INFERENCIA EN EL MODELO DE REGRESIÓN LINEAL CLÁSICO

Notas de clase de Econometría I

Material estrictamente académico

ESQUEMA

1. Introducción
2. Estimación por Intervalos
3. Pruebas de Hipótesis
4. Pruebas de Hipótesis Individuales sobre "β": t-estadístico
5. Pruebas de Hipótesis Conjuntas sobre "β": el estadístico *F*
6. Aplicación: La Demanda por Dinero en el Perú

1. INTRODUCCIÓN

- **Supuestos explícitos**:
  - Se cumplen los supuestos clásicos.
  - Los supuestos 2 y 5 se resumen en el supuesto de que las perturbaciones son Ruido Blanco o White Noise independientes:

$$u_{\mathrm{i}} \sim \mathrm{iid}(0, \sigma_u^2)$$

- **Supuesto adicional**:
  - En muestras pequeñas, cada perturbación se distribuye Normal:

$$u_{\mathrm{i}} \sim iid\ N(0, \sigma_u^2)$$

1. INTRODUCCIÓN

- Bajo los supuestos clásicos y el supuesto de normalidad:

$$\hat{\alpha} \sim N\left(\alpha, \sigma_u^2\left(\frac{1}{n} + \frac{\bar{X}}{\sum_{i=1}^{n} x_i^2}\right)\right)$$

$$\hat{\beta} \sim N\left(\beta, \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n} x_i^2}\right)$$

2. ESTIMACIÓN POR INTERVALOS

- Los estimadores "puntuales" toman valores diferentes en muestreo repetido (dependen de la muestra analizada).
- Error de estimación, que es una variable aleatoria, se define como:

$$e \equiv \left|\hat{\theta} - \theta\right| \qquad , \qquad e \equiv \left|\hat{\beta}_{MCO} - \beta\right|$$

- La probabilidad $P(e < a)$
  - Proporciona una medida de la bondad de una sola estimación.
  - Proporción de veces que, en muestreo repetido, el estimado obtenido con el estimador $\hat{\theta}$ es diferente de θ.

2. ESTIMACIÓN POR INTERVALOS

- Para un problema de estimación, si se conociera la distribución muestral del estimador, sería sencillo calcular $P(e<a)$:

$$\int_{\theta-a}^{\theta+a} f(\hat{\theta})d\hat{\theta}$$

- En un contexto de estimación, es importante encontrar un valor de ***a*** tal que:

$$P(e<a) \equiv P(\left|\hat{\theta}-\theta\right|<a) = \int_{\theta-a}^{\theta+a} f(\hat{\theta})d\hat{\theta} = 0.95$$

- De esta manera, se puede obtener un ***estimador por intervalos*** para θ

$$P(\left|\hat{\theta}-\theta\right|<a) = 0.95$$

$$\theta \in [\hat{\theta}-a, \hat{\theta}+a] \quad con \quad 95\% \; de \;\; probabilidad$$

2. ESTIMACIÓN POR INTERVALOS

- La mayoría de variables aleatorias observadas en la naturaleza caen en un intervalo de dos desvíos estándar alrededor de su media con una probabilidad aproximada de 0.95.
- Esto significa que: $Y \in [\mu_Y - 2\sigma, \mu_Y + 2\sigma] \quad con \quad 95\%\ de\ probabilidad$
- En la práctica esto significa que una buena aproximación del error de predicción de un estimador insesgado es $a = 2\sigma_{\hat{\theta}}$, lo cual implica que:

$$\theta \in [\hat{\theta} - 2\sigma_{\hat{\theta}}, \hat{\theta} + 2\sigma_{\hat{\theta}}] \quad con \quad 95\%\ de\ probabilidad$$

2. ESTIMACIÓN POR INTERVALOS

- El objetivo es encontrar un estimador por intervalo que genere intervalos angostos que contengan a θ con una alta probabilidad.
- Los estimadores por intervalo se denominan usualmente ***intervalos de confianza***.
- La probabilidad de que un intervalo de confianza contenga a θ se denomina ***coeficiente de confianza***:
  - En muestreo repetido: proporción de veces que los intervalos construidos para cada muestra contienen al parámetro θ:

$$P(\hat{\theta}_L < \theta < \hat{\theta}_U) = 1 - \varepsilon$$

$$P(\hat{\theta}_L < \theta) = 1 - \varepsilon$$

$$P(\theta < \hat{\theta}_U) = 1 - \varepsilon$$

Método Pivote

- Útil para obtener intervalos de confianza.
- Depende de la determinación de una *variable pivote*, la cual presenta dos características:
  - Es función de los datos muestrales (estadístico) y de θ .
  - Su distribución de probabilidades es conocida y no depende de θ .

Ejemplo: Bajo el supuesto de normalidad, se tiene que en el MRLC2:

$$Z = \frac{\hat{\beta} - \beta}{\sigma_u \Big/ \sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2}} \sim N(0,1)$$

2. ESTIMACIÓN POR INTERVALOS

- ***Si conocemos*** $\boldsymbol{\sigma^2}$, podemos encontrar un intervalo de confianza para β a partir del intervalo de confianza para esta variable pivote:

$$P(|Z| < Z_{\varepsilon/2}) = 1 - \varepsilon$$

$$P(-Z_{\varepsilon/2} < Z < Z_{\varepsilon/2}) = 1 - \varepsilon$$

$$P\left(\hat{\beta} - Z_{\varepsilon/2}\sigma\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} < \beta < \hat{\beta} + Z_{\varepsilon/2}\sigma\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right) = 1 - \varepsilon$$

$$P\left(\hat{\beta} - Z_{\varepsilon/2} s.e.(\hat{\beta}) < \beta < \hat{\beta} + Z_{\varepsilon/2} s.e.(\hat{\beta})\right) = 1 - \varepsilon$$

- Entonces, el estimador por intervalos para β es:

$$\beta \in [\,\hat{\beta} - Z_{\varepsilon/2} s.e.(\,\hat{\beta}\,), \hat{\beta} + Z_{\varepsilon/2} s.e.(\,\hat{\beta}\,)]$$

el cual tiene un ***coeficiente de confianza de*** $1-\varepsilon$ ***o*** $1-\alpha$

2. ESTIMACIÓN POR INTERVALOS

- La variable pivote anterior puede usarse a pesar que no se cumpla el supuesto de normalidad en muestras pequeñas apelando al TLC.
- Si ***no conocemos*** $\boldsymbol{\sigma^2}$, se usa como pivote una variable cuya distribución no dependa de éste parámetro desconocido.
- Se elige una variable pivote con distribución "t-Student", definida como el cociente de una variable normal estándar (Z) y la raíz cuadrada de una variable chi-cuadrado con "S" grados de libertad, independientes entre sí:

$$t = \frac{Z}{\sqrt{\chi^2(S)/S}}$$

2. ESTIMACIÓN POR INTERVALOS

- Para esto, en el MRLC2 se requieren dos resultados adicionales que permitirán obtener dos variables aleatorias independientes entre sí, una con distribución chi-cuadrado y una con distribución normal estándar:

$$\frac{\sum_{i=1}^{n} e_i^2}{\sigma_u^2} = \frac{(n-2)S^2}{\sigma_u^2} \sim \chi^2(n-2)$$

$$\frac{\sum_{i=1}^{n} e_i^2}{\sigma_u^2} \textit{ independiente de } \hat{\alpha}, \hat{\beta} \textit{ y } f(\hat{\alpha}, \hat{\beta})$$

2. ESTIMACIÓN POR INTERVALOS

- Para el MRLC2:

$$t = \frac{\hat{\beta} - \beta}{S \Big/ \sqrt{\sum_{i=1}^{n} x_i^2}} \sim t(n-2)$$

- Esta variable aleatoria *t* se usa para la estimación por intervalos y las pruebas de hipótesis. En particular, se puede establecer que:

$$P(|t| \leq t_{\varepsilon/2}) = (1-\varepsilon)$$

$$P(-t_{\varepsilon/2} \leq t \leq t_{\varepsilon/2}) = (1-\varepsilon)$$

2. ESTIMACIÓN POR INTERVALOS

- El estimador por intervalos para β es:

$$\beta \in [\,\hat{\beta} - t_{\varepsilon/2}\, \widehat{s.e.}(\hat{\beta}) \leq \beta \leq \hat{\beta} + t_{\varepsilon/2}\, \widehat{s.e.}(\hat{\beta})\,] = 1 - \varepsilon$$

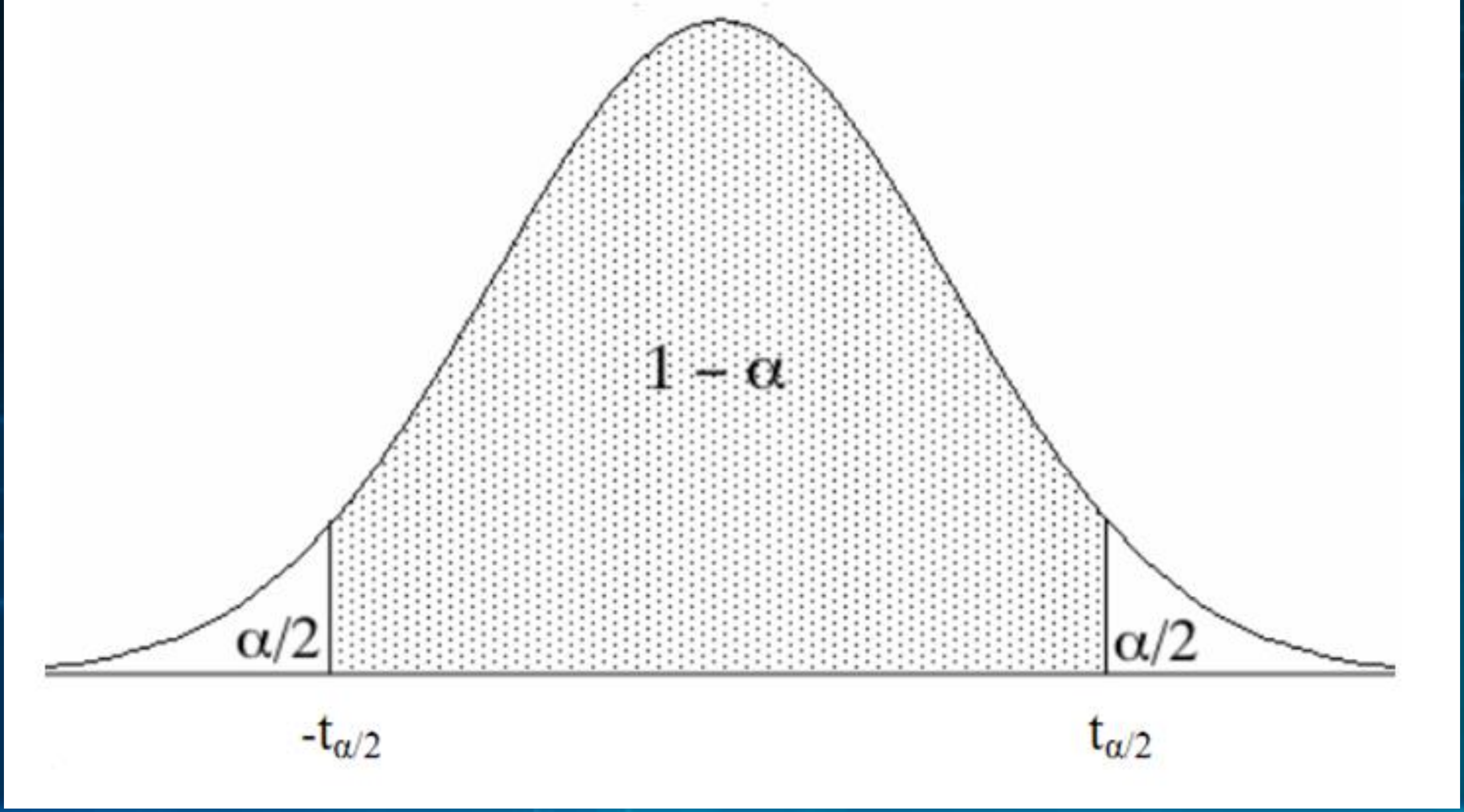
- El estimador por intervalos de β está dado por:

$$\beta \in \left[\hat{\beta} \pm t_{\varepsilon/2} \frac{S}{\sqrt{\sum_{i=1}^{n} x_i^2}} \right] = \left[\hat{\beta} \pm t_{\varepsilon/2}\, \widehat{s.e.}(\hat{\beta}) \right]$$

- Este intervalo, con puntos aleatorios, tiene **1- ε** de probabilidad de contener el verdadero pero desconocido parámetro poblacional **β**. Cuando se conocen exactamente los valores límites del intervalo, se denomina intervalo de confianza al **100(1-ε)** por ciento para **β**

2. ESTIMACIÓN POR INTERVALOS

- Significado de: $P(|t| \leq t_{\varepsilon/2}) = (1-\alpha)$



2. ESTIMACIÓN POR INTERVALOS

- La interpretación del estimador por intervalos, que para cada muestra específica genera un intervalo de confianza, se hace en el contexto del ***muestreo repetido***:
  - Si se eligen muchas muestras de tamaño *S* cada una, se estima β y luego se calcula el intervalo estimado o intervalo de confianza para cada muestra, entonces el $(1-\varepsilon)100$ por ciento de todos los intervalos deberían contener el verdadero parámetro poblacional.

- Así, la justificación del uso de los estimadores por intervalos se basa en las propiedades de éste procedimiento en muestras repetidas y no en un intervalo de confianza específico para una muestra.

2. ESTIMACIÓN POR INTERVALOS

- ***Para el caso del parámetro poblacional "α", tenemos:***

$$t = \frac{\hat{\alpha} - \alpha}{S\sqrt{\frac{1}{n} + \left(\bar{X} \Big/ \sum_{i=1}^{n} x_i^2 \right)}} \sim t(\text{n-2})$$

$$P\left(-t_{\varepsilon/2} < t < t_{\varepsilon/2}\right) = (1 - \varepsilon)$$

El intervalo de confianza al 100(1-ε) por ciento para α es:

$$\alpha \in \left[\hat{\alpha} \pm t_{\varepsilon/2} S \sqrt{\frac{1}{n} + \frac{\bar{X}}{\sum_{i=1}^{n} x_i^2}} \right] = \left[\hat{\alpha} \pm t_{\varepsilon/2}\, \widehat{s.e.}(\hat{\alpha}) \right]$$

2. ESTIMACIÓN POR INTERVALOS

- Para el caso del parámetro de la varianza poblacional del término de perturbación σ^2, se tiene*:*

$$\chi = \frac{S^2(n-2)}{\sigma_u^2} \sim \chi^2(\text{n-2})$$

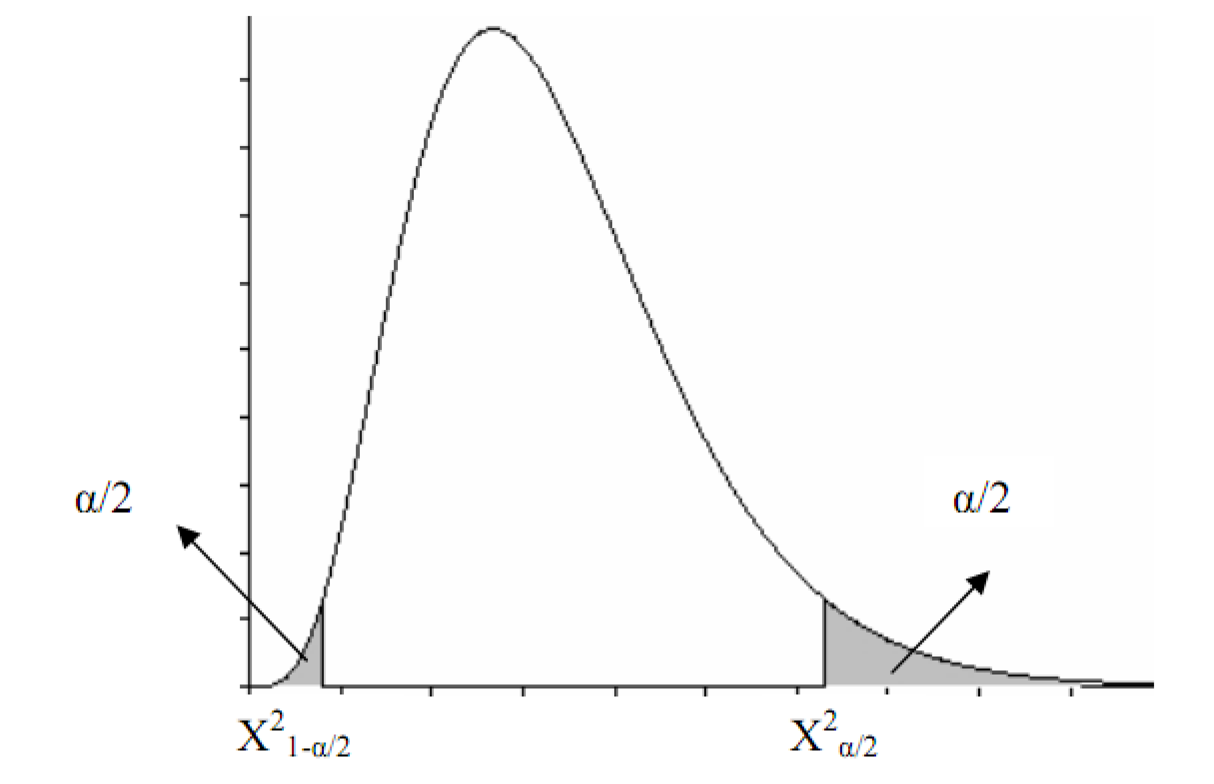
$$P\left(\chi^2_{\varepsilon/2} < \chi < \chi^2_{1-(\varepsilon/2)}\right) = (1-\varepsilon)$$

- El intervalo de confianza al 100(1-ε) por ciento para σ^2 es:

$$\sigma_u^2 \in \left[\frac{(n-2)S^2}{\chi^2_{1-(\varepsilon/2)}}, \frac{(n-2)S^2}{\chi^2_{\varepsilon/2}}\right]$$

2. ESTIMACIÓN POR INTERVALOS

- Significado de: $P\left(\chi^2_{\varepsilon/2} < \chi < \chi^2_{1-(\varepsilon/2)}\right) = (1-\varepsilon)$



2. ESTIMACIÓN POR INTERVALOS

- Para el *MRLCK*, bajo el supuesto de normalidad de las perturbaciones en muestras pequeñas: $\hat{\beta} \sim N[\beta, \sigma^2 (X'X)^{-1}]$. Así:

$$\frac{(\hat{\beta}-\beta)'(X'X)(\hat{\beta}-\beta)}{\sigma_u^2} = \frac{u'Qu}{\sigma_u^2} \sim \chi^2(k)$$

$$\frac{e'e}{\sigma_u^2} = \frac{u'Mu}{\sigma_u^2} \sim \chi^2(n-k)$$

- Dado que M y Q son ortogonales, entonces:

$$\frac{u'Qu}{\sigma_u^2} \; y \; \frac{u'Mu}{\sigma_u^2} \; son \; independientes$$

2. ESTIMACIÓN CONJUNTA POR INTERVALOS

- Se construye un intervalo *k* dimensional para los parámetros utilizando como variable pivote un estadístico *F*.

$$F = \frac{\dfrac{\dfrac{(\hat{\beta}-\beta)'(X'X)(\hat{\beta}-\beta)/k}{\sigma_u^2} \sim \chi^2(k)}{\dfrac{e'e/n-k}{\sigma_u^2} \sim \chi^2(n-k)} = \frac{u'Qu/k}{u'Mu/(n-k)} \sim F(k,n\text{-}k)$$

- A partir de dos estimadores por intervalos, se obtiene un estimador conjunto al 0.95*0.95=0.9025 de confianza. Por ello, es necesario “agrandar” la región de confianza de este estimador.
- Esto genera una región de confianza en forma de una ***elipse.***

2. ESTIMACIÓN CONJUNTA POR INTERVALOS

- Los resultados de una prueba de hipótesis conjunta pueden ser diferentes a los de cada una de las hipótesis individuales que conforman la hipótesis conjunta.

3. PRUEBAS DE HIPÓTESIS

- Dado el MRLCK $y = X\beta + u$, si se estima por MCO se obtiene:

$$\hat{y} = X\hat{\beta}_{MCO}$$

- El siguiente paso es analizar la ***significancia de los parámetros estimados*** a través de:
  - Pruebas de ***Hipótesis Individuales***, que pueden involucrar un solo parámetro o una combinación lineal de parámetros poblacionales.
  - Pruebas de ***Hipótesis Conjuntas***, que involucran más de una Hipótesis Individual.

3. PRUEBAS DE HIPÓTESIS

1.1. <u>Hipótesis Individuales Lineales sobre "β"</u>.

- Pueden ser agrupadas en los siguientes casos generales:

1. $H_0 : \beta_i = 0$
2. $H_0 : \beta_i = \beta_{i0}$
3. $H_0 : \beta_2 + \beta_3 = 1$
4. $H_0 : \beta_3 = \beta_4 \quad o \quad H_0 : \beta_3 - \beta_4 = 0$

- En todos los casos, cada una de estas hipótesis puede ser evaluada a través de un estadístico ***t-student***.

3. PRUEBAS DE HIPÓTESIS

1.2. <u>Hipótesis Conjuntas Lineales sobre "β"</u> .

- Las diversas hipótesis lineales conjuntas que pueden realizarse sobre los β's del *MRLCK* abarcan los siguientes casos:

  5. Prueba de significancia conjunta (excepto el intercepto).

$$H_0 : \begin{bmatrix} \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

3. PRUEBAS DE HIPÓTESIS

6. Prueba de significancia parcial:

$$H_0 : \begin{bmatrix} \beta_{k_1+1} \\ \beta_{k_1+2} \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

7. Hipótesis que establezcan simultáneamente dos o más hipótesis individuales como las descritas antes:

$$H_0 : \begin{bmatrix} \beta_2 \\ \beta_2 + \beta_3 \\ \beta_3 - \beta_4 \\ \beta_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

3. PRUEBAS DE HIPÓTESIS

- Todas estas hipótesis conjuntas lineales, así como las hipótesis individuales planteadas antes, pueden expresarse como:

$$H_0 : R\beta - r = 0 \quad ó \quad H_0 : R\beta = r$$

- donde ***R*** es una matriz “q x k” (q= número de hipótesis individuales) y ***r*** un vector de constantes de orden “k x 1”.
- El tipo de hipótesis lineal a evaluar determinará la forma específica de ***R*** y ***r***.

3. PRUEBAS DE HIPÓTESIS

- La estimación por intervalos o construcción de intervalos de confianza es una herramienta de estimación. Una técnica para hacer inferencia estadística está dada por las ***pruebas de hipótesis***.
- Elementos: ***Una hipótesis nula, una hipótesis alternativa, un contraste (test) o prueba estadística*** y una ***región de rechazo***.

Hipótesis Nula

- Es una creencia que se mantiene hasta que la evidencia muestral demuestre que no es así. Por ejemplo:

$$H_0 : \beta = 0$$

3. PRUEBAS DE HIPÓTESIS

Hipótesis Alternativa

- Puede ser a una cola o a dos colas.

Contraste (Test) Estadístico

- La información muestral sobre la H_0 está contenida en el valor muestral de un ***estadístico*** (que es una variable aleatoria).

- Se caracteriza porque su ***función*** de ***densidad de probabilidades debe conocerse completamente cuando la hipótesis nula es cierta***, y debe tener ***otra forma*** si la hipótesis nula no es cierta.

3. PRUEBAS DE HIPÓTESIS

La Región de Rechazo

- En la práctica, se define como el conjunto de valores del estadístico tales que ***cuando la H_0 es cierta***, tienen ***poca probabilidad*** de ser observados.
- Solamente es posible construirla si se conoce la función de densidad del estadístico cuando es cierta la hipótesis nula.
- La interpretación del resultado de una prueba estadística debe considerar el siguiente criterio: ***la hipótesis nula no necesariamente es cierta si para el contraste estadístico se obtiene un valor muestral que se encuentre en la región de no rechazo***.

3. PRUEBAS DE HIPÓTESIS

- Por ejemplo, si el verdadero valor del parámetro poblacional fuera muy cercano del establecido por la hipótesis, la probabilidad de que el valor del contraste estadístico pertenezca a la ***región de no rechazo*** es ***alta***.
- En este caso, no se rechaza la Hipótesis Nula a pesar de que sea falsa (Error Tipo II).

- ***<u>REGLA</u>: Si el valor del estadístico cae en la región de rechazo (colas de la distribución t), se rechaza la hipótesis nula. En caso contrario, no se puede rechaza la hipótesis nula.***

3. PRUEBAS DE HIPÓTESIS

Errores Tipo I y II

Error Tipo I: La H_0 es cierta y se decide ***rechazarla. Tamaño del Test***

Error Tipo II: La H_0 es falsa y se decide ***aceptarla. Poder del Test***

- La ***probabilidad del Error Tipo I*** se denomina ***nivel de significancia o*** ε**.** Cuando la H_0 es cierta, el estadístico toma un valor en la región de rechazo (colas) con una probabilidad igual a ε**.**

- Así, este procedimiento ***rechazará una*** H_0 ***"verdadera" con una probabilidad (o nivel de significancia) igual a*** ε**.**

- El investigador controla la probabilidad del Error Tipo I eligiendo el nivel de significancia del contraste: 0.01, 0.05 y 0.10.

3. PRUEBAS DE HIPÓTESIS

- ***El Error Tipo II no es controlable ni calculable***, y depende del verdadero valor (desconocido) del parámetro poblacional:
  - La probabilidad del Error Tipo II varía inversamente con el nivel de significancia del test o prueba estadística.
  - El ***test pierde poder*** para discriminar entre el verdadero valor del parámetro y el valor establecido erróneamente en la hipótesis nula si estos valores son muy parecidos; es decir, el Error Tipo II es mayor.
  - Dado un nivel de significancia, a mayor tamaño de muestra, menor la probabilidad de cometer el Error Tipo II.
  - No existe un mejor test, en el sentido que tenga el menor Error Tipo II para cada nivel de Error Tipo I.

3. PRUEBAS DE HIPÓTESIS

P-value o Probabilidad

- Es la probabilidad de que el valor del estadístico ***t*** pueda ser mayor o igual al ***valor absoluto del valor muestral del estadístico***.

- ***Regla***:
  - Cuando el ***p-value*** *de una prueba de hipótesis sea menor al nivel de significancia, entonces se rechaza la hipótesis nula.*
  - *En caso contrario,* no podemos rechazar la hipótesis nula.

4. HIPÓTESIS INDIVIDUALES: ESTADÍSTICO "t"

- Para probar la ***hipótesis nula*** de que β es igual a β_0 , H_0: $\beta = \beta_0$, utilizamos el ***estadístico calculado***:

$$|t| = \left| \frac{\hat{\beta} - \beta_0}{\overset{\wedge}{s.e.}(\hat{\beta})} \right| > t_{\varepsilon/2}(S\text{-}2)$$

- Si se ***cumple*** la ***desigualdad***, ***rechazamos la hipótesis nula*** a un ***nivel de significancia de 100ε %***; es decir, con una probabilidad de cometer el Error Tipo I de **100ε %**.
- Si ***no se cumple***, ***no podemos rechazar la hipótesis nula*** a un ***nivel de significancia de 100ε %***.

4. HIPÓTESIS INDIVIDUALES: ESTADÍSTICO "t"

- El Intervalo de Confianza y una prueba de hipótesis son las "caras opuestas de una misma moneda":

  - Si el Test Estadístico nos dice que rechacemos la H_0 a un nivel de 100 ε, de significancia, entonces β está fuera del intervalo construido al $100(1-\varepsilon)$ por ciento de confianza.

  - Si β está dentro del intervalo construido al $100(1-\varepsilon)$ por ciento de confianza, entonces el test estadístico nos dirá que ***no podemos rechazar H_0 a un nivel de 100 ε de significancia***.

4. HIPÓTESIS INDIVIDUALES: ESTADÍSTICO "t"

- Para probar la hipótesis nula de que α es igual a α_0 , H_0: $\alpha = \alpha_0$, utilizamos:

$$|t| = \left| \frac{\hat{\alpha} - \alpha_0}{\widehat{s.e.}(\hat{\alpha})} \right| > t_{\varepsilon/2}(S - 2)$$

- Si se ***cumple*** la ***desigualdad***, ***rechazamos la hipótesis nula*** a un ***nivel de significancia de 100ε %***; es decir, con una probabilidad de cometer el Error Tipo I de **100ε %.**
- Si ***no se cumple, no podemos rechazar la hipótesis nula*** a un ***nivel de significancia de 100ε %.***

4. HIPÓTESIS INDIVIDUALES: ESTADÍSTICO "t"

- Los estadísticos t-student y F de Fisher están relacionados:

$$t^2(m) = F(1,m)$$

- Rechazar una hipótesis nula es una conclusión más fuerte que no poder rechazarla:
  - En este último caso solo es posible afirmar que la información muestral es compatible con la hipótesis nula.
  - Usualmente, la hipótesis nula se establece de tal forma que, si la teoría que se utiliza es cierta, entonces se rechazará la hipótesis nula.

4. HIPÓTESIS INDIVIDUALES: ESTADÍSTICO "t"

Aplicación: La Demanda por Dinero en el Perú

Dependent Variable: LCR
Method: Least Squares
Sample: 1991:06 1999:12
Included observations: 103

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LY94	1.760349	0.055706	31.60064	0.0000
C	-5.241600	0.258403	-20.28463	0.0000

R-squared	0.908148	Mean dependent var	2.920373
Adjusted R-squared	0.907239	S.D. dependent var	0.259639
S.E. of regression	0.079077	Akaike info criterion	-2.217554
Sum squared resid	0.631576	Schwarz criterion	-2.166394
Log likelihood	116.2040	F-statistic	998.6004
Durbin-Watson stat	0.503445	Prob(F-statistic)	0.000000

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

5.1. Hipótesis Individuales Lineales sobre "β".

- Pueden ser agrupadas en los siguientes casos generales:

1. $H_0 : \beta_i = 0$
2. $H_0 : \beta_i = \beta_{i0}$
3. $H_0 : \beta_2 + \beta_3 = 1$
4. $H_0 : \beta_3 = \beta_4 \quad o \quad H_0 : \beta_3 - \beta_4 = 0$

- En todos los casos, cada una de estas hipótesis puede ser evaluada a través de un estadístico ***t-student***.

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

5.2. Hipótesis Conjuntas Lineales sobre "β" .

- Las diversas hipótesis lineales conjuntas que pueden realizarse sobre los β's del *MRLCK* abarcan los siguientes casos:

  5. Prueba de significancia conjunta (excepto el intercepto).

$$H_0 : \begin{bmatrix} \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

6. Prueba de significancia parcial:

$$H_0 : \begin{bmatrix} \beta_{k_1+1} \\ \beta_{k_1+2} \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

7. Hipótesis que establezcan simultáneamente dos o más hipótesis individuales como las descritas antes:

$$H_0 : \begin{bmatrix} \beta_2 \\ \beta_2 + \beta_3 \\ \beta_3 - \beta_4 \\ \beta_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

- Todas estas hipótesis conjuntas lineales, así como las hipótesis individuales planteadas antes, pueden expresarse como:

$$H_0 : R\beta - r = 0 \quad ó \quad H_0 : R\beta = r$$

- donde ***R*** es una matriz “q x k” (q= número de hipótesis individuales) y ***r*** un vector de constantes de orden “q x 1”.
- El tipo de hipótesis lineal a evaluar determinará la forma específica de ***R*** y ***r***.

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

Caso 1: $H_0 : \beta_i = 0$

$$R\beta = r$$

$$\begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$$

$$\beta_i = 0$$

Caso 2: $H_0 : \beta_i = \beta_{i0}$

$$R\beta = r$$

$$\begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} \beta_{i0} \end{bmatrix}$$

$$\beta_i = \beta_{i0}$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

Caso 3: $H_0 : \beta_2 + \beta_3 = 1$

$$R\beta = r$$

$$\begin{bmatrix} 0 & 1 & 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix}$$

$$\beta_2 + \beta_3 = 1$$

Caso 4: $H_0 : \beta_3 = \beta_4 \quad o \quad H_0 : \beta_3 - \beta_4 = 0$

$$R\beta = r$$

$$\begin{bmatrix} 0 & 0 & 1 & -1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$$

$$\beta_3 - \beta_4 = 0$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

Caso 5:

$$H_0 : \begin{bmatrix} \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$R\beta = r$$

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \vdots \\ \vdots & \vdots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

Caso 6:

$$H_0 : \begin{bmatrix} \beta_{k_1+1} \\ \beta_{k_1+2} \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$R\beta = r$$

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & \vdots & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & \vdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \vdots & 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & 1 & 0 \\ 0 & 0 & \cdots & 0 & \vdots & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{k_1} \\ \beta_{k_1+1} \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

Caso 7:

$$H_0 : \begin{bmatrix} \beta_2 \\ \beta_2 + \beta_3 \\ \beta_3 - \beta_4 \\ \beta_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

$$R\beta = r$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

- Para probar cada hipótesis podemos usar como ***pivote***: $R\hat{\beta} - r$
- Para encontrar el estadístico adecuado para probar las hipótesis, necesitamos encontrar la **distribución muestral relevante** de la variable pivote ***bajo la hipótesis nula***. Así, tenemos que: $\hat{\beta} \sim \mathrm{N}\left(\beta, \sigma_u^2 (X'X)^{-1}\right)$
- A partir de lo cual deducimos:

$$R(\hat{\beta} - \beta) \sim \mathrm{N}\left(0, \sigma_u^2 R(X'X)^{-1} R'\right)$$

- Si ***se cumple la hipótesis nula***, $H_0 : R\beta = r$, tenemos:

$$(R\hat{\beta} - r) \sim \mathrm{N}\left(0, \sigma_u^2 R(X'X)^{-1} R'\right)$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

- Para evitar la presencia de la varianza poblacional, se puede construir un estadístico F a partir de dos estadísticos ***chi-cuadrado***:

$$(R\hat{\beta}-r)'\left[\sigma_u^2 R(X'X)^{-1}R'\right]^{-1}(R\hat{\beta}-r) \sim \chi^2(q)$$

$$\frac{e'e}{\sigma_u^2} \sim \chi^2(n-k)$$

- El estadístico *F* tendrá la siguiente forma final:

$$F=(R\hat{\beta}-r)'\left[\hat{\sigma}^2 R(X'X)^{-1}R'\right]^{-1}(R\hat{\beta}-r)/q \sim \mathrm{F}(q,n-k)$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

- En la fórmula del estadístico F tenemos la matriz de varianzas y covarianzas del vector de parámetros estimados:

$$\hat{Var}(\hat{\beta}) = \hat{\sigma}^2 (X'X)^{-1} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & \cdots & c_{kk} \end{bmatrix}_{(kxk)}$$

- El estadístico *F* tomará diferentes formas, dependiendo el caso que se analice:

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

Caso 1:

$$F = \frac{\hat{\beta}_i^2}{c_{ii}} = \frac{\hat{\beta}_i^2}{\overset{\wedge}{Var}(\hat{\beta}_i)} \sim F(1, n-k)$$

$$F = \left(\frac{\hat{\beta}_i}{\sqrt{c_{ii}}}\right)^2 = \left(\frac{\hat{\beta}_i}{\overset{\wedge}{s.e.}(\hat{\beta}_i)}\right)^2 = t^2$$

Caso 2:

$$F = \frac{(\hat{\beta}_i - \beta_{i0})^2}{c_{ii}} = \frac{(\hat{\beta}_i - \beta_{i0})^2}{\overset{\wedge}{Var}(\hat{\beta}_i)} \sim F(1, n-k)$$

$$F = \left(\frac{\hat{\beta}_i - \beta_{i0}}{\sqrt{c_{ii}}}\right)^2 = t^2$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

Caso 3:

$$F = \frac{(\hat{\beta}_2 + \hat{\beta}_3 - 1)^2}{\widehat{Var}(\hat{\beta}_2 + \hat{\beta}_3)} \sim F(1, n-k)$$

$$F = \left(\frac{\hat{\beta}_2 + \hat{\beta}_3 - 1}{\widehat{s.e.}(\hat{\beta}_2 + \hat{\beta}_3)} \right)^2 = t^2$$

Caso 4:

$$F = \frac{(\hat{\beta}_3 - \hat{\beta}_4)^2}{\widehat{Var}(\hat{\beta}_3 - \hat{\beta}_4)} \sim F(1, n-k)$$

$$F = \left(\frac{\hat{\beta}_3 - \hat{\beta}_4}{\widehat{s.e.}(\hat{\beta}_3 - \hat{\beta}_4)} \right)^2 = t^2$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

Caso 5:

$$F = (\hat{\beta}_2' [\hat{\sigma}^2 (X_*' X_*)^{-1}]^{-1} \hat{\beta}_2)/(k-1)$$

$$F = \frac{(\hat{\beta}_2' X_*' X_* \hat{\beta}_2)/(k-1)}{\hat{\sigma}^2}$$

$$F = \frac{SCE/(k-1)}{e'e/(n-k)} = \frac{SCE/(k-1)}{SCR/(n-k)} \sim F(k-1, n-k)$$

$$F = \frac{(SCT)R^2/(k-1)}{SCR/(n-k)} = \frac{R^2/(k-1)}{\left(\frac{SCR}{SCT}\right)/(n-k)} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} \sim F(k-1, n-k)$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

Caso 6:

$$F = \frac{\hat{\boldsymbol{\beta}}_2^{'}[(\mathbf{X}_2^{'}\mathbf{M}_1\mathbf{X}_2)^{-1}]^{-1}\hat{\boldsymbol{\beta}}_2 / k_2}{e'e/(n-k)} = \frac{\hat{\boldsymbol{\beta}}_2^{'}\mathbf{X}_2^{'}\mathbf{M}_1\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 / k_2}{e'e/(n-k)}$$

$$y = \hat{y} + e$$

$$y = X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + e$$

$$M_1 y = M_1 X_1 \hat{\beta}_1 + M_1 X_2 \hat{\beta}_2 + M_1 e$$

$$M_1 y = M_1 X_2 \hat{\beta}_2 + e$$

$$y' M_1 y = \hat{\beta}_2^{'} X_2^{'} M_1^{'} M_1 X_2 \hat{\beta}_2 + e'e$$

$$e_R^{'} e_R = \hat{\beta}_2^{'} X_2^{'} M_1 X_2 \hat{\beta}_2 + e'e$$

$$F = \frac{e_R^{'}e_R - e'e / k_2}{e^{'}e/(n-k)} \sim F(k_2, n-k)$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

- En la medida que el ***F calculado tome valores muy grandes***, se concluye que las ***restricciones impuestas bajo la hipótesis nula no son válidas***, pues el modelo restringido presentaría una SCR muy grande en comparación a la del modelo sin restricciones o modelo irrestricto.

- El *F* estadístico relevante toma la siguiente forma general:

$$F = \frac{e_R' e_R - e_I' e_I / q}{e_I' e_I / (S - k)}$$

5. HIPÓTESIS CONJUNTAS: ESTADÍSTICO F

- Cuando las restricciones (establecidas en las hipótesis) son verdaderas, estas no son sustentadas exactamente por los datos debido a la presencia de las perturbaciones:
  - Por ello, ***la SCR aumenta cuando se imponen restricciones a un modelo.***
  - ***A pesar de esto, el incremento en la SCR no será grande en relación a la influencia de las perturbaciones.***
  - Por ello, si las restricciones son verdaderas, el F tomaría valores pequeños.

6. APLICACIONES

Dependent Variable: LOG(EARNINGS)
Method: Least Squares
Date: 08/02/11 Time: 13:47
Sample: 1 540
Included observations: 540

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.751657	0.575209	-3.045252	0.0024
LOG(S)	1.291564	0.160356	8.054358	0.0000
LOG(ASVABC)	0.548691	0.144172	3.805811	0.0002
AGE	-0.024005	0.010297	-2.331279	0.0201

R-squared	0.257606	Mean dependent var	2.789022
Adjusted R-squared	0.253451	S.D. dependent var	0.609847
S.E. of regression	0.526927	Akaike info criterion	1.563870
Sum squared resid	148.8215	Schwarz criterion	1.595660
Log likelihood	-418.2450	Hannan-Quinn criter.	1.576303
F-statistic	61.99612	Durbin-Watson stat	1.764502
Prob(F-statistic)	0.000000		

6. APLICACIONES

Coefficient Confidence Intervals
Date: 08/02/11 Time: 13:50
Sample: 1 540
Included observations: 540

		90% CI		95% CI		99% CI	
Variable	Coefficient	Low	High	Low	High	Low	High
C	-1.751657	-2.699430	-0.803884	-2.881598	-0.621716	-3.238592	-0.2647
LOG(S)	1.291564	1.027345	1.555782	0.976561	1.606567	0.877038	1.7060
LOG(ASVABC)	0.548691	0.311139	0.786243	0.265480	0.831902	0.176002	0.9213
AGE	-0.024005	-0.040971	-0.007039	-0.044232	-0.003778	-0.050623	0.0026

Wald Test:
Equation: EQ1

Test Statistic	Value	df	Probability
F-statistic	92.37585	(2, 536)	0.0000
Chi-square	184.7517	2	0.0000

Null Hypothesis: C(2)=C(3)=0
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2)	1.291564	0.160356
C(3)	0.548691	0.144172

Restrictions are linear in coefficients.