# Exercises Chapter 5

## 5.1 :

$Y \sim Bin(n, \pi)$

$E(Y) = n\pi$

$Var(Y) = n\pi(1-\pi)$

(a) • $f(y;\pi) = \binom{n}{y}\pi^{y}(1-\pi)^{n-y}$ (for one observation)

$$\ell(\pi) = y\ln(\pi) + (n-y)\ln(1-\pi) + \ln\left[\binom{n}{y}\right]$$

$$U = U(\pi) = \frac{\partial \ell}{\partial \pi} = \frac{y}{\pi} - \frac{n-y}{1-\pi} = \frac{y(1-\pi)-(n-y)\pi}{\pi(1-\pi)}$$

$$= \frac{y - y\pi - n\pi + y\pi}{\pi(1-\pi)} = \frac{y-n\pi}{\pi(1-\pi)}$$

$$\boxed{U = \frac{y-n\pi}{\pi(1-\pi)}}$$

• The MLE of $\pi$ is $\boxed{\hat{\pi} = \frac{y}{n}}$

Note that:

$$E(U) = E\left[\frac{y-n\pi}{\pi(1-\pi)}\right] = \frac{E(Y)-n\pi}{\pi(1-\pi)} = \frac{n\pi - n\pi}{\pi(1-\pi)} = 0$$

• Information:

$$\mathcal{J} = Var(U) = Var\left[\frac{Y-n\pi}{\pi(1-\pi)}\right]$$

$$= \frac{Var(Y)}{\pi^2(1-\pi)^2} = \frac{n\pi(1-\pi)}{\pi^2(1-\pi)^2} = \frac{n}{\pi(1-\pi)}$$

$$\boxed{\mathcal{J} = \frac{n}{\pi(1-\pi)}}$$

- Now:

$$(\hat{\pi}-\pi)^t J (\hat{\pi}-\pi) = \left(\frac{y}{n}-\pi\right)^2 \left[\frac{n}{\pi(1-\pi)}\right]$$

$$= \left(\frac{y-n\pi}{n}\right)^2 \left(\frac{n}{\pi(1-\pi)}\right)$$

$$= \frac{(y-n\pi)^2}{n\pi(1-\pi)}$$

$$= \left[\frac{y-n\pi}{\sqrt{n\pi(1-\pi)}}\right]^2$$

Since $Y \sim Bin(n,\pi)$, we have:

$$Z = \frac{y-E(Y)}{\sqrt{Var(Y)}} = \frac{y-n\pi}{\sqrt{n\pi(1-\pi)}} \stackrel{\cdot}{\sim} N(0,1) \text{ (approximately)}$$

(Normal approximation to binomial)

$$Z^2 = \left[\frac{y-n\pi}{\sqrt{n\pi(1-\pi)}}\right]^2 \stackrel{\cdot}{\sim} \chi^2_{(1)} \text{ (approximately)}$$

(b) • Wald statistic is:

$$Z^2 = \left[\frac{y-n\pi}{\sqrt{n\pi(1-\pi)}}\right]^2 = \frac{(y-n\pi)^2}{n\pi(1-\pi)} \text{ ——— } (*)$$

- $$U^t J^{-1} U = \left[\frac{y-n\pi}{\pi(1-\pi)}\right]^2 \frac{1}{\frac{n}{\pi(1-\pi)}} = \frac{(y-n\pi)^2}{\pi^2(1-\pi)^2}\,\frac{\pi(1-\pi)}{n}$$

$$= \frac{(y-n\pi)^2}{n\pi(1-\pi)} \text{ ——— } (**)$$

$(*) = (**)$

$\therefore$ Wald statistic $= U^t J^{-1} U$

3

(c)

- $\ell(\pi) = y\ln(\pi) + (n-y)\ln(1-\pi) + \ln\left[\binom{n}{y}\right]$
- For the saturated model, we have:
  - MLE of $\pi$ is $\hat{\pi} = \frac{y}{n}$
  - The fitted value $\hat{y} = n\hat{\pi} = n\left(\frac{y}{n}\right) = y$
  - The maximum value of $\ell(\pi)$ for the saturated model is:

$$\ell(b_{max};y) = \ell(\hat{\pi}) = y\ln(\hat{\pi}) + (n-y)\ln(1-\hat{\pi}) + \ln\left[\binom{n}{y}\right]$$
$$= y\ln\left(\frac{y}{n}\right) + (n-y)\ln\left(1-\frac{y}{n}\right) + \ln\left[\binom{n}{y}\right]$$

- For the model of interest:
  - Let $\tilde{\pi}$ be the MLE of $\pi$, and $\tilde{y} = n\tilde{\pi}$ be the fitted value.
  - The maximum value of $\ell(\pi)$ for the model of interest is:

$$\ell(b;y) = \ell(\tilde{\pi}) = y\ln(\tilde{\pi}) + (n-y)\ln(1-\tilde{\pi}) + \ln\left[\binom{n}{y}\right]$$
$$= y\ln\left(\frac{\tilde{y}}{n}\right) + (n-y)\ln\left(1-\frac{\tilde{y}}{n}\right) + \ln\left[\binom{n}{y}\right]$$

- The deviance is:

$$D = 2\left[\ell(b_{max}) - \ell(b)\right]$$
$$= 2\left\{\left[y\ln\left(\frac{y}{n}\right) + (n-y)\ln\left(1-\frac{y}{n}\right) + \ln\left[\binom{n}{y}\right]\right] - \left[y\ln\left(\frac{\tilde{y}}{n}\right) + (n-y)\ln\left(1-\frac{\tilde{y}}{n}\right) + \ln\left[\binom{n}{y}\right]\right]\right\}$$
$$= 2\left\{y\ln\left(\frac{y}{\tilde{y}}\right) + (n-y)\ln\left(\frac{1-\frac{y}{n}}{1-\frac{\tilde{y}}{n}}\right)\right]$$
$$= 2\left[y\ln\left(\frac{y}{\tilde{y}}\right) + (n-y)\ln\left(\frac{n-y}{n-\tilde{y}}\right)\right]$$

(d) Wald statistic: $= (\hat{\pi}-\pi)^t J (\hat{\pi}-\pi)$ $\quad \frac{(y-n\pi)^2}{n\pi(1-\pi)} \dot\sim \chi^2_{(1)}$

Score statistic: $u^t J^{-1} u \dot\sim \chi^2_{(1)}$

Deviance: $D = 2\left[y \ln\left(\frac{y}{\tilde{y}}\right) + (n-y)\ln\left(\frac{n-y}{n-\tilde{y}}\right)\right]$

- We have shown that:

$$\text{Wald Statistic} = \text{Score Statistic} = \frac{(y-n\pi)^2}{n\pi(1-\pi)}$$

$n = 10$ , $y = 3$ , $\chi^2_{0.05(1)} = 3.84$

(i) For the model with $\pi = \pi_0 = 0.1$ : $\begin{Bmatrix} H_0: \pi = 0.1 \\ H_1: \pi \neq 0.1 \end{Bmatrix}$

$\tilde{y} = n\pi_0 = (10)(0.1) = 1$

- Wald Statistic $= \frac{(y-n\pi_0)^2}{n\pi_0(1-\pi_0)} = \frac{(3-1)^2}{(10)(0.1)(0.9)}$

$= 4.444$

Since $4.444 > \chi^2_{0.05(1)}$ , we reject $H_0: \pi = 0.1$

- Wald Statistic = Score statistic = 4.444

we will have the same calculations and the same conclusion.

Therefore, the model is not adequate based on these two statistics.

- Deviance

$$D = 2\left[y\ln\left(\frac{y}{n\pi_0}\right) + (n-y)\ln\left(\frac{n-y}{n-n\pi_0}\right)\right]$$

$$= 2\left[(3)\ln\left(\frac{3}{1}\right) + (10-3)\ln\left(\frac{10-3}{10-1}\right)\right]$$

$$= 3.07327$$

Since $3.07327 < \chi^2_{0.05(1)}$ , we do not reject Ho based on the deviance statistic, and we conclude that the model is adequate.

We notice that these statistics do not lead to the same conclusions.

(ii) Exercise for the students.

(iii) 〃 〃 〃 〃

5.2 :

$Y_1, Y_2, \dots, Y_N$ are independent.

$Y_i \sim \text{Exp}(\theta_i)$ for all $i = 1, 2, \dots, N$

$$f(y_i, \theta_i) = \theta_i e^{-\theta_i y_i} \quad \begin{pmatrix} y_i > 0 \\ \theta_i > 0 \end{pmatrix}$$

- For maximal model, the $\theta_i$'s are different.
- For the model of interest, the $\theta_i$'s are the same $\theta_1 = \theta_2 = \dots = \theta_N = \theta$.

(1) For the maximal model:

The log-likelihood function is:

$$l = l(\theta_1, \dots, \theta_N) = \sum_{i=1}^{N} \ln(\theta_i e^{-\theta_i y_i}) = \sum_{i=1}^{N} \ln(\theta_i) - \sum_{i=1}^{N} \theta_i y_i$$

$$\frac{\partial l}{\partial \theta_i} = \frac{\partial l(\theta_1, \dots, \theta_N)}{\partial \theta_i} = \frac{1}{\theta_i} - y_i$$

$$\frac{\partial l}{\partial \theta_i} \overset{\text{set}}{=} 0 \Leftrightarrow \frac{1}{\theta_i} - y_i = 0 \Leftrightarrow \boxed{\hat{\theta}_i = \frac{1}{y_i}}$$

The MLE of $\theta_i$ is $\hat{\theta}_i = \frac{1}{y_i}$ $\left(\text{Note! } \frac{\partial^2 l}{\partial \theta_i^2} = -\frac{1}{\theta_i^2} < 0\right)$

The Maximum value of $l(\theta_1, \dots, \theta_N)$ is:

$$l(b_{max}) = l(\hat{\theta}_1, \dots, \hat{\theta}_N) = \sum_{i=1}^{N} \ln\left(\frac{1}{y_i}\right) - \sum_{i=1}^{N} \frac{1}{y_i}(y_i)$$

$$= \sum_{i=1}^{N} \ln\left(\frac{1}{y_i}\right) - N$$

(2) For the model of interest $(\theta_1 = \theta_2 = \dots = \theta_N = \theta)$:

The log-likelihood function is:

$$l = l(\theta) = \sum_{i=1}^{N} \ln(\theta e^{-\theta y_i}) = N \ln(\theta) - \theta \sum_{i=1}^{N} y_i$$

$$\frac{\partial l}{\partial \theta} = \frac{N}{\theta} - \sum_{i=1}^{N} y_i$$

$$\frac{\partial l}{\partial \theta} \overset{\text{set}}{=} 0 \iff \frac{N}{\theta} - \sum_{i=1}^{N} y_i = 0 \iff \hat{\theta} = \frac{N}{\sum_{i=1}^{N} y_i} = \frac{1}{\bar{y}}$$

The MLE of $\theta$ is $\hat{\theta} = \frac{1}{\bar{y}}$ $\left(\frac{\partial l^2}{\partial \theta^2} = -\frac{N}{\theta^2} < 0\right)$

The maximum value of $l(\theta)$ is

$$l(\underset{\sim}{b}) = l(\hat{\theta}) = N \ln(\hat{\theta}) - \hat{\theta} \sum y_i$$
$$= N \ln\left(\frac{1}{\bar{y}}\right) - \frac{1}{\bar{y}} \sum_{i=1}^{N} y_i = N \ln\left(\frac{1}{\bar{y}}\right) - N \quad \left(\sum y_i = N\bar{y}\right)$$

For (1) and (2):

The deviance is

$$D = 2\left[l(\underset{\sim}{b}_{max}) - l(\underset{\sim}{b})\right]$$
$$= 2\left\{\left[\sum_{i=1}^{N} \ln\left(\frac{1}{y_i}\right) - N\right] - \left[N \ln\left(\frac{1}{\bar{y}}\right) - N\right]\right\}$$
$$= 2\left[\sum_{i=1}^{N} \ln\left(\frac{1}{y_i}\right) - N \ln\left(\frac{1}{\bar{y}}\right)\right]$$
$$= 2\left[\sum_{i=1}^{N} \ln\left(\frac{1}{y_i}\right) - \sum_{i=1}^{N} \ln\left(\frac{1}{\bar{y}}\right)\right]$$
$$= 2\sum_{i=1}^{N}\left[\ln\left(\frac{1}{y_i}\right) - \ln\left(\frac{1}{\bar{y}}\right)\right]$$
$$= 2\sum_{i=1}^{N} \ln\left[\frac{1/y_i}{1/\bar{y}}\right]$$
$$= 2\sum_{i=1}^{N} \ln\left(\frac{\bar{y}}{y_i}\right)$$

5.3 :

$Y_1, Y_2, \ldots, Y_N$ are independent.

$Y_i \sim \text{Pareto}(\theta)$ for all $i = 1, 2, \ldots, N$

$$f(y_i;\theta) = \frac{\theta}{y_i^{\theta+1}} = \theta y_i^{-\theta-1} \quad ; \begin{pmatrix} y_i > 1 \\ \theta > 0 \end{pmatrix}$$

(a) The log-likelihood function is:

$$l = l(\theta) = \sum_{i=1}^{N} \ln(\theta y_i^{-\theta-1}) = N \ln(\theta) - (\theta+1) \sum_{i=1}^{N} \ln(y_i)$$

$$\frac{\partial l}{\partial \theta} = \frac{N}{\theta} - \sum \ln(y_i)$$

$$\frac{\partial l}{\partial \theta} \overset{\text{set}}{=} 0 \iff \frac{N}{\theta} - \sum \ln(y_i) \iff \hat{\theta} = \frac{N}{\sum \ln(y_i)}$$

The MLE of $\theta$ is $\hat{\theta} = \dfrac{N}{\sum_{i=1}^{N} \ln(y_i)}$ $\left( \dfrac{\partial^2 l}{\partial \theta^2} = -\dfrac{N}{\theta^2} < 0 \right)$

(b) Recall From Exercise 3.11 :

- Score Statistic is : $U = \dfrac{N}{\theta} - \sum_{i=1}^{N} \ln(y_i)$
- Information is : $\mathcal{J} = \text{Var}(U) = \dfrac{N}{\theta^2}$

The variance of $\hat{\theta} = \dfrac{N}{\sum \ln(y_i)}$ is :

$$\text{Var}(\hat{\theta}) = \mathcal{J}^{-1} = \frac{\theta^2}{N}$$

$$\hat{\theta} \overset{\cdot}{\sim} N(\theta, \mathcal{J}^{-1}) = N\left(\theta, \frac{\theta^2}{N}\right) \quad \text{(asymptotically)}$$

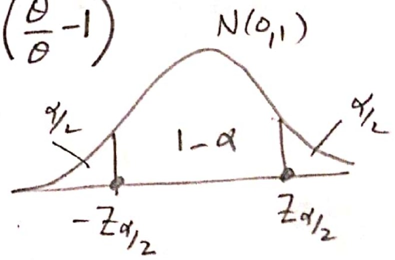
Wald Statistic is:

$$Z = \frac{\hat{\theta} - E(\hat{\theta})}{\sqrt{Var(\hat{\theta})}} = \frac{\hat{\theta} - \theta}{\sqrt{\frac{\theta^2}{N}}} \stackrel{.}{\sim} N(0,1) \quad \text{(approximately)}$$

$$Z^2 = \frac{(\hat{\theta} - \theta)^2}{\theta^2/N} = \frac{N(\hat{\theta} - \theta)^2}{\theta^2} = N\left(\frac{\hat{\theta} - \theta}{\theta}\right)^2 \stackrel{.}{\sim} \chi^2_{(1)}$$

(c)

$$Z = \frac{\hat{\theta} - \theta}{\sqrt{\frac{\theta^2}{N}}} = \frac{\sqrt{N}(\hat{\theta} - \theta)}{\theta} = \sqrt{N}\left(\frac{\hat{\theta}}{\theta} - 1\right)$$




$$1 - \alpha = P\left[-Z_{\alpha/2} < Z < Z_{\alpha/2}\right]$$

$$= P\left[-Z_{\alpha/2} < \sqrt{N}\left(\frac{\hat{\theta}}{\theta} - 1\right) < Z_{\alpha/2}\right]$$

$$= P\left[-Z_{\alpha/2}\frac{1}{\sqrt{N}} < \frac{\hat{\theta}}{\theta} - 1 < Z_{\alpha/2}\frac{1}{\sqrt{N}}\right]$$

$$= P\left[1 - Z_{\alpha/2}\frac{1}{\sqrt{N}} < \frac{\hat{\theta}}{\theta} < 1 + Z_{\alpha/2}\frac{1}{\sqrt{N}}\right]$$

$$= P\left[\frac{1}{1 + Z_{\alpha/2}\frac{1}{\sqrt{N}}} < \frac{\theta}{\hat{\theta}} < \frac{1}{1 - Z_{\alpha/2}\frac{1}{\sqrt{N}}}\right]$$

$$= P\left[\frac{\hat{\theta}}{1 + \frac{Z_{\alpha/2}}{\sqrt{N}}} < \theta < \frac{\hat{\theta}}{1 - \frac{Z_{\alpha/2}}{\sqrt{N}}}\right]$$

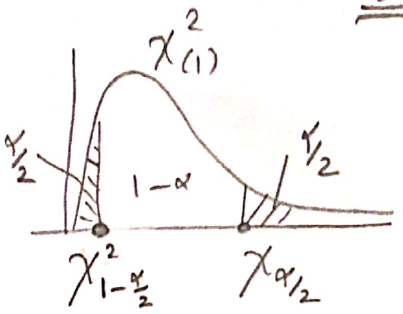
$\therefore$ $(1-\alpha)100\%$ C.I for $\theta$ is

$$\frac{\hat{\theta}}{1 + \frac{Z_{\alpha/2}}{\sqrt{N}}} < \theta < \frac{\hat{\theta}}{1 - \frac{Z_{\alpha/2}}{\sqrt{N}}}$$

The 95% C.I. for $\theta$ is $\left(\frac{\hat{\theta}}{1 + \frac{1.96}{\sqrt{N}}}, \frac{\hat{\theta}}{1 - \frac{1.96}{\sqrt{N}}}\right)$.

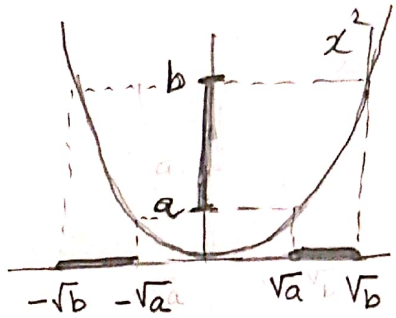
2nd Method:

$$\frac{N(\hat{\theta}-\theta)^2}{\theta^2} = N\left(\frac{\hat{\theta}}{\theta}-1\right)^2 \overset{\cdot}{\sim} \chi^2_{(1)}$$




$$1-\alpha = P\left[\chi^2_{1-\frac{\alpha}{2}} < N\left(\frac{\hat{\theta}}{\theta}-1\right)^2 < \chi^2_{\alpha/2}\right]$$

$$= P\left[\frac{\chi^2_{1-\frac{\alpha}{2}}}{N} < \left(\frac{\hat{\theta}}{\theta}-1\right)^2 < \frac{\chi^2_{\alpha/2}}{N}\right]$$




$$= P\left[\left(\sqrt{\frac{\chi^2_{1-\alpha/2}}{N}} < \frac{\hat{\theta}}{\theta}-1 < \sqrt{\frac{\chi^2_{\alpha/2}}{N}}\right) \cup \left(-\sqrt{\frac{\chi^2_{\alpha/2}}{N}} < \frac{\hat{\theta}}{\theta}-1 < -\sqrt{\frac{\chi^2_{1-\alpha/2}}{N}}\right)\right]$$

$$= P\left\{\left[\frac{1}{\hat{\theta}}\left(1+\sqrt{\frac{\chi^2_{1-\alpha/2}}{N}}\right) < \frac{1}{\theta} < \frac{1}{\hat{\theta}}\left(1+\sqrt{\frac{\chi^2_{\alpha/2}}{N}}\right)\right] \cup \left[\frac{1}{\hat{\theta}}\left(1-\sqrt{\frac{\chi^2_{\alpha/2}}{N}}\right) < \frac{1}{\theta} < \frac{1}{\hat{\theta}}\left(1-\sqrt{\frac{\chi^2_{1-\alpha/2}}{N}}\right)\right]\right\}$$

$$= P\left\{\left(\frac{\hat{\theta}}{1+\sqrt{\frac{\chi^2_{\alpha/2}}{N}}} < \theta < \frac{\hat{\theta}}{1+\sqrt{\frac{\chi^2_{1-\alpha/2}}{N}}}\right) \cup \left(\frac{\hat{\theta}}{1-\sqrt{\frac{\chi^2_{1-\alpha/2}}{N}}} < \theta < \frac{\hat{\theta}}{1-\sqrt{\frac{\chi^2_{\alpha/2}}{N}}}\right)\right\}$$

$\therefore$ $(1-\alpha)100\%$ C.I. for $\theta$ is (the region):

$$\left(\frac{\hat{\theta}}{1+\sqrt{\frac{\chi^2_{\alpha/2}}{N}}}, \frac{\hat{\theta}}{1+\sqrt{\frac{\chi^2_{1-\alpha/2}}{N}}}\right) \cup \left(\frac{\hat{\theta}}{1-\sqrt{\frac{\chi^2_{1-\alpha/2}}{N}}}, \frac{\hat{\theta}}{1-\sqrt{\frac{\chi^2_{\alpha/2}}{N}}}\right)$$

(d) $Y \sim \text{Pareto}(\theta)$ ; $f(y;\theta) = \theta y^{-\theta-1}$ ; $y > 1$, $\theta > 0$

The cumulative distⁿ function is:

$$F(y) = \int_1^y \theta x^{-\theta-1}\, dx = \theta \left[ \frac{x^{-\theta-1+1}}{-\theta-1+1} \right]_{x=1}^{y}$$

$$= \theta \left[ \frac{x^{-\theta}}{-\theta} \right]_{x=1}^{y} = \left[ -x^{-\theta} \right]_{x=1}^{y}$$

$$= -\left[ y^{-\theta} - 1^{-\theta} \right] = 1 - y^{-\theta} \quad ; \; y > 1 \quad (\theta > 0)$$

Suppose $U \sim \text{Uniform}(0,1)$.

To generate $Y$, we first generate a random number $U$ from Uniform$(0,1)$, and then equate $u = F(y)$ and then solve for $y = F^{-1}(u)$.

For Pareto,

$$u = 1 - y^{-\theta} \iff y^{-\theta} = 1 - u \iff \left(\frac{1}{y}\right)^{\theta} = 1 - u$$

$$\iff \frac{1}{y} = (1-u)^{1/\theta} \iff y = (1-u)^{-1/\theta}$$

$$\therefore \; y = \left(\frac{1}{1-u}\right)^{\frac{1}{\theta}}$$

Note: $U \sim \text{Uniform}(0,1) \iff 1-U \sim \text{Uniform}(0,1)$

Thus, we can use:

$$y = \left(\frac{1}{u}\right)^{1/\theta}$$

- We will use computer to do that. (See Minitab output) For $\theta = 2$.

• For $\theta = 2$: Generating Process:

• Using Minitab, we generate 100 random numbers from uniform(0,1):

$$y = \left(\frac{1}{u}\right)^{1/\theta} = \left(\frac{1}{u}\right)^{\frac{1}{2}} = \sqrt{\frac{1}{u}}.$$

| $u_i$ | $y_i$ |
|---|---|
| 0.1290 | 2.7842 |
| 0.4867 | 1.4335 |
| ⋮ | ⋮ |
| 0.7293 | 1.1710 |

$$\hat{\theta} = \frac{N}{\sum \ln(y_i)} \quad , N = 100$$

$$\sum_{i=1}^{N} \ln(y_i) = 51.786$$

$$\hat{\theta} = \frac{100}{51.786} = 1.931$$

$\hat{\theta} = 1.931$

95% C.I. for $\theta$ is
$$\frac{\hat{\theta}}{1 + \frac{1.96}{\sqrt{N}}} < \theta < \frac{\hat{\theta}}{1 - \frac{1.96}{\sqrt{N}}}$$

$$= \frac{1.931}{1 + \frac{1.96}{10}} < \theta < \frac{1.931}{1 - \frac{1.96}{10}}$$

$$= 1.7247 < \theta < 2.4017$$

• We repeat this process 20 times using Minitab Macro, and found that:

① $\overline{\hat{\theta}} = 1.99$

② out of 20 95% C.I.'s for $\theta$, 19 intervals contain $\theta = 2.0$ (See Minitab Output)

5.4 :

Using data in Exercise 4.2 :

| $y_i$ | $x_i$ |
|---|---|
| 65 | 3.36 |
| 156 | 2.88 |
| ⋮ | ⋮ |
| 65 | 5 |

$y_i$ = time to death (in weeks)

$x_i = \log_{10}$ (initial white blood white cell)

$\mu_i = E(Y_i) = e^{\beta_1 + \beta_2 x_i}$

$= e^{\eta(x_i)} = e^{\eta_i}$ (This ensure that $\mu_i > 0$)

$\Leftrightarrow \ln(\mu_i) = \beta_1 + \beta_2 x_i$

$= \eta(x_i)$

$= \eta_i$

link function

$g(\mu_i) = \ln(\mu_i)$

We will use the following form for Exp(θ)

$f(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ ; $y>0$ $(\theta>0)$

$E(Y_i) = \theta_i$

$Var(Y_i) = \theta_i^2$

(a) Wald Statistic :

$$(\underset{\sim}{b} - \underset{\sim}{\beta})^t \underline{J} (\underset{\sim}{b} - \underset{\sim}{\beta}) \sim \chi^2_{(p)}$$

From Exercies 4,2, We have :

$b_1 = 8.4775$

$b_2 = -1.1093$

$$\underline{J}(\underset{\sim}{b}) = \begin{bmatrix} 17 & 69.63 \\ 69.63 & 291.457 \end{bmatrix} = Var(\underset{\sim}{U})$$

$$\underline{J}^{-1}(\underset{\sim}{b}) = \begin{bmatrix} 2.73839 & -0.654209 \\ -0.654209 & 0.159724 \end{bmatrix} = Var(\underset{\sim}{b})$$

$$\underset{\sim}{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 8.4775 \\ -1.1093 \end{pmatrix}$$

- Individuals C.I. for $\beta_1$ and $\beta_2$:

95% C.I. for $\beta_1$ is:

$$b_1 \pm 1.96\, se(b_1) = 8.4775 \pm 1.96 \sqrt{2.73839}$$

$$5.23408 < \beta_1 < 11.72092$$

95% C.I. for $\beta_2$ is:

$$b_2 \pm 1.96\, se(b_2) = -1.1093 \pm 1.96 \sqrt{0.159724}$$

$$-1.89262 < \beta_2 < -0.32598$$

- Simultaneous C.I. for $\beta_1$ and $\beta_2$:

A $(1-\alpha)100\%$ confidence <u>Region</u> for $\underset{\sim}{\beta}$ is

$$\left\{ \underset{\sim}{\beta} : \chi^2_{1-\alpha/2(2)} < (\underset{\sim}{b} - \underset{\sim}{\beta})^t \, \underline{J} \, (\underset{\sim}{b} - \underset{\sim}{\beta}) < \chi_{\alpha/2(2)} \right\}$$

where:

$$(\underset{\sim}{b} - \underset{\sim}{\beta})^t \, \underline{J} \, (\underset{\sim}{b} - \underset{\sim}{\beta}) = (8.4775 - \beta_1, -1.1093 - \beta_2) \begin{bmatrix} 17 & 69.63 \\ 69.63 & 291.457 \end{bmatrix} * \begin{pmatrix} 8.4775 - \beta_1 \\ -1.1093 - \beta_2 \end{pmatrix}$$

$$= 17(8.4775 - \beta_1)^2 + 291.457(-1.1093 - \beta_2)^2 + 2(69.63)(8.4775 - \beta_1)(-1.1093 - \beta_2)$$

(b) $\begin{cases} H_0: \beta_2 = 0 \\ H_1: \beta_2 \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0: \mu_i = e^{\beta_1} \\ H_1: \mu_i = e^{\beta_1 + \beta_2 x_i} \end{cases}$

(1) Model under $H_0$:

$\mu_i = e^{\beta_1} \quad i = 1, \dots, N$ $\left( \begin{matrix} \text{Recall} \\ \mu_i = \theta_i \end{matrix} \right)$

For this model

$$\mu_1 = \mu_2 = \dots = \mu_N = \mu = \theta$$

The log-likelihood function is:

$$l_0 = l(\theta) = \sum_{i=1}^{N} \ln\left(\frac{1}{\theta} e^{-y_i/\theta}\right) \qquad \left(f(y;\theta) = \frac{1}{\theta} e^{-y/\theta}\right)$$

$$= -N \ln(\theta) - \frac{1}{\theta} \sum_{i=1}^{N} y_i$$

$$= -N \ln(\theta) - \frac{N}{\theta} \bar{y}$$

$$\frac{\partial l_0}{\partial \theta} = -\frac{N}{\theta} + \frac{N\bar{y}}{\theta^2}$$

$$\frac{\partial l_0}{\partial \theta} \overset{\text{set}}{=} 0 \Leftrightarrow \frac{N\bar{Y}}{\theta^2} = \frac{N}{\theta} \Leftrightarrow \frac{\bar{Y}}{\theta} = 1 \Leftrightarrow \hat{\theta} = \bar{Y}$$

The MLE of $\theta$ is $\hat{\theta} = \bar{y}$ $\left( \left.\frac{\partial^2 l_0}{\partial \theta^2}\right|_{\theta = \hat{\theta}} = \dots = \frac{-N}{(\bar{y})^2} < 0 \right)$

The Maximum value of the log-likelihood function is

$$\hat{l}_0 = l(\hat{\theta}) = -N \ln(\hat{\theta}) - \frac{N}{\hat{\theta}} \bar{y} = -N \ln(\bar{y}) - N\left(\frac{\bar{y}}{\bar{y}}\right)$$

$$= -N\left[\ln(\bar{y}) + 1\right] \qquad (\bar{y} = 62.4706)$$

$$= -17\left[\ln(62.4706) + 1\right]$$

$$= -87.289832$$

(2) Model under $H_1$:

$$\mu_i = e^{\beta_1+\beta_2 x_i} \qquad (\mu_i = \theta_i)$$

We have found that the MLE's of $\beta_1$ and $\beta_2$ are: $b_1 = 8.4775$ and $b_2 = -1.1093$

Therefore, the MLE of $\mu_i = \theta_i$ is

$$\hat{\theta}_i = e^{b_1+\beta_2 x_i} = e^{8.4775-1.1093\,x_i}$$

The log-likelihood function is:

$$l_1 = l(\theta_1, \ldots, \theta_N) = \sum_{i=1}^{N} \ln\left(\frac{1}{\theta_i} e^{-y_i/\theta_i}\right)$$

$$= \sum_{i=1}^{N} \left[-\ln(\theta_i) - \frac{y_i}{\theta_i}\right]$$

$$= -\sum_{i=1}^{N} \ln(\theta_i) - \sum_{i=1}^{N} \frac{y_i}{\theta_i}$$

The maximum value of the log-likelihood function is:

$$\hat{l}_1 = l(\hat{\theta}_1, \ldots, \hat{\theta}_N) = -\sum_{i=1}^{N} \ln(\hat{\theta}_i) - \sum_{i=1}^{N} \frac{y_i}{\hat{\theta}_i}$$

$$= -\sum_{i=1}^{N} \ln\left(e^{8.4775-1.1093 x_i}\right) - \sum_{i=1}^{N} \frac{y_i}{e^{8.4775-1.1093 x_i}}$$

$$= -\sum_{i=1}^{N} (8.4775-1.1093 x_i) - \sum_{i=1}^{N} \frac{y_i}{e^{8.4775-1.1093 x_i}}$$

$$= -(66.8769) - (17.0001) \qquad \text{(Calculated from the data)}$$

$$= -83.877$$

The test statistic is :

$$\Delta D = 2[\hat{l}_1 - \hat{l}_0]$$
$$= 2[-83.877 - (-87.289832)]$$
$$= 6.825664$$

under Ho, $\Delta D \sim \chi^2_{(2-1)} = \chi^2_{(1)}$

$$\chi^2_{0.05}(1) = 3.842$$

Since $\Delta D > \chi^2_{0.05(1)}$, we reject $H_0: \beta_2 = 0$ and we conclude that $x$ has a significant importance in predicting $y$.