

Problem 1

$$Y \sim \text{Bernoulli}(\mu)$$

$$p(y|\mu) = \mu^y(1 - \mu^y) \quad y \in \{0, 1\} \quad 0 < \mu < 1$$

$$P(Y = 1) = \mu = 1 - P(Y = 0)$$

$$\mu^y(1 - \mu^y) = \left(\frac{\mu}{1-\mu}\right)^y (1 - \mu)$$

Getting the distribution in the form of an exponential family:

$$p(y|\theta) = h(y) \exp\{\theta T(y) - k(\theta)\}$$

$$\theta = \log\left\{\frac{\mu}{1-\mu}\right\}$$

$$e^\theta = \frac{\mu}{1-\mu} \Rightarrow \mu = \frac{1}{1+e^{-\theta}}$$

$$p(y|\mu) = \exp\left\{y \log\left(\frac{\mu}{1-\mu}\right) + \log(1 - \mu)\right\} \text{ where the range of } y \text{ does not depend on } \mu$$

Thus the distribution belong to the exponential family of distributions.

$$p(y|\theta) = \exp\{\theta y - \log(1 + e^{-\theta})\}$$

Where θ is the canonical parameter, y is the canonical statistic, and

$k(\theta) = \log(1 + e^{-\theta})$ is the cumulant function.

canonical link function: $\log(\frac{\mu}{1-\mu}) = \theta$

$$\text{Var}(y) = k''(\theta) = \frac{e^{-\theta}}{(1+e^{-\theta})^2}$$

$$\frac{e^{-\theta}}{(1+e^{-\theta})^2} = \frac{(1-\mu)/\mu}{[1+(1-\mu)/\mu]^2} = \mu(1 - \mu)$$

$$V(\mu) = \mu(1 - \mu)$$

$$\text{Deviance: } D(y, \hat{\mu}) = -2[y \log(\mu) + (1 - y) \log(1 - \mu)]$$

Problem 2

$$\text{Var}(Y_i) = 1 \quad \text{cov}(Y_i, Y_j) = \rho, i \neq j \quad -1/(n-1) < \rho < 1$$

$$\text{covariance matrix } \Sigma = aI + bJ$$

$$\text{Since } \text{Var}(Y_i) = 1, \quad a + b = 1$$

$$\text{Since covariance} = \rho, \quad b = \rho$$

$$\text{Then } a = 1 - \rho$$

$$\text{The weight matrix } W = \Sigma^{-1} = cI + dJ$$

$$\Sigma * \Sigma^{-1} = I \Rightarrow [(1 - \rho)I + \rho J] * [cI + dJ] = I$$

$$= (1 - \rho)cI + (1 - \rho)dJ + \rho cJ + \rho d n J = I$$

$$= [(1 - \rho)d + \rho c + \rho d n]J + (1 - \rho)cI = I$$

$$\text{Then } (1 - \rho)c = 1 \Rightarrow c = 1/(1 - \rho)$$

$$\text{And } (1 - \rho)d + \rho c + \rho d n = 0$$

Plugging in $c = 1/(1 - \rho)$ we have:

$$(1 - \rho)d + \rho[1/(1 - \rho)] + \rho d n = 0$$

Solving for d we have:

$$d = \frac{-\rho}{(1 - \rho)(1 + \rho(n - 1))}$$

The $n \times n$ covariance matrix has the form of:
$$\begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots \\ \rho & \rho & 1 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

The $n \times n$ weight matrix has the form:
$$\begin{bmatrix} c+d & d & \dots & d \\ d & c+d & d & \dots \\ d & d & c+d & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

Since Variance is all one and the covariances are all homogeneous ρ the weights would all be the same thus the OLS estimates and WLS estimates are equivalent. That is we have homoscedasity. OLS seeks to minimize residuals and produce the smallest possible standard errors. OLS gives equal weight to all observations and one of the assumptions of OLS is homoscedasity. With heteroscadesity observations with larger disturbances have more effect than other observations. Weighted least squares seeks to weight down these larger disturbances, hence the weight matrix. However with homoscadestiy, all of the weights will be the same thus OLS and WLS estimates are equivalent

Problem 11.2

Part 7

Adjusting for overdispersion affects the standard errors only, the parameter estimates are the same as in 11.2.5. The standard errors are different.

Using a robust variance estimator to adjust for overdispersion.

$$\hat{\beta}_1 = 2.5046 \quad se(\hat{\beta}_1) = .5974$$

$$\hat{\beta}_2 = -.1797 \quad se(\hat{\beta}_2) = .3317$$

$$\hat{\beta}_3 = -.0148 \quad se(\hat{\beta}_3) = .0209$$

$$\ln(\hat{\mu}_i) = 2.5046 + -.1797 * Treatment_i + -.0148 * age_i$$

Part 8

$\exp(-.1797) = .8355$ is the age-adjusted rate ratio comparing progabide to placebo, taking into account overdispersion

$$95\% \text{ CI: } (\exp(-0.8299), \exp(0.4704)) = (.4361, 1.601)$$