Pruebas de Hipótesis Lineal y Estimación con Restricciones Lineal

Fundamentos de Econometría

Juan Palomino[1]

[1]Magister en Economía Aplicada con Mención Estudios Regionales
juan.palominoh@pucp.pe

Departamento de Economía

Índice

1 Pruebas de Hipótesis Lineales

- Prueba de hipótesis sobre un coeficiente
- Prueba de hipótesis lineales
- Test F

2 Estimación sujeto a Restricciones Lineales

- Modelos Restringidos y No Restringidos
- Estimación sujeto a Restricciones Lineales

Objetivo

- Realizar pruebas de hipótesis que involucren a más de un parámetro al mismo tiempo.
- Estimación del modelo de regresión lineal cuando se sujeta a restricciones lineales sobre los parámetros.

Índice

1 Pruebas de Hipótesis Lineales
- Prueba de hipótesis sobre un coeficiente
- Prueba de hipótesis lineales
- Test F

2 Estimación sujeto a Restricciones Lineales
- Modelos Restringidos y No Restringidos
- Estimación sujeto a Restricciones Lineales

Prueba de hipótesis sobre un coeficiente

- Hipótesis sobre parámetros individuales, en particular, sobre alguno de los parámetros en el vector β.
- Supongamos que tenemos una hipótesis sobre el coeficiente del $j-\acute{e}simo$ regresor X_j.
- La hipótesis nula $H_0 : \beta_j = a$ y se usa el estadístico t así:

$$t_j = \frac{\hat{\beta}_j - a}{\sqrt{\hat{\sigma}^2 c_{jj}}} \sim t_{(n-k)}$$

donde c_{jj} es el $j-\acute{e}simo$ elemento de la diagonal de la matriz $(X'X)^{-1}$, y por lo tanto $\sqrt{\hat{\sigma}^2 c_{jj}}$ es la desviación estándar de $\hat{\beta}_j$.

- Si $\left|t_j\right| > t_{1-\frac{\alpha}{2}}(n-K)$, se rechaza la $H_0 : \beta_j = a$.

Índice

1 Pruebas de Hipótesis Lineales
- Prueba de hipótesis sobre un coeficiente
- Prueba de hipótesis lineales
- Test F

2 Estimación sujeto a Restricciones Lineales
- Modelos Restringidos y No Restringidos
- Estimación sujeto a Restricciones Lineales

Prueba de Hipótesis Lineal

- Hipótesis que involucren combinaciones lineales de varios parámetros a la vez.

Prueba de Hipótesis Lineal

Ejemplo 1: Prueba de Rendimientos Constantes

Tenemos la siguiente función de producción Cobb-Douglas:

$$Q = AK^{\beta_2}L^{\beta_3}$$

donde K y L son las cantidades de capital y trabajo que utiliza una empresa, mientras que A, β_2 y β_3 son parámetros tecnológicos.
Tomando logaritmos, la función de producción queda como:

$$lnQ_i = \beta_1 + \beta_2 lnK_i + \beta_3 lnL_i + \varepsilon_i$$

donde $\beta_1 = lnA$.

Prueba de Hipótesis Lineal

Ejemplo 1: Prueba de Rendimientos Constantes

En Microeconomía, la función de producción de Cobb Douglas exhibe rendimientos constantes a escala cuando $\beta_2 + \beta_3 = 1$.
La hipótesis sería:

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

Esta expresión es lineal en los parámetros. En notación matricial, esto sería:

$$H_0 : \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = [1]$$

Prueba de Hipótesis Lineal

Ejemplo 2: Ecuación de Salarios

Tenemos la siguiente ecuación:

$$ln(w_i) = \beta_1 + \beta_2 sch_i + \beta_3 exper_i + \beta_4 tenure_i + \varepsilon_i$$

donde w_i es el salario del individuo i, $exper$ son los años de experiencia en el mercado laboral, $tenure$ son los años de permanencia en el trabajo actual. Probar la hipótesis que los años de experiencia y años de permanencia tienen el mismo impacto sobre los salarios y que los años de educación no tienen efecto. Entonces, la hipótesis contiene dos ecuaciones

$$H_0 : \beta_3 - \beta_4 = 0, \, \beta_2 = 0$$

Prueba de Hipótesis Lineal

Ejemplo 2: Ecuación de Salarios

En notación matricial

$$H_0 : \begin{bmatrix} 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Prueba de Hipótesis Lineal

Ejemplo 3: Ecuación de Salarios

Queremos probar la "significancia de la regresión" en conjunto. Es decir, probemos si todos los parámetros en β, excepto el intercepto son iguales a cero. La hipótesis nula es:

$$H_0 : \beta_2 = 0,\ \beta_3 = 0,\ \beta_4 = 0$$

En notación matricial

$$H_0 : \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

La hipótesis alternativa es que al menos alguno de los parámetros β es estadísticamente distinto de cero.

Prueba de Hipótesis Lineal

- Se puede generalizar la prueba de hipótesis de significancia conjunta en una regresión con k variables:

$$H_0 : \beta_1 = 0,\ \beta_2 = 0, \cdots,\ \beta_k = 0$$
$$H_1 : al\ menos\ alguno\ de\ los\ \beta\ es\ distinto\ de\ cero$$

Prueba de Hipótesis Lineal

Hipótesis Lineal

Una hipótesis nula que incluye a q ecuaciones lineales de los parámetros:

$$H_0 : \underbrace{R}_{(q\times K)} \underbrace{\beta_0}_{(K\times 1)} = \underbrace{r}_{q\times 1}$$

donde valores de R y r son expresiones matriciales que contienen a números fijos, no a variables aleatorias ni parámetros.

Denotamos el número de ecuaciones, que es la dimensión de r, es decir, q. Entonces, R es de dimensión $q\times K$.

Hipótesis Lineal

- Podemos escribir $R\beta_0 = r$ como:

$$\begin{array}{rcl} r_{11}\beta_1 + r_{12}\beta_2 + \cdots r_{1K}\beta_K & = & r_1 \\ r_{21}\beta_1 + r_{22}\beta_2 + \cdots r_{2K}\beta_K & = & r_2 \\ & & \\ r_{q1}\beta_1 + r_{q2}\beta_2 + \cdots r_{qK}\beta_K & = & r_q \end{array}$$

- Las hipótesis nula y alternativa para un test de dos colas de restricciones lineales en los parámetros de regresión en el modelo de regresión lineal $y = X\beta + \varepsilon$ son:

$$H_0: \; R\beta_0 - r = 0$$
$$H_a: \; R\beta_0 - r \neq 0$$

- Se requiere que el rango de R es q.

Índice

1 Pruebas de Hipótesis Lineales
- Prueba de hipótesis sobre un coeficiente
- Prueba de hipótesis lineales
- Test F

2 Estimación sujeto a Restricciones Lineales
- Modelos Restringidos y No Restringidos
- Estimación sujeto a Restricciones Lineales

Test F

- La prueba t es insuficiente.
- Se requiere el estadístico F, el cual sigue una distribución F de Fisher con q grados de libertad en el numerador y $n-k$ grados de libertad en el denominador.

Test F

- Bajo la hipótesis nula, $R\beta_0 - r = 0$, el valor esperado de $R\hat{\beta} - r$ es:

$$E(R\hat{\beta} - r \mid X) = RE(\hat{\beta} \mid X) - r = R\beta_0 - r = 0$$

- Asimismo, la varianza de $R\hat{\beta} - r$ es:

$$\begin{aligned} Var(R\hat{\beta} - r \mid X) &= Var(R\hat{\beta} \mid X) \\ &= R[Var(\hat{\beta} \mid X)]R' \\ &= \sigma_0^2 R(X'X)^{-1}R' \end{aligned}$$

Test F

- Bajo el supuesto de normalidad, sabemos que:

$$\hat{\beta} \sim N(\beta_0, \sigma_0^2 (X^{'}X)^{-1})$$

- Asimismo, usando el Teorema de combinación lineal de normales:

$$R\hat{\beta} - r \sim N[0, \sigma_0^2 R(X^{'}X)^{-1}R^{'}]$$

- Explicaremos el Teorema para la distribución del estadístico F.

Test F

Teorema: Distribución del estadístico F

Bajo la hipótesis nula $H_0 : R\beta = r$, donde R es de dimensión $q \times K$ con $rango(R) = q$, el estadístico F se define como:

$$
\begin{aligned}
F &\equiv \frac{(R\hat{\beta} - r)^{'}[R(X^{'}X)^{-1}R^{'}]^{-1}(R\hat{\beta} - r)/q}{\hat{\sigma}^2} \\
&\equiv (R\hat{\beta} - r)^{'}[RVar(\hat{\beta} \mid X)R^{'}]^{-1}(R\hat{\beta} - r)/q
\end{aligned}
$$

es distribuido como $F(q, n-K)$.

Demostración del Test F

- Ya que $\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{(n-K)}$, escribimos:

$$F = \frac{(R\hat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\hat{\beta} - r)/q}{\hat{\varepsilon}'\hat{\varepsilon}/(n-K)}$$

- Dividiendo en el numerador y denominador por σ_0^2, tenemos:

$$\begin{aligned} F &= \frac{(R\hat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\hat{\beta} - r)/\sigma_0^2 q}{\hat{\varepsilon}'\hat{\varepsilon}/\sigma_0^2(n-K)} \\ &= \frac{w/q}{s/(n-K)} \end{aligned}$$

donde $w = (R\hat{\beta} - r)'[\sigma_0^2 R(X'X)^{-1}R']^{-1}(R\hat{\beta} - r)$ y $s = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma_0^2}$

Demostración del Test F

- Necesitamos mostrar que:

1. $w \mid X \sim \chi^2(q)$
2. $s \mid X \sim \chi^2(n-K)$
3. w y s son distribuidos independientemente condicionado sobre X.

Demostración del Test F

- Entonces, por el siguiente Teorema:

Teorema: Ratio de Variables chi-cuadrado

Si x_1 y x_2 son variables independientes chi-cuadrado con n_1 y n_2 grados de libertad, respectivamente. Entonces el ratio

$$F(n_1,n_2) = \frac{x_1/n_1}{x_2/n_2}$$

tiene la distribución F con n_1 y n_2 grados de libertad.

Demostración del Test F

- El estadístico F se distribuye como $F(q, n-K)$.

1. Sea $v = R\hat{\beta} - r$. Ya se ha mostrado que $R\hat{\beta} - r \sim N[0, \sigma_0^2 R(X^{'}X)^{-1}R^{'}]$. Por lo tanto, w puede ser escrito como $v^{'}Var(v \mid X)^{-1}v$, ya que R es de rango completo y $X^{'}X$ es no singular, $\sigma^2 R(X^{'}X)^{-1}R^{'}$ es no singular. Asimismo, por la definición de distribución χ^2, $w \mid X \sim \chi^2(q)$.
2. w es una función de $\hat{\beta}$ y s es una función de $\hat{\varepsilon}$. Pero $\hat{\beta}$ y $\hat{\varepsilon}$ son distribuidos independientes condicionado sobre X. Por lo tanto, w y s son distribuidos independientes condicionado sobre X.

Test F

- Si la H_0 es cierta, se esperaría que $R\hat{\beta} \approx r$ y por lo tanto el valor estadístico F debería ser cercano a cero.
- Si la H_0 no es cierta, entonces $R\hat{\beta} \neq r$ y también sería muy probable que $R\hat{\beta}$ sea muy distinto de r, con lo cual el F estadístico podría tomar valores bastantes grandes.

Por ello, grandes valores de F serían una señal del no cumplimiento de la hipótesis nula.

Test F

- La regla de decisión del test F al nivel de significancia de α es como sigue:

1. Calcular el estadístico F
2. Ir a la tabla de la distribución F y buscar la entrada para q (el numerador de los grados de libertad) y $n-K$ (el denominador de los grados de libertad).
3. Encontrar el valor crítico $F_{\alpha}(q,n-K)$ que toma α para la cola superior de la distribución F:
    1. Si $F > F_{\alpha}(q,n-K)$, se rechaza la H_0 con $\alpha\%$ de significancia.
    2. Si $F < F_{\alpha}(q,n-K)$, no se rechaza la H_0 con $\alpha\%$ de significancia.

Estimación sujeto a Restricciones Lineales

- Estimar los coeficientes imponiendo la restricción $R\hat{\beta} = r$
- Dos formas:
  - Reemplazando $R\hat{\beta} = r$ en $y = X\hat{\beta} + \varepsilon$ y minimizando la SCR
  - Minimizando la SCR sujeta a $R\hat{\beta} = r$

Índice

1 Pruebas de Hipótesis Lineales
- Prueba de hipótesis sobre un coeficiente
- Prueba de hipótesis lineales
- Test F

2 Estimación sujeto a Restricciones Lineales
- Modelos Restringidos y No Restringidos
- Estimación sujeto a Restricciones Lineales

Modelos Restringidos y No Restringidos

Ejemplo 1: Prueba de Rendimientos Constantes

Del modelo linealizado de la función de producción Cobb-Douglas. Llamando $Y_i = lnQ_i$, $X_{2i} = lnK_i$, $X_{3i} = lnL_i$, el modelo irrestricto es:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \hat{\varepsilon}_i$$

Ahora, estimamos el modelo imponiendo una restricción: $\hat{\beta}_2 + \hat{\beta}_3 = 1$. Pensamos en un modelo estimado como:

$$Y_i = \tilde{\beta}_1 + \tilde{\beta}_2 X_{2i} + (1 - \tilde{\beta}_2) X_{3i} + \tilde{\varepsilon}_i$$

Reordenando, obtenemos el modelo restringido:

$$Y_i - X_{3i} = \tilde{\beta}_1 + \tilde{\beta}_2 (X_{2i} - X_{3i}) + \tilde{\varepsilon}_i$$

Esto es un modelo bivariado en donde la variable endógena es $Y_i - X_{3i} = lnQ_i - lnL_i = ln(Q_i/L_i)$ y la variable exógena es igual a la expresión $X_{2i} - X_{3i} = ln(K_i/L_i)$.

Modelos Restringidos y No Restringidos

Ejemplo 1: Prueba de Rendimientos Constantes

La estimación por MCO entrega los valores $\tilde{\beta}_1$ y $\tilde{\beta}_2$, que son los estimadores sujetos a la restricciones. El parámetro $\tilde{\beta}_3$ se obtiene de la condición $\tilde{\beta}_2 + \tilde{\beta}_3 = 1$.

Alternativamente, otro método es la minimización de SCR sujeta a la restricción:

$$\min_{\tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3} \sum_{i=1}^{n} \tilde{\varepsilon}_i^2 = \sum_{i=1}^{n} (Y_i - \tilde{\beta}_1 - \tilde{\beta}_2 X_{2i} - \tilde{\beta}_3 X_{3i})^2$$

$$s.a.\ \tilde{\beta}_2 + \tilde{\beta}_3 = 1$$

Índice

1 Pruebas de Hipótesis Lineales
- Prueba de hipótesis sobre un coeficiente
- Prueba de hipótesis lineales
- Test F

2 Estimación sujeto a Restricciones Lineales
- Modelos Restringidos y No Restringidos
- Estimación sujeto a Restricciones Lineales

Estimación sujeto a Restricciones Lineales

- **Objetivo:** obtener una expresión del estimador MCO con restricciones $\tilde{\beta}$ y su relación con los estimadores MCO sin restricciones $\hat{\beta}$.
- El problema es:

$$min\, \tilde{\varepsilon}' \tilde{\varepsilon} = (y - X\tilde{\beta})'(y - X\tilde{\beta})$$
$$s.a.\, R\tilde{\beta} = r$$

Estimación sujeto a Restricciones Lineales

- Planteamos el Lagrangeano:

$$\begin{aligned}\mathscr{L} &= (y - X\tilde{\beta})'(y - X\tilde{\beta}) - 2\lambda'(R\tilde{\beta} - r) \\ &= y'y - y'X\tilde{\beta} - \tilde{\beta}X'y + \tilde{\beta}X'X\tilde{\beta} - \underbrace{2\lambda'R\tilde{\beta}}_{2\tilde{\beta}R'\lambda} + 2\lambda'r\end{aligned}$$

$$\mathscr{L} = y'y - y'X\tilde{\beta} - \tilde{\beta}X'y + \tilde{\beta}X'X\tilde{\beta} - 2\tilde{\beta}R'\lambda + 2\lambda'r$$

Estimación sujeto a Restricciones Lineales

- Derivando el lagrangiano con respecto a $\tilde{\beta}$ y λ:

$$\frac{\partial \mathscr{L}}{\partial \tilde{\beta}} = -2X^{'}y + 2X^{'}X\tilde{\beta} - 2R^{'}\lambda = 0$$

$$\frac{\partial \mathscr{L}}{\partial \lambda} = -2(R\tilde{\beta} - r) = 0$$

Estimación sujeto a Restricciones Lineales

- De la primera derivada se obtiene:

$$
\begin{aligned}
X'X\tilde{\beta} &= X'y + R'\lambda \\
\tilde{\beta} &= \underbrace{(X'X)^{-1}X'y}_{\hat{\beta}} + (X'X)^{-1}R'\lambda
\end{aligned}
$$

- Multiplicando R:

$$
\begin{aligned}
R\tilde{\beta} &= R\hat{\beta} + R(X'X)^{-1}R'\lambda \\
\lambda &= (R(X'X)^{-1}R')^{-1}(R\tilde{\beta} - R\hat{\beta})
\end{aligned}
$$

Estimación sujeto a Restricciones Lineales

- De la segunda derivada, se obtiene $R\tilde{\beta} = r$. Entonces:

$$\lambda = (R(X^{'}X)^{-1}R^{'})^{-1}(r - R\hat{\beta})$$

- Reemplazando en $\tilde{\beta}$:

$$\tilde{\beta} = \hat{\beta} + (X^{'}X)^{-1}R^{'}(R(X^{'}X)^{-1}R^{'})^{-1}(r - R\hat{\beta})$$

- Esta es la expresión general de los estimados de β por MCO sujeto a cualquier restricción lineal del tipo $R\hat{\beta} = r$.

Estimación sujeto a Restricciones Lineales

- Aplicando el valor esperado

$$\begin{aligned} E[\tilde{\beta}] &= E[\hat{\beta}] + (X^{'}X)^{-1}R^{'}(R(X^{'}X)^{-1}R^{'})^{-1}(r - RE[\hat{\beta}]) \\ &= \beta + (X^{'}X)^{-1}R^{'}(R(X^{'}X)^{-1}R^{'})^{-1}(r - R\beta) \end{aligned}$$

- Los estimadores restringidos son insesgados únicamente si la restricción es correcta; es decir, si es cierto que $R\beta = r$. En cualquier otro caso, los estimadores serán sesgados.

Referencias

- Capítulo 7.3-7.5. Greene, W. (2000). Análisis Econométrico. Tercera Edición. Madrid.
- Capítulo 1.4. Hayashi, F. (2000). Econometrics. Princeton University Press.