MATH 503: Mathematical Statistics
Dr. Kimberly F. Sellers, Instructor
Homework 7 Solutions

1. **Let $X_1, X_2, \ldots, X_n$ be a random sample from a $\mathbf{N}(\mu_0, \sigma^2 = \theta)$ distribution, where $0 < \theta < \infty$ and μ_0 is known. Show that the likelihood ratio test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ can be based upon the statistic $W = \sum_{i=1}^n (X_i - \mu_0)^2/\theta_0$. Determine the null distribution of W and give, explicitly, the rejection rule for a level α test.**

$$\begin{aligned}
f(x_i;\theta) &= \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}(x_i-\mu_0)^2} \\
L(\theta;\boldsymbol{x}) &= (2\pi)^{-n/2}\,\theta^{-n/2} e^{-\frac{1}{2\theta}\sum_{i=1}^n (x_i-\mu_0)^2} \\
\ln L(\theta;\boldsymbol{x}) &= \frac{-n}{2}\ln(2\pi) - \frac{n}{2}\ln(\theta) - \frac{1}{2\theta}\sum_{i=1}^n (x_i-\mu_0)^2 \\
\frac{\partial \ln L(\theta;\boldsymbol{x})}{\partial\theta} &= \frac{-n}{2\theta} + \frac{\sum_{i=1}^n (x_i-\mu_0)^2}{2\theta^2} = 0 \qquad (1)
\end{aligned}$$

where Equation (1) implies that

$$\begin{aligned}
-n\theta + \sum_{i=1}^n (x_i-\mu_0)^2 &= 0 \\
\hat{\theta} &= \frac{1}{n}\sum_{i=1}^n (x_i-\mu_0)^2.
\end{aligned}$$

Thus, the likelihood ratio test statistic is

$$\begin{aligned}
\Lambda &= \frac{L(\theta_0;\boldsymbol{x})}{L(\hat{\theta};\boldsymbol{x})} = \frac{(2\pi)^{-n/2}\,\theta_0^{-n/2} e^{-\frac{1}{2\theta_0}\sum_{i=1}^n (x_i-\mu_0)^2}}{(2\pi)^{-n/2}\left(\frac{1}{n}\sum_{i=1}^n (x_i-\mu_0)^2\right)^{-n/2} e^{-\frac{1}{2\left(\frac{1}{n}\sum_{i=1}^n (x_i-\mu_0)^2\right)}\sum_{i=1}^n (x_i-\mu_0)^2}} \\
&= \left(\frac{1}{n}\right)^{n/2}\left(\frac{\sum_{i=1}^n (x_i-\mu_0)^2}{\theta_0}\right)^{n/2} e^{-\frac{1}{2}\left(-n+\frac{\sum_{i=1}^n (x_i-\mu_0)^2}{\theta_0}\right)} \\
&= \left(\frac{W}{n}\right)^{n/2} e^{-\frac{1}{2}(W-n)},
\end{aligned}$$

therefore $\Lambda = g(W)$ is a function in terms of the test statistic $W = \frac{\sum_{i=1}^n (x_i-\mu_0)^2}{\theta_0}$ where

$$\begin{aligned}
\Lambda = g(W) &= \left(\frac{W}{n}\right)^{n/2} e^{-\frac{1}{2}(W-n)} \\
g'(W) &= \frac{1}{2}\left(\frac{W}{n}\right)^{\frac{n}{2}-1} e^{-\frac{1}{2}(W-n)}\left(1 - W^{\frac{n}{2}+1}\right) = 0,
\end{aligned}$$

so $W = 1$ is the critical point.

$\Lambda \le c$ implies that $W \le c_1$ or $W \ge c_2$, where c_1, c_2 are chosen so that $P_{\theta_0}[W \le c_1 \text{ or } W \ge c_2] = \alpha$. Under H_0, $W = \sum_{i=1}^n \left(\frac{x_i - \mu_0}{\sqrt{\theta_0}}\right)^2 \sim \chi^2_n$.

2. **Let $X_1, X_2, \ldots, X_n$ be a random sample form a Poisson distribution with mean $\theta > 0$.**

   (a) Show that the likelihood ratio test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ is based upon the statistic $Y = \sum_{i=1}^n X_i$. Obtain the null distribution of Y.

   (b) For $\theta_0 = 2$ and $n = 5$, find the significance level of the test that rejects H_0 if $Y \le 4$ or $Y \ge 17$.

(a)

$$\begin{aligned}
p(x) &= \frac{e^{-\theta}\theta^x}{x!} \\
L(\theta; \boldsymbol{x}) &= \frac{e^{-n\theta}\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \\
\ln L(\theta; \boldsymbol{x}) &= -n\theta + \left(\sum_{i=1}^n x_i\right)\ln\theta - \sum_{i=1}^n \ln(x_i!) \\
\frac{\partial \ln L(\theta)}{\partial\theta} &= -n + \frac{\sum_{i=1}^n x_i}{\theta} = 0
\end{aligned}$$

thus, $\hat{\theta} = \bar{x}$ is the MLE. Thus,

$$\begin{aligned}
\Lambda &= \frac{L(\theta_0)}{L(\hat{\theta})} = \frac{\frac{e^{-n\theta_0}\theta_0^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}}{\frac{e^{-n\bar{x}}(\bar{x})^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}} = \left(\frac{\theta_0}{\bar{x}}\right)^{\sum_{i=1}^n x_i} e^{-n(\theta_0 - \bar{x})} \\
&= \left(\frac{n\theta_0}{\sum_{i=1}^n x_i}\right)^{\sum_{i=1}^n x_i} e^{-n\theta_0 + \sum_{i=1}^n x_i} = \left(\frac{n\theta_0}{Y}\right)^Y e^{-n\theta_0 + Y},
\end{aligned}$$

where $Y = \sum_{i=1}^n x_i$. So, $\Lambda \le k$ for some k implies

$$\begin{aligned}
\left(\frac{n\theta_0}{Y}\right)^Y e^{-n\theta_0} e^Y &\le k \\
\left(\frac{n\theta_0}{Y}\right)^Y e^Y &\le k_1 \\
Y\ln(n\theta_0) - Y\ln Y + Y &\le \ln(k_1) \doteq k_2.
\end{aligned}$$

Let

$$\begin{aligned}
f(Y) &= Y\ln(n\theta_0) - Y\ln Y + Y \\
f'(Y) &= \ln(n\theta_0) - \ln Y - 1 + 1 = \ln(n\theta_0) - \ln Y = 0,
\end{aligned}$$

so $Y = n\theta_0$ is a critical point (maximum). So $\Lambda \le k$ implies $Y \le c_1$ or $Y \ge c_2$ where c_1, c_2 are chosen so that $P_{\theta_0}[Y \le c_1 \text{ or } Y \ge c_2] = \alpha$. Under H_0, $X_i \sim$Poisson(θ_0), therefore $Y \sim$Poisson$(n\theta_0)$.

(b) $\theta_0 = 2$ and $n = 5$, thus $Y \sim$Poisson(10). This implies that

$$P_{\theta_0=2}(Y \le 4 \text{ or } Y \ge 17) = P(Y \le 4) + [1 - P(Y \le 16)] = 0.029 + (1 - 0.973) = 0.056$$

3. **Let $X_1, X_2, \ldots, X_n$ be a random sample from the Beta distribution with $\alpha = \beta = \theta$ and $\Omega = \{\theta : \theta = 1, 2\}$. Show that the likelihood ratio test statistic Λ for testing $H_0 : \theta = 1$ versus $H_1 : \theta = 2$ is a function of the statistic $W = \sum_{i=1}^n \log X_i + \sum_{i=1}^n \log(1 - X_i)$.**

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

In particular under $H_0 : \theta = 1$ versus $H_1 : \theta = 2$, for $0 < x < 1$

$$\begin{aligned} f(x; \theta = 1) &= \frac{\Gamma(2)}{\Gamma(1)\Gamma(1)} x^{1-1}(1-x)^{1-1} = 1 \\ f(x; \theta = 2) &= \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)} x^{2-1}(1-x)^{2-1} = 6x(1-x), \end{aligned}$$

so the likelihood ratio test statistic is

$$\Lambda = \frac{1}{6^n \left(\prod_{i=1}^n x_i\right)\left(\prod_{i=1}^n (1-x_i)\right)} = 6^{-n} \left(\prod_{i=1}^n x_i\right)^{-1} \left(\prod_{i=1}^n (1-x_i)\right)^{-1}$$

and the decision rule considers $\Lambda \le c$, i.e.

$$\begin{aligned} \log \Lambda &= -n \log 6 - \sum_{i=1}^n \log x_i - \sum_{i=1}^n log(1-x_i) \le \log c \\ -2 \log \Lambda &= 2n \log 6 + 2\left(\sum_{i=1}^n \log x_i + \sum_{i=1}^n \log(1-x_i)\right) = 2n \log 6 + 2W \ge -2 \log c, \end{aligned}$$

so the LRT statistic is a function of W.

4. **Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with pmf $p(x; \theta) = \theta^x(1-\theta)^{1-x}$, $x = 0, 1$, where $0 < \theta < 1$. We wish to test $H_0 : \theta = \frac{1}{3}$ versus $H_1 : \theta \ne \frac{1}{3}$.**

   (a) Find Λ and $-2 \log \Lambda$.

   (b) Determine the Wald-type test.

   (c) What is Rao's score statistic?

(a)

$$\begin{aligned} L(\theta; \boldsymbol{x}) &= \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n - \sum_{i=1}^n x_i} \\ \log L(\theta; \boldsymbol{x}) &= \log \theta \sum_{i=1}^n x_i + \left(n - \sum_{i=1}^n x_i\right) \log(1-\theta) \\ \frac{\partial \log L(\theta; \boldsymbol{x})}{\partial \theta} &= \frac{\sum_{i=1}^n x_i}{\theta} - \frac{n - \sum_{i=1}^n x_i}{1-\theta} = 0, \end{aligned}$$

which implies that $\hat{\theta} = \bar{X}$ is the MLE. Thus, the likelihood ratio test statistic is

$$
\begin{aligned}
\Lambda &= \frac{\left(\frac{1}{3}\right)^{\sum_{i=1}^n x_i}\left(\frac{2}{3}\right)^{n-\sum_{i=1}^n x_i}}{\bar{x}^{\sum_{i=1}^n x_i}(1-\bar{x})^{n-\sum_{i=1}^n x_i}} = \left(\frac{1}{3\bar{x}}\right)^{\sum_{i=1}^n x_i}\left(\frac{2}{3(1-\bar{x})}\right)^{n-\sum_{i=1}^n x_i} = \left(\frac{1}{3\bar{x}}\right)^{n\bar{x}}\left(\frac{2}{3(1-\bar{x})}\right)^{n-n\bar{x}} \\
-2\log\Lambda &= -2\log(3\bar{x})^{-\sum_{i=1}^n x_i} - 2\log\left(\frac{2}{3(1-\bar{x})}\right)^{n-\sum_{i=1}^n x_i} \\
&= 2\left(\sum_{i=1}^n x_i\right)\log 3 + 2\left(\sum_{i=1}^n x_i\right)\log\bar{x} - 2\left(n-\sum_{i=1}^n x_i\right)\log 2 + 2\left(n-\sum_{i=1}^n x_i\right)\log 3 \\
&\quad +2\left(n-\sum_{i=1}^n x_i\right)\log(1-\bar{x}) \\
&= 2n\bar{x}\log 3 + 2n\bar{x}\log\bar{x} - 2n\log 2 + 2n\bar{x}\log 2 + 2n\log 3 - 2n\log 3 - 2n\bar{x}\log 3 \\
&\quad +2n\log(1-\bar{x}) - 2n\bar{x}\log(1-\bar{x})
\end{aligned}
$$

or

$$-2\log\Lambda = 2n\bar{x}\log(3\bar{x}) - 2(n-n\bar{x})\log 2 + 2(n-n\bar{x})\log(3(1-\bar{x}))$$

(b) We want to find $\chi^2_W = \left[\sqrt{nI(\hat{\theta})}\left(\hat{\theta}-\theta_0\right)\right]^2$ where

$$
\begin{aligned}
p(x;\theta) &= \theta^x(1-\theta)^{1-x} \\
\log p(x;\theta) &= x\log\theta + (1-x)\log(1-\theta) \\
\frac{\partial \log p(x;\theta)}{\partial\theta} &= \frac{x}{\theta} - \frac{1-x}{1-\theta} \\
\frac{\partial^2 \log p(x;\theta)}{\partial\theta^2} &= -\frac{x}{\theta^2} - \frac{1-x}{(1-\theta)^2} \\
I(\theta) &= -E\left(\frac{\partial^2 \log p(x;\theta)}{\partial\theta^2}\right) = \frac{E(X)}{\theta^2} + \frac{1-E(X)}{(1-\theta)^2} = \frac{\theta}{\theta^2} + \frac{1-\theta}{(1-\theta)^2} = \frac{1}{\theta(1-\theta)},
\end{aligned}
$$

thus the Wald statistic is $\chi^2_W = \left[\sqrt{nI(\hat{\theta})}\left(\hat{\theta}-\theta_0\right)\right]^2 = \left[\sqrt{\frac{n}{\bar{x}(1-\bar{x})}}\left(\bar{x}-\frac{1}{3}\right)\right]^2$.

(c) Rao's statistic is $\chi^2_R = \left(\frac{l'(\theta_0)}{\sqrt{nI(\theta_0)}}\right)^2$, where

$$
\begin{aligned}
\frac{l'(\theta_0)}{\sqrt{n}} &= \frac{1}{\sqrt{n}}\sum_{i=1}^n\left(\frac{x_i}{\theta_0} - \frac{1-x_i}{1-\theta_0}\right) = \frac{1}{\sqrt{n}\theta_0}\sum_{i=1}^n x_i - \frac{n}{\sqrt{n}(1-\theta_0)} + \frac{\sum_{i=1}^n x_i}{\sqrt{n}(1-\theta_0)} \\
&= \frac{\bar{x}\sqrt{n}}{\theta_0} - \frac{\sqrt{n}}{1-\theta_0} + \frac{\bar{x}\sqrt{n}}{1-\theta_0} = \frac{(\bar{x}-\theta_0)\sqrt{n}}{\theta_0(1-\theta_0)}.
\end{aligned}
$$

This implies that $\chi^2_R = \left(\frac{l'(\theta_0)}{\sqrt{nI(\theta_0)}}\right)^2 = \left(\frac{(\bar{x}-\theta_0)\sqrt{n}}{\theta_0(1-\theta_0)}\cdot\sqrt{\theta_0(1-\theta_0)}\right)^2 = \frac{n(\bar{x}-\theta_0)^2}{\theta_0(1-\theta_0)} = \left(\frac{\bar{x}-\frac{1}{3}}{\sqrt{\frac{2}{9n}}}\right)^2$

5. **Let $X_1, X_2, \ldots, X_n$ be a random sample from a $\Gamma(\alpha,\beta)$-distribution where α is known and $\beta > 0$. Determine the likelihood ratio test for $H_0: \beta = \beta_0$ against $H_1: \beta \neq \beta_0$.**

$$\begin{aligned}
L(\beta;\boldsymbol{x}) &= \left(\frac{1}{\Gamma(\alpha)\beta^\alpha}\right)^n \left(\prod_{i=1}^n x_i\right)^{\alpha-1} e^{-\frac{\sum_{i=1}^n x_i}{\beta}} \\
\ln L(\beta;\boldsymbol{x}) &= \ln\left(\frac{1}{\Gamma(\alpha)\beta^\alpha}\right)^n + (\alpha-1)\sum_{i=1}^n \log x_i - \frac{\sum_{i=1}^n x_i}{\beta} \\
&= -n\ln\Gamma(\alpha) - n\alpha\ln\beta + (\alpha-1)\sum_{i=1}^n \ln x_i - \frac{\sum_{i=1}^n x_i}{\beta} \\
\frac{\partial \ln L(\beta;\boldsymbol{x})}{\partial\beta} &= \frac{-n\alpha}{\beta} + \frac{\sum_{i=1}^n x_i}{\beta^2} = 0,
\end{aligned}$$

thus $\hat{\beta} = \frac{\bar{x}}{\alpha}$. Accordingly, the likelihood ratio test statistic is

$$\Lambda = \frac{L(\beta_0;\boldsymbol{x})}{L(\hat{\beta};\boldsymbol{x})} = \left(\frac{\bar{x}/\alpha}{\beta_0}\right)^{\alpha n} e^{-\sum_{i=1}^n x_i\left(\frac{1}{\beta_0}-\frac{\alpha}{\bar{x}}\right)}$$

which can be represented as a function $g(t)$ where $t = \frac{\bar{x}}{\alpha\beta_0}$, and

$$g'(t) = \alpha n t^{\alpha n-1} e^{-n\bar{x}\left(\frac{1}{\beta_0}-t\beta_0\right)} + t^{\alpha n}\left(-\beta_0 e^{-n\bar{x}\left(\frac{1}{\beta_0}-t\beta_0\right)}\right) = t^{\alpha n-1}(\alpha n - \beta_0 t)e^{-n\bar{x}\left(\frac{1}{\beta_0}-t\beta_0\right)} = 0$$

so Λ has a critical point at $t = \alpha n/\beta_0$. Thus, $\Lambda \le c$ iff $T \le c_1$ or $T \ge c_2$ where c_1, c_2 are chosen so that $P_{\beta_0}[T \le c_1 \text{ or } T \ge c_2] = \alpha$.

6. **Let $X_1, X_2, \ldots, X_n$ be a random sample from a $\mathbf{N}(0, \sigma^2 = \theta)$ distribution, where $\theta > 0$ unknown. Consider $H_0 : \theta = \theta'$ versus $H_1 : \theta < \theta'$. Show that the set $\{(x_1, \ldots, x_n) : \sum_{i=1}^n x_i^2 \le c\}$ is a uniformly most powerful critical region for testing H_0 versus H_1.**

   The density and corresponding likelihood function are

$$\begin{aligned}
f(x;\theta) &= \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}x^2} \\
L(\theta;\boldsymbol{x}) &= (2\pi\theta)^{-n/2} e^{-\frac{1}{2\theta}\sum_{i=1}^n x_i^2}.
\end{aligned}$$

   Consider $H_0 : \theta = \theta'$ versus $H_1 : \theta = \theta''$ where $\theta'' < \theta'$. By the Neymann-Pearson Theorem, for some constant $k > 0$, the most powerful (MP) critical region is

$$\begin{aligned}
\frac{L(\theta';\boldsymbol{x})}{L(\theta'';\boldsymbol{x})} = \frac{(2\pi\theta')^{-n/2} e^{-\frac{1}{2\theta'}\sum_{i=1}^n x_i^2}}{(2\pi\theta'')^{-n/2} e^{-\frac{1}{2\theta''}\sum_{i=1}^n x_i^2}} &\le k \\
\left(\frac{\theta''}{\theta'}\right)^{n/2} e^{\frac{1}{2}\left(\frac{1}{\theta''}-\frac{1}{\theta'}\right)\sum_{i=1}^n x_i^2} &\le k \\
e^{\frac{1}{2}\left(\frac{1}{\theta''}-\frac{1}{\theta'}\right)\sum_{i=1}^n x_i^2} &\le k\left(\frac{\theta'}{\theta''}\right)^{n/2} = k_1
\end{aligned}$$

$$\frac{1}{2}\underbrace{\left(\frac{1}{\theta''}-\frac{1}{\theta'}\right)}_{>0}\sum_{i=1}^n x_i^2 \le \ln(k_1) = k_2$$

$$\sum_{i=1}^n x_i^2 \le \frac{k_2}{\frac{1}{2}\left(\frac{1}{\theta''}-\frac{1}{\theta'}\right)} = c.$$

Further, because $\{\sum_{i=1}^n x_i^2 \leq c\}$ is the MP critical region for $H_0 : \theta = \theta'$ versus $H_1 : \theta = \theta''$ for any $\theta'' < \theta'$, this critical region is a UMP critical region for $H_0 : \theta = \theta'$ versus $H_1 : \theta < \theta'$.

7. **Let $X_1, X_2, \ldots, X_n$ be a random sample from a $\mathbf{N}(0, \sigma^2 = \theta)$ distribution, where $\theta > 0$ unknown. Consider $H_0 : \theta = \theta'$ versus $H_1 : \theta \neq \theta'$. Show that there is no uniformly most powerful test for testing H_0 versus H_1.**

   The density and corresponding likelihood function are

$$\begin{aligned} f(x;\theta) &= \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}x^2} \\ L(\theta;\boldsymbol{x}) &= (2\pi\theta)^{-n/2} e^{-\frac{1}{2\theta}\sum_{i=1}^n x_i^2}. \end{aligned}$$

   Consider $H_0 : \theta = \theta'$ versus $H_1 : \theta = \theta''$ where $\theta'' > \theta'$. By the Neymann-Pearson Theorem, for some constant $k > 0$, the most powerful (MP) critical region is

$$\begin{aligned} \frac{L(\theta';\boldsymbol{x})}{L(\theta'';\boldsymbol{x})} = \frac{(2\pi\theta')^{-n/2} e^{-\frac{1}{2\theta'}\sum_{i=1}^n x_i^2}}{(2\pi\theta'')^{-n/2} e^{-\frac{1}{2\theta''}\sum_{i=1}^n x_i^2}} &\leq k \\ \left(\frac{\theta''}{\theta'}\right)^{n/2} e^{\frac{1}{2}\left(\frac{1}{\theta''}-\frac{1}{\theta'}\right)\sum_{i=1}^n x_i^2} &\leq k \\ e^{\frac{1}{2}\left(\frac{1}{\theta''}-\frac{1}{\theta'}\right)\sum_{i=1}^n x_i^2} &\leq k\left(\frac{\theta'}{\theta''}\right)^{n/2} = k_1 \end{aligned}$$

$$\frac{1}{2}\underbrace{\left(\frac{1}{\theta''}-\frac{1}{\theta'}\right)}_{<0}\sum_{i=1}^n x_i^2 \leq \ln(k_1) = k_2$$

$$\sum_{i=1}^n x_i^2 \geq \frac{k_2}{\frac{1}{2}\left(\frac{1}{\theta''}-\frac{1}{\theta'}\right)} = c.$$

   Thus for any $\theta'' > \theta'$, $\{\sum_{i=1}^n x_i^2 \geq c\}$ is the MP critical region for $H_0 : \theta = \theta'$ versus $H_1 : \theta = \theta''$. However, as shown in the solution to Problem 6, for any $\theta'' < \theta'$, $\{\sum_{i=1}^n x_i^2 \leq c\}$ is the MP critical region for $H_0 : \theta = \theta'$ versus $H_1 : \theta = \theta''$. Thus, there is no UMP test for $H_0 : \theta = \theta'$ versus $H_1 : \theta \neq \theta'$, because such a test must produce the same critical region under both scenarios.

8. **Let $X_1, X_2, \ldots, X_n$ be a random sample from a $\mathbf{N}(\theta, \sigma^2 = 16)$ distribution. Find the sample size n and a uniformly most powerful test of $H_0 : \theta = 25$ against $H_1 : \theta < 25$ with power function $\gamma(\theta)$ so that approximately $\gamma(25) = 0.10$ and $\gamma(23) = 0.90$.**

   The density and corresponding likelihood function are

$$\begin{aligned} f(x;\theta) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\theta)^2} \\ L(\theta;\boldsymbol{x}) &= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\theta)^2}. \end{aligned}$$

Consider $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$, where $\theta_1 < \theta_0$. By Neymann-Pearson Theorem,

$$
\begin{aligned}
\frac{L(\theta_0; \boldsymbol{x})}{L(\theta_1; \boldsymbol{x})} = \frac{(2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\theta_0)^2}}{(2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\theta_1)^2}} &\leq k \\
\exp\left(-\frac{1}{2\sigma^2}\left[\sum_{i=1}^n (x_i-\theta_0)^2 - \sum_{i=1}^n (x_i-\theta_1)^2\right]\right) &\leq k \\
\sum_{i=1}^n (x_i-\theta_0)^2 - \sum_{i=1}^n (x_i-\theta_1)^2 &\geq -2\sigma^2 \ln k = k_1 \\
2(\theta_1-\theta_0)\sum_{i=1}^n x_i + n(\theta_0^2-\theta_1^2) &\geq k_1 \\
\sum_{i=1}^n x_i &\leq \frac{k_1 - n(\theta_0^2-\theta_1^2)}{2(\theta_1-\theta_0)} = c, \text{ because } \theta_1-\theta_0<0.
\end{aligned}
$$

Note that, because $X_1, X_2, \ldots, X_n$ are a random sample from a $\mathrm{N}(\theta, \sigma^2 = 16)$ distribution, $\sum_{i=1}^n X_i$ has a $\mathrm{N}(n\theta, n\sigma^2 = 16n)$ distribution; alternatively, note that $\bar{X} \sim N(\theta, \sigma_{\bar{X}}^2 = 16/n)$. We want to find c and n so that $\gamma(25) = P_{\theta_0=25}\left(\sum_{i=1}^n x_i \leq c\right) = 0.10$ and $\gamma(23) = P_{\theta_1=23}\left(\sum_{i=1}^n x_i \leq c\right) = 0.90$, i.e.

$$
\begin{aligned}
P_{\theta_0=25}\left(\sum_{i=1}^n x_i \leq c\right) &= P\left(Z \leq \frac{c-25n}{4\sqrt{n}}\right) = 0.10 \\
P_{\theta_1=23}\left(\sum_{i=1}^n x_i \leq c\right) &= P\left(Z \leq \frac{c-23n}{4\sqrt{n}}\right) = 0.90,
\end{aligned}
$$

thus

$$
\frac{c-25n}{4\sqrt{n}} = -1.281 \qquad \text{and} \qquad \frac{c-23n}{4\sqrt{n}} = 1.281.
$$

Alternatively, we can represent the problem to find c and n so that $\gamma(25) = P_{\theta_0=25}\left(\bar{X} \leq c\right) = 0.10$ and $\gamma(23) = P_{\theta_1=23}\left(\bar{X} \leq c\right) = 0.90$, i.e.

$$
\begin{aligned}
P_{\theta_0=25}\left(\bar{X} \leq c\right) &= P\left(Z \leq \frac{c-25}{4/\sqrt{n}}\right) = 0.10 \\
P_{\theta_1=23}\left(\bar{X} \leq c\right) &= P\left(Z \leq \frac{c-23}{4/\sqrt{n}}\right) = 0.90,
\end{aligned}
$$

thus

$$
\frac{c-25}{4/\sqrt{n}} = -1.281 \qquad \text{and} \qquad \frac{c-23}{4/\sqrt{n}} = 1.281.
$$

Either way, we find that $n = 26.255 \approx 27$ and $c = 23.986$.

9. **Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with pdf**

$$
f(x;\theta) = \theta x^{\theta-1}, 0 < x < 1, \ \textbf{zero elsewhere},
$$

where $\theta > 0$. Find a sufficient statistic for θ and show that a uniformly most powerful test of $H_0 : \theta = 6$ against $H_1 : \theta < 6$ is based on this statistic.

By the Neymann-Fisher Factorization Theorem,

$$\prod_{i=1}^{n} f(x_i;\theta) = \underbrace{\theta^n (\prod_{i=1}^{n} x_i)^{\theta-1}}_{k_1(\prod_{i=1}^{n} x_i;\theta)} \cdot \underbrace{1}_{k_2(\boldsymbol{x})},$$

$Y = \prod_{i=1}^{n} x_i$ is a sufficient statistic for θ.
Consider $H_0 : \theta = \theta' = 6$ against $H_1 : \theta = \theta''$ where $\theta'' < 6$. Recognizing that $L(\theta;\boldsymbol{x}) = \prod_{i=1}^{n} f(x_i;\theta)$

$$\frac{L(\theta' = 6;\boldsymbol{x})}{L(\theta'';\boldsymbol{x})} = \frac{6^n (\prod_{i=1}^{n} x_i)^{6-1}}{(\theta'')^n (\prod_{i=1}^{n} x_i)^{\theta''-1}} = \left(\frac{6}{\theta''}\right)^n \left(\prod_{i=1}^{n} x_i\right)^{6-\theta''}$$

is monotone increasing in $Y = \prod_{i=1}^{n} x_i$. This implies that $\{\prod_{i=1}^{n} x_i \leq k\}$ is the UMP critical region, which is clearly based on the sufficient statistic, $Y = \prod_{i=1}^{n} x_i$.