Modelos Lineales Generalizados: aplicaciones en R

Modelos Lineales Generalizados: un enfoque aplicado

Ana M. Bianco Jemina García
(anambianco@gmail.com) (jeminagarcia@gmail.com)

4. Regresión Logística: Estimación e Inferencia

Inferencia: Intervalos de Confianza y Tests de Hipótesis

Fahrmeir y Kaufmann (1985) estudiaron el comportamiento de los estimadores de máxima verosimilitud para GLM bajo condiciones de regularidad.

Para n sufcientemente grande, una aproximación razonable es

$$(\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) \overset{(a)}{\sim} N(\mathbf{0}, \mathbf{V}(\widehat{\boldsymbol{\beta}}_n)),$$

siendo

$$\mathbf{V}(\widehat{\boldsymbol{\beta}}_n) = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}.$$

y estimaremos por

$$\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}}_n) = (\mathbf{X}'\mathbf{W}(\widehat{\boldsymbol{\beta}}_n)\mathbf{X})^{-1}.$$

Inferencia sobre una función de los coeficientes

Para una función lineal de los parámetros $\Psi = \mathbf{a}'\boldsymbol{\beta}$, una aproximación razonable para n suficientemente grande es

$$(\mathbf{a}^t\widehat{\boldsymbol{\beta}} - \mathbf{a}^t\boldsymbol{\beta}) \overset{(a)}{\sim} N(\mathbf{0}, \mathbf{a}^t\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}})\mathbf{a})\,.$$

Por lo tanto, para n grande tendremos que

$$\mathbf{a}^t\widehat{\boldsymbol{\beta}} \pm z_{\alpha/2}\sqrt{\mathbf{a}^t\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}})\mathbf{a}}$$

es un intervalo de nivel aproximado $1-\alpha$ para $\mathbf{a}^t\boldsymbol{\beta}$.

Inferencia sobre una función de los coeficientes

Así por ejemplo, si $\mathbf{a} = \mathbf{e}_j$, $j-$ésimo vector de la base canónica, tenemos que $\mathbf{a}^t\boldsymbol{\beta} = \mathbf{e}_j^t\boldsymbol{\beta} = \beta_j$ y por lo tanto, para n grande tendremos que

$$\widehat{\beta}_j \pm z_{\alpha/2}\sqrt{\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}_j})}$$

es un intervalo de nivel aproximado $1-\alpha$ para β_j.

Ahora, si $\mathbf{a} = \mathbf{x}_o$ es un nuevo punto,

- ¿Cómo se puede estimar la probabilidad $p(\mathbf{x}_o, \boldsymbol{\beta})$?

- Hallar un intervalo de confianza de nivel aproximado $1 - \alpha$ para $p(\mathbf{x}_o, \boldsymbol{\beta})$.

Ahora, si $\mathbf{a} = \mathbf{x}_o$ es un nuevo punto,

- ¿Cómo se puede estimar la probabilidad $p(\mathbf{x}_o, \boldsymbol{\beta})$?

- Hallar un intervalo de confianza de nivel aproximado $1-\alpha$ para $p(\mathbf{x}_o, \boldsymbol{\beta})$.

La probabilidad deseada puede estimarse por $\dfrac{1}{1+e^{-\mathbf{x}_o^t\widehat{\boldsymbol{\beta}}}}$

Opción 1: Para n grande tenemos

$$\left(\mathbf{x}_o^t\widehat{\boldsymbol{\beta}} - z_{\alpha/2}\sqrt{\mathbf{x}_o^t\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}})\mathbf{x}_o}, \mathbf{x}_o^t\widehat{\boldsymbol{\beta}} + z_{\alpha/2}\sqrt{\mathbf{x}_o^t\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}})\mathbf{x}_o}\right)$$

es un intervalo de nivel aproximado $1-\alpha$ para $\mathbf{x}_o^t\boldsymbol{\beta}$. Como $h(t) = \dfrac{1}{1+e^{-t}}$ es estrictamente creciente, entonces

$$\left(\frac{1}{1+e^{-\left(\mathbf{x}_o^t\widehat{\boldsymbol{\beta}} - z_{\alpha/2}\sqrt{\mathbf{x}_o^t\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}})\mathbf{x}_o}\right)}}, \frac{1}{1+e^{-\left(\mathbf{x}_o^t\widehat{\boldsymbol{\beta}} + z_{\alpha/2}\sqrt{\mathbf{x}_o^t\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}})\mathbf{x}_o}\right)}}\right)$$

es un intervalo de nivel aproximado $1-\alpha$ para la probabilidad buscada.

Opción 2: Método Delta

Otra posibilidad es la de aproximar la distribución de nuestro estimador $\dfrac{1}{1+e^{-\mathbf{x}_o^t \widehat{\boldsymbol{\beta}}}}$.

Este método se basa en aproximar en forma lineal a la función de interés y luego computar la varianza de esta expresión sencilla lineal.

Por lo que, bajo condiciones de regularidad, tendríamos que si

$$(\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) \overset{(a)}{\sim} N(\mathbf{0}, \mathbf{V}(\widehat{\boldsymbol{\beta}}_n)),$$

entonces

$$(g(\widehat{\boldsymbol{\beta}}_n) - g(\boldsymbol{\beta})) \overset{(a)}{\sim} N\left(\mathbf{0}, \frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}^t \mathbf{V}(\widehat{\boldsymbol{\beta}}_n) \frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right) \qquad (1)$$

Opción 2: Método Delta

Ejemplificamos con la situación del modelo más sencillo en el que tenemos una sola variable y sin intercept: el modelo es

$$p(x, \beta) = \frac{1}{1 + e^{-x\beta}}$$

Sea

$$g(b) = h(xb) = \frac{1}{1 + e^{-xb}}$$

Opción 2: Método Delta

En nuestro caso concreto $g(b) = h(xb) = \dfrac{1}{1+e^{-xb}}$, luego:

$$\frac{\partial g(\beta)}{\partial \beta} = \frac{\partial h(x\beta)}{\partial \beta} = h(x\beta)(1 - h(x\beta))x$$

Entonces, si nos interesa un intervalo de confianza de nivel aproximado $1-\alpha$ para la probabilidad en x_o, esto nos lleva al intervalo **simétrico**:

$$p(\mathbf{x}_o, \widehat{\beta}) \pm z_{\alpha/2} \sqrt{\left[x_o\, p(x_o, \widehat{\beta})(1 - p(x_o, \widehat{\beta}))\right]^2 \widehat{V}(\widehat{\beta})}$$

A diferencia del la opción 1, estos intervalos podrían incluir valores fuera del intervalo $[0, 1]$.

Opción 2: Método Delta

Usamos una aproximación mediante un desarrollo de Taylor de primer orden:

$$h(x\widehat{\beta}) \approx h(x\widehat{\beta}) + (\widehat{\beta} - \beta)\frac{\partial h(x\beta)}{\partial \beta}$$

donde $\frac{\partial h(x\beta)}{\partial \beta}$ es la derivada de $h(xb)$ respecto de b y evaluada en β.

$$V(h(x\widehat{\beta})) \approx V\left(h(x\beta) + (\widehat{\beta} - \beta)\frac{\partial h(x\beta)}{\partial \beta}\right)$$

$$\begin{aligned}
V\left(h(x\beta) + (\widehat{\beta} - \beta)\frac{\partial h(x\beta)}{\partial \beta}\right) &= V(h(x\beta)) + V\left((\widehat{\beta} - \beta)\frac{\partial h(x\beta)}{\partial \beta}\right) + \\
&+ 2Cov\left(h(x\beta), (\widehat{\beta} - \beta)\frac{\partial h(x\beta)}{\partial \beta}\right) \\
&= 0 + V((\widehat{\beta} - \beta)\frac{\partial h(x\beta)}{\partial \beta})) + 0 \\
&= \left(\frac{\partial h(x\beta)}{\partial \beta}\right)^2 V\left(\widehat{\beta}\right)
\end{aligned}$$

Opción 2: Método Delta

En nuestro caso concreto $h(xb) = \dfrac{1}{1+e^{-xb}}$, luego:

$$\frac{\partial g(\beta)}{\partial \beta} = \frac{\partial h(x\beta)}{\partial \beta} = h(x\beta)(1-h(x\beta))x$$

Entonces:

$$\begin{aligned} V\left(h(x\beta)+(\widehat{\beta}-\beta)\frac{\partial h(x\beta)}{\partial \beta}\right) &= [x\,h(x\beta)(1-h(x\beta))]^2\, V\left(\widehat{\beta}\right) \\ &= [x\,p(x,\beta)(1-p(x,\beta))]^2\, V\left(\widehat{\beta}\right) \end{aligned}$$

Por lo tanto la $V(h(x\widehat{\beta}))$ la aproximamos por

$$[x\,p(x,\beta)(1-p(x,\beta))]^2\, V\left(\widehat{\beta}\right)$$

Opción 2: Método Delta

En el caso general en que tenemos más covariables

$$g(\mathbf{b}) = h(\mathbf{x}^t\mathbf{b}) = p(x, \mathbf{b}) = \frac{1}{1 + e^{-\mathbf{x}^t\mathbf{b}}}$$

luego:

$$\begin{aligned} \frac{\partial h(\mathbf{x}^t\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \left(\frac{\partial h(\mathbf{x}^t\mathbf{b})}{\partial b_o}, \frac{\partial h(\mathbf{x}^t\mathbf{b})}{\partial b_1}, \ldots, \frac{\partial h(\mathbf{x}^t\mathbf{b})}{\partial b_p}\right)\Bigg|_{\mathbf{b}=\boldsymbol{\beta}} \\ &= h(\mathbf{x}^t\boldsymbol{\beta})(1 - h(\mathbf{x}^t\boldsymbol{\beta}))\,\mathbf{x} \end{aligned}$$

Entonces, en (1) tendremos

$$\left[p(\mathbf{x}^t, \boldsymbol{\beta})(1 - p(\mathbf{x}, \boldsymbol{\beta})\right]^2 \mathbf{x}^t\mathbf{V}(\widehat{\boldsymbol{\beta}})\,\mathbf{x}$$