

ECONOMETRIA I - CURSO 2015

PRACTICO 0

Breve Introducción al GRET

Repaso de Estadística II

GRET (Gnu Regression Econometric and Time Series) es un software libre que se puede descargar de varios sitios, como por ejemplo:

<http://gretl.sourceforge.net/>

LOS DATOS:

1. **Importar datos.** El *gretl* permite importar datos que están en archivos de extensión *txt* y *xls* (permite otras extensiones, pero esas dos son las que trabajaremos en el curso).

EJEMPLO: en el *eva* (Entorno Virtual de Aprendizaje) del curso, se halla disponible el archivo *votos2014.txt*, en el que se incluyen datos del apoyo a un candidato presidencial de los USA (A) y a su principal adversario (B). Los datos corresponden a 173 distritos electorales y las variables son:

- *voteA*: porcentaje de votos correspondientes al candidato A.
- *state*: código correspondiente al estado correspondiente al distrito.
- *distrit*: distrito electoral.
- *expendA*: gastos de campaña correspondientes al candidato A en miles de dólares americanos.
- *expendB*: gastos de campaña correspondientes al candidato B en miles de dólares americanos.
- *prtystrA*: porcentaje de adhesiones en las últimas elecciones pasadas correspondientes al candidato A.

Grábalo en el escritorio de su pc o laptop.

Para importarlo:

- a) Abrimos sesión del *gretl*
- b) La ventana que se despliega en pantalla tiene únicamente habilitados tres ítems: Archivo – Herramientas – Ayuda (el resto se habilitará una vez se haya cargado un archivo de datos).
- c) Se eligen las opciones:
Archivo – Abrir archivo de datos – Archivo de usuario – Escritorio – Archivos ASCII
- d) Se selecciona el archivo *votos2014.txt*, y Abrir
- e) El *gretl* preguntará si se deben interpretar los datos como series temporales, a lo que respondemos que “no”, dado que son datos de sección cruzada.
- f) El resultado en pantalla es:
 - se habilitan el resto de los ítems de la barra de comandos
 - aparecen las variables con su respectivo nombre (porque el archivo contenía una línea con esa información: no es obligatorio).
 - podemos asociar a cada variable una etiqueta descriptiva que aparecerá, por ejemplo, en los gráficos.
- g) Para adjudicar etiquetas a las variables el procedimiento es:
 - posicionarse en la variable cuya información se quiere complementar

- seleccionar Variable – Editar atributos (Se desplegarán los campos para describir la variable, el nombre a mostrar en los gráficos y la posibilidad de definirla como discreta).
- h) La constante es generada automáticamente.
- i) Añadir variables que son función de algunas variables:
Añadir – Logaritmos de la variable seleccionada (hacerlo para las variables asociadas a los gastos de campaña de A y B).
- j) Archivo – Guardar datos (se genera un archivo en formato .gdt para usos futuros).

2. **Ingresar valores uno a uno:** Para ello, iniciar el programa *gretl* y luego:

Archivo – Nuevo conjunto de datos.

Aparece una ventana preguntando el número de datos de la base que se está creando. Luego, otra ventana solicita se indique el tipo de datos que se introducirán: sección cruzada, series temporales o datos panel. Finalmente, una última ventana solicita se confirmen las características de la base de datos que se va a crear: marcamos la casilla correspondiente al inicio de inserción de los datos.

Automáticamente aparecen dos nuevas ventanas: una solicita el nombre de la variable, y otra es una planilla en la que se introducirán los valores observados correspondientes a la variable.

En la planilla, para introducir una nueva variable, basta marcar el signo de '+', con lo que se solicitará el nombre de la nueva variable... y así sucesivamente.

Haciendo clic en el “visto bueno”, quedan registrados los datos de la base. Se puede cerrar la planilla editora que se usó para ingresar los datos.

En la pantalla principal, aparecerán las variables ingresadas.

Posicionándose en el nombre de una de las variables de la base, y marcando en el menú: Variable – Mostrar valores se despliegan los valores de la variable en cuyo nombre se posicionó el cursor.

Como antes, Archivo – Guardar datos (se genera un archivo en formato .gdt para usos futuros).

Nota: Por más detalles, referirse a “*Econometría básica Aplicada con Gretl*” (link indicado en el *eva*).

EJEMPLO: Ingrese los datos de la página 2 del referido manual.

ANÁLISIS DESCRIPTIVO DE LOS DATOS:

- i) abrimos la sesión de *gretl*.
- ii) Abrir – Archivo de datos – Archivo de usuario – *votos2014.gdt* (ubicarlo en el directorio en que fue creado)
- iii) Elegimos la variable *expndA*.
- iv) **Métodos Gráficos:** frecuencias, histograma (por default se toman clases de igual amplitud). Luego de elegir el número de clases y haber indicado que se despliegue el gráfico, éste se desplegará en otra pantalla. Con el botón derecho en el área del gráfico, y eligiendo la opción “editar gráfico”, se puede agregar un título al gráfico, incorporar la rejilla, cambiar el color de relleno, etc.
- v) En otra pantalla, aparecerán simultáneamente las frecuencias absolutas y relativas de la variable *expndA*.

- vi) **Medidas resumen:** Siempre considerando la misma variable, haciendo Variable – Estadísticos principales, aparecerán las medidas resumen usuales para el caso de variables unidimensionales.
- vii) **Ajuste de Normalidad:** Variable – Contraste de Normalidad, para obtener los resultados de la Bondad de Ajuste utilizando los estadísticos usuales de Jarque-Bera, Lilliefors, etc.
- viii) **Concentración:** Variable – Índice de Gini permite obtener el valor del índice y la Curva de Lorenz de la variable en estudio.

CUESTIONARIO: Estadística Descriptiva Univariada

1. De acuerdo con las medidas resumen obtenidas, **expendA** presenta
 - a) asimetría a la izquierda
 - b) asimetría a la derecha
 - c) simetría
2. De acuerdo con las medidas obtenidas,
 - a) la variable **expendA** es más aplanada que una normal
 - b) la variable **expendA** es menos aplanada que una normal
 - c) la variable **expendA** es igual de aplanada que una normal



Considere el Diagrama de Caja de la variable **expendA** para contestar las siguientes preguntas:

3. La distancia entre F y C es:
 - 387,8650
 - 193,9325
 - 581,7975
4. El valor asociado a D es:
 - 310,611
 - 280,986
 - 242,780
5. El valor asociado a E es:
 - 145,085
 - 242,780
 - 280,986
6. Los cuartiles son:
 - C, D y E
 - C, D y F
 - C, E y F
7. Si la primera cuartila es 71,150, G es:
 - 0,30000000
 - -510,64750
 - -1092,4450
8. Si la tercer cuartila es 459,015, el punto A es:
 - un valor atípico
 - un valor atípico extremo

9. Considere ahora las variables `expendA` y `expendB`. Calcule el Índice de Gini de ambas variables y grafique la Curva de Lorenz. (Variable – Coeficiente de Gini). De acuerdo con los resultados que obtuvo, la distribución del gasto del candidato A en los distritos considerados es:
 - más uniforme que la del candidato B
 - igual de uniforme que la del candidato B
 - menos uniforme que la del candidato B
10. La correlación lineal (Ver – Matriz de correlación) entre las variables `expendA` y `expendB`
 - no es estadísticamente significativa
 - es estadísticamente significativa
11. La covarianza entre las variables `expendA` y `expendB` es:
 - 32183,6
 - 9857043,7791
 - 10744591,6743
12. La variable `expendA` es
 - más dispersa que la variable `expendB`
 - menos dispersa que la variable `expendB`
 - tan dispersa como la variable `expendB`

OBSERVACIÓN 1: No siempre todos los resultados que el usuario precisa se obtienen a partir de íconos incluidos en los menús establecidos en las pantallas interactivas de *gretl*. El abanico de posibilidades crece si se crea un archivo (código, guión) que incluya las instrucciones que conduzcan a los resultados deseados.

1. Archivo – Archivos de guión – Nuevo guión – Guión de Gretl
2. se despliega una ventana en la que se escribirá el código (*gretl*: sin nombre)
3. Para elegir el comando adecuado, en la pantalla *gretl*: Ayuda – Guía de funciones
4. Se despliega una pantalla (*gretl*: Guía de Funciones)
5. A la izquierda de esa nueva pantalla se elige la opción Estadística y, entre ellas, se elige la instrucción adecuada a las necesidades del momento.
6. Entonces se tiene la sintaxis del comando adecuado a los resultados buscados.
7. Ejemplo: `quantile`
 Output: scalar or matrix
 Arguments: `y` (series or matrix)

`p` (scalar between 0 and 1)

 If `y` is a series, returns the `p`-quantile for the series. For example, when `p = 0.5`, the median is returned.
8. Una vez escrita la instrucción con la sintaxis adecuada, se debe hacer clic en el ícono ejecutar del menú de la pantalla (*gretl*: sin nombre). En una nueva pantalla (*gretl*: resultados de guión) aparecerán los resultados deseados.

OBSERVACIÓN 2:

- El punto (.) es el separador de la parte entera de las cifras decimales.
- La coma (,) separa los argumentos de las funciones.
- Los guiones no admiten los signos de punto y coma (;).

PRUEBAS DE HIPÓTESIS:

- a) **Para la media:** $H_0) \mu = c$ vs. $H_1) \mu \neq c (\mu > c \text{ ó } \mu < 0)$
Herramientas – Calculadora de estadísticos de contraste - Media
Se despliega una pantalla (gretl: calculadora de contrastes) en la que se indica la variable a considerar y el valor c. Automáticamente se generan los valores de la media muestral, del desvío y el número de observaciones.
Haciendo clic en Aceptar, se despliega la pantalla con los resultados (gretl: contraste de hipótesis) en la que se da el valor del estadístico y el p-valor a una cola y a dos colas. También se despliega otra pantalla con el gráfico de la densidad del estadístico de prueba.
- b) **Para la varianza:** $H_0) \sigma^2 = c$ vs. $H_1) \sigma^2 \neq c (\sigma^2 > c \text{ ó } \sigma^2 < 0)$
Herramientas – Calculadora de estadísticos de contraste – Varianza
Se despliega una pantalla (gretl: calculadora de contrastes) en la que se indica la variable a considerar y el valor c. Automáticamente se generan los valores de la media muestral, del desvío y el número de observaciones.
Haciendo clic en Aceptar, se despliega la pantalla con los resultados (gretl: contraste de hipótesis) en la que se da el valor del estadístico y el p-valor a una cola y a dos colas. También se despliega otra pantalla con el gráfico de la densidad del estadístico de prueba.
- c) **Para la proporción:** $H_0) p = p_0$ vs. $H_1) p \neq p_0 (p > p_0 \text{ ó } p < p_0)$
Herramientas – Calculadora de estadísticos de contraste – Proporción
Se despliega una pantalla (gretl: calculadora de contrastes) en la que se indica la variable a considerar y el valor c. Ya no se generan automáticamente los valores de la proporción muestral, del desvío y el número de observaciones. Se deben calcular previamente.
Haciendo clic en Aceptar, se despliega la pantalla con los resultados (gretl: contraste de hipótesis) en la que se da el valor del estadístico y el p-valor a una cola y a dos colas. También se despliega otra pantalla con el gráfico de la densidad del estadístico de prueba.
Ojo!: - el recorrido de la variable en estudio debe ser [0; 1]
- se debe calcular previamente el valor muestral de la proporción.
- d) **Comparación de medias, varianzas y proporciones:** la forma de proceder en estos casos es similar a la de los casos para un solo parámetro.

CUESTIONARIO: Inferencia Estadística Paramétrica

13. Realice la prueba para la media de la variable `expendA` (varianza desconocida), considerando $H_0) \mu = 300$:
- a) Si $H_1) \mu \neq 300$, el p-valor de la prueba es 0.31
 - b) Si $H_1) \mu > 300$, el p-valor de la prueba es 0.31
 - c) Ninguna de las dos opciones propuestas.
14. Realice la prueba para la media de la variable `expendA` (varianza desconocida), considerando $H_0) \mu = 310$ y $\alpha = 0.05$:
- a) Si $H_1) \mu \neq 310$ y dado que el valor del estadístico de prueba es 0.0286055 (menor que 0.05), se rechaza H_0 .

- b) Si $H_1) \mu > 310$ y dado que el valor del estadístico de prueba es 0.0286055 (menor que 0.05), se rechaza H_0 .
- c) Si $H_1) \mu \neq 310$ y dado que el p-valor es 0.9772 (mayor que 0.05), no se rechaza H_0 .
15. Realice una prueba para determinar si la diferencia de medias de las variables `expendA` y `expendB` es 10 (se asume que las varianzas de las dos variables son iguales), considerando $\alpha = 0.05$. De acuerdo con sus cálculos:
- a) $H_1)$ es $\mu_A - \mu_B \neq 300$ y el p-valor de la prueba es $0.8874 > \alpha$ por lo que no rechazo H_0 .
- b) $H_1)$ es $\mu_A - \mu_B > 300$ y el p-valor de la prueba es $0.8874 > \alpha$ por lo que no rechazo H_0 .
- c) $H_1)$ es $\mu_A - \mu_B < 300$ y el p-valor de la prueba es $0.4437 > \alpha$ por lo que no rechazo H_0 .
16. Realice una prueba para determinar si las varianzas de las variables `expendA` y `expendB` son iguales, considerando $\alpha = 0.05$. De acuerdo con sus cálculos:
- a) como la distribución del estadístico de prueba es una F (Snédecor) de recorrido positivo, la Región Crítica de la prueba es “a una cola” y el p-valor es 0.1297.
- b) la distribución del estadístico de prueba no es una F (Snédecor), por lo que la Región Crítica de la prueba es “a dos colas” y el p-valor es 0.2593
- c) la distribución del estadístico de prueba es una F (Snédecor) y la Región Crítica de la prueba es “a dos colas”. El p-valor es 0.2593.
17. Considere la prueba para determinar si las varianzas de las variables `expendA` y `expendB` son iguales ($\alpha = 0.05$) realizada anteriormente.
- a) Los valores del estadístico de prueba que definen la Región Crítica son iguales y opuestos.
- b) Los valores del estadístico de prueba que definen la Región Crítica son los cuantiles correspondientes a las probabilidades 0.025 y 0.975 de una distribución $F_{k,m}$ con $k = m = 172$.
- c) Los valores del estadístico de prueba que definen la Región Crítica son los cuantiles correspondientes a las probabilidades 0.025 y 0.975 de una distribución $F_{k,m}$ con $k = m = 173$.
18. Defina una nueva variable (`prtystrAp`) cuyos valores se obtienen dividiendo por 100 los valores de la variable `prtystrA`. Realice la prueba para determinar si la proporción de votos que adhirieron al candidato A en las elecciones pasadas fue 0.50, considerando todos los distritos electorales.
- a) La $H_1)$ correspondiente es $p \neq 0.5$, el p-valor de la prueba es 0.9490
- b) La $H_1)$ correspondiente es $p \neq 0.5$, el p-valor de la prueba es 0.4745
- c) No se puede calcular el p-valor de la prueba, dado que no se conoce la varianza de la variable `prtystrAp` por desconocer el verdadero valor de p .
19. Defina una nueva variable (`pvoteA`) cuyos valores se obtienen dividiendo por 100 los valores de la variable `voteA`. Realice la prueba para determinar si las proporciones asociadas a `pvoteA` y `prtystrAp` son iguales ($\alpha = 0.05$).
- a) La $H_1)$ correspondiente es del tipo $p_1 - p_2 \neq 0$ y el p-valor de la prueba es $0.8896 > 0.5$, por lo que no se rechaza la igualdad de proporciones.
- b) La $H_1)$ correspondiente es del tipo $p_1 - p_2 \neq 0$ y el p-valor de la prueba es $0.8896 > 0.05$, por lo que no se rechaza la igualdad de proporciones.
- c) La $H_1)$ correspondiente es del tipo $p_1 - p_2 > 0$ y el p-valor de la prueba es $0.4448 > 0.05$, por lo que no se rechaza la igualdad de proporciones.