STAT 442 Sept. 30

Brief review

Poisson models · y · $P(y=y) = \dfrac{\mu^y e^{-\mu}}{y!}$

$\mu = E(y) = Var(y)$!

Large sample inference

$\hat{\mu} = \bar{y}$ s.e.$(\hat{\mu}) = \sqrt{\bar{y}/n}$

Rates. often $\lambda = \mu / T \rightarrow$ persons

more relevant person-years

base inference on logs

↳ for relative risk

$\theta = \log(\mu_1 / \mu_2)$ ↳ λ_1 / λ_2

observe counts y_1, y_2

$\hat{\theta} = \log(y_1 / y_2)$

s.e.$(\hat{\theta}) = \sqrt{\dfrac{1}{y_1} + \dfrac{1}{y_2}}$

conf. for $\hat{\theta} \rightarrow$ exponentiate to

get C.I for $\mu_1 / \mu_2 \rightarrow \times \dfrac{T_2}{T_1} \rightarrow \lambda_1 / \lambda_2$

Poisson Regression

$Y_i \to$ set of counts $i = 1 \ldots n$

$K_i \to$ corresponding population size values - non-random

$X_i \to$ explanatory - year

= ex. Y: shark attacks, K_i - Florida pop^n

$E(Y_i) = \mu_i$ $\quad \lambda_i = \mu_i / K_i$ = attack rate

$r_i = \hat{\lambda}_i = y_i / K_i \to$ raw or crude rate

Want a regression type model

?? $\quad r_i = \beta_0 + \beta_1 X_i + \underbrace{\epsilon_i}$??

$Var(r_i) = Var(y_i) / K_i^2 \to$ not constant over i

ϵ_i - "peculiar dist^n"

$\to$ discrete with fractional + z - values

Need to work with $y_i \to \text{Poisson}(\mu_i)$

$\Rightarrow$?? $\to E(y_i) = (k_i(\beta_0 + \beta_1 X))$

so we turn to logarithms

work with $\log(\mu_i) = \eta_i$

$$\eta_i = \underbrace{\log k_i}_{\text{offset}} + \beta_0 + \beta_1 X_1$$

$\to$ adjustment for population size
- persons
- person-years

- $\log(\mu_i) = \eta_i =$ linear expression
  - $\to$ log-linear model
  - $\to$ extends multiple linear regression

$$\eta_i = \log k_i + \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

- example of generalized linear model
  - 1970's development

- include logistic regression + more

GLIM - generalized linear interactive modelling

→ in general, situation is

$y \sim$ distribution (Poisson, Expon, Binary...)

$\mu = E(y)$ - function called a link function - g

$$g(\mu) = \beta_0 + \beta_1 X_1 \ldots + \beta_p X_p$$

→ estimation done by maximum likelihood
→ iterative numerical solutions
- no algebraic expression for $\hat{\beta}$.
+ no simple exact procedures
→ asymptotic likelihood theory

Maximum likelihood for Poisson Regression

$Y_i \sim \text{Poisson}(\mu_i)$ independent

likelihood $= \prod_{i=1}^{n} \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!}$

log-lik $= \sum_{i=1}^{n} y_i \eta_i - e^{\eta_i} - \underbrace{\log(y_i!)}_{\text{no } \mu_i}$

$\eta_i = \log \mu_i$

what about β's?

$\eta_i = \text{offset} + \beta_0 + \beta_1 X_1 \ldots \beta_k X_k$

use chain rule $\frac{\partial \cdot}{\partial \beta_j} = \frac{\partial \cdot}{\partial \eta} \cdot \frac{\partial \eta_i}{\partial \beta_j}$

$\frac{\partial \eta_i}{\partial \beta_j} = (x_{ij})$

$\frac{\partial \log lik}{\partial \eta_i} = (y_i - e^{\eta_i})$

$X = (x_{ij})$ - x matrix . with $x_{0j} = 1$ for β_0

$$* \quad X^t(\underset{\sim}{y} - e^{\underset{\sim}{\eta}}) = 0 \qquad \underset{\sim}{y} = \begin{pmatrix} y_1 \\ y_n \end{pmatrix} \quad \underset{\sim}{\eta} = \begin{pmatrix} \eta_1 \\ \eta_n \end{pmatrix}$$

LHS → score function

LHS = 0 → max. lik. equation

- $e^{\underset{\sim}{\eta}}$: non-linear

→ non-linear equation algorithm

- Newton-Raphson
- by re-arranging * slightly can form iterative solution as iterative weighted least squares

→ do iterated weighted regressions

→ $\underset{\sim}{\hat{\beta}}$ - as m.l.e.

- asymptotically unbiased
- " normal.

- asymptotic variance $= I^{-1}$ → inverse of Fisher information

$$\mathcal{I} = E\left[\frac{\partial \log lik}{\partial \beta_i \partial \beta_j}\right] \quad (k+1)\times(k+1) \quad ?$$

$\hookrightarrow$ $\dfrac{\partial \text{ LHS}}{\partial \underline{\beta}}$ + score function

$$\left[\frac{\partial \log}{\partial \beta_i \beta_j}\right] = X^t \left(-e^{\underline{\eta}_i}\right) X$$

diag matrix with entries e^{η}

$$I = (X^t V X) \quad \text{observed info}$$

$$\mathcal{I} = E(I) = I$$ → since log canonical link function for Poisson

To make inference for $\underline{\beta}$ - or individual β_js

have estimated variance-covariance matrix for $\hat{\beta}$: $(\hat{\mathcal{I}})^{-1}$

$$\hat{I} = \left(X^t \hat{V} X \right)$$

↓ plug in

$\hat{\beta} \rightarrow \hat{m}_i$

8/8

→ in R → max. lik. for GLIM's implemented in glm

- syntax like lm - regression
  → specify - family → poisson
  - offset if necessary

→ familiar output - regression format

→ based on s.e.$(\hat{\beta}_j)$: from $\sqrt{\operatorname{diag}(\hat{I})^{-1}}$

→ Test for $\beta_j = 0$

$$z = \hat{\beta}_j / \text{s.e.}(\hat{\beta}_j)$$