20170058 **Keonwoo Kim** · June 7, 2020
MAS555 Advanced Statistics

Homework 5

8.12. (a) The likelihood ratio test statistic is given by

$$\begin{aligned}\lambda(\mathbf{x}) &= \frac{\sup_{\mu\le 0}\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \prod_{i=1}^n \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]}{\sup_{\mu\in\mathbb{R}}\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \prod_{i=1}^n \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]} \\ &= \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n\left[(x_i-\hat{\mu})^2-(x_i-\bar{x})\right]\right]\end{aligned}$$

where

$$\hat{\mu}=\hat{\mu}(\mathbf{x})=\min(\bar{x},0)=\begin{cases}\bar{x}, & \bar{x}\le 0\\ 0, & \bar{x}>0\end{cases}.$$

Therefore,

$$\lambda(\mathbf{x})=\begin{cases}1, & \bar{x}\le 0\\ \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n\left[x_i^2-(x_i-\bar{x})\right]\right]=\exp\left[-\frac{n\bar{x}^2}{2\sigma^2}\right], & \bar{x}>0\end{cases}.$$

So, the LRT has the following rejection region: for $0\le c'<1$,

$$\begin{aligned}\{\mathbf{x}:\lambda(\mathbf{x})\le c'\} &= \left\{\mathbf{x}:\bar{x}>0 \text{ and } \exp\left[-\frac{n\bar{x}^2}{2\sigma^2}\right]\le c'\right\}\\ &= \left\{\mathbf{x}:\bar{x}>0 \text{ and } \left|\frac{\bar{x}}{\sigma/\sqrt{n}}\right|\ge\sqrt{-2\log c'}\right\}\\ &= \left\{\mathbf{x}:\frac{\bar{x}}{\sigma/\sqrt{n}}\ge\sqrt{-2\log c'}=:c\right\}.\end{aligned}$$

Then, the power function is as follows:

$$\begin{aligned}\beta(\mu) &= P\left(\frac{\bar{X}}{\sigma/\sqrt{n}}\ge c\right)\\ &= P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\ge c-\frac{\mu}{\sigma/\sqrt{n}}\right)\\ &= 1-\Phi\left(c-\frac{\mu}{\sigma/\sqrt{n}}\right)\end{aligned}$$

where $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathrm{n}(0,1)$. Since the size of the test is $\alpha = .05$, we have

$$\sup_{\mu \leq 0} \beta(\mu) = 1 - \Phi(c) = .05 \implies c = z_{.05} \approx 1.645.$$

Finally,

$$\beta(\mu) = 1 - \Phi\left(z_{.05} - \frac{\mu}{\sigma/\sqrt{n}}\right) = 1 - \Phi\left(z_{.05} - \frac{\mu}{\sigma/\sqrt{n}}\right)$$

(b) The likelihood ratio test statistic is given by

$$\begin{aligned}
\lambda(\mathbf{x}) &= \frac{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \prod_{i=1}^n \exp\left[-\frac{x_i^2}{2\sigma^2}\right]}{\sup_{\mu \in \mathbb{R}} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \prod_{i=1}^n \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]} \\
&= \exp\left[-\frac{n\bar{x}^2}{2\sigma^2}\right]
\end{aligned}$$

whence the rejection region is of the following form:

$$\begin{aligned}
\left\{\mathbf{x} : \lambda(\mathbf{x}) \leq c'\right\} &= \left\{\mathbf{x} : \exp\left[-\frac{n\bar{x}^2}{2\sigma^2}\right] \leq c'\right\} \\
&= \left\{\mathbf{x} : \left|\frac{\bar{x}}{\sigma/\sqrt{n}}\right| \geq \sqrt{-2\log c'}\right\}.
\end{aligned}$$

Then, the power function is as follows:

$$\begin{aligned}
\beta(\mu) &= P\left(\left|\frac{\bar{X}}{\sigma/\sqrt{n}}\right| \geq c\right) \\
&= P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \geq c - \frac{\mu}{\sigma/\sqrt{n}}\right) + P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq -c - \frac{\mu}{\sigma/\sqrt{n}}\right) \\
&= 1 - \Phi\left(c - \frac{\mu}{\sigma/\sqrt{n}}\right) + \Phi\left(-c - \frac{\mu}{\sigma/\sqrt{n}}\right)
\end{aligned}$$

where $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathrm{n}(0,1)$. Since the size of the test is $\alpha = .05$, we have

$$\beta(0) = 1 - \Phi(c) + \Phi(-c) = .05 \implies c = z_{.025} \approx 1.960.$$

Finally,

$$\beta(\mu) = 1 - \Phi\left(z_{.025} - \frac{\mu}{\sigma/\sqrt{n}}\right) + \Phi\left(-z_{.025} - \frac{\mu}{\sigma/\sqrt{n}}\right).$$

8.14. The CLT implies that

$$Z \coloneqq \frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \dot\sim \mathrm{n}(0,1)$$

where $\dot\sim$ means the asymptotic distribution as $n \to \infty$. We have a test that rejects H_0 if $\sum X_i > c$. For the type I error,

$$\begin{aligned}\beta(.49) &= P\left(\sum_{i=1}^n X_i > c \,\middle|\, p = .49\right) \\ &= P\left(Z > \frac{c - np}{\sqrt{np(1-p)}} \,\middle|\, p = .49\right) \\ &= 1 - \Phi\left(\frac{c - .49\,n}{\sqrt{.49 \times .51\,n}}\right) \le .01;\end{aligned}$$

and for the type II error,

$$\begin{aligned}1 - \beta(.51) &= P\left(\sum_{i=1}^n X_i \le c \,\middle|\, p = .51\right) \\ &= P\left(Z \le \frac{c - np}{\sqrt{np(1-p)}} \,\middle|\, p = .51\right) \\ &= \Phi\left(\frac{c - .51\,n}{\sqrt{.49 \times .51\,n}}\right) \le .01.\end{aligned}$$

These imply

$$\frac{c - .49\,n}{\sqrt{.49 \times .51}} \ge z_{.01} \approx 2.32635, \qquad \frac{c - .51\,n}{\sqrt{.49 \times .51\,n}} \le -z_{.01} \approx -2.32635.$$

With $c = n/2$, we have

$$\frac{.01\,n}{\sqrt{.49 \times .51\,n}} \ge z_{.01} \implies n \ge 13525.$$

Taking $n = 13525$, c becomes 6762.5.

8.18. (a)

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

$$= P\left(\frac{\bar{X}-\theta}{\sigma/\sqrt{n}} > c + \frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right) + P\left(\frac{\bar{X}-\theta}{\sigma/\sqrt{n}} < -c + \frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right)$$
$$= 1 - \Phi\left(c + \frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right) + \Phi\left(-c + \frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right).$$

(b) The type I error probability is

$$\beta(\theta_0) = 1 - \Phi(c) + \Phi(-c) = .05$$

so that $c = z_{.025} \approx 1.960$. The type II error probability at $\theta = \theta_0 + \sigma$ is

$$1 - \beta(\theta_0+\sigma) = 1 - \Phi(z_{.025} - \sqrt{n}) + \Phi(-z_{.025} - \sqrt{n}).$$

We may assume that $\Phi(-z_{.025} - \sqrt{n})$ is negligible so that

$$1 - \Phi(z_{.025} - \sqrt{n}) \le .25 \implies z_{.025} - \sqrt{n} \le 0.67449 \implies n \ge 7.$$

8.20. Choose $k = 2.5$ in the Neyman–Pearson lemma, where the test rejects H_0 if $f(x|H_1)/f(x|H_0) > k$. Since

x	1	2	3	4	5	6	7
$f(x\|H_1)/f(x\|H_0)$	6	5	4	3	2	1	79/94

,

the rejection region is $R = \{1,2,3,4\}$ and the size of the test is $\alpha = P(X \in R \mid H_0) = .04$. The type II error probability is

$$P(X \notin R \mid H_1) = P(X \in \{5,6,7\}) = .02 + .01 + .79 = .82.$$

8.22. (a) Neyman–Pearson lemma says the most powerful test is a test which rejects H_0 iff

$$\frac{f(\mathbf{x} \mid p = \frac{1}{4})}{f(\mathbf{x} \mid p = \frac{1}{2})} = \left(\frac{3}{2}\right)^{10}\left(\frac{1}{3}\right)^{\sum_{i=1}^{10} X_i} > k'$$

for some k'. Since the likelihood ratio above is decreasing in $\sum_{i=1}^{10} X_i$, the most powerful test is a test which rejects H_0 iff $\sum_{i=1}^{10} X_i \le k$ for some k. Observing that

$$P\left(\sum_{i=1}^{10} X_i \le k \,\middle|\, X_i \sim \text{Bernoulli}\left(\frac{1}{2}\right)\right)$$

$$= \alpha = .0547 = P\left(X \leq 2 \,\middle|\, X \sim \text{Binomial}\left(10, \frac{1}{2}\right)\right),$$

$k = 2$ is a valid choice. So the rejection region is

$$\left\{\mathbf{x} : \sum_{i=1}^{10} X_i \leq 2\right\}.$$

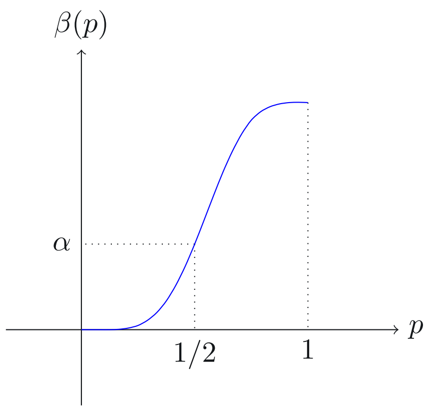
Therefore, the power function β satisfies $\beta(1/2) = 0.0547$ and

$$\beta\left(\frac{1}{4}\right) = P\left(X \leq 2 \,\middle|\, X \sim \text{Binomial}\left(10, \frac{1}{4}\right)\right) = 0.5256.$$

(b) The power function β is

$$\beta(p) = P\left(\sum_{i=1}^{10} X_i \geq 6\right) = P\left(X \geq 6 \,\middle|\, X = \sum_{i=1}^{10} X_i \sim \text{Binomal}(10, p)\right)$$

and the graph of β is as follows:



The size of the test is

$$\alpha = \sup_{p \leq 1/2} \beta(p) = \beta\left(\frac{1}{2}\right) = \frac{193}{512} \approx 0.37695.$$

(c) Neyman–Pearson lemma tells us that the rejection region is of the form $\left\{\mathbf{x} : \sum_{i=1}^{10} X_i \leq k\right\}$ as in (a). Thus, α should be matched with one of the rejection region above:

$$\alpha = \alpha(k) = P\left(\sum_{i=1}^{10} X_i \le k \,\middle|\, p = \frac{1}{2}\right),$$

where $k = -1, 0, \ldots, 10$ and

$$\begin{aligned}
&\alpha(-1) = 0, \quad \alpha(0) = \frac{1}{1024}, \quad \alpha(1) = \frac{11}{1024}, \quad \alpha(2) = \frac{56}{1024}, \\
&\alpha(3) = \frac{176}{1024}, \quad \alpha(4) = \frac{386}{1024}, \quad \alpha(5) = \frac{638}{1024}, \quad \alpha(6) = \frac{848}{1024}, \\
&\alpha(7) = \frac{968}{1024}, \quad \alpha(8) = \frac{1013}{1024}, \quad \alpha(9) = \frac{1023}{1024}, \quad \alpha(10) = 1.
\end{aligned}$$

Those values are the desired values for α.

8.28. (a) If $\theta_1 < \theta_2$,

$$\begin{aligned}
\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} &= \frac{e^{x-\theta_2}/e^{x-\theta_1}}{\left((1+e^{x-\theta_2})/(1+e^{x-\theta_1})\right)^2} \\
&= e^{\theta_1-\theta_2}\left(\frac{e^{-x}+e^{-\theta_1}}{e^{-x}+e^{-\theta_2}}\right)^2
\end{aligned}$$

is nondecreasing in x. Thus, the logistic location family has an MLR.

(b) The most powerful test rejects H_0 if $\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} > k'$ for some k'. Since $\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)}$ is nondecreasing in x, it is equivalent to $x > k$ for some k. Then the size of this test is

$$\alpha = P\left(X > k \mid \theta = 0\right) = \int_k^\infty \frac{e^x}{(1+e^x)^2}\,dx = \frac{1}{e^k+1},$$

hence $k = \log 4$. Then the size of the type II error is

$$\beta = P\left(X \le k \mid \theta = 1\right) = \int_{-\infty}^{\log 4} \frac{e^{x-1}}{(1+e^{x-1})^2}\,dx = \frac{4}{4+e}.$$

(c) X is sufficient for θ by definition of the sufficiency. Also, the logistic location family has an MLR in X. Therefore, for any k, Karlin–Rubin theorem (Theorem 8.3.17 in Casella–Berger) implies that the test rejecting $H_0\colon \theta \le 0$ iff $X > k$ is a UMP level α test, where $\alpha = P\left(X > k \mid \theta = 0\right)$.