Solutions to exercises - Week 45

Likelihood ratio tests

- Exercises 8.3, 8.6 and 8.37.c

Exercise 8.3

Let $Y_1,....,Y_m$ be iid Bernoulli variables with pmf

$$f(y_i\,|\,\theta)=\theta^{y_i}(1-\theta)^{1-y_i} \quad \text{for} \quad y_i=0,1; \quad 0<\theta<1$$

We will test $H_0:\theta\le\theta_0$ versus $H_1:\theta>\theta_0$

The likelihood is given by $[\,\mathbf{y}=(y_1,...,y_m)\,]$

$$L(\theta\,|\,\mathbf{y})=\prod_{i=1}^{m}\theta^{y_i}(1-\theta)^{1-y_i}=\theta^{T(\mathbf{y})}(1-\theta)^{m-T(\mathbf{y})}$$

where

$$T=T(\mathbf{y})=\sum_{i=1}^{m}y_i$$

The unrestricted MLE for θ is $\hat{\theta}=T/m$

The MLE for θ under $H_0:\theta\le\theta_0$ is

$$\hat{\theta}_0=\begin{cases} T/m & \text{if } T/m\le\theta_0 \\ \theta_0 & \text{if } T/m>\theta_0 \end{cases}$$

Thus the LRT statistic becomes

$$\lambda(\mathbf{y})=\frac{L(\hat{\theta}_0\,|\,\mathbf{y})}{L(\hat{\theta}\,|\,\mathbf{y})}$$

$$=\begin{cases} 1 & \text{if } T/m\le\theta_0 \\ \dfrac{\theta_0^T(1-\theta_0)^{m-T}}{(T/m)^T(1-T/m)^{m-T}} & \text{if } T/m>\theta_0 \end{cases}$$

The likelihood ratio test rejects the null hypothesis if

$$\lambda(\mathbf{y})=\frac{\theta_0^T(1-\theta_0)^{m-T}}{(T/m)^T(1-T/m)^{m-T}}\le c$$

i.e. if (by taking logarithms and multiplying by -1)

$$g(T)\overset{\text{def}}{=}T\log\left(\frac{T/m}{\theta_0}\right)+(m-T)\log\left(\frac{1-T/m}{1-\theta_0}\right)\ge-\log c$$

One may show that $g(T)$ is an increasing function of T when $T/m>\theta_0$

Hence the test rejects H_0 if $T\ge b$

Exercise 8.6

We have independent random variables:

$$X_1, \ldots, X_n \sim \text{exponential}(\theta)$$

$$Y_1, \ldots, Y_m \sim \text{exponential}(\mu)$$

We will test $H_0: \theta = \mu$ versus $H_1: \theta \neq \mu$

a) The likelihood takes the form

$$L(\theta, \mu \mid \mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} f(x_i \mid \theta) \prod_{j=1}^{m} f(y_j \mid \mu)$$

$$= \prod_{i=1}^{n} \left(\frac{1}{\theta} e^{-x_i/\theta} \right) \prod_{j=1}^{m} \left(\frac{1}{\mu} e^{-y_j/\mu} \right)$$

$$= \frac{1}{\theta^n \mu^m} \exp\left\{ -\frac{1}{\theta} \sum_{i=1}^{n} x_i - \frac{1}{\mu} \sum_{j=1}^{m} y_j \right\}$$

The unrestricted MLE for θ and μ are

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \hat{\mu} = \frac{1}{m} \sum_{j=1}^{m} y_j$$

The MLE for $\theta = \mu$ under H_0 is

$$\hat{\theta}_0 = \hat{\mu}_0 = \frac{1}{n+m} \left(\sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_j \right)$$

Thus the LRT statistic becomes

$$\lambda(\mathbf{x}, \mathbf{y}) = \frac{L(\hat{\theta}_0, \hat{\mu}_0 \mid \mathbf{x}, \mathbf{y})}{L(\hat{\theta}, \hat{\mu} \mid \mathbf{x}, \mathbf{y})} = \frac{\frac{1}{\hat{\theta}_0^n \hat{\mu}_0^m} \exp\{-(n+m)\}}{\frac{1}{\hat{\theta}^n \hat{\mu}^m} \exp\{-n-m\}} = \frac{\hat{\theta}^n \hat{\mu}^m}{\hat{\theta}_0^n \hat{\mu}_0^m}$$

This gives

$$\lambda(\mathbf{x}, \mathbf{y}) = \frac{(n+m)^{n+m} \left(\sum_{i=1}^{n} x_i \right)^n \left(\sum_{j=1}^{m} y_j \right)^m}{n^n m^m \left(\sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_j \right)^{n+m}}$$

The likelihood ratio test rejects the null hypothesis if $\lambda(\mathbf{x}, \mathbf{y}) \leq c$

b) The LRT statistic may be rewritten as

$$\lambda(\mathbf{x}, \mathbf{y}) = \frac{(n+m)^{n+m}}{n^n m^m} \left(\frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_j} \right)^n \left(\frac{\sum_{j=1}^{m} y_j}{\sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_j} \right)^m$$

If we introduce

$$T = T(\mathbf{x}, \mathbf{y}) = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_j}$$

we may write

$$\lambda(\mathbf{x}, \mathbf{y}) = \frac{(n+m)^{n+m}}{n^n m^m} T^n (1-T)^m$$

Rejection for $\lambda(\mathbf{x}, \mathbf{y}) \leq c$ is equivalent to rejection for $T \leq a$ or $T \geq b$, where $0 < a < b$ are constants that satisfy $a^n (1-a)^m = b^n (1-b)^m$

c) When H_0 is true we have:

$$\sum_{i=1}^{n} X_i \sim \text{gamma}(n,\theta) \qquad \sum_{j=1}^{m} Y_j \sim \text{gamma}(m,\theta)$$

By the example on slides 19-20 for the lectures of week 35 (August 31st), we then have that

$$T = \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j}$$

is $\text{beta}(n,m)$ distributed

Exercise 8.37.c

Let $X_1, X_2, \ldots, X_n$ be iid $n(\theta,\sigma^2)$

We will test $H_0: \theta \le \theta_0$ versus $H_1: \theta > \theta_0$

Consider the test that rejects H_0 when

$$\bar{X} > \theta_0 + t_{n-1,\alpha} S/\sqrt{n} \qquad (*)$$

The power function of this test is

$$\beta(\theta,\sigma^2) = P_{\theta,\sigma^2}\left(\bar{X} > \theta_0 + t_{n-1,\alpha} S/\sqrt{n}\right)$$

When $\theta \le \theta_0$ we have (with $T \sim t_{n-1}$-distributed)

$$\beta(\theta,\sigma^2) \le P_{\theta_0,\sigma^2}\left(\bar{X} > \theta_0 + t_{n-1,\alpha} S/\sqrt{n}\right)$$

$$= P_{\theta_0,\sigma^2}\left(\frac{\bar{X}-\theta}{S/\sqrt{n}} > t_{n-1,\alpha}\right) = P\left(T > t_{n-1,\alpha}\right) = \alpha$$

Thus the test given by (*) has size α

We will show that the test corresponds to the LRT

The likelihood takes the form

$$L(\mu,\sigma^2 \mid \mathbf{x}) = (2\pi\sigma^2)^{-n/2} \exp\left\{-(1/2)\sum (x_i-\theta)^2/\sigma^2\right\}$$

The unrestricted MLEs are given by

$$\hat{\theta} = \bar{x}$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\sum (x_i - \bar{x})^2$$

If $\hat{\theta} = \bar{x} \le \theta_0$, the restricted MLEs (i.e. under H_0) are the same as the unrestricted MLEs, while if $\hat{\theta} > \theta_0$ the restricted MLEs are

$$\hat{\theta}_0 = \theta_0$$

and

$$\hat{\sigma}_0^2 = \frac{1}{n}\sum (x_i - \theta_0)^2$$

Thus the LRT statistic becomes

$$\lambda(\mathbf{x}) = \begin{cases} 1 & \text{if } \bar{x} \le \theta_0 \\ \dfrac{L(\theta_0, \hat{\sigma}_0^2 \mid \mathbf{x})}{L(\hat{\theta}, \hat{\sigma}^2 \mid \mathbf{x})} & \text{if } \bar{x} > \theta_0 \end{cases}$$

Now we have

$$\frac{L(\theta_0,\hat{\sigma}_0^2 \mid \mathbf{x})}{L(\hat{\theta},\hat{\sigma}^2 \mid \mathbf{x})} = \frac{(2\pi\hat{\sigma}_0^2)^{-n/2}\exp\left\{-(1/2)\sum(x_i-\theta_0)^2/\hat{\sigma}_0^2\right\}}{(2\pi\hat{\sigma}^2)^{-n/2}\exp\left\{-(1/2)\sum(x_i-\overline{x})^2/\hat{\sigma}^2\right\}}$$

$$= \frac{(2\pi\hat{\sigma}_0^2)^{-n/2}\exp\left\{-(n/2)\right\}}{(2\pi\hat{\sigma}^2)^{-n/2}\exp\left\{-(n/2)\right\}}$$

$$= \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2}$$

$$= \left(\frac{\sum(x_i-\overline{x})^2}{\sum(x_i-\theta_0)^2}\right)^{n/2}$$

The likelihood ratio test rejects H_0 if (assuming $\overline{x} > \theta_0$)

$$\lambda(\mathbf{x}) = \left(\frac{\sum(x_i-\overline{x})^2}{\sum(x_i-\theta_0)^2}\right)^{n/2} \le c$$

$\Leftrightarrow$

$$\frac{\sum(x_i-\overline{x})^2}{\sum(x_i-\theta_0)^2} \le c^{2/n}$$

$\Leftrightarrow$

$$\frac{\sum(x_i-\overline{x})^2 + n(\overline{x}-\theta_0)^2}{\sum(x_i-\overline{x})^2} \ge c^{-2/n}$$

$\Leftrightarrow$

$$1 + \frac{n(\overline{x}-\theta_0)^2}{\sum(x_i-\overline{x})^2} \ge c^{-2/n}$$

Thus (assuming $\overline{x} > \theta_0$)

$$\lambda(\mathbf{x}) \le c \quad \Leftrightarrow \quad \frac{n(\overline{x}-\theta_0)^2}{\sum(x_i-\overline{x})^2} \ge c^{-2/n} - 1 = k$$

Thus the LRT rejects H_0 if

$$\frac{\overline{x}-\theta_0}{\sqrt{\sum(x_i-\overline{x})^2/n}} \ge \sqrt{k}$$

which is equivalent to

$$\frac{\overline{x}-\theta_0}{s/\sqrt{n}} > K$$

for a suitably chosen constant *K*

Thus the test (*) corresponds to the LRT