intel Mac son apropiadas para Stata / MP. Stata / MP es compatible con OS X 10.6.8 o posterior.

Ver el informe de ejecución Stata / MP en el link: http://www.stata.com/statamp/report.pdf