

Introducción al estudio de la desigualdad social a partir del programa STATA.

Jose Rodríguez de la Fuente.

Cita:

Jose Rodríguez de la Fuente (2017). *Introducción al estudio de la desigualdad social a partir del programa STATA*. Material Didáctico Sistematizado.

Dirección estable: <https://www.aacademica.org/joserodriguez/62>

ARK: <https://n2t.net/ark:/13683/pq7B/ezD>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

Introducción al estudio de la desigualdad social a partir del programa STATA

JOSE RODRÍGUEZ DE LA FUENTE (IIGG – UBA –
CONICET – INCASI)

IV SEMINARIO INTERNACIONAL DE DESIGUALDAD Y
MOVILIDAD SOCIAL EN AMÉRICA LATINA



Propuesta del taller

- Introducir características básicas y centrales del programa STATA
- Cálculo del coeficiente de GINI
- Construcción de un índice a partir del análisis factorial
- Reflexiones finales, dudas, interrogantes

¿Por qué STATA? ¿Por qué estudiar la desigualdad social con STATA?

	Potencia	Dificultad de uso	Precio	Uso prioritario
SPSS	Alta	Baja*	Muy alto	Docencia / profesional
Stata	Muy alta	Media	Medio/Alto	Investigación aplicada
R	Máxima	Alta	Gratuito	Investigación básica y aplicada

Fuente: Curso de introducción a Stata – Jordi Muñoz (UAB)

¿Cómo usar STATA?

- Interfaz
- Línea de comando
- Sintaxis (do-files)

- Automatización de tareas complejas
- Generar comandos propios (ej: gini)
- Capacidad de reproducir manipulaciones de datos y resultados (¡lo más importante!)



#	Command	_rc
1	findit fastgini	
2	findit fastgini	
3	findit ineqerr	
4	use "C:\Users\José\...	
5	describe region	
6	codebook region	
7	sum region	

```
13.0 Copyright 1985-2013 StataCorp LP
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC http://www.stata.com
979-696-4600 stata@stata.com
979-696-4601 (fax)

3-user 8-core Stata network perpetual license:
Serial number: 501306208483
Licensed to: IDRE-UCLA
IDRE-UCLA

Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables

. findit fastgini
. findit fastgini
. findit ineqerr
. use "C:\Users\José\OneDrive\Bases\EPH\STATA\Individual_t207.dta", clear
. describe region

      storage   display   value
variable name  type     format   label      variable label
-----
region         double   %10.0g   region     Código de Región
```

Data editor (base)

Do-file

Ventana de resultados

Ventana de revisión

Ventana de comando

Variable	Label
CODUSU	Código para distin...
nro_hogar	Código para distin...
componente	Número de comp...
h15	Entrevista individu...
ano4	Año de relevamie...
	mes...
	gión
	seg...
aglomerado	Agglomerado
pondera	Ponderación
ch03	Relación de paren...
ch04	Sexo
ch06	Edad en años cum...

Ventana de variables

Properties

Name	region
Label	Código de Región
Type	double
Form	
Valu	
Not	

Files	Labels
Notes	
Variables	176
Observations	63.235
Size	83,70M

Ventana de propiedades de variables



Data editor (base)

CODUSU[1] 125009

	CODUSU	nro_hogar	componente	h15	ano4	trimestre	region	mas_500	aglomerado	pondera	ch03	ch04	ch
1	125009	1	2	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	931	Jefe	Varón	
2	125009	1	3	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	931	Hermano	Mujer	
3	125009	1	4	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	931	Hermano	Varón	
4	125009	1	5	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	931	Hermano	Mujer	
5	125250	1	1	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1408	Jefe	Varón	
6	125250	1	2	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1408	Cónyuge/Pa	Mujer	
7	125250	1	3	0	2007	2do. Trime	Gran Bueno	S	Ciudad de	1408	Hijo/Hijas	Mujer	
8	125250	1	4	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1408	Suegro	Mujer	
9	125963	1	1	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1154	Jefe	Mujer	
10	125963	1	2	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1154	Hijo/Hijas	Mujer	
11	125963	1	3	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1154	Madre/Padr	Mujer	
12	126002	1	1	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	457	Jefe	Varón	
13	126002	1	2	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	457	Cónyuge/Pa	Mujer	
14	126002	1	3	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	457	Hijo/Hijas	Varón	
15	126002	1	4	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	457	Hijo/Hijas	Varón	
16	126002	2	servicio d	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	457	Jefe	Mujer	
17	126289	1	1	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1393	Jefe	Mujer	
18	126289	2	servicio d	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1393	Jefe	Mujer	
19	126442	1	1	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1834	Jefe	Varón	
20	126442	1	2	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1834	Cónyuge/Pa	Mujer	
21	126467	1	1	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1767	Jefe	Mujer	
22	126467	1	2	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1767	Cónyuge/Pa	Varón	
23	126467	1	3	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1767	Hijo/Hijas	Mujer	
24	126467	1	4	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1767	Hijo/Hijas	Mujer	
25	126683	1	1	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1569	Jefe	Varón	
26	126683	1	2	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1569	Cónyuge/Pa	Mujer	
27	126683	1	3	0	2007	2do. Trime	Gran Bueno	S	Ciudad de	1569	Hijo/Hijas	Mujer	
28	126716	1	1	Sí	2007	2do. Trime	Gran Bueno	S	Ciudad de	1568	Jefe	Varón	

Variables

Filter variables here

Variable	Label
<input checked="" type="checkbox"/>	CODUSU Código para distin...
<input checked="" type="checkbox"/>	nro_hogar Código para distin...
<input checked="" type="checkbox"/>	componente Número de comp...
<input checked="" type="checkbox"/>	h15 Entrevista individu...
<input checked="" type="checkbox"/>	ano4 Año de relevamie...
<input checked="" type="checkbox"/>	trimestre Número de trimes...
<input checked="" type="checkbox"/>	region Código de Región
<input checked="" type="checkbox"/>	mas_500 Aglomerados seg...
<input checked="" type="checkbox"/>	aglomerado Aglomerado
<input checked="" type="checkbox"/>	pondera Ponderación
<input checked="" type="checkbox"/>	ch03 Relación de paren...
<input checked="" type="checkbox"/>	ch04 Sexo

Properties

Variables

Name	CODUSU
Label	Código para disti...
Type	str8
Format	%8s
Value Label	
Notes	

Data

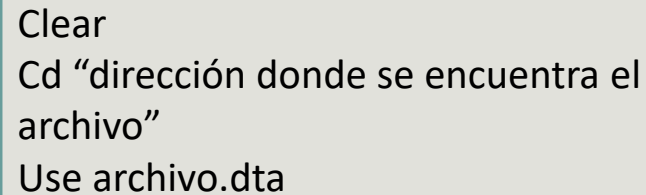
Filename	Individual_t207.dta
Label	
Notes	
Variables	176
Observations	63,235
Size	83,70M
Memory	128M

Funciones de ayuda de STATA

- Elemento central que provee STATA
- Comando “**help**”: brinda información completa acerca del comando sobre el que tenemos dudas (descripción, opciones, ejemplos, videos, etc.)
 - Sintaxis: help [nombre del comand] (ej: help tab)
- Comando “**Findit**”: buscar programas (rutinas) creadas por otros usuarios para agregar al STATA (ej: gini, multiple correspondence)
- Importante **comunidad de usuarios** que consultan y responden a interrogantes en tono al programa (foros, blogs, videos, STATALIST, página UCLA, etc.)

Abriendo la base de datos

- Formatos que soporta STATA: dta, xls, csv, txt, etc.
- Si queremos importar archivos de .sav, debemos primero guardarlos como .dta en el SPSS, para luego abrirlo con el STATA.
- Opciones para abrir una base:
 - A través de la interfaz grafica
 - Con sintaxis



```
Clear  
Cd "dirección donde se encuentra el  
archivo"  
Use archivo.dta
```

Lenguaje STATA

- comando [variable/s], [opciones]
- se puede abreviar (ej: summarize/sum; tabulate/tab)
- agregar comentarios en sintaxis (*), agregar comentarios en una línea donde hay órdenes (//)
- ¡Ojo! Hay confusiones con respecto al (=). El (=) funciona como operador aritmético y el (==) como operador lógico (prueba de equidad).

Inspeccionando las variables (comandos)

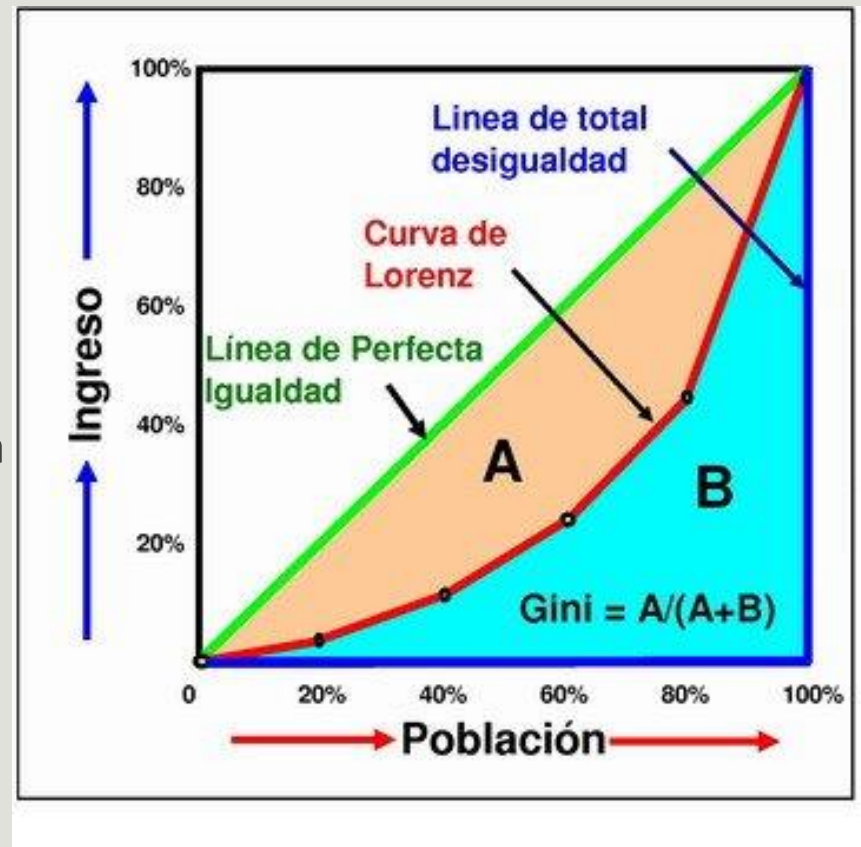
- DESCRIBE: brinda información sobre el tipo de variable que seleccionemos
- SUMMARIZE: información básica sobre como se distribuyen los casos según la variable (obs, media, desvío estándar, min, max)
- CODEBOOK: es como el summarize pero brinda mayor información

Distribución de frecuencias

- **Comando TAB:** nos permite ver como se distribuyen las frecuencias según variable.
- También permite realizar tablas de contingencias de dos o más variables.
- Permite el uso de “condicionantes” (IF) o “segmentaciones” (BYSORT).

El análisis de la desigualdad a partir del coeficiente de GINI

- ¿Qué necesito conocer?
 - Frecuencia acumulada de ingresos
 - Frecuencia acumulada de personas
- Es una medida de dispersión
 - 1 → Máxima desigualdad
 - 0 → Máxima igualdad
- La definición mas corriente se formula en términos de la curva de Lorenz
- Puede ser calculado a partir de información agregada o individual.
- Puede ser descompuesto algebraicamente.



Comandos de STATA (GINI)

- Buscaremos la palabra GINI en el repositorio de STATA (findit o search)
- Buscaremos los siguientes programas específicos e instalarlos:
 - Fastgini
 - Ineqerr
 - Inequal
 - Rspread
- Calcularemos el índice para el IPCF. Revisemos la particularidad de cada uno de los programas y si presentan diferencias en el cálculo.

Comandos de STATA (GINI)

Programa	Características
fastgini	Es rápido Por default, elimina los casos \neq a 0
rspread	Calcula varias medidas de dispersión Permite graficar la curva de Lorenz
ineqerr	Es más lento Calcula el GINI, THEIL y Var Logs Calcula los errores standard
Inequal	Brinda diversas medidas de dispersión

Comandos de STATA (GINI)

- Comandos para la descomposición del GINI
 - Diginig (instalar DASP)
 - Diginis (instalar DASP)
 - Descogini

Decomposition of the Gini Index by Groups
Group variable : ch04

Group	Gini index	Population Share	Income Share	Absolute Contrib.	Relative Contrib.
1: Varón	0,620	0,479	0,582	0,173	0,265
	0,002	0,002	0,003	0,002	0,002
2: Mujer	0,675	0,521	0,418	0,147	0,226
	0,002	0,002	0,003	0,002	0,002
Within	---	---	---	0,320	0,490
	---	---	---	---	---
Beetwin	---	---	---	0,102	0,157
	---	---	---	---	---
Overlap	---	---	---	0,230	0,353
	---	---	---	---	---
Population	0,652	1,000	1,000	0,652	1,000
	0,002	0,000	0,000	0,002	0,000

Construcción de un índice factorial con STATA

- El índice es una medida resumen construida a partir de un conjunto de indicadores (bienes, ingresos, vivienda, etc.)
- Puede construirse de forma “teórica” o con criterios estipulados y a partir de técnicas factoriales
- Algunos ejemplos de utilización:
 - MINUJIN, A. y J. H. BANG (2002) "Indicadores de inequidad social. Acerca del uso del «índice de bienes» para la distribución de los hogares". En *Desarrollo Económico*. Año 42, N° 165, abril. Págs. 129.
 - FILMER, D. y L. H. PRITCHETT (2001) "Estimating Wealth Effects without Expenditure Data-or Tears: An Application to Educational Enrollments in States of India". En *Demography*. Año 38, N° 1, febrero. Págs. 115.
 - BARANGER, D. (2009) "Construcción y análisis de datos. Introducción al uso de técnicas cuantitativas en la investigación social".

Construcción de un índice factorial con STATA

TABLA 2
Coeficientes para el cálculo de puntajes factoriales de la componente principal según indicador ordenados de forma decreciente

Indicadores	Coeficiente
Retrete con descarga de agua	0.144
Agua por cañería dentro de la vivienda	0.137
Vivienda con materiales de alta calidad	0.122
Teléfono	0.119
Televisor color	0.104
Televisión por cable	0.100
Calefacción por sistema central, artefactos fijos, estufas eléctricas	0.099
Lavarropa automático	0.096
Cantidad de habitaciones de la vivienda	0.093
Video reproductor/grabador	0.093
Heladera con freezer o freezer independiente	0.084
Auto de hasta 5 años de antigüedad	0.079
Computadora personal	0.066
Electricidad	0.060
Sólo auto de entre 6 a 10 años de antigüedad	0.031
Sólo auto de más de 10 años de antigüedad	0.021
Motocicleta	0.015
Bicicleta	0.005
Agua por bomba	-0.020
Sólo heladera sin freezer	-0.039
Agua por fuente no entubada	-0.053
Vivienda con materiales de baja calidad	-0.066
Sólo televisor blanco y negro	-0.069
Cocina con leña, carbón o materiales orgánicos	-0.075
Sin retrete	-0.095
Retrete sin descarga de agua	-0.102

Fuente: Minujín y Bang (2002)



+ apropiación de activos

- apropiación de activos (privación)

Construcción de un índice factorial con STATA

1. Elección de la técnica factorial (Análisis de correspondencias múltiples “ACM”)
2. Principal característica ACM: resumir un conjunto de variables en un nuevo conjunto más reducido (factores)
3. Usar el comando MCA (ya viene en STATA) con las variables consideradas

Ejemplo ACM

