Stat 346/446: Theoretical Statistics II: Homework 5 Solutions

Textbook Exercises

8.3 (346 & 446 : 2 pts.)Here, the LRT alluded to in Example 8.2.9 will be derived. Suppose that we observe m iid Bernoulli(θ) random variables, denoted by $Y_1, \ldots, Y_m$. Show that the LRT of $H_0\colon \theta \le \theta_0$ versus $H_1\colon \theta > \theta_0$ will reject H_0 if $\sum_{i=1}^m Y_i > b$.

First: $Y = \sum_{i=1}^m Y_i$ is a sufficient statistic (see e.g. Example 7.57, complete $\Rightarrow$ also sufficient). The LRT rejects H_0 if $\lambda(y) < c$ for some c, where

$$\lambda(y) = \frac{\sup_{\theta \le \theta_0} L(\theta \mid y)}{\sup_{\theta\in\Theta} L(\theta \mid y)} = \frac{\sup_{\theta \le \theta_0} \binom{m}{y}\theta^y(1-\theta)^{m-y}}{\sup_{\theta\in\Theta} \binom{m}{y}\theta^y(1-\theta)^{m-y}}.$$

We know that the MLE for θ is $\hat\theta = \frac{y}{m}$. Thus,

$$\sup_{\theta\in\Theta} L(\theta \mid y) = \begin{cases} L(\theta_0 \mid y) & \text{if } \theta_0 \le \hat\theta, \\ L(\hat\theta \mid y) & \text{if } \theta_0 > \hat\theta. \end{cases}$$

So,

$$\lambda(y) = \begin{cases} \frac{L(\theta_0|y)}{L(\hat\theta|y)} = \frac{\theta_0^y(1-\theta_0)^{m-y}}{(y/m)^y(1-y/m)^{m-y}} & \text{if } \theta_0 \le \frac{y}{m}, \\ 1 & \text{if } \theta_0 > \frac{y}{m}. \end{cases}$$

So the LRT rejects H_0 if $\frac{y}{m} \ge \theta_0$ and

$$\lambda(y) = \frac{\theta_0^y(1-\theta_0)^{m-y}}{(y/m)^y(1-y/m)^{m-y}} \le c \quad \text{for some } c.$$

We want to show that this is equivalent to rejecting if $y > b$ for some b. If $\lambda(y)$ is a decreasing function of y, then $\lambda(y) < c$ means $y > b$. To show that, we examine:

$$\lambda(y) \text{ decreasing} \Leftrightarrow \log(\lambda(y)) \text{ decreasing,}$$

$$\Rightarrow \text{ show that } \frac{d}{dy}\log(\lambda(y)) < 0 \text{ for } \theta_0 \le \frac{y}{m}.$$

$$\begin{aligned}
\frac{d}{dy}\log(\lambda(y)) &= \frac{d}{dy}\left[y\log\theta_0 + (m-y)\log(1-\theta_0) - y\log\left(\frac{y}{m}\right) - (m-y)\log\left(\frac{m-y}{m}\right)\right] \\
&= \log\theta_0 - \log(1-\theta_0) - \log\left(\frac{y}{m}\right) - y\cdot\frac{1}{y} + \log\left(\frac{m-y}{m}\right) + (m-y)\cdot\left(\frac{-1}{m-y}\right) \\
&= \log\left(\frac{\theta_0}{1-\theta_0}\cdot\frac{m-y}{y}\right) - 1 + 1 \\
&= \log\left(\frac{\theta_0}{y/m}\cdot\frac{1-y/m}{1-\theta_0}\right) = \log\left(\frac{\theta_0}{y/m}\cdot\frac{1-y/m}{1-\theta_0}\right).
\end{aligned}$$

If $\theta_0 \le \frac{y}{m} \Rightarrow \frac{\theta_0}{y/m} \le 1$, and similarly

$$1 - \theta_0 \ge 1 - \frac{y}{m} \Rightarrow \frac{1-y/m}{1-\theta_0} \le 1.$$

Hence,

$$\frac{d}{dy}\log(\lambda(y)) < 0.$$

which shows that λ is decreasing in y and $\lambda(y) < c$ if and only if $y > b$.

8.6 Suppose that we have two independent random samples: $X_1, \ldots, X_n$ are exponential(θ), and $Y_1, \ldots, Y_m$ are exponential(μ).

(a) (346 : 2 pts, 446 : 1 pts.) Find the LRT of $H_0 : \theta = \mu$ versus $H_1 : \theta \neq \mu$.

$$\lambda(\mathbf{x}, \mathbf{y}) = \frac{\sup_{\theta_0} L(\theta|\mathbf{x}, \mathbf{y})}{\sup_{\theta,\mu} L(\theta|\mathbf{x}, \mathbf{y})} = \frac{\sup_\theta \prod_{i=1}^n \frac{1}{\theta} e^{-x_i/\theta} \cdot \prod_{j=1}^m \frac{1}{\theta} e^{-y_j/\theta}}{\sup_{\theta,\mu} \prod_{i=1}^n \frac{1}{\theta} e^{-x_i/\theta} \cdot \prod_{j=1}^m \frac{1}{\mu} e^{-y_j/\mu}}$$

$$= \frac{\sup_\theta \frac{1}{\theta^{n+m}} \exp\left\{-\left(\sum_{i=1}^n x_i + \sum_{j=1}^m y_j\right)/\theta\right\}}{\sup_{\theta,\mu} \frac{1}{\theta^n} \exp\left\{-\sum_{i=1}^n x_i/\theta\right\} \cdot \frac{1}{\mu^m} \exp\left\{-\sum_{j=1}^m y_j/\mu\right\}}$$

Differentiation will show that in the numerator

$$\hat{\theta}_0 = \frac{\sum_i x_i + \sum_j y_j}{n+m},$$

while in the denominator

$$\hat{\theta} = \bar{x}, \quad \hat{\mu} = \bar{y}.$$

Therefore,

$$\lambda(\mathbf{x}, \mathbf{y}) = \frac{\left(\frac{n+m}{\sum x_i + \sum y_j}\right)^{n+m} \exp\left\{-\left(\frac{n+m}{\sum x_i + \sum y_j}\right)\left(\sum x_i + \sum y_j\right)\right\}}{\left(\frac{n}{\sum x_i}\right)^n \exp\left\{-\left(\frac{n}{\sum x_i}\right)\sum x_i\right\} \cdot \left(\frac{m}{\sum y_j}\right)^m \exp\left\{-\left(\frac{m}{\sum y_j}\right)\sum y_j\right\}}$$

$$= \frac{(n+m)^{n+m}\left(\sum x_i\right)^n \left(\sum y_j\right)^m}{n^n m^m \left(\sum x_i + \sum y_j\right)^{n+m}}$$

And the LRT is to reject H_0 if $\lambda(\mathbf{x}, \mathbf{y}) \leq c$.

(b) (346 : 2 pts, 446 : 1 pts.) Show that the test in part (a) can be based on the statistic

$$T = \frac{\sum X_i}{\sum X_i + \sum Y_i}.$$

$$\lambda = \frac{(n+m)^{n+m}}{n^n m^m}\left(\frac{\sum_j x_i}{\sum_i x_i + \sum_j y_j}\right)^n \left(\frac{\sum_j y_j}{\sum_i x_i + \sum_j y_j}\right)^m = \frac{(n+m)^{n+m}}{n^n m^m} T^n (1-T)^m.$$

Therefore λ is a function of T. λ is a unimodal function of T which is maximized when

$$T = \frac{n}{n+m}.$$

Rejection for $\lambda \leq c$ is equivalent to rejection for $T \leq a$ or $T \geq b$, where a and b are constants that satisfy

$$a^n(1-a)^m = b^n(1-b)^m.$$

(c) (446 : 1 pts.) Find the distribution of T when H_0 is true.

Under $H_0 : \theta = \mu$, we have:

$$\sum X_i \sim \text{Gamma}(n, \theta), \quad \sum Y_i \sim \text{Gamma}(m, \theta),$$

and these two sums are independent. Define

$$T = \frac{\sum X_i}{\sum X_i + \sum Y_i}.$$

It is a known result that if $U \sim \text{Gamma}(n, \theta)$ and $V \sim \text{Gamma}(m, \theta)$ are independent, then

$$\frac{U}{U+V} \sim \text{Beta}(n, m).$$

Therefore, under H_0, the statistic T follows a Beta distribution:

$$T \sim \text{Beta}(n, m).$$

8.7 (446 : 2 pts.)We have already seen the usefulness of the LRT in dealing with problems with nuisance parameters. We now look at some other nuisance parameter problems.

(a) Find the LRT of

$$H_0 : \theta \leq 0 \qquad \text{versus} \qquad H_1 : \theta > 0$$

based on a sample $X_1, \ldots, X_n$ from a population with probability density function

$$f(x \mid \theta, \lambda) = \frac{1}{\lambda} e^{-(x-\theta)/\lambda} \mathbb{I}_{[\theta,\infty)}(x),$$

where both θ and λ are unknown.

$$L(\theta, \lambda \mid \mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\lambda} e^{-(x_i-\theta)/\lambda} \mathbb{I}_{[\theta,\infty)}(x_i) = \left(\frac{1}{\lambda}\right)^n e^{-(\sum x_i - n\theta)/\lambda} \mathbb{I}_{[\theta,\infty)}(x_{(1)}),$$

which is increasing in θ if $x_{(1)} \geq \theta$ (regardless of λ). So the MLE of θ is $\hat{\theta} = x_{(1)}$. Then

$$\frac{\partial \log L}{\partial \lambda} = -\frac{n}{\lambda} + \frac{\sum x_i - n\hat{\theta}}{\lambda^2} \quad \text{set } = 0 \quad \Rightarrow \quad n\hat{\lambda} = \sum x_i - n\hat{\theta} \quad \Rightarrow \quad \hat{\lambda} = \bar{x} - x_{(1)}.$$

Because

$$\frac{\partial^2 \log L}{\partial \lambda^2} = \frac{n}{\lambda^2} - 2 \cdot \frac{\sum x_i - n\hat{\theta}}{\lambda^3} = \frac{n}{(\bar{x} - x_{(1)})^2} - \frac{2n(\bar{x} - x_{(1)})}{(\bar{x} - x_{(1)})^3} = \frac{-n}{(\bar{x} - x_{(1)})^2} < 0,$$

we have $\hat{\theta} = x_{(1)}$ and $\hat{\lambda} = \bar{x} - x_{(1)}$ as the unrestricted MLEs of θ and λ. Under the restriction $\theta \leq 0$, the MLE of θ (regardless of λ) is

$$\hat{\theta}_0 = \begin{cases} 0 & \text{if } x_{(1)} > 0, \\ x_{(1)} & \text{if } x_{(1)} \leq 0. \end{cases}$$

For $x_{(1)} > 0$, substituting $\hat{\theta}_0 = 0$ and maximizing with respect to λ, as above, yields $\hat{\lambda}_0 = \bar{x}$. Therefore,

$$\lambda(\mathbf{x}) = \frac{\sup_{\Theta_0} L(\theta, \lambda \mid \mathbf{x})}{\sup_{\Theta} L(\theta, \lambda \mid \mathbf{x})} = \frac{\sup_{(\lambda,\theta):\theta \leq 0} L(\lambda, \theta \mid \mathbf{x})}{L(\hat{\lambda}, \hat{\theta} \mid \mathbf{x})} = \begin{cases} 1 & \text{if } x_{(1)} \leq 0, \\ \frac{L(\bar{x}, 0 \mid \mathbf{x})}{L(\hat{\lambda}, \hat{\theta} \mid \mathbf{x})} & \text{if } x_{(1)} > 0. \end{cases}$$

where

$$L(\bar{x}, 0 \mid \mathbf{x}) = \left(\frac{1}{\bar{x}}\right)^n e^{-n\bar{x}/\bar{x}} = \left(\frac{1}{\bar{x}}\right)^n e^{-n},$$

$$L(\hat{\lambda}, \hat{\theta} \mid \mathbf{x}) = \left(\frac{1}{\hat{\lambda}}\right)^n e^{-n(\bar{x}-x_{(1)})/(\bar{x}-x_{(1)})} = \left(\frac{1}{\bar{x}-x_{(1)}}\right)^n e^{-n}.$$

So

$$\lambda(\mathbf{x}) = \left(\frac{\bar{x}-x_{(1)}}{\bar{x}}\right)^n = \left(1 - \frac{x_{(1)}}{\bar{x}}\right)^n.$$

So rejecting if $\lambda(\mathbf{x}) \le c$ is equivalent to rejecting if $\frac{x_{(1)}}{\bar{x}} \ge c^*$, where c^* is some constant.

8.12 For samples of size $n = 1, 4, 16, 64, 100$ from a normal population with mean μ and known variance σ^2, plot the power function of the following LRTs. Take $\alpha = .05$. (Special note: set $\sigma^2 = 1$)

(a) (346 & 446 : 1 pts.) $H_0 : \mu \le 0$ versus $H_1 : \mu > 0$

For $H_0 : \mu \le 0$ vs. $H_1 : \mu > 0$ the LRT is to reject H_0 if $\bar{x} > c\sigma/\sqrt{n}$ (Example 8.3.3). For $\alpha = .05$ take $c = 1.645$. The power function is

$$\beta(\mu) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > 1.645 - \frac{\mu}{\sigma/\sqrt{n}}\right) = P\left(Z > 1.645 - \frac{\sqrt{n}\mu}{\sigma}\right).$$

Note that the power will equal .5 when $\mu = 1.645\sigma/\sqrt{n}$.

(b) (346 & 446 : 1 pts.) $H_0 : \mu = 0$ versus $H_1 : \mu \ne 0$

For $H_0 : \mu = 0$ vs. $H_A : \mu \ne 0$ the LRT is to reject H_0 if $|\bar{x}| > c\sigma/\sqrt{n}$ (Example 8.2.2). For $\alpha = .05$ take $c = 1.96$. The power function is

$$\beta(\mu) = P\left(-1.96 - \frac{\sqrt{n}\mu}{\sigma} \le Z \le 1.96 + \frac{\sqrt{n}\mu}{\sigma}\right).$$

In this case, $\mu = \pm 1.96\sigma/\sqrt{n}$ gives power of approximately .5.

8.16 (346 : 2 pts, 446 : 1 pts.)One very striking abuse of α levels is to choose them *after* seeing the data and to choose them in such a way as to force rejection (or acceptance) of a null hypothesis. To see what the *true* Type I and Type II Error probabilities of such a procedure are, calculate size and power of the following two trivial tests:

(a) Always reject H_0, no matter what data are obtained (equivalent to the practice of choosing the α level to force rejection of H_0).

Rule: Always reject
Power function:

$$\beta(\theta) = P_\theta(\text{reject}) = 1 \quad \text{for all } \theta$$

$$\begin{aligned} \text{Size} &= P(\text{reject } H_0 \mid H_0 \text{ is true}) = 1 \Rightarrow \text{Type I error} = 1. \\ \text{Power} &= P(\text{reject } H_0 \mid H_A \text{ is true}) = 1 \Rightarrow \text{Type II error} = 0. \end{aligned}$$

(b) Always accept H_0, no matter what data are obtained (equivalent to the practice of choosing the α level to force acceptance of H_0).

$$\begin{aligned} \text{Size} &= P(\text{reject } H_0 \mid H_0 \text{ is true}) = 0 \Rightarrow \text{Type I error} = 0. \\ \text{Power} &= P(\text{reject } H_0 \mid H_A \text{ is true}) = 0 \Rightarrow \text{Type II error} = 1. \end{aligned}$$