M4: Poisson and gamma regression

THA4315 H2018
27.09.2018

Poisson process → number of events in time and/or space follows a Poisson distribution $\underbrace{}_{Y_i}$ $i=1,\dots,n$

$$Y_i \sim \text{Poisson}(\lambda_i)$$

$$f(y_i) = \frac{\lambda_i^{y_i}}{y_i!} e^{-\lambda_i} \qquad y_i = 0,1,2,\dots$$

$E(Y_i) = \lambda_i$, $Var(Y_i) = \lambda_i$ ← NB: $E(Y_i) = Var(Y_i)$

<u>Univ. exponential family</u>: $f(y_i;\theta_i,\phi) = \exp\left(\frac{y_i \cdot \theta_i - b(\theta_i)}{\phi} w_i + c(y_i,\phi,w_i)\right)$

$\theta_i = \ln \lambda_i$, $\phi = 1$, $w_i = 1$

Canonical link: $\eta_i = \theta_i$

$\theta_i = \eta_i = \ln(\lambda_i)$

Poisson regression

1) Random component: $Y_i \sim \text{Poisson}(\lambda_i)$
   with $\mu_i = E(Y_i) = \lambda_i$ and $Var(Y_i) = \lambda_i$

2) Systematic component: $\eta_i = x_i^T \beta$

3) Link:

   canonical link

   $\eta_i = \ln(\lambda_i)$ log link

   $\lambda_i = \exp(\eta_i)$ exponential response function

Interpreting regression coefficients

$$\lambda_i = \exp(\underset{\substack{\uparrow \\ x_i^T\beta = \beta_0 + \beta_1 x_{i1} + \dots}}{\eta_i}) = \exp(\beta_0) \cdot \exp(\beta_1)^{x_{i1}} \cdots \exp(\beta_k)^{x_{ik}}$$

What happens if x_{i1} increase to $x_{i1}+1$:

$$\lambda_i = \exp(\beta_0) \cdot \exp(\beta_1) \cdot \exp(\beta_1)^{x_{i1}} \cdot \dots \quad \Big\} \text{ multiplicative effect of covariates}$$

$\Downarrow$

$E(Y_i)$ will change with a factor $\exp(\beta_1)$

Ex: crabs & satellites:

1) $\eta_i = \beta_0$: $\lambda_i = \exp(\beta_0)$

and $\hat{\lambda}_i = \exp(\hat{\beta}_0)$ = average number of Sa in data

$= 2.91$

2) $\eta_i = \beta_0 + \beta_1 \underset{\substack{\uparrow \\ \text{Width}}}{x_{i1}}$ $\qquad \hat{\lambda}_i$ "increase" by $e^{\hat{\beta}_1} = 1.17$
($\uparrow$ multiply by)

if W increase by one unit

3) $\overline{x_1}$ = average width $= 26.3$

$\hat{\lambda} = \exp(\hat{\beta}_0 + \hat{\beta}_1 \cdot \overline{x}_1) \approx 2.74$

Maximum likelihood estimation

$$L(\beta) = \prod_{i=1}^{n} \frac{\lambda_i^{y_i}}{y_i!} e^{-\lambda_i} \qquad \begin{cases} \eta_i = \ln \lambda_i \\ \lambda_i = \exp(\eta_i) \end{cases}$$

$$l(\beta) = \sum_{i=1}^{n} \left(y_i \ln \lambda_i - \lambda_i - \ln(y_i!) \right)$$

$$= \sum_{i=1}^{n} \left[y_i \eta_i - e^{\eta_i} - \ln(y_i!) \right] = \sum_{i=1}^{n} l_i(\beta)$$

$$S(\beta) = \sum_{i=1}^{n} s_i(\beta) = \sum_{i=1}^{n} \underbrace{\frac{\partial l_i}{\partial \eta_i}}_{1\times 1} \underbrace{\frac{\partial \eta_i}{\partial \beta}}_{p\times 1} \qquad \frac{\partial (x_i^T \beta)}{\partial \beta} = \underset{p\times 1}{x_i}$$

$$\frac{\partial l_i}{\partial \eta_i} = \frac{\partial}{\partial \eta_i} \left(y_i \eta_i - e^{\eta_i} - \ln(y_i!) \right) = y_i - e^{\eta_i}$$

$$S(\beta) = \sum_{i=1}^{n} (y_i - e^{\eta_i}) x_i = \sum_{i=1}^{n} \overbrace{(y_i - \lambda_i) x_i}^{s_i(\beta)}$$

Observe $E(S(\beta)) = \sum_{i=1}^{n} (\underbrace{E(Y_i)}_{\lambda_i} - \lambda_i) \cdot x_i = 0$

$$F(\beta) = Cov(S(\beta)) = \sum_{i=1}^{n} Cov(s_i(\beta)) = \sum_{i=1}^{n} F_i(\beta)$$

$$F_i(\beta) = Cov\left([Y_i - \lambda_i] x_i \right) \qquad Cov(Z) = E\left((Z-\mu_Z)(Z-\mu_Z)^T \right)$$

$$= E\left(\left((Y_i - \lambda_i) x_i - 0 \right) \left((Y_i - \lambda_i) x_i - 0 \right)^T \right)$$

$$= E\left((Y_i - \lambda_i)^2 \right) x_i x_i^T = Var(Y_i) \cdot x_i x_i^T = \lambda_i x_i x_i^T$$

$$\underline{\underline{F(\beta) = \sum_{i=1}^{n} \lambda_i x_i x_i^T}} \qquad (p \times p)$$

Hypothesis testing $\left\{ \begin{array}{l} \text{Wald} \quad C\beta = d \\ \text{LRT} \end{array} \right\}$ "same" as for binary regression

The score test

A: large model : p_A param. $\Rightarrow$ ML $\hat{\beta}$ Ex: $Sa \sim W + C$ $p_A = 5$

B: small model : p_B param $\Rightarrow$ ML $\tilde{\beta}$ $Sa \sim W$ $p_B = 2$

H_0: large and small model is equally good
" the β's in β_A but not in β_B are all zero "

H_1: not so

$$\underbrace{U}_{\text{test statistics}} = \underbrace{\big(s(\tilde{\beta}) - \overset{0}{s(\hat{\beta})}\big)^T}_{[(p_A-p_B)\times 1]^T} \underbrace{F^{-1}(\tilde{\beta})}_{(p_A-p_B)\times(p_A-p_B)} \underbrace{\big(s(\tilde{\beta}) - \overset{0}{s(\hat{\beta})}\big)}_{(p_A-p_B)\times 1}$$

If U is large $\Rightarrow$ "H_0 is not correct" $\rightarrow$ reject H_0
$\Rightarrow$ the B model is not as good as the A model

When n is large $U \sim \chi^2_{p_A - p_B}$

See R-code on module page for details on calculating U.