Week 1

Chapter 1.2

- **Binomial distribution:**

$$p(y) = \binom{n}{y}\pi^y(1-\pi)^{n-y}, \quad y = 0, 1, 2, \dots, n$$

$$\mu = E(Y) = n\pi \quad \sigma^2 = \mathrm{var}(Y) = n\pi(1-\pi)$$

- **Multinomial distribution:**

$$p(n_1, n_2, \dots, n_{c-1}) = \frac{n!}{n_1! n_2! \dots n_c!}\pi_1^{n_1}\pi_2^{n_2}\dots\pi_c^{n_c},$$

$$\sum_j n_j = n, \quad \text{where} \quad n_c = n - (n_1 + \dots + n_{c-1})$$

$$E(n_j) = n\pi_j, \quad \mathrm{var}(n_j) = n\pi_j(1-\pi_j), \quad \mathrm{cov}(n_j, n_k) = -n\pi_j\pi_k$$

- **Poisson distribution:**

$$p(y) = \frac{e^{-\mu}\mu^y}{y!}, \quad y = 0, 1, 2, \dots$$

$$E(Y) = \mathrm{var}(Y) = \mu$$

- **Poisson multinomial distribution:**

$$P\left(Y_1 = n_1, Y_2 = n_2, \dots, Y_c = n_c \mid \sum_j Y_j = n\right) = \frac{\prod_i [\exp(-\mu_i)\mu_i^{n_i}/n_i!]}{\exp\left(-\sum_j \mu_j\right)\left(\sum_j \mu_j\right)^n/n!}$$
$$= \frac{n!}{\prod_i n_i!}\prod_i \pi_i^{n_i}$$

Chapter 1.3

- **Log-likelihood function:** $L(\theta)$
- **Score function:** $L'(\theta)$
- **Fisher information function:** $-\mathbb{E}(L''(\theta))$
- **Wald test:**

$$z_w = \frac{\hat{\beta} - \beta_0}{SE}, \quad \text{where} \quad SE = \frac{1}{\sqrt{-\mathbb{E}(L''(\theta))}}$$

- **Likelihood ratio test:**

$$z_l = -2\log\Lambda = -2\log\left(\frac{\ell_0}{\ell_1}\right) = -2(L_0 - L_1)$$

- **Score test:**

$$z_s = \frac{[u(\beta_0)]^2}{\iota(\beta_0)} = \frac{\left[\frac{\partial L(\beta)}{\partial \beta_0}\right]^2}{-E\left[\frac{\partial^2 L(\beta)}{\partial \beta_0^2}\right]}$$

Chapter 1.4

- **Binomial Wald:**

$$Z_w = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}}$$

- **Binomial Score:**

$$z_S = \frac{u(\pi_0)}{\sqrt{I(\pi_0)}} = \frac{y - n\pi_0}{\sqrt{n\pi_0(1-\pi_0)}} = \frac{\hat{\pi} - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$$

- **Binomial Likelihood ratio:**

$$-2(L_0 - L_1) = 2\left[y \log \frac{y}{n\pi_0} + (n-y)\log \frac{n-y}{n-n\pi_0}\right]$$

- **Binomial Wald CI:**

$$\hat{\pi} \pm z_{\alpha/2}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$$

Chapter 1.5

- **MLE Multinomial:**

$$\hat{\pi}_j = \frac{n_j}{n}$$

- **Pearson Chi-Squared for multinomial:**

$$X^2 = \sum_j \frac{(n_j - \mu_j)^2}{\mu_j} \quad \text{where} \quad \mu_j = n\pi_{j0} \quad \text{df} = c-1$$

- **Likelihood-Ratio Chi-Squared for multinomial:**

$$G^2 = -2\log\Lambda = 2\sum_j n_j \log\left(\frac{n_j}{n\pi_{j0}}\right) \quad \text{df} = c-1$$

Week 2

Chapter 2.1

- **Independence:** $\pi_{ij} = \pi_{i+}\pi_{+j}$ for $i = 1, \ldots, I \quad j = 1 \ldots, J$
- **Independent multinomial sampling:**

$$\prod_i \left[\frac{n_i!}{\prod_j n_{ij}!} \prod_j \pi_{j|i}^{n_{ij}} \right]$$

Chapter 2.2

- **Relative Risk:**

$$\text{rr} = \frac{\pi_1}{\pi_2} \quad \text{where} \quad \pi_i = \pi_{1|i}$$

  Note: Not meaningful in case control studies where marginal is fixed for X (row)

- **Odds Ratio:**

$$\hat{\theta} = \frac{\pi_1(1-\pi_1)}{\pi_2(1-\pi_2)} = \frac{\pi_{11}/\pi_{12}}{\pi_{21}/\pi_{22}} = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}$$

- **OR and RR:**

$$\text{Odds Ratio} = \text{relative risk} \left(\frac{1-\pi_2}{1-\pi_1} \right)$$

  Note: Similar whenever π_i of the outcome of interest is close to zero for both groups.

- **Risk Difference:**

$$\text{Relative Risk} = \pi_2 - \pi_1$$

- **Odds ratio are invariant under row/column switching:**

$$\begin{aligned} \theta = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} &= \frac{P(Y=1 \mid X=1)/P(Y=2 \mid X=1)}{P(Y=1 \mid X=2)/P(Y=2 \mid X=2)} \\ &= \frac{P(X=1 \mid Y=1)/P(X=2 \mid Y=1)}{P(X=1 \mid Y=2)/P(X=2 \mid Y=2)}. \end{aligned}$$

Chapter 2.3

For a $2 \times 2 \times K$ table

- **Conditional odds ratio:**

$$\theta_{XY(k)} = \frac{\mu_{11k}\mu_{22k}}{\mu_{12k}\mu_{21k}} \quad \text{for a fixed category k of variable Z}$$

- **Marginal odds ratio:**

$$\theta_{XY} = \frac{\mu_{11+}\mu_{22+}}{\mu_{12+}\mu_{21+}}$$

- **Conditional independence:**

$$\pi_{ijk} = \frac{\pi_{i+k} \cdot \pi_{+jk}}{\pi_{++k}} \quad \text{for all } i, j, \text{ and } k$$

  Note: Need to check each cell

- **Marginal independence:**

$$\pi_{ij+} = \pi_{i++}\pi_{+j+}$$

Chapter 2.4

For a $I \times J$ table, we can compute

- **Local odds ratios:**

$$\theta_{ij} = \frac{\pi_{ij}\pi_{i+1,j+1}}{\pi_{i,j+1}\pi_{i+1,j}}, \quad i = 1, \ldots, I-1, \quad j = 1, \ldots, J-1$$

- **Another way to construct odds ratio:**

$$\alpha_{ij} = \frac{\pi_{ij}\pi_{IJ}}{\pi_{Ij}\pi_{iJ}}, \quad i = 1, \ldots, I-1, \quad j = 1, \ldots, J-1$$

- **Proportional Reduction in Variation:**

$$\frac{V(Y) - E[V(Y|X)]}{V(Y)}$$

- **Uncertainty Coefficiente:**

$$U = -\frac{\sum_i \sum_j \pi_{ij} \log(\pi_{ij}/\pi_{i+}\pi_{+j})}{\sum_j \pi_{+j} \log \pi_{+j}}$$

Week 3

Chapter 3.1

- **Sample odds ratio:**
$$\hat{\theta} = \frac{n_{11}n_{22}}{n_{12}n_{22}}$$
- **SE:**
$$\hat{\sigma}(\log \hat{\theta}) = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$$
- **Wald CI:**
$$\log \hat{\theta} \pm z_{\alpha/2}\hat{\sigma}(\log \hat{\theta})$$

- **Sample proportion**
$$\hat{\pi}_i = \frac{y_i}{n_i}$$
- **Expectation and SE for difference of sample proportion:**
$$E(\hat{\pi}_1 - \hat{\pi}_2) = \pi_1 - \pi_2 \quad \text{and} \quad \sigma(\hat{\pi}_1 - \hat{\pi}_2) = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}$$
- **Wald CI:** $(\hat{\pi}_1 - \hat{\pi}_2) \pm z_{\alpha/2}\hat{\sigma}(\hat{\pi}_1 - \hat{\pi}_2)$
- **Sample relative risk:** $r = \dfrac{\hat{\pi}_1}{\hat{\pi}_2} = \dfrac{\left(\dfrac{y_1}{n_1}\right)}{\left(\dfrac{y_2}{n_2}\right)}$
- **SE:** $\hat{\sigma}(\log r) = \sqrt{\dfrac{1-\hat{\pi}_1}{y_1} + \dfrac{1-\hat{\pi}_2}{y_2}}$
- **Wald CI:** $\log r \pm z_{\alpha/2}\hat{\sigma}(\log r)$
- **Delta Method:** Let T_n be a sequence of estimators for a parameter θ such that:
$$\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, \sigma^2)$$
as $n \to \infty$, where $\xrightarrow{d}$ denotes convergence in distribution. Let $g(\cdot)$ be a function that is differentiable at θ, and let $g'(\theta)$ be the derivative of $g(\cdot)$ evaluated at θ. Then, the distribution of $g(T_n)$ can be approximated for large n by:
$$\sqrt{n}(g(T_n) - g(\theta)) \xrightarrow{d} N(0, (g'(\theta))^2\sigma^2),$$
or equivalently,
$$g(T_n) \xrightarrow{d} N\left(g(\theta), \frac{(g'(\theta))^2\sigma^2}{n}\right).$$

Chapter 3.2

- **Testing for independence of Two-way tables:**

$$H_0 : \pi_{ij} = \pi_{i+}\pi_{+j} \forall i, j$$

- **Pearson chi-squared test for contengency table:**

$$X^2 = \sum_i \sum_j \frac{(n_{ij} - \hat{\mu}_{ij})^2}{\hat{\mu}_{ij}} \quad \text{where} \quad \hat{\mu}_{ij} = n\hat{\pi}_{i+}\hat{\pi}_{+j} = \frac{n_{i+} \cdot n_{+j}}{n} \quad \text{with df} = (I-1)(J-1)$$

- **Likelihood-ratio test for multinomial sampling:**

$$G^2 = -2 \log \Lambda = 2 \sum_i \sum_j n_{ij} \log \left(\frac{n_{ij}}{\hat{\mu}_{ij}} \right) \text{with df} = (I-1)(J-1)$$

- **Expected value and variance of cell (2x2):**

$$E(n_{11}) = \frac{n_{1+}n_{+1}}{n} \quad \text{and} \quad \text{var}(n_{11}) = \frac{n_{1+}n_{+1}n_{2+}n_{+2}}{n^2(n-1)}$$

- **Testing independence in 2x2:**

$$X^2 = \frac{n(n_{11}n_{22} - n_{12}n_{21})^2}{n_{1+}n_{2+}n_{+1}n_{+2}} \quad \textbf{with df =}(I-1)(J-1)$$

or

$$z = \frac{\hat{\pi}_1 - \hat{\pi}_2}{\sqrt{\hat{\pi}(1-\hat{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \textbf{where} \quad \hat{\pi} = \frac{y_1 + y_2}{n_1 + n_2}$$

Chapter 3.3

- **Pearson Residual:**

$$H_0 : \text{independence} \quad e_{ij} = \frac{n_{ij} - \hat{\mu}_{ij}}{\sqrt{\hat{\mu}_{ij}}}$$

- **Standardized Pearson residual:**

$$H_0 : \text{independence} \quad r_{ij} = \frac{n_{ij} - \hat{\mu}_{ij}}{\sqrt{\hat{\mu}_{ij}(1 - p_{i+})(1 - p_{+j})}}$$

- **Sample association factor:**

$$a_{ij} = \frac{p_{ij}}{p_{i+}p_{+j}}$$

Week 4

Chapter 4.1

- **Natural exponential family:**

$$f(y_i;\theta_i) = a(\theta_i)b(y_i)\exp\left[y_i Q(\theta_i)\right]$$

- **Link function:**

$$\eta_i = g(\mu_i) = \sum_j \beta_j x_{ij}, \quad i = 1, \ldots, N$$

- **PMF Bernoulli:**

$$f(y;\pi) = \pi^y(1-\pi)^{1-y} = (1-\pi)\left[\frac{\pi}{1-\pi}\right]^y = (1-\pi)\exp\left[y\left(\log\frac{\pi}{1-\pi}\right)\right], \quad y = 0, 1$$

- **PMF Poisson:**

$$\log \mu_i = \sum_j \beta_j x_{ij}, \quad i = 1, \ldots, N.$$

$$f(y;\mu) = \frac{e^{-\mu}\mu^y}{y!} = \exp(-\mu)\left(\frac{1}{y!}\right)\exp\left[y(\log\mu)\right], \quad y = 0, 1, 2, \ldots$$

- **Deviance:** $-2[L(\hat{\mu};y) - L(y;y)]$

Chapter 4.2

- **Linear probability model:** $\pi(\mathbf{x}) = \alpha + \beta_1 x_1 + \cdots + \beta_p x_p$
- **Odds with multiple predictors and the logit functions:**

$$\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})} = \exp\left(\alpha + \beta_1 x_1 + \cdots + \beta_p x_p\right)$$

$$\log\left(\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})}\right) = \alpha + \beta_1 x_1 + \cdots + \beta_p x_p$$

- **Probit model:** $\Phi^{-1}[\pi(x)] = \alpha + \beta x$ where $\Phi()$ is the standard cdf of a continuous rv

Chapter 4.3

- **Poisson loglinear models:**

$$\mu(\mathbf{x}) = \exp(\alpha + \beta_1 x_1 + \cdots + \beta_p x_p) = e^{\alpha}(e^{\beta_1})^{x_1} \cdots (e^{\beta_p})^{x_p}$$

$$\log \mu(\mathbf{x}) = \alpha + \beta_1 x_1 + \cdots + \beta_p x_p$$

- **Negative binomial:**

$$f(y;k,\mu) = \frac{\Gamma(y+k)}{\Gamma(k)\Gamma(y+1)} \left(\frac{k}{\mu+k}\right)^k \left(\frac{\mu}{\mu+k}\right)^y, \quad y = 0, 1, 2, \ldots$$

$$E(Y) = \mu, \quad \text{var}(Y) = \mu + \gamma\mu^2.$$

- **Poisson with Offset:**

$$\log\left(\frac{\mu_i}{t_i}\right) = \alpha + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}$$

Chapter 4.4

- **Exponential dispersion family:**

$$f(y_i;\theta_i,\phi) = \exp\left(\frac{[y_i\theta_i - b(\theta_i)]}{a(\phi)} + c(y_i,\phi)\right)$$

- **Mean and variance for random component:**

$$\mu_i = E(Y_i) = b'(\theta_i)$$

$$\text{var}(Y_i) = b''(\theta_i)a(\phi)$$