Unidad 4. RLM: Inferencia

Erika R. Badillo
erika.badilloen@unaula.edu.co

Facultad de Economía
Universidad Autónoma Latinoamericana

Inferencia sobre parámetros individuales

- Verificación de hipótesis

$$
\begin{aligned}
H_0 &: \beta_j = \beta_{jo} \\
H_A &: \beta_j < \beta_{jo} \text{ ó} \\
&\quad \beta_j \neq \beta_{jo} \text{ ó} \\
&\quad \beta_j > \beta_{jo}
\end{aligned}
$$

Bajo H_o $\frac{\hat{\beta}_j - \beta_{jo}}{\hat{\sigma}_{\hat{\beta}_j}} \sim t_{N-k}$ gdl, y la regla de decisión se establece teniendo en cuenta la hipótesis alternativa, H_A

Por ejemplo, si $H_A : \beta_j < \beta_{jo}$, la RD es rechazar H_o al nivel de significancia α, si

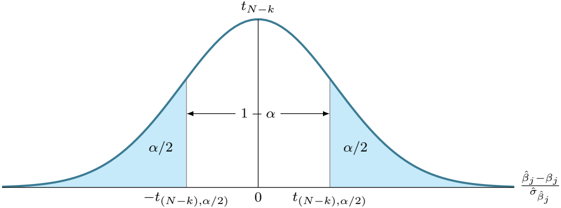
$$t_o = \frac{\hat{\beta}_j - \beta_{jo}}{\hat{\sigma}_{\hat{\beta}_j}} < -t_{(N-k),\alpha}$$

o cuando

$$\textit{p-value} < \alpha$$

con *p-value*$= \int_{-\infty}^{t_o} t_{N-k} dt$

- Se construye el IC(1-α) para β_j



La lectura probabilística de la gráfica dice:

$$Prob\left[-t_{(N-k),\alpha/2)} \leq \frac{\hat{\beta}_j-\beta_j}{\hat{\sigma}_{\hat{\beta}_j}} \leq t_{(N-k),\alpha/2)}\right] = 1-\alpha$$

$$Prob\left[\hat{\beta}_j - \hat{\sigma}_{\hat{\beta}_j} t_{(N-k),\alpha/2)} \leq \beta_j \leq \hat{\beta}_j + \hat{\sigma}_{\hat{\beta}_j} t_{(N-k),\alpha/2)}\right] = 1-\alpha$$

Teniendo en cuenta que en este mundo la aleatoriedad son los límites del intervalo, tenemos que

$$IC(1-\alpha)(\beta_j) = \left[\hat{\beta}_j \pm \hat{\sigma}_{\hat{\beta}_j} t_{(N-k),\alpha/2)}\right]$$

Inferencia sobre combinaciones lineales paramétricas

Es una posibilidad que aumenta el potencial del modelo de RLM. Por ejemplo, una función de producción, en donde Y_i es log de la producción, X_{2i} es el log del capital y X_{3i} es el log del trabajo. Una pregunta que surge, entonces, es si existen rendimientos constantes en la producción. La función tipo Cobb-Douglas sería

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

La hipótesis de rendimientos constantes (la suma de las elasticidades producto-capital y producto-trabajo son iguales a 1) equivale a

$$\beta_2 + \beta_3 = 1$$

Se trata de una combinación lineal y en este caso

$$\mathbf{B} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \quad \mathbf{R} = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$$

y $\mathbf{RB} = 1$ representa la hipótesis a contrastar

Ahora supongamos que la producción sea intensiva en mano de obra, esto podría implicar que la elasticidad del capital es menor que la elasticidad del trabajo, esto implica contrastar la siguiente hipótesis

$$\beta_2 < \beta_3$$
$$\beta_2 - \beta_3 < 0$$

Inferencia sobre combinaciones lineales paramétricas

Otro ejemplo, en una ecuación de ingresos laboral o ecuación salarial de la forma:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{3i}^2 + u_i$$

donde Y_i : log del salario del individuo i
X_{2i} : años de educación
X_{3i} : años de experiencia

Tasa de retorno de la educación $\Longrightarrow \frac{\partial Y_i}{\partial X_{2i}} = \beta_2$

Tasa de retorno de la experiencia $\Longrightarrow \frac{\partial Y_i}{\partial X_{3i}} = \beta_3 + 2\beta_4 X_{3i}$

Se observa entonces que el retorno de la experiencia es una combinación lineal entre los parámetros. Esta combinación lineal se puede representar en términos matriciales:

$$\beta_3 + 2\beta_4 X_{3i} = 0 \Longrightarrow \mathbf{RB} = 0 \Longrightarrow \begin{bmatrix} 0 & 0 & 1 & 2X_{3i} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = 0$$

Inferencia sobre combinaciones lineales paramétricas

Miremos ahora cómo sería el estadístico. Se sabe que $\widehat{\mathbf{B}} \sim N(\mathbf{B}, Cov(\widehat{\mathbf{B}}))$, bajo el supuesto de que $\mathbf{u} \sim N(\mathbf{0}, \sigma_u^2\mathbf{I})$. En consecuencia $\mathbf{R}\hat{\mathbf{B}} \sim N(\mathbf{RB}, \mathbf{R}Cov(\widehat{\mathbf{B}})\mathbf{R}')$. Se trata de una combinación lineal de normales.

Se tiene que:

- $E(\mathbf{R}\widehat{\mathbf{B}}) = \mathbf{R}E(\widehat{\mathbf{B}}) = \mathbf{RB}$
- $$\begin{aligned} Cov(\mathbf{R}\widehat{\mathbf{B}}) &= E[(\mathbf{R}\widehat{\mathbf{B}} - E(\mathbf{R}\widehat{\mathbf{B}}))(\mathbf{R}\widehat{\mathbf{B}} - E(\mathbf{R}\widehat{\mathbf{B}}))'] \\ &= E[(\mathbf{R}\widehat{\mathbf{B}} - \mathbf{RB})(\mathbf{R}\widehat{\mathbf{B}} - \mathbf{RB})'] \\ &= E[\mathbf{R}(\widehat{\mathbf{B}} - \mathbf{B})(\widehat{\mathbf{B}} - \mathbf{B})'\mathbf{R}'] \\ &= \mathbf{R}E[(\widehat{\mathbf{B}} - \mathbf{B})(\widehat{\mathbf{B}} - \mathbf{B})']\mathbf{R}' \\ &= \mathbf{R}E[(\widehat{\mathbf{B}} - \mathbf{B})(\widehat{\mathbf{B}} - \mathbf{B})']\mathbf{R}' \\ &= \mathbf{R}Cov(\widehat{\mathbf{B}})\mathbf{R}' \end{aligned}$$

Estandarizando tenemos

$$\frac{\mathbf{R}\hat{\mathbf{B}} - \mathbf{RB}}{\sqrt{\mathbf{R}Cov(\widehat{\mathbf{B}})\mathbf{R}'}} \sim N(0,1)$$

$$\frac{\mathbf{R}\hat{\mathbf{B}} - \mathbf{RB}}{\sigma_u\sqrt{\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'}} \sim N(0,1)$$

Ya que desconocemos σ_u construimos la t-student. Teniendo en cuenta que

$$\frac{\sum \widehat{u}_i^2}{\sigma_u^2} = \frac{(N-k)\widehat{\sigma}_u^2}{\sigma_u^2} \sim \chi^2_{N-k}$$

Se tiene que

$$\frac{\frac{\mathbf{R}\hat{\mathbf{B}} - \mathbf{RB}}{\sigma_u\sqrt{\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'}}}{\sqrt{\frac{(N-k)\hat{\sigma}_u^2}{\sigma_u^2(N-k)}}} = \frac{\mathbf{R}\hat{\mathbf{B}} - \mathbf{RB}}{\sqrt{\mathbf{R}C\hat{o}v(\widehat{\mathbf{B}})\mathbf{R}'}} \sim t_{N-k} \text{ gdl}$$

Prueba de hipótesis
H_o : $\mathbf{RB} = \mathbf{r}$

$$H_A : \mathbf{RB} \neq \mathbf{r}$$

Bajo H_o $\frac{\mathbf{R\hat{B}}-\mathbf{RB}}{\sqrt{\mathbf{R}\hat{Cov}(\widehat{\mathbf{B}})\mathbf{R}'}} \sim t_{N-k}$ gdl, y la regla de decisión (RD) es rechazar H_o al nivel de significancia α si

$$|t_0| > t_{(N-k),\alpha/2},$$

o cuando

$$p\text{-value} < \alpha$$

Inferencia sobre combinaciones lineales paramétricas

- Prueba de hipótesis H_o : $\mathbf{RB} = \mathbf{r}$

  En el ejemplo de la función de producción Cobb-Douglas y la hipótesis de rendimientos constantes

$$H_o : \beta_2 + \beta_3 = 1$$
$$H_A : \beta_2 + \beta_3 \neq 1$$

  En este caso entonces

$$\mathbf{RB} = \mathbf{r}, \text{ con } \mathbf{R} = [0, 1, 1] \text{ y } \mathbf{r} = 1$$

  El estadístico es de la forma

$$t_o = \frac{\mathbf{R}\hat{\mathbf{B}} - \mathbf{r}}{\sqrt{\mathbf{R}\hat{Cov}(\widehat{\mathbf{B}})\mathbf{R}'}} \text{ y } \mathbf{R}\hat{\mathbf{B}} = \hat{\beta}_2 + \hat{\beta}_3$$

  La matriz var-cov será

$$\mathbf{R}\hat{Cov}(\widehat{\mathbf{B}})\mathbf{R}' = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \hat{\sigma}^2_{\hat{\beta}_1} & \hat{Cov}(\hat{\beta}_1, \hat{\beta}_2) & \hat{Cov}(\hat{\beta}_1, \hat{\beta}_3) \\ \hat{Cov}(\hat{\beta}_1, \hat{\beta}_2) & \hat{\sigma}^2_{\hat{\beta}_2} & \hat{Cov}(\hat{\beta}_2, \hat{\beta}_3) \\ \hat{Cov}(\hat{\beta}_1, \hat{\beta}_3) & \hat{Cov}(\hat{\beta}_2, \hat{\beta}_3) & \hat{\sigma}^2_{\hat{\beta}_3} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{R}\hat{Cov}(\widehat{\mathbf{B}})\mathbf{R}' = \hat{\sigma}^2_{\hat{\beta}_2} + \hat{\sigma}^2_{\hat{\beta}_3} + 2\hat{Cov}(\hat{\beta}_2, \hat{\beta}_3)$$

- Prueba de hipótesis H_o : $\mathbf{RB} = \mathbf{r}$

  El estadístico queda de la forma:

$$t_o = \frac{\hat{\beta}_2 + \hat{\beta}_3 - 1}{\sqrt{\hat{\sigma}^2_{\hat{\beta}_2} + \hat{\sigma}^2_{\hat{\beta}_3} + 2C\hat{o}v(\hat{\beta}_2, \hat{\beta}_3)}} \sim t_{N-k} \text{ gdl}$$

  La RD será rechazar H_o si $|t_o| > t_{(N-k),\alpha/2}$ o si *p-value*$< \alpha$

Inferencia general

Hasta ahora las pruebas de hipótesis involucran una sola restricción lineal (ej. $\beta_1 = 0$ o $\beta_1 = \beta_2$)

Puede suceder que el interés esté en testear conjuntamente múltiples hipótesis sobre nuestros parámetros

Un ejemplo típico es testear "restricciones de exclusión", saber si un grupo de parámetros son todos iguales a cero

Inferencia general

Volviendo a la ecuación salarial de la forma:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{3i}^2 + u_i$$

donde Y_i : log del salario del individuo i
X_{2i} : años de educación
X_{3i} : años de experiencia

Se quiere verificar la hipótesis acerca de la significancia de la experiencia

$$H_o : \beta_3 = 0 \text{ y } \beta_4 = 0 \text{ la experiencia no es significativa}$$
$$H_A : \beta_3 \neq 0 \text{ y } \beta_4 \neq 0 \text{ la experiencia es significativa}$$

En H_o hay dos hipótesis $\beta_3 = 0$ y $\beta_4 = 0$. La representación matricial donde q son el número de hipótesis a contrastar, es

$$\underbrace{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}_{q \times k}} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}}_{\mathbf{B}_{k \times 1}} = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\mathbf{r}_{q \times 1}}$$

No se puede simplemente chequear cada estadístico t separadamente, porque se quiere saber si los q parámetros son **conjuntamente** significativos a un nivel dado (es posible que ninguno sea individualmente significativo a ese nivel)

Inferencia general

El contraste más general:

R es una matriz de dimensión qxk, q representa el número de restricciones lineales que queremos contrastar, k el número de parámetros del modelo.

Así, **R** tendrá tantas columnas como parámetros tenga el modelo y tantas filas como restricciones lineales contenga la hipótesis nula.

Prueba de hipótesis

$$H_o : \mathbf{RB} = \mathbf{r}$$

Donde **r** es un vector de dimensión q que está formado por los términos independientes de las restricciones lineales

$$H_A : \mathbf{RB} \neq \mathbf{r}$$

En este contraste están incluidas las hipótesis acerca de

- parámetros únicos
- combinaciones lineales
- subconjunto de parámetros
- el modelo en su conjunto

Inferencia general

¿Cuál es el estadístico de prueba que permite rechazar o no rechazar la hipóteis nula?

Se puede demostrar que el estadístico de prueba para contrastar las restricciones lineales exactas de igualdad sobre los parámetros es el siguiente:

$$F_o = \frac{\frac{(\mathbf{R}\hat{\mathbf{B}}-\mathbf{r})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\mathbf{B}}-\mathbf{r})}{q}}{\frac{\mathbf{u}'\mathbf{u}}{N-k}} \sim F_{q,N-k}$$

Regla de decisión (RD):

- Si el valor del estadístico de prueba F_o es mayor que el valor crítico de la F de Snedecor con q y $N-k$ gl para un nivel de significancia, α, se rechaza la H_o. Por tanto, las restricciones lineales no son ciertas en el ámbito de la población

  Se rechaza H_o al nivel de significancia α si $F_o > F_{q,N-k}$ o si *p-value*$< \alpha$

- Si F_o es inferior al valor crítico de la tabla, no se puede rechazar la hipótesis nula, lo cual quiere decir que las restricciones son ciertas

Inferencia general

Es posible demostrar que

$$(\mathbf{R}\hat{\mathbf{B}} - \mathbf{r})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\mathbf{B}} - \mathbf{r}) = SCR_o - SCR$$

donde

SCR_o: es la SCR del estimado bajo H_o, es decir, restringido por la hipótesis nula

SCR: es la SCR del estimado bajo H_A, es decir, no restringido, el modelo general

$$\text{Bajo } H_o,\ F_c = \frac{(SCR_o - SCR)/q}{SCR/N-k} \sim F^q_{N-k}$$

Rechazo H_o al nivel de significancia α si $F_c > F^q_{N-k}$ o si *p-value*$< \alpha$

Se requiere estimar tanto el "modelo restringido", como el "modelo no restringido"

El estadístico F siempre es positivo, dado que la SCR del modelo restringido no puede ser menor que SCR del modelo no restringido

Intuitivamente, se está evaluando si el aumento de la SCR, al pasar del modelo no restringido al modelo restringido es suficientemente grande para rechazar H_o

Note que $q =$ número de restricciones, o $df_R - df_{NR}$

La significancia de la regresión globalmente considerada

En el modelo

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + ... + \beta_k X_{2k} + u_i$$

interesa verificar si el conjunto de variables es globalmente significativas, esto es

$$H_o : \beta_2 = \beta_3 = ... = \beta_k = 0 \text{ modelo ingenuo}$$
$$H_A : \text{al menos un } \beta_j \neq 0 \text{ modelo de RLM}$$

Aplicando la fórmula general
$SCR_o = \sum (Y_i - \overline{Y})^2 = SCT$ y $q = k - 1$, por lo tanto

$$\text{Bajo } H_o,\ F_c = \frac{(SCR_o - SCR)/k-1}{SCR/N-k} \sim F_{N-k}^{k-1}$$

Esta expresión es equivalente a

$$F_c = \frac{SEC/k-1}{SCR/N-k} \sim F_{N-k}^{k-1}$$

Regla de decisión rechazar H_o si $F_c > F_{N-k}^{k-1}(\alpha)$ o si *p-value*$< \alpha$

La significancia de la regresión globalmente considerada

Existe una relación entre el anterior F_c y el R^2

$$R^2 = \frac{SEC}{SCT} \Longrightarrow SEC = R^2 SCT$$

$$R^2 = 1 - \frac{SCR}{SCT} \Longrightarrow SCR = (1 - R^2)SCT$$

Entonces

$$F_c = \frac{(N-k)}{(k-1)} \frac{R^2 SCT}{(1-R^2)SCT} = \frac{(N-k)}{(k-1)} \frac{R^2}{1-R^2} \sim F^{k-1}_{N-k}$$

ToDo

Estime el modelo de regresión lineal siguiente a partir de los quince datos dados:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + U_i, \; \forall i = 1, ..., 15$$

i	Y_i	X_{1i}	X_{2i}
1	48.6	10.0	2.3
2	48.8	10.1	2.2
3	49.0	10.4	2.9
4	49.3	10.4	2.7
5	49.8	10.2	2.3
6	52.8	10.8	2.0
7	51.5	10.5	2.1
8	53.1	11.0	2.4
9	49.1	10.3	2.7
10	50.1	10.4	2.6
11	48.5	10.0	2.0
12	49.5	10.3	2.3
13	49.9	10.3	2.0
14	49.0	10.0	2.0
15	53.0	10.8	2.1

1. Cuáles son los elementos en **Y** y **X** cuando el modelo se escribe en la notación matricial $\mathbf{Y} = \mathbf{XB} + \mathbf{u}$?
2. Calcule en términos numéricos lo siguiente (hágalo a mano y muestre todos los pasos)
   - a. $\mathbf{X'X}$, $\mathbf{X'Y}$ y $(\mathbf{X'X})^{-1}$
   - b. El estimador MCO de **B**
   - c. Plantee el modelo estimado $\widehat{\mathbf{Y}_i}$
   - d. La varianza estimada del termino de perturbación $\hat{\sigma}_u^2$
   - e. La matriz de varianzas-covarianzas de $\hat{\mathbf{B}}$, calcule los errores estándar para los coeficientes estimados, los t estadísticos y los intervalos de confianza al 95 %. Analice si los coeficientes son estadísticamente significativos
   - f. El coeficiente de determinación R^2 y el ajustado
   - g. Contraste la significancia conjunta del modelo
3. Use R (o Stata) para verificar los resultados obtenidos anteriormente