Solutions to Assignment #12

1. Suppose that you observe n iid Bernoulli(p) random variables, denoted by $X_1, X_2, \ldots, X_n$. Find the LRT rejection region for the test of $\mathrm{H}_o\colon p \leqslant p_o$ versus $\mathrm{H}_1\colon p > p_o$ in terms of the test statistic $Y = \displaystyle\sum_{i=1}^{n} X_i$.

   ***Solution**:* The likelihood ratio statistic is

$$\Lambda(x_1, x_2, \ldots, x_n) = \frac{\displaystyle\sup_{p \leqslant p_o} L(p \mid x_1, x_2, \ldots, x_n)}{L(\widehat{p} \mid x_1, x_2, \ldots, x_n)},$$

   where

$$L(p \mid x_1, x_2, \ldots, x_n) = p^y(1-p)^{n-y}, \quad \text{for} \quad y = \sum_{i=1}^{n} x_i,$$

   is the likelihood function, and

$$\widehat{p} = \frac{1}{n} y = \overline{x}$$

   is the MLE for p.

   Observe that if $p_o \geqslant \widehat{p}$, then

$$\sup_{p \leqslant p_o} L(p \mid x_1, x_2, \ldots, x_n) = L(\widehat{p} \mid x_1, x_2, \ldots, x_n)$$

   and so $\Lambda(x_1, x_2, \ldots, x_n) = 1$; thus, in this case we would not get a rejection region for the LRT,

$$R\colon \quad \Lambda(x_1, x_2, \ldots, x_n) \leqslant c,$$

   for some $0 < c < 1$. We therefore have that $p_o < \widehat{p}$ so that

$$\begin{aligned} \Lambda(x_1, x_2, \ldots, x_n) &= \frac{L(p_o \mid x_1, x_2, \ldots, x_n)}{L(\widehat{p} \mid x_1, x_2, \ldots, x_n)} \\ &= \frac{p_o^y(1-p_o)^{n-y}}{\widehat{p}^y(1-\widehat{p}_1)^{n-y}}, \end{aligned}$$

where $\widehat{p} > p_o$, which can in turn be written as

$$\Lambda(x_1, x_2, \ldots, x_n) = \frac{1}{\left(\dfrac{\widehat{p}}{p_o}\right)^y} \left(\frac{\dfrac{1}{p_o} - 1}{\dfrac{1}{p_o} - \dfrac{\widehat{p}}{p_o}}\right)^{n-y}.$$

Setting $t = \dfrac{\widehat{p}}{p_o}$, we see that Λ can be written as a function of t as follows

$$\Lambda(t) = \begin{cases} 1 & \text{for } 0 \leqslant t \leqslant 1, \\ \dfrac{1}{t^{np_ot}} \cdot \left(\dfrac{1-p_o}{1-p_ot}\right)^{n-np_ot}, & \text{for } 1 < t \leqslant \frac{1}{p_o}, \end{cases}$$

since $\widehat{p} > p_o$, where we have used the fact that $\widehat{p} = \dfrac{1}{n}y$ so that $y = np_ot$.

The graph of $\Lambda(t)$ can be shown to be like the one sketched as the one shown in Figure 1 on page 3, where we have sketched the case $p_o = 1/4$ and $n = 20$ for $0 \leqslant t \leqslant 4$. The sketch in Figure 1 shows that $\Lambda(t)$ decreases for $t > 1$; thus, given any positive value of c such that $c < 1$ and $c > p_o^n$, there exists a positive value t_2 such that $t_2 > 1$, and

$$\Lambda(t_2) = c,$$

and

$$\Lambda(t) \leqslant c \quad \text{for} \quad t \geqslant t_2.$$

Thus, the LRT rejection region for the test of $\mathrm{H}_o\colon\ p \leqslant p_o$ versus $\mathrm{H}_1\colon\ p > p_o$ is equivalent to

$$\frac{\widehat{p}}{p_o} \geqslant t_2,$$

which we could rephrase in terms of $Y = \displaystyle\sum_{i=1}^n X_i$ as

$$R\colon \quad Y \geqslant t_2 np_o,$$

for some t_2 with $t_2 > 1$. This rejection region can also be phrased as

$$R\colon \quad Y > np_o + b,$$

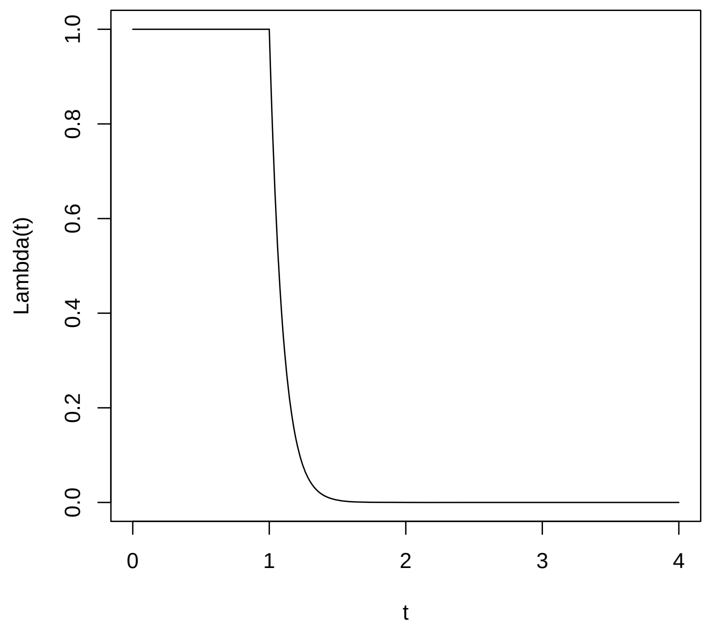

Figure 1: Sketch of graph of $\Lambda(t)$ for $p_o = 1/4$, $n = 20$, and $0 \leqslant t \leqslant 4$

for some $b > 0$. The value of b will then be determined by the significance level that we want to impose on the test, and Y is the statistic

$$Y = \sum_{i=1}^{n} X_i,$$

which counts the number of successes in the sample. $\square$

2. Consider the likelihood ratio test for $\text{H}_o\colon\ p = p_o$ versus $\text{H}_1\colon\ p = p_1$, where $p_o \neq p_1$, based on a random sample $X_1, X_2, \ldots, X_n$ from a Bernoulli(p) distribution for $0 < p < 1$. Show that, if $p_1 > p_o$, then the likelihood ratio statistic for the test is is a monotonically decreasing function of $Y = \sum_{i=1}^{n} X_i$. . Conclude, therefore, that if the test rejects H_o at the significance level α for an observed value $\widehat{y}$ of Y, it will also rejects H_o at that level for $Y > \widehat{y}$.

***Solution**:* The likelihood ratio statistic in this case is

$$\Lambda(x_1, x_2, \ldots, x_n) = \frac{p_o^y(1-p_o)^{n-y}}{p_1^y(1-p_1)^{n-y}}, \quad \text{for} \ \ y = \sum_{i=1}^{n} x_i,$$

which can be written as

$$\Lambda(x_1, x_2, \ldots, x_n) = a^n r^y,$$

where

$$a = \frac{1-p_o}{1-p_1} > 0 \quad \text{and} \quad r = \frac{p_o(1-p_1)}{p_1(1-p_o)} < 1,$$

since $p_1 > p_o$. It then follows that the likelihood ratio statistic for the test is is a monotonically decreasing function of $Y = \sum_{i=1}^{n} X_i$, since $r < 1$.

Suppose now that the test rejects H_o at the significance level α for an observed value $\widehat{y}$ of Y; that is

$$\Lambda(\widehat{y}) \leqslant c,$$

for some c in $(0, 1)$ determined by α. Then, for any value of y which is bigger than $\widehat{y}$,

$$\Lambda(y) < \Lambda(\widehat{y}) \leqslant c, \tag{1}$$

since Λ is a decreasing function of y. It then follows from (1) that the LRT will also rejects H_o at that level for $Y > \widehat{y}$. $\square$

3. We wish to use an LRT to test the hypothesis $\mathrm{H}_o\colon\ \mu = \mu_o$ against the alternative $\mathrm{H}_1\colon\ \mu \neq \mu_o$ based on a random sample, $X_1, X_2, \ldots, X_n$, from a normal$(\mu, 1)$ distribution.

   (a) Give the maximum likelihood estimator, $\widehat{\mu}$, for μ based on the sample.

   Solution: The likelihood function in this problem is

$$L(\mu \mid x_1, x_2, \ldots, x_n) = \frac{e^{-\sum_{i=1}^{n}(x_i-\mu)^2/2}}{(2\pi)^{n/2}} \quad \text{for } \mu \in \mathbb{R}.$$

   To find the MLE for μ, it suffices to maximize the natural logarithm of the likelihood function

$$\ell(\mu) = -\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2 - \frac{n}{2}\ln(2\pi), \quad \text{for } \ \mu \in \mathbb{R}.$$

   Taking derivatives we get

$$\ell'(\mu) = \sum_{i=1}^{n}(x_i - \mu) = \sum_{i=1}^{n} x_i - n\mu,$$

   and

$$\ell''(\mu) = \sum_{i=1}^{n}(x_i - \mu) = -n < 0.$$

   It then follows that $\widehat{\mu} = \dfrac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$ is the only critical point of ℓ and $\ell''(\widehat{\mu}) < 0$. Thus, the likelihood function is maximized at $\widehat{\mu} = \overline{x}$, the sample mean. □

   (b) Give the likelihood ratio statistic for the test.

   Solution: The likelihood ratio statistic is

$$\Lambda(x_1, x_2, \ldots, x_n) = \frac{L(\mu_o \mid x_1, x_2, \ldots, x_n)}{L(\widehat{\mu} \mid x_1, x_2, \ldots, x_n)},$$

   where $\widehat{\mu} = \overline{x}$ is the MLE for μ. We then have that

$$\Lambda(x_1, x_2, \ldots, x_n) = \frac{e^{-\sum_{i=1}^{n}(x_i-\mu_o)^2/2}}{e^{-\sum_{i=1}^{n}(x_i-\overline{x})^2/2}}, \tag{2}$$

where

$$\begin{aligned}\sum_{i=1}^{n}(x_i-\mu_o)^2 &= \sum_{i=1}^{n}(x_i-\overline{x}+\overline{x}-\mu_o)^2\\ &= \sum_{i=1}^{n}(x_i-\overline{x})^2+\sum_{i=1}^{n}(\overline{x}-\mu_o)^2\\ &= \sum_{i=1}^{n}(x_i-\overline{x})^2+n(\overline{x}-\mu_o)^2,\end{aligned}$$

since

$$\sum_{i=1}^{n}2(x_i-\overline{x})(\overline{x}-\mu_o)=2(\overline{x}-\mu_o)\sum_{i=1}^{n}(x_i-\overline{x})=0.$$

Hence, from (2) we have that

$$\Lambda(x_1,x_2,\ldots,x_n)=e^{-n(\overline{x}-\mu_o)^2/2}, \tag{3}$$

where $\overline{x}$ denotes the sample mean. □

(c) Express the LRT rejection region in terms of the sample mean $\overline{X}_n$.

***Solution**:* The LRT rejection region is given by

$$R\colon\quad \Lambda(x_1,x_2,\ldots,x_n)\leqslant c,$$

for some c with $0<c<1$. It then follows from equation (3) in part (b) of this problem that H_o is rejected if

$$e^{-n(\overline{x}-\mu_o)^2/2}\leqslant c,$$

or, taking natural logarithm on both sides of the last inequality,

$$-n(\overline{x}-\mu_o)^2/2\leqslant \ln c,$$

or

$$n(\overline{x}-\mu_o)^2\geqslant -2\ln c,$$

or

$$\sqrt{n}|\overline{x}-\mu_o|\geqslant\sqrt{-2\ln c}\equiv b>0.$$

Thus, the LRT will reject H_o if

$$\sqrt{n}|\overline{X}_n-\mu_o|\geqslant b,$$

for some $b>0$ determined by the significance level α. □

4. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a uniform$(0, \theta)$ distribution for some parameter $\theta > 0$.

   (a) Give the likelihood function $L(\theta \mid x_1, x_2, \ldots, x_n)$.

   Solution: The pdf for each of the X_is is

$$f(x \mid \theta) = \begin{cases} \dfrac{1}{\theta} & \text{if } 0 < x < \theta; \\ \\ 0 & \text{otherwise.} \end{cases}$$

   Thus, the likelihood function is

$$L(\theta \mid x_1, x_2, \ldots, x_n) = \begin{cases} \dfrac{1}{\theta^n} & \text{if } 0 < x_1, x_2, \ldots, x_n < \theta; \\ \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

   □

   (b) Give the maximum likelihood estimator for θ.

   Solution: Observe that the likelihood function in (4) is a decreasing function of θ. Thus, $L(\theta \mid x_1, x_2, \ldots, x_n)$ will the largest when θ is the smallest value it can take, $\widehat{\theta}$, based on the sample. This value is the maximum of the values $x_1, x_2, \ldots, x_n$, because, if $x_i > \theta$ for some i, then $L(\theta \mid x_1, x_2, \ldots, x_n) = 0$ according to the definition of the likelihood function given (4). It then follows that the MLE for θ is

$$\widehat{\theta} = \max\{X_1, X_2, \ldots, X_n\}.$$

   □

5. Let R denote the rejection region for an LRT of $\mathrm{H}_o\colon\ \theta = \theta_o$ versus $\mathrm{H}_1\colon\ \theta = \theta_1$ based on a random sample, $X_1, X_2, \ldots, X_n$, from continuous distribution with pdf $f(x \mid \theta)$. Let $L(\theta \mid x_1, x_2, \ldots, x_n)$ denote the likelihood function. Suppose the LRT has significance level α.

   (a) Explain why

$$\alpha = \int_R L(\theta_o \mid x_1, x_2, \ldots, x_n)\ \mathrm{d}x_1\ \mathrm{d}x_2 \cdots\ \mathrm{d}x_n.$$

Answer: The significance level α is the probability the the LRT will reject H_o when H_o is true; in other words, when $\theta = \theta_o$. Thus,

$$\begin{aligned}
\alpha &= \text{P}((x_1, x_2, \ldots, x_n) \in R \\
&= \int_R f(x_1, x_2, \ldots, x_n \mid \theta_o) \ \, \mathrm{d}x_1 \, \mathrm{d}x_2 \cdots \, \mathrm{d}x_n,
\end{aligned}$$

where R is the rejection region of the LRT and $f(x_1, x_2, \ldots, x_n \mid \theta_o)$ is the joint distribution of the sample for the case in which H_o is true. It then follows that

$$\alpha = \int_R L(\theta_o \mid x_1, x_2, \ldots, x_n) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \cdots \, \mathrm{d}x_n,$$

by the definition of the likelihood function. □

(b) Explain why the power of the test is

$$\gamma(\theta_1) = \int_R L(\theta_1 \mid x_1, x_2, \ldots, x_n) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \cdots \, \mathrm{d}x_n.$$

Answer: The power of the test, $\gamma(\theta_1)$, is the probability the the LRT will reject H_o when H_o is false; in other words, when $\theta = \theta_1$. Thus,

$$\gamma(\theta_1) \ = \ \textstyle\int_R f(x_1, x_2, \ldots, x_n \mid \theta_1) \ \, \mathrm{d}x_1 \, \mathrm{d}x_2 \cdots \, \mathrm{d}x_n,$$

which yields

$$\gamma(\theta_1) = \int_R L(\theta_1 \mid x_1, x_2, \ldots, x_n) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \cdots \, \mathrm{d}x_n.$$

by the definition of the likelihood function. □

(c) Explain why

$$\alpha \leqslant c\gamma(\theta_1),$$

where c is the critical value used in the definition of the rejection region, R, for the LRT.

Solution: Since, $\Lambda(x_1, x_2, \ldots, x_n) \leqslant c$ on R, , for some $c \in (0, 1)$ determined by α, it follows that

$$L(\theta_0 \mid x_1, x_2, \ldots, x_n) \leqslant cL(\theta_1 \mid x_1, x_2, \ldots, x_n)$$

for all $(x_1, x_2, \ldots, x_n) \in R$. Consequently,

$$\int_R L(\theta_0 \mid x_1, x_2, \ldots, x_n) \leqslant \int_R cL(\theta_1 \mid x_1, x_2, \ldots, x_n)$$

from which the result follows in view of parts (a) and (b) of this problem. $\square$