(Parcial 1) Estadística Aplicada III

David Isaac López Romero
CU: 173993

①

1) $Y_i = \beta_0 + \beta_1 X_i + U_i$

Utilizando $V(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$

Hallar varianza y covarianza de $\hat{\beta}_0$ y $\hat{\beta}_1$.

Dem:

Escribimos el modelo $Y_i = \beta_0 + \beta_1 X_i + U_i$ en su forma matricial

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{pmatrix} \iff \underline{Y} = X \underline{\beta} + \underline{U}$$

Usamos $V(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$ y obtenemos que

$$V(\hat{\beta}) = \sigma^2 (X^T X)^{-1} = \sigma^2 \left[\begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{pmatrix} \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix} \right]^{-1}$$

$$= \sigma^2 \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1} = \sigma^2 \frac{1}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} \begin{pmatrix} \sum_{i=1}^n X_i^2 & -\sum_{i=1}^n X_i \\ -\sum_{i=1}^n X_i & n \end{pmatrix}$$

Notamos que $\displaystyle Var(\hat{\beta}_0) = \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}$

$$Var(\hat{\beta}_1) = \frac{n \sigma^2}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}$$

Notamos que

$$Var(\hat{\beta}_1) = \frac{\cancel{n} \sigma^2}{\cancel{n} \left\{ \sum_{i=1}^n X_i^2 - n \bar{X}^2 \right\}} \overset{(*)}{=} \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

Asimismo

$$Var(\hat{\beta}_0) = \frac{\sigma^2 \sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \overset{(**)}{=} \frac{\sigma^2 \left\{ \sum_{i=1}^{n}(x_i-\bar{x})^2 + n\bar{x}^2 \right\}}{n\left\{ \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \right\}} = A$$

$$\sum_{i=1}^{n}(x_i-\bar{x})^2 \underset{(*)}{=} \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \Leftrightarrow \sum_{i=1}^{n} x_i^2 \underset{(**)}{=} \sum_{i=1}^{n}(x_i-\bar{x})^2 + n\bar{x}^2$$

continuando de A

$$A \overset{(*)}{=} \frac{\sigma^2 \cancel{\sum_{i=1}^{n}(x_i-\bar{x})^2}}{n\cancel{\sum_{i=1}^{n}(x_i-\bar{x})^2}} + \frac{\sigma^2 \cancel{n}\bar{x}^2}{\cancel{n}\sum_{i=1}^{n}(x_i-\bar{x})^2} = \sigma^2\left\{ \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \right\}$$

Como conclusión,

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad y \quad Var(\hat{\beta}_0) = \sigma^2\left\{ \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \right\}$$

Asimismo, falta $Cov(\beta_0, \beta_1) = Cov(\beta_1, \beta_0)$ que de las expresiones matriciales tenemos que

$$Cov(\beta_0, \beta_1) = \sigma^2 \frac{-\sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \sigma^2 \frac{-\frac{1}{n}\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

$$= \frac{\sigma^2 \bar{x}}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

Finalmente,

$$V(\hat{\beta}) = \begin{pmatrix} \dfrac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} & \dfrac{\sigma^2\bar{x}}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \\ \dfrac{\sigma^2\bar{x}}{\sum_{i=1}^{n}(x_i-\bar{x})^2} & \sigma^2\left\{\dfrac{1}{n} + \dfrac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right\} \end{pmatrix}$$

2) a) $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + U_i$

$SCReg = 80$, $SCT = 100$, $n = 34$, $\alpha = 0.05$.

¿Es significativo el modelo?

Para ver si el modelo es significativo, hacemos prueba de hipótesis

$H_0: \beta_1 = \beta_2 = \beta_3 = 0$ v.s. $H_1: \beta_1^2 + \beta_2^2 + \beta_3^2 \neq 0$

Calculamos el estadístico F^* y

$$F^* = \frac{SCReg / k}{SCRes / (n-(k+1))} = \frac{80/3}{20/(34-4)} = 40$$

// $SCT = SCReg + SCRes \iff SCRes = SCT - SCReg$ //

lo comparamos con $F_{(k, n-(k+1)), \alpha} = F_{(3, 40), 0.05}$

$F_{(3,40), 0.05} = 2.922277$

Vemos que $F^* = 40 \geq 2.922277 = F_{(3,40), 0.05}$ por lo que se rechaza $H_0: \beta_1 = \beta_2 = \beta_3 = 0$, así el modelo explica a la Y.

Si obtuvimos el p-value $= \mathbb{P}(F_{(3,30)} \geq 40)$

p-value $= 1.322736 \times 10^{-10} \approx 0 \leq 0.05 = \alpha$

Entonces, también se rechaza H_0.

b) Si ahora tenemos $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_3 X_{3i} + U_i$ con $SCRes = 25$, $n = 34$ y $\alpha = 0.05$. ¿Cuál es el mejor modelo?

Tenemos la prueba de hipótesis

$H_0: \beta_2 = 0$ v.s. $H_1: \beta_2 \neq 0$

$SCR_{NR} = 20$, $SCR_R = 25$

Modelo no Restringido   Modelo Restringido

Para la prueba

$$F^* = \frac{SCR_R - SCR_{NR} / (k-r)}{SCR_{NR} / (n-(k+1))} = \frac{25-20/(3-1)}{20/(34-4)}$$

$$= \frac{5/2}{20/30} = \frac{15}{4} = 3.75$$

y

$F_{(k-r,\, n-(k+1)),\, \alpha} = F_{(2,30),\, 0.05} = 3.31583$

Como $F^* = 3.75 \geq 3.31583 = F_{(2,30),0.05}$, resulta que rechazamos la hipótesis nula, así $\beta_2 \neq 0$ y tenemos que el modelo no restringido explica mejor a la variable Y.

El mejor modelo es el <u>no</u> <u>restringido</u>, el del inciso (a)

Con el p-value $= \mathbb{P}(F_{(2,30)} \geq 3,75) = 0,03518$

Como p-value $= 0.03518 \leq 0.05 = \alpha$

también es argumento para rechazar Ho y decir que el mejor modelo es el no restringido.

3) $X \sim N(\mu, \sigma^2 I)$, $\mu = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

a) $Y = \dfrac{x^t A x}{\sigma^2}$, $A = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix}$

Vemos que

I) A es simétrica pues,

$$A^T = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix} = A$$

II) A es idempotente pues

$$A^2 = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix} \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix}$$

$$= \begin{pmatrix} 6/9 & -3/9 & -3/9 \\ -3/9 & 6/9 & -3/9 \\ -3/9 & -3/9 & 6/9 \end{pmatrix} = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix} = A$$

III) El rango de A se da por la traza pues A es simétrica e idempotente

$$\text{rango}(A) = tr(A) = \frac{2}{3} + \frac{2}{3} + \frac{2}{3} = \frac{6}{3} = 2 \leq 3$$

IV) Por último, vemos que

$$A\underline{\mu} = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0$$

Como A es simétrica, idempotente, rango $(A) = 2$ y $A\mu = \underline{0}$, por teorema visto en clase resulta que

$$Y = \frac{x^T A x}{\sigma^2} \sim \chi^2_{(2)}$$

3b) $W = \dfrac{x^T(I-A)x}{\sigma^2}$ P.D. W es independiente a Y y hallar distribución de W.

Las hipótesis para $Y = \frac{x^T A x}{\sigma^2}$ ya se cumplen, por (3a), resta ver que W satisface las hipótesis,

$$I - A = \begin{pmatrix} -1/3 & -1/3 & -1/3 \\ -1/3 & -1/3 & -1/3 \\ -1/3 & -1/3 & -1/3 \end{pmatrix}$$

I) $I-A$ es simétrica pues,

$$(I-A)^T = I^T - A^T = I - A$$

II) $I-A$ es idempotente pues

$$(I-A)^2 = I^2 - IA - IA + A^2 \underset{\substack{\uparrow \\ A \text{ idempotente}}}{=} I - A - A + A = I - A$$

III) Como $I-A$ es simétrica e idempotente, el rango de $(I-A)$ se da por la traza

$$\text{rango}(I-A) = tr(I-A) = tr(I) - tr(A) = 3 - 2 = 1$$

IV) También

$$(I-A)\mu = \begin{pmatrix} -1/3 & -1/3 & -1/3 \\ -1/3 & -1/3 & -1/3 \\ -1/3 & -1/3 & -1/3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \underline{0}$$

$\therefore (I-A)\underline{\mu} = \underline{0}$, de aquí obtenemos que $W = \dfrac{x^T(I-A)x}{\sigma^2} \sim \chi^2_{(1)}$

Falta ver el producto entre A e $I-A$

$$A(I-A) = A - A^2 \underset{\substack{\uparrow \\ A \text{ es idempotente}}}{=} A - A = 0$$

$\therefore A(I-A) = 0$

Como se cumplen todas las hipótesis del teorema visto en clase, resulta que

$$Y = \frac{x^T A x}{\sigma^2} \sim \chi^2_{(2)} \quad \text{y} \quad W = \frac{x^T(I-A)x}{\sigma^2} \sim \chi^2_{(1)}$$

las cuales son independientes por teorema

4) $Y \sim N(\mu, \sigma^2 I)$, $\mu = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, H matriz con

$h_{ij} = \dfrac{x_i x_j}{\sum_{k=1}^{n} x_k^2}$, $i,j = 1, \dots, n$ con $\sum_{k=1}^{n} x_k^2 \neq 0$. ¿Distribución de $W = \dfrac{Y^T H Y}{\sigma^2}$?

Veamos lo siguiente

1) H es simétrica porque

$$[H^T]_{ij} = h_{ij} = \frac{x_i x_j}{\sum_{k=1}^{n} x_k^2} = \frac{x_j x_i}{\sum_{k=1}^{n} x_k^2} = h_{ji} = [H_{ji}]^T \quad \forall\, i=1,\dots,n,\ j=1,\dots,n$$

2) H es idempotente porque

$[H^2]_{ij} = H_{(i)} H^{(j)}$, donde $H_{(i)}$ denota el renglón i-ésimo
$H^{(j)}$ denota la columna j-ésima

Entonces, veamos el producto, si $i \neq j$

$$H_i H^j = \left(\frac{x_i x_1}{\sum x_k^2}, \frac{x_i x_2}{\sum x_k^2}, \dots, \frac{x_i^2}{\sum x_k^2}, \dots, \frac{x_i x_n}{\sum x_k^2} \right) \begin{pmatrix} \frac{x_1 x_j}{\sum x_k^2} \\ \frac{x_2 x_j}{\sum x_k^2} \\ \vdots \\ \frac{x_j^2}{\sum x_k^2} \\ \vdots \\ \frac{x_n x_j}{\sum x_k^2} \end{pmatrix}$$

$$= \frac{1}{(\sum x_k^2)^2} \{ x_i x_1^2 x_j + x_i x_2^2 x_j + \dots + x_i^2 x_i x_j + \dots + x_i x_j x_j^2 + \dots + x_i x_n^2 x_j \}$$

$$= \frac{x_i x_j}{(\sum x_k^2)^2} \{ x_1^2 + x_2^2 + \dots + x_i^2 + \dots + x_j^2 + \dots + x_n^2 \}$$

$$= \frac{x_i x_j}{(\sum x_k^2)^{\cancel{2}}} (\cancel{\sum x_k^2}) = \frac{x_i x_j}{\sum x_k^2} = H_{ij}$$

Pero, si $i = j$ resulta que

$$H_i H^j = \left(\frac{x_i x_1}{\sum x_k^2}, \dots, \frac{x_i^2}{\sum x_k^2}, \dots, \frac{x_i x_n}{\sum x_k^2} \right) \cdot \begin{pmatrix} \frac{x_1 x_i}{\sum x_k^2} \\ \vdots \\ \frac{x_i^2}{\sum x_k^2} \\ \vdots \\ \frac{x_n x_i^2}{\sum x_k^2} \end{pmatrix}$$

(9)

$$= \frac{1}{(\sum x_k^2)^2}\{x_1^2x_i^2 + x_2^2x_i^2 + \dots + x_i^4 + \dots + x_n^2x_i^2\}$$

$$= \frac{x_i^2}{(\sum x_k^2)^2}\{x_1^2 + x_2^2 + \dots + x_i^2 + \dots + x_n^2\} = \frac{x_i^2}{(\sum x_k^2)^{\cancel{2}}}\cancel{\sum x_k^2}$$

$$= \frac{x_i^2}{\sum x_k^2} = H_{ii}$$

De este modo, tenemos que H es simétrica e idempotente

3) Así, su rango se da por

$$\text{rango}(H) = tr(H) = \sum_{i=1}^{n}\left(\frac{x_i^2}{\sum_{k=1}^{n} x_k^2}\right) = \frac{\cancel{\sum_{i=1}^{n} x_i^2}}{\cancel{\sum_{k=1}^{n} x_k^2}} = 1$$

4) Falta ver qué ocurre con $H\mu$, Así

$$[H]_{ij}\mu = \frac{x_i x_j}{\sum x_k^2}\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix} = \begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix} = \underline{0}$$

Así, $H\underline{\mu} = \underline{0}$

De este modo, por el teorema de clase, tenemos que

$$W = \frac{Y^T H Y}{\sigma^2} \sim \chi^2_{(1)}$$

5) $Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + u_i$

$\hat{Y}_i$ pronóstico de Y_i con $\hat{u}_i$ su residuo.

El modelo de forma matricial se escribe por $\underline{Y} = X\underline{\beta} + \underline{u}$

Consideramos al pronóstico $\hat{Y}_i$ que se da por

$\hat{Y}_i = \underline{x}_i' \hat{\underline{\beta}}$ pues se obtiene del enfoque $\mathbb{E}(Y_i) = \underline{x}_i' \underline{\beta}$

Entonces, el residuo $\hat{u}_i = Y_i - \hat{Y}_i$

a) $\hat{\underline{Y}} = \underline{X}\beta = X(X^TX)^{-1}X^T\underline{Y} = H\underline{Y}$

Así,

$u_i = \underline{Y} - \hat{\underline{Y}} = Y - HY = (I-H)Y = MY$

Notemos que $\sum u_i = n\frac{1}{n}\sum u_i = n\bar{u}$

$\Rightarrow u = MY = MX\beta$

Notemos que

$MX\beta = (I-H)X\beta = (X-HX)\beta = (X-X)\beta = 0$

$\therefore \mathbb{E}(u) = 0$

Así, $\mathbb{E}(u) = \frac{1}{n}\sum_{i=1}^{n} u_i = 0 \Leftrightarrow \sum_{i=1}^{n} u_i = 0$

b)

PD $\sum \hat{u}_j x_{ji} = 0$, $j = 1, \dots, k$

Tenemos que

$u x_j = MY x_j = (I-H)Y x_j = Y x_j - HY x_j$

$= Y x_j - Hx_j Y = Y x_j - x_j Y = Y - x_j - Y x_j = 0$

$x_j \in \mathbb{R}$

$\therefore u x_j = 0 \Rightarrow \sum \hat{u}_j x_{ji} = 0$

c) PD $\sum \hat{u}_i \hat{Y}_i$

Veamos que

$u HY = (Y-\hat{Y})HY = (Y-\hat{Y})HY = (Y-HY)HY = (I-H)YHY$

$= YHY - YH^2Y = YHY - YHY = 0$

H idempotente

De este modo