IE 605: Engineering Statistics

Solutions of tutorial 10

Solution 1

Power of the LRT,

$$\begin{aligned}
\beta(\theta) &= P_\theta\left(\frac{|\bar{X}-\theta_0|}{\sigma/\sqrt{n}} > c\right) \\
&= 1 - P_\theta\left(\frac{|\bar{X}-\theta_0|}{\sigma/\sqrt{n}} \le c\right) \\
&= 1 - P_\theta\left(-\frac{c\sigma}{\sqrt{n}} \le \bar{X}-\theta_0 \le \frac{c\sigma}{\sqrt{n}}\right) \\
&= 1 - P_\theta\left(\frac{-c\sigma/\sqrt{n}+\theta_0-\theta}{\sigma/\sqrt{n}} \le \frac{\bar{X}-\theta}{\sigma/\sqrt{n}} \le \frac{c\sigma/\sqrt{n}+\theta_0-\theta}{\sigma/\sqrt{n}}\right) \\
&= 1 - P\left(-c+\frac{\theta_0-\theta}{\sigma/\sqrt{n}} \le Z \le c+\frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right) \\
&= 1 + \Phi\left(-c+\frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right) - \Phi\left(c+\frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right)
\end{aligned}$$

where $Z \sim N(0,1)$ and Φ is the standard normal CDF.

Solution 2

The LRT statistic is,

$$\lambda(y) = \frac{sup_{\theta\le\theta_0} L(\theta|y_1,\dots,y_m)}{sup_{\Theta} L(\theta|y_1,\dots,y_m)}.$$

Let $y = \sum_{i=1}^{m} y_i$, and note that the MLE in the numerator is $\min\{y/m, \theta_0\}$ (see below NOTE*) while the denominator has y/m as the MLE (Check). Thus

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

and we reject H_0 if $\quad \dfrac{\theta_0^y(1-\theta_0)^{m-y}}{(y/m)^y(1-y/m)^{m-y}} \le c$

To show that this is equivalent to rejecting if $y > b$, we could show $\lambda(y)$ is decreasing in y so that $\lambda(y) < c$ occurs for $y > b > m\theta_0$. It is easier to work with $\log \lambda(y)$, and we have

$$\log \lambda(y) = y \log \theta_0 + (m-y)\log(1-\theta_0) - y \log\left(\frac{y}{m}\right) - (m-y)\log\left(\frac{m-y}{m}\right),$$

and $$\frac{d}{dy}\log \lambda(y) = \log \theta_0 - \log(1-\theta_0) - \log\left(\frac{y}{m}\right) - y\frac{1}{y} + \log\left(\frac{m-y}{m}\right) + (m-y)\frac{1}{m-y}$$
$$= \log\left(\frac{\theta_0\left(\frac{m-y}{m}\right)}{\frac{y}{m}(1-\theta_0)}\right)$$

For $y/m > \theta_0, 1 - y/m = (m-y)/m < 1-\theta_0$, so each fraction above is less than 1, and the log is less than 0. Thus $\frac{d}{dy}\log\lambda < 0$ which shows that λ is decreasing in y and $\lambda(y) < c$ if and only if $y > b$.

*NOTE: Maximum Likelihood Estimator of θ:

The likelihood function is

$$L(\theta|y) = \theta^y(1-\theta)^{n-y}, \quad \text{where } y = \sum_{i=1}^{m} y_i.$$

We can show that $L(\theta|y)$ is increasing for $\theta \le \bar{y}$ and is decreasing for $\theta \ge \bar{y}$.

Here $0 < \theta < \theta_0$. Therefore, when $\bar{y} \le \theta_o$, $\bar{Y}$ is the MLE of θ, because $\bar{Y}$ is the overall maximum of $L(\theta|y)$. When $\bar{Y} > \theta_0$, $L(\theta|y)$ is an increasing function of θ on $[0,\theta_0]$ and obtains its maximum at the upper bound of θ which is θ_0. So the MLE is $\hat{\theta} = \min\{\bar{Y}, \theta_0\}$.

Solution 3

For discrete random variables, $L(\theta|x) = f(x|\theta) = P(X = x|\theta)$. So the numerator and denominator of $\lambda(x)$ are the supremum of this probability over the indicated sets.

Solution 4

1. The log-likelihood is,

$$\log L(\theta, \nu|x) = n\log\theta + n\theta\log\nu - (\theta+1)\log\left(\prod_i x_i\right), \quad \nu \le x_{(1)}$$

where $x_{(1)} = \min_i x_i$. For any value of θ, this is an increasing function of ν for $\nu \le x_{(1)}$. So both the restricted and unrestricted MLEs of ν are $\hat{\nu} = x_{(1)}$. To find the MLE of θ, set

$$\frac{\partial}{\partial\theta}\log L(\theta, x_{(1)}|x) = \frac{n}{\theta} + n\log x_{(1)} - \log\left(\prod_i x_i\right) = 0,$$

and solve for θ yielding

$$\hat{\theta} = \frac{n}{\log(\prod_i x_i / x_{(1)}^n)} = \frac{n}{T}$$

$$(\partial^2/\partial\theta^2) \log L(\theta, x_{(1)}|x) = -n/\theta^2 < 0, \quad \text{for all } \theta.$$

So $\hat{\theta}$ is a maximum.

2. Under H_0, the MLE of θ is $\hat{\theta_0} = 1$, and the MLE of ν is still $\hat{\nu} = x_{(1)}$. So the likelihood ratio statistic is

$$\begin{aligned}\lambda(x) &= \frac{x_{(1)}^n / (\prod_i x_i)^2}{(n/T)^n x_{(1)}^{n^2/T} / (\prod_i x_i)^{n/T+1}} \\ &= \left(\frac{T}{n}\right)^n \frac{e^{-T}}{(e^{-T})^{n/T}} \\ &= \left(\frac{T}{n}\right)^n e^{-T+n}. \\ \frac{\partial}{\partial T} \log \lambda(X) &= \left(\frac{n}{T} - 1\right).\end{aligned}$$

Hence, $\lambda(x)$ is increasing if $T \le n$ and decreasing if $T \ge n$. Thus, $T \le c$ is equivalent to $T \le c_1$ or $T \ge c_2$, for appropriately chosen constants c_1 and c_2.

3. We will not use the hint, although the problem can be solved that way. Instead, make the following three transformations. First, let $Y_i = \log X_i, i = 1, \ldots, n$. Next, make the n-to-1 transformation that sets $Z_1 = \min_i Y_i$ and sets $Z_2, \ldots, Z_n$ equal to the remaining Y_is, with their order unchanged. Finally, let $W_1 = Z_1$ and $W_i = Z_i - Z_1, i = 2, \ldots, n$. Then you find that the W_is are independent with $W_1 \sim f_{W_1}(w) = n\nu^n e^{-nw}, w > \log \nu$, and $W_i \sim exponential(1), i = 2, \ldots, n$. Now $T = \sum_{i=2}^{n} W_i \sim Gamma(n-1, 1)$, and, hence, $2T \sim gamma(n-1, 2) = \chi^2_{2(n-1)}$.

Solution 5

1. Suppose that we have two independent samples $X_1, \ldots, X_n$ are Exponential(θ), and $Y_1, \ldots, Y_n$ are Exponential(μ).

$$\begin{aligned}\lambda(x, y) &= \frac{sup_{\Theta_0} L(\theta|x, y)}{sup_{\Theta} L(\theta|x, y)} \\ &= \frac{\sup_\theta \prod_{i=1}^n \frac{1}{\theta} e^{-x_i/\theta} \prod_{j=1}^m \frac{1}{\theta} e^{-y_j/\theta}}{\sup_{\theta,\mu} \prod_{i=1}^n \frac{1}{\theta} e^{-x_i/\theta} \prod_{j=1}^m \frac{1}{\mu} e^{-y_j/\mu}}\end{aligned}$$

Differentiation will show that the numerator $\hat{\theta}_0 = \frac{\sum_i x_i + \sum_j y_j}{n+m}$, while in the denominator $\hat{\theta} = \bar{x}$ and $\hat{\mu} = \bar{y}$. Therefore,

$$\begin{aligned}\lambda(x,y) &= \frac{\left(\frac{n+m}{\sum_i x_i + \sum_j y_j}\right)^{n+m} \exp\left\{-\left(\frac{n+m}{\sum_i x_i + \sum_j y_j}\right)\left(\sum_i x_i + \sum_j y_j\right)\right\}}{\left(\frac{n}{\sum_i x_i}\right)^n \exp\left\{-\left(\frac{n}{\sum_i x_i}\right)\sum_i x_i\right\}\left(\frac{m}{\sum_j y_j}\right)^m \exp\left\{-\left(\frac{m}{\sum_j y_j}\right)\sum_j y_j\right\}} \\ &= \frac{(n+m)^{n+m}(\sum_i x_i)^n(\sum_j y_j)^m}{n^n m^m (\sum_i x_i + \sum_j y_j)^{n+m}}\end{aligned}$$

And the LRT is to reject H_0 if $\lambda(x,y) \le c$.

2. The LRT is given by,

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

Therefore λ is a function of T. λ is a unimodal function of T which is maximized when $T = \frac{n}{m+n}$. Rejection for $\lambda \le c$ is equivalent to rejection for $T \le a$ or $T \ge b$, where a and b are constants that satisfy $a^n(1-a)^m = b^n(1-b)^m$.

3. When H_0 is true, $\sum_i X_i \sim Gamma(n,\theta)$ and $\sum_j Y_j \sim Gamma(m,\theta)$ and they are independent. So by an extension of the following exercise, $T \sim Beta(n,m)$.

Example 1. *If $X_i, i = 1,2$ are independent Gamma$(\alpha_i, 1)$ random variables. Find the marginal distribution of $Y_1 = \frac{X_1}{X_1+X_2}$ and $Y_2 = \frac{X_2}{X_1+X_2}$.*

Proof: Make the transformation $y_1 = \frac{x_1}{x_1+x_2}, y_2 = x_1 + x_2$ then $x_1 = y_1 y_2, \quad x_2 = y_2(1-y_1)$ and $|J| = y_2$. Then

$$f(y_1,y_2) = \left[\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y_1^{\alpha_1-1}(1-y_1)^{\alpha_2-1}\right]\left[\frac{1}{\Gamma(\alpha_1+\alpha_2)} y_2^{\alpha_1+\alpha_2-1} e^{-y_2}\right]$$

thus $Y_1 \sim Beta(\alpha_1,\alpha_2), Y_2 \sim Gamma(\alpha_1+\alpha_2, 1)$ and are independent.

Solution 6

1. Let $Y = \sum_i X_i$. The posterior distribution of $\lambda|y$ is Gamma$(y+\alpha, \beta/(n\beta+1))$. [Verify]

2. (a)

$$\mathbb{P}\{\lambda \le \lambda_0 | y\} = \frac{(n\beta+1)^{y+\alpha}}{\Gamma(y+\alpha)\beta^{y+\alpha}} \int_0^{\lambda_0} t^{y+\alpha-1} e^{-t(n\beta+1)/\beta} dt.$$

$$\mathbb{P}\{\lambda > \lambda_0|y\} = 1 - \mathbb{P}\{\lambda \le \lambda_0|y\}.$$

(b) Because $\beta/(n\beta+1)$ is a scale parameter in the posterior distribution, $(2(n\beta+1)\lambda/\beta)|y$ has a Gamma$(y+\alpha, 2)$ distribution. If 2α is an integer, this is a $\chi^2_{2y+2\alpha}$ distribution. So, for $\alpha = 5/2$ and $\beta = 2$,

$$\begin{aligned}\mathbb{P}\{\lambda \le \lambda_0|y\} &= \mathbb{P}\left\{\frac{2(n\beta+1)\lambda}{\beta} \le \frac{2(n\beta+1)\lambda_0}{\beta}|y\right\} \\ &= \mathbb{P}\left\{\chi^2_{2y+5} \le 3\lambda_0\right\}.\end{aligned}$$

For given value of λ_0, we can apply chi-square table to get the answer.

Solution 7

1. For $H_0 : \mu \le 0$ vs. $H_1 : \mu > 0$ the LRT is to reject H0 if $\bar{x} > c\sigma/\sqrt{n}$ (Example 8.3.3). For $\alpha = 0.05$ take $c = 1.645$. The power function is

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

Note that the power will equal 0.5 when $\mu = 1.645\sigma/\sqrt{n}$.

2. For $H_0 : \mu = 0$ vs. $H_A : \mu \ne 0$ the LRT is to reject H0 if $|\bar{x}| > c\sigma/\sqrt{n}$ (Example 8.2.2). For $\alpha = 0.05$ take $c = 1.96$. The power function is

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

In this case, $\mu = \pm 1.96\sigma/\sqrt{n}$ gives power of approximately 0.5.

Solution 8

The pdf of Y is

$$f(y|\theta) = \frac{1}{\theta} y^{\frac{1}{\theta}-1} e^{-y^{\frac{1}{\theta}}} \quad ,y > 0$$

By the Neyman-Pearson Lemma, the UMP test will reject if

$$\frac{1}{2} y^{-\frac{1}{2}} e^{y-y^{1/2}} = \frac{f(y|2)}{f(y|1)} > k$$

To see the form of this rejection region, we compute

$$\frac{d}{dy}\left(\frac{1}{2} y^{-\frac{1}{2}} e^{y-y^{1/2}}\right) = \frac{1}{2} y^{-3/2} e^{y-y^{1/2}} \left(y - \frac{y^{1/2}}{2} - \frac{1}{2}\right)$$

which is negative for y < 1 and positive for y > 1. Thus $\frac{f(y|2)}{f(y|1)}$ is decreasing for

$y \leq 1$ and increasing for y > 1. Hence, rejecting for $\frac{f(y|2)}{f(y|1)} > k$ is equivalent to rejecting for $y \leq c_0$ or $y \geq c_1$. To obtain a size α test, the constants c_0 and c_1 must satisfy

$$\alpha = P(Y \leq c_0|\theta = 1) + P(Y \geq c_1|\theta = 1) = 1 - e^{-c_0} + e^{-c_1} \text{ and } \frac{f(c_0|2)}{f(c_0|1)} = \frac{f(c_1|2)}{f(c_1|1)}$$

Solving these two equations numerically, for α = .10, yields c_0 = .076546 and c_1 = 3.637798. The Type II error probability is

$$P(c_0 < Y < c_1|\theta = 2) = \int_{c_0}^{c_1} \frac{1}{2} y^{-1/2} e^{-y^{1/2}}\, dy == .609824.$$

Solution 9

By the Neyman-Pearson Lemma, the UMP test rejects for large values of $\frac{f(x|H_1)}{f(x|H_0)}$ Computing this ratio we obtain

x	1	2	3	4	5	6	7
$\frac{f(x\|H_1)}{f(x\|H_0)}$	6	5	4	3	2	1	0.84

The ratio is decreasing in x. So rejecting for large values of $\frac{f(x|H_1)}{f(x|H_0)}$ corresponds to rejecting for small values of x. To get a size test, we need to choose c so that $P(X \leq c|H_0) = \alpha$. The value c = 4 gives the UMP size α = .04 test. The Type II error probability is $P(X = 5, 6, 7|H_1) = .82$

Solution 10