MAT-466: Estimación en GLM

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Función de log-verosimulitud

Para $Y_1, \dots, Y_n$ siguiendo un GLM tenemos que

$$\begin{aligned}
\ell(\boldsymbol{\psi}) &= \sum_{i=1}^{n} \log f(y_i; \theta_i, \phi) \\
&= \sum_{i=1}^{n} \{\phi(y_i\theta_i - b(\theta_i)) + c(y_i; \phi)\} \\
&= \phi \sum_{i=1}^{n} (y_i\theta_i - b(\theta_i)) + \sum_{i=1}^{n} c(y_i; \phi),
\end{aligned}$$

donde $\boldsymbol{\psi} = (\boldsymbol{\beta}^\top, \phi)^\top$.

Observación:

Debemos destacar que $\boldsymbol{\beta}$ y ϕ son parametros ortogonales, y por tanto la inferencia puede ser realizada de manera independiente.

Función score

Para GLM la función score $\boldsymbol{U}(\boldsymbol{\beta}) = \dot{\ell}(\boldsymbol{\beta})$ adopta la forma:

$$
\begin{aligned}
\frac{\partial \ell(\boldsymbol{\psi})}{\partial \boldsymbol{\beta}} &= \phi \sum_{i=1}^{n} \left\{ Y_i \frac{\mathrm{d}\,\theta_i}{\mathrm{d}\,\mu_i} \frac{\mathrm{d}\,\mu_i}{\mathrm{d}\,\eta_i} \frac{\partial \eta_i}{\partial \boldsymbol{\beta}} - \frac{\mathrm{d}\,b(\theta_i)}{\mathrm{d}\,\theta_i} \frac{\mathrm{d}\,\theta_i}{\mathrm{d}\,\mu_i} \frac{\mathrm{d}\,\mu_i}{\mathrm{d}\,\eta_i} \frac{\partial \eta_i}{\partial \boldsymbol{\beta}} \right\} \\
&= \phi \sum_{i=1}^{n} \left\{ Y_i - \frac{\mathrm{d}\,b(\theta_i)}{\mathrm{d}\,\theta_i} \right\} \frac{\mathrm{d}\,\theta_i}{\mathrm{d}\,\mu_i} \frac{\mathrm{d}\,\mu_i}{\mathrm{d}\,\eta_i} \frac{\partial \eta_i}{\partial \boldsymbol{\beta}} \\
&= \phi \sum_{i=1}^{n} (Y_i - b'(\theta_i)) \left\{ \frac{\mathrm{d}\,\mu_i}{\mathrm{d}\,\theta_i} \right\}^{-1} \frac{\mathrm{d}\,\mu_i}{\mathrm{d}\,\eta_i} \frac{\partial \eta_i}{\partial \boldsymbol{\beta}} \\
&= \phi \sum_{i=1}^{n} (Y_i - \mu_i) V(\mu_i)^{-1} \frac{\mathrm{d}\,\mu_i}{\mathrm{d}\,\eta_i} \boldsymbol{x}_i \\
&= \phi \sum_{i=1}^{n} \omega_i^{1/2} \frac{(Y_i - \mu_i)}{\sqrt{V_i}} \boldsymbol{x}_i,
\end{aligned}
$$

donde $\omega_i = (\mathrm{d}\,\mu_i / \mathrm{d}\,\eta_i)^2 / V_i$ y $V_i = V(\mu_i)$, para $i = 1, \dots, n$.

Función score

Una manera mucho más compacta de escribir la función score en GLM es:

$$\boldsymbol{U}(\boldsymbol{\beta}) = \phi \boldsymbol{X}^\top \boldsymbol{W}^{1/2} \boldsymbol{V}^{-1/2} (\boldsymbol{Y} - \boldsymbol{\mu}),$$

donde $\boldsymbol{W} = \operatorname{diag}(\omega_1, \dots, \omega_n)$, $\boldsymbol{V} = \operatorname{diag}(V_1, \dots, V_n)$ y $\boldsymbol{Y} = (Y_1, \dots, Y_n)^\top$.

Vamos a suponer que $\boldsymbol{X}$ es matriz de rango completo cuya i-ésima fila es dada por $\boldsymbol{x}_i^\top$, para $i = 1, \dots, n$. Además, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top$ con $\mu_i = \mu_i(\boldsymbol{\beta})$.

Matriz de información de Fisher

La matriz Hessiana en GLM es dada por:

$$\begin{aligned}\frac{\partial^2 \ell(\boldsymbol{\psi})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = {} & \phi \sum_{i=1}^n (Y_i - \mu_i) \frac{\mathsf{d}^2\, \theta_i}{\mathsf{d}\, \mu_i^2} \Big(\frac{\mathsf{d}\, \mu_i}{\mathsf{d}\, \eta_i} \Big)^2 \boldsymbol{x}_i \boldsymbol{x}_i^\top \\ & + \phi \sum_{i=1}^n (Y_i - \mu_i) \frac{\mathsf{d}\, \theta_i}{\mathsf{d}\, \mu_i} \Big(\frac{\mathsf{d}^2\, \mu_i}{\mathsf{d}\, \eta_i^2} \Big)^2 \boldsymbol{x}_i \boldsymbol{x}_i^\top \\ & - \phi \sum_{i=1}^n \frac{\mathsf{d}\, \theta_i}{\mathsf{d}\, \mu_i} \Big(\frac{\mathsf{d}\, \mu_i}{\mathsf{d}\, \eta_i} \Big)^2 \boldsymbol{x}_i \boldsymbol{x}_i^\top\end{aligned}$$

De este modo, la matriz de información de Fisher asume la forma,

$$\boldsymbol{\mathcal{F}}(\boldsymbol{\beta}) = \mathsf{E}\Big\{ -\frac{\partial^2 \ell(\boldsymbol{\psi})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \Big\} = \phi \sum_{i=1}^n \frac{(\mathsf{d}\, \mu_i / \mathsf{d}\, \eta_i)^2}{V_i} \boldsymbol{x}_i \boldsymbol{x}_i^\top = \phi \boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{X}.$$

Algoritmo Fisher-scoring en GLM

El algoritmo Fisher-scoring para $\boldsymbol{\beta}$ es dado por:

$$\boldsymbol{\beta}^{(r+1)} = \boldsymbol{\beta}^{(r)} + \boldsymbol{\mathcal{F}}^{-1}(\boldsymbol{\beta}^{(r)})\boldsymbol{U}(\boldsymbol{\beta}^{(r)}),$$

y para el caso de GLM, adopta la forma:

$$\begin{aligned}
\boldsymbol{\beta}^{(r+1)} &= \boldsymbol{\beta}^{(r)} + (\boldsymbol{X}^\top \boldsymbol{W}^{(r)}\boldsymbol{X})^{-1}\boldsymbol{X}^\top {\boldsymbol{W}^{(r)}}^{1/2}{\boldsymbol{V}^{(r)}}^{-1/2}(\boldsymbol{Y} - \boldsymbol{\mu}^{(r)}) \\
&= (\boldsymbol{X}^\top \boldsymbol{W}^{(r)}\boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{W}^{(r)}\boldsymbol{X}\boldsymbol{\beta}^{(r)} \\
&\quad + (\boldsymbol{X}^\top \boldsymbol{W}^{(r)}\boldsymbol{X})^{-1}\boldsymbol{X}^\top {\boldsymbol{W}^{(r)}}^{1/2}{\boldsymbol{V}^{(r)}}^{-1/2}(\boldsymbol{Y} - \boldsymbol{\mu}^{(r)}) \\
&= (\boldsymbol{X}^\top \boldsymbol{W}^{(r)}\boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{W}^{(r)}\boldsymbol{Z}^{(r)},
\end{aligned}$$

donde

$$\boldsymbol{Z} = \boldsymbol{\eta} + \boldsymbol{W}^{-1/2}\boldsymbol{V}^{-1/2}(\boldsymbol{Y} - \boldsymbol{\mu}),$$

denota la respuesta de trabajo.

Estimación de ϕ

Tenemos que la función score asociada a ϕ, adopta la forma:

$$U(\phi) = \frac{\partial \ell(\boldsymbol{\psi})}{\partial \phi} = \sum_{i=1}^{n} (y_i \theta_i - b(\theta_i)) + \sum_{i=1}^{n} c'(y_i; \phi),$$

mientras que la información de Fisher, es dada por

$$\mathcal{F}(\phi) = -\sum_{i=1}^{n} \mathsf{E}\{c''(Y_i; \phi)\}.$$

De este modo, podemos considerar el procedimiento de estimación:

$$\phi^{(r+1)} = \phi^{(r)} + \frac{U(\phi^{(r)})}{\mathcal{F}(\phi^{(r)})}, \qquad r = 0, 1, \dots .$$

Alternativamente, podemos resolver la ecuación $U(\phi) = 0$, obteniendo

$$\sum_{i=1}^{n} c'(y_i; \widehat{\phi}) = \frac{1}{2} D(\boldsymbol{y}; \widehat{\boldsymbol{\mu}}) - \sum_{i=1}^{n} (y_i \widehat{\theta}_i - b(\widehat{\theta}_i)).$$

Estimación de ϕ

Casos particulares: Modelo normal

Es fácil notar que

$$c'(y_i;\phi) = \frac{1}{2\phi} - \frac{y_i^2}{2}, \qquad c''(y_i;\phi) = -\frac{1}{2\phi^2}.$$

Así, $\mathcal{F}(\phi) = n/(2\phi^2)$. Evidentemente,

$$\begin{aligned} U(\phi) &= \sum_{i=1}^n \Big(y_i\mu_i - \frac{\mu_i^2}{2}\Big) + \sum_{i=1}^n \Big(\frac{1}{2\phi} - \frac{y_i^2}{2}\Big) \\ &= -\frac{1}{2}\sum_{i=1}^n (y_i^2 - 2y_i\mu_i + \mu_i^2) + \frac{n}{2\phi}, \end{aligned}$$

desde $U(\phi) = 0$, sigue que

$$\widehat{\phi}^{-1} = \frac{1}{n}\sum_{i=1}^n (y_i - \widehat{\mu}_i)^2.$$

Casos particulares: Modelo normal inversa

Tenemos,

$$c'(y_i;\phi) = \frac{1}{2\phi} - \frac{y_i^2}{2}, \qquad c''(y_i;\phi) = -\frac{1}{2\phi^2}.$$

y portanto, $\mathcal{F}(\phi) = n/(2\phi^2)$. De este modo,

$$U(\phi) = \sum_{i=1}^n \Big(\frac{1}{\mu_i} - \frac{y_i}{2\mu_i^2}\Big) + \sum_{i=1}^n \Big(\frac{1}{2\phi} - \frac{y_i^2}{2}\Big).$$

Evidentemente, desde $U(\phi) = 0$ obtenemos

$$\widehat{\phi}^{-1} = \frac{1}{n} \sum_{i=1}^n (y_i - \widehat{\mu}_i)^2.$$

EX UMBRA IN SOLEM

Casos particulares: Modelo Gama

Es este caso

$$c'(y_i;\phi) = \log y_i + \log\phi + 1 - \psi(\phi), \qquad c''(y_i;\phi) = \frac{1}{\phi} - \psi'(\phi),$$

donde $\psi(z) = \Gamma'(z)/\Gamma(z)$ denota la función digama. De ahí que

$$2\{\log\widehat{\phi} - \psi(\widehat{\phi})\} = \frac{1}{n}\,D(\boldsymbol{y};\widehat{\boldsymbol{\mu}}).$$

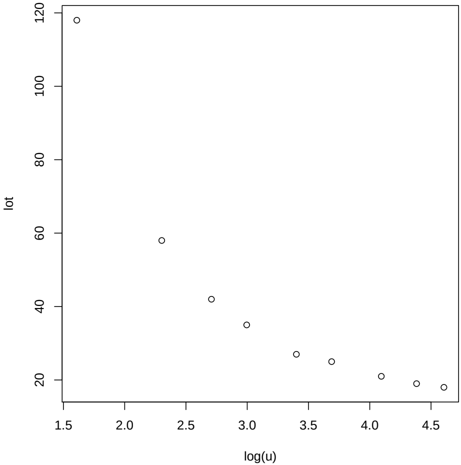
Usando la aproximación

$$\psi(\phi) \approx \log\phi - \frac{1}{2\phi} - \frac{1}{12\phi^2},$$

lleva a

$$\widehat{\phi} \approx \frac{1+\sqrt{1+2\overline{D}/3}}{2\overline{D}}, \qquad \overline{D} = \frac{1}{n}D(\boldsymbol{y};\widehat{\boldsymbol{\mu}}).$$

Datos de coagulación



Datos de coagulación de la sangre (McCullagh y Nelder, 1989, pp. 300-2)

Conjunto de datos: Coagulación inducida por dos lotes de tromboplastina.

```
clotting <- data.frame(
+    u = c(5,10,15,20,30,40,60,80,100),
+    lot = c(118,58,42,35,27,25,21,19,18))
```

Exploramos el conjunto de datos por medio del gráfico:

```
> plot(lot ~ log(u), data = clotting)
```

Datos de coagulación de la sangre

```
> fit <- glm(lot ~ log(u), data = clotting, family = Gamma)
> summary(fit)

Call:
glm(formula = lot ~ log(u), family = Gamma, data = clotting)

Deviance Residuals:
     Min         1Q    Median         3Q        Max
-0.04008   -0.03756  -0.02637    0.02905    0.08641

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.0165544  0.0009275  -17.85 4.28e-07 ***
log(u)       0.0153431  0.0004150   36.98 2.75e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 0.002446059)

    Null deviance: 3.51283  on 8  degrees of freedom
Residual deviance: 0.01673  on 7  degrees of freedom
AIC: 37.99

Number of Fisher Scoring iterations: 3
```