Equations
STAT 475/875

Section 1.1

- Binomial distribution: $P(W = w) = \dfrac{n!}{w!(n-w)!}\pi^{w}(1-\pi)^{n-w}$ for w = 0,1,…,n with $E(W) = n\pi$ and $Var(W) = n\pi(1-\pi)$
- $\hat{\pi} = w/n$
- Wald confidence interval for π is $\hat{\pi} \pm Z_{1-\alpha/2}\sqrt{\dfrac{\hat{\pi}(1-\hat{\pi})}{n}}$
- Wilson interval for π is $\tilde{\pi} \pm \dfrac{Z_{1-\alpha/2}n^{1/2}}{n+Z_{1-\alpha/2}^{2}}\sqrt{\hat{\pi}(1-\hat{\pi})+\dfrac{Z_{1-\alpha/2}^{2}}{4n}}$ with $\tilde{\pi} = \dfrac{w + Z_{1-\alpha/2}^{2}/2}{n+Z_{1-\alpha/2}^{2}}$
- Agresti-Coull interval for π is $\tilde{\pi} \pm Z_{1-\alpha/2}\sqrt{\dfrac{\tilde{\pi}(1-\tilde{\pi})}{n+Z_{1-\alpha/2}^{2}}}$
- True confidence level is $\sum_{w=0}^{n} I(w)\binom{n}{w}\pi^{w}(1-\pi)^{n-w}$ where I(w) = 1 if the interval for a w contains π and 0 otherwise
- Expected length is $\sum_{w=0}^{n} L(w)\binom{n}{w}\pi^{w}(1-\pi)^{n-w}$ where L(w) is the length of an interval at a particular value of w
- Clopper-Pearson confidence interval for π is Beta(α/2; w, n-w+1) < π < Beta(1-α/2; w+1, n-w)
- Score test statistic for testing $\pi = \pi_0$ is $Z_S = \dfrac{\hat{\pi}-\pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}}$ where a standard normal distribution approximation is used
- General form of LR statistic: $\Lambda = \dfrac{\text{Max. lik. when parameters satisfy } H_0}{\text{Max. lik. when parameters satisfy } H_0 \text{ or } H_a}$ and $-2\log(\Lambda)$ is approximately a χ^2 random variable for large samples under H_0 with degrees of freedom equal to the difference in dimension between the alternative and null hypothesis parameter spaces

Section 1.2

- $\hat{\pi}_j = w_j / n_j$
- Wald confidence interval for $\pi_1 - \pi_2$ is $\hat{\pi}_1 - \hat{\pi}_2 \pm Z_{1-\alpha/2}\sqrt{\dfrac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1}+\dfrac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}$
- Agresti-Caffo confidence interval is $\tilde{\pi}_1 - \tilde{\pi}_2 \pm Z_{1-\alpha/2}\sqrt{\dfrac{\tilde{\pi}_1(1-\tilde{\pi}_1)}{n_1+2}+\dfrac{\tilde{\pi}_2(1-\tilde{\pi}_2)}{n_2+2}}$ where $\tilde{\pi}_1 = \dfrac{w_1+1}{n_1+2}$ and $\tilde{\pi}_2 = \dfrac{w_2+1}{n_2+2}$
- Score interval inverts

$$\frac{|\hat{\pi}_1 - \hat{\pi}_2 - d|}{\sqrt{\hat{\pi}_1^{(0)}(1-\hat{\pi}_1^{(0)})/n_1 + \hat{\pi}_2^{(0)}(1-\hat{\pi}_2^{(0)})/n_2}} < Z_{1-\alpha/2}$$

- Score test statistic for testing $\pi_1 - \pi_2 = 0$ is $Z_S = \dfrac{\hat{\pi}_1 - \hat{\pi}_2}{\sqrt{\bar{\pi}(1-\bar{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ where a standard normal distribution approximation is used and $\bar{\pi} = w_+ / n_+$
- Pearson chi-square test statistic for testing $\pi_1 - \pi_2 = 0$ is $X^2 = \sum_{j=1}^{2} \dfrac{(w_j - n_j\bar{\pi})^2}{n_j\bar{\pi}(1-\bar{\pi})}$ where a χ^2_1 distribution approximation is used
- LRT statistic for testing $\pi_1 - \pi_2 = 0$ is
$$-2\log(\Lambda) = -2\left[w_1\log\left(\frac{\bar{\pi}}{\hat{\pi}_1}\right) + (n_1 - w_1)\log\left(\frac{1-\bar{\pi}}{1-\hat{\pi}_1}\right) + w_2\log\left(\frac{\bar{\pi}}{\hat{\pi}_2}\right) + (n_2 - w_2)\log\left(\frac{1-\bar{\pi}}{1-\hat{\pi}_2}\right)\right]$$
- $RR = \dfrac{\pi_1}{\pi_2}$ and $RR = \dfrac{\hat{\pi}_1}{\hat{\pi}_2} = \dfrac{w_1 n_2}{w_2 n_1}$
- Wald confidence interval for RR is $\exp\left[\log(\hat{\pi}_1/\hat{\pi}_2) \pm Z_{1-\alpha/2}\sqrt{Var(\log(\hat{\pi}_1/\hat{\pi}_2))}\right]$ where
$Var(\log(\hat{\pi}_1/\hat{\pi}_2)) = \dfrac{1}{w_1} - \dfrac{1}{n_1} + \dfrac{1}{w_2} - \dfrac{1}{n_2}$
- $OR = \dfrac{odds_1}{odds_2} = \dfrac{\pi_1/(1-\pi_1)}{\pi_2/(1-\pi_2)} = \dfrac{\pi_1(1-\pi_2)}{\pi_2(1-\pi_1)}$ and $OR = \dfrac{odds_1}{odds_2} = \dfrac{\hat{\pi}_1(1-\hat{\pi}_2)}{\hat{\pi}_2(1-\hat{\pi}_1)} = \dfrac{w_1(n_2 - w_2)}{w_2(n_1 - w_1)}$
- Wald confidence interval for OR is $\exp\left[\log(OR) \pm Z_{1-\alpha/2}\sqrt{\frac{1}{w_1} + \frac{1}{n_1 - w_1} + \frac{1}{w_2} + \frac{1}{n_2 - w_2}}\right]$
- $OR = \dfrac{(w_1 + 0.5)(n_2 - w_2 + 0.5)}{(w_2 + 0.5)(n_1 - w_1 + 0.5)}$

Chapter 2

- Logistic regression model is $\log\left(\dfrac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$
- Estimated covariance matrix form when there is only one explanatory variable:
$$-E\left(\begin{bmatrix} \frac{\partial^2 \log[L(\beta_0,\beta_1 \mid y_1,\ldots,y_n)]}{\partial\beta_0^2} & \frac{\partial^2 \log[L(\beta_0,\beta_1 \mid y_1,\ldots,y_n)]}{\partial\beta_0\partial\beta_1} \\ \frac{\partial^2 \log[L(\beta_0,\beta_1 \mid y_1,\ldots,y_n)]}{\partial\beta_0\partial\beta_1} & \frac{\partial^2 \log[L(\beta_0,\beta_1 \mid y_1,\ldots,y_n)]}{\partial\beta_1^2} \end{bmatrix}\right)^{-1}\Bigg|_{\beta_0=\hat{\beta}_0,\beta_1=\hat{\beta}_1}$$
- Wald test statistic for testing $\beta_r = 0$ is $Z_W = \dfrac{\hat{\beta}_r}{\sqrt{Var(\hat{\beta}_r)}}$ where a standard normal approximation is used
- LRT statistic for testing a set of β's are equal to 0 is
$-2\log(\Lambda) = -2\log\left(\dfrac{L(\boldsymbol{\beta}^{(0)} \mid y_1,\ldots,y_n)}{L(\boldsymbol{\beta}^{(a)} \mid y_1,\ldots,y_n)}\right) = -2\left[\sum_{i=1}^{n} y_i\log\left(\dfrac{\hat{\pi}_i^{(0)}}{\hat{\pi}_i^{(a)}}\right) + (1-y_i)\log\left(\dfrac{1-\hat{\pi}_i^{(0)}}{1-\hat{\pi}_i^{(a)}}\right)\right]$ where a χ^2_q distribution approximation is used (q is number of β's set to 0 in the null hypothesis)

- $OR = \exp(c\beta_1)$ for an one explanatory variable logistic regression model and a c-unit increase in x_1; Wald confidence interval for OR is $\exp\left(c\hat{\beta}_1 \pm cZ_{1-\alpha/2}\sqrt{Var(\hat{\beta}_1)}\right)$
- Profile likelihood interval for β_1 uses $-2\log\left(\frac{L(\tilde{\beta}_0, \beta_1 \mid y_1, \ldots, y_n)}{L(\hat{\beta}_0, \hat{\beta}_1 \mid y_1, \ldots, y_n)}\right)$
- In general for two random variables W_1 and W_2 and constants a_1 and a_2, we have $Var(a_1W_1 + a_2W_2) = a_1^2Var(W_1) + a_2^2Var(W_2) + 2a_1a_2Cov(W_1, W_2)$
- Wald confidence interval for π is

$$\frac{\exp\left(\hat{\beta}_0 + \hat{\beta}_1x_1 + \cdots + \hat{\beta}_px_p \pm Z_{1-\alpha/2}\sqrt{Var(\hat{\beta}_0 + \hat{\beta}_1x_1 + \cdots + \hat{\beta}_px_p)}\right)}{1 + \exp\left(\hat{\beta}_0 + \hat{\beta}_1x_1 + \cdots + \hat{\beta}_px_p \pm Z_{1-\alpha/2}\sqrt{Var(\hat{\beta}_0 + \hat{\beta}_1x_1 + \cdots + \hat{\beta}_px_p)}\right)}$$

  where $Var(\hat{\beta}_0 + \hat{\beta}_1x_1 + \cdots + \hat{\beta}_px_p) = \sum_{i=0}^{p} Var(\hat{\beta}_i) + 2\sum_{i=0}^{p-1}\sum_{j=i+1}^{p} x_ix_jCov(\hat{\beta}_i, \hat{\beta}_j)$
- Convergence of the parameter estimates occurs when $\frac{\left|G^{(k)} - G^{(k-1)}\right|}{0.1 + \left|G^{(k)}\right|} < \epsilon$ where $G^{(k)}$ denotes the residual deviance at iteration i
- Probit regression model is probit(π) = $\beta_0 + \beta_1x_1 + \cdots \beta_px_p$
- Complementary log-log regression model is log[-log(1 – π)] = $\beta_0 + \beta_1x_1 + \cdots \beta_px_p$

Chapter 3

- Multinomial PMF for $n_1, \ldots, n_J$: $\frac{n!}{\prod_{j=1}^{J} n_j!}\prod_{j=1}^{J}\pi_j^{n_j}$
- $P(X = i, Y = j) = \pi_{ij}$
- Independence: $\pi_{ij} = \pi_{i+}\pi_{+j}$
- $P(Y = j \mid X = i) = \pi_{j|i}$
- $X^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(n_{ij} - n_{i+}n_{+j}/n)^2}{n_{i+}n_{+j}/n}$; under independence, X^2 has a $\chi^2_{(I-1)(J-1)}$ distribution for a large sample
- $-2\log(\Lambda) = 2\sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij}\log\left(\frac{n_{ij}}{n_{i+}n_{+j}/n}\right)$; under independence, $-2\log(\Lambda)$ has a $\chi^2_{(I-1)(J-1)}$ distribution for a large sample
- P-value for Monte Carlo test involving X^2: $(\text{Number of } X^{2*} \geq X^2)/B$
- Multinomial regression model: $\log(\pi_j/\pi_1) = \beta_{j0} + \beta_{j1}x_1 + \ldots + \beta_{jp}x_p$ for j = 2, …, J
- For one explanatory variable: $\pi_1 = \frac{1}{1 + \sum_{j=2}^{J}\exp(\beta_{j0} + \beta_{j1}x)}$ and $\pi_j = \frac{\exp(\beta_{j0} + \beta_{j1}x)}{1 + \sum_{j=2}^{J}\exp(\beta_{j0} + \beta_{j1}x)}$
- Proportional odds model: $\text{logit}[P(Y \leq j)] = \beta_{j0} + \beta_1x_1 + \cdots + \beta_px_p$

- For one explanatory variable: $\pi_j = \frac{\exp(\beta_{j0} + \beta_1 x)}{1 + \exp(\beta_{j0} + \beta_1 x)} - \frac{\exp(\beta_{j-1,0} + \beta_1 x)}{1 + \exp(\beta_{j-1,0} + \beta_1 x)}$ for j = 2, …, J – 1, π_1 = $\exp(\beta_{10} + \beta_1 x) \big/ \left[1 + \exp(\beta_{10} + \beta_1 x)\right]$, and π_J = $1 - \exp(\beta_{J-1,0} + \beta_1 x) \big/ \left[1 + \exp(\beta_{J-1,0} + \beta_1 x)\right]$
- Nonproportional odds model: $\text{logit}(P(Y \le j)) = \beta_{j0} + \beta_{j1}x_1 + \cdots + \beta_{jp}x_p$
- Adjacent-categories model: $\log(\pi_j / \pi_{j+1}) = \beta_{j0} + \beta_{j1}x_1 + \cdots + \beta_{jp}x_p$

Chapter 4

- Poisson PMF: $P(Y = y) = \frac{\exp(-\mu)\mu^y}{y!}$ where E(Y) = μ and Var(Y) = μ
- $L(\mu; y_1 \ldots, y_n) = \prod_{k=1}^{n} \frac{\exp(-\mu)\mu^{y_k}}{y_k!}$
- $\hat{\mu} = n^{-1}\sum_{k=1}^{n} y_k$
- $\text{Var}(\hat{\mu}) = \hat{\mu} / n$
- Score test statistic for testing $\mu = \mu_0$ is $Z_S = \frac{\hat{\mu} - \mu_0}{\sqrt{\mu_0 / n}}$ where a standard normal distribution approximation is used
- Score confidence interval for μ: $\left(\hat{\mu} + \frac{Z_{1-\alpha/2}^2}{2n}\right) \pm Z_{1-\alpha/2}\sqrt{\frac{\hat{\mu} + Z_{1-\alpha/2}^2 / 4n}{n}}$
- Wald interval for μ using $\log(\mu)$ transformation: $\exp\left(\log(\hat{\mu}) \pm Z_{1-\alpha/2}\sqrt{1/(\hat{\mu}n)}\right)$
- Poisson regression model: $\log(\mu) = \beta_0 + \beta_1 x_1 + \cdots \beta_p x_p$
- $\mu(x + c) / \mu(x) = \exp(c\beta_1)$ for the model $\log(\mu) = \beta_0 + \beta_1 x_1$
- PC = $100\left[\exp(c\beta_1) - 1\right]\%$
- Poisson rate regression model: $\mu = t\exp(\beta_0 + \beta_1 x)$

Chapter 5

- IC(k) = -2log(Likelihood function evaluated at parameter estimates) + kr
- $AIC = IC(2)$
- $AIC_c = IC(2n / (n - r - 1)) = AIC + \frac{2r(r+1)}{n - r - 1}$
- $BIC = IC(\log(n))$
- $\Delta_m = BIC_m - BIC_0$
- $\hat{\tau}_A \approx \frac{\exp(-0.5\Delta_A)}{\sum_{m=1}^{M} \exp(-0.5\Delta_m)}$
- Model averaged estimate: $\hat{\theta}_{MA} = \sum_{m=1}^{M} \hat{\tau}_m \hat{\theta}_m$
- $\text{Var}(\hat{\theta}_{MA}) = \sum_{m=1}^{M} \hat{\tau}_m [(\hat{\theta}_m - \hat{\theta}_{MA})^2 + \text{Var}(\hat{\theta}_m)]$

- Pearson residual: $e_m = \dfrac{y_m - \hat{y}_m}{\sqrt{\text{Var}(Y_m)}}$
- Standardized Pearson residual: $r_m = \dfrac{y_m - \hat{y}_m}{\sqrt{\text{Var}(Y_m - \hat{Y}_m)}} = \dfrac{y_m - \hat{y}_m}{\sqrt{\text{Var}(Y_m)(1-h_m)}}$
- Deviance residual: $e_m^D = \text{sign}(y_m - \hat{y}_m)\sqrt{d_m}$
- Standardized deviance residual: $r_m^D = e_m^D / \sqrt{(1-h_m)}$
- Pearson statistic: $X^2 = \sum_{m=1}^{M} e_m^2 = \sum_{m=1}^{M} \dfrac{(y_m - \hat{y}_m)^2}{\text{Var}(y_m)}$
- $D/(M-\tilde{p})$
- Cook's distance: $CD_m = \dfrac{r_m^2 h_m}{(p+1)(1-h_m)^2}$
- $\Delta X_m^2 = r_m^2$
- $\Delta D_m = (e_m^D)^2 + h_m r_m^2$
- Var(Y) = $\gamma\mu$
- $\dfrac{1}{\gamma}\sum_{m=1}^{M}(Y_m - \mu_m)x_{mr}$
- $r_m = \dfrac{y_m - \hat{y}_m}{\sqrt{\text{Var}(Y_m)(1-h_m)\gamma}}$
- $\hat{\gamma}$ = $X^2/(M-\tilde{p})$
- Negative binomial PMF: $\dbinom{y+k-1}{y}\left(\dfrac{k}{\mu+k}\right)^k\left(1-\dfrac{k}{\mu+k}\right)^y$ where E(Y) = μ and Var(Y) = $\mu + \mu^2/k$

Chapter 6

- Hypergeometric distribution: $P(M=m) = \dbinom{a}{m}\dbinom{b}{k-m}\Big/\dbinom{n}{k}$
- Multiple hypergeometric distribution: $\dfrac{\left(\prod_{i=1}^{I} n_{i+}!\right)\left(\prod_{j=1}^{J} n_{+j}!\right)}{n!\left(\prod_{i=1}^{I}\prod_{j=1}^{J} n_{ij}!\right)}$
- P-value for permutation test involving X^2: (Number of $X^{2*} \geq X^2$) / B
- Poisson regression model with random effect: $\log(\mu_{ik}) = \beta_0 + \beta_1 x_i + b_{0i}$
- $L(\beta_0, \beta_1, \sigma_{b0} \mid y_{11}, \ldots, y_{nt}) = \prod_{i=1}^{n}\int_{-\infty}^{\infty}\prod_{k=1}^{t}\dfrac{\exp(-\mu_{ik})\mu_{ik}^{y_{ik}}}{y_{ik}!}\dfrac{1}{\sigma_{b0}\sqrt{2\pi}}\exp\left(-b_{0i}^2/(2\sigma_{b0}^2)\right)db_{0i}$
- Predicted value of b_{0i}: $E(B_{0i} \mid y_{i1}, \ldots, y_{it})$ or mode of $f(b_{0i} \mid y_{i1}, \ldots, y_{it}) = \dfrac{f(y_{i1}, \ldots, y_{it} \mid b_{0i}) \times g(b_{0i})}{f(y_{i1}, \ldots, y_{it})}$; replace parameters with estimates
- P-value for bootstrap test: $\dfrac{1}{B}\sum_{b=1}^{B} I(W_b^* \geq W)$