Solutions for Homework 4
Due: Wednesday, March 28

1. (10 pt) Let \( X \) be a random variable having probability density
\[ f(x|\theta) = \exp\{\eta(\theta)Y(x) - (\theta)\}h(x), \]
where \( \eta \) is an increasing and differentiable function of \( \theta \in \Theta \subset \mathbb{R} \).

(a) Show that \( \log \ell(\hat{\theta}) - \log \ell(\theta_0) \) is increasing(or decreasing) in \( Y \), when \( \hat{\theta} > \theta_0 \)(or \( \hat{\theta} < \theta_0 \)), where \( \ell(\theta) = f(x|\theta) \), \( \hat{\theta} \) is an MLE of \( \theta \), and \( \theta_0 \in \Theta \).

(b) For testing \( H_0 : \theta_1 \leq \theta \leq \theta_2 \) versus \( H_1 : \theta < \theta_1 \) or \( \theta > \theta_2 \) or for testing \( H_0 : \theta = \theta_0 \) versus \( H_1 : \theta \neq \theta_0 \), show that there is a likelihood ratio test whose rejection region is equivalent to \( Y(X) < c_1 \) or \( Y(x) > c_2 \) for some constants \( c_1 \) and \( c_2 \).

Solution:

(a) From the property of exponential families, \( \hat{\theta} \) is a solution of the likelihood equation:
\[ \frac{\partial \log \ell(\theta)}{\partial \theta} = \eta'(\theta)Y(X) - \xi'(\theta) \]
Since \( \eta'(\hat{\theta})Y - \xi'(\hat{\theta}) = 0 \). \( \hat{\theta} \) is an increasing function of \( Y \) and \( \frac{\partial \hat{\theta}}{\partial Y} > 0 \). Consequently, for \( \forall \theta_0 \in \Theta \),
\[
\frac{d}{dY} \left[ \log \ell(\hat{\theta}) - \log \ell(\theta_0) \right] = \frac{d}{dY} \left[ \eta(\hat{\theta})Y - \xi(\hat{\theta}) - \eta(\theta_0)Y + \xi(\theta_0) \right]
= \frac{d\hat{\theta}}{dY} \eta'(\hat{\theta})Y + \eta'(\hat{\theta}) - \frac{d\hat{\theta}}{dY} \xi'(\hat{\theta}) - \eta(\theta_0)
= \frac{d\hat{\theta}}{dY} \left[ \eta'(\hat{\theta})Y - \xi'(\hat{\theta}) \right] + \eta(\hat{\theta}) - \eta(\theta_0)
= \eta(\hat{\theta}) - \eta(\theta_0)
\]
which is positive if \( \hat{\theta} > \theta_0 \).

(b) Since \( \ell(\theta) \) is increasing when \( \theta \leq \hat{\theta} \) and decreasing when \( \theta > \hat{\theta} \),
\[
\lambda(X) = \frac{\sup_{\theta_1 \leq \theta \leq \theta_2} \ell(\theta)}{\sup_{\theta \in \Theta} \ell(\theta)} = \begin{cases} \ell(\theta_1) & \hat{\theta} < \theta_1 \\ \ell(\theta) & \theta_1 \leq \hat{\theta} \leq \theta_2 \\ \ell(\theta_2) & \hat{\theta} > \theta_2 \end{cases}
\]
2. (10 pt) Let an unknown. 
(b) show that the test 
(c) derive the likelihood ratio 

\[ T(x) = \begin{cases} 1 & \text{if } h(x) < c \\ 0 & \text{if } h(x) \geq c \end{cases} \]

Hence the family of densities of \( X \) has MLR in nonincreasing in \( h(x) = \frac{f_0}{g_0} \). Hence the family of densities of \( X \) has MLR in

(a) Let \( f(x) \) and \( g(x) \) be the corresponding pdfs. The probability density of \( X \) is then \( \theta f(x) + (1 - \theta)g(x) \). For \( 0 \leq \theta_1 < \theta \leq 1 \), the power is

\[ \beta_r(\theta) = \frac{\theta f(x) + (1 - \theta)g(x)}{\theta f(x) + (1 - \theta)g(x)} \]

(b) For any test \( T \), its power is

\[ \beta_r(\theta) = \int T(x) \left( \frac{f(x)}{g(x)} \right) dx \]

\[ \int \left( \frac{f(x)}{g(x)} \right) dx = \int f(x) \left( \frac{1}{g(x)} \right) dx + \int g(x) \left( \frac{1}{g(x)} \right) dx \]

\[ \int f(x) \left( \frac{1}{g(x)} \right) dx = \int f(x) \left( \frac{1}{g(x)} \right) dx + \int g(x) \left( \frac{1}{g(x)} \right) dx \]

Therefore, the power is

\[ \beta_r(\theta) = \int \left( \frac{f(x)}{g(x)} \right) dx \]

(c) derive the likelihood ratio \( \lambda(X) \) for \( H_0 : \theta = \theta_0 \) or \( \theta \geq \theta_2 \).

(a) Find a UMP test of size \( \alpha \) for testing \( H_0 : \theta \leq \theta_0 \) versus \( H_1 : \theta \geq \theta_2 \), where \( \theta \in [0, 1] \) is unknown.

(b) show that the test \( T(X) = \alpha \) is a UMP test of size \( \alpha \) for testing \( H_0 : \theta = \theta_0 \) or \( \theta \geq \theta_2 \) versus \( H_1 : \theta < \theta_2 \), where \( \theta \in [0, 1] \) is known.

(c) derive the likelihood ration \( \lambda(X) \) for \( H_0 : \theta \leq \theta_1 \) or \( \theta > \theta_2 \).

And with the result obtained from (i), we have \( \lambda(X) < c \) or \( \lambda(X) > d_1 \), for some constant \( c_1 \) and \( d_1 \).
(c) \[ \lambda(x) = \frac{\sup_{\theta \in \Theta_1} \theta(f(x) - g(x)) + g(x)}{\sup_{\theta \in \Theta_0} \theta(f(x) - g(x)) + g(x)} \]

Discussing different cases for \( f(x) - g(x) \) will give us \( \text{LRT} = \{ x : c_1 \leq h(x) \leq c_2 \} \), for some \( c_1, c_2 \).

3. (10 pt) A family of probability density functions \( f(x \mid \theta) \) on \( \mathbb{R} \), indexed by \( \theta \in \Theta \subset \mathbb{R} \), is said to have a \textit{monotone likelihood ratio} (MLR) if, for each \( \theta_0 \neq \theta_1 \), the ratio \( f(x \mid \theta_1)/f(x \mid \theta_0) \) is monotonic in \( x \). Assume that \( \frac{\partial^2}{\partial \theta \partial x} \log f(x \mid \theta) \) exists.

(a) Show that a family of density functions \( \{ f(x \mid \theta) : \theta \in \Theta \subset \mathbb{R} \} \) has MLR in \( x \) is equivalent to one of the following conditions:

(a) \( \frac{\partial^2}{\partial \theta \partial x} \log f(x \mid \theta) \geq 0 \) for all \( x \) and \( \theta \);
(b) \( f(x \mid \theta) \frac{\partial^2}{\partial \theta \partial x} \log f(x \mid \theta) \geq \frac{\partial}{\partial x} f(x \mid \theta) \frac{\partial}{\partial x} f(x \mid \theta) \) for all \( x \) and \( \theta \)

(b) Let \( Z \sim \text{No}(\sqrt{\theta}, 1) \), then \( Z^2 \) has the non-central chi-square distribution \( \chi^2_1(\theta) \).

Show that \( \{ \chi^2_1(\theta) \} \) has MLR in \( x \).

Solution: Note: there’s a typo in the original problem set. So for (b), you might get that it’s not a sufficient and necessary condition.

(a) Note that

\[ \frac{\partial^2}{\partial \theta \partial x} \log f_\theta(x) = \frac{\partial}{\partial \theta} \frac{\partial^2}{\partial x^2} f_\theta(x) = \frac{\partial^2}{\partial \theta \partial x} f_\theta(x) \frac{\partial}{\partial x} f_\theta(x) - \frac{\partial}{\partial x} f_\theta(x) \frac{\partial^2}{\partial x \partial \theta} f_\theta(x) \]

Since \( f_\theta(x) > 0 \), conditions (a) and (b) are equivalent. (a) \( \iff \frac{\partial}{\partial \theta} \log f_\theta(x) = \frac{\partial}{\partial x} f_\theta(x) \) is nondecreasing in \( \theta \) for any fixed \( x \). Hence, it is equivalent to, for \( \forall \theta_1 < \theta_2 \), and \( \forall x \),

\[ \frac{\partial}{\partial x} f_{\theta_2}(x) \leq \frac{\partial}{\partial x} f_{\theta_1}(x) \]

, i.e., for \( \forall \theta_1 < \theta_2 \), and \( \forall x \),

\[ \frac{\partial}{\partial x} f_{\theta_2}(x) = \frac{\partial}{\partial x} f_{\theta_1}(x) \frac{\partial^2}{\partial x^2} f_{\theta_1}(x) - \frac{\partial^2}{\partial x \partial \theta} f_{\theta_1}(x) \frac{\partial^2}{\partial x^2} f_{\theta_1}(x) \]

, which is equivalent to, the family of density functions \( \{ f(x \mid \theta) : \theta \in \Theta \subset \mathbb{R} \} \) has MLR in \( x \).
(b) Let \( f_\theta(x) \) be the Lebesgue density of \( \chi^2_1(\theta) \).

\[
f_\theta(x) = \frac{1}{2\sqrt{2\pi x}}[e^{-(\sqrt{x}-\theta)^2/2} + e^{-(\sqrt{x}+\theta)^2/2}]
\]

for \( 0 \leq \theta_1 < \theta_2 \)

\[
f_{\theta_2}(x) = \frac{e^{-\theta_2^2/2}(e^{\theta_2\sqrt{x}} + e^{-\theta_2\sqrt{x}})}{e^{-\theta_1^2/2}(e^{\theta_1\sqrt{x}} + e^{-\theta_1\sqrt{x}})}
\]

Let \( g_\theta(y) = e^{\theta y} + e^{-\theta y} \). Note that

\[
\frac{\partial}{\partial y} g_\theta(y) = \theta(e^{\theta y} + e^{-\theta y})
\]

\[
\frac{\partial}{\partial \theta} g_\theta(y) = y(e^{\theta y} + e^{-\theta y})
\]

\[
\frac{\partial^2}{\partial \theta \partial y} g_\theta(y) = \theta y(e^{\theta y} + e^{-\theta y})
\]

Hence,

\[
g_\theta(y) \frac{\partial^2}{\partial \theta \partial y} g_\theta(y) = \theta y(e^{\theta y} + e^{-\theta y})^2 \\
\geq \theta y(e^{\theta y} - e^{-\theta y})^2 \\
= \frac{\partial}{\partial y} g_\theta(y) \frac{\partial}{\partial \theta} g_\theta(y)
\]

Thus condition (b) in (i) holds, \( \{\chi^2_1(\theta) \ \theta \geq 0\} \) has MLR in \( x \).

---

4. (10 pt) A family of probability distributions on \( \mathbb{R} \), indexed by \( \theta \in \Theta \subset \mathbb{R} \), is called stochastically increasing if their CDFs \( \{F(x \mid \theta), \ \theta \in \Theta\} \) satisfy

\[
\theta_1 < \theta_2 \implies F(x \mid \theta_1) \geq F(x \mid \theta_2) \quad \forall x \in \mathbb{R}.
\]

(Intuitively this says that larger values of the parameter \( \theta \) are associated with larger values of the random variable \( X \)).

(a) Show that if a family of pdfs \( \{f(x \mid \theta) : \theta \in \Theta\} \) has an MLR, then the corresponding family of CDFs is stochastically increasing in \( \theta \).

(b) Show that the converse of part (a) is false.

solution:
(a) A smart way to show (a) From Wing Sven (Aaron Silver found it)
MLR ⇒ \( \frac{f(x_1 | \theta_1)}{f(x_1 | \theta_0)} \geq \frac{f(x_0 | \theta_1)}{f(x_0 | \theta_0)} \forall x_1 > x_0, \theta_1 > \theta_0 \)
\[
\int_{-\infty}^{x_1} f(x_1 | \theta_1)f(x_0 | \theta_0)dx_0 \geq \int_{-\infty}^{x_1} f(x_1 | \theta_0)f(x_0 | \theta_1)dx_0
\]
so, \( f(x_1 | \theta_1)F(x_1 | \theta_0) \geq f(x_1 | \theta_0)F(x_1 | \theta_1) \)
which indicates \( \forall x \)
\[
\frac{f(x | \theta_1)}{f(x | \theta_0)} \leq \frac{F(x | \theta_1)}{F(x | \theta_0)}
\]
also similary, we can get \( \int_{-\infty}^{x_0} f(x_1 | \theta_1)f(x_0 | \theta_0)dx_1 \geq \int_{-\infty}^{x_0} f(x_1 | \theta_0)f(x_0 | \theta_1)dx_1 \)
so \( \forall x \)
\[
\frac{f(x | \theta_1)}{f(x | \theta_0)} \leq \frac{1 - F(x | \theta_1)}{1 - F(x | \theta_0)}
\]
It follows that
\[
\frac{F(x | \theta_0)}{1 - F(x | \theta_0)} \leq \frac{F(x | \theta_1)}{1 - F(x | \theta_1)}
\]
\( \Rightarrow F(x | \theta_0) \geq F(x | \theta_1) \)

(b) A family of pdfs \( \{f(x | \theta) : \theta \in \Theta\} \) has an MLR, for any \( \theta_2 > \theta_1 \), \( \frac{f(x | \theta_2)}{f(x | \theta_1)} \) is monotone in \( x \). WLOG, consider \( \frac{f(x | \theta_2)}{f(x | \theta_1)} > \lim_{x \to -\infty} \frac{f(x | \theta_2)}{f(x | \theta_1)} > 1 \), otherwise, \( \int f(x | \theta_2)dx < \int f(x | \theta_1)dx = 1 \), however, \( \int f(x | \theta_2)dx \) should be 1 since it’s a pdf. Set
\[
k = \inf\{x : \frac{f(x | \theta_2)}{f(x | \theta_1))} \geq 1\}
\]
Want to show \( F(x | \theta_1) \geq F(x | \theta_2) \) \( \forall x \in \mathbb{R} \), we only need to show \( P(X > x | \theta_2) \geq P(X | \theta_1) \), that is
\[
\int_{x}^{\infty} f(x | \theta_2)dx \geq \int_{x}^{\infty} f(x | \theta_1)dx \quad (*)
\]
i. \( \forall x \geq k \), \( (*) \) obviously holds since \( f(x | \theta_2) \geq f(x | \theta_1) \).
ii. \( \forall x < k \), \( \int_{x}^{\infty} f(x | \theta_2) - \int_{x}^{\infty} f(x | \theta_1) = - \int_{-\infty}^{x}[f(x | \theta_2) - f(x | \theta_1)]dx \geq 0 \),
since \( \forall x < k, 0 \leq f(x | \theta_2) < f(x | \theta_1) \).
Together (i),(ii), we have \( F(x | \theta_1) \geq F(x | \theta_2) \) \( \forall x \in \mathbb{R} \), \( \forall \theta_2 > \theta_1 \), i.e, \( \{F(x | \theta), \theta \in \Theta\} \) stochastically increase.

(c) A counter example: let \( X \sim f(x | \sigma^2) = \mathcal{N}(0, \sigma^2) \), \( f(x | \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{x^2}{2\sigma^2}} \)
\[
\frac{f(x | \sigma_2^2)}{f(x | \sigma_1^2)} = \frac{\sigma_1}{\sigma_2}e^{-\frac{x^2(\frac{1}{2\sigma_2^2})}{2\sigma_1^2}} - \frac{1}{2\sigma_1^2}
\]
, which is not nonotone in \( x \). But
\[
P(X > t| \sigma_2) \quad \frac{t}{\sigma_2} < \frac{t}{\sigma_1}
\]
\[
= P \left( \frac{X}{\sigma_2} > \frac{t}{\sigma_2} \right) = 1 - \Phi \left( \frac{t}{\sigma_2} \right)
\]
\[
> 1 - \Phi \left( \frac{t}{\sigma_1} \right) = P(X > t| \sigma_1), \forall \sigma_2 > \sigma_1.
\]
\[\therefore \{F(x \mid \sigma), \sigma > 0\} \text{ is stocastically increasing in } \sigma.\]

Another counter example, the Cauchy family with pdfs
\[
f(x \mid \theta) = \frac{1}{\pi(1 + (x - \theta)^2)} \quad -\infty < x < \infty, \quad \theta > 0
\]

---

5. (10 pt) Bickel & Doksum, page 270: 4.1.3

Solution:

(a) \( X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Po}(\theta) \). The MLE for \( \theta \) is \( \bar{X} = \frac{1}{n} \sum X_i \equiv \frac{1}{n} S_n \), where \( S_n \equiv \sum X_i \).

It will be slightly easier to work with \( S_n \) than with \( \bar{X} \), since the distribution of \( S_n \) is over the integers rather than the integer multiples of \( \frac{1}{n} \). A useful property of Poisson random variables is that if \( X_i \overset{iid}{\sim} \text{Po}(\gamma_i) \) are all independent, then \( \sum X_i \overset{iid}{\sim} \text{Po}(\sum \gamma_i) \). In our case, that means that \( S_n \overset{iid}{\sim} \text{Po}(n\theta) \).

We have \( H : \theta \leq \theta_0 \) vs. \( K : \theta > \theta_0 \). We will reject \( H \) in favor of \( K \) if \( S_n \) is sufficiently large. In order to have a level (size) \( \alpha \) test, we choose the lower endpoint \( c \) of our critical region to be the smallest integer \( s \) for which the probability of getting \( S_n \) at least \( s \) is less than \( \alpha \): \( c = \min \{ s : P(S_n \geq s) \leq \alpha \} \).

We then reject \( H \) in favor of \( K \) if we observe \( S_n \geq c \) or, equivalently, if we observe \( \bar{X} \geq \frac{c}{n} \).

(b) The power function is:

\[
\beta(\theta) = P(S_n \geq c \mid \theta)
\]
\[
= \sum_{s=c}^{\infty} P(S_n = s \mid \theta)
\]
\[
= \sum_{s=c}^{\infty} \frac{(n\theta)^s}{s!} e^{-n\theta}
\]
\[
= 1 - \sum_{s=0}^{c-1} \frac{(n\theta)^s}{s!} e^{-n\theta}
\]
We wish to show that this function is increasing in $\theta$. First consider the special case where $c = 0$:

$$\beta(\theta) = P(S_n \geq 0|\theta) = 1, \forall \theta$$

This trivial case hardly deserves to be called a test. But at any rate, it is “increasing” in $\theta$ in that it is not decreasing in $\theta$. Now consider the special case where $c = 1$:

$$\beta(\theta) = P(S_n \geq 1|\theta) = 1 - P(S_n = 0|\theta) = 1 - e^{-n\theta}$$

Here we have $\frac{d}{d\theta}[1 - e^{-n\theta}] = ne^{-n\theta} > 0$, so again the power function is increasing in $\theta$. Now we consider all the cases where $c \geq 2$:

$$\frac{d}{d\theta} \beta(\theta) = -\frac{d}{d\theta} \left[ \sum_{s=0}^{c-1} \frac{(n\theta)^s}{s!} e^{-n\theta} \right]$$

$$= \sum_{s=0}^{c-1} \left[ \frac{(n\theta)^s}{s!} \right] (ne^{-n\theta}) - (e^{-n\theta}) \left( \frac{n s (n\theta)^{s-1}}{s!} \right)$$

$$= ne^{-n\theta} \left[ \sum_{s=0}^{c-1} \frac{(n\theta)^s}{s!} - \sum_{s=0}^{c-1} \frac{s (n\theta)^{s-1}}{s!} \right]$$

$$= ne^{-n\theta} \left[ \sum_{s=0}^{c-1} \frac{(n\theta)^s}{s!} - 0 - \sum_{s=1}^{c-1} \frac{(n\theta)^{s-1}}{(s-1)!} \right]$$

$$= ne^{-n\theta} \left[ \sum_{s=0}^{c-1} \frac{(n\theta)^s}{s!} - \sum_{s'=0}^{c-2} \frac{(n\theta)^{s'}}{s'!} \right]$$

$$= ne^{-n\theta} \left[ \frac{(n\theta)^{c-1}}{(c-1)!} \right]$$

$$> 0, \forall \theta$$

Since this function’s derivative is positive for all $\theta$, the power function must be increasing.

Additionally, we note that the above expression for $\beta(\theta)$ resembles a $\text{Ga}(c, 1)$ distribution. Define $g(x) = \frac{1}{(c-1)!} x^{c-1} e^{-x}$. (This is the pdf of the $\text{Ga}(c, 1)$ distribution.)
Define $G(x) = \int_{-\infty}^{x} g(t)\,dt$ so that $\frac{d}{dx}G(x) = g(x)$. \( \frac{d}{d\theta}G(n\theta) = ng(n\theta) = \frac{d}{d\theta}\beta(\theta) \). Integrating both sides of this expression with respect to \( \theta \) gives $G(n\theta) = \beta(\theta)$, where the integration constant must be zero since these functions both have limits of 0 and 1 at $-\infty$ and $+\infty$, respectively. Thus, the power function is the cdf of a $\text{Ga}(c, 1)$ distribution, evaluated at $n\theta$.

(c) If we assume that $n$ is large, and $\theta$ is not too close to 0, then by the Central Limit Theorem, the distribution of $S_n$ (or of $\bar{X}$) will be approximately normal under the null. We know that $\mathbb{E}S_n = n\theta_0$ and $\text{Var}S_n = n\theta_0$, so the critical value $c$, under the normal approximation, should take the value $c \approx \Phi^{-1}(1 - \alpha)\sqrt{n\theta} + n\theta$. Using $\bar{X}$ as a test statistic, we have a critical value of $\frac{\bar{X}}{n} \approx \Phi^{-1}(1 - \alpha)\sqrt{\frac{2}{n}} + \theta$.

6. (10 pt) Bickel & Doksum, page 272: 4.2.3

Solution:

(a) The gambler would like to test his hypothesis, so we let $H$ be the hypothesis that the die has his perceived probability distribution of outcomes, and $K$ be that the die is fair. The most powerful hypothesis test will be the likelihood ratio test where the test statistic is:

$$T = \frac{(\frac{1}{6})^{N_1}(\frac{1}{6})^{N_2}(\frac{1}{6})^{N_3}(\frac{1}{6})^{N_4}(\frac{1}{6})^{N_5}(\frac{1}{6})^{N_6}}{(0.17)^{N_1}(0.17)^{N_2}(0.17)^{N_3}(0.17)^{N_4}(0.14)^{N_5}(0.18)^{N_6}}$$

or simply

$$T = \prod_{j=1}^{6} (6\theta_j)^{-N_j}$$

where $N_i$ be the number of times the die rolls face $i$ after $n$ rolls, $\sum_{i=1}^{6} N_i = N$, and $\theta_i$ is the probability of face $i$ showing under the gambler’s null hypothesis $H$.

The gambler will reject $H$ (unfair die) in favor of $K$ (fair die) if $T$ is sufficiently large. For simplicity’s sake, it is worth transforming the test into an equivalent one with a simpler test statistic. Since $\log(\cdot)$ is a strictly increasing function, then we may create an equivalent test that rejects $H$ in favor of $K$ when $\log T$ is sufficiently large.
\[ \log T = \sum_{i=1}^{6} N_j \log(6 \theta_j) \]

\[ = \sum_{i=1}^{6} -N_j \log \theta_j - (\log 6) \sum_{i=1}^{6} N_j \]

\[ = \sum_{i=1}^{6} -N_j \log \theta_j - (\log 6)N \]

Since \(-(\log 6)N\) is a constant, subtracting it from the test statistic will produce an equivalent test statistic:

\[ T' = \sum_{i=1}^{6} -N_j \log \theta_j \]

\[ = \left[ \sum_{i=1}^{4} -N_j \log(0.17) \right] - N_5 \log(0.14) - N_6 \log(0.18) \]

\[ = -\log(0.17)[N - N_5 - N_6] \log(0.17) - N_5 \log(0.14) - N_6 \log(0.18) \]

\[ = -\log(0.17)N + N_5[\log(0.17) - \log(0.14)] + N_6[\log(0.17) - \log(0.18)] \]

Since \(-N \log(0.17)\) is a constant, it may also be subtracted, giving a simpler test statistic:

\[ T'' = [\log(0.17) - \log(0.14)]N_5 + [\log(0.17) - \log(0.18)]N_6 \]

\[ \approx 0.194N_5 - 0.057N_6 \]

(b) The following table shows the possible values of the pair \((N_5, N_6)\) when \(n = 2\), their respective probabilities under \(H\), and the approximate value for each pair of \(T'\). For convenience, the rows in the table have been arranged in ascending order of \(T'\).

| \((N_5, N_6)\) | \(P(N_5, N_6|H)\) | \(T' \approx\) |
|----------------|-----------------|-------------|
| (0, 2)         | \((0.18)^2 = 0.0324\) | -0.114      |
| (0, 1)         | \(2(0.68)(0.18) = 0.2448\) | -0.057      |
| (0, 0)         | \((0.68)^2 = 0.4624\) | 0           |
| (1, 1)         | \(2(0.14)(0.18) = 0.0504\) | 0.137       |
| (1, 0)         | \(2(0.68)(0.14) = 0.1904\) | 0.194       |
| (2, 0)         | \((0.14)^2 = 0.0196\) | 0.388       |

Recall that the gambler will reject \(H\) (unfair die) in favor of \(K\) (fair die) if the test statistic \(T'\) is sufficiently large. From the table we see that to have a level
\( \alpha = 0.0196 \) test, the gambler will reject \( H \) if and only if \( N_5 = 2 \) and \( N_6 = 0 \), i.e., if he rolls two fives. (Note that determining the entire distribution of the test statistic is not necessary, but it is necessary to show that the likelihood ratio takes no value larger than it does when two fives are rolled. Merely showing that the probability of two fives occurring under the null is 0.0196 does not complete the proof that such a rejection region constitutes the most powerful test.)

(c) We are interested in the approximate distribution of the test statistic \( T' \approx 0.194N_5 - 0.057N_6 \). If \( n \) is sufficiently large, then (see the problem statement) the distribution of \( T' \) will be approximately \( N_0(n\mu, n\sigma^2) \), where \( \mu \) and \( \sigma^2 \) are as follows:

\[
\mu = (0.194)(0.14) + (-0.057)(0.18) = 0.0169
\]

\[
\sigma^2 = 4(0.17)(0 - 0.0169)^2 + (0.14)(0.194 - 0.0169)^2 + (0.18)(-0.057 - 0.0169)^2 \\approx 0.0056
\]

Thus, if \( n \) is sufficiently large, then the gambler should reject \( H \) in favor of \( K \) if the computed test statistic \( T' \) exceeds the critical value \( \Phi^{-1}(1 - \alpha)\sqrt{0.0056n} + 0.0169n \approx \Phi^{-1}(1 - \alpha)0.0747\sqrt{n} + 0.0169n \).

7. (10 pt) Bickel & Doksum, page 276: 4.3.9

**Solution:**

for any particular value of \( \theta_1 > 1 \), consider the pair of hypotheses \( H : \theta = 1 \) \((X_i \sim \text{Ex}(1))\) vs. \( K : \theta = \theta_1 \) \((X_i \sim \text{We}(\theta_1, 1))\). From section §4.2 (B&D) we know that the likelihood ratio test is the most powerful test of \( H \) vs. \( K \).

Since the only two hypotheses being considered restrict the observations to \( \mathbb{R}_+ \), we will continue with the likelihood functions over that domain. If in fact \( x_i < 0 \) were observed for any \( i \), then of course both hypotheses would have likelihood 0 and one would probably want to reconsider one’s models altogether.

Our test statistic should be:

\[
T = \frac{L(H)}{L(K)} = \frac{\prod \left( \theta_1 e^{-x_i^{\theta_1}} x_i^{\theta_1 - 1} \right)}{\prod (e^{-x_i})}
\]

10
Taking the log of the likelihood ratio will give an equivalent MP test.

\[ \log T = \sum \left[ \log \theta_i - x_i^\theta_1 + (\theta_1 - 1) \log x_i \right] - \sum [-x_i] \]

\[ = n \log \theta_1 - \sum x_i^\theta_1 + (\theta_1 - 1) \sum \log x_i + \sum x_i \]

And since \( n \log \theta_1 \) is a constant in \( x \), it may be dropped out of the test statistic, leaving:

\[ T' = - \sum x_i^\theta_1 + (\theta_1 - 1) \sum \log x_i + \sum x_i \]

The most powerful (MP) test of \( H \) against \( K \) will reject \( H \) in favor of \( K \) when \( T' \) (defined above) is sufficiently large, where “sufficiently large” is determined by its distribution under \( H \), which is not derived here.

In order for this test to be uniformly most powerful (UMP), it must have the same rejection region for all possible values of \( \theta_1 \). (Otherwise, you’d have different MP tests for \( \theta_1 \) and for \( \theta_2 \) rather than a single UMP test.) But the statistic \( T' \) depends upon \( \theta_1 \) in such a way that a different value of \( \theta_1 \) will give a different rejection region. So it can’t be UMP.

---

8. (10 pt) Bickel & Doksum, page 290: 4.9.1

Solution: The MLE of \( \theta \) in the binomial setting is \( \frac{X}{n} \) (not proven here). Thus, for testing \( H: \theta = \frac{1}{2} \) versus \( K: \theta \neq \frac{1}{2} \), the likelihood ratio is:

\[ T(x) = \frac{\binom{n}{x} \left( \frac{x}{n} \right)^x \left( 1 - \frac{x}{n} \right)^{n-x}}{\binom{n}{x} \left( \frac{1}{2} \right)^x \left( \frac{1}{2} \right)^{n-x}} \]

Clearly the common term \( \binom{n}{x} \) in the numerator and denominator of this expression may cancel. Furthermore, because \( \log \) is a strictly increasing function, we may take the log of this test statistic to find an equivalent test. Similarly, we may subtract any constants in \( x \).

\[ \log T(x) = x \log \left( \frac{x}{n} \right) + (n - x) \log \left( 1 - \frac{x}{n} \right) - n \log 2 \]

\[ = x [\log x - \log n] + (n - x) [\log(n - x) - \log n] - n \log 2 \]

\[ = x \log x - x \log n + n \log(n - x) - n \log n - x \log(n - x) + x \log n - n \log 2 \]

\[ = x \log x + (n - x) \log(n - x) + C \]
Thus, we may consider a test using the statistic \( T'(x) = x \log x + (n - x) \log(n - x) \). We wish to show that this test statistic is equivalent to \( |2x - n| \). It will suffice to show that (1) \( T'(x) \) is symmetric about \( x = \frac{n}{2} \) and (2) for \( x > \frac{n}{2} \), \( T'(x) \) is strictly increasing in \( x \). Once those two things are established, it follows that each rejection region \( \{ x : T'(x) > c \} \) is equivalent to exactly one rejection region \( \{ x : |x - \frac{n}{2}| > d \} \) for some bijection between \( c \) and \( d \). Multiplying by 2 gives the given test statistic \( |2x - n| \).

(1). \( T'(n - x) = (n - x) \log(n - x) + x \log(x) = T'(x) \). So symmetry is established.
(2). Consider the continuous extension of the statistic to all reals. Then we have:

\[
\frac{dT'}{dx} = x \cdot \frac{1}{x} + \log x + (n - x) \cdot -\frac{1}{n - x} - \log(n - x) \\
= \log x - \log(n - x) \\
= \log \left( \frac{x}{n - x} \right) 
\]

This derivative will be positive whenever \( \frac{x}{n-x} > 1 \), or equivalently, whenever \( x > n-x \), or \( x > \frac{n}{2} \). So our test statistic \( T' \) increases for \( x > \frac{n}{2} \).

Since (1) and (2) are both true, then our MP test is equivalent to a test on \( |2x - n| \).