7. Continuous covariates may also be used
   a. As with simpler regression models, one should consider the proper scale for continuous covariates
   b. Consider adding polynomial terms
   c. Estimates could be seriously impacted by other variables in model
      i. Mild effect of collinearity
      ii. Impact can be minimized by subtracting out mean.

St: 11.2

8. You can use another function instead of logit
   a. Must still map \( \mathbb{R} \) into \([0,1]\)
   b. Logit has some mathematical properties we will discuss later
   c. Normal CDF is sometimes used
      i. Called the probit
      ii. Results from discretizing standard multiple regression.
         a. Suppose \( Y_j = x_j \beta + \sigma \epsilon_j \), \( \epsilon_j \sim N(0,1) \)
         b. \( Z_j = \begin{cases} 1 & \text{if } Y_j > c \\ 0 & \text{otherwise} \end{cases} \)
         c. \( P[Z_j = 1] = P[Y_j > c] = P[\epsilon_j > (c - x_j \beta)/\sigma] = 1 - \Phi((c - x_j \beta)/\sigma) = \Phi((x_j \beta - c)/\sigma) \)
         d. After rescaling, probit and logit are very close. See Figure 6.
   Se: 9 pp. 298–310

G. Regression Models for case–control studies
   1. We want to model \( P(Y_j = y_j | x_j) \)
      a. Suppose there are \( j \) strata
      iii. sum is over all rearrangements \( w_j \) of the \( x_j \)
   3. Then \( P[X_j = x_j \forall j | Y_j = y_j \forall j, \{X_j\}] = \prod_j \frac{\exp(x_j \beta y_j)}{1 + \exp(x_j \beta)} / \prod_j \frac{\exp(x_j \beta)}{1 + \exp(x_j \beta)} = \frac{\prod_j \exp(x_j \beta y_j)}{\prod_j \exp(x_j \beta)} \)

B&D1: 7.0, 7.2

4. In a stratified study
   a. Suppose there are \( K \) strata
   b. \( P[Y_{jk} = 1] = \exp(x_{jk} \beta + \alpha_k)/(1 + \exp(x_{jk} \beta + \alpha_k)) \)
   c. \( \prod_{jk} P[Y_{jk} = y_{jk} \forall j, k | X_{jk} = x_{jk}] = \prod_{jk} \exp[(x_{jk} \beta + \alpha_k)y_{jk}]/(1 + \exp(x_{jk} \beta + \alpha_k)) \)
   d. Condition on the number of cases and control in each strata, and on the collection of covariate patterns \( \{x_{jk}\} \):

\[
\prod_{jk} P[Y_{jk} = y_{jk} \forall j, k | \{X_{jk}\} = \{x_{jk}\}, \sum_{j} Y_{jk} = \sum_{j} y_{jk} \forall j] = \sum_{y_{jk}} \exp[(x_{jk} \beta + \alpha_k)y_{jk}]/(1 + \exp(x_{jk} \beta + \alpha_k))
\]

B&D1: 7.1, 7.5

f. Suppose we fit a logistic regression model to the stratified tables
   i. \( O_{jk} \sim \text{Bin}(\pi_{jk}, \gamma_{1+}) \)
   ii. \( \logit(\pi_{jk}) = \alpha_k + \gamma_{jk} \text{ with } \gamma_1 = 0 \)
      a. Probability only depends on data through \( \sum_{jk} x_{jk} \gamma_{jk} \)
      b. Best fit will make \( \sum_{jk} x_{jk} \gamma_{jk} \) exactly
         i. Both exposed: \( \hat{\gamma}_{1+} = \hat{\gamma}_{2+} \)
         ii. Both unexposed: \( \hat{\gamma}_{1-} = \hat{\gamma}_{2-} \)
   iv. discordant pairs:
      a. \( P[\text{diseased}|\text{exposed}] + P[\text{diseased}|\text{unexposed}] = \pi_{jk} = 1/\text{stratum} \)
      b. Hence \( \exp(\hat{\alpha}_j)/(1 + \exp(\hat{\alpha}_j)) + \exp(\hat{\gamma}_{2+} + \hat{\gamma}_{2-} + \hat{\gamma}_{1+})/(1 + \exp(\hat{\gamma}_{2+} + \hat{\alpha}_j)) = 1 \)
Lecture 9

D. One-way analyses

C. Approximately,

1. External standards based on SMR

- Sum of fitted values among exposed

   \( (n_{10} + n_{01}) \exp(\gamma_2/2) / (1 + \exp(\gamma_2/2)) = n_{10} \)

- Hence \( \gamma_2 = 2 \log(n_{10} / n_{01}) \)

B&D1: 7.6

g. Suppose we ignore stratification:

i. Recall \( \psi = \frac{\sum_i (1 - \pi_i)}{\sum_i \pi_i} \)

ii. Use \( \hat{\psi}^* = \frac{\sum_i O_{2i} \pi_i}{\sum_i O_{1i} \pi_i} \)

iii. Estimate

\[
\psi^* = \frac{\sum_i (1 - \pi_i) \pi_i}{\sum_i \pi_i} = \frac{\sum_i (1 - \pi_i) \pi_i}{\sum_i (1 - \pi_i) \pi_i}
\]

for \( \hat{\psi} = \pi_{11} / (1 - \pi_{11}) \).

vi. Then the power is

\[
\frac{1 - \pi_{11}}{\sum_j (1 - \pi_{2j})} - \frac{1 - \pi_{11}}{\sum_j (1 - \pi_{1j})}
\]

Note: 

- Suppose \( H_0 : T \sim N(\mu_0, \sigma_0^2) \)
- \( H_A : T \sim N(\mu_A, \sigma_0^2) \)

2. Critical value: \( C \)

a. Reject \( H_0 \) if \( (T - \mu_0) / \sigma_0 \geq \alpha \)

b. \( 1 - \Phi(z_\alpha) = \alpha \)

c. Reject \( H_0 \) if \( T \geq \mu_0 + z_\alpha \sigma_0 \)

3. Power is \( P_A [T \geq C] = \Phi((\mu_A - \mu_0 - \sigma_0 z_\alpha) / \sigma_0) \)

a. Special Case: \( \sigma_0 = \sigma_A \), power is \( \Phi((\mu_A - \mu_0) / \sigma_0 - z_\alpha) \)

Sample size:

a. Assume that \( \sigma_0 = \sigma_A = \tau_A / \sqrt{n} \)

b. Require power \( 1 - \beta \)

i. Typically, \( \beta = 0.8 \)

ii. Then \( -z_\beta = (\mu_0 + \sigma_0 z_\alpha - \mu_A) / \sigma_A \)

iii. \( (\tau_A z_\beta + \sigma_0 z_\alpha - \mu_A) / \sqrt{n} = \mu_A - \mu_0 \)

iv. \( (\tau_A z_\beta + \sigma_0 z_\alpha)^2 / (\mu_A - \mu_0)^2 = n \)

v. When \( \tau_A = \tau_0 \), \( n = \tau_0^2 (z_\beta + z_\alpha)^2 / (\mu_A - \mu_0)^2 = n \)

Se: 3 pp. 87–90

D. One-way analyses

1. External standards based on SMR

a. We could in principal perform calculations exactly

i. This is typically not considered necessary

b. Assume \( O \sim P(e) \)

ii. Then the power is

\[
\frac{1 - \pi_{11}}{\sum_j (1 - \pi_{2j})} - \frac{1 - \pi_{11}}{\sum_j (1 - \pi_{1j})}
\]

i. Big values are associated with big \( \theta_i \)

H. Summary:

<table>
<thead>
<tr>
<th>Number of Covar-</th>
<th>Logist.</th>
<th>Cond.</th>
<th>Mantel –</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strata</td>
<td>Regr.</td>
<td>Logist.</td>
<td>Regr.</td>
</tr>
<tr>
<td>1</td>
<td>Large</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>2</td>
<td>Small</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>3</td>
<td>Large</td>
<td>yes</td>
<td>√</td>
</tr>
<tr>
<td>4</td>
<td>Small</td>
<td>yes</td>
<td>√</td>
</tr>
</tbody>
</table>

Case–Control:

- Any | no | √ | √ |
- Any | yes | √ | √ |

- Good
- √ Test OK; estimator suboptimal except for paired data
- Inappropriate
- Slow

1. Mantel–Haentzel test and estimator reduces to McNemar’s test and estimator when strata are pairs.

Se: 3 pp. 83–86

IV. Sample Size Calculations

A. Preliminaries

1. We’ll do power for 1-sided tests

a. Conceptually easier (as we shall see)

b. Get power for 2-sided tests by doubling \( \alpha \)

B. Exactly:

1. Select smallest \( C \) such that \( P_0 [T \geq C] \leq \alpha \)

2. Power is \( P_A [T \geq C] \)

C. Approximately

1. Suppose

- \( H_0 \)

2. Critical value: \( C \)

3. Power is \( P_A [T \geq C] = \Phi((\mu_A - \mu_0 - \sigma_0 z_\alpha) / \sigma_0) \)

a. Special Case: \( \sigma_0 = \sigma_A \), power is \( \Phi((\mu_A - \mu_0) / \sigma_0 - z_\alpha) \)

Sample size:

a. Assume that \( \sigma_0 = \sigma_A = \tau_A / \sqrt{n} \)

b. Require power \( 1 - \beta \)

i. Typically, \( \beta = 0.8 \)

2. Dichotomous exposures

a. Without age stratification

b. \( H_0 : \text{SMR}=1 \) vs. \( H_A : \text{SMR}=\varsigma \)
c. Inference is based on $O_1|O_+ \sim \text{Bin}(\pi_1, O_+)$

d. $T = O_1/O_+.$

e. $\mu_0 = Q_1/(Q_0 + Q_1), \sigma_0 = \sqrt{\mu_0(1 - \mu_0)/O_+}.$
f. $\mu_A = \varsigma Q_1/(Q_0 + \varsigma Q_1) = \frac{\pi_1^\varsigma}{1 + (\varsigma - 1)\pi_1^\varsigma}$ and $\sigma_A = \sqrt{\mu_A(1 - \mu_A)/O_+}.$
g. No serious simplification to power and sample size formulae

h. Example:
i. For shipping example, suppose that we want to test $H_0: \text{two ship classes have same accident rate}$

ii. Want 80% power to detect difference if one class has 50% more accidents

iii. Suppose that both kinds of ships have equal months at risk

iv. Need $O_+ = (\frac{z_{.8} - z_{.025}}{1.645})^2 = 194$ accidents.

v. Median rate was .002 accidents per month

vi. Hence need to follow ships for $194/0.002 = 95000$ total months.

i. Power is approximate

ii. Better approximation for Poisson uses fact that when $O \sim \mathcal{P}(\mu)$ then $\text{Var}(\sqrt{O}) \approx \mu \times (\frac{1}{2}\mu^{-1/2})^2 = \frac{1}{4}.$

iii. Better approximation for binomial uses fact that $\text{arcsin}(\sqrt{O_1/O_+}) \sim N(\text{arcsin}(\sqrt{\frac{Q_1}{Q_0 + Q_1}}), 1/(4O_+))$

- $\frac{d}{dx} \text{arcsin}(x) = 1/\sqrt{1 - x^2}$
- $\frac{d}{dx} \text{arcsin}(\sqrt{p}) = 1/(2\sqrt{p}(1 - x))$ See Figures 7 and 8.

iii. $\mu_0 = \text{arcsin}(\sqrt{Q_1/(Q_0 + Q_1)}),$ $\mu_A = \text{arcsin}(\varsigma \sqrt{Q_1/(Q_0 + \varsigma Q_1)}),$ $\sigma_A = \sigma_0 = \sqrt{1/(4O_+)}$

iv. Power and sample size as before.

v. Shipping example:

$$\frac{(z_{.8} - z_{.025})^2}{\left(\text{arcsin}\left(\frac{\sqrt{.5}}{1.5}\right) - \text{arcsin}\left(\frac{\sqrt{.5}}{1.5}\right)\right)^2} = 194$$

j. Can also calculate minimal sample size to give a desired CI width.