MAT-466: Test de hipótesis en GLM

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Test de hipótesis

Considere la hipótesis simple,

$$H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0 \quad \text{versus} \quad H_1 : \boldsymbol{\beta} \neq \boldsymbol{\beta}_0,$$

con $\boldsymbol{\beta}_0$ un vector conocido.[1]

- Test de razón de verosimilitudes (LRT): Para el caso de hipótesis simples, el estadístico de razón de verosimilitudes es dado por:

  $$LR = 2\{\ell(\widehat{\boldsymbol{\beta}}) - \ell(\boldsymbol{\beta}_0)\} \xrightarrow{\text{D}} \chi^2(p),$$

  para el caso de GLMs, el estadístico LRT adopta la forma:

  $$LR = \phi\{D(\boldsymbol{y}; \boldsymbol{\mu}_0) - D(\boldsymbol{y}; \widehat{\boldsymbol{\mu}})\},$$

  donde $\boldsymbol{\mu}_0 = \boldsymbol{g}^{-1}(\boldsymbol{\eta}_0)$, $\boldsymbol{\eta}_0 = \boldsymbol{X}\boldsymbol{\beta}_0$, y análogamente para $\widehat{\boldsymbol{\mu}} = \boldsymbol{g}^{-1}(\widehat{\boldsymbol{\eta}})$, con $\widehat{\boldsymbol{\eta}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}$.

[1] asumiremos ϕ conocido.

Test de hipótesis

- Test de Wald: El estadístico de Wald es definido como

$$W = (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top \{\mathsf{Cov}(\widehat{\boldsymbol{\beta}})\}^{-1} (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0),$$

donde $\mathsf{Cov}(\widehat{\boldsymbol{\beta}})$ corresponde a la matriz de covarianza de $\widehat{\boldsymbol{\beta}}$.[2]

Para GLMs, sabemos que

$$\mathsf{Cov}(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\mathcal{F}}(\boldsymbol{\beta})\big|_{\beta=\widehat{\beta}} = \phi(\boldsymbol{X}^\top \widehat{\boldsymbol{W}} \boldsymbol{X})^{-1},$$

donde $\widehat{\boldsymbol{W}} = \operatorname{diag}(\widehat{\omega}_1, \dots, \widehat{\omega}_n)$ con $\widehat{\omega}_i = \omega_i(\widehat{\mu}_i)$. De este modo,

$$W = \phi(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top (\boldsymbol{X}^\top \widehat{\boldsymbol{W}} \boldsymbol{X})(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \xrightarrow{\mathsf{D}} \chi^2(p).$$

[2] que debe ser estimada bajo H_1.

Observación:

Suponga $p = 1$, entonces

$$W = \frac{(\widehat{\beta} - \beta_0)^2}{\mathsf{var}(\widehat{\beta})} \xrightarrow{\mathsf{D}} \chi^2(1),$$

que es equivalente al estadístico

$$T = \frac{\widehat{\beta} - \beta_0}{\sqrt{\mathsf{var}(\widehat{\beta})}} \xrightarrow{\mathsf{D}} \mathsf{N}(0, 1).$$

Es conocido que, cuando $\boldsymbol{\eta}(\boldsymbol{\beta})$ es no lineal en $\boldsymbol{\beta}$, el estadístico W no es invariante a la parametrización, y por tanto, formas equivalentes para $\boldsymbol{\eta}(\boldsymbol{\beta})$, puede llevar a valores diferentes de W.

- Test score: El estadístico score (de Rao) es definido por

$$R = \boldsymbol{U}(\boldsymbol{\beta}_0)^\top \operatorname{Cov}(\boldsymbol{\beta}_0)\boldsymbol{U}(\boldsymbol{\beta}_0),$$

donde $\operatorname{Cov}(\boldsymbol{\beta}_0)$ corresponde a la matriz de covarianza de $\widehat{\boldsymbol{\beta}}$ que debe ser estimada bajo H_0.

Para GLMs, obtenemos

$$R = \phi^{-1}\boldsymbol{U}(\boldsymbol{\beta}_0)^\top(\boldsymbol{X}^\top\boldsymbol{W}_0\boldsymbol{X})^{-1}\boldsymbol{U}(\boldsymbol{\beta}_0) \xrightarrow{\mathsf{D}} \chi^2(p).$$

donde $\boldsymbol{W} = \operatorname{diag}(\omega_1, \dots, \omega_n)$ es evaluada en $\boldsymbol{\beta} = \boldsymbol{\beta}_0$. De este modo, también podemos escribir

$$R = \phi(\boldsymbol{Y} - \boldsymbol{\mu}_0)^\top \boldsymbol{V}_0^{-1/2}\boldsymbol{H}_0\boldsymbol{V}_0^{-1/2}(\boldsymbol{Y} - \boldsymbol{\mu}_0),$$

con $\boldsymbol{H} = \boldsymbol{W}^{1/2}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{W}^{1/2}$.

- Test F: Para el caso de hipótesis simples el estadístico F adopta la forma:

$$F = \frac{\{D(\boldsymbol{y};\boldsymbol{\mu}_0) - D(\boldsymbol{y};\widehat{\boldsymbol{\mu}})\}/p}{D(\boldsymbol{y};\widehat{\boldsymbol{\mu}})/(n-p)}$$

que, bajo H_0 y cuando $\phi \to \infty$, tiene una distribución $F(p, n-p)$.

Observaciones:

- Tenemos también que

$$F \xrightarrow{\mathsf{D}} F(p, n-p),$$

cuando usamos un estimador consistente de ϕ^{-1}.

- El estadístico F también es invariante por reparametrizaciones.
- Finalmente, el estadístico F no depende del parámetro de dispersión ϕ^{-1}.

EX UMBRA IN SOLEM

Ejemplo (caso normal):

Para el caso normal, tenemos

$$\boldsymbol{\mathcal{F}}(\boldsymbol{\beta}) = \frac{1}{\sigma^2}\boldsymbol{X}^\top\boldsymbol{X}, \qquad \boldsymbol{U}(\boldsymbol{\beta}) = \frac{1}{\sigma^2}\boldsymbol{X}^\top(\boldsymbol{Y} - \boldsymbol{\mu}).$$

De este modo:

$$\begin{aligned}
LR &= \frac{1}{\sigma^2}\Big\{\sum_{i=1}^n (y_i - \mu_{0i})^2 - \sum_{i=1}^n (y_i - \widehat{\mu}_i)^2\Big\}, \\
W &= \frac{1}{\sigma^2}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top\boldsymbol{X}^\top\boldsymbol{X}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0), \\
R &= \frac{1}{\sigma^2}(\boldsymbol{Y} - \boldsymbol{\mu}_0)^\top\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top(\boldsymbol{Y} - \boldsymbol{\mu}_0), \\
F &= \frac{1}{qs^2}\Big\{\sum_{i=1}^n (y_i - \mu_{0i})^2 - \sum_{i=1}^n (y_i - \widehat{\mu}_i)^2\Big\}.
\end{aligned}$$

Una región de confianza asintótica de tamaño $1-\alpha$ para $\boldsymbol{\beta}$ basada en el estadístico de Wald, es dada por:

$$RC_W(\boldsymbol{\beta}) = \{\boldsymbol{\beta} : (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})^\top \boldsymbol{X}^\top \widehat{\boldsymbol{W}} \boldsymbol{X} (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \leq \phi^{-1} \chi^2_{1-\alpha}(p)\}.$$

donde $\chi^2_{1-\alpha}(p)$ denota el valor cuantil $1-\alpha$ de la distribución chi-cuadrado con p grados de libertad.

La región de confianza basada en el estadístico de Wald, puede depender de la parametrización.[3] En cuyo caso, puede ser recomendable usar estadísticas invariantes. Por ejemplo,

$$RC_{LR}(\boldsymbol{\beta}) = \{\boldsymbol{\beta} : 2[\ell(\widehat{\boldsymbol{\beta}}) - \ell(\widehat{\boldsymbol{\beta}})] \leq \chi^2_{1-\alpha}(p)\}.$$

[3] Cuando $\boldsymbol{\eta}$ es no lineal.

Considere hipótesis lineales del tipo

$$H_0 : \boldsymbol{C\beta} = \mathbf{0} \qquad \text{versus} \qquad H_1 : \boldsymbol{C\beta} \neq \mathbf{0},$$

donde $\boldsymbol{C}$ es matriz de contrastes de orden $k \times p$ con $\operatorname{rk}(\boldsymbol{C}) = k$.

Evidentemente $\widehat{\boldsymbol{\beta}}$ coincide con el estimador bajo $H_1 : \boldsymbol{C\beta} \neq \mathbf{0}$. Mientras que la estimación bajo H_0 requiere, en general, de un procedimiento iterativo

$$\widetilde{\boldsymbol{\beta}}^{(r+1)} = \widetilde{\boldsymbol{\beta}}^{(r)} - (\boldsymbol{X}^\top \boldsymbol{W}^{(r)} \boldsymbol{X})^{-1} \boldsymbol{C}^\top \{\boldsymbol{C}(\boldsymbol{X}^\top \boldsymbol{W}^{(r)} \boldsymbol{X})^{-1} \boldsymbol{C}^\top\}^{-1} \boldsymbol{C} \widetilde{\boldsymbol{\beta}}^{(r)},$$

para $r = 0, 1, \dots$ con matriz de covarianza

$$\mathsf{Cov}(\widetilde{\boldsymbol{\beta}}) = \phi^{-1} (\boldsymbol{X}^\top \boldsymbol{W X})^{-1} [\boldsymbol{I} - \boldsymbol{C}^\top \{\boldsymbol{C}(\boldsymbol{X}^\top \boldsymbol{W X})^{-1} \boldsymbol{C}^\top\}^{-1} \boldsymbol{C}]^{-1} (\boldsymbol{X}^\top \boldsymbol{W X})^{-1}.$$

Hipótesis lineales

Para hipótesis lineales del tipo $H_0 : \boldsymbol{C\beta} = \boldsymbol{0}$, tenemos

$$LR = \phi\{D(\boldsymbol{y}; \widetilde{\boldsymbol{\mu}}) - D(\boldsymbol{y}; \widehat{\boldsymbol{\mu}})\},$$

donde $\widetilde{\boldsymbol{\mu}}$ denota el MLE de $\boldsymbol{\mu}$ bajo $H_0 : \boldsymbol{C\beta} = \boldsymbol{0}$. Mientras que

$$R = \phi^{-1}\boldsymbol{U}^\top(\widetilde{\boldsymbol{\beta}})(\boldsymbol{X}^\top\widetilde{\boldsymbol{W}}\boldsymbol{X})^{-1}\boldsymbol{U}(\widetilde{\boldsymbol{\beta}}),$$

con $\widetilde{\boldsymbol{W}}$ es evaluada en $\widetilde{\boldsymbol{\beta}}$. Por otro lado,

$$\begin{aligned} W &= (\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{0})^\top\{\mathsf{Cov}(\boldsymbol{C}\widehat{\boldsymbol{\beta}})\}^{-1}(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{0}) \\ &= \phi(\boldsymbol{C}\widehat{\boldsymbol{\beta}})^\top\{\boldsymbol{C}(\boldsymbol{X}^\top\widehat{\boldsymbol{W}}\boldsymbol{X})^{-1}\boldsymbol{C}^\top\}^{-1}\boldsymbol{C}\widehat{\boldsymbol{\beta}}. \end{aligned}$$

Además,

$$LR \xrightarrow{\mathsf{D}} \chi^2(k), \qquad W \xrightarrow{\mathsf{D}} \chi^2(k), \qquad R \xrightarrow{\mathsf{D}} \chi^2(k).$$

Hipótesis lineales

El estadístico F para hipótesis lineales de la forma $H_0 : \boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{0}$ adopta la forma:

$$F = \frac{\{D(\boldsymbol{y}; \widetilde{\boldsymbol{\mu}}) - D(\boldsymbol{y}; \widehat{\boldsymbol{\mu}})\}/k}{D(\boldsymbol{y}; \widehat{\boldsymbol{\mu}})/(n-p)},$$

que tiene una distribución $F(k, n-p)$ cuando $\phi \to \infty$. Es posible mostrar que

$$D(\boldsymbol{y}; \widetilde{\boldsymbol{\mu}}) - D(\boldsymbol{y}; \widehat{\boldsymbol{\mu}}) = (\boldsymbol{C}\widehat{\boldsymbol{\beta}})^\top \{\boldsymbol{C}(\boldsymbol{X}^\top \widehat{\boldsymbol{W}} \boldsymbol{X})^{-1} \boldsymbol{C}^\top\}^{-1} \boldsymbol{C}\widehat{\boldsymbol{\beta}}.$$

Observación:

- Suponga la hipótesis

$$H_0 : \boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{0}, \qquad \text{versus} \qquad H_1 : \boldsymbol{C}\boldsymbol{\beta} \geq \boldsymbol{0},$$

  en ese caso los estadísticos LR, W, R tienen una distribución de mezcla de variables chi-cuadrado.

- Suponga el modelo:

$$\boldsymbol{\eta} = \boldsymbol{X}\boldsymbol{\beta} + \gamma \boldsymbol{d}_i,$$

  con $\boldsymbol{d}_i = (\boldsymbol{0}, 1, \boldsymbol{0})^\top$. La hipótesis

$$H_0 : \gamma = 0, \qquad \text{versus} \qquad H_1 : \gamma \neq 0,$$

  ofrece una interesante interpretación para el estadístico t.