MATH 503: Mathematical Statistics
Dr. Kimberly F. Sellers, Instructor
Homework 6 Solutions

1. **Let X have a pdf of the form $f(x;\theta) = \theta x^{\theta-1}, 0 < x < 1$, zero elsewhere, where $\theta \in \{\theta : \theta = 1, 2\}$. To test the simple hypothesis $H_0 : \theta = 1$ against the alternative simple hypothesis $H_1 : \theta = 2$, use the random sample X_1, X_2 of size $n = 2$ and define the critical region $C = \left\{(x_1, x_2) : \frac{3}{4} \le x_1 x_2\right\}$. Find the power function of the test.**

   Solution: We consider the hypothesis test $H_0 : \theta = 1$ vs $H_1 : \theta = 2$ where $f(x;\theta) = \theta x^{\theta-1}, 0 < x < 1$ is the pdf and $C = \left\{(x_1, x_2) : \frac{3}{4} \le x_1 x_2\right\}$ is the critical region (i.e. the rejection region). Thus, the power function is

$$
\begin{aligned}
P_\theta\left(x_1 x_2 \ge \frac{3}{4}\right) &= \int_{x_1 x_2 \ge \frac{3}{4}} f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 \\
&= \int_{3/4}^{1}\int_{3/4x_2}^{1} \theta^2 x_1^{\theta-1} x_2^{\theta-1}\mathrm{d}x_1\mathrm{d}x_2 = \int_{3/4}^{1} \theta x_2^{\theta-1}\left(\int_{3/4x_2}^{1} \theta x_1^{\theta-1}\mathrm{d}x_1\right)\mathrm{d}x_2 \\
&= \int_{3/4}^{1} \theta x_2^{\theta-1}\left(x_1^{\theta} \mid_{3/4x_2}^{1}\right)\mathrm{d}x_2 = \int_{3/4}^{1} \theta x_2^{\theta-1}\left[1 - \left(\frac{3}{4}\right)^{\theta}\left(\frac{1}{x_2}\right)^{\theta}\right]\mathrm{d}x_2 \\
&= \int_{3/4}^{1} \theta x_2^{\theta-1}\mathrm{d}x_2 - \left(\frac{3}{4}\right)^{\theta}\theta\int_{3/4}^{1}\frac{1}{x_2}\mathrm{d}x_2 \\
&= x_2^{\theta} \mid_{3/4}^{1} - \left(\frac{3}{4}\right)^{\theta}\theta\ln(x_2)\mid_{3/4}^{1} = 1 - \left(\frac{3}{4}\right)^{\theta} - \theta\left(\frac{3}{4}\right)^{\theta}(\ln 1 - \ln(3/4)) \\
&= 1 - \left(\frac{3}{4}\right)^{\theta} + \theta\left(\frac{3}{4}\right)^{\theta}(\ln(3/4))
\end{aligned}
$$

   So, for $\theta = 1$, the power function equals 0.03434 (which is the significance level), and for $\theta = 2$, the power function equals 0.11386 (which is the power of the test).

2. **Let us say the life of a tire in miles, say X, is normally distributed with mean θ and standard deviation 5000. Past experience indicates that $\theta = 30{,}000$. The manufacturer claims that the tires made by a new process have mean $\theta > 30{,}000$. It is possible that $\theta = 35{,}000$. Check his claim by testing $H_0 : \theta = 30{,}000$ against $H_1 : \theta > 30{,}000$. We shall observe n independent values of X, say $x_1, \ldots, x_n$, and we shall reject H_0 (thus accept H_1) if and only if $x \ge c$. Determine n and c so that the power function $\gamma(\theta)$ of the test has the values $\gamma(30{,}000) = 0.01$ and $\gamma(35{,}000) = 0.98$.**

   Solution: Let $X \sim \mathrm{N}(\theta, \sigma = 5000)$ and consider the hypotheses, $H_0 : \theta = 30{,}000$ vs $H_1 : \theta > 30{,}000$.

The rejection region is $\bar{X} \geq c$ where $\gamma(30,000) = 0.01$ and $\gamma(35,000) = 0.98$, where

$$
\begin{aligned}
\alpha &= \gamma(30,000) = P_{\theta=30,000}(\bar{X} \geq c) = P_{\theta=30,000}\left(\frac{\bar{X}-30,000}{5000/\sqrt{n}} \geq \frac{c-30,000}{5000/\sqrt{n}}\right) \\
&= P_{\theta=30,000}\left(Z \geq \frac{c-30,000}{5000/\sqrt{n}}\right) = 0.01 \\
&\Rightarrow \frac{c-30,000}{5000/\sqrt{n}} = 2.33 \text{ (using chart or R) and} \\
\text{Power} &= \gamma(35,000) = P_{\theta=35,000}(\bar{X} \geq c) = P_{\theta=35,000}\left(\frac{\bar{X}-35,000}{5000/\sqrt{n}} \geq \frac{c-35,000}{5000/\sqrt{n}}\right) \\
&= P_{\theta=35,000}\left(Z \geq \frac{c-35,000}{5000/\sqrt{n}}\right) = 0.98 \\
&\Rightarrow \frac{c-35,000}{5000/\sqrt{n}} = -2.055 \text{ (using chart or R)}
\end{aligned}
$$

thus, we want to find n, c that solves the system of equations

$$\frac{c-30,000}{5000/\sqrt{n}} \approx 2.328 \qquad \text{and} \qquad \frac{c-35,000}{5000/\sqrt{n}} \approx -2.058.$$

In solving the equations, resulting answers will vary (due to various contributing factors – approximation of the z-score, roundoff error, etc). Answers resulting in $n \approx 19$ will produce c anywhere around 32639.312 to 32670.399. Answers resulting in $n \approx 20$ will produce c around 32602.78 to 32699.086.

The "exact" answers for n, c are $n \approx 19.2370$ and $capprox32653.8988$. Recall, however, that n denotes a sample size; technically speaking, it doesn't make sense to have a "fraction/decimal part" of a specimen, thus n should be an integer. One can debate which value for n is more appropriate.

3. **Let $Y_1 < Y_2 < Y_3 < Y_4$ be the order statistics of a random sample X_1, X_2, X_3, X_4 of size $n = 4$ from a distribution with pdf $f(x;\theta) = \frac{1}{\theta}, 0 < x < \theta$, zero elsewhere, where $\theta > 0$. The hypothesis $H_0 : \theta = 1$ is rejected and $H_1 : \theta > 1$ is accepted if the observed $Y_4 \geq c$.**

   (a) Find the constant c so that the significance level is $\alpha = 0.05$.

   (b) Determine the power function of the test.

Solution: $Y_1 < Y_2 < Y_3 < Y_4$ are order statistics of a random sample X_1, X_2, X_3, X_4 with pdf $f(x;\theta) = \frac{1}{\theta}, 0 < x < \theta$; $F(x) = \frac{x}{\theta}$.

(a) $F_{Y_4}(y) = P(Y_4 \leq y) = F^4(y) = \left(\frac{y}{\theta}\right)^4$ and $f_{Y_4}(y) = \frac{4y^3}{\theta^4}$, thus

$$
\begin{aligned}
\alpha = 0.05 = P_{\theta=1}(Y_4 \geq c) = 1 - P_{\theta=1}(Y_4 \leq c) &= 1 - c^4 \\
c^4 &= 1 - \alpha = 0.95 \\
c &= \sqrt[4]{0.95}
\end{aligned}
$$

(b)

$$\gamma(\theta) = P_\theta(Y_4 \geq \sqrt[4]{0.95}) = 1 - P_\theta(Y_4 \leq \sqrt[4]{0.95}) = 1 - \left(\frac{\sqrt[4]{0.95}}{\theta}\right)^4 = 1 - \frac{0.95}{\theta^4}$$

4. **Assume that the weight of cereal in a "10-ounce box" is $\mathbf{N}(\mu, \sigma^2)$. To test $H_0 : \mu = 10.1$ against $H_1 : \mu > 10.1$, we take a random sample of size $n = 16$ and observe $x = 10.4$ and $s = 0.4$.**

(a) Do we reject or fail to reject H_0 at the 5% significance level?

(b) What is the approximate p-value of this test?

Solution: $H_0 : \mu = 10.1$ vs. $H_1 : \mu > 10.1$. The test statistic is

$$T = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{10.4 - 10.1}{0.4/\sqrt{16}} = 3 \text{ with 15 df}$$

(a) The critical region is $\{T : T > 1.753\}$, thus reject H_0 because $3 > 1.753$.

(b) The p-value is $P(T > 3) \approx 0.005$ with 15 df.

5. **Each of 51 golfers hit three golf blls of brand X and three golf balls of brand Y in a random order. Let X_i and Y_i equal the averages of the distances traveled by the brand X and brand Y golf balls hit by the ith golfer, $i = 1, 2, \ldots, 51$. Let $W_i = X_i - Y_i$, $i = 1, 2, \ldots, 51$ and test $H_0 : \mu_W = 0$ against $H_1 : \mu_W > 0$, where μ_W is the mean of the differences. If $\bar{W} = 2.07$ and $s_W^2 = 84.63$, would H_0 be rejected at the 5% significance level? What is the p-value of this test?**

Solution: $n = 51$, $\bar{W} = 2.07$ and $s_W^2 = 84.63$; $H_0 : \mu_W = 0$ vs $H_1 : \mu_W > 0$, and $\alpha = 0.05$. Reject H_0 if

$$Z = \frac{\bar{W} - 0}{s_W/\sqrt{n}} \geq 1.645$$

The test statistic is

$$Z = \frac{2.07}{\sqrt{84.63/51}} \approx 1.607,$$

which is less than 1.645, inferring that we fail to reject H_0 at the 5% significance level. The p-value is $P(Z > 1.607) \approx P(Z > 1.61) = 1 - 0.9463 = 0.0537$.

6. **Let the random variable X have the pdf $f(x;\theta) = \frac{1}{\theta}e^{-x/\theta}$, $0 < x < \infty$, zero elsewhere. Consider the simple hypothesis $H_0 : \theta = \theta' = 2$ and the alternative hypothesis $H_1 : \theta = \theta'' = 4$. Let X_1, X_2 denote a random sample of size 2 from this distribution. Show that the best test of H_0 against H_1 may be carried out by use of the statistic $X_1 + X_2$.**

Solution:

$$\begin{aligned} f(x;\theta) &= \frac{1}{\theta}e^{-x/\theta} \\ L(\theta; x_1, x_2) &= \frac{1}{\theta^2}e^{-(x_1+x_2)/\theta} \end{aligned}$$

Consider $H_0 : \theta = \theta' = 2$ vs. $H_1 : \theta = \theta'' = 4$.

$$\begin{aligned} \frac{L(\theta = 2; x_1, x_2)}{L(\theta = 4; x_1, x_2)} = \frac{\frac{1}{4}e^{-(x_1+x_2)/2}}{\frac{1}{16}e^{-(x_1+x_2)/4}} = 4e^{-(x_1+x_2)/4} &\leq k \\ e^{-(x_1+x_2)/4} &\leq \frac{k}{4} = k_1 \\ x_1 + x_2 &\geq -4\ln(k_1) = c, \end{aligned}$$

therefore the best critical region is $C = \{\boldsymbol{x} = (x_1, x_2) : x_1 + x_2 \geq c\}$ for some value c.

7. **If $X_1, X_2, \ldots, X_n$ is a random sample from a distribution having pdf of the form $f(x;\theta) = \theta x^{\theta-1}$, $0 < x < 1$, zero elsewhere, show that a best critical region for testing $H_0 : \theta = 1$ against $H_1 : \theta = 2$ is $C = \{\boldsymbol{x} = (x_1, x_2, \ldots, x_n) : c \leq \prod_{i=1}^n x_i\}$.**

   Solution:

$$\begin{aligned} f(x;\theta) &= \theta x^{\theta-1} \\ L(\theta;\mathbf{x}) &= \theta^n \left(\prod_{i=1}^n x_i\right)^{\theta-1} \end{aligned}$$

   Consider $H_0 : \theta = 1$ against $H_1 : \theta = 2$. Thus

$$\begin{aligned} \frac{L(\theta=1;\mathbf{x})}{L(\theta=2;\mathbf{x})} = \frac{1}{2^n \prod_{i=1}^n x_i} &\leq k \\ 2^n \prod_{i=1}^n x_i &\geq \frac{1}{k} = k_1 \\ \prod_{i=1}^n x_i &\geq \frac{k_1}{2^n} = c, \end{aligned}$$

   thus the best critical region has the form $C = \{\boldsymbol{x} = (x_1, x_2, \ldots, x_n) : \prod_{i=1}^n x_i \geq c\}$ for some value c.

8. **If $X_1, X_2, \ldots, X_n$ is a random sample from a beta distribution with parameters $\alpha = \beta = \theta > 0$, find a best critical region for testing $H_0 : \theta = 1$ against $H_1 : \theta = 2$.**

   Solution:

$$\begin{aligned} f(x;\theta) &= \frac{\Gamma(2\theta)}{\Gamma(\theta)\Gamma(\theta)} x^{\theta-1}(1-x)^{\theta-1} \\ L(\theta;\boldsymbol{x}) &= \left(\frac{\Gamma(2\theta)}{\Gamma(\theta)\Gamma(\theta)}\right)^n \left[\prod_{i=1}^n x_i(1-x_i)\right]^{\theta-1} = \frac{\Gamma^n(2\theta)}{\Gamma^{2n}(\theta)} \left[\prod_{i=1}^n x_i(1-x_i)\right]^{\theta-1} \end{aligned}$$

   Consider hypothesis tests $H_0 : \theta = 1$ against $H_1 : \theta = 2$. Then

$$\begin{aligned} \frac{L(\theta=1;\boldsymbol{x})}{L(\theta=2;\boldsymbol{x})} = \frac{1}{\frac{\Gamma^n(4)}{\Gamma^{2n}(2)} \prod_{i=1}^n x_i(1-x_i)} = \frac{1}{6^n \prod_{i=1}^n x_i(1-x_i)} &\leq k \\ 6^n \prod_{i=1}^n x_i(1-x_i) &\geq \frac{1}{k} = k_1 \\ \prod_{i=1}^n x_i(1-x_i) &\geq \frac{k_1}{6^n} = c, \end{aligned}$$

   thus the best critical region has the form $C = \{\boldsymbol{x} : \prod_{i=1}^n x_i(1-x_i) \geq c\}$ for some value c.