NORMALIDAD ASINTÓTICA E INFERENCIA CON MUESTRAS GRANDES
MODELO CLÁSICO DE REGRESIÓN NORMAL (MCRLN).

Teoría clásica de la inferencia estadística

- Estimación
- Prueba de hipótesis

OBJETIVO: ESTIMAR LA FRM Y UTILIZARLA PARA HACER INFERENCIA ESTADÍSTICA SOBRE LA FRP.

SUPUESTO DE NORMALIDAD[1]

Hasta los momentos sobre las (residuos) perturbaciones se ha supuesto:

- $E(u_i) = 0$
- $E[u_i - E(u_i)]^2 = E({u_i}^2) = \sigma^2$
- $E\{[u_i - E(u_i)][u_j - E(u_j)]\} = E(u_i u_j) = 0 \quad i \neq j$

[1] La distribución muestral de los estimadores MCO depende de la distribución subyacente de los errores o perturbaciones estocásticas, porque aun bajo el teorema de Gauss Markov la distribución de los estimadores puede tener casi cualquier forma.

El MCRLN supone que cada u_i está normalmente distribuida para poder hacer inferencia.

$$u_i \sim N(0, \sigma^2)$$

PARA EFECTOS DE ESTIMACIÓN NO SE REQUIERE EL SUPUESTO DE NORMALIDAD DE RESIDUOS, TAL SUPUESTO SE NECESITA SOLAMENTE PARA HACER INFERENCIA ESTADÍSTICA.

$$Y|X \approx Normal$$

Para dos variables normalmente distribuidas, una covarianza o correlación cero significa independencia entre las dos variables. Así, haciendo uso del supuesto de normalidad, significa que u_i y u_j no sólo están correlacionadas sino también están independientemente distribuidas

$$u_i \sim NID(0, \sigma^2)$$

¿POR QUÉ UNA DISTRIBUCIÓN NORMAL?[2]

[2] Como los residuos son la suma de muchos factores que afectan a Y, se puede apelar por el teorema central del límite para concluir que los U tienen una distribución aproximadamente normal. Este argumento es cierto pero no carece de debilidades. 1) Los factores incluido en los residuos pueden tener distribuciones muy diferentes en la población. 2) Esto supone que todos los factores incluidos en U afectan a Y de forma aditiva. 3) Suponer normalidad es una cuestión empírica.

- Es muy conocida y de uso muy sencillo.

- Permite obtener distribuciones muestrales exactas de los estimadores MCO.

- ***Teorema del Límite Central (TLC),*** mediante el cual se demuestra que si existe un gran número de variables aleatorias independientes e idénticamente distribuidas, su suma tiende a ser normal a medida que el número de variables se incrementa[3].
    - ***TLC:*** Sea $\{Y_1,\ Y_2,\ Y_3, \dots, Y_n\}$ con media μ y varianza σ^2. Entonces, $Z_n = \frac{\overline{Y_n}-\mu}{\sigma/\sqrt{n}}$ tiene una distribución normal asintótica[4].

- Aunque el número de variables no sea muy grande o si no son estrictamente independientes, su suma tiende a una normal.

- La distribución de probabilidad de los estimadores MCO también será una normal porque cualquier función lineal de variables normalmente distribuidas estará también normalmente.

- Cuando se trabaja con muestras finitas la suposición de normalidad es muy importante ya que, ayuda a derivar las distribuciones de probabilidad exacta de los estimadores MCO permitiendo así el uso de las pruebas estadísticas t, F y X^2 que son las más frecuentes en econometría aplicada.

---

[3] Una excepción a este teorema es la disteribución Cauchy , la cual no tiene media ni momentos más altos.

[4] Sea $\{Z_n: n = 1,2, \dots\}$ una secuencia de variables aleatorias tal que para todos los números z, $P(Z_n \leq z) \to \Phi_z\ \ cuando\ n \to \infty$. Donde Φ_z es una función de distribución acumulada normal estándar.

PROPIEDADES DE LOS ESTIMADORES MCO BAJO EL SUPUESTO DE NORMALIDAD

- Son insesgados. $E(\hat{\beta}_i) = \beta_i$.

- Son eficientes. Tienen varianza mínima.

- Son consistentes $\hat{\beta}_i \xrightarrow{n} \beta_i$

Sea θ_n un estimador de θ basado en una muestra Y_1, Y_2, $Y_3, \dots, Y_n$, del tamaño *n*. Entonces, θ_n es un estimador consistente de θ si para todo $\varepsilon > 0$, $P(|\theta_n - \theta| > \varepsilon) \rightarrow 0, cuando\ n \rightarrow \infty$, de lo contrario es un estimador inconsistente.

Respecto a la consistencia:

"Si usted no puede obtener un resultado correcto a medida que el tamaño de muestra tiende a infinito, es mejor que se dedique a otra cosa" Clive W.J Granger. Premio Nobel de Economía.

$$\hat{\beta}_i \sim N(\beta_i, \sigma^2{}_{\hat{\beta}_i})$$

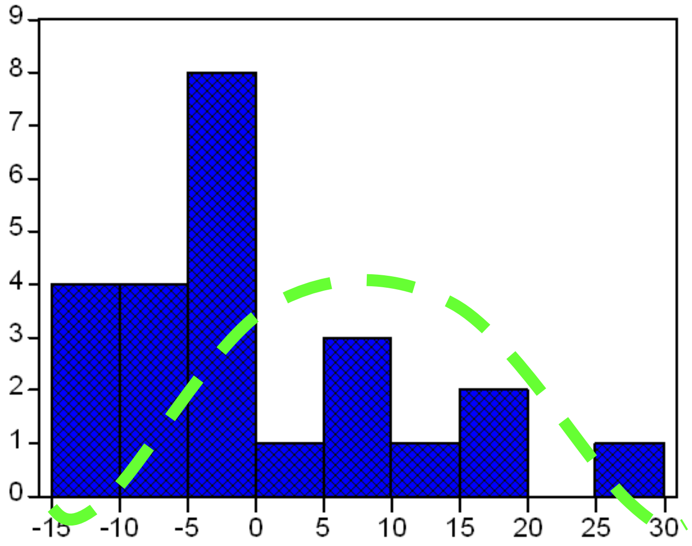
$$Y_i \sim N(\beta_1 + \beta_2 X_i, \sigma^2)$$

TEOREMA DE RAO: los estimadores MCO tienen varianza mínima entre todas las clases de estimadores insesgados, lineales o no. Esto significa que los estimadores MCO son **MEI**

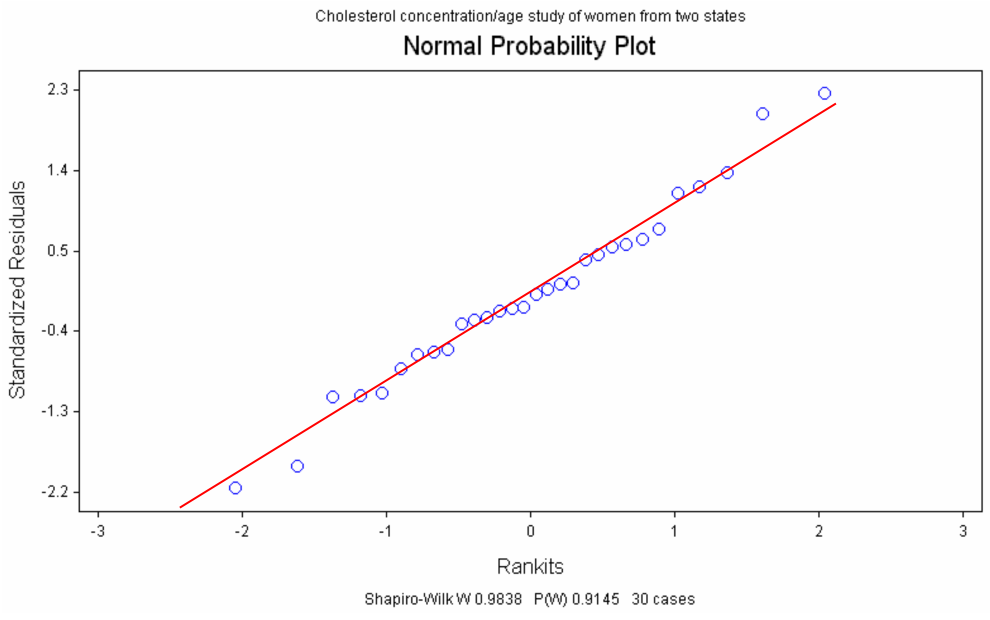
TEOREMA DE CONSISTENCIA DE LOS MCO: Bajo los supuestos del MCRL, estimador $\hat{\beta}_i$ es consistente para β_i, para toda i= 0,1,2,3,…,n

¿CÓMO SE PRUEBA QUE LOS RESIDUOS SON NORMALES?

- ***Histograma de residuos:***



- ***Gráfica de probabilidad normal (GPN):*** estudia la forma de la función de densidad de probabilidad



- ***Prueba de normalidad de Jarque-Bera (JB)***
- ***Prueba de normalidad de Shapiro-Wilk (SW)***
- ***Prueba de Kolmogorov- Smirnov (K-S)***

Metodología:

a) Planteamiento de la hipótesis:

Ho: Los residuos son normales.
H_1: Los residuos no son normales.

b) Selección del estadístico de prueba: $W = \frac{1}{n\,s^2}\left(\sum_{i=1}^{h} a_{i\,n}\left(x_{(n-i+1)} - x_{(i)}\right)\right)^2$

$JB = n\left[\frac{A^2}{6} + \frac{(K-3)^2}{24}\right] \sim \chi_k^2$; Donde A^2: $asimetría$; K: $kurtosis$

$SW = \frac{1}{nS^2}\left[\sum_{i=1}^{n} a_{in}(X_{n-i+1} - X_i)\right]^2$; Donde a_{in}: $valores\ tabulado$; S^2: $varianza\ de\ las\ variables$

$$KS = \frac{1}{n}\sum_{i=1}^{n}\begin{cases} 1 & si \quad y_i \leq x \\ 0 & alternativa \end{cases}$$

c) Aplicar regla de decisión:

Valor calculado > valor tabulado RECHAZAR Ho

Valor calculado < valor tabulado NO RECHAZAR Ho

P-v > α NO RECHAZAR Ho

P-v < α RECHAZAR Ho

```
jarqueberaTest(res2)
```

```
##
## Title:
##  Jarque - Bera Normalality Test
##
## Test Results:
##   STATISTIC:
##     X-squared: 32.6719
##   P VALUE:
##     Asymptotic p Value: 8.042e-08
##
## Description:
##  Mon Sep 25 17:27:45 2017 by user: Laura
```

```
shapiroTest(res2)
```

```
##
## Title:
##  Shapiro - Wilk Normality Test
##
## Test Results:
##   STATISTIC:
##     W: 0.9547
##   P VALUE:
##     6.064e-06
##
## Description:
##  Mon Sep 25 17:27:45 2017 by user: Laura
```

```
ksnormTest(res2)
```

```
##
## Title:
##  One-sample Kolmogorov-Smirnov test
##
## Test Results:
##   STATISTIC:
##     D: 0.262
##   P VALUE:
##     Alternative Two-Sided: 3.157e-12
##     Alternative      Less: 1.579e-12
##     Alternative   Greater: 6.021e-08
##
## Description:
##  Mon Sep 25 17:27:45 2017 by user: Laura
```

¿QUÉ SUCEDE SI LOS RESIDUOS NO SON NORMALES?

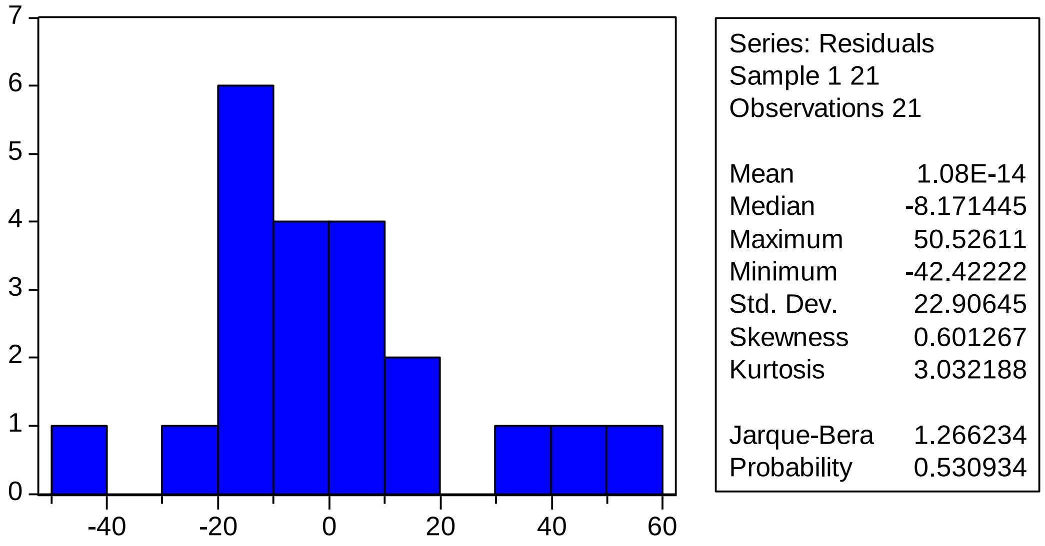
- Los estimadores MCO siguen siendo MELI bajo los supuestos MCRL.
- Las pruebas de hipótesis no serán confiables, por tanto los estimadores MCO tampoco lo serán. Recuerde que son función lineal de los residuos.

LA NO NORMALIDAD DE LOS RESIDUOS NO VIOLA NINGUNO DE LOS SUPUESTOS DEL MCRL.

¿QUÈ HACER SI LOS RESIDUOS NO SON NORMALES?

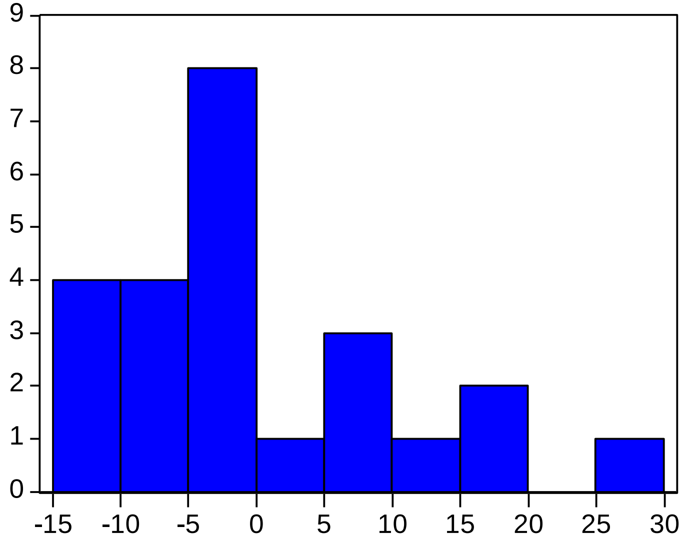
- Aumentar el tamaño de la muestra.
- Transformar las variables.

Ejemplo 1: Impacto de los gastos en publicidad. Wall Street Journal datos de 1984.



¿Son normales estos residuos?

Ejemplo 2: Relación tasa de interés e inflación en Venezuela 1982-2005.

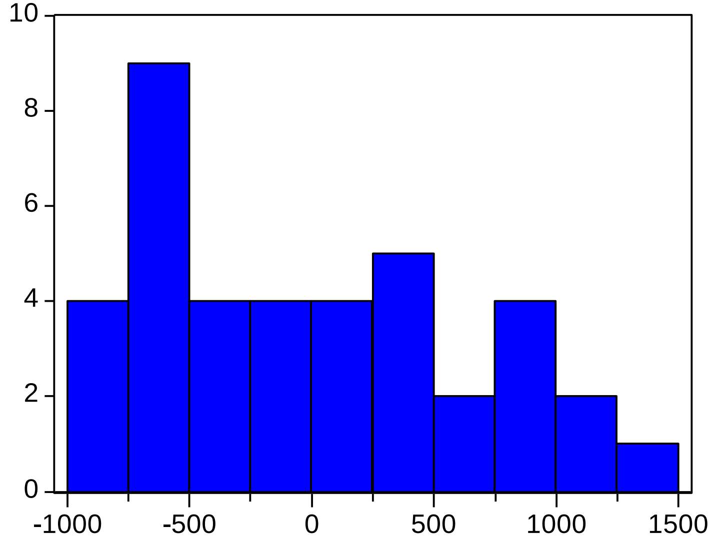


Series: Residuals
Sample 1982 2005
Observations 24

Mean	3.55E-15
Median	-2.373125
Maximum	29.53232
Minimum	-11.83558
Std. Dev.	10.03610
Skewness	1.213423
Kurtosis	4.305520
Jarque-Bera	7.593967
Probability	0.022438

¿Son normales estos residuos?

Ejemplo 3: PIB nominal de los EEUU 1959-1997.



Series: Residuals
Sample 1959 1997
Observations 39

Mean	2.53E-11
Median	-122.8406
Maximum	1291.554
Minimum	-830.2584
Std. Dev.	642.7390
Skewness	0.422858
Kurtosis	1.985332
Jarque-Bera	2.835277
Probability	0.242285

¿Son normales estos residuos?

Lecturas recomendadas:

- ✓ Gujarati, D. (2004). Econometría. 4ta. Edición McGraw Hill. Capítulo 4.
- ✓ Novales, A. (1993). Econometría. 2da. Edición. McGraw Hill Interamericana. Capítulo 3.
- ✓ Woodldridge, J. (2010). Introducción a la econometría. Un enfoque moderno. Capítulos 4 y 5.