3/8/22

Model Comparison Tests Based on Deviance (Sec. 4.4.2 + 3.4.4)

- General GLM

H_0: Reduced model M_R

H_a: Full model M_F

M_R is nested within M_F → Can obtain M_R by setting a subset of coefs. in M_F to zero.

Likelihood ratio test:

$$\text{Statistic (LRT)} = -2\,[\,L_R - L_F\,]$$

L_F: maximized log-likelihood assuming full model

L_R: maximized log-likelihood assuming reduced model

$$= \underbrace{-2\,[\,L_R - L_S\,]}_{\text{(Residual) Deviance for reduced model}} - \underbrace{[\,-2\,[\,L_F - L_S\,]\,]}_{\text{Deviance for full model}}$$

L_S: "saturated model" = perfect fit

R: anova(M_R, M_F)

For normal linear models: residual deviance

$$= SSE$$

$$= \sum_{i=1} (y_i - \hat{y}_i)^2$$

GLMs for Count Data (Poisson Regression) Sec. 3.3

Response = Y $\longrightarrow$ count # of ___

e.g. $Y =$ # of siblings
# of atoms emitted from radioactive source in 1 min.
customers that arrive in 1 hour.

GLM: ① Distribution (random component)

$$Y \sim \text{Pois}(\mu) \qquad \mu > 0$$

i.e.
$$P(Y=y) = \begin{cases} \dfrac{e^{-\mu}\mu^{y}}{y!} & y = 0,1,2,\ldots \\ 0 & \text{else} \end{cases}$$

② Link function: $g(\mu) = \log \mu$

③ Systematic component: $\eta = \alpha + \beta_1 x_1 + \cdots \beta_k x_k$

Model: $\log \mu = \alpha + \beta_1 x_1 + \cdots + \beta_k x_k$

where $Y \sim \text{Pois}(\mu)$

Properties of Poisson distribution: $Y \sim \text{Pois}(\mu)$

① $E(Y) = \mu$

② $\text{Var}(Y) = \mu \implies SD(Y) = \sqrt{\mu}$

If in observed data: Variance > mean

$\longrightarrow$ overdispersion

Why this may occur?

- unmeasured predictors.

<u>Reality</u>:

Group A: $Y \sim Pois(\mu_A)$

Group B: $Y \sim Pois(\mu_B)$

<u>Interpreting</u>:

$$\hat{\mu} = e^{a+bx} \qquad e^{a+b(x+1)} = e^{a+bx} \cdot \underline{e^{b}}$$

multiplicative effect of x increasing by 1 unit on $\hat{\mu}$