Chapter Five

Ryan Honea

9/27/2017

Exercise One

Question

Consider the single response variable Y with $Y \sim \text{Bin}(n, \pi)$.

Solution

(a): Find the Wald statistic $(\hat{\pi} - \pi)^T \boldsymbol{\mathcal{J}}(\hat{\pi} - \pi)$, where $\hat{\pi}$ is the maximum likelihood estimator of π and $\boldsymbol{\mathcal{J}}$ is the information.

Solution: The most likely estimator is $\frac{Y}{n}$, and as this is a single response variable, we know the Wald statistic is $(\hat{\pi} - \pi)^2 \boldsymbol{\mathcal{J}}$. We also saw from Example 5.22 that $\boldsymbol{\mathcal{J}} = \frac{n}{\pi(1-\pi)}$.

So, we have $W = \left(\frac{Y}{n} - \pi\right)^2 \left(\frac{n}{\pi(1-\pi)}\right)$

(b): Verify that the Wald statistic is the same as the score statistic $\boldsymbol{U}^T \boldsymbol{\mathcal{J}}^{-1} \boldsymbol{U}$ in this case (see Example 5.22).

Solution:

$$
\begin{aligned}
\frac{U^2}{\boldsymbol{\mathcal{J}}} &= \frac{(Y - n\pi)^2}{n\pi(1-\pi)} \\
W &= \left(\frac{Y}{n} - \pi\right)^2 \left(\frac{n}{\pi(1-\pi)}\right) \\
&= \left(\frac{Y^2}{n^2} - \frac{2Y\pi}{n} + \pi^2\right) \left(\frac{n}{\pi(1-\pi)}\right) \\
&= \frac{1}{n^2}\left(Y^2 - 2Yn\pi + \pi^2 n^2\right) \left(\frac{n}{\pi(1-\pi)}\right) \\
&= \frac{1}{n^2}(Y - n\pi)^2 \left(\frac{n}{\pi(1-\pi)}\right) \\
&= \frac{(Y - n\pi)^2}{n\pi(1-\pi)} = \frac{U^2}{\boldsymbol{\mathcal{J}}}
\end{aligned}
$$

(c): Find the deviance

$$2[l(\hat{\pi}; y) - l(\pi; y)].$$

$$
\begin{aligned}
l(\hat{\pi}; Y) &= \log \binom{n}{Y} + Y \log \hat{\pi} + (n - Y)\log(1 - \hat{\pi}) \\
l(\pi; Y) &= \log \binom{n}{Y} + Y \log \pi + (n - Y)\log(1 - \pi) \\
2[l(\hat{\pi}; y) - l(\pi; y)] &= 2\left[Y \log\left(\frac{\hat{\pi}}{\pi}\right) + (n - Y)\log\left(\frac{1 - \hat{\pi}}{1 - \pi}\right)\right]
\end{aligned}
$$

(d): For large samples, both the Wald/score statistic and the deviance approximatley have the $\chi^2(1)$ distribution. For $n = 10$ and $y = 3$, use both statistics to assess the adequacy of the models:

(1) $\pi = 0.1$
(2) $\pi = 0.3$
(3) $\pi = 0.5$

Do the two statistics lead to the same conclusions?

Solution:

I use the following functions to find my results.

```
wald <- function(y, n, p) {
  return (((y - n*p)^2.0)/(n*p*(1.0-p)))
}

dev <- function(y, n, p) {
  p_hat = y/n
  return (2*(y*log(p_hat/p) + (n - y)*log((1-p_hat)/(1-p))))
}
```

$$
\begin{array}{lll}
(1) & W = 4.\bar{4} & D = 3.073 \\
(2) & W = 0 & D = 0 \\
(3) & W = 1.6 & D = 1.65
\end{array}
$$

The 95th percentile of the χ^2 test is 3.84 and so our rejection region is $D > 3.84$ or $W > 3.84$. In the case of (1), W and D disagree, where W leads us to reject our hypothesis that $\pi = 0.1$ while D leads us to fail to reject. For (2) and (3), we reject the null hypothesis for both D and W.

Exercise Two

Question

Consider a random sample $Y_1, ..., Y_N$ with the exponential distributon

$$f(y_i, \theta_i) = \theta_i \exp(-y_i\theta_i)$$

Derive the deviance by comparing the maximal model with different values for θ_i for each Y_i and the model with $\theta_i = \theta$ for all i.

Solution

Note that $y_i = 1/\theta_i$ and $\hat{y}_i = 1/\theta$.

$$\begin{aligned}
l(f(y_i;\theta_i)) &= \sum_{i=1}^{n}(\log\theta_i - y_i\theta_i) \\
l(f(y_i;\theta)) &= \sum_{i=1}^{n}(\log\theta - y_i\theta) \\
2\left[l(f(y_i;\theta_i)) - l(f(y_i;\theta))\right] &= 2\left[\sum_{i=1}^{n}\log\theta_i - y_i\theta_i - \log\theta + y_i\theta\right] \\
&= 2\left[\sum_{i=1}^{n}\log(1/y_i) - 1 - \log(1/\hat{y}_i) + (y_i/\hat{y}_i)\right] \\
&= 2\left[\sum_{i=1}^{n}\log\left(\frac{\hat{y}_i}{y_i}\right) + \frac{y_i}{\hat{y}_i} - 1\right] \\
&= 2\left[\sum_{i=1}^{n}\frac{y_i}{\hat{y}_i} - \log\left(\frac{y_i}{\hat{y}_i}\right) - 1\right]
\end{aligned}$$

Exercise Four

Question

For the leukemia survival data in 4.2:

Solution

(a): Use the Wald statistic to obtain an approximate 95% confidence interval for the paramter β_1.

(b): By comparing the deviances for two appropriate models, test the null hypothesis $\beta_2 = 0$ against the alternative hypothesis $\beta_2 \neq 0$. What can you conclude about the use of the initial white blood cell count as a predictor of survival time.