STA 312f2022 Formulas [1]

$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)}$ $\quad$ $Pr(A) = \sum_{j=1}^k Pr(A|B_j)Pr(B_j)$ $\quad$ $Pr(B|A) = \frac{Pr(A|B)Pr(B)}{Pr(A|B)Pr(B)+Pr(A|B^c)Pr(B^c)}$

Bernoulli $\quad P(y) = \pi^y(1-\pi)^{1-y}$ for $y = 0, 1$

Binomial $\quad P(y) = \binom{n}{y}\pi^y(1-\pi)^{n-y}$ for $y = 0, \ldots, n$

Poisson $\quad P(y) = \frac{e^{-\lambda}\lambda^y}{y!}$ for $y = 0, \ldots$

Hypergeometric $\quad P(y) = \frac{\binom{M}{y}\binom{N-M}{n-y}}{\binom{N}{n}}$, where $\binom{a}{b}$ must make sense.

$Z_1 = \frac{\sqrt{n}(p-\pi_0)}{\sqrt{\pi_0(1-\pi_0)}}$ $\quad Z_2 = \frac{\sqrt{n}(p-\pi_0)}{\sqrt{p(1-p)}}$ $\quad p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$

```
> qnorm(0.975)
[1] 1.959964
> qnorm(0.995)
[1] 2.575829
```

Multinomial

$P(n_1, \ldots, n_c) = \binom{n}{n_1 \cdots n_c}\pi_1^{n_1}\cdots\pi_c^{n_c}$ $\quad$ $\ell(\boldsymbol{\pi}) = \prod_{i=1}^n \pi_1^{y_{i,1}}\pi_2^{y_{i,2}}\cdots\pi_c^{y_{i,c}} = \pi_1^{n_1}\pi_2^{n_2}\cdots\pi_c^{n_c}$

$G^2 = -2\log\left(\frac{\max_{\beta\in\mathscr{B}_0}\ell(\beta)}{\max_{\beta\in\mathscr{B}}\ell(\beta)}\right) = -2\log\left(\frac{\ell(\widehat{\beta}_0)}{\ell(\widehat{\beta})}\right)$ $\quad$ $G^2 = 2\sum_{j=1}^c n_j \log\left(\frac{n_j}{n\widehat{\pi}_j}\right) = 2\sum_{j=1}^c n_j\log\left(\frac{n_j}{\widehat{\mu}_j}\right)$

$X^2 = \sum_{j=1}^c \frac{(n_j-\widehat{\mu}_j)^2}{\widehat{\mu}_j}$ $\quad$ $n_{i+} = \sum_{j=1}^J n_{ij} \quad n_{+j} = \sum_{i=1}^I n_{ij} \quad \widehat{\mu}_{ij} = \frac{n_{i+}n_{+j}}{n}$

Odds $= \frac{\pi}{1-\pi}$ $\quad$ $\theta = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}$

Logistic regression

$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_{i,1} + \ldots + \beta_k x_{i,k} = \mathbf{x}_i'\boldsymbol{\beta}$ $\quad$ $\pi_i = \frac{e^{\mathbf{x}_i'\boldsymbol{\beta}}}{1+e^{\mathbf{x}_i'\boldsymbol{\beta}}}$

If $\mathbf{z} \sim N_p(\boldsymbol{\mu}, \Sigma)$, then $\mathbf{Az} \sim N_q(\mathbf{A}\boldsymbol{\mu}, \mathbf{A\Sigma A}')$, and $w = (\mathbf{z}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{z}-\boldsymbol{\mu}) \sim \chi^2(p)$

$\widehat{\boldsymbol{\beta}}_n \stackrel{\cdot}{\sim} N_{k+1}(\boldsymbol{\beta}, \mathbf{V}_n)$ $\quad$ $W_n = (\mathbf{L}\widehat{\boldsymbol{\beta}}_n - \mathbf{h})'\left(\mathbf{L}\widehat{\mathbf{V}}_n\mathbf{L}'\right)^{-1}(\mathbf{L}\widehat{\boldsymbol{\beta}}_n - \mathbf{h}) \stackrel{\cdot}{\sim} \chi^2(r)$
if $H_0 : \mathbf{L}\boldsymbol{\beta} = \mathbf{h}$ is true.

[1] This formula sheet was prepared by Jerry Brunner, Department of Statistics, University of Toronto. The LaTeX source code is available from the course website: http://www.utstat.toronto.edu/brunner/oldclass/312f22

Poisson regression

$\log(\lambda_i) = \beta_0 + \beta_1 x_{i,1} + \ldots + \beta_{p-1} x_{i,p-1}$

Multinomial logit regression

$$\log\left(\frac{\pi_1}{\pi_3}\right) = \beta_{0,1} + \beta_{1,1}x_1 + \ldots + \beta_{k,1}x_k = L_1$$

$$\log\left(\frac{\pi_2}{\pi_3}\right) = \beta_{0,2} + \beta_{1,2}x_1 + \ldots + \beta_{k,2}x_k = L_2$$

$$\pi_1 = \frac{e^{L_1}}{1 + e^{L_1} + e^{L_2}}$$

$$\pi_2 = \frac{e^{L_2}}{1 + e^{L_1} + e^{L_2}}$$

$$\pi_3 = \frac{1}{1 + e^{L_1} + e^{L_2}}$$

```
> df = 1:8
> CriticalValue = qchisq(0.95,df)
> round(rbind(df,CriticalValue),3)
               [,1]  [,2]  [,3]  [,4]  [,5]   [,6]   [,7]   [,8]
df            1.000 2.000 3.000 4.000  5.00  6.000  7.000  8.000
CriticalValue 3.841 5.991 7.815 9.488 11.07 12.592 14.067 15.507
```