Problem 1 (8.16)

\(X_1, \ldots, X_n\) i.i.d. with density function 
\[ f(x|\sigma) = \frac{1}{2\sigma} \exp \left( -\frac{|x|}{\sigma} \right) \]

(a) − (c) (See HW 4 Solutions)

(d) According to Corollary A on page 309 of the text, the maximum likelihood estimate is a function of a sufficient statistic \(T\). In part (b), the maximum likelihood estimate was found to be
\[ \hat{\sigma}_{MLE} = \frac{1}{n} \sum_{i=1}^{n} |x_i| \]

Therefore, a sufficient statistic \(T(X_1, X_2, \ldots, X_n)\) is given by:
\[ T(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} |X_i| \]

Problem 2 (8.52)

\(X_1, \ldots, X_n\) i.i.d. with density function 
\[ f(x|\theta) = (\theta + 1)x^\theta, \quad 0 \leq x \leq 1 \]

(a)

\[ E[X] = \int_0^1 x f(x|\theta)dx = \int_0^1 x(\theta + 1)x^\theta dx = \left. \frac{\theta + 1}{\theta + 2} x^{\theta + 1} \right|_0^1 = \frac{\theta + 1}{\theta + 2} \]
Therefore, a method of moments estimate of $\theta$ is given by:

\[
\hat{\mu}_1 = \frac{\hat{\theta}_{MM} + 1}{\hat{\theta}_{MM} + 2}
\]

\[
\Rightarrow \hat{\mu}_1 (\hat{\theta}_{MM} + 2) = \hat{\theta}_{MM} + 1
\]

\[
\Rightarrow (\hat{\mu}_1 - 1) \hat{\theta}_{MM} = 1 - 2\hat{\mu}_1
\]

\[
\Rightarrow \hat{\theta}_{MM} = \frac{1 - 2\hat{\mu}_1}{\hat{\mu}_1 - 1}
\]

where $\hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}_n$.

(b)

\[
l(\theta) = \sum_{i=1}^{n} [\log(\theta + 1) + \theta \log(x_i)]
\]

\[
\Rightarrow \frac{d}{d\theta} l(\theta) = \frac{n}{\theta + 1} + \sum_{i=1}^{n} \log(x_i)
\]

\[
\Rightarrow 0 = \frac{n}{\hat{\theta}_{MLE} + 1} + \sum_{i=1}^{n} \log(x_i)
\]

\[
\Rightarrow \hat{\theta}_{MLE} = -\frac{n}{\sum_{i=1}^{n} \log(x_i)} - 1
\]

(c)

\[
I(\theta) = -E \left[ \frac{\partial^2}{\partial \theta^2} \log f(X|\theta) \right]
\]

\[
= -E \left[ \frac{\partial^2}{\partial \theta^2} (\log(\theta + 1) + \theta \log X) \right]
\]

\[
= -E \left[ \frac{\partial}{\partial \theta} \left( \frac{1}{\theta + 1} + \log X \right) \right]
\]

\[
= -E \left[ -\frac{1}{(\theta + 1)^2} \right]
\]

\[
\approx -\frac{1}{(\theta + 1)^2}
\]

\[
\text{Var} \left[ \hat{\theta}_{MLE} \right] \approx \frac{1}{nI(\theta)}
\]

\[
= \frac{(\theta + 1)^2}{n}
\]

(d) According to Corollary A on page 309 of the text, the maximum likelihood estimate is a function of a sufficient statistic $T$. In part (b), the maximum likelihood estimate was found to be

\[
\hat{\theta}_{MLE} = -\frac{n}{\sum_{i=1}^{n} \log(x_i)} - 1
\]
Therefore, a sufficient statistic \( T(X_1, X_2, \ldots, X_n) \) is given by:

\[
T(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} \log(X_i)
\]

**Problem 3 (8.59)**

Let \( I \) denote the event that a pair of twins is identical, so \( P(I) = \alpha \).

(a) \[
P(MM) = P(MM|I)P(I) + P(MM|I^C)P(I^C)
= \frac{1}{2}\alpha + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)(1-\alpha)
= \frac{1 + \alpha}{4}
\]
\[
P(FF) = P(FF|I)P(I) + P(FF|I^C)P(I^C)
= \frac{1}{2}\alpha + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)(1-\alpha)
= \frac{1 + \alpha}{4}
\]
\[
P(MF) = 1 - (P(MM) + P(FF))
= 1 - \frac{1 + \alpha}{2}
= \frac{1 - \alpha}{2}
\]

(b) We will assume that \( n \) sets of twins are sampled, so \( n_1 + n_2 + n_3 = n \).

\[
\text{lik}(\alpha) = \left(\frac{1 + \alpha}{4}\right)^{n_1} + \left(\frac{1 + \alpha}{4}\right)^{n_2} + \left(\frac{1 - \alpha}{2}\right)^{n_3}
\Rightarrow l(\alpha) = (n_1 + n_2)\log\left(\frac{1 + \alpha}{4}\right) + n_3\log\left(\frac{1 - \alpha}{2}\right)
\Rightarrow \frac{d}{d\alpha} l(\alpha) = \frac{n_1 + n_2}{1 + \alpha} - \frac{n_3}{1 - \alpha}
\Rightarrow 0 = \frac{n_1 + n_2}{1 + \hat{\alpha}_{MLE}} - \frac{n_3}{1 - \hat{\alpha}_{MLE}}
\Rightarrow \hat{\alpha}_{MLE} = \frac{n_1 + n_2 - n_3}{n}
\]

Now to compute the variance of \( \hat{\alpha}_{MLE} \), we will rewrite \( \hat{\alpha}_{MLE} \) as

\[
\hat{\alpha}_{MLE} = \frac{n_1 + n_2 - n_3}{n}
= \frac{n_1 + n_2 - (n - n_1 - n_2)}{n}
= \frac{2(n_1 + n_2) - n}{n}
\]
Then the variance of the MLE can be computed as

\[
\text{Var}[\hat{\alpha}_{MLE}] = \text{Var}\left[\frac{2(n_1 + n_2) - n}{n}\right] = \frac{4}{n^2}\text{Var}[n_1 + n_2] = \frac{4}{n^2}(\text{Var}[n_1] + \text{Var}[n_2] + 2\text{Cov}(n_1, n_2))
\]

We note that \(n_1\) and \(n_2\) are both Binomial random variables with \(n\) trials and success probability \(\frac{1+\alpha}{4}\), so

\[
\text{Var}[n_1] = \text{Var}[n_2] = n\left(\frac{1+\alpha}{4}\right)\left(\frac{3-\alpha}{4}\right)
\]

Now we define \(Y_i = 1\{i^{th} \text{ set of twins is } MM\}\) and \(X_i = 1\{i^{th} \text{ set of twins is } FF\}\). Clearly \(n_1 = \sum_{i=1}^{n} Y_i\) and \(n_2 = \sum_{i=1}^{n} X_i\), and also \(Y_i X_i = 0\) since a given set of twins cannot be both two males and two females. Using these definitions, we have

\[
\text{Cov}(n_1, n_2) = E[n_1 n_2] - E[n_1]E[n_2]
\]

\[
= E\left[\sum_{i=1}^{n} Y_i \left(\sum_{j=1}^{n} X_j\right)\right] - \frac{n(1+\alpha)n(1+\alpha)}{4} = E\left[\sum_{i=1}^{n} Y_i X_i + \sum_{i\neq j} Y_i X_j\right] - n^2\left(\frac{1+\alpha}{4}\right)^2
\]

\[
= E\left[\sum_{i=1}^{n} 0 + \sum_{i\neq j} Y_i X_j\right] - n^2\left(\frac{1+\alpha}{4}\right)^2
\]

\[
= 0 + (n^2 - n)E[Y_i]E[X_j] - n^2\left(\frac{1+\alpha}{4}\right)^2
\]

\[
= (n^2 - n)\left(\frac{1+\alpha}{4}\right)^2 - n^2\left(\frac{1+\alpha}{4}\right)^2
\]

\[
= -n\left(\frac{1+\alpha}{4}\right)^2
\]

Substituting these results back into the expression for \(\text{Var}[\hat{\alpha}_{MLE}]\), we have

\[
\text{Var}[\hat{\alpha}_{MLE}] = \frac{4}{n^2}\left[n\left(\frac{1+\alpha}{4}\right)\left(\frac{3-\alpha}{4}\right) + n\left(\frac{1+\alpha}{4}\right)\left(\frac{3-\alpha}{4}\right) + 2\left(-n\left(\frac{1+\alpha}{4}\right)^2\right)\right]
\]

\[
= \frac{1}{n}\left[\frac{(1+\alpha)(3-\alpha)}{2} + \frac{(1+\alpha)^2}{2}\right]
\]

\[
= \frac{(1+\alpha)}{2n}
\]

\[
= \frac{2(1+\alpha)}{n}
\]
Problem 4 (8.68)

\( X_1, \ldots, X_n \) i.i.d. with probability mass function function  \( p(x|\lambda) = \frac{1}{x!} \lambda^x e^{-\lambda} \)

(a) To show that \( T = \sum_{i=1}^n X_i \) is sufficient for \( \lambda \), we first note that \( T \) has a Poisson distribution with parameter \( n\lambda \), so we have:

\[
P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n | T = t) = \frac{P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n, T = t)}{P(T = t)}
\]

\[
= \frac{P(X_1 = x_1, X_2 = x_2, \ldots, X_n = t - \sum_{i=1}^{n-1} x_i)}{P(T = t)}
\]

\[
= \frac{\left[ \prod_{i=1}^{n-1} \lambda x_i e^{-\lambda / x_i} \right] \left[ \lambda^{(t-\sum_{i=1}^{n-1} x_i)} e^{-\lambda / \left( t - \sum_{i=1}^{n-1} x_i \right)} \right]}{(n\lambda)^t e^{-n\lambda / t!}}
\]

\[
e^{-n\lambda \sum_{i=1}^{n} x_i} \sum_{i=1}^{n-1} \left( \prod_{i=1}^{n-1} 1 / x_i! \right) \lambda^{(t-\sum_{i=1}^{n-1} x_i)} e^{-\lambda / \left( t - \sum_{i=1}^{n-1} x_i \right)}
\]

\[
e^{-n\lambda} \lambda^{t} \left( \prod_{i=1}^{n-1} 1 / x_i! \right) \left[ \lambda^{(t-\sum_{i=1}^{n-1} x_i)} e^{-\lambda / \left( t - \sum_{i=1}^{n-1} x_i \right)} \right]
\]

\[
= \frac{\lambda^{t} \left( \prod_{i=1}^{n-1} 1 / x_i! \right) \left[ \lambda^{(t-\sum_{i=1}^{n-1} x_i)} e^{-\lambda / \left( t - \sum_{i=1}^{n-1} x_i \right)} \right]}{n^t / t!}
\]

Since the distribution of \( X_1, \ldots, X_n \) given \( T \) does not depend on \( \lambda \), \( T = \sum_{i=1}^n X_i \) is sufficient.

(b) To show that \( X_1 \) is not sufficient, we again compute the distribution of \( X_1, \ldots, X_n \) given \( X_1 \):

\[
P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n | X_1 = x_1) = \frac{P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n, X_1 = x_1)}{P(X_1 = x_1)}
\]

\[
= \frac{P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)}{P(X_1 = x_1)}
\]

\[
= \frac{\prod_{i=1}^{n} \lambda^{x_i} e^{-\lambda / x_i!}}{\lambda^{x_1} e^{-\lambda / x_1!}}
\]

\[
= \prod_{i=2}^{n} \lambda^{x_i} e^{-\lambda / x_i!}
\]

Since this distribution still depends on \( \lambda \), \( X_1 \) is not sufficient.
(c) According to Theorem A of Section 8.8.1, the statistic $T$ is sufficient if and only if the density $f(x_1, \ldots, x_n|\lambda)$ can be factored as

$$f(x_1, \ldots, x_n|\lambda) = g(T(x_1, \ldots, x_n), \lambda)h(x_1, \ldots, x_n)$$

For the Poisson density and the statistic $T = \sum_{i=1}^{n} X_i$, we can write

$$f(x_1, \ldots, x_n|\lambda) = \prod_{i=1}^{n} \frac{\lambda x_i e^{-\lambda}}{x_i!}$$

$$= \left[ \lambda \sum_{i=1}^{n} x_i e^{-\lambda} \right] \left[ \prod_{i=1}^{n} \frac{1}{x_i!} \right]$$

$$= \left[ \lambda^T e^{-n\lambda} \right] \left[ \prod_{i=1}^{n} \frac{1}{x_i!} \right]$$

$$= g(T(x_1, \ldots, x_n), \lambda)h(x_1, \ldots, x_n)$$

where

$$g(T, \lambda) = \lambda^T e^{-n\lambda}$$

$$h(x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{x_i!}$$

Problem 5

$X_1, \ldots, X_n$ i.i.d. with density function $f(x|\mu, \tau^2, p) = pf_1(x|\mu) + (1-p)f_2(x|\mu, \tau^2)$, where

$$f_1(x|\mu) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{(x - \mu)^2}{2} \right\}$$

is the $N(\mu, 1)$ density, and

$$f_2(x|\mu, \tau) = \frac{1}{\sqrt{2\pi\tau^2}} \exp \left\{ -\frac{(x - \mu)^2}{2\tau^2} \right\}$$

is the $N(\mu, \tau^2)$ density. Then the expectation of a random variable with this mixture density is given by:

$$E[X_i] = \int_{-\infty}^{\infty} x f(x|\mu, \tau^2, p)dx$$

$$= \int_{-\infty}^{\infty} x \left( pf_1(x|\mu) + (1-p)f_2(x|\mu, \tau^2) \right) dx$$

$$= p \int_{-\infty}^{\infty} x f_1(x|\mu)dx + (1-p) \int_{-\infty}^{\infty} x f_2(x|\mu, \tau^2)dx$$

$$= p\mu + (1-p)\mu$$

$$= \mu$$
To calculate the variance of a random variable with this mixture density, we use the fact that $\text{Var}[X] = E[X^2] - (E[X])^2$, where $E[X^2]$ is given by:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x|\mu, \tau^2, p) dx$$

$$= \int_{-\infty}^{\infty} x^2 \left( pf_1(x|\mu) + (1-p)f_2(x|\mu, \tau^2) \right) dx$$

$$= p \int_{-\infty}^{\infty} x^2 f_1(x|\mu) dx + (1-p) \int_{-\infty}^{\infty} x^2 f_2(x|\mu, \tau^2) dx$$

$$= p(1 + \mu^2) + (1-p)(\tau^2 + \mu^2)$$

$$= \mu^2 + p + (1-p)\tau^2$$

So we have

$$\text{Var}[X] = E[X^2] - (E[X])^2$$

$$= \mu^2 + p + (1-p)\tau^2 - \mu^2$$

$$= p + (1-p)\tau^2$$

**Problem 6**

For the mixture density of problem 5, the variance of the sample mean is given by

$$\text{Var}[\bar{X}_n] = \frac{\text{Var}[X_i]}{n}$$

$$= \frac{p + (1-p)\tau^2}{n}$$

And for large $n$, we have that

$$\sqrt{n} (M_n - M) \rightarrow_d N \left(0, \frac{1}{4f^2(M)} \right)$$

where $M$ is the true median of the distribution. Since this distribution is symmetric about $\mu$, we have that $M = \mu$, and therefore

$$f^2(M) = f^2(\mu)$$

$$= \left( p\frac{1}{\sqrt{2\pi}} + (1-p)\frac{1}{\sqrt{2\pi}\tau^2} \right)^2$$

$$= p^2 \frac{1}{2\pi} + p(1-p)2 \frac{1}{2\pi\tau^2} + (1-p)^2 \frac{1}{2\pi\tau^2}$$
Thus we have
\[
\text{Var}[\sqrt{n}(M_n - M)] \approx \frac{1}{4f^2(M)}
\]
\[
\Rightarrow \quad \text{Var}[\sqrt{n}M_n] \approx \frac{1}{4f^2(M)}
\]
\[
\Rightarrow \quad n\text{Var}[M_n] \approx \frac{1}{4f^2(M)}
\]
\[
\Rightarrow \quad \text{Var}[M_n] \approx \frac{1}{4nf^2(M)}
\]
\[
\Rightarrow \quad \text{Var}[M_n] \approx \frac{1}{4n \left(p^2 \frac{1}{m^2} + p(1-p)2\frac{1}{m^2} + (1-p)^2 \frac{1}{m^2}\right)}
\]

When \(p = 0.9\) and \(\tau = 5\),
\[
\text{Var}[X_n] = \frac{0.9 + 0.1(25)}{n} \approx \frac{1}{3.4}
\]
\[
\Rightarrow \quad \text{Var}[M_n] \approx \frac{1}{4n \left(0.92 \frac{1}{m^2} + 0.9(0.1)2\frac{1}{m^2} + 0.1^2 \frac{1}{m^2}\right)}
\]
\[
\approx \frac{1}{n} \times 1.8559
\]

So the ratio of the asymptotic variances for \(p = 0.9\) and \(\tau = 5\) is
\[
\frac{\text{Var}[X_n]}{\text{Var}[M_n]} = \frac{3.4/n}{1.8559/n} = \frac{1.8320}{n}
\]

A confidence interval for \(\mu\) based on \(X_n\) is given by
\[
X_n \pm z_{1-\alpha/2} \sqrt{\text{Var}[X_n]}
\]

In order for a 95% confidence interval to have length 0.1, we must have
\[
z_{1-0.05/2} \sqrt{\text{Var}[X_n]} = 0.05
\]
\[
\Rightarrow \quad 1.96 \sqrt{\frac{3.4}{n}} = 0.05
\]
\[
\Rightarrow \quad 39.2 = \sqrt{n}
\]
\[
\Rightarrow \quad 1536.64 = \frac{n}{3.4}
\]
\[
\Rightarrow \quad n = 5224.576
\]

Thus a sample of size 5225 is needed to give a 95% confidence interval based on \(X_n\) with length \(\leq 0.1\).

Similarly, a confidence interval for \(\mu\) based on \(M_n\) is given by
\[
M_n \pm z_{1-\alpha/2} \sqrt{\text{Var}[M_n]}
\]

8
In order for this 95% confidence interval to have length 0.1, we must have
\[ z_{1-0.05/2} \sqrt{\text{Var}[M_n]} = 0.05 \]
\[ \Rightarrow 1.96 \sqrt{\frac{1.8559}{n}} = 0.05 \]
\[ \Rightarrow 39.2 = \sqrt{\frac{n}{1.8559}} \]
\[ \Rightarrow 1536.64 = \frac{n}{1.8559} \]
\[ \Rightarrow n = 2851.85 \]

So a sample of size 2852 is needed to give a 95% confidence interval based on \( M_n \) with length \( \leq 0.1 \).

**Problem 7**

\( X_1, \ldots, X_n \) i.i.d. according to the Cauchy distribution with density function
\[ f(x|\theta) = \frac{1}{\pi} \cdot \frac{1}{1 + (x - \theta)^2} \]

If \( \hat{\theta}_n \) is the sample median, we have for large \( n \) that
\[ \sqrt{n} \left( \hat{\theta}_n - M \right) \rightarrow_{d} \mathcal{N} \left( 0, \frac{1}{4f^2(M)} \right) \]

where \( M \) is the true median of the distribution. Since this distribution is symmetric about \( \theta \), we have that \( M = \theta \), and therefore \( f^2(M) = f^2(\theta) = \left( \frac{1}{\pi} \right)^2 \). This gives
\[ \sqrt{n} \left( \hat{\theta}_n - \theta \right) \rightarrow_{d} \mathcal{N} \left( 0, \frac{\pi^2}{4} \right) \]
\[ \Rightarrow \frac{\sqrt{n} \left( \hat{\theta}_n - \theta \right)}{\pi/2} \rightarrow_{d} \mathcal{N}(0, 1) \]

Using this limiting distribution, we have
\[ P \left( \left| \hat{\theta}_n - \theta \right| \leq \frac{1}{5} \right) = P \left( -\frac{1}{5} \leq \hat{\theta}_n - \theta \leq \frac{1}{5} \right) \]
\[ = P \left( -\frac{\sqrt{n}}{5} \leq \sqrt{n} \left( \hat{\theta}_n - \theta \right) \leq \frac{\sqrt{n}}{5} \right) \]
\[ = P \left( -\frac{\sqrt{n} \frac{2}{5}}{\pi} \leq \frac{\sqrt{n} \left( \hat{\theta}_n - \theta \right)}{\pi/2} \leq \frac{\sqrt{n} \frac{2}{5}}{\pi} \right) \]
\[ = \Phi \left( \frac{2\sqrt{n}}{5\pi} \right) - \Phi \left( \frac{-2\sqrt{n}}{5\pi} \right) \]
\[ = \Phi \left( \frac{2\sqrt{101}}{5\pi} \right) - \Phi \left( \frac{-2\sqrt{101}}{5\pi} \right) \]
\[ = 0.7993 \]
If instead we use an efficient estimator $\hat{\theta}_n$ that satisfies

$$\sqrt{n} \left( \hat{\theta}_n - \theta \right) \rightarrow_d N \left( 0, \frac{1}{I(\theta)} \right)$$

$$\Rightarrow \quad \sqrt{n} \left( \hat{\theta}_n - \theta \right) \rightarrow_d N (0, 2)$$

$$\Rightarrow \quad \frac{\sqrt{n} \left( \hat{\theta}_n - \theta \right)}{\sqrt{2}} \rightarrow_d N (0, 1)$$

then we have

$$P \left( \left| \hat{\theta}_n - \theta \right| \leq \frac{1}{5} \right) = P \left( -\frac{1}{5} \leq \hat{\theta}_n - \theta \leq \frac{1}{5} \right)$$

$$= P \left( -\frac{\sqrt{n}}{5} \leq \sqrt{n} \left( \hat{\theta}_n - \theta \right) \leq \frac{\sqrt{n}}{5} \right)$$

$$= P \left( -\frac{\sqrt{n}}{5} \frac{1}{\sqrt{2}} \leq \frac{\sqrt{n} \left( \hat{\theta}_n - \theta \right)}{\sqrt{2}} \leq \frac{\sqrt{n}}{5} \frac{1}{\sqrt{2}} \right)$$

$$= \Phi \left( \frac{\sqrt{n}}{5\sqrt{2}} \right) - \Phi \left( -\frac{\sqrt{n}}{5\sqrt{2}} \right)$$

$$= \Phi \left( \frac{\sqrt{101}}{5\sqrt{2}} \right) - \Phi \left( -\frac{\sqrt{101}}{5\sqrt{2}} \right)$$

$$= 0.8448$$

Thus the efficient estimator $\hat{\theta}_n$ has a higher probability of being within 0.2 of the true value $\theta$, as expected.