Regresión

HIPÓTESIS LINEAL GENERAL $T\beta = 0$

Supongamos que estamos interesados en probar la hipótesis de interés de la forma:

$$H_0: T\beta = 0$$

donde:

T_{mxp} es una matriz de constantes, tal que solamente r de las m ecuaciones de $T\beta = 0$ son independientes.

El modelo completo es $y = X\beta + \varepsilon$,

con $\hat{\beta} = (X'X)^{-1}X'Y$,

y la suma de cuadrados de residuales, para este modelo es

$$SCR_{es}(MC) = Y'Y - \hat{\beta}'X'Y \quad \text{con} \quad (n-p \text{ g.l.})$$

Para obtener el modelo reducido, se usan las r ecuaciones independientes en $T\beta = 0$, para solucionar los "r" coeficientes de regresión en el modelo completo, en términos de los $p-r$ coeficientes de regresión restantes.

Esto conduce al modelo reducido $y = Z\gamma + \varepsilon$,
donde

Z es una matriz $n\times(p-r)$ y
γ es un vector $(p-r)\times 1$, de coeficientes de regresión desconocidos.

Podemos ahora estimar el vector de parámetros γ en el nuevo modelo por

$$\hat{\gamma} = (Z'Z)^{-1} Z'Y$$

Si $Z'Z$ es no singular, podemos obtener una nueva suma de cuadrados de cuadrados de los residuales para este modelo reducido definido como:

$$SCR_{es}(MR) = Y'Y - \hat{\gamma}'Z'Y \quad (n-p+r \text{ g.l.})$$

El modelo reducido contiene menos parámetros que el modelo completo, entonces

$$SCR_{es}(MR) \geq SCR_{es}(MC).$$

Para probar la hipótesis $H_0 : T\beta = 0$ se emplea la diferencia de sumas de cuadrados de residuales

$$SCH = SCR_{es}(MR) - SCR_{es}(MC) \quad \text{con } n-p+r-(n-p) = r \text{ g.l.}$$

Aquí, esta diferencia se llama ***suma de cuadrados debida a la hipótesis H_0: $T\beta = 0$.***

El estadístico de prueba para esta hipótesis es

$$F_0 = (SC_H/r) \,/\, (SCR_{es}(MC)/(n-p)) \quad \sim \quad F_{(\alpha;\, r;\, n-p)} \qquad {**}$$

Se rechaza $H_0 : T\beta = 0$ si $F_0 > F_{(\alpha;\, r;\, n-p)}$

La prueba estadística ** puede también ser escrita como:

$$F_0 = \hat{\beta}' T' (T(X'X)^{-1}T')^{-1} T\hat{\beta} / r \,/\, SCR_{es}(MC)/(n-p)$$

Es muy útil hacer una ***extensión para la hipótesis lineal general*** y la presentamos a continuación:

H_0: $T\beta = c$
H_1: $T\beta = c$

$$F_0 = (T\hat{\beta} - c)' \left(T(X'X)^{-1}T'\right)^{-1} (T\hat{\beta} - c) / r \,/\, SCR_{es}(MC)/(n-p)$$

Con distribución:

$$\boldsymbol{F_0} = \frac{(T\hat{\beta}-c)'\left[T(X'X)^{-1}T'\right]^{-1}(T\hat{\beta}-c)/r}{SC\,Re\,s/(n-p)} \sim \frac{\chi_r^2/r}{\chi_{n-p}^2/(n-p)} \sim \boldsymbol{F_{r,n-p}}$$

Se rechaza H_0: $T\beta = c$ si $F_0 > F_{(\alpha;\, r,\, n-p)}$

Ejemplo:

Supongamos que el modelo es: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$
y que se desea probar

$$H_0: \beta_1 = \beta_3, \quad \beta_2 = 0$$