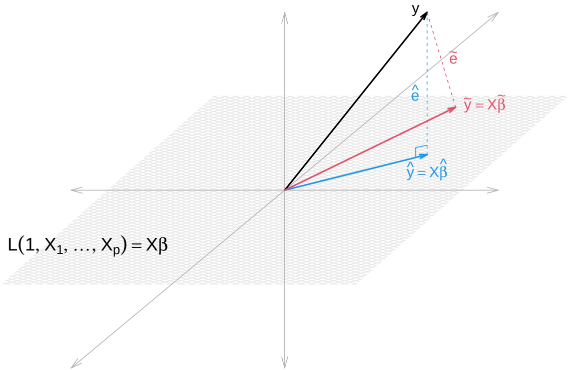
4b - H_0: $A\beta = c$



−ebz

Contenido

Resultados. Álgebra lineal y distribuciones[1]

Proposición

Sea $x_{n\times 1}$ un vector aleatorio con media $\mathbb{E}(x) = \mu_{n\times 1}$, y matriz de covarianzas $\text{var}(x) = \Sigma_{n\times n}$. Sean $A_{m\times n}$ y $B_{n\times n} > 0$ y $C_{s\times r}$ matrices de constantes y $y_{r\times 1}$ un vector aleatorio. Entonces,

1. $\mathbb{E}[Ax] = A\mathbb{E}[x] = A\mu$.
2. $\text{var}(Ax) = A\text{var}(x)A' = A\Sigma A'$.
3. $\text{cov}(Ax, Cy) = A\,\text{cov}(x, y)\, C'$.
4. $\mathbb{E}[x' Bx] = \text{tr}[B\Sigma] + \mu' B\mu$.

Proposición

Sea $x \sim N_n(0, \Sigma)$. Entonces, $x'\Sigma^{-1}x \sim \chi^2_n$.

Proposición

Sea H una matriz $n \times n$, simétrica e idempotente, entonces H es una matriz de proyección sobre $\mathcal{L}(H)$, subespacio generado por las columnas de H. Luego, $(I - H)$ es la matriz de proyección sobre el subespacio ortogonal a $\mathcal{L}(H)$.

[1] Schott (1997)

−ebz

Resultados (cont.)

Proposición

Sean $X_{n\times q}$ con rango$(X) = q$, y las matrices de proyección $H = X(X'X)^{-1}X'$ y $M = (I - H)$. Entonces, rango$(H) = q$ y rango$(M) = n - q$.

Proposición

Sea L una matriz definida positiva. Entonces, para todo vector b se tiene que

$$\sup_{h\neq 0}\left\{\frac{(h'b)^2}{h'Lh}\right\} = b'L^{-1}b$$

Modelo de Regresión Lineal Múltiple [2]

Considere el modelo de regresión lineal

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \epsilon_i, \quad i = 1, \ldots, n$$

Modelo de Regresión Lineal Múltiple

$$y = X\beta + \epsilon \tag{1}$$

con $y_{n\times 1}$, $\epsilon_{n\times 1}$, $\beta_{q\times 1}$, y $X_{n\times q}$, donde $q = 1 + p$.

La *suma de cuadrados* $\mathrm{SC}(\beta) = \sum_{i=1}^n (y_i - \beta' x_i)^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_p x_{ip})^2$ queda matricialmente

Suma Cuadrados

$$\mathrm{SC}(\beta) \quad = \quad ||y - X\beta||^2 = (y - X\beta)'(y - X\beta)$$

Si el criterio de estimación es *mínimos cuadrados*, $\frac{\partial \mathrm{SC}_{(}\beta)}{\partial \beta} = 2X'X\beta - 2X'y = 0$,

Ecuaciones Normales de Mínimos Cuadrados

$$X'X\beta = X'y \tag{2}$$

[2] Trabajo apoyado principalmente en Seber (1977) y Seber and Lee (2003).

Estimadores de Mínimos Cuadrados (EMC)

$$\hat{\beta} = (X'X)^{-1}X'y \tag{3}$$

Se define la *respuesta ajustada* como $\hat{y} = X\hat{\beta}$ y su correspondiente vector de *residuales* por $\hat{e} = y - \hat{y}$.
La *suma de cuadrados de los residuales* (SC_{Res}), también llamada *suma de cuadrados de los errores*: $\text{SC}(\hat{\beta}) = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$.

Suma de Cuadrado de los Residuales (SC_{Res})

$$\text{SC}_{\text{Res}} = ||y - X\hat{\beta}||^2 = y'y - \hat{\beta}'X'y$$

utilizando las ecuaciones normales (2). Note

$$\begin{aligned} \hat{y} &= X\hat{\beta} = X(X'X)^{-1}X'y = Hy \\ \hat{e} &= y - X(X'X)^{-1}X'y = (I - H)y = My \end{aligned} \tag{4}$$

donde H y M son matrices *simétricas e idempotentes*[3]. Luego, la suma de cuadrados de los errores se puede expresar también en términos de M como

$$\text{SC}_{\text{Res}} = ||\hat{e}||^2 = \hat{e}'\hat{e} = y'M^2y = \epsilon'M\epsilon$$

$y = X\beta + \epsilon$ y M es la matrix de proyección ortogonal al espacio generado por las columnas de X.

−ebz

[3]Para algunas propiedades de las matrices de proyección H y M refiérase a (Hoaglin and Welsch 1978).

Proposición

Considere el modelo de regresión lineal

$$y = X\beta + \epsilon$$

con vector de medias y matriz de covarianzas

$$\mathbb{E}[\epsilon] = 0 \quad \text{y} \quad \text{cov}(\epsilon) = \sigma^2 I_n$$

Entonces se cumplen los siguientes resultados:

a. $\mathbb{E}[y] = X\beta$. Es decir, el modelo *es correcto*.
b. $\text{cov}(y) = \sigma^2 I_n$.
c. $\text{cov}(\hat{e}) = \sigma^2 M$.
d. $\mathbb{E}[\hat{\beta}] = \beta$. El EMC es *insesgado*.
e. $\text{cov}(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \sigma^2 C$.

Proposición (cont.)

- f) Si $\operatorname{tr}(X'X)^{-1} \longrightarrow 0$, conforme $n \nearrow \infty$, entonces $\hat{\beta}$ es además un estimador consistente de β.
- g) $S(\hat{\beta}) = \sum_i m_{ii}\epsilon_i^2 + \sum_{i\neq j} m_{ij}\epsilon_i\epsilon_j$
- h) $\mathbb{E}[S(\hat{\beta})] = (n-q)\sigma^2$.
- i) $\mathbb{E}[s^2] = \sigma^2$.
- j) s^2 es un estimador insesgado y consistente de σ^2 [a].
- k) Un estimador insesgado y consistente de $\operatorname{cov}(\hat{\beta})$ es $s^2(X'X)^{-1} = s^2C$.
- l) Un estimador de la matriz de correlación del parámetro $\hat{\beta}$,

$$R = \widehat{\operatorname{corr}(\hat{\beta})} = \left[\frac{c_{ij}}{\sqrt{c_{ii}c_{jj}}}\right]$$

[a] Refiérase a (Sen and Srivastava 1990) Sec. 2A.2.

Teorema de Gauss-Markov[4]

Función estimable

La función paramétrica $a'\beta$ es una *función estimable* si tiene un estimador lineal insesgado, por ejemplo, $c'y$.

Teorema Gauss-Markov –caso escalar

Considere el modelo de regresión lineal $y = X\beta + \epsilon$ y el EMC del vector de coeficientes $\hat{\beta} = (X'X)^{-1}X'y$. Bajo la condiciones usuales de la regresión (7), el estimador de mínimos cuadrados $\hat{\theta} = \ell'\hat{\beta}$ de la función estimable $\theta = \ell'\beta$, tiene mínima varianza entre todos los estimadores lineales insesgados de θ.

Teorema Gauss-Markov –caso vectorial

Bajo la condiciones usuales de la regresión lineal, el estimador $L'\hat{\beta}$ de mínimos cuadrados de la función estimable $\Theta = L'\beta$ es de varianza mínima entre todos los estimadores lineales insesgados. Esto es, la matriz $[\text{cov}(Cy) - \text{cov}(L\hat{\beta})]$ es semidefinida positiva. Luego, es BLUE.

[4] Sen and Srivastava (1990) Sec. 2.9.

Distribuciones

Proposicion

Sea $y \sim \mathrm{N}_n(X\beta, \sigma^2 I)$, con $X_{n\times q}$, rango$(X) = q$, y rango$(M) = n - q$. Entonces,

1. $\hat{\beta} \sim N_q(\beta, \sigma^2(X'X)^{-1})$.
2. $||X(\hat{\beta} - \beta)||^2 = (\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta) \sim \sigma^2\chi^2_q$.
3. $\hat{\beta}$ es independiente de s^2.
4. SCE $= (n - q)s^2 \sim \sigma^2\chi^2_\nu$, con $\nu = n - q$.

–ebz

Mínimos Cuadrados con Restricciones

Considere nuevamente el modelo de regresión lineal múltiple (1)

$$y = X\beta + \epsilon$$

y suponga que se desea ajustar el modelo que satisface la hipótesis

$$H_0 : A\beta = c \tag{6}$$

donde $A_{r\times q}$ y rango$(A) = r(\leq q)$. Esto es, estimar el valor del parámetro β que satisface las restricciones $A\beta = c$.

Mínimos cuadrados con restricciones

El problema de mínimos cuadrados con restricciones puede plantearse entonces como

$$\begin{aligned} \text{Minimizar:} \quad & SC(\beta) = y'y - 2\beta'X'y + \beta'X'X\beta \\ \text{Sujeto a:} \quad & A\beta = c \end{aligned} \tag{7}$$

–ebz

Estimadores

Para resolver el problema de optimización con las restricciones (7), construimos el correspondiente lagrangiano

$$L(\beta, \lambda) = y'y - 2\beta'X'y + \beta'X'X\beta + \lambda'(A\beta - c)$$

Si se deriva el lagrangiano anterior y se iguala a 0,

$$\frac{\partial L}{\partial \beta} = 2(X'X)\beta - 2X'y - A'\lambda = 0; \qquad \frac{\partial L}{\partial \lambda} = c - A\beta = 0$$

se da lugar a

$$\begin{aligned} \hat{\beta}_H &= (X'X)^{-1}[X'y + \frac{1}{2}A'\lambda] = \hat{\beta} + \frac{1}{2}(X'X)^{-1}A'\hat{\lambda} \\ c &= A\hat{\beta}_H = A\hat{\beta} + \frac{1}{2}A(X'X)^{-1}A'\hat{\lambda} \end{aligned}$$

donde $\hat{\beta}$ es el estimador de mínimos cuadrados sin restricciones (3). Luego, eliminando λ en la expresión anterior, se tiene el *estimador de mínimos cuadrados con restricciones*,

Estimadores Mínimos Cuadrados con Restricciones

$$\hat{\beta}_H = \hat{\beta} + (X'X)^{-1}A'\left[A(X'X)^{-1}A'\right]^{-1}(c - A\hat{\beta}) \tag{8}$$

solución del problema de minimización de la suma de cuadrados de los residuales cuando los coeficientes satisfacen ciertas condiciones lineales (7).

Ejemplo

Considere la hipótesis o restricción $H_0 : \beta_1 = g_1$. Se sigue de (6) que en este caso $A = (1, 0, \ldots, 0)$, y $c = g_1$. Luego, se sigue de (8) que

$$\hat{\beta}_H = \hat{\beta} + C \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} C_{11}^{-1}(g_1 - \hat{\beta}_1) = \begin{bmatrix} g_1 \\ \hat{\beta}_2 + (g_1 - \hat{\beta}_1)C_{21}/C_{11} \\ \vdots \\ \hat{\beta}_p + (g_1 - \hat{\beta}_1)C_{p1}/C_{11} \end{bmatrix} = \begin{bmatrix} g_1 \\ \hat{\beta}_2 + r_{12}\frac{s_2}{s_1}\,\delta_1 \\ \vdots \\ \hat{\beta}_p + r_{1p}\frac{s_p}{s_1}\,\delta_1 \end{bmatrix}$$

donde $C = (X'X)^{-1} = (C_{ij})$, $\delta_1 = (g_1 - \beta_1)$, $s_i = \mathrm{ee}(\hat{\beta}_i)$ es el error estándar de $\hat{\beta}_i$, y $r_{ij} = \mathrm{corr}(\hat{\beta}_i, \hat{\beta}_j)$ es la correlación lineal entre $\hat{\beta}_i$ y $\hat{\beta}_j$, por lo que $C_{i1}/C_{11} = r_{1i}s_i/s_1$.

Distribuciones

Considere nuevamente el modelo de regresión lineal múltiple con el supuesto esférico sobre el error, $\epsilon \sim \mathrm{N}_n(0, \sigma^2 I)$. Se desea probar la hipótesis

$$H_0 : A\beta = c$$

Sea SCE $= S(\hat{\beta})$ la suma de cuadrados de los errores sin restricciones y $\mathrm{SCE}_H = S(\hat{\beta}_H)$, la suma de cuadrados de los errores del problema *con restricciones*

$$\mathrm{SCE}_H = ||y - X\hat{\beta}_H||^2 = (y - X\hat{\beta}_H)'(y - X\hat{\beta}_H)$$

Proposición

$$\begin{aligned}
\mathrm{SCE} - \mathrm{SCE}_H &= (A\hat{\beta} - c)' \left[A(X'X)^{-1}A'\right]^{-1} (A\hat{\beta} - c) \geq 0 \\
\mathbb{E}\left[\mathrm{SCE}_H - \mathrm{SCE}\right] &= r\sigma^2 + (A\beta - c)' \left[A(X'X)^{-1}A'\right]^{-1} (A\beta - c)
\end{aligned}$$

−ebz

Prueba de Hipótesis

Terorema

Bajo H$_0$

$$F = \frac{(\mathrm{SCE}_H - \mathrm{SCE})/r}{\mathrm{SCE}/(n-q)} = \frac{(c - A\hat{\beta})' \left[A(X'X)^{-1}A'\right]^{-1} (c - A\hat{\beta})}{rs^2} \sim F_{r,n-q}$$

Teorema

Cuando $c = 0$, bajo H_0

$$F = \frac{n-q}{r} \; \frac{y'(H - P_H)y}{y(I-H)y}$$

donde P_H es la matriz de proyección (matriz simétrica e idempotente) tal que $HP_H = P_H H = P_H$.

Notas

1. Por el teorema de Gauss-Markov, $A\hat{\beta}$ es el mejor estimador lineal insesgado de $A\beta$ estimable, que bajo H_0, sería cercano a c. Luego,

$$\mathrm{SCE}_H - \mathrm{SCE} = (A\hat{\beta} - c)' \left[A(X'X)^{-1}A' \right]^{-1} (A\hat{\beta} - c)$$

   es pequeño. En caso contrario, si $A\hat{\beta}$ no es cercano a c, la diferencia $\mathrm{SCE}_H - \mathrm{SCE}$ tendería a ser grande. Entonces, la "prueba F" tiene que ser de una cola. Esto es, la regla de decisión sería: *Rechazar H_0 si $F > F_{1-\alpha}$*, donde $F_{1-\alpha}$ es el $(1-\alpha)$-cuantil de la distribución F.

2. Si $r > 2$ resulta más práctico calcular SCE y SCE_H mediante la minimización de $\epsilon'\epsilon$, *con* y *sin* las restricciones (6).

3. Puesto que SCE_H es única, no importa como se obtiene. Se puede usar $A\beta = c$ para eliminar algunos de los β_j's y después minimizar $||\epsilon||^2$ con respecto a los β's restantes.

4. El procedimiento basado en los 2 incisos anteriores se conoce como *Suma de Cuadrados Extra*, que es la diferencia entre las suma de cuadrados que no alcanza a explicar el modelo reducido, que satisface H_0, comparado con el modelo completo, sin restricciones.

Cociente de verosimilitudes

Considere la función de verosimilitud asociada con el modelo de regresión lineal *sin* restricciones

$$L(\beta, \sigma^2 | X, y) = (2\pi\sigma^2)^{-n/2} \exp\left\{ -\frac{1}{2\sigma^2} (y - X\beta)'(y - X\beta) \right\}$$

La maximización de L respecto a β y σ^2 da lugar a los *estimadores de máxima verosimilitud*, dados por

$$\hat{\beta} = (X'X)^{-1} X'y \qquad \text{y} \qquad \hat{\sigma}^2 = \frac{\text{SCE}}{n}$$

Entonces el máximo de L está dado por $L(\hat{\beta}, \hat{\sigma}^2) = (2\pi\hat{\sigma}^2)^{-n/2} e^{-n/2}$.

De manera similar, si se satisface H_0, los correspondientes estimadores de máxima verosimilitud maximizan $L(\beta, \sigma^2 | X, y)$, sujeta a las restricciones impuestas por H_0. Luego

$$\hat{\beta}_H = \hat{\beta} + (X'X)^{-1} A' \left[A(X'X)^{-1} \right]^{-1} (c - A\hat{\beta}) \qquad \text{y} \qquad \hat{\sigma}^2 = \frac{\text{SCE}_H}{n}$$

En este caso, bajo H_0, $L(\hat{\beta}_H, \hat{\sigma}^2_H | X, y) = (2\pi\hat{\sigma}^2_H)^{-n/2} e^{-n/2}$.

Cociente de verosimilitudes (cont.)

Luego, la *razón de verosimilitud* está dada por

$$\Lambda = \frac{L(\hat{\beta}_H, \hat{\sigma}_H^2)}{L(\hat{\beta}, \hat{\sigma}^2)} = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_H^2}\right)^{n/2}$$

Ahora bien, el *principio de la razón de verosimilitud* indica rechazar H_0 para valores pequeños de Λ, que en este caso es una función monótona decreciente del estadístico F. A saber,

$$F = \frac{n-p}{q}\left(\Lambda^{-2/n} - 1\right)$$

Luego, se rechaza $H_0 : A\beta = c$, para valores grandes del estadístico de prueba F.

Considere el modelo de regresión lineal múltiple

$$y = X\beta + \epsilon$$

Hipótesis sobre la significancia de la regresión

$$H_0 : \beta = 0, \quad vs. \quad H_a : \text{Al menos un } \beta_j \neq 0,$$

En el contexto de la prueba general de hipótesis, si $H_0 : A\beta = c$, entonces, $A = I_p$, la identidad de orden p, y $c = 0$. Luego, $A\hat{\beta} - c = \hat{\beta}$. Luego, el estadístico F toma la forma

$$F = \frac{(X\hat{\beta})'(X\hat{\beta})/p}{\hat{e}'\hat{e}/(n-q)} = \frac{\hat{\beta}'X'y/p}{s^2} = \frac{\text{CM}_{\text{Reg}}}{\text{CM}_{\text{Res}}} \sim F_{(p,n-q)} \tag{9}$$

que es como se calcula el estadístico F a partir de la tabla de ANOVA. El estadístico F de la expresión anterior (9) es el que reportan los programas estadísticos para verificar la significancia de la regresión.

–ebz

Considere nuevamente el modelo

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + \epsilon$$

conocido en este contexto como *modelo completo* (MC) escrito matricialmente

Modelo completo (MC)

$$y = X\beta + \epsilon = X_1 B_1 + X_2 B_2 + \epsilon$$

donde, $B_{1\ q_1 \times 1}$, $B_{2\ q_2 \times 1}$ y X_i las matrices $n \times q_i$ correspondientes para $i = 1, 2$.
Ahora bien, se considera la posibilidad de simplificar el modelo al *modelo restringido* (MR)

Modelo restringido (MR)

$$y = X_1 B_1 + \epsilon$$

Para esto, se contrastan las hipótesis

$$H_0 : \; B_2 = 0, \quad vs. \quad H_a : B_2 \neq 0$$

Estadístico de prueba para suma de cuadrados extra

$$F = \frac{(\mathrm{SC}_{\mathrm{Res}}(\mathrm{MR}) - \mathrm{SC}_{\mathrm{Res}}(\mathrm{MC}))/q_2}{\mathrm{SC}_{\mathrm{Res}}(\mathrm{MC})/n - q} \sim F_{q_2, n-q}$$

y se determina su significancia.

Región de confianza para el vector de parámetros β

Recuerde que

$$(c - A\hat{\beta})' \left[A(X'X)^{-1}A'\right]^{-1} (c - A\hat{\beta}) \sim rs^2 F_{r,\nu}$$

donde rango$(A) = r$, y $\nu = n - q$. Si se toma $A = I_q$ y $c = \beta$,

$$\mathbb{P}\left\{(\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta) \leq rs^2 F(1 - \alpha; r, \nu)\right\} = 1 - \alpha$$

Región de confianza para β

$$\mathcal{R}_{1-\alpha} = \left\{\beta : (\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta) \leq rs^2 F(1 - \alpha; r, \nu)\right\}$$

constituye una región del $100(1 - \alpha)\,\%$ confianza para el vector de parámetros β.

Método de Scheffé

Considere la restricción $A\beta = c$. Entonces, por el teorema Gauss-Markov $A\hat{\beta} = \hat{c}$ es un buen estimador de c y como se vio

$$\begin{aligned} 1-\alpha &= \mathbb{P}\left\{(A\hat{\beta}-c)'\left[A(X'X)^{-1}A'\right]^{-1}(A\hat{\beta}-c) \leq rs^2 F_{r,\nu}\right\} \\ &= \mathbb{P}\left\{(\hat{c}-c)'L^{-1}(\hat{c}-c) \leq \xi\right\} \end{aligned}$$

donde $\hat{c} = A\hat{\beta}$, $L = \left[A(X'X)^{-1}A'\right]$ y $\xi = rs^2 F(1-\alpha; r, \nu)$. Luego, tomando $\gamma = \hat{c} - c$ se puede mostrar que

$$\begin{aligned} 1-\alpha &= \mathbb{P}\left\{\gamma' L^{-1}\gamma \leq \xi\right\} \\ &= \mathbb{P}\left\{\sup_{h\neq 0}\left\{\frac{(h'\gamma)^2}{h'Lh}\right\} \leq \xi\right\} \\ &= \mathbb{P}\left\{\frac{(h'\gamma)^2}{h'Lh} \leq \xi, \ \text{ para todo } h \neq 0\right\} \\ &= \mathbb{P}\left\{\frac{(|h'(\hat{c}-c)|}{s\sqrt{h'Lh}} \leq \sqrt{rF(1-\alpha; r, \nu)}\right\} \end{aligned}$$

−ebz

Método de Scheffé (cont.)

Por lo que podemos construir un intervalo del $100(1-\alpha)\,\%$ de confianza para *cualquier* combinación lineal $h'c = h'A\beta$. A saber,

Intervalo de confianza

$$h'\hat{c} \pm (rF(1-\alpha; r, \nu))^{1/2}\, s\sqrt{h'Lh} \tag{10}$$

y la probabilidad sobre *toda* la familia $h'c$ es precisamente $1-\alpha$.

Notas

1. El denominador $s^2 h'Lh$ es un estimador insesgado de la varianza de $h'\hat{c}$. En efecto,

$$\begin{aligned} \mathrm{var}(h'\hat{c}) &= h'\mathrm{var}(\hat{c})h = h'\mathrm{var}(A\hat{\beta})h \\ &= \sigma^2 h'\left[A(X'X)^{-1}A'\right]h \\ &= \sigma^2 h'Lh \end{aligned}$$

2. Puesto que $h'c = c_i$ para algún h, entonces el intervalo de confianza para cualquier $a_i'\beta = c_i$ $(i = 1, \ldots, r)$ está incluido en la expresión general (10) anterior.

Estimación de respuesta media y nuevas observaciones

Intervalos de confianza para la respuesta media a nivel $x = x_0$

$$\hat{y}(x_0) \pm t_{(1-\alpha/2,\nu)} s \sqrt{x_0'(X'X)^{-1}x_0}$$

Intervalos de confianza Bonferroni a nivel $\dot{x}_1, \ldots, \dot{x}_m$

$$\hat{y}(\dot{x}_i) \pm t_{(1-\alpha/2m,\nu)} s \sqrt{\dot{x}_i'(X'X)^{-1}\dot{x}_i}$$

Intervalos de confianza Scheffé a nivel $\dot{x}_i$

$$\hat{y}(\dot{x}_i) \pm (mF(1-\alpha; m, \nu))^{1/2}\, s \sqrt{\dot{x}_i'(X'X)^{-1}\dot{x}_i}$$

Similarmente pueden construirse regiones simultáneas para predicciones de observaciones nuevas. Utilizando por ejemplo el método de Scheffé, una región de predicción conjunta para niveles de los regresores $\dot{x}_1, \ldots, \dot{x}_\ell$, está dado por

Intervalos de confianza conjunto a niveles $\dot{x}_1, \ldots, \dot{x}_\ell$

$$\hat{y}(\dot{x}_i) \pm (mF(1-\alpha; \ell, \nu))^{1/2}\, s \sqrt{1 + \dot{x}_i'(X'X)^{-1}\dot{x}_i}$$

–ebz

Referencias

Hoaglin, D. C. and R. E. Welsch (1978).
The Hat Matrix in Regression and ANOVA.
The American Statistician 32(1), 17–22.

Schott, J. R. (1997).
Matrix Analysis for Statistics.
New York: Wiley.

Seber, G. A. F. (1977).
Linear Regression Analysis.
New York: Wiley.

Seber, G. A. F. and A. J. Lee (2003).
Linear Regression Analysis (2 ed.).
New York: Wiley.

Sen, A. and M. Srivastava (1990).
Regression Analysis.
New York: Springer-Verlag.

−ebz