7.11 (a) The likelihood function is

\[ L(\theta|x) = f(x|\theta) = \prod_{i=1}^{n} \theta x_i^{\theta-1} = \theta^n \left( \prod_{i=1}^{n} x_i \right)^{\theta-1} \]

and the loglikelihood function is

\[ l = \log L(\theta|x) = n \log \theta + (\theta - 1) \log \prod_{i=1}^{n} x_i = n \log \theta + (\theta - 1) \sum_{i=1}^{n} \log x_i \]

Thus

\[ \frac{dl}{d\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i. \]

We set \( l' = 0 \) and solve for \( \theta \). This gives

\[ \hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \log x_i} = n \sum_{i=1}^{n} x_i^{-1} \]

The second derivative is \(-n/\theta^2 < 0\), so this is the MLE.

Let \( Y = \log X^{-1} \). What is this random variable \( Y \)? We do a transformation with \( X = e^{-Y} \) and conclude that

\[ f_Y(y) = \theta e^{-\theta y} \quad y > 0, \]

i.e., \( Y \sim \text{exponential with parameter } 1/\theta \). Thence \( T = \sum_i Y_i \sim \text{gamma}(n, 1/\theta) \) and \( \hat{\theta} = n/T \), where

\[ f_T(t) = \frac{1}{\Gamma(n) (\frac{1}{\theta})^n} t^{n-1} e^{-\theta t} = \frac{\theta^n}{\Gamma(n)} t^{n-1} e^{-\theta t} \quad t \geq 0. \]

Now we find the first two moments of the random variable \( 1/T \):

\[ E\left( \frac{1}{T} \right) = \int_0^\infty \frac{1}{t \Gamma(n)} t^{n-1} e^{-\theta t} dt = \frac{\theta^n}{\Gamma(n)} \int_0^\infty t^{n-2} e^{-\theta t} dt \]

Recall the complete gamma function (e.g., p. 99 of the text):

\[ \Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt \]

Now we do a change of variable using \( u = \theta t \). This gives

\[ \hat{\theta} = \frac{\theta^n}{\Gamma(n)} \theta^{-(n-1)} \Gamma(n-1) = \frac{\theta}{n-1}. \]
Similarly, we have
\[
E\left( \frac{1}{T} \right)^2 = \int_0^{\infty} \frac{1}{t^2 \Gamma(n)} t^{n-1} e^{-\theta t} dt = \frac{\theta^n}{\Gamma(n)} \int_0^{\infty} t^n e^{-\theta t} dt = \frac{\theta^2}{(n-1)(n-2)}
\]

Using the above, we find
\[
E\left( \frac{\theta}{1} \right) = nE\left( \frac{1}{T} \right) = \frac{n\theta}{n-1}
\]
and
\[
Var\left( \frac{\theta}{1} \right) = \frac{n^2\theta^2}{(n-1)(n-2)} - \left( \frac{n\theta}{n-1} \right)^2
\]
\[
= \frac{n^2\theta^2(n-1) - n^2\theta^2(n-2)}{(n-1)^2(n-2)} = \frac{n^2\theta^2}{(n-1)^2(n-2)} \to 0 \quad \text{as } n \to \infty.
\]

\[(b)\] We recognize that the density for \(X_i\) is beta \((\theta, 1)\). Hence its mean is given by \(\frac{\theta}{1+\theta}\). So we equate the sample moment and its expectation:
\[
\frac{1}{n} \sum_{i=1}^{n} X_i = \frac{\theta}{1+\theta} \quad \text{or} \quad \bar{\theta} = \frac{\bar{X}}{1-\bar{X}} \quad \text{or} \quad \bar{\theta} = \frac{\sum_{i=1}^{n} X_i}{n - \sum_{i=1}^{n} X_i} \Box
\]

7.24 Let \(Y = \sum_i X_i\). We know that \(Y \sim Poisson(n\lambda)\).
\[(a)\] The marginal mass function of \(Y\) is
\[
m(y) = \int_0^{\infty} \frac{(n\lambda)^y e^{-n\lambda}}{y!} \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta} d\lambda
\]
\[
= \frac{n^y}{y!\Gamma(\alpha)\beta^\alpha} \int_0^{\infty} \lambda^y e^{-n\lambda} \lambda^{\alpha-1} e^{-\lambda/\beta} d\lambda
\]
\[
= \frac{n^y}{y!\Gamma(\alpha)\beta^\alpha} \int_0^{\infty} \lambda^{(y+\alpha)-1} \exp\left(-\frac{\lambda}{\beta/(n\beta+1)}\right) d\lambda
\]
Now we use the gamma function
\[
\int_0^{\infty} t^{\alpha-1} e^{-t} dt = \Gamma(\alpha)
\]
to clean up \(m(y)\). Let
\[
t = \frac{\lambda}{\beta/(n\beta+1)} \implies \lambda = \frac{\beta}{n\beta+1} t \implies d\lambda = \frac{\beta}{n\beta+1} dt
\]
Hence
\[
m(y) = \frac{n^y}{y!\Gamma(\alpha)\beta^\alpha} \int_0^{\infty} \left(\frac{\beta}{n\beta+1}\right)^{y+\alpha} t^{y+\alpha-1} e^{-t} dt
\]
\[
= \frac{n^y}{y!\Gamma(\alpha)\beta^\alpha} \left(\frac{\beta}{n\beta+1}\right)^{y+\alpha} \Gamma(y+\alpha)
\]
Now the joint density \( f(y, \lambda) \) is given by
\[
f(y, \lambda) = f(y|\lambda)\pi(\lambda) = \frac{(n\lambda)^y e^{-n\lambda}}{y!} \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta}
\]
The posterior density for \( \lambda \) is
\[
f(\lambda | y) = \frac{f(y, \lambda)}{m(y)} = \frac{(n\lambda)^y e^{-n\lambda}}{y!} \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta}
\]
Let \( m(y) = \frac{n\lambda}{\Gamma(\alpha)\beta^{\alpha+1}} \left( \frac{\beta}{n\beta+1} \right)^{\frac{y+\alpha}{\beta}} \Gamma(y+\alpha) \)
\[
= \left( \frac{\beta}{n\beta+1} \right)^{\frac{y+\alpha}{\beta}} \Gamma(y+\alpha)
\]
\[
\exp \left( -\frac{\lambda}{\beta/(n\beta+1)} \right)
\]
Let \( B = \beta/(n\beta+1) \) and \( Z = y + \alpha \). Then the above can be written as
\[
f(\lambda | y) = \frac{1}{B^\alpha \Gamma(Z)} \lambda^{Z-1} \exp \left( -\frac{\lambda}{B} \right) \quad \lambda > 0
\]
We now conclude that \( \lambda|y \sim \text{gamma}(Z, B) \).

(b) From p. 624 of the text, we know
\[
E[\lambda|Y] = Z \times B = (Y + \alpha) \frac{\beta}{n\beta+1} = \frac{Y + \alpha}{n + \frac{1}{\beta}}
\]
and
\[
Var[\lambda|Y] = Z \times B^2 = \frac{Y + \alpha}{\left( n + \frac{1}{\beta} \right)^2}
\]

7.37 To find a best unbiased estimator of \( \theta \), we first find a complete sufficient statistic. The joint density is
\[
f(x|\theta) = \left( \frac{1}{2\theta} \right)^n \prod_{i=1}^{n} I_{(-\theta, \theta)}(x_i) = \left( \frac{1}{2\theta} \right)^n I_{(0, \theta)}(\max_i |x_i|)
\]
Based on the Factorization Theorem, we conclude that \( Y = \max_i |X_i| \) is a sufficient statistic. In order to check for completeness, we first need to know the density of \( Y \). Let \( Z_i = |X_i| \). What is the density of \( Z_i \)? This is given by Example 7c of Ross (p. 24), or it is easily derived, namely, \( Z_i \sim U(0, \theta) \). Now \( Y = \max_i Z_i = Z(n) \). From Example 7.3.13, we know
\[
f_Y(y|\theta) = \frac{ny^{n-1}}{\theta^n} \quad 0 < y < \theta.
\]
Suppose \( g(y) \) is a function such that
\[
E[g(Y)] = \int_0^\theta \frac{ny^{n-1}}{\theta^n} g(y) dy = 0 \quad \forall \theta
\]
Following the argument identical to that given in Example 6.2.23, we conclude that $Y$ is a complete statistic. Now

$$E[Y] = \int_{0}^{\theta} y \frac{n y^{n-1}}{\theta^n} dy = \frac{n}{\theta^n} \int_{0}^{\theta} y^n dy = \frac{n \theta^{n+1}}{\theta^n(n+1)} = \frac{n}{n+1} \theta$$

This implies

$$E\left(\frac{n+1}{n} Y\right) = \theta$$

Applying Theorem 7.3.23, we conclude that 

$$\frac{n+1}{n} \max \{|X_i|\}$$

is the unique best unbiased estimator of its expected value $\theta$. 

\[ \square \]

7.38 We will use Corollary 7.3.15 to solve the problem. To apply the Corollary, we note that the two densities given in (a) and (b) are all of exponential family. Hence the condition for applying the Cramer-Rao Theorem is met (specifically, the interchange of differentiation and integration is permissible, cf. p. 339 of the text).

(a) Note

$$\frac{\partial}{\partial \theta} \log L(\theta|x) = \frac{\partial}{\partial \theta} \log \prod_{i=1}^{n} \theta x_i^{\theta-1} = \frac{\partial}{\partial \theta} \sum_{i=1}^{n} \left[ \log \theta + (\theta - 1) \log x_i \right]$$

$$= \sum_{i=1}^{n} \left[ \frac{1}{\theta} + \log x_i \right] = -n \left[ -\sum_{i=1}^{n} \frac{\log x_i}{n} + \frac{1}{\theta} \right]$$

Let $a(\theta) = -n$, and

$$W(X) = -\sum_{i=1}^{n} \frac{\log X_i}{n}.$$ 

Now

$$E[\log X_i] = \int_{0}^{1} (\log x) (\theta x^{\theta-1}) dx = \int_{0}^{1} (\log x) d(x^\theta)$$

$$= \left[ (\log x) x^\theta \right]_0^1 - \int_{0}^{1} x^\theta d(\log x) = 0 - \frac{1}{\theta}$$

Hence

$$E(W(X)) = -\frac{1}{n} \sum_{i=1}^{n} E \log X_i = \frac{1}{\theta}$$

Hence $W(X)$ is an unbiased estimator of $1/\theta$. By Corollary 7.3.15, we conclude that $W(X)$ attains the CRLB.
(b) Note

\[ \frac{\partial}{\partial \theta} \log L(\theta|x) = \frac{\partial}{\partial \theta} \log \prod_{i=1}^{n} \frac{\log \theta}{\theta - 1} \theta^{x_i} \]

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

\[ = \sum_{i=1}^{n} \left( \frac{1}{\theta \log \theta} - \frac{1}{\theta - 1} \right) + \frac{1}{\theta} \sum_{i=1}^{n} x_i \]

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

Let \( a(\theta) = n/\theta \) and \( W(X) = \bar{x} \). Since \( E(X_i) = E(\bar{X}) \), if we can show that

\[ E(X_i) = \frac{\theta}{\theta - 1} - \frac{1}{\log \theta} \]  

(*)

then we are done. Note

\[ E(X_i) = \int_{0}^{1} x \left( \frac{\log \theta}{\theta - 1} \theta^x \right) dx \]

\[ = \frac{\log \theta}{\theta - 1} \int_{0}^{1} x \theta^x dx \]

We see

\[ \int_{0}^{1} x \theta^x dx = \int_{0}^{1} x d \left( \frac{e^{\log \theta x}}{\log \theta} \right) = \frac{1}{\log \theta} \left[ (x \theta^x)_{0}^{1} - \int_{0}^{1} \theta^x dx \right] \]

\[ = \frac{1}{\log \theta} \left[ \theta - \int_{0}^{1} \theta d \left( \frac{e^{\log \theta x}}{\log \theta} \right) \right] = \frac{1}{\log \theta} \left[ \theta - \frac{1}{\log \theta} [\theta^x]_{0}^{1} \right] \]

\[ = \frac{1}{\log \theta} \left[ \theta - \frac{1}{\log \theta} (\theta - 1) \right] \]

Thus

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

\[ = \frac{1}{\theta - 1} \left[ \theta - \frac{1}{\log \theta} (\theta - 1) \right] = \frac{\theta}{\theta - 1} - \frac{1}{\log \theta} \]

This shows that \( W(X) \) is an unbiased estimator of

\[ g(\theta) = \frac{\theta}{\theta - 1} - \frac{1}{\log \theta} \]

By Corollary 7.3.15, we conclude that \( W(X) \) attains the CRLB.
Problem E. (a) In Example 7.2.7, we have already obtained the MLE of $\theta$:

$$\hat{\theta} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{y}{n} = \frac{13}{16} = 0.8125$$

(b) The following is a plot of the three priors. If you choose beta(1,1) as the prior, this means you believe that prior to your looking at the data, you believe that $\theta$ is equally likely to be anywhere between 0 and 1. If you choose beta(1/2, 1/2) as the prior, it implies that you think that, prior to sampling, you believe that $\theta$ is more likely to lie in either extreme. Finally, for a beta(2,2) prior, you believe that $\theta$ is equally likely to be on either side of 0.5 with the most likely value being 0.5. In each one of the three cases, there is a symmetry in the density around 0.5.

(c) From Example 7.2.14, we have

$$\hat{p}_B = \frac{y + \alpha}{\alpha + \beta + n}$$

Hence

- Under beta(1, 1) : $\hat{p}_B = \frac{13 + 1}{1 + 1 + 16} = 0.7778$
- Under beta(1/2, 1/2) : $\hat{p}_B = \frac{13 + 1/2}{1/2 + 1/2 + 16} = 0.7941$
- Under beta(2, 2) : $\hat{p}_B = \frac{13 + 2}{2 + 2 + 16} = 0.75$

By looking at these results, what conclusion can be drawn in terms of the relation between the posterior probabilities and their respective priors?
(d) Under $\text{beta}(1, 1)$, the posterior density is $\text{beta}(13 + 1, 16 - 13 + 1) = \text{beta}(14, 4)$ (e.g., see p. 325 of the text). Thus

$$\frac{1}{B(14, 4)} \int_{0.6}^{1} \theta^{14-1}(1 - \theta)^{4-1} d\theta = \frac{17!}{13!3!} \int_{0.6}^{1} \theta^{13}(1 - \theta)^3 d\theta = 0.95358$$

Under $\text{beta}(0.5, 0.5)$, the posterior density is $\text{beta}(13 + 0.5, 16 - 13 + 0.5) = \text{beta}(13.5, 3.5)$. Thus

$$\frac{1}{B(13.5, 3.5)} \int_{0.6}^{1} \theta^{13.5-1}(1 - \theta)^{3.5-1} d\theta = \frac{\Gamma(17)}{\Gamma(13.5)\Gamma(3.5)} \int_{0.6}^{1} \theta^{12.5}(1 - \theta)^{2.5} d\theta = 0.96378$$

Under $\text{beta}(2, 2)$, the posterior density is $\text{beta}(13+2, 16-13+2) = \text{beta}(15, 5)$. Thus

$$\frac{1}{B(15, 5)} \int_{0.6}^{1} \theta^{15-1}(1 - \theta)^{5-1} d\theta = \frac{19!}{14!4!} \int_{0.6}^{1} \theta^{14}(1 - \theta)^4 d\theta = 0.93039$$

(e) (f)

The two parts are combined. We see that when the sample size is doubled, the posterior densities have smaller dispersions. This means that for a larger sample, our knowledge about the location of $\theta$ becomes more reliable. Moreover, the effect of prior is diminished as evidenced by the observation that the three posterior densities under $n = 32$ with different priors are staying closer to one another than those under $n = 16$. $\square$