GLM Resum I

Densitat de la distribució exponencial dels mlg

$$f_{Y_i|(\theta_i,\Phi)}(y) \sim e^{\frac{y_i\theta_i - b(\theta_i)}{\Phi} + c(y,\phi)}$$

$$b'(\theta_i) = E[Y_i] \qquad b''(\theta_i)\,\Phi = Var(Y_i)$$

Paràmetre de dispersió

$$\Phi$$

$\sqrt{\Phi}$ també s'anomena paràmetre d'escala.

GLM Resum II

Predictor lineal

$$\eta_i = X_i\beta$$

Funció d'enllaç: link function

$$\eta_i = g(\mu_i)$$

inversa de la funció link

$$\mu_i = g^{-1}(\eta_i)$$

Inversa de $b'(\theta_i) = \mu_i$: q function

$$\theta_i = q(\mu_i)$$

GLM Resum III

Variance function

$$V(\mu_i) \quad : \quad Var(y_i) = \Phi V(\mu_i)$$

$$V(\mu_i) = b''(\theta_i)$$

Deviància i Estadístic de Pearson generalitzat I

Deviància escalada: scaled deviance

$$D^s = -2\log\left(\frac{\mathcal{L}(fitted)}{\mathcal{L}(saturated)}\right)$$
$$D^s = 2\log\left(\mathcal{L}\left(saturated\right)\right) - 2\log\left(\mathcal{L}\left(fitted\right)\right)$$

fitted $\longrightarrow \ell\left(\hat{\theta}, \Phi | y\right) = \sum_{i=1}^{n} \frac{y_i\hat{\theta}_i - b\left(\hat{\theta}_i\right)}{\Phi} + c\left(y_i, \phi\right)$

saturated $\longrightarrow \ell\left(\tilde{\theta}, \Phi | y\right) = \sum_{i=1}^{n} \frac{y_i\tilde{\theta}_i - b\left(\tilde{\theta}_i\right)}{\Phi} + c\left(y_i, \phi\right)$

on $\tilde{\eta}_i = y_i \Rightarrow \tilde{\mu}_i = g^{-1}\left(y_i\right) \Rightarrow \tilde{\theta}_i = q\left(g^{-1}\left(y_i\right)\right)$

$$D^s = 2\sum_{i=1}^{n} \frac{y_i\left(\tilde{\theta}_i - \hat{\theta}_i\right) - \left(b\left(\tilde{\theta}_i\right) - b\left(\hat{\theta}_i\right)\right)}{\Phi} \sim \chi^2_{n-k}$$

Deviància i Estadístic de Pearson generalitzat II

Deviància, o deviància no escalada: unscaled deviance

$$D =_{\Phi D^s}= 2\sum_{i=1}^{n}\left(y_i\left(\tilde{\theta}_i - \hat{\theta}_i\right) - \left(b\left(\tilde{\theta}_i\right) - b\left(\hat{\theta}_i\right)\right)\right) \sim \Phi\chi^2_{n-k}$$

Estadístic de Pearson generalitzat

$$\chi^2 = \sum_{i=1}^{n}\left(\frac{y_i - \hat{\mu}}{V\left(\hat{\mu}_i\right)}\right) = \Phi\sum_{i=1}^{n}\left(\frac{y_i - \hat{\mu}}{Var\left(y_i\right)}\right) \sim \Phi\chi^2_{n-k}$$

Paràmetre de dispersion & Residuals I

Φ a partir de la deviància

$D \sim \Phi\chi^2_{n-k} \Rightarrow E\left[D\right] = \Phi\left(n-k\right)$

$$\hat{\Phi}_{deviance} = \frac{D}{n-k}$$

Φ a partir de Pearson `(per defecte a R)`

$\chi^2 \sim \Phi\chi^2_{n-k} \Rightarrow E\left[\chi^2\right] = \Phi\left(n-k\right)$

$$\hat{\Phi}_{pearson} = \frac{\chi^2}{n-k}$$

Paràmetre de dispersion & Residuals II

Residuals

- Deviance `(per defecte a R)`

$$r_{D,i} = sign\left(y_i - \hat{\mu}_i\right)\sqrt{d_i}$$

on $d_i = 2\left(y_i\left(\tilde{\theta}_i - \hat{\theta}_i\right) - \left(b\left(\tilde{\theta}_i\right) - b\left(\hat{\theta}_i\right)\right)\right)$

Nota: $D\left(y;\mu\right) = \sum_{i=1}^{n} r_{D,i}^2$

- Pearson

$$r_{P,i} = \frac{y_i - \hat{\mu}_i}{\sqrt{V\left(\hat{\mu}_i\right)}}$$

Nota: $\chi^2 = \sum_{i=1}^{n} r_{D,i}^2$

Quasiversemblança

Funció de quasiversemblança

$$Q\left(\mu, \Phi; y\right) = \frac{1}{\Phi} \int \frac{y-\mu}{V\left(\mu\right)} \mathrm{d}\mu$$

Funció U "score de la quasiversemblança"

$$U\left(\mu, \Phi; y\right) = \frac{1}{\Phi} \frac{y-\mu}{V\left(\mu\right)}$$

Quasi-deviance

$$\int_{\mu}^{y} \frac{y-s}{V\left(s\right)} \mathrm{d}s = \Phi\left(Q\left(y, \Phi; y\right) - Q\left(\mu, \Phi; y\right)\right)$$