Chapter 3

Unbiased Estimation

Exercise 1. Let $X$ be a sample from $P \in \mathcal{P}$ and $\theta$ be a parameter. Show that if both $T_1(X)$ and $T_2(X)$ are UMVUE’s (uniformly minimum variance unbiased estimators) of $\theta$ with finite variances, then $T_1(X) = T_2(X)$ a.s. $P$ for any $P \in \mathcal{P}$.

Solution. Since both $T_1$ and $T_2$ are unbiased, $T_1 - T_2$ is unbiased for 0. By the necessary and sufficient condition for UMVUE (e.g., Theorem 3.2 in Shao, 2003),

$$E[T_1(T_1 - T_2)] = 0 \quad \text{and} \quad E[T_2(T_1 - T_2)] = 0$$

for any $P$. Then, for any $P \in \mathcal{P}$,

$$E(T_1 - T_2)^2 = E[T_1(T_1 - T_2)] - E[T_2(T_1 - T_2)] = 0,$$

which implies that $T_1 = T_2$ a.s. $P$. □

Exercise 2 (#3.1). Let $(X_1, \ldots, X_n)$ be a sample of binary random variables with $P(X_i = 1) = p \in (0, 1)$.

(i) Find the UMVUE of $p^m$, where $m$ is a positive integer and $m \leq n$.
(ii) Find the UMVUE of $P(X_1 + \cdots + X_m = k)$, where $m$ and $k$ are positive integers and $k \leq m \leq n$.
(iii) Find the UMVUE of $P(X_1 + \cdots + X_{n-1} > X_n)$.

Solution. (i) Let $T = \sum_{i=1}^{n} X_i$. Then $T$ is a complete and sufficient statistic for $p$. By Lehmann-Scheffé’s theorem (e.g., Theorem 3.1 in Shao, 2003), the UMVUE should be $h_m(T)$ with a Borel $h_m$ satisfying $E[h_m(T)] = p^m$.

We now try to find such a function $h_m$. Note that $T$ has the binomial distribution with size $n$ and probability $p$. Hence

$$E[h_m(T)] = \sum_{k=0}^{n} \binom{n}{k} h_m(k)p^k(1-p)^{n-k}.$$
Chapter 3. Unbiased Estimation

Setting $E[h_m(T)] = p^m$, we obtain that

$$
\sum_{k=0}^{n} \binom{n}{k} h_m(k) p^{k-m} (1 - p)^{n-m-(k-m)} = 1
$$

for all $p$. If $m < k$, $p^{k-m} \to \infty$ as $p \to 0$. Hence, we must have $h_m(k) = 0$ for $k = 0, 1, \ldots, m - 1$. Then

$$
\sum_{k=m}^{n} \binom{n}{k} h_m(k) p^{k-m} (1 - p)^{n-m-(k-m)} = 1
$$

for all $p$. On the other hand, from the property of a binomial distribution,

$$
\sum_{k=m}^{n} \binom{n-m}{k-m} p^{k-m} (1 - p)^{n-m-(k-m)} = 1
$$

for all $p$. Hence, $\binom{n}{k} h_m(k) = \binom{n-m}{k-m}$ for $k = m, \ldots, n$. The UMVUE of $p^m$ is

$$
h_m(T) = \begin{cases} 
\frac{(n-m)}{(n)} & T = m, \ldots, n \\
0 & T = 0, 1, \ldots, m - 1.
\end{cases}
$$

(ii) Note that

$$
P(X_1 + \cdots + X_m = k) = \binom{m}{k} p^k (1 - p)^{m-k}$$

$$
= \binom{m}{k} p^k \sum_{j=0}^{m-k} \binom{m-k}{j} (-1)^j p^j$$

$$
= \binom{m}{k} \sum_{j=0}^{m-k} \binom{m-k}{j} (-1)^j p^{j+k}.
$$

By the result in part (i), the UMVUE of $p^{j+k}$ is $h_{j+k}(T)$, where the function $h_{j+k}$ is given in part (i) of the solution, $j = 0, 1, \ldots, m - k$. By Corollary 3.1 in Shao (2003), the UMVUE of $P(X_1 + \cdots + X_m = k)$ is

$$
\binom{m}{k} \sum_{j=0}^{m-k} \binom{m-k}{j} (-1)^j h_{j+k}(T).
$$

(iii) Let $S_{n-1} = X_1 + \cdots + X_{n-1}$. Then $S_{n-1}$ and $X_n$ are independent and $S_{n-1}$ has the binomial distribution with size $n - 1$ and probability $p$. 
Hence,

\[ P(S_{n-1} > X_n) = P(X_n = 0)P(S_{n-1} > 0) + P(X_n = 1)P(S_{n-1} > 1) \]
\[ = P(S_{n-1} > 0) - P(X_n = 1)P(S_{n-1} = 1) \]
\[ = 1 - (1 - p)^{n-1} - (n - 1)p^2(1 - p)^{n-2} \]
\[ = \sum_{j=1}^{n-1} \binom{n-1}{j}(-1)^{j+1}p^j - (n - 1)\sum_{j=0}^{n-2} \binom{n-2}{j}(-1)^{j}p^{j+2} \]
\[ = \sum_{j=1}^{n} c_j p^j, \]

where \( c_1 = n - 1, \ c_n = (-1)^{n+1}(n - 1), \) and

\[ c_j = (-1)^{j+1} \left[ \binom{n-1}{j} + (n - 1)\binom{n-2}{j-2} \right], \quad j = 2, ..., n - 1. \]

The UMVUE of \( P(S_{n-1} > X_n) \) is \( \sum_{j=1}^{n} c_j h_j(T) \) with \( h_j \) defined in part (i) of the solution.

**Exercise 3 (#3.2).** Let \((X_1, ..., X_n)\) be a random sample from \( N(\mu, \sigma^2) \) with an unknown \( \mu \in \mathcal{R} \) and a known \( \sigma^2 > 0 \).

(i) Find the UMVUE’s of \( \mu^3 \) and \( \mu^4 \).

(ii) Find the UMVUE’s of \( P(X_1 \leq t) \) and \( \frac{d}{dt}P(X_1 \leq t) \) with a fixed \( t \in \mathcal{R} \).

**Solution.** (i) Let \( \bar{X} \) be the sample mean, which is complete and sufficient for \( \mu \). Since

\[ 0 = E(\bar{X} - \mu) = E(\bar{X}^3 - 3\mu\bar{X}^2 + 3\mu^2\bar{X} - \mu^3) \]
\[ = E(\bar{X}^3) - 3\mu\sigma^2/n - \mu^3, \]

we obtain that

\[ E[\bar{X}^3 - (3\sigma^2/n)\bar{X}] = E(\bar{X}^3) - 3\mu\sigma^2/n = \mu^3 \]

for all \( \mu \). By Lehmann-Scheffé’s theorem, the UMVUE of \( \mu^3 \) is \( X^3 - (3\sigma^2/n)\bar{X} \). Similarly,

\[ 3\sigma^4 = E(\bar{X} - \mu)^4 \]
\[ = E[\bar{X}(\bar{X} - \mu)^3] \]
\[ = E[\bar{X}^4 - 3\mu\bar{X}^3 + 3\mu^2\bar{X}^2 - \mu^3\bar{X}] \]
\[ = E(\bar{X}^4) - 3\mu(3\mu\sigma^2/n + \mu^3) + 3\mu^2(\sigma^2/n + \mu^2) - \mu^4 \]
\[ = E(\bar{X}^4) - 6\mu^2\sigma^2/n - 4\mu^4 \]
\[ = E(\bar{X}^4) - (6\sigma^2/n)E(\bar{X}^2 - \sigma^2/n) - 4\mu^4. \]

Hence, the UMVUE of \( \mu^4 \) is \( [\bar{X}^4 - (6\sigma^2/n)(\bar{X}^2 - \sigma^2/n) - 3\sigma^4]/4 \).

(ii) Since \( E[P(X_1 \leq t|\bar{X})] = P(X_1 \leq t) \), the UMVUE of \( P(X_1 \leq t) \) is
\( P(X_1 \leq t | \bar{X}) \). From the properties of normal distributions, \((X_1, \bar{X})\) is bivariate normal with mean \((\mu, \mu)\) and covariance matrix

\[
\sigma^2 \begin{pmatrix}
1 & n^{-1} \\
n^{-1} & n^{-1}
\end{pmatrix}.
\]

Consequently, the conditional distribution of \(X_1\) given \(\bar{X}\) is the normal distribution \(N(\bar{X}, (1 - n^{-1})\sigma^2)\). Then, the UMVUE of \(P(X_1 \leq t)\) is

\[
\Phi \left( \frac{t - \bar{X}}{\sigma \sqrt{1 - n^{-1}}} \right),
\]

where \(\Phi\) is the cumulative distribution function of \(N(0, 1)\). By the dominated convergence theorem,

\[
\frac{d}{dt} P(X_1 \leq t) = E \left[ \frac{d}{dt} \Phi \left( \frac{t - X}{\sigma \sqrt{1 - n^{-1}}} \right) \right] = E \left[ \frac{d}{dt} \Phi \left( \frac{t - \bar{X}}{\sigma \sqrt{1 - n^{-1}}} \right) \right].
\]

Hence, the UMVUE of \(\frac{d}{dt} P(X_1 \leq t)\) is

\[
\frac{d}{dt} \Phi \left( \frac{t - \bar{X}}{\sigma \sqrt{1 - n^{-1}}} \right) = \frac{1}{\sigma \sqrt{1 - n^{-1}}} \Phi' \left( \frac{t - \bar{X}}{\sigma \sqrt{1 - n^{-1}}} \right).
\]

Exercise 4 (#3.4). Let \((X_1, ..., X_m)\) be a random sample from \(N(\mu_x, \sigma_x^2)\) and let \(Y_1, ..., Y_n\) be a random sample from \(N(\mu_y, \sigma_y^2)\). Assume that \(X_i\)'s and \(Y_j\)'s are independent.

(i) Assume that \(\mu_x \in \mathbb{R}, \mu_y \in \mathbb{R}, \sigma_x^2 > 0, \text{ and } \sigma_y^2 > 0\). Find the UMVUE's of \(\mu_x - \mu_y\) and \((\sigma_x/\sigma_y)^r\), where \(r > 0\) and \(r < n\).

(ii) Assume that \(\mu_x \in \mathbb{R}, \mu_y \in \mathbb{R}, \text{ and } \sigma_x^2 = \sigma_y^2 > 0\). Find the UMVUE's of \(\sigma_x^2\) and \((\mu_x - \mu_y)/\sigma_x\).

(iii) Assume that \(\mu_x = \mu_y \in \mathbb{R}, \sigma_x^2 > 0, \sigma_y^2 > 0, \text{ and } \sigma_x^2/\sigma_y^2 = \gamma\) is known. Find the UMVUE of \(\mu_x\).

(iv) Assume that \(\mu_x = \mu_y \in \mathbb{R}, \sigma_x^2 > 0, \text{ and } \sigma_y^2 > 0\). Show that a UMVUE of \(\mu_x\) does not exist.

(v) Assume that \(\mu_x \in \mathbb{R}, \mu_y \in \mathbb{R}, \sigma_x^2 > 0, \text{ and } \sigma_y^2 > 0\). Find the UMVUE of \(P(X_1 \leq Y_1)\).

(vi) Repeat (v) under the assumption that \(\sigma_x = \sigma_y\).

Solution: (i) The complete and sufficient statistic for \((\mu_x, \mu_y, \sigma_x^2, \sigma_y^2)\) is \((\bar{X}, \bar{Y}, S_X^2, S_Y^2)\), where \(\bar{X}\) and \(S_X^2\) are the sample mean and variance based on \(X_i\)'s and \(Y\) and \(S_Y^2\) are the sample mean and variance based on \(Y_i\)'s. Therefore \(\bar{X} - \bar{Y}\) is the UMVUE of \(\mu_x - \mu_y\). A direct calculation shows that

\[
E(S_X^r) = \frac{\sigma_x^r}{\kappa_{m-1,r}}.
\]
where
\[ \kappa_{m,r} = \frac{m^r/2 \Gamma\left(\frac{m}{2}\right)}{2^r/2 \Gamma\left(\frac{m+r}{2}\right)}. \]

Hence, the UMVUE of \( \sigma_x^r \) is \( \kappa_{m-1,r}S_X^r \). Similarly, the UMVUE of \( \sigma_y^{-r} \) is \( \kappa_{n-1,-r}S_Y^{-r} \). Since \( S_X \) and \( S_Y \) are independent, the UMVUE of \( (\sigma_x/\sigma_y)^r \) is \( \kappa_{m-1,r}\kappa_{n-1,-r}S_X^r S_Y^{-r} \).

(ii) The complete and sufficient statistic for \( (\mu, \mu_y, \sigma_x^2) \) is \( (\bar{X}, \bar{Y}, S^2) \), where
\[ S^2 = \frac{1}{m+n-2} \left[ \sum_{i=1}^{m} (X_i - \bar{X})^2 + \sum_{j=1}^{n} (Y_j - \bar{Y})^2 \right]. \]

Since \( (m+n-2)S^2/\sigma_x^2 \) has the chi-square distribution \( \chi^2_{m+n-2} \), the UMVUE of \( \sigma_x^2 \) is \( S^2 \) and the UMVUE of \( \sigma_x^{-1} \) is \( \kappa_{m+n-2,-1}S^{-1} \). Since \( \bar{X} - \bar{Y} \) and \( S^2 \) are independent, \( \kappa_{m+n-2,-1}(\bar{X} - \bar{Y})/S \) is the UMVUE of \( (\mu_x - \mu_y)/\sigma_x \).

(iii) The joint distribution of \( X_i \)'s and \( Y_j \)'s is from an exponential family with \( (m\bar{X} + \gamma n\bar{Y}, \sum_{i=1}^{m} X_i^2 + \gamma \sum_{j=1}^{n} Y_j^2) \) as the complete and sufficient statistic for \( (\mu_x, \sigma_x^2) \). Hence, the UMVUE of \( \mu_x \) is \( (m\bar{X} + \gamma n\bar{Y})/(m + \gamma n) \) is a UMVUE of \( \mu_x \) when \( \mathcal{P}_\gamma \) is considered as the family of distributions for \( (X_1, \ldots, X_m, Y_1, \ldots, Y_n) \). Since \( E(T - T_\gamma) = 0 \) for any \( P \in \mathcal{P}_\gamma \) and \( T \) is a UMVUE, \( E[T(T - T_\gamma)] = 0 \) for any \( P \in \mathcal{P}_\gamma \). Similarly, \( E[T_\gamma(T - T_\gamma)] = 0 \) for any \( P \in \mathcal{P}_\gamma \). Then, \( E(T - T_\gamma)^2 = 0 \) for any \( P \in \mathcal{P}_\gamma \) and, thus, \( T = T_\gamma \) a.s. \( \mathcal{P}_\gamma \). Since a.s. \( \mathcal{P}_\gamma \) implies a.s. \( \mathcal{P} \), \( T = T_\gamma \) a.s. \( \mathcal{P} \) for any \( \gamma > 0 \). This shows that \( T \) depends on \( \gamma = \sigma_x^2/\sigma_y^2 \), which is impossible.

(iv) Since \( U = (\bar{X}, \bar{Y}, S_X^2, S_Y^2) \) is complete and sufficient for \( (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2) \), \( P(X_1 \leq Y_1 |U) \) is UMVUE for \( P(X_1 \leq Y_1) \). Note that
\[ P(X_1 \leq t, Y_1 \leq v |U = (\bar{x}, \bar{y}, s_x^2, s_y^2)) = P\left(Z \leq \frac{t - \bar{x}}{s_x}, W \leq \frac{v - \bar{y}}{s_y}\right), \]
where
\[ Z = (X_1 - \bar{X})/S_X \] and
\[ W = (Y_1 - \bar{Y})/S_Y. \] From Example 3.4 in Shao (2003), \( Z \) has Lebesgue density \( f_m(z) \) and \( W \) has Lebesgue density \( f_n(w) \), where
\[ f_k(z) = \frac{\sqrt{k} \Gamma\left(\frac{k-1}{2}\right)}{\sqrt{\pi} (k-1) \Gamma\left(\frac{k-2}{2}\right)} \left[ 1 - \frac{kz^2}{(k-1)^2} \right]^{(k/2)-2} I_{(0, (k-1)/\sqrt{k})}(|z|). \]

Since \( Z \) and \( W \) are independent, the conditional density of \( (X_1, Y_1) \) given \( U \) is
\[ \frac{1}{S_X} f_m\left(\frac{t - \bar{X}}{S_X}\right) \frac{1}{S_Y} f_n\left(\frac{v - \bar{Y}}{S_Y}\right). \]
Hence, the UMVUE is
\[ P(X_1 \leq Y_1 | U) = \frac{1}{S_X S_Y} \int_{-\infty}^{0} \int_{-\infty}^{\infty} f_m \left( \frac{v - \bar{X}}{S_X} \right) f_n \left( \frac{t - v - \bar{Y}}{S_Y} \right) dv. \]

(vi) In this case, \( U = (\bar{X}, \bar{Y}, S^2) \) with \( S^2 \) defined in (ii) is complete and sufficient for \((\mu_x, \mu_y, \sigma_x^2)\). Similar to part (v) of the solution, we have
\[ P(X_1 \leq Y_1 | U = u) = P \left( \frac{(X_1 - \bar{X}) - (Y_1 - \bar{Y})}{\sqrt{m + n - 2S}} \leq r \right), \]
where \( r \) is the observed value of \( R = -(\bar{X} - \bar{Y})/(\sqrt{m + n - 2S}) \). If we denote the Lebesgue density of \( T = [(X_1 - \bar{X}) - (Y_1 - \bar{Y})]/(\sqrt{m + n - 2S}) \) by \( f(t) \), then the UMVUE of \( P(X_1 \leq Y_1) \) is \( \int_{-\infty}^{R} f(t) dt \). To determine \( f \), we consider the orthogonal transformation
\[ (Z_1, ..., Z_{m+n})^\tau = A(X_1, ..., X_m, Y_1, ..., Y_n)^\tau, \]
where \( A \) is an orthogonal matrix of order \( m + n \) whose first three rows are
\[ (m^{-1/2} J_m, 0 J_n), \]
\[ (0 J_m, n^{-1/2} J_n), \]
and
\[ (2 - m^{-1} - n^{-1})^{-1/2}(1 - m^{-1}, -m^{-1} J_{m-1}, n^{-1} - 1, n^{-1} J_{n-1}), \]
and \( J_k \) denotes a row of 1’s with dimension \( k \). Then \( Z_1 = \sqrt{m} \bar{X}, Z_2 = \sqrt{n} \bar{Y}, Z_3 = (2 - m^{-1} - n^{-1})^{-1}[(X_1 - \bar{X}) - (Y_1 - \bar{Y})], (m + n - 2)S^2 = \sum_{i=3}^{m+n} Z_i^2 \), and \( Z_i, i = 3, ..., m + n \), are independent and identically distributed as \( N(0, \sigma^2_x) \). Note that
\[ T = \frac{\sqrt{2 - m^{-1} - n^{-1}} Z_3}{\sqrt{Z_3^2 + Z_4^2 + \cdots + Z_{m+n}^2}}. \]

Then, a direct calculation shows that
\[ f(t) = c_{m,n} \left( 1 - \frac{t^2}{2 - m^{-1} - n^{-1}} \right)^{(m+n-5)/2} I_{(0,\sqrt{2 - m^{-1} - n^{-1}})}(|t|), \]
where
\[ c_{m,n} = \frac{\Gamma \left( \frac{m+n-2}{2} \right)}{\sqrt{\pi} (2 - m^{-1} - n^{-1}) \Gamma \left( \frac{m+n-3}{2} \right)}. \]
Exercise 5 (#3.5). Let \((X_1, \ldots, X_n), n > 2,\) be a random sample from the uniform distribution on the interval \((\theta_1 - \theta_2, \theta_1 + \theta_2),\) where \(\theta_1 \in \mathbb{R}\) and \(\theta_2 > 0.\) Find the UMVUE’s of \(\theta_j, j = 1, 2,\) and \(\theta_1/\theta_2.\)

Solution. Let \(X_{(j)}\) be the \(j\)th order statistic. Then \((X_{(1)}, X_{(n)})\) is complete and sufficient for \((\theta_1, \theta_2).\) Hence, it suffices to find a function of \((X_{(1)}, X_{(n)})\) that is unbiased for the parameter of interest. Let \(Y_i = \left[ X_i - (\theta_1 - \theta_2) \right] / (2\theta_2),\) \(i = 1, \ldots, n.\) Then \(Y_i’s\) are independent and identically distributed as the uniform distribution on the interval \((0, 1).\) Let \(Y_{(j)}\) be the \(j\)th order statistic of \(Y_i’s.\) Then,

\[
E(X_{(n)}) = 2\theta_2 E(Y_{(n)}) + \theta_1 - \theta_2
\]

\[
= 2\theta_2 n \int_0^1 y^n dy + \theta_1 - \theta_2
\]

\[
= \frac{2\theta_2 n}{n + 1} + \theta_1 - \theta_2
\]

and

\[
E(X_{(1)}) = 2\theta_2 E(Y_{(1)}) + \theta_1 - \theta_2
\]

\[
= 2\theta_2 n \int_0^1 y(1-y)^{n-1} dy + \theta_1 - \theta_2
\]

\[
= -\frac{2\theta_2 n}{n + 1} + \theta_1 + \theta_2.
\]

Hence, \(E(X_{(n)} + X_{(1)})/2 = \theta_1\) and \(E(X_{(n)} - X_{(1)}) = 2\theta_2 (n - 1)/(n + 1).\) Therefore, the UMVUE’s of \(\theta_1\) and \(\theta_2\) are, respectively, \((X_{(n)} + X_{(1)})/2\) and \((n + 1)(X_{(n)} + X_{(1)})/[2(n - 1)].\) Furthermore,

\[
E\left( \frac{X_{(n)} + X_{(1)}}{X_{(n)} - X_{(1)}} \right) = E\left( \frac{Y_{(n)} + Y_{(1)}}{Y_{(n)} - Y_{(1)}} \right) + \frac{\theta_1 - \theta_2}{\theta_2} E\left( \frac{1}{Y_{(n)} - Y_{(1)}} \right)
\]

\[
= n(n - 1) \int_0^1 \int_0^y (x+y)(y-x)^{n-3} dxdy + \frac{\theta_1 - \theta_2}{\theta_2} n(n - 1) \int_0^1 \int_0^y (y-x)^{n-3} dxdy
\]

\[
= \frac{n}{n - 2} + \frac{\theta_1 - \theta_2}{\theta_2} \frac{n}{n - 2}
\]

\[
= \frac{n}{n - 2} \frac{\theta_1}{\theta_2}.
\]

Hence the UMVUE of \(\theta_1/\theta_2\) is \(n(n - 2)/(n - 2)(X_{(n)} + X_{(1)})/(X_{(n)} - X_{(1)}).\)

Exercise 6 (#3.6). Let \((X_1, \ldots, X_n)\) be a random sample from the exponential distribution on \((a, \infty)\) with scale parameter \(\theta,\) where \(\theta > 0\) and
(i) Find the UMVUE of \( a \) when \( \theta \) is known.
(ii) Find the UMVUE of \( \theta \) when \( a \) is known.
(iii) Find the UMVUE’s of \( \theta \) and \( a \).
(iv) Assume that \( \theta \) is known. Find the UMVUE of \( P(X_1 \geq t) \) and the UMVUE of \( \frac{d}{dt}P(X_1 \geq t) \) for a fixed \( t > a \).
(v) Find the UMVUE of \( P(X_1 \geq t) \) for a fixed \( t > a \).

Solution: (i) When \( \theta \) is known, the smallest order statistic \( X(1) \) is complete and sufficient for \( a \). Since \( EX(1) = a + \theta/n \), \( X(1) - \theta/n \) is the UMVUE of \( a \).

(ii) When \( a \) is known, \( T = \sum_{i=1}^{n} X_i \) is complete and sufficient for \( \theta \). Since \( ET = n(a + \theta) \), \( T/n - a \) is the UMVUE of \( \theta \).

(iii) Note that \((X(1), T - nX(1))\) is complete and sufficient for \((a, \theta)\) and \(2(T - nX(1))/\theta \) has the chi-square distribution \( \chi^2_{2(n-1)} \). Then \( E(T - nX(1)) = (n - 1)\theta \) and the UMVUE of \( \theta \) is \((T - nX(1))/(n - 1) \). Since \( EX(1) = a + \theta/n \), the UMVUE of \( a \) is \( X(1) - (T - nX(1))/(n(n - 1)) \).

(iv) Since \( X(1) \) is complete and sufficient for \( a \), the UMVUE of \( P(X_1 \geq t) \) is \( g(X(1)) \) satisfying

\[
P(X_1 \geq t) = \begin{cases} 
  e^{(a-t)/\theta} & t > a \\
  1 & t \leq a
\end{cases}
\]

for any \( a \), which is the same as

\[
\frac{ne^{t/\theta}}{\theta} \int_{a}^{\infty} g(x)e^{-nx/\theta} dx = e^{-(n-1)a/\theta}
\]

for any \( a < t \) and \( g(a) = 1 \) for \( a \geq t \). Differentiating both sides of the above expression with respect to \( a \), we obtain that

\[
ne^{t/\theta}g(a)e^{-na/\theta} = (n-1)e^{-(n-1)a/\theta}.
\]

Hence,

\[
g(x) = \begin{cases} 
(1-n^{-1})e^{(x-t)/\theta} & x < t \\
1 & x \geq t
\end{cases}
\]

and the UMVUE of \( P(X_1 > t) \) is \( g(X(1)) \). The UMVUE of \( \frac{d}{dt}P(X_1 \geq t) \) is then \( -\theta^{-1}e^{(a-t)/\theta} \) and is then \( -\theta^{-1}g(X(1)) \).

(v) The complete and sufficient statistic for \((a, \theta)\) is \( U = (X(1), T - nX(1)) \). The UMVUE is \( P(X_1 \geq t | U) \). Let \( Y = T - nX(1) \) and \( A_j = \{X(1) = X_j\} \).
Then \( P(A_j) = n^{-1} \). If \( t < X_{(1)} \), obviously \( P(X_1 \geq t | U) = 1 \). For \( t \geq X_{(1)} \), consider \( U = u = (x_{(1)}, y) \) and
\[
P(X_1 \geq t | U = u) = P \left( \frac{X_1 - X_{(1)}}{Y} \geq \frac{t - x_{(1)}}{y} \bigg| U = u \right)
\]
\[
= P \left( \frac{X_1 - X_{(1)}}{Y} \geq \frac{t - x_{(1)}}{y} \right)
\]
\[
= \sum_{j=1}^{n} P(A_j) P \left( \frac{X_1 - X_{(1)}}{Y} \geq \frac{t - x_{(1)}}{y} \bigg| A_j \right)
\]
\[
= \frac{n - 1}{n} P \left( \frac{X_1 - X_{(1)}}{Y} \geq \frac{t - x_{(1)}}{y} \bigg| A_n \right)
\]
\[
= \frac{n - 1}{n} P \left( \frac{X_1 - X_{(1)}}{ \sum_{i=1}^{n-1} (X_i - X_{(1)}) } \geq \frac{t - x_{(1)}}{y} \bigg| A_n \right)
\]
\[
= \frac{n - 1}{n} \left( 1 - \frac{t - x_{(1)}}{y} \right)^{n-2},
\]
where the second equality follows from the fact that \( U \) and \( (X_1 - X_{(1)})/Y \) are independent (Basu’s theorem), the fourth equality follows from the fact that the conditional probability given \( A_1 \) is 0 and the conditional probabilities given \( A_j, j = 2, \ldots, n \), are all the same, the fifth equality follows from the fact that \( Y = \sum_{i=1}^{n-1} (X_i - X_{(1)}) \) on the event \( A_n \), and the last equality follows from the fact that conditional on \( A_n \), \( X_i - X_{(1)}, i = 1, \ldots, n-1 \), are independent and identically distributed as the exponential distribution on \((0, \infty)\) with scale parameter \( \theta \) and \( (X_1 - X_{(1)})/\sum_{i=1}^{n-1} (X_i - X_{(1)}) \) has the beta distribution with density \((n - 2)(1 - x)^{n-3}I_{(0,1)}(x)\). Therefore, the UMVUE is equal to 1 when \( t < X_{(1)} \) and
\[
\left( 1 - \frac{1}{n} \right) \left[ 1 - \frac{t - X_{(1)}}{ \sum_{i=1}^{n} (X_i - X_{(1)}) } \right]^{n-2}
\]
when \( X_{(1)} \leq t \). \( \blacksquare \)

**Exercise 7 (#3.7).** Let \((X_1, \ldots, X_n)\) be a random sample from the Pareto distribution with Lebesgue density \( \theta a^\theta x^{-(\theta + 1)}I_{(a,\infty)}(x) \), where \( \theta > 0 \) and \( a > 0 \).

(i) Find the UMVUE of \( \theta \) when \( a \) is known.

(ii) Find the UMVUE of \( a \) when \( \theta \) is known.

(iii) Find the UMVUE’s of \( a \) and \( \theta \).

**Solution:** (i) The joint Lebesgue density of \( X_1, \ldots, X_n \) is
\[
f(x_1, \ldots, x_n) = \theta^n a^\theta \exp \left\{ -(\theta + 1) \sum_{i=1}^{n} \log x_i \right\} I_{(a,\infty)}(x_{(1)}),
\]
where \( x_{(1)} = \min_{1 \leq i \leq n} x_i \). When \( a \) is known, \( T = \sum_{i=1}^{n} \log X_i \) is complete and sufficient for \( \theta \) and \( T - n \log a \) has the gamma distribution with shape parameter \( n \) and scale parameter \( \theta^{-1} \). Hence, \( ET^{-1} = \theta/(n-1) \) and, thus, \((n-1)/T\) is the UMVUE of \( \theta \).

(ii) When \( \theta \) is known, \( X_{(1)} \) is complete and sufficient for \( a \). Since \( X_{(1)} \) has the Lebesgue density \( n \theta a^{n \theta} x^{-(n \theta + 1)} I_{(a, \infty)}(x) \), \( EX_{(1)} = n \theta a/(n \theta - 1) \).

Therefore, \((1 - n \theta)X_{(1)}/(n \theta)\) is the UMVUE of \( a \).

(iii) When both \( a \) and \( \theta \) are unknown, \((Y, X_{(1)})\) is complete and sufficient for \((a, \theta)\), where \( Y = \sum_i (\log X_i - \log X_{(1)}) \). Also, \( Y \) has the gamma distribution with shape parameter \( n - 1 \) and scale parameter \( \theta^{-1} \) and \( X_{(1)} \) and \( Y \) are independent. Since \( EY^{-1} = \theta/(n - 2) \), \((n - 2)/Y\) is the UMVUE of \( \theta \). Since

\[
E \left\{ \left[ 1 - \frac{Y}{n(n-1)} \right] X_{(1)} \right\} = \left[ 1 - \frac{EY}{n(n-1)} \right] EX_{(1)} = \left( 1 - \frac{1}{n \theta} \right) \frac{n \theta a}{n \theta - 1} = a,
\]

\( \left[ 1 - \frac{Y}{n(n-1)} \right] X_{(1)} \) is the UMVUE of \( a \).

Exercise 8 (#3.11). Let \( X \) be a random variable having the negative binomial distribution with

\[
P(X = x) = \binom{x - 1}{r - 1} p^r (1 - p)^{x-r}, \quad x = r, r + 1, \ldots,
\]

where \( p \in (0, 1) \) and \( r \) is a known positive integer.

(i) Find the UMVUE of \( p^t \), where \( t \) is a positive integer and \( t < r \).

(ii) Find the UMVUE of \( \text{Var}(X) \).

(iii) Find the UMVUE of \( \log p \).

Solution. (i) Since \( X \) is complete and sufficient for \( p \), the UMVUE of \( p^t \) is \( h(X) \) with a function \( h \) satisfying \( E[h(X)] = p^t \) for any \( p \), i.e.,

\[
\sum_{x=r}^{\infty} h(x) \binom{x - 1}{r - 1} p^r (1 - p)^{x-r} = p^t
\]

for any \( p \). Let \( q = 1 - p \). Then

\[
\sum_{x=r}^{\infty} h(x) \binom{x - 1}{r - 1} q^x = \frac{q^r}{(1 - q)^{r-t}}
\]

for any \( q \in (0, 1) \). From the negative binomial identity

\[
\sum_{x=j}^{\infty} \binom{x - 1}{j - 1} q^x = \frac{q^j}{(1 - q)^j}
\]
with any positive integer \( j \), we obtain that
\[
\sum_{x=r}^{\infty} h(x) \binom{x-1}{r-1} q^x = \sum_{x=r-t}^{\infty} \binom{x-1}{r-t-1} q^{x+t} = \sum_{x=r}^{\infty} \binom{x-t-1}{r-t-1} q^x
\]
for any \( q \). Comparing the coefficients of \( q^x \), we obtain that
\[
h(x) = \frac{\binom{x-t-1}{r-t-1}}{\binom{x-1}{r-1}}, \quad x = r, r+1, \ldots.
\]

(ii) Note that \( \text{Var}(X) = r(1-p)/p^2 = rq/(1-q)^2 \). The UMVUE of \( \text{Var}(X) \) is \( h(X) \) with \( E[h(X)] = rq/(1-q)^2 \) for any \( q \in (0, 1) \). That is,
\[
\sum_{x=r}^{\infty} h(x) \binom{x-1}{r-1} q^x = \frac{q^r}{(1-q)^r} \text{Var}(X) = r \frac{q^{r+1}}{(1-q)^{r+2}}
\]
for any \( q \). Using the negative binomial identity, this means that
\[
\sum_{x=r}^{\infty} h(x) \binom{x-1}{r-1} q^x = r \sum_{x=r+2}^{\infty} \binom{x-1}{r+1} q^{x-1} = r \sum_{x=r+1}^{\infty} \binom{x}{r+1} q^x
\]
for any \( q \), which yields
\[
h(x) = \begin{cases} 
0 & x = r \\
\frac{r(x+1)}{(r-1)} & x = r+1, r+2, \ldots 
\end{cases}
\]

(iii) Let \( h(X) \) be the UMVUE of \( \log p = \log(1-q) \). Then, for any \( q \in (0, 1) \),
\[
\sum_{x=r}^{\infty} h(x) \binom{x-1}{r-1} q^x = \frac{q^r}{(1-q)^r} \log(1-q)
\]
\[
= -\sum_{x=r}^{\infty} \binom{x-1}{r-1} q^x \sum_{i=1}^{\infty} \frac{q^i}{i}
\]
\[
= \sum_{x=r+1}^{\infty} \sum_{k=0}^{x-r-1} \binom{r+k-1}{k} \frac{q^x}{k+r-x}.
\]

Hence \( h(r) = 0 \) and
\[
h(x) = \frac{1}{\binom{x-1}{r-1}} \sum_{k=0}^{x-r-1} \binom{r+k-1}{k} \frac{1}{k+r-x}
\]
for \( x = r+1, r+2, \ldots. \) \( \blacksquare \)
Exercise 9 (#3.12). Let \((X_1, ..., X_n)\) be a random sample from the Poisson distribution truncated at 0, i.e., \(P(X_i = x) = (e^\theta - 1)^{-1}e^{\theta x}/x!, \ x = 1, 2, ..., \theta > 0\). Find the UMVUE of \(\theta\) when \(n = 1, 2\).

Solution. Assume \(n = 1\). Then \(X\) is complete and sufficient for \(\theta\) and the UMVUE of \(\theta\) is \(h(X)\) with \(E[h(X)] = \theta\) for any \(\theta\).

On the other hand,

\[
\theta(e^\theta - 1)^2 = \left(\sum_{i=1}^\infty \frac{\theta^i}{i!}\right)^2 = \sum_{i=1}^\infty \sum_{j=1}^\infty \frac{\theta^i \theta^{i+1}}{i! j!} = \sum_{t=3}^\infty \frac{1}{t! (t - 1)!}.
\]

Comparing the coefficient of \(\theta^x\) leads to \(h(2) = 0\) and

\[
h(t) = \sum_{i=0}^{t-2} \frac{1}{i! (t - 1 - i)!} / \sum_{i=0}^{t-1} \frac{1}{i! (t - i)!}
\]

for \(t = 3, 4, ...\).

Exercise 10 (#3.14). Let \(X_1, ..., X_n\) be a random sample from the log-distribution with

\[
P(X_1 = x) = -(1 - p)^x/(x \log p), \ x = 1, 2, ...
\]

\(p \in (0, 1)\). Let \(k\) be a fixed positive integer.

(i) For \(n = 1, 2, 3\), find the UMVUE of \(p^k\).

(ii) For \(n = 1, 2, 3\), find the UMVUE of \(P(X = k)\).

Solution. (i) Let \(\theta = 1 - p\). Then \(p^k = \sum_{r=0}^k \binom{k}{r} (-1)^r \theta^r\). Hence, it suffices to obtain the UMVUE for \(\theta^r\). Note that the distribution of \(X_1\) is from a
Chapter 3. Unbiased Estimation

Power series distribution with \(\gamma(x) = x^{-1}\) and \(e(\theta) = -\log(1 - \theta)\) (see Example 3.5 in Shao, 2003). The statistic \(T = \sum_{i=1}^{n} X_i\) is complete and sufficient for \(\theta\). By the result in Example 3.5 of Shao (2003), the UMVUE of \(\theta^r\) is

\[
\frac{\gamma_n(T - r)}{\gamma_n(T)} I_{\{r, r+1, \ldots\}}(T),
\]

where \(\gamma_n(t)\) is the coefficient of \(\theta^t\) in \((\sum_{y=1}^{\infty} \frac{\theta^y}{y})^n\), i.e., \(\gamma_n(t) = 0\) for \(t < n\) and

\[
\gamma_n(t) = \sum_{y_1 + \ldots + y_n = t - n, y_i \geq 0} \frac{1}{(y_1 + 1) \cdots (y_n + 1)}
\]

for \(t = n, n + 1, \ldots\). When \(n = 1, 2, 3\), \(\gamma_n(t)\) has a simpler form. In fact,

\[
\gamma_1(1) = 0 \quad \text{and} \quad \gamma_1(t) = t^{-1}, \quad t = 2, 3, \ldots;
\]

\[
\gamma_2(1) = \gamma_2(2) = 0 \quad \text{and} \quad \gamma_2(t) = \sum_{l=0}^{t-2} \frac{1}{(l+1)(t-l-1)}, \quad t = 3, 4, \ldots;
\]

\[
\gamma_3(1) = \gamma_3(2) = \gamma_3(3) = 0 \quad \text{and} \quad \gamma_3(t) = \sum_{l_1=0}^{t-3} \sum_{l_2=0}^{t-3} \frac{1}{(l_1 + 1)(l_2 + 1)(t - l_1 - l_2 - 2)}, \quad t = 4, 5, \ldots.
\]

(ii) By Example 3.5 in Shao (2003), the UMVUE of \(P(X_1 = k)\) is

\[
\frac{\gamma_{n-1}(T - k)}{k\gamma_n(T)} I_{\{k, k+1, \ldots\}}(T),
\]

where \(\gamma_n(t)\) is given in the solution of part (i).

Exercise 11 (#3.19). Let \(Y_1, \ldots, Y_n\) be a random sample from the uniform distribution on the interval \((0, \theta)\) with an unknown \(\theta \in (1, \infty)\).

(i) Suppose that we only observe

\[
X_i = \begin{cases} 
Y_i & \text{if } Y_i \geq 1 \\
1 & \text{if } Y_i < 1
\end{cases}, \quad i = 1, \ldots, n.
\]

Derive a UMVUE of \(\theta\).

(ii) Suppose that we only observe

\[
X_i = \begin{cases} 
Y_i & \text{if } Y_i \leq 1 \\
1 & \text{if } Y_i > 1
\end{cases}, \quad i = 1, \ldots, n.
\]
Derive a UMVUE of the probability $P(Y_1 > 1)$.

**Solution.** (i) Let $m$ be the Lebesgue measure and $\delta$ be the point mass on $\{1\}$. The joint probability density of $X_1, \ldots, X_n$ with respect to $\delta + m$ is (see, e.g., Exercise 16 in Chapter 1) $\theta^{-n}I_{[0, \theta]}(X_{(n)})$, where $X_{(n)} = \max_{1 \leq i \leq n} X_i$. Hence $X_{(n)}$ is complete and sufficient for $\theta$ and the UMVUE of $\theta$ is $h(X_{(n)})$ satisfying $E[h(X_{(n)})] = \theta$ for all $\theta > 1$. The probability density of $X_{(n)}$ with respect to $\delta + m$ is $\theta^{-n}I_{[1]}(x) + n\theta^{-n}x^{n-1}I_{(1, \theta)}(x)$. Hence

$$E[h(X_{(n)})] = \frac{h(1)}{\theta^n} + \frac{n}{\theta^n} \int_1^{\theta} h(x)x^{n-1}dx.$$  

Then

$$\theta^{n+1} = h(1) + n \int_1^{\theta} h(x)x^{n-1}dx$$

for all $\theta > 1$. Letting $\theta \to 1$ we obtain that $h(1) = 1$. Differentiating both sides of the previous expression with respect to $\theta$ we obtain that

$$(n + 1)\theta^n = nh(\theta)\theta^{n-1} \quad \theta > 1.$$  

Hence $h(x) = (n + 1)x/n$ when $x > 1$.

(ii) The joint probability density of $X_1, \ldots, X_n$ with respect to $\delta + m$ is $\theta^{-r}(1 - \theta)^{n-r}$, where $r$ is the observed value of $R =$ the number of $X_i$'s that are less than 1. Hence, $R$ is complete and sufficient for $\theta$. Note that $R$ has the binomial distribution with size $n$ and probability $\theta^{-1}$ and $P(Y_1 > 1) = 1 - \theta^{-1}$. Hence, the UMVUE of $P(Y_1 > 1)$ is $1 - R/n$.

**Exercise 12 (#3.22).** Let $(X_1, \ldots, X_n)$ be a random sample from $P \in \mathcal{P}$ containing all symmetric distributions with finite means and with Lebesgue densities on $\mathcal{R}$.

(i) When $n = 1$, show that $X_1$ is the UMVUE of $\mu$.

(ii) When $n > 1$, show that there is no UMVUE of $\mu = EX_1$.

**Solution.** (i) Consider the sub-family $\mathcal{P}_1 = \{N(\mu, 1) : \mu \in \mathcal{R}\}$. Then $X_1$ is complete for $P \in \mathcal{P}_1$. Hence, $E[h(X_1)] = 0$ for any $P \in \mathcal{P}$ implies that $E[h(X_1)] = 0$ for any $P \in \mathcal{P}_1$ and, thus, $h = 0$ a.e. Lebesgue measure. This shows that 0 is the unique estimator of 0 when the family $\mathcal{P}$ is considered. Since $EX_1 = \mu$, $X_1$ is the unique unbiased estimator of $\mu$ and, hence, it is the UMVUE of $\mu$.

(ii) Suppose that $T$ is a UMVUE of $\mu$. Let $\mathcal{P}_1 = \{N(\mu, 1) : \mu \in \mathcal{R}\}$. Since the sample mean $\bar{X}$ is UMVUE when $\mathcal{P}_1$ is considered, by using the same argument in the solution for Exercise 4(iv), we can show that $T = \bar{X}$ a.s. $P$ for any $P \in \mathcal{P}_1$. Since the Lebesgue measure is dominated by any $P \in \mathcal{P}_1$, we conclude that $T = \bar{X}$ a.e. Lebesgue measure. Let $\mathcal{P}_2$ be the family given in Exercise 5. Then $(X_{(1)} + X_{(n)})/2$ is the UMVUE when $\mathcal{P}_2$ is considered, where $X_{(j)}$ is the $j$th order statistic. Then $\bar{X} = (X_{(1)} + X_{(n)})/2$ a.s. $P$ for any $P \in \mathcal{P}_2$, which is impossible. Hence, there is no UMVUE of $\mu$. □
Exercise 13 (#3.24). Suppose that $T$ is a UMVUE of an unknown parameter $\theta$. Show that $T^k$ is a UMVUE of $E(T^k)$, where $k$ is any positive integer for which $E(T^{2k}) < \infty$.

Solution. Let $U$ be an unbiased estimator of 0. Since $T$ is a UMVUE of $\theta$, $E(TU) = 0$ for any $P$, which means that $TU$ is an unbiased estimator of 0. Then $E(T^2U) = E[T(TU)] = 0$ if $ET^4 < \infty$. By Theorem 3.2 in Shao (2003), $T^2$ is a UMVUE of $ET^2$. Similarly, we can show that $T^3$ is a UMVUE of $ET^3$. .

Exercise 14 (#3.27). Let $X$ be a random variable having the Lebesgue density $[(1 - \theta) + \theta/(2\sqrt{x})]I_{(0,1)}(x)$, where $\theta \in [0, 1]$. Show that there is no UMVUE of $\theta$ based on an observation $X$.

Solution. Consider estimators of the form $h(X) = a(X^{-1/2} + b)I_{(c, 1)}(X)$ for some real numbers $a$ and $b$, and $c \in (0, 1)$. Note that

$$\int_0^1 h(x)dx = a \int_c^1 x^{-1/2}dx + ab \int_c^1 dx = 2a(1 - \sqrt{c}) + ab(1 - c).$$

If $b = -2/(1 + \sqrt{c})$, then $\int_0^1 h(x)dx = 0$ for any $a$ and $c$. Also,

$$\int_0^1 \frac{h(x)}{2\sqrt{x}}dx = \frac{a}{2} \int_c^1 x^{-1}dx + \frac{ab}{2} \int_c^1 x^{-1/2}dx = -\frac{a}{2} \log c + ab(1 - \sqrt{c}).$$

If $a = [b(1 - \sqrt{c}) - 2^{-1}\log c]^{-1}$, then $\int_0^1 \frac{h(x)}{2\sqrt{x}}dx = 1$ for any $b$ and $c$. Let $g_c = h$ with $b = -2/(1 + \sqrt{c})$ and $a = [b(1 - \sqrt{c}) - 2^{-1}\log c]^{-1}$, $c \in (0, 1)$. Then

$$E[g_c(X)] = (1 - \theta) \int_0^1 g_c(x)dx + \theta \int_0^1 \frac{g_c(x)}{2\sqrt{x}}dx = \theta$$

for any $\theta$, i.e., $g_c(X)$ is unbiased for $\theta$ for any $c \in (0, 1)$. The variance of $g_c(X)$ when $\theta = 0$ is

$$E[g_c(X)]^2 = a^2 \int_c^1 (x^{-1} + b^2 + 2bx^{-1/2})dx$$

$$= a^2 \left[- \log c + b^2(1 - c) + 4b(1 - \sqrt{c})\right]$$

$$= \frac{- \log c + b^2(1 - c) + 4b(1 - \sqrt{c})}{[b(1 - \sqrt{c}) - 2^{-1}\log c]^2},$$

where $b = -2/(1 + \sqrt{c})$. Letting $c \to 0$, we obtain that $b \to -2$ and, thus, $E[g_c(X)]^2 \to 0$. This means that no minimum variance estimator within the class of estimators $g_c(X)$. Hence, there is no UMVUE of $\theta$. ■

Exercise 15 (#3.28). Let $X$ be a random sample with $P(X = -1) = 2p(1 - p)$ and $P(X = k) = p^k(1 - p)^{3-k}$, $k = 0, 1, 2, 3$, where $p \in (0, 1)$.

(i) Determine whether there is a UMVUE of $p$. 


(ii) Determine whether there is a UMVUE of \( p(1 - p) \).

**Solution.** (i) Suppose that \( f(X) \) is an unbiased estimator of \( p \). Then

\[
p = 2f(-1)p(1-p) + f(0)(1-p)^3 + f(1)p(1-p)^2 + f(2)p^2(1-p) + f(3)p^3
\]

for any \( p \). Letting \( p \to 0 \), we obtain that \( f(0) = 0 \). Letting \( p \to 1 \), we obtain that \( f(3) = 1 \). Then

\[
1 = 2f(-1)(1-p) + f(1)(1-p)^2 + f(2)p(1-p) + p^2
\]

Thus, \( 2f(-1) + f(1) = 0, f(2) - 2f(-1) - 2f(1) = 0 \), and \( f(1) - f(2) + 1 = 0 \).

These three equations are not independent; in fact the second equation is a consequence of the first and the last equations. Let \( f(2) = c \). Then \( f(1) = c - 1 \) and \( f(-1) = 1 - c/2 \). Let \( g_c(2) = c, g_c(1) = c - 1, g_c(-1) = 1 - c/2, g_c(0) = 0 \), and \( g_c(3) = 1 \). Then the class of unbiased estimators of \( p \) is \( \{g_c(X) : c \in \mathbb{R} \} \). The variance of \( g_c(X) \) is

\[
E[g_c(X)]^2 - p^2 = 2(1-c/2)^2p(1-p) + (c-1)^2p(1-p)^2 + c^2p^2(1-p) + p^3 - p^2.
\]

Denote the right hand side of the above equation by \( h(c) \). Then

\[
h'(c) = -(2-c)p(1-p) + 2(c-1)p(1-p)^2 + 2cp^2(1-p).
\]

Setting \( h'(c) = 0 \) we obtain that

\[
0 = c - 2 + 2(c-1)(1-p) + 2cp = c - 2 + 2c - 2(1-p).
\]

Hence, the function \( h(c) \) reaches its minimum at \( c = (4 - 2p)/3 \), which depends on \( p \). Therefore, there is no UMVUE of \( p \).

(ii) Suppose that \( f(X) \) is an unbiased estimator of \( p(1-p) \). Then

\[
p(1-p) = 2f(-1)p(1-p) + f(0)(1-p)^3 + f(1)p(1-p)^2
\]

\[
+ f(2)p^2(1-p) + f(3)p^3
\]

for any \( p \). Letting \( p \to 0 \) we obtain that \( f(0) = 0 \). Letting \( p \to 1 \) we obtain that \( f(3) = 0 \). Then

\[
1 = 2f(-1) + f(1)(1-p) + f(2)p
\]

for any \( p \), which implies that \( f(2) = f(1) \) and \( 2f(-1) + f(1) = 1 \). Let \( f(-1) = c \). Then \( f(1) = f(2) = 1 - 2c \). Let \( g_c(-1) = c, g_c(0) = g_c(3) = 0 \), and \( g_c(1) = g_c(2) = 1 - 2c \). Then the class of unbiased estimators of \( p(1-p) \) is \( \{g_c(X) : c \in \mathbb{R} \} \). The variance of \( g_c(X) \) is

\[
E[g_c(X)]^2 - p^2 = 2c^2p(1-p) + (1-2c)^2p(1-p)^2
\]

\[
+ (1-2c)^2p^2(1-p) - p^2
\]

\[
= 2c^2p(1-p) + (1-2c)^2p(1-p) - p^2
\]

\[
= [2c^2 + (1-2c)^2]p(1-p) - p^2,
\]
which reaches its minimum at $c = 1/3$ for any $p$. Thus, the UMVUE of $p(1 - p)$ is $g_{1/3}(X)$.

**Exercise 16 (#3.29(a)).** Let $(X_1, ..., X_n)$ be a random sample from the exponential distribution with density $\theta^{-1}e^{-(x-a)/\theta}I_{(a,\infty)}(x)$, where $a \leq 0$ and $\theta$ is known. Obtain a UMVUE of $p(1 - p)$.

**Note.** The minimum order statistic, $X_{(1)}$, is sufficient for $a$ but not complete because $a \leq 0$.

**Solution.** Let $U(X_{(1)})$ be an unbiased estimator of 0. Then $E[U(X_{(1)})] = 0$ implies

$$\int_a^0 U(x)e^{-x/\theta} dx + \int_0^\infty U(x)e^{-x/\theta} dx = 0$$

for all $a \leq 0$. Hence, $U(x) = 0$ a.e. for $x \leq 0$ and $\int_0^\infty U(x)e^{-x/\theta} dx = 0$. Consider

$$h(X_{(1)}) = (bX_{(1)} + c)I_{(-\infty,0)}(X_{(1)})$$

with constants $b$ and $c$. Then $E[h(X_{(1)})U(X_{(1)})] = 0$ for any $a$. By Theorem 3.2 in Shao (2003), $h(X_{(1)})$ is a UMVUE of its expectation

$$E[h(X_{(1)})] = \frac{ne^{na/\theta}}{\theta} \int_a^0 (bx + c)e^{-nx/\theta} dx$$

$$= c \left(1 - e^{na/\theta}\right) + ab + \frac{b\theta}{n} \left(1 - e^{na/\theta}\right),$$

which equals $a$ when $b = 1$ and $c = -\theta/n$. Therefore, the UMVUE of $a$ is

$$h(X_{(1)}) = (X_{(1)} - \theta/n)I_{(-\infty,0]}(X_{(1)})$$

**Exercise 17 (#3.29(b)).** Let $(X_1, ..., X_n)$ be a random sample from the distribution on $\mathcal{R}$ with Lebesgue density $\theta a^\theta x^{-(\theta+1)}I_{(a,\infty)}(x)$, where $a \in (0, 1]$ and $\theta$ is known. Obtain a UMVUE of $a$.

**Solution.** The minimum order statistic $X_{(1)}$ is sufficient for $a$ and has Lebesgue density $n\theta a^{n\theta}x^{-(n\theta+1)}I_{(a,\infty)}(x)$. Let $U(X_{(1)})$ be an unbiased estimator of 0. Then $E[U(X_{(1)})] = 0$ implies

$$\int_a^1 U(x)x^{-(n\theta+1)} dx + \int_1^\infty U(x)x^{-(n\theta+1)} dx = 0$$

for all $a \in (0, 1]$. Hence, $U(x) = 0$ a.e. for $x \in (0, 1]$ and $\int_1^\infty U(x)x^{-(n\theta+1)} dx = 0$. Let

$$h(X_{(1)}) = cI_{(1,\infty)}(X_{(1)}) + bX_{(1)}I_{(0,1]}(X_{(1)})$$

with some constants $b$ and $c$. Then

$$E[h(X_{(1)})U(X_{(1)})] = c \int_1^\infty U(x)x^{-(n\theta+1)} dx = 0.$$
By Theorem 3.2 in Shao (2003), \( h(X_{(1)}) \) is a UMVUE of its expectation

\[
E[h(X_{(1)})] = bn\theta a^n \theta \int_a^1 x^{-n\theta} dx + cn\theta a^n \theta \int_1^{\infty} x^{-(n\theta+1)} dx
\]

which equals \( a \) when \( b = 1 - \frac{1}{n\theta} \) and \( c = 1 \). Hence, the UMVUE of \( a \) is

\[ h(X_{(1)}) = I_{(1,\infty)}(X_{(1)}) + \left( 1 - \frac{1}{n\theta} \right) X_{(1)} I_{(0,1]}(X_{(1)}) . \]

**Exercise 18 (#3.30).** Let \((X_1, ..., X_n)\) be a random sample from the population in a family \( \mathcal{P} \) as described in Exercise 18 of Chapter 2. Find a UMVUE of \( \theta \).

**Solution.** Note that \( \mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2 \), where \( \mathcal{P}_1 \) is the family of Poisson distributions with the mean parameter \( \theta \in (0, 1) \) and \( \mathcal{P}_2 \) is the family of binomial distributions with size 1 and probability \( \theta \). The sample mean \( \bar{X} \) is the UMVUE of \( \theta \) when either \( \mathcal{P}_1 \) or \( \mathcal{P}_2 \) is considered as the family of distributions. Hence \( \bar{X} \) is the UMVUE of \( \theta \) when \( \mathcal{P} \) is considered as the family of distributions.

**Exercise 19 (#3.33).** Find a function of \( \theta \) for which the amount of information is independent of \( \theta \), when \( \mathcal{P}_\theta \) is

(i) the Poisson distribution with unknown mean \( \theta > 0 \);

(ii) the binomial distribution with known size \( r \) and unknown probability \( \theta \in (0, 1) \);

(iii) the gamma distribution with known shape parameter \( \alpha \) and unknown scale parameter \( \theta > 0 \).

**Solution.** (i) The Fisher information about \( \theta \) is \( I(\theta) = \frac{1}{\theta} \). Let \( \eta = \eta(\theta) \). If the Fisher information about \( \eta \) is

\[
\tilde{I}(\eta) = \left( \frac{d\theta}{d\eta} \right)^2 I(\theta) = \left( \frac{d\theta}{d\eta} \right)^2 \frac{1}{\theta} = c
\]

not depending on \( \theta \), then \( \frac{d\eta}{d\theta} = 1/\sqrt{c\theta} \). Hence, \( \eta(\theta) = 2\sqrt{\theta}/\sqrt{c} \).

(ii) The Fisher information about \( \theta \) is \( I(\theta) = \frac{r}{\theta(1-\theta)} \). Let \( \eta = \eta(\theta) \). If the Fisher information about \( \eta \) is

\[
\tilde{I}(\eta) = \left( \frac{d\theta}{d\eta} \right)^2 I(\theta) = \left( \frac{d\theta}{d\eta} \right)^2 \frac{r}{\theta(1-\theta)} = c
\]

not depending on \( \theta \), then \( \frac{d\eta}{d\theta} = \sqrt{r}/\sqrt{c\theta(1-\theta)} \). Choose \( c = 4r \). Then \( \eta(\theta) = \arcsin(\sqrt{\theta}) \).
(iii) The Fisher information about $\theta$ is $I(\theta) = \frac{\alpha}{\theta^2}$. Let $\eta = \eta(\theta)$. If the Fisher information about $\eta$ is 

$$\tilde{I}(\eta) = \left( \frac{d\theta}{d\eta} \right)^2 I(\theta) = \left( \frac{d\theta}{d\eta} \right)^2 \frac{\alpha}{\theta^2} = \alpha,$$

then $\frac{d\eta}{d\theta} = \theta^{-1}$ and, hence, $\eta(\theta) = \log \theta$. 

**Exercise 20 (#3.34).** Let $(X_1, ..., X_n)$ be a random sample from a distribution on $\mathcal{R}$ with the Lebesgue density $\frac{1}{\sigma} f \left( \frac{x-\mu}{\sigma} \right)$, where $f(x) > 0$ is a known Lebesgue density and $f'(x)$ exists for all $x \in \mathcal{R}$, $\mu \in \mathcal{R}$, and $\sigma > 0$. Let $\theta = (\mu, \sigma)$. Show that the Fisher information about $\theta$ contained in $X_1, ..., X_n$ is

$$I(\theta) = \frac{n}{\sigma^2} \begin{pmatrix} \int \frac{|f'(x)|^2}{f(x)} dx & \int \frac{f'(x)[xf'(x)+f(x)]}{f(x)} dx \\ \int \frac{f'(x)[xf'(x)+f(x)]}{f(x)} dx & \int \frac{|xf'(x)+f(x)|^2}{f(x)} dx \end{pmatrix},$$

assuming that all integrals are finite.

**Solution.** Let $g(\mu, \sigma, x) = \log \frac{1}{\sigma} f \left( \frac{x-\mu}{\sigma} \right)$. Then

$$\frac{\partial}{\partial \mu} g(\mu, \sigma, x) = -\frac{f'(x)}{\sigma f \left( \frac{x-\mu}{\sigma} \right)}$$

and

$$\frac{\partial}{\partial \sigma} g(\mu, \sigma, x) = -\frac{(x-\mu)f'(x)}{\sigma f \left( \frac{x-\mu}{\sigma} \right)} - \frac{1}{\sigma}.$$

Then

$$E \left[ \frac{\partial}{\partial \mu} g(\mu, \sigma, X_1) \right]^2 = \frac{1}{\sigma^2} \int \left[ \frac{f'(x)}{f \left( \frac{x-\mu}{\sigma} \right)} \right]^2 \frac{1}{\sigma} f \left( \frac{x-\mu}{\sigma} \right) dx$$

$$= \frac{1}{\sigma^2} \int \frac{|f'(x)|^2}{f \left( \frac{x-\mu}{\sigma} \right)} d \left( \frac{x}{\sigma} \right)$$

$$= \frac{1}{\sigma^2} \int \frac{|f'(x)|^2}{f(x)} dx,$$

$$E \left[ \frac{\partial}{\partial \sigma} g(\mu, \sigma, X_1) \right]^2 = \frac{1}{\sigma^2} \int \left[ \frac{x-\mu}{\sigma} f' \left( \frac{x-\mu}{\sigma} \right) + 1 \right] \frac{1}{\sigma} f \left( \frac{x-\mu}{\sigma} \right) dx$$

$$= \frac{1}{\sigma^2} \int \left[ \frac{xf'(x)}{f(x)} + 1 \right]^2 f(x) dx$$

$$= \frac{1}{\sigma^2} \int \frac{|xf'(x)+f(x)|^2}{f(x)} dx.$$
and
\[
E \left[ \frac{\partial}{\partial \mu} g(\mu, \sigma, X_1) \frac{\partial}{\partial \sigma} g(\mu, \sigma, X_1) \right] = \frac{1}{\sigma^2} \int \frac{f'(x - \mu)}{f(x - \mu)} \left[ -\frac{x - \mu}{\sigma} \frac{f'(x - \mu)}{f(x - \mu)} + 1 \right] \frac{1}{\sigma} f \left( \frac{x - \mu}{\sigma} \right) dx
\]
\[
= \int \frac{f'(x) [xf'(x) + f(x)]}{f(x)} dx.
\]

The result follows since
\[
I(\theta) = nE \left[ \frac{\partial}{\partial \theta} \log \frac{1}{\sigma} f \left( \frac{X_1 - \mu}{\sigma} \right) \right] \left[ \frac{\partial}{\partial \theta} \log \frac{1}{\sigma} f \left( \frac{X_1 - \mu}{\sigma} \right) \right]^T.
\]

**Exercise 21 (#3.36).** Let \( X \) be a sample having a probability density \( f_{\theta}(x) \) with respect to \( \nu \), where \( \theta \) is a \( k \)-vector of unknown parameters. Let \( T(X) \) be a statistic having a probability density \( g_{\theta}(t) \) with respect to \( \lambda \). Suppose that \( \frac{\partial}{\partial \theta} f_{\theta}(x) \) and \( \frac{\partial}{\partial \theta} g_{\theta}(t) \) exist for any \( x \) and \( t \) and that, on any set \( \{ || \theta || \leq c \} \), there are functions \( u_c(x) \) and \( v_c(t) \) such that \( | \frac{\partial}{\partial \theta} f_{\theta}(x) | \leq u_c(x), | \frac{\partial}{\partial \theta} g_{\theta}(t) | \leq v_c(t), \int u_c(x) d\nu < \infty, \) and \( \int v_c(t) d\lambda < \infty. \) Show that
(i) \( I_X(\theta) - I_T(\theta) \) is nonnegative definite, where \( I_X(\theta) \) is the Fisher information about \( \theta \) contained in \( X \) and \( I_T(\theta) \) is the Fisher information about \( \theta \) contained in \( T \);
(ii) \( I_X(\theta) = I_T(\theta) \) if \( T \) is sufficient for \( \theta \).

**Solution.** (i) For any event \( T^{-1}(B) \),
\[
\int_{T^{-1}(B)} \frac{\partial}{\partial \theta} \log f_{\theta}(X) dP = \int_{T^{-1}(B)} \frac{\partial}{\partial \theta} f_{\theta}(x) d\nu
\]
\[
= \frac{\partial}{\partial \theta} \int_{T^{-1}(B)} f_{\theta}(x) d\nu
\]
\[
= \frac{\partial}{\partial \theta} P(T^{-1}(B))
\]
\[
= \frac{\partial}{\partial \theta} \int_B g_{\theta}(t) d\lambda
\]
\[
= \int_B \frac{\partial}{\partial \theta} g_{\theta}(t) d\lambda
\]
\[
= \int_B \left[ \frac{\partial}{\partial \theta} \log g_{\theta}(t) \right] g_{\theta}(t) d\lambda
\]
\[
= \int_{T^{-1}(B)} \frac{\partial}{\partial \theta} \log g_{\theta}(T) dP,
\]

where the exchange of differentiation and integration is justified by the dominated convergence theorem under the given conditions. This shows
that
\[ E \left[ \frac{\partial}{\partial \theta} \log f_\theta(X) \right] = \frac{\partial}{\partial \theta} \log g_\theta(T) \quad \text{a.s.} \]

Then
\[
\begin{align*}
E & \left[ \frac{\partial}{\partial \theta} \log f_\theta(X) \right] \left[ \frac{\partial}{\partial \theta} \log g_\theta(T) \right] \\
& = \mathbb{E} \left\{ \mathbb{E} \left[ \frac{\partial}{\partial \theta} \log f_\theta(X) \right] \left[ \frac{\partial}{\partial \theta} \log g_\theta(T) \right] \right\} \\
& = \mathbb{E} \left[ \frac{\partial}{\partial \theta} \log g_\theta(T) \right] \left[ \frac{\partial}{\partial \theta} \log g_\theta(T) \right]^T \\
& = I_T(\theta).
\end{align*}
\]

Then the nonnegative definite matrix
\[
E \left[ \frac{\partial}{\partial \theta} \log f_\theta(X) - \frac{\partial}{\partial \theta} \log g_\theta(T) \right] \left[ \frac{\partial}{\partial \theta} \log f_\theta(X) - \frac{\partial}{\partial \theta} \log g_\theta(T) \right]^T
\]
is equal to \( I_X(\theta) + I_T(\theta) - 2I_T(\theta) = I_X(\theta) - I_T(\theta) \). Hence \( I_X(\theta) - I_T(\theta) \) is nonnegative definite.

(i) If \( T \) is sufficient, then by the factorization theorem, \( f_\theta(x) = \tilde{g}_\theta(t) h(x) \).

Since \( \frac{\partial}{\partial \theta} \log f_\theta(x) = \frac{\partial}{\partial \theta} \log \tilde{g}_\theta(t) \), the result in part (i) of the solution implies that
\[
\frac{\partial}{\partial \theta} \log \tilde{g}_\theta(T) = \frac{\partial}{\partial \theta} \log g_\theta(T) \quad \text{a.s.}
\]

Therefore, \( I_X(\theta) = I_T(\theta) \).

**Exercise 22 (\#3.37).** Let \((X_1, ..., X_n)\) be a random sample from the uniform distribution on the interval \((0, \theta)\) with \( \theta > 0 \).

(i) Show that \( \frac{d}{d \theta} \int x f_\theta(x) dx \neq \int x \frac{d}{d \theta} f_\theta(x) dx \), where \( f_\theta \) is the density of \( X_{(n)} \), the largest order statistic.

(ii) Show that the Fisher information inequality does not hold for the UMVUE of \( \theta \).

**Solution.** (i) Note that \( f_\theta(x) = n \theta^{-n} x^{n-1} I_{(0, \theta)}(x) \). Then
\[
\begin{aligned}
\int x \frac{d}{d \theta} f_\theta(x) dx &= \frac{n}{\theta^n+1} \int_0^\theta x^n dx \\
&= \frac{n^2}{n+1}.
\end{aligned}
\]

On the other hand,
\[
\begin{aligned}
\frac{d}{d \theta} \int x f_\theta(x) dx &= \frac{d}{d \theta} \left( \frac{n}{\theta^n} \int_0^\theta x^n dx \right) \\
&= \frac{d}{d \theta} \left( \frac{n \theta}{n+1} \right) \\
&= \frac{n}{n+1}.
\end{aligned}
\]

(ii) The UMVUE of \( \theta \) is \((n + 1)\overline{X}_{(n)}/n \) with variance \( \theta^2/[n(n+2)] \). On the other hand, the Fisher information is \( I(\theta) = n \theta^{-2} \). Hence \([I(\theta)]^{-1} = \theta^2/n > \theta^2/[n(n+2)]\).
Exercise 23 (#3.39). Let $X$ be an observation with Lebesgue density $(2\theta)^{-1}e^{-|x|/\theta}$ with unknown $\theta > 0$. Find the UMVE’s of the parameters $\theta, \theta^r$ ($r > 1$), and $(1 + \theta)^{-1}$ and, in each case, determine whether the variance of the UMVE attains the Cramér-Rao lower bound.

**Solution.** For $\theta$, Cramér-Rao lower bound is $\theta^2$ and $|X|$ is the UMVE of $\theta$ with $\text{Var}(|X|) = \theta^2$, which attains the lower bound.

For $\theta^r$, Cramér-Rao lower bound is $r^2\theta^{2r}$. Since $E[|X|^r/\Gamma(r+1)] = \theta^r$, $|X|^r/\Gamma(r+1)$ is the UMVE of $\theta^r$ with

$$\text{Var}\left(\frac{|X|^r}{\Gamma(r+1)}\right) = \theta^{2r} \left[\frac{\Gamma(2r+1)}{\Gamma(r+1)\Gamma(r+1)} - 1\right] > r^2\theta^{2r}$$

when $r > 1$.

For $(1+\theta)^{-1}$, Cramér-Rao lower bound is $\theta^2/(1+\theta)^4$. Since $E(e^{-|X|}) = (1+\theta)^{-1}$, $e^{-|X|}$ is the UMVE of $(1 + \theta)^{-1}$ with

$$\text{Var}(e^{-|X|}) = \frac{1}{1+2\theta} - \frac{1}{(1+\theta)^2} > \frac{\theta^2}{(1+\theta)^4}.$$  

Exercise 24 (#3.42). Let $(X_1, \ldots, X_n)$ be a random sample from $N(\mu, \sigma^2)$ with an unknown $\mu \in \mathcal{R}$ and a known $\sigma^2 > 0$. Find the UMVE of $e^{t\mu}$ with a fixed $t \neq 0$ and show that the variance of the UMVE is larger than the Cramér-Rao lower bound but the ratio of the variance of the UMVE over the Cramér-Rao lower bound converges to 1 as $n \to \infty$.

**Solution.** The sample mean $\bar{X}$ is complete and sufficient for $\mu$. Since

$$E(e^{t\bar{X}}) = e^{t\mu t + \sigma^2 t^2/(2n)},$$

the UMVE of $e^{t\mu}$ is $T(X) = e^{-\sigma^2 t^2/(2n)} + t\bar{X}$.

The Fisher information $I(\mu) = n/\sigma^2$. Then the Cramér-Rao lower bound is $\left(\frac{d}{d\mu}e^{t\mu}\right)^2 / I(\mu) = \sigma^2 t^2 e^{2t\mu}/n$. On the other hand,

$$\text{Var}(T) = e^{-\sigma^2 t^2/n} E[e^{2t\bar{X}} - e^{2t\mu}] = \left(e^{\sigma^2 t^2/n} - 1\right) e^{2t\mu} > \frac{\sigma^2 t^2 e^{2t\mu}}{n},$$

the Cramér-Rao lower bound. The ratio of the variance of the UMVE over the Cramér-Rao lower bound is $(e^{\sigma^2 t^2/n} - 1)/\left(\sigma^2 t^2/n\right)$, which converges to 1 as $n \to \infty$, since $\lim_{x \to 0}(e^x - 1)/x = 1$. 

Exercise 25 (#3.46, #3.47). Let $X_1, X_2, \ldots$ be independent and identically distributed random variables, $m$ be a positive integer, and $h(x_1, \ldots, x_m)$ be a function on $\mathcal{R}^m$ such that $E[h(X_1, \ldots, X_m)]^2 < \infty$ and $h$ is symmetric in its $m$ arguments. A U-statistic with kernel $h$ (of order $m$) is defined as

$$U_n = \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} h(X_{i_1}, \ldots, X_{i_m}),$$
where $\sum_{1 \leq i_1 < \cdots < i_m \leq n}$ denotes the summation over the $\binom{n}{m}$ combinations of $m$ distinct elements $\{i_1, \ldots, i_m\}$ from $\{1, \ldots, n\}$. For $k = 1, \ldots, m$, define $h_k(x_1, \ldots, x_k) = E[h(x_1, \ldots, x_k, X_{k+1}, \ldots, X_m)]$ and $\zeta_k = \text{Var}(h_k(X_1, \ldots, X_k))$. Show that

(i) $\zeta_1 \leq \zeta_2 \leq \cdots \leq \zeta_m$;
(ii) $(n + 1)\text{Var}(U_{n+1}) \leq n\text{Var}(U_n)$ for any $n \geq m$;
(iii) if $\zeta_j = 0$ for $j < k$ and $\zeta_k > 0$, where $1 \leq k \leq m$, then
\[
\text{Var}(U_n) = \frac{k! \binom{m}{k}^2 \zeta_k}{n^k} + O\left(\frac{1}{n^{k+1}}\right);
\]
(iv) $m^2 \zeta_1 \leq n\text{Var}(U_n) \leq m\zeta_m$ for any $n \geq m$.

**Solution.** (i) For any $k = 1, \ldots, m - 1$, let $W = h_{k+1}(X_1, \ldots, X_k, X_{k+1})$ and $Y = (X_1, \ldots, X_k)$. Then $\zeta_{k+1} = \text{Var}(W)$ and $\zeta_k = \text{Var}(E(W|Y))$, since
\[
E(W|Y) = E[h_{k+1}(X_1, \ldots, X_k, X_{k+1})|X_1, \ldots, X_k] = h_k(X_1, \ldots, X_k).
\]
The result follows from
\[
\text{Var}(W) = E\{E[(W - EW)^2|Y]\} \geq E\{[E(W|Y) - EW]^2\} = \text{Var}(E(W|Y)),
\]
where the inequality follows from Jensen’s inequality for conditional expectations and the equality follows from $EW = E[E(W|Y)]$.

(ii) We use induction. The result is obvious when $m = 1$, since $U$ is an average of independent and identically distributed random variables when $m = 1$. Assume that the result holds for any U-statistic with a kernel of order $m - 1$. From Hoeffding’s representation (e.g., Serfling, 1980, p. 178),
\[
U_n - EU_n = W_n + S_n,
\]
where $W_n$ is a U-statistic with a kernel of order $m - 1$, $S_n$ is a U-statistic with variance $\binom{n}{m}^{-1} \eta_m$, $\eta_m$ is a constant not depending on $n$, and $\text{Var}(U_n) = \text{Var}(W_n) + \text{Var}(S_n)$. By the induction assumption, $(n + 1)\text{Var}(W_{n+1}) \leq n\text{Var}(W_n)$. Then, for any $n \geq m$,
\[
n\text{Var}(U_n) = n\text{Var}(W_n) + n\text{Var}(S_n)
= n\text{Var}(W_n) + n\binom{n}{m}^{-1} \eta_m
\geq (n + 1)\text{Var}(W_{n+1}) + \frac{m! \eta_m}{(n - 1)(n - 2) \cdots (n - m + 1)}
\geq (n + 1)\text{Var}(W_{n+1}) + \frac{(n + 1)(n + 1)}{n(n - 1) \cdots (n - m + 2)} \eta_m
= (n + 1)\text{Var}(W_{n+1}) + (n + 1)\binom{n}{m}^{-1} \eta_m
= (n + 1)\text{Var}(U_{n+1}).
(iii) From Hoeffding’s theorem (e.g., Theorem 3.4 in Shao, 2003),

\[
\text{Var}(U_n) = \sum_{l=1}^{m} \frac{(m)}{l} \frac{(n-m)}{m-l} \zeta_l.
\]

For any \(l = 1, ..., m,\)

\[
\frac{(m)}{l} \frac{(n-m)}{m-l} = l! \left( \frac{m}{l} \right)^2 \frac{(n-m)(n-m-1) \cdots [n-m-(m-l-1)]}{n(n-1) \cdots [n-(m-1)]}
\]

\[
= l! \left( \frac{m}{l} \right)^2 \left[ \frac{1}{n^l} + O \left( \frac{1}{n^{l+1}} \right) \right]
\]

\[
= O \left( \frac{1}{n^l} \right).
\]

If \(\zeta_j = 0\) for \(j < k\) and \(\zeta_k > 0,\) where \(1 \leq k \leq m,\) then

\[
\text{Var}(U_n) = \sum_{l=k}^{m} \frac{(m)}{l} \frac{(n-m)}{m-l} \zeta_l
\]

\[
= \frac{(m)}{k} \frac{(n-m)}{m-k} \zeta_k + \sum_{l=k+1}^{m} \frac{(m)}{l} \frac{(n-m)}{m-l} \zeta_l
\]

\[
= k! \left( \frac{m}{k} \right)^2 \zeta_k + \sum_{l=k+1}^{m} O \left( \frac{1}{n^{l+1}} \right)
\]

\[
= k! \left( \frac{m}{k} \right)^2 \zeta_k + O \left( \frac{1}{n^{k+1}} \right).
\]

(iv) From the result in (ii), \(n \text{Var}(U_n)\) is nonincreasing in \(n.\) Hence \(n \text{Var}(U_n) \leq m \text{Var}(U_m) = m \zeta_m\) for any \(n \geq m.\) Also, \(\lim_n [n \text{Var}(U_n)] \leq n \text{Var}(U_n)\) for any \(n \geq m.\) If \(\zeta_1 > 0,\) from the result in (iii), \(\lim_n [n \text{Var}(U_n)] = m^2 \zeta_1.\) Hence, \(m^2 \zeta_1 \leq n \text{Var}(U_n)\) for any \(n \geq m,\) which obviously also holds if \(\zeta_1 = 0.\)

**Exercise 26 (#3.53).** Let \(h(x_1, x_2, x_3) = I_{(-\infty, 0)}(x_1 + x_2 + x_3).\) Find \(h_k\) and \(\zeta_k,\) \(k = 1, 2, 3,\) for the U-statistic with kernel \(h\) based on independent random variables \(X_1, X_2, ...\) with a common cumulative distribution function \(F.\)

**Solution.** Let \(G * H\) denote the convolution of the two cumulative distribution functions \(G\) and \(H.\) Then

\[
h_1(x_1) = E[I_{(-\infty, 0)}(x_1 + X_2 + X_3)] = F * F(-x_1),
\]

\[
h_2(x_1, x_2) = E[I_{(-\infty, 0)}(x_1 + x_2 + X_3)] = F(-x_1 - x_2),
\]
Chapter 3. Unbiased Estimation

\[ h_3(x_1, x_2, x_3) = I_{(-\infty, 0)}(x_1 + x_2 + x_3), \]
\[ \zeta_1 = \text{Var}(F \ast F(-X_1)), \]
\[ \zeta_2 = \text{Var}(F(-X_1 - X_2)), \]

and
\[ \zeta_3 = F \ast F \ast F(0)[1 - F \ast F \ast F(0)]. \]

Exercise 27 (#3.54). Let \( X_1, \ldots, X_n \) be a random sample of random variables having finite \( EX_1^2 \) and \( EX_1^{-2} \). Let \( \mu = EX_1 \) and \( \bar{\mu} = E X_1^{-1} \). Find a U-statistic that is an unbiased estimator of \( \mu \bar{\mu} \) and derive its variance and asymptotic distribution.

Solution. Consider \( h(x_1, x_2) = (\frac{x_1}{x_2} + \frac{x_2}{x_1})/2 \). Then the U-statistic
\[ U_n = \frac{1}{n(n-1)} \sum_{1 \leq i < j \leq n} \left( \frac{X_i}{X_j} + \frac{X_j}{X_i} \right) \]
is unbiased for \( E[h(X_1, X_2)] = \mu \bar{\mu} \). Define \( h_1(x) = (x \mu + x^{-1} \mu)/2 \). Then
\[ \zeta_1 = \text{Var}(h(X_1)) = \frac{\bar{\mu}^2 V(X_1) + \mu^2 \text{Var}(X_1^{-1}) + 2 \mu \bar{\mu}(1 - \mu \bar{\mu})}{4}. \]

By Theorem 3.5 in Shao (2003),
\[ \sqrt{n}(U_n - \mu \bar{\mu}) \rightarrow_d N(0, 4\zeta_1). \]

Using the formula for the variance of U-statistics given in the solution of the previous exercise, we obtain the variance of \( U_n \) as \( [4(n-2)\zeta_1 + 2\zeta_2]/[n(n-1)] \), where \( \zeta_2 = \text{Var}(h(X_1, X_2)) \).

Exercise 28 (#3.58). Suppose that
\[ X_{ij} = \alpha_i + \theta t_{ij} + \varepsilon_{ij}, \quad i = 1, \ldots, a, \quad j = 1, \ldots, b, \]
where \( \alpha_i \) and \( \theta \) are unknown parameters, \( t_{ij} \) are known constants, and \( \varepsilon_{ij} \) are independent and identically distributed random variables with mean 0. Find explicit forms for the least squares estimators (LSE’s) of \( \theta, \alpha_i, \) \( i = 1, \ldots, a \).

Solution. Write the model in the form of \( X = Z \beta + \varepsilon \), where
\[ X = (X_{11}, \ldots, X_{1b}, \ldots, X_{a1}, \ldots, X_{ab}), \]
\[ \beta = (\alpha_1, \ldots, \alpha_a, \theta), \]
and
\[ \varepsilon = (\varepsilon_{11}, \ldots, \varepsilon_{1b}, \ldots, \varepsilon_{a1}, \ldots, \varepsilon_{ab}). \]
Then the design matrix $Z$ is

$$Z = \begin{pmatrix} J_b & 0 & 0 & t_1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & J_b & t_a \end{pmatrix},$$

where $t_i = (t_{i1}, \ldots, t_{ib})$ and $J_b$ is the $b$-vector of 1's. Solving the normal equation $(Z^t au Z) \hat{\beta} = Z^t X$, we obtain the LSE's

$$\hat{\theta} = \frac{\sum_{i=1}^a \sum_{j=1}^b t_{ij} X_{ij} - b \sum_i \bar{t}_i \bar{X}_i}{\sum_{j=1}^b (t_{ij} - \bar{t}_i)^2},$$

where $\bar{t}_i = \frac{1}{b} \sum_{j=1}^b t_{ij}$, $\bar{X}_i = \frac{1}{b} \sum_{j=1}^b X_{ij}$, and

$$\hat{\alpha}_i = \bar{X}_i - \hat{\theta} \bar{t}_i, \quad i = 1, \ldots, a.$$

**Exercise 29 (#3.59).** Consider the polynomial model

$$X_i = \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \beta_3 t_i^3 + \varepsilon_i, \quad i = 1, \ldots, n,$$

where $\varepsilon_i$'s are independent and identically distributed random variables with mean 0. Suppose that $n = 12$, $t_i = -1$, $i = 1, \ldots, 4$, $t_i = 0$, $i = 5, \ldots, 8$, and $t_i = 1$, $i = 9, \ldots, 12$. Show whether the following parameters are estimable (i.e., they can be unbiasedly estimated): $\beta_0 + \beta_2$, $\beta_1$, $\beta_0 - \beta_1$, $\beta_1 + \beta_3$, and $\beta_0 + \beta_1 + \beta_2 + \beta_3$.

**Solution.** Let $X = (X_1, \ldots, X_{12})$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{12})$, and $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$. Then $X = Z \beta + \varepsilon$ with

$$Z^\tau = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

and

$$Z^\tau Z = \begin{pmatrix} 12 & 0 & 8 & 0 \\ 0 & 8 & 0 & 8 \\ 8 & 0 & 8 & 0 \\ 0 & 8 & 0 & 8 \end{pmatrix}.$$

From the theory of linear models (e.g., Theorem 3.6 in Shao, 2003), a parameter $l^\tau \beta$ with a known vector $l$ is estimable if and only if $l \in \mathcal{R}(Z^\tau Z)$. Note that $\beta_0 + \beta_2 = l^\tau \beta$ with $l = (1, 0, 1, 0)$, which is the third row of $Z^\tau Z$ divided by 8. Hence $\beta_0 + \beta_2$ is estimable. Similarly, $\beta_1 + \beta_3 = l^\tau \beta$...
with \( l = (0, 1, 0, 1) \), which is the second row of \( Z^\top Z \) divided by 8 and, hence, \( \beta_1 + \beta_3 \) is estimable. Then \( \beta_0 + \beta_1 + \beta_2 + \beta_3 \) is estimable, since any linear combination of estimable functions is estimable. We now show that \( \beta_0 - \beta_1 = l^\top \beta \) with \( l = (1, -1, 0, 0) \) is not estimable. If \( \beta_0 - \beta_1 \) is estimable, then there is \( c = (c_1, \ldots, c_4) \) such that \( l = Z^\top Z c \), i.e.,

\[
\begin{align*}
12c_1 + 8c_3 &= 1 \\
8c_2 + 8c_4 &= -1 \\
8c_1 + 8c_3 &= 0 \\
8c_2 + 8c_4 &= 0,
\end{align*}
\]

where the second and the last equations have no solution. Similarly, the parameter \( \beta_1 \) is not estimable, since \( 8c_2 + 8c_4 = 1 \) and \( 8c_2 + 8c_4 = 0 \) cannot hold at the same time.

Exercise 30 (#3.60). Consider the one-way ANOVA model

\[
X_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad j = 1, \ldots, n_i, i = 1, \ldots, m,
\]

where \( \mu \) and \( \alpha_i \) are unknown parameters and \( \varepsilon_{ij} \) are independent and identically distributed random variables with mean 0. Let

\[
X = (X_{11}, \ldots, X_{1n_1}, \ldots, X_{m1}, \ldots, X_{mn_m}),
\]

\[
\varepsilon = (\varepsilon_{11}, \ldots, \varepsilon_{1n_1}, \ldots, \varepsilon_{m1}, \ldots, \varepsilon_{mn_m}),
\]

and \( \beta = (\mu, \alpha_1, \ldots, \alpha_m) \). Find the matrix \( Z \) in the linear model \( X = Z\beta + \varepsilon \), the matrix \( Z^\top Z \), and the form of \( l \) for estimable \( l^\top \beta \).

Solution. Let \( n = n_1 + \cdots + n_m \) and \( J_a \) be the \( a \)-vector of 1’s. Then

\[
Z = \begin{pmatrix}
J_{n_1} & J_{n_1} & 0 & \cdots & 0 \\
J_{n_2} & 0 & J_{n_2} & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
J_{n_m} & 0 & 0 & \cdots & J_{n_m}
\end{pmatrix}
\]

and

\[
Z^\top Z = \begin{pmatrix}
n & n_1 & n_2 & \cdots & n_m \\
n_1 & n_1 & 0 & \cdots & 0 \\
n_2 & 0 & n_2 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
n_m & 0 & 0 & \cdots & n_m
\end{pmatrix}.
\]

Note that \( l^\top \beta \) is estimable if and only if \( l \in \mathcal{R}(Z^\top Z) \), the linear space generated by the rows of \( Z^\top Z \). We now show that \( l^\top \beta \) is estimable if and only if \( l_0 = l_1 + \cdots + l_m \) for \( l = (l_0, l_1, \ldots, l_m) \in \mathcal{R}^{m+1} \).
If \( l \in \mathcal{R}(Z^\tau Z) \), then there is a \( c = (c_0, c_1, \ldots, c_m) \in \mathcal{R}^{m+1} \) such that
\[
nc_0 + n_1 c_1 + \cdots + n_m c_m = l_0
\]
\[
n_1 c_0 + n_1 c_1 = l_1
\]
..............
\[
n_m c_0 + n_m c_m = l_m
\]
holds. Then \( l_0 = l_1 + \cdots + l_m \). On the other hand, if \( l_0 = l_1 + \cdots + l_m \), then the previous \( m+1 \) equations with \( c_0, c_1, \ldots, c_m \) considered as variables have infinitely many solutions. Hence \( l \in \mathcal{R}(Z^\tau Z) \).

**Exercise 31 (#3.61).** Consider the two-way balanced ANOVA model
\[
X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, \quad i = 1, \ldots, a, j = 1, \ldots, b, k = 1, \ldots, c,
\]
where \( a, b, \) and \( c \) are some positive integers, \( \varepsilon_{ijk} \)'s are independent and identically distributed random variables with mean 0, and \( \mu, \alpha_i, \beta_j, \) and \( \gamma_{ij} \)'s are unknown parameters. Let \( X \) be the vector of \( X_{ijk} \)'s, \( \varepsilon \) be the vector of \( \varepsilon_{ijk} \)'s, and \( \beta = (\mu, \alpha_1, \ldots, \alpha_a, \beta_1, \ldots, \beta_b, \gamma_{11}, \ldots, \gamma_{1b}, \ldots, \gamma_{a1}, \ldots, \gamma_{ab}) \).

(i) Obtain the design matrix \( Z \) in the model \( X = Z\beta + \varepsilon \) and show that the rank of \( Z \) is \( ab \).

(ii) Find the form of estimable \( l^\tau \beta, l \in \mathcal{R}^{1+a+b+ab} \).

(iii) Obtain an LSE of \( \beta \).

**Solution.** (i) Let \( J_t \) be the \( t \)-vector of 1’s, \( I_t \) be the identity matrix of order \( t \), \( A \) be the \( ab \times b \) block diagonal matrix whose \( j \)th diagonal block is \( J_a, j = 1, \ldots, b, \)
\[
B = (I_b I_b \cdots I_b),
\]
and
\[
\Lambda = (J_{ab} A B^\tau I_{ab}),
\]
which is an \( ab \times (1+a+b+ab) \) matrix. Then \( Z \) is the \( (1+a+b+ab) \times abc \) matrix whose transpose is
\[
Z^\tau = (A^\tau A^\tau \cdots A^\tau)
\]
and
\[
Z^\tau Z = c\Lambda^\tau \Lambda = c \begin{pmatrix} \Lambda_0^\tau \Lambda_0 & \Lambda_0^\tau \\ \Lambda_0 & I_{ab} \end{pmatrix},
\]
where \( \Lambda_0 = (J_{ab} A B^\tau) \). Clearly, the last \( ab \) rows of \( Z^\tau Z \) are linearly independent. Hence the rank of \( Z \), which is the same as the rank of \( Z^\tau Z \), is no smaller than \( ab \). On the other hand, the rank of \( \Lambda \) is no larger than \( ab \) and, hence, the rank of \( Z^\tau Z \) is no larger than \( ab \). Thus, the rank of \( Z \) is \( ab \).

(ii) A function \( l^\tau \beta \) with \( l \in \mathcal{R}^{1+a+b+ab} \) is estimable if and only if \( l \) is a linear combination of the rows of \( Z^\tau Z \). From the discussion in part (i)
of the solution, we know that \( l'\beta \) is estimable if and only if \( l \) is a linear combination of the rows in the matrix \((A_0^\top I_{ab})\).

(iii) Any solution of \( Z\tau Z\beta = Z\tau X \) is an LSE of \( \beta \). A direct calculation shows that an LSE of \( \beta \) is \((\hat{\mu}, \hat{\alpha}_1, ..., \hat{\alpha}_a, \hat{\beta}_1, ..., \hat{\beta}_b, \hat{\gamma}_{11}, ..., \hat{\gamma}_{1b}, ..., \hat{\gamma}_{a1}, ..., \hat{\gamma}_{ab})\), where \( \hat{\mu} = X_i, ..., \hat{\alpha}_i = X_i - X, ..., \hat{\beta}_j = X_j - X, ..., \hat{\gamma}_{ij} = X_{ij} - X_i - X_j + X \), and a dot is used to denote averaging over the indicated subscript.

\[ \text{Exercise 32 (#3.63).} \]
Assume that \( X \) is a random \( n \)-vector from the multivariate normal distribution \( N_n(Z\beta, \sigma^2 I_n) \), where \( Z \) is an \( n \times p \) known matrix of rank \( r \leq p < n \), \( \beta \) is a \( p \)-vector of unknown parameters, \( I_n \) is the identity matrix of order \( n \), and \( \sigma^2 > 0 \) is unknown. Find the UMVUE’s of \((l'\beta)^2, \frac{l'\beta}{\sigma}, \frac{(l'\beta)^2}{\sigma^2}\) for an estimable \( l'\beta \).

\[ \text{Solution.} \]
Let \( \beta \) be the LSE of \( \beta \) and \( \hat{\sigma}^2 = \frac{||X - Z\hat{\beta}||^2}{(n - r)} \). Note that \((Z'X, \hat{\sigma}^2)\) is complete and sufficient for \((\beta, \sigma^2)\), \( \hat{\beta} \) has the normal distribution \( N(I'\beta, \sigma^2 I'(Z'Z)^{-1}) \), and \((n - r)\hat{\sigma}^2/\sigma^2 \) has the chi-square distribution \( \chi^2_{n-r} \), where \( A^- \) is a generalized inverse of \( A \). Since \( E(I'\hat{\beta}) = [E(I'\beta)] + \text{Var}(I'\hat{\beta}) = (I'\beta)^2 + \sigma^2 I'(Z'Z)^{-1} \), the UMVUE of \((l'\beta)^2\) is \((I'\beta)^2 - \hat{\sigma}^2 I'(Z'Z)^{-1} \). Since \( \kappa_{n-r,1}\hat{\sigma}^{-1} \) is the UMVUE of \( \sigma^{-1} \), where \( \kappa_{n-r,1} \) is given in Exercise 4, and \( I'\beta \) is independent of \( \hat{\sigma}^2, \kappa_{n-r,1} I'\hat{\beta}\hat{\sigma}^{-1} \) is the UMVUE of \( I'\beta/\sigma \). A similar argument yields the UMVUE of \((I'\beta/\sigma)^2\) as \((\kappa_{n-r,2}(I'\hat{\beta})^2\hat{\sigma}^{-2} - l'(Z'Z)^{-1}) \).

\[ \text{Exercise 33 (#3.65).} \]
Consider the one-way random effects model

\[ X_{ij} = \mu + A_i + e_{ij}, \quad j = 1, ..., n, i = 1, ..., m, \]

where \( \mu \in \mathcal{R} \) is an unknown parameter, \( A_i \)'s are independent and identically distributed as \( N(0, \sigma^2_a) \), \( e_{ij} \)'s are independent and identically distributed as \( N(0, \sigma^2) \), and \( A_i \)'s and \( e_{ij} \)'s are independent. Based on observed \( X_{ij} \)'s, show that the family of populations is an exponential family with sufficient and complete statistics \( \bar{X}_., S_A = n \sum_{i=1}^m (\bar{X}_i - \bar{X}_.)^2, \) and \( S_E = \sum_{i=1}^m \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2, \) where \( \bar{X}_i = (nm)^{-1} \sum_{j=1}^n X_{ij} \) and \( \bar{X}_i = n^{-1} \sum_{j=1}^n X_{ij} \). Find the UMVUE’s of \( \mu, \sigma^2_a, \) and \( \sigma^2 \).

\[ \text{Solution.} \]
Let \( X_i = (X_{i1}, ..., X_{in}) \), \( i = 1, ..., m \). Then \( X_1, ..., X_m \) are independent and identically distributed as the multivariate normal distribution \( N_n(\mu J_n, \Sigma) \), where \( J_n \) is the \( n \)-vector of 1’s and \( \Sigma = \sigma^2_a J_n J_n^\top + \sigma^2 I_n \). The joint Lebesgue density of \( X_i \)'s is

\[ (2\pi)^{-\frac{mn}{2}} |\Sigma|^{-\frac{m}{2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (X_i - \mu J_n)^\top \Sigma^{-1} (X_i - \mu J_n) \right\}. \]

Note that

\[ \Sigma^{-1} = (\sigma^2_a J_n J_n^\top + \sigma^2 I_n)^{-1} = \frac{1}{\sigma^2} I_n - \frac{\sigma^2_a}{\sigma^2(\sigma^2 + n\sigma^2_a)} J_n J_n^\top. \]
Hence, the sum in the exponent of the joint density is equal to

\[
\sum_{i=1}^{m} (X_i - \mu J_n)^{\top} \Sigma^{-1} (X_i - \mu J_n)
\]

\[
= \frac{1}{\sigma^2} \sum_{i=1}^{m} \sum_{j=1}^{n} (X_{ij} - \mu)^2 - \frac{n^2 \sigma_a^2}{\sigma^2 (\sigma^2 + n \sigma_a^2)} \sum_{i=1}^{m} (\bar{X}_i - \mu)^2
\]

\[
= \frac{1}{\sigma^2} \sum_{i=1}^{m} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2 + \frac{n}{\sigma^2 + n \sigma_a^2} \sum_{i=1}^{m} (\bar{X}_i - \mu)^2
\]

\[
= \frac{S_E}{\sigma^2} + \frac{S_A}{\sigma^2 + n \sigma_a^2} + \frac{n m}{\sigma^2 + n \sigma_a^2} \sum_{i=1}^{m} (\bar{X}_i - \mu)^2.
\]

Therefore, the joint density of $X_{ij}$’s is from an exponential family with $(\bar{X}_., S_A, S_E)$ as the sufficient and complete statistics for $(\mu, \sigma_a^2, \sigma^2)$. The UMVUE of $\mu$ is $\bar{X}_.$, since $EX_\cdot = \mu$. Since $E(S_E) = m(n - 1)\sigma^2$, the UMVUE of $\sigma^2$ is $SE/[m(n - 1)]$. Since $\bar{X}_i$, $i = 1, ..., m$, are independently from $N(\mu, \sigma_a^2 + \sigma^2/n)$, $E(S_A) = (m - 1)(\sigma^2 + n \sigma_a^2)$ and, thus, the UMVUE of $\sigma_a^2$ is $S_A/[n(m - 1)] - SE/[mn(n - 1)]$. \[\blacksquare\]

**Exercise 34 (#3.66).** Consider the linear model $X = Z\beta + \varepsilon$, where $Z$ is a known $n \times p$ matrix, $\beta$ is a $p$-vector of unknown parameters, and $\varepsilon$ is a random $n$-vector whose components are independent and identically distributed with mean 0 and Lebesgue density $\sigma^{-1} f(x/\sigma)$, where $f$ is a known Lebesgue density and $\sigma > 0$ is unknown. Find the Fisher information about $(\beta, \sigma)$ contained in $X$.

**Solution.** Let $Z_i$ be the $i$th row of $Z$, $i = 1, ..., n$. Consider a fixed $i$ and let $\theta = (Z_i^\top \beta, \sigma^2)$. The Lebesgue density of $X_i$, the $i$th component of $X$, is $\sigma^{-1} f((x - \theta)/\sigma)$. From Exercise 20, the Fisher information about $(\theta, \sigma)$ contained in $X_i$ is

\[I(\theta) = \frac{1}{\sigma^2} \begin{pmatrix}
\int \frac{[f'(x)]^2}{f(x)} dx & \int \frac{f'(x) [xf'(x) + f(x)]}{f(x)} dx \\
\int \frac{f'(x) [xf'(x) + f(x)]}{f(x)} dx & \int \frac{[xf'(x) + f(x)]^2}{f(x)} dx
\end{pmatrix}.
\]

Let $a_{ij}$ be the $(i,j)$th element of the matrix $\sigma^2 I(\theta)$. Since $X_i$’s are independent, $\frac{\partial \theta}{\partial \beta} = Z_i^\top$ and $\frac{\partial \theta}{\partial \sigma} = 1$, the Fisher information about $\eta = (\beta, \sigma)$ contained in $X$ is

\[
\sum_{i=1}^{n} \frac{\partial \theta}{\partial \eta} I(\theta) \frac{\partial \theta^*}{\partial \eta} = \frac{1}{\sigma^2} \sum_{i=1}^{n} \begin{pmatrix}
Z_i & 0 \\
0 & 1
\end{pmatrix} \begin{pmatrix}
a_{11} & a_{12} & Z_i^\top \\
a_{21} & a_{22} & 0
\end{pmatrix} \begin{pmatrix}
a_{11} Z_i Z_i^\top & a_{12} Z_i \\
a_{21} Z_i^\top & a_{22}
\end{pmatrix} = \frac{1}{\sigma^2} \sum_{i=1}^{n} \begin{pmatrix}
a_{11} Z_i Z_i^\top & a_{12} Z_i \\
a_{21} Z_i^\top & a_{22}
\end{pmatrix}. \[\blacksquare\]
Exercise 35 (#3.67). Consider the linear model \( X = Z\beta + \epsilon \), where \( Z \) is a known \( n \times p \) matrix, \( \beta \) is a \( p \)-vector of unknown parameters, and \( \epsilon \) is a random \( n \)-vector whose components are independent and identically distributed with mean 0 and variance \( \sigma^2 \). Let \( c \in \mathbb{R}^p \). Show that if the equation \( c = Z^\tau y \) has a solution, then there is a unique solution \( y_0 \in \mathcal{R}(Z^\tau) \) such that \( \text{Var}(y_0^\tau X) \leq \text{Var}(y^\tau X) \) for any other solution of \( c = Z^\tau y \).

Solution. Since \( c = Z^\tau y \) has a solution, \( c \in \mathcal{R}(Z) = \mathcal{R}(Z^\tau Z) \). Then, there is \( \lambda \in \mathbb{R}^p \) such that \( c = (Z^\tau Z)\lambda = Z^\tau y_0 \) with \( y_0 = Z\lambda \in \mathcal{R}(Z) \). This shows that \( c = Z^\tau y \) has a solution in \( \mathcal{R}(Z^\tau) \). Suppose that there is another \( y_1 \in \mathcal{R}(Z^\tau) \) such that \( c = Z^\tau y_1 \). Then \( y_0^\tau Z\beta = c^\tau \beta = y_1^\tau Z\beta \) for all \( \beta \in \mathbb{R}^p \). Since \( \mathcal{R}(Z^\tau) = \{Z\beta : \beta \in \mathbb{R}^p \} \), \( y_0 = y_1 \), i.e., the solution of \( c = Z^\tau y \) in \( \mathcal{R}(Z^\tau) \) is unique. For any \( y \in \mathbb{R}^n \) satisfying \( c = Z^\tau y \),

\[
\text{Var}(y^\tau X) = \text{Var}(y^\tau X - y_0^\tau X + y_0^\tau X) \\
= \text{Var}(y^\tau X - y_0^\tau X) + \text{Var}(y_0^\tau X) + 2\text{Cov}((y - y_0)^\tau X, y_0^\tau X) \\
= \text{Var}(y^\tau X - y_0^\tau X) + \text{Var}(y_0^\tau X) + 2E[(y - y_0)^\tau XX^\tau y_0] \\
= \text{Var}(y^\tau X - y_0^\tau X) + \text{Var}(y_0^\tau X) + 2\sigma^2(y - y_0)^\tau y_0 \\
= \text{Var}(y^\tau X - y_0^\tau X) + \text{Var}(y_0^\tau X) + 2(y_0^\tau X)\lambda \\
= \text{Var}(y^\tau X - y_0^\tau X) + \text{Var}(y_0^\tau X) + 2(c^\tau - c_0^\tau)\lambda \\
= \text{Var}(y^\tau X - y_0^\tau X) + \text{Var}(y_0^\tau X) \\
\geq \text{Var}(y_0^\tau X). 
\]

Exercise 36 (#3.69). Consider the linear model \( X = Z\beta + \epsilon \), where \( Z \) is a known \( n \times p \) matrix, \( \beta \) is a \( p \)-vector of unknown parameters, and \( \epsilon \) is a random \( n \)-vector whose components are independent and identically distributed with mean 0 and variance \( \sigma^2 \). Let \( X_i \) be the \( i \)-th component of \( X \), \( Z_i \) be the \( i \)-th row of \( Z \), \( h_{ij} \) be the \( (i,j) \)-th element of \( Z(Z^\tau Z)^{-1}Z^\tau \), \( h_i = h_{ii} \), \( \hat{\beta} \) be an LSE of \( \beta \), and \( \hat{X}_i = Z_i^\tau \hat{\beta} \). Show that

(i) \( \text{Var}(\hat{X}_i) = \sigma^2 h_i \);

(ii) \( \text{Var}(X_i - \hat{X}_i) = \sigma^2(1 - h_i) \);

(iii) \( \text{Cov}(\hat{X}_i, \hat{X}_j) = \sigma^2 h_{ij} \);

(iv) \( \text{Cov}(X_i - \hat{X}_i, X_j - \hat{X}_j) = -\sigma^2 h_{ij}, i \neq j \);

(v) \( \text{Cov}(\hat{X}_i, X_j - \hat{X}_j) = 0 \).

Solution. (i) Since \( Z_i \in \mathcal{R}(Z) \), \( Z_i^\tau \hat{\beta} \) is estimable and

\[
\text{Var}(Z_i^\tau \hat{\beta}) = \sigma^2 Z_i^\tau (Z^\tau Z)^{-1} Z_i = \sigma^2 h_i.
\]

(ii) Note that

\[
\hat{X}_i = Z_i^\tau \hat{\beta} = Z_i^\tau (Z^\tau Z)^{-1} Z^\tau X = \sum_{j=1}^{n} h_{ij} X_j.
\]
Hence,
\[ X_i - \hat{X}_i = (1 - h_i)X_i - \sum_{j \neq i} h_{ij}X_j. \]

Since \( X_i \)'s are independent and \( \text{Var}(X_i) = \sigma^2 \), we obtain that
\[ \text{Var}(X_i - \hat{X}_i) = (1 - h_i)^2 \sigma^2 + \sigma^2 \sum_{j \neq i} h_{ij}^2 \]
\[ = (1 - h_i)^2 \sigma^2 + (h_i - h_i^2)\sigma^2 \]
\[ = (1 - h_i)\sigma^2, \]
where the second equality follows from the fact that \( \sum_{j=1}^n h_{ij}^2 = h_{ii} = h_i \), a property of the projection matrix \( \mathbf{Z}(\mathbf{Z}^\tau\mathbf{Z})^{-1}\mathbf{Z}^\tau \).

(iii) Using the formula for \( \hat{X}_i \) in part (ii) of the solution and the independence of \( X_i \)'s,
\[ \text{Cov}(\hat{X}_i, \hat{X}_j) = \text{Cov} \left( \sum_{k=1}^n h_{ik}X_k, \sum_{l=1}^n h_{ji}X_l \right) = \sigma^2 \sum_{k=1}^m h_{ik}h_{jk} = \sigma^2 h_{ij}, \]
where the last equality follows from the fact that \( \mathbf{Z}(\mathbf{Z}^\tau\mathbf{Z})^{-1}\mathbf{Z}^\tau \) is a projection matrix.

(iv) For \( i \neq j \),
\[ \text{Cov}(X_i, \hat{X}_j) = \text{Cov} \left( X_i, \sum_{k=1}^n h_{jk}X_k \right) = \sigma^2 h_{ij} \]
and, thus,
\[ \text{Cov}(X_i - \hat{X}_i, X_j - \hat{X}_j) = -\text{Cov}(X_i, \hat{X}_j) - \text{Cov}(X_j, \hat{X}_i) + \text{Cov}(\hat{X}_i, \hat{X}_j) \]
\[ = -\sigma^2 h_{ij} - \sigma^2 h_{ji} + \sigma^2 h_{ij} \]
\[ = -\sigma^2 h_{ij}. \]

(v) From part (iii) and part (iv) of the solution,
\[ \text{Cov}(\hat{X}_i, X_j - \hat{X}_j) = \text{Cov}(\hat{X}_i, X_j) - \text{Cov}(\hat{X}_i, \hat{X}_j) = \sigma^2 h_{ij} - \sigma^2 h_{ij} = 0. \]

**Exercise 37 (#3.70).** Consider the linear model \( X = Z\beta + \varepsilon \), where \( Z \) is a known \( n \times p \) matrix, \( \beta \) is a \( p \)-vector of unknown parameters, and \( \varepsilon \) is a random \( n \)-vector whose components are independent and identically distributed with mean 0 and variance \( \sigma^2 \). Let \( Z = (Z_1, Z_2) \) and \( \beta = (\beta_1, \beta_2) \), where \( Z_j \) is \( n \times p_j \) and \( \beta_j \) is a \( p_j \)-vector, \( j = 1, 2 \). Assume that \( (Z_1^\tau Z_1)^{-1} \) and \( [Z_2^\tau Z_2 - Z_2^\tau Z_1(Z_1^\tau Z_1)^{-1}Z_1^\tau Z_2]^{-1} \) exist.

(i) Derive the LSE of \( \beta \) in terms of \( Z_1, Z_2, \) and \( X \).
(ii) Let $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)$ be the LSE in (i). Calculate the covariance between $\hat{\beta}_1$ and $\hat{\beta}_2$.

(iii) Suppose that it is known that $\beta_2 = 0$. Let $\tilde{\beta}_1$ be the LSE of $\beta_1$ under the reduced model $X = Z_1 \beta_1 + \varepsilon$. Show that, for any $l \in \mathbb{R}^{p_1}$, $l^\tau \tilde{\beta}_1$ is better than $l^\tau \hat{\beta}_1$ in terms of their variances.

**Solution.** (i) Note that
\[
Z^\tau Z = \begin{pmatrix}
Z^\tau_1 Z_1 & Z^\tau_1 Z_2 \\
Z^\tau_2 Z_1 & Z^\tau_2 Z_2
\end{pmatrix}.
\]

From matrix algebra,
\[
(Z^\tau Z)^{-1} = \begin{pmatrix}
A & B \\
B^\tau & C
\end{pmatrix},
\]
where
\[
C = [Z^\tau_2 Z_2 - Z^\tau_2 Z_1 (Z^\tau_1 Z_1)^{-1} Z^\tau_1 Z_2]^{-1},
B = -(Z^\tau_1 Z_1)^{-1} C
\]
and
\[
A = (Z^\tau_1 Z_1)^{-1} + (Z^\tau_1 Z_1)^{-1} Z^\tau_1 Z_2 C Z^\tau_2 Z_1 (Z^\tau_1 Z_1)^{-1}.
\]
The LSE of $\beta$ is
\[
\hat{\beta} = (Z^\tau Z)^{-1} Z^\tau X = \begin{pmatrix}
A & B \\
B^\tau & C
\end{pmatrix} \begin{pmatrix}
Z^\tau_1 X \\
Z^\tau_2 X
\end{pmatrix} = \begin{pmatrix}
AZ^\tau_1 X + BZ^\tau_2 X \\
B^\tau Z^\tau_1 X + CZ^\tau_2 X
\end{pmatrix}.
\]

(ii) Since $\text{Var}(\hat{\beta}) = \sigma^2 (Z^\tau Z)^{-1}$, $\text{Cov}(\hat{\beta}_1, \hat{\beta}_2) = \sigma^2 B$.

(iii) Note that $\text{Var}(l^\tau \hat{\beta}_1) = \sigma^2 l^\tau (Z^\tau_1 Z_1)^{-1} l$. From part (i) of the solution,
\[
\text{Var}(l^\tau \hat{\beta}_1) = \sigma^2 l^\tau A l \geq \sigma^2 l^\tau (Z^\tau_1 Z_1)^{-1} l.
\]

**Exercise 38** (#3.71, #3.72). Consider the linear model $X = Z \beta + \varepsilon$, where $Z$ is a known $n \times p$ matrix, $\beta$ is a $p$-vector of unknown parameters, and $\varepsilon$ is a random $n$-vector with $E(\varepsilon) = 0$ and finite $\text{Var}(\varepsilon) = \Sigma$. Show the following statements are equivalent:

(a) The LSE $l^\tau \beta$ is the best linear unbiased estimator (BLUE) of $l^\tau \beta$.
(b) $\text{Var}(\varepsilon) = Z \Lambda_1 Z^\tau + U \Lambda_2 U^\tau$ for some matrices $\Lambda_1$ and $\Lambda_2$, where $U$ is a matrix such that $Z^\tau U = 0$ and $R(U^\tau) + R(Z^\tau) = \mathbb{R}^n$.
(c) $\text{Var}(\varepsilon) = ZB$ for some matrix $B$.
(d) $R(Z^\tau)$ is generated by $r$ eigenvectors of $\text{Var}(\varepsilon)$, where $r$ is the rank of $Z$.

**Solution.** (i) From the proof in Shao (2003, p. 191), (a) is equivalent to (c) $Z^\tau \text{Var}(\varepsilon) U = 0$ and (c) implies (e). Hence, to show that (a) and (e) are
equivalent, it suffices to show that (e) implies (c). Since \(Z(Z^T Z)^{-1} Z^T Z = Z\), (e) implies that
\[
Z^T \text{Var}(\varepsilon) U = Z^T ZZ(Z^T Z)^{-1} Z^T \text{Var}(\varepsilon) U = Z^T \text{Var}(\varepsilon) Z(Z^T Z)^{-1} Z^T U = 0.
\]

(ii) We now show that (f) and (c) are equivalent. If (f) holds, then \(\text{Var}(\varepsilon) Z = ZB\) for some matrix \(B\) and
\[
Z^T \text{Var}(\varepsilon) U = B^T Z^T U = 0.
\]
If (c) holds, then (e) holds. Then
\[
\text{Var}(\varepsilon) Z = \text{Var}(\varepsilon) Z(Z^T Z)^{-1} Z^T Z = Z(Z^T Z)^{-1} Z^T \text{Var}(\varepsilon) Z
\]
and (f) holds with \(B = (Z^T Z)^{-1} Z^T \text{Var}(\varepsilon) Z\).

(iii) Assume that (g) holds. Then \(\mathcal{R}(Z^T) = \mathcal{R}(\xi_1, \ldots, \xi_r)\), the linear space generated by \(r\) linearly independent eigenvectors \(\xi_1, \ldots, \xi_r\) of \(\text{Var}(\varepsilon)\). Let \(\xi_{r+1}, \ldots, \xi_n\) be the other \(n-r\) linearly independent eigenvectors of \(\text{Var}(\varepsilon)\) that are orthogonal to \(\xi_1, \ldots, \xi_r\). Then \(\mathcal{R}(U^T) = \mathcal{R}(\xi_{r+1}, \ldots, \xi_n)\). For \(j \leq r\), \(\text{Var}(\varepsilon) \xi_j = a_j \xi_j\) for some constant \(a_j\). For \(k \geq r+1\), \(\xi_j^T \text{Var}(\varepsilon) \xi_k = a \xi_j^T \xi_k = 0\). Hence, \(Z^T \text{Var}(\varepsilon) U = 0\), i.e., (c) holds.

Now, assume (c) holds. Let \(\xi_1, \ldots, \xi_n\) be \(n\) orthogonal eigenvectors of \(\text{Var}(\varepsilon)\) and \(M\) be the matrix with \(\xi_i\) as the \(i\)th column. Decompose \(M\) as \(M = M_Z + M_U\), where columns of \(M_Z\) are in \(\mathcal{R}(Z^T)\) and columns of \(M_U\) are in \(\mathcal{R}(U^T)\). Then
\[
\text{Var}(\varepsilon) M_Z + \text{Var}(\varepsilon) M_U = M_Z D + M_U D,
\]
where \(D\) is a diagonal matrix. Multiplying the transposes of both sides of the above equation by \(M_U^T\) from the right, we obtain that, by (c),
\[
M_U^T \text{Var}(\varepsilon) M_U = D M_U^T M_U
\]
which is the same as
\[
\text{Var}(\varepsilon) M_U = M_U D,
\]
and, hence,
\[
\text{Var}(\varepsilon) M_Z = M_Z D.
\]
This means that column vectors of \(M_Z\) are eigenvectors of \(\text{Var}(\varepsilon)\). Then (g) follows from \(\mathcal{R}(Z) = \mathcal{R}(M_Z)\).

**Exercise 39 (#3.74).** Suppose that
\[
X = \mu J_n + H \xi + e,
\]
where \(\mu \in \mathcal{R}\) is an unknown parameter, \(J_n\) is the \(n\)-vector of 1’s, \(H\) is an \(n \times p\) known matrix of full rank, \(\xi\) is a random \(p\)-vector with \(E(\xi) = 0\) and
\[ \text{Var}(\xi) = \sigma_\xi^2 I_p, \text{ e is a random } n\text{-vector with } E(e) = 0 \text{ and } \text{Var}(e) = \sigma^2 I_n, \text{ and } \xi \text{ and } e \text{ are independent. Show that the LSE of } \mu \text{ is the BLUE if and only if the row totals of } HH^\tau \text{ are the same.} \]

**Solution.** From the result in the previous exercise, it suffices to show that the LSE of \( \mu \) is the BLUE if and only if \( J_n \) is an eigenvector of \( \text{Var}(H\xi + e) = \sigma^2 HH^\tau + \sigma^2 I_n \). Since

\[ (\sigma^2 HH^\tau + \sigma^2 I_n)J_n = \sigma^2 \eta + \sigma^2 J_n, \]

where \( \eta \) is the vector of row totals of \( HH^\tau \), \( J_n \) is an eigenvector of the matrix \( \text{Var}(H\xi + e) \) if and only if \( \eta = cJ_n \) for some constant. ☐

**Exercise 40 (\#3.75).** Consider a linear model

\[ X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \quad i = 1, \ldots, a, j = 1, \ldots, b, \]

where \( \mu, \alpha_i \)'s, and \( \beta_j \)'s are unknown parameters, \( E(\varepsilon_{ij}) = 0 \), \( \text{Var}(\varepsilon_{ij}) = \sigma^2 \), \( \text{Cov}(\varepsilon_{ij}, \varepsilon_{i'j'}) = 0 \) if \( i \neq i' \), and \( \text{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) = \sigma^2 \rho \) if \( j \neq j' \). Show that the LSE of \( l^\tau \beta \) is the BLUE for any \( l \in \mathcal{R}(Z) \).

**Solution.** Write the model in the form of \( X = Z\beta + \varepsilon \). Then \( \text{Var}(\varepsilon) \) is a block diagonal matrix whose \( j \)th diagonal block is \( \sigma^2 (1 - \rho) I_a + \sigma^2 \rho J_a J_a^\tau \), \( j = 1, \ldots, b \), where \( I_a \) is the identity matrix of order \( a \) and \( J_a \) is the \( a \)-vector of 1's. Let \( A \) and \( B \) be as defined in Exercise 31. Then \( Z = (J_{ab} \ A \ B^\tau) \).

Let \( \Lambda \) be the \((1 + a + b) \times (1 + a + b) \) matrix whose first element is \( \sigma^2 \rho \) and all the other elements are 0. Then, \( Z\Lambda Z^\tau \) is a block diagonal matrix whose \( j \)th diagonal block is \( \sigma^2 \rho J_a J_a^\tau \), \( j = 1, \ldots, b \). Thus,

\[ \text{Var}(\varepsilon) = \sigma^2 (1 - \rho) I_{ab} + Z\Lambda Z^\tau. \]

This shows that (c) in Exercise 38 holds. Hence, the LSE of \( l^\tau \beta \) is the BLUE for any \( l \in \mathcal{R}(Z) \). ☐

**Exercise 41 (\#3.76).** Consider the linear model \( X = Z\beta + \varepsilon \), where \( Z \) is a known \( n \times p \) matrix, \( \beta \) is a \( p \)-vector of unknown parameters, and \( \varepsilon \) is a random \( n \)-vector with \( E(\varepsilon) = 0 \) and \( \text{Var}(\varepsilon) \) is a block diagonal matrix whose \( i \)th block diagonal \( V_i \) is \( n_i \times n_i \) and has a single eigenvalue \( \lambda_i \) with eigenvector \( J_{n_i} \) (the \( n_i \)-vector of 1's) and a repeated eigenvalue \( \rho_i \) with multiplicity \( n_i - 1, i = 1, \ldots, k, \sum_{i=1}^k n_i = n \). Let \( U \) be the \( n \times k \) matrix whose \( i \)th column is \( U_i \), where \( U_1 = (J_{n_1}^\tau, 0, \ldots, 0), U_2 = (0, J_{n_2}^\tau, \ldots, 0), \ldots, U_k = (0, 0, \ldots, J_{n_k}^\tau) \), and let \( \hat{\beta} \) be the LSE of \( \beta \).

(i) If \( \mathcal{R}(Z^\tau) \subset \mathcal{R}(U^\tau) \) and \( \lambda_i \equiv \lambda \), show that \( l^\tau \hat{\beta} \) is the BLUE of \( l^\tau \beta \) for any \( l \in \mathcal{R}(Z) \).

(ii) If \( Z^\tau U_i = 0 \) for all \( i \) and \( \rho_i \equiv \rho \), show that \( l^\tau \hat{\beta} \) is the BLUE of \( l^\tau \beta \) for any \( l \in \mathcal{R}(Z) \).
Solution. (i) Condition $\mathcal{R}(Z^\tau) \subset \mathcal{R}(U^\tau)$ implies that there exists a matrix $B$ such that $Z = UB$. Then

$$\text{Var}(\varepsilon)Z = \text{Var}(\varepsilon)UB = \lambda UB = \lambda Z$$

and, thus,

$$Z(Z^\tau Z)^{-} Z^\tau \text{Var}(\varepsilon) = \lambda Z(Z^\tau Z)^{-} Z^\tau,$$

which is symmetric. Hence the result follows from the result in Exercise 38.

(ii) Let $\Lambda_\rho$ be the $(n-k) \times (n-k)$ matrix whose columns are the $n-k$ eigenvectors corresponding to the eigenvalue $\rho$. Then $Z^\tau U_i = 0$ for all $i$ implies that $\mathcal{R}(Z^\tau) \subset \mathcal{R}(\Lambda_\rho^\tau)$ and there exists a matrix $C$ such that $Z = \Lambda_\rho C$. Since

$$\text{Var}(\varepsilon)Z = \text{Var}(\varepsilon)\Lambda_\rho C = \rho \Lambda_\rho C = \rho Z,$$

we obtain that

$$Z(Z^\tau Z)^{-} Z^\tau \text{Var}(\varepsilon) = \rho Z(Z^\tau Z)^{-} Z^\tau,$$

which is symmetric. Hence the result follows from the result in Exercise 38.

Exercise 42 (#3.80). Consider the linear model $X = Z\beta + \varepsilon$, where $Z$ is a known $n \times p$ matrix, $\beta$ is a $p$-vector of unknown parameters, and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ with independent and identically distributed $\varepsilon_1, \ldots, \varepsilon_n$ having $E(\varepsilon_i) = 0$ and $\text{Var}(\varepsilon_i) = \sigma^2$. Let $Z_i$ be the $i$th row of $Z$, $\hat X_i = Z_i^\tau \hat \beta$; $\hat \beta$ be the LSE of $\beta$, and $h_i = Z_i^\tau (Z^\tau Z)^{-} Z_i$.

(i) Show that for any $\epsilon > 0$,

$$P(|\hat X_i - E\hat X_i| \geq \epsilon) \geq \min\{P(\varepsilon_i \geq \epsilon/h_i), P(\varepsilon_i \leq -\epsilon/h_i)\}.$$ 

(ii) Show that $\hat X_i - E\hat X_i \to_p 0$ if and only if $\lim_n h_i = 0$.

Solution. (i) For independent random variables $U$ and $Y$ and $\epsilon > 0$,

$$P(|U + Y| \geq \epsilon) \geq P(U \geq \epsilon)P(Y \geq 0) + P(U \leq -\epsilon)P(Y < 0) \geq \min\{P(U \geq \epsilon), P(U \leq -\epsilon)\}.$$ 

Using the result in the solution of Exercise 36,

$$\hat X_i - E\hat X_i = \sum_{j=1}^n h_{ij} (X_j - EX_j) = \sum_{j=1}^n h_{ij}\varepsilon_j = h_i\varepsilon_i + \sum_{j \neq i} h_{ij}\varepsilon_j.$$ 

Then the result follows by taking $U = h_i\varepsilon_i$ and $Y = \sum_{j \neq i} h_{ij}\varepsilon_j$.

(ii) If $\hat X_i - E\hat X_i \to_p 0$, then it follows from the result in (i) that

$$\lim_{n} \min\{P(\varepsilon_i \geq \epsilon/h_i), P(\varepsilon_i \leq -\epsilon/h_i)\} = 0,$$
which holds only if \( \lim n h_i = 0 \). Suppose now that \( \lim h_i = 0 \). From Exercise 36, \( \lim n \text{Var}(\hat{X}_i) = \lim n \sigma^2 h_i = 0 \). Therefore, \( \hat{X}_i - E\hat{X}_i \to_p 0 \).

**Exercise 43 (#3.81).** Let \( Z \) be an \( n \times p \) matrix, \( Z_i \) be the \( i \)th row of \( Z \), \( h_i = Z_i^\tau (Z^\tau Z)^{-} Z_i \), and \( \lambda_n \) be the largest eigenvalue of \((Z^\tau Z)^{-}\). Show that if \( \lim n \lambda_n = 0 \) and \( \lim n Z_n^\tau (Z^\tau Z)^{-} Z_n = 0 \), then \( \lim n \max_{1 \leq i \leq n} h_i = 0 \).

**Solution.** Since \((Z^\tau Z)^{-}\) depends on \( n \), we denote \((Z^\tau Z)^{-}\) by \( A_n \). Let \( i_n \) be the integer such that \( h_{i_n} = \max_{1 \leq i \leq n} h_i \). If \( \lim n i_n = \infty \), then

\[
\lim_n h_{i_n} = \lim_n Z_{i_n}^\tau A_n Z_{i_n} = \lim_n Z_{i_n}^\tau A_n Z_{i_n} = 0,
\]

where the inequality follows from \( i_n \leq n \) and, thus, \( A_{i_n} - A_n \) is nonnegative definite. If \( i_n \leq c \) for all \( n \), then

\[
\lim_n h_{i_n} = \lim_n Z_{i_n}^\tau A_n Z_{i_n} \leq \lim_n \max_{1 \leq i \leq c} \|Z_i\|^2 = 0.
\]

Therefore, for any subsequence \( \{j_n\} \subset \{i_n\} \) with \( \lim n j_n = a \in (0, \infty] \), \( \lim n h_{j_n} = 0 \). This shows that \( \lim n h_{i_n} = 0 \).

**Exercise 44 (#3.84).** Consider the one-way random effects model

\[
X_{ij} = \mu + A_i + e_{ij}, \quad j = 1, \ldots, n_i, i = 1, \ldots, m,
\]

where \( \mu \in \mathcal{R} \) is an unknown parameter, \( A_i \)'s are independent and identically distributed with mean 0 and variance \( \sigma^2_a \), \( e_{ij} \)'s are independent with mean 0, and \( A_i \)'s and \( e_{ij} \)'s are independent. Assume that \( \{n_i\} \) is bounded and \( E|e_{ij}|^{2+\delta} < \infty \) for some \( \delta > 0 \). Show that the LSE \( \hat{\mu} \) of \( \mu \) is asymptotically normal and derive an explicit form of \( \text{Var}(\hat{\mu}) \).

**Solution.** The LSE of \( \mu \) is \( \hat{\mu} = \hat{X}_n \), the average of \( X_{ij} \)'s. The model under consideration can be written as \( X = Z\mu + \varepsilon \) with \( Z = J_n, Z^\tau Z = n \), and

\[
\lim_n \max_{1 \leq i \leq n} Z_i^\tau (Z^\tau Z)^{-} Z_i = \lim_n \frac{1}{n} = 0.
\]

Since we also have \( E|e_{ij}|^{2+\delta} < \infty \) and \( \{n_i\} \) is bounded, by Theorem 3.12(i) in Shao (2003),

\[
\frac{\hat{\mu} - \mu}{\sqrt{\text{Var}(\hat{\mu})}} \to d N(0, 1),
\]

where \( \text{Var}(\hat{\mu}) = \text{Var}(\hat{X}_n) = n^{-2} \sum_{i=1}^m (n_i^2 \sigma^2_a + n_i \sigma^2) \).

**Exercise 45 (#3.85).** Suppose that

\[
X_i = \rho t_i + \varepsilon_i, \quad i = 1, \ldots, n,
\]

where \( \rho \in \mathcal{R} \) is an unknown parameter, \( t_i \)'s are known and in \((a, b)\), \( a \) and \( b \) are known positive constants, and \( \varepsilon_i \)'s are independent random variables
satisfying $E(\varepsilon_i) = 0$, $E|\varepsilon_i|^{2+\delta} < \infty$ for some $\delta > 0$, and $\text{Var}(\varepsilon_i) = \sigma^2 t_i$ with an unknown $\sigma^2 > 0$.

(i) Obtain the LSE of $\rho$.

(ii) Obtain the BLUE of $\rho$.

(iii) Show that both the LSE and BLUE are asymptotically normal and obtain the asymptotic relative efficiency of the BLUE with respect to the LSE.

**Solution.**

(i) The LSE of $\rho$ is

$$\hat{\rho} = \frac{\sum_{i=1}^n t_i X_i}{\sum_{i=1}^n t_i^2}.$$

(iii) Let $X = (X_1, ..., X_n)$ and $c = (c_1, ..., c_n)$. Consider minimizing

$$E(c^\top X - \rho)^2 = \sum_{i=1}^n t_i c_i^2$$

under the constraint $\sum_{i=1}^n c_i t_i = 1$ (to ensure unbiasedness), which yields $c_i = (\sum_{i=1}^n t_i)^{-1}$. Hence, the BLUE of $\rho$ is

$$\tilde{\rho} = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n t_i}.$$

(iii) The asymptotic normality of the LSE and BLUE follows directly from Lindeberg’s central limit theorem. Since

$$\text{Var}(\hat{\rho}) = \frac{\sigma^2 \sum_{i=1}^n t_i^3}{(\sum_{i=1}^n t_i^2)^2},$$

and

$$\text{Var}(\tilde{\rho}) = \frac{\sigma^2}{\sum_{i=1}^n t_i},$$

the asymptotic relative efficiency of the BLUE with respect to the LSE is

$$\frac{(\sum_{i=1}^n t_i^3)^2}{(\sum_{i=1}^n t_i^2)(\sum_{i=1}^n t_i)}.$$

**Exercise 46 (#3.87).** Suppose that $X = (X_1, ..., X_n)$ is a simple random sample without replacement from a finite population $P = \{y_1, ..., y_N\}$ with all $y_i \in \mathcal{R}$.

(i) Show that a necessary condition for $h(y_1, ..., y_N)$ to be estimable is that $h$ is symmetric in its $N$ arguments.

(ii) Find the UMVUE of $P(X_i \leq X_j)$, $i \neq j$.

(iii) Find the UMVUE of $\text{Cov}(X_i, X_j)$, $i \neq j$. 
Solution. (i) If \( h(y_1, \ldots, y_N) \) is estimable, then there exists a function \( u(x_1, \ldots, x_n) \) that is symmetric in its arguments and satisfies

\[
    h(y_1, \ldots, y_N) = E[u(X_1, \ldots, X_n)] = \frac{1}{\binom{N}{n}} \sum_{1 \leq i_1 < \cdots < i_n \leq N} u(y_{i_1}, \ldots, y_{i_n}).
\]

Hence, \( h \) is symmetric in its arguments.

(ii) From Watson-Royall’s theorem (e.g., Theorem 3.13 in Shao, 2003), the order statistics are complete and sufficient. Hence, for any estimable parameter, its UMVUE is the unbiased estimator \( g(X_1, \ldots, X_n) \) that is symmetric in its arguments. Thus, the UMVUE of \( P(X_i \leq X_j), i \neq j \), is

\[
    U_1 = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} I_{(-\infty, X_i]}(X_j) + I_{(-\infty, X_j]}(X_i) \cdot \frac{1}{2}.
\]

(iii) From the argument in part (ii) of the solution, the UMVUE of \( E(X_iX_j) \) when \( i \neq j \) is

\[
    U_1 = \frac{1}{nN} \sum_{i=1}^{N} X_iX_j.
\]

Let \( \bar{X} \) be the sample mean. Since

\[
    E(\bar{X}^2) = \frac{1}{nN} \sum_{i=1}^{N} y_i^2 + \frac{2(n-1)}{nN(N-1)} \sum_{1 \leq i < j \leq N} y_iy_j
\]

and

\[
    E \left( \frac{1}{n} \sum_{i=1}^{n} X_i^2 \right) = \frac{1}{N} \sum_{i=1}^{N} y_i^2,
\]

the UMVUE of \( 2 \sum_{1 \leq i < j \leq N} y_iy_j \) is

\[
    U_2 = \frac{nN(N-1)}{n-1} \left( \bar{X}^2 - \frac{1}{n^2} \sum_{i=1}^{n} X_i^2 \right).
\]

From

\[
    \text{Cov}(X_i, X_j) = E(X_iX_j) - \left( \frac{1}{N} \sum_{i=1}^{N} y_i \right)^2
\]

\[
    = E(X_iX_j) - \frac{1}{N^2} \sum_{i=1}^{N} y_i^2 - \frac{2}{N^2} \sum_{1 \leq i < j \leq N} y_iy_j,
\]

the UMVUE of \( \text{Cov}(X_i, X_j), i \neq j \), is

\[
    U_1 - \frac{1}{nN} \sum_{i=1}^{n} X_i^2 - \frac{U_2}{N^2}.
\]
Exercise 47 (\#3.100). Let \((X_1, \ldots, X_n)\) be a random sample from the normal distribution \(N(\mu, \sigma^2)\), where \(\mu \in \mathbb{R}\) and \(\sigma^2 > 0\). Consider the estimation of \(\vartheta = E[\Phi(a + bX_1)]\), where \(\Phi\) is the cumulative distribution function of \(N(0, 1)\) and \(a\) and \(b\) are known constants. Obtain an explicit form of a function \(g(\mu, \sigma^2) = \vartheta\) and the asymptotic mean squared error of \(\hat{\vartheta} = g(\bar{X}, S^2)\), where \(\bar{X}\) and \(S^2\) are the sample mean and variance.

Solution. Let \(Z\) be a random variable that has distribution \(N(0, 1)\) and is independent of \(X_1\). Define \(Y = Z - bX_1\). Then \(Y\) has distribution \(N(-b\mu, 1 + b^2\sigma^2)\) and \(E[\Phi(a + bX_1)] = E[\Phi(Z - bX_1)] = E[\Phi(Y)] = \Phi\left(\frac{a + b\mu}{\sqrt{1 + b^2\sigma^2}}\right)\).

Hence
\[
g(\mu, \sigma^2) = \Phi\left(\frac{a + b\mu}{\sqrt{1 + b^2\sigma^2}}\right).
\]

From Example 2.8 in Shao (2003),
\[
\sqrt{n}\left(\begin{array}{c}
\bar{X} - \mu \\
S^2 - \sigma^2
\end{array}\right) \rightarrow_d N_2\left(\left(\begin{array}{c}0 \\
0
\end{array}\right), \left(\begin{array}{cc}\sigma^2 & 0 \\
0 & 2\sigma^4
\end{array}\right)\right).
\]

Then, by the \(\delta\)-method,
\[
\sqrt{n}(\hat{\vartheta} - \vartheta) = \sqrt{n}[g(\bar{X}, S^2) - \vartheta] \rightarrow_d N(0, \kappa),
\]

where
\[
\kappa = \left[\frac{b^2\sigma^2}{1 + b^2\sigma^2} + \frac{(a + b\mu)^2b^4\sigma^2}{2(1 + b^2\sigma^2)}\right]\left[\Phi'\left(\frac{a + b\mu}{\sqrt{1 + b^2\sigma^2}}\right)\right]^2.
\]

The asymptotic mean squared error of \(\hat{\vartheta}\) is \(\kappa/n\).

Exercise 48 (\#3.103). Let \((X_1, \ldots, X_n)\) be a random sample from \(P\) in a parametric family. Obtain moment estimators of parameters in the following cases.

(i) \(P\) is the gamma distribution with shape parameter \(\alpha > 0\) and scale parameter \(\gamma > 0\).

(ii) \(P\) has Lebesgue density \(\theta^{-1} e^{-(x-a)/\theta} I_{(a, \infty)}(x), a \in \mathbb{R}, \theta > 0\).

(iii) \(P\) has Lebesgue density \(\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}I_{(0,1)}(x), \alpha > 0, \beta > 0\).

(iv) \(P\) is the log-normal distribution with parameter \((\mu, \sigma^2)\) (i.e., log \(X_1\).
has distribution \(N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma > 0\).

(v) \(P^n\) is the negative binomial distribution with discrete probability density 
\[
\frac{p^n(1-p)^{x-r}}{\binom{x-1}{r-1}}, \quad x = r, r+1, \ldots, \quad p \in (0, 1), \quad r = 1, 2, \ldots
\]

**Solution.** Let \(\mu_k = E(X_1^k)\) and \(\hat{\mu}_k = n^{-1}\sum_{i=1}^n X_i^k\).

(i) Note that \(\hat{\mu} = \alpha \gamma\) and \(\mu_2 - \mu_1^2 = \alpha \gamma^2\). Hence, the moment estimators are \(\hat{\gamma} = (\hat{\mu}_2 - \hat{\mu}_1^2)/\hat{\mu}_1\) and \(\hat{\alpha} = \hat{\mu}_1^2/(\hat{\mu}_2 - \hat{\mu}_1^2)\).

(ii) Note that \(\hat{\mu}_1 = \alpha + \theta\) and \(\mu_2 - \mu_1^2 = \theta^2\). Hence, the moment estimators are \(\hat{\theta} = \sqrt{\hat{\mu}_2 - \hat{\mu}_1^2}\) and \(\hat{\alpha} = \hat{\mu}_1 - \hat{\theta}\).

(iii) Note that \(\hat{\mu}_1 = \alpha/(\alpha + \beta)\) and \(\mu_2 = \alpha(\alpha + 1)/[(\alpha + \beta)(\alpha + \beta + 1)]\). Then \(1 + \beta/\alpha = \mu_1^{-1}\), which leads to \(\mu_2 = \mu_1(1 + \alpha^{-1})/\mu_1^{-1} + \alpha^{-1})\). Then the moment estimators are \(\hat{\alpha} = \hat{\mu}_1(\hat{\mu}_1 - \hat{\mu}_2)/(\hat{\mu}_2 - \hat{\mu}_1^2)\) and \(\hat{\beta} = (\hat{\mu}_1 - \hat{\mu}_2)(1 - \hat{\mu}_1)/(\hat{\mu}_2 - \hat{\mu}_1^2)\).

(iv) Note that \(\hat{\mu}_1 = e^{\mu + \sigma^2/2}\) and \(\mu_2 = e^{2\mu + 2\sigma^2}\). Then \(\mu_2/\mu_1 = e^{\sigma^2}\), i.e., \(\sigma^2 = \log(\mu_2/\mu_1^2)\). Then \(\mu = \log(\mu_1 + \sigma^2/2)\). Hence, the moment estimators are \(\hat{\sigma}^2 = \log(\hat{\mu}_2/\hat{\mu}_1^2)\) and \(\hat{\mu} = \hat{\mu}_1 - \frac{1}{2}\log(\hat{\mu}_2/\hat{\mu}_1^2)\).

(v) Note that \(\mu_1 = r/p\) and \(\mu_2 - \mu_1^2 = r(1-p)/p^2\). Then \(r = p\mu_1\) and \((\mu_2 - \mu_1^2)p = \mu_1(1-p)\). Hence, the moment estimators are \(\hat{\rho} = \hat{\mu}_1/(\hat{\mu}_2 - \hat{\mu}_1^2 + \hat{\mu}_1)\) and \(\hat{\sigma} = \hat{\mu}_1^2/(\hat{\mu}_2 - \hat{\mu}_1^2 + \hat{\mu}_1)\).

**Exercise 49 (#3.106).** In Exercise 11(i), find a moment estimator of \(\theta\) and derive its asymptotic distribution. In Exercise 11(ii), obtain a moment estimator of \(\theta^{-1}\) and its asymptotic relative efficiency with respect to the UMVUE of \(\theta^{-1}\).

**Solution.** (i) From Exercise 11(i),

\[
\mu_1 = EX_1 = P(Y_1 < 1) + \frac{1}{\theta} \int_1^\theta x dx = \frac{1}{\theta} + \frac{\theta^2 - 1}{2\theta} = \frac{\theta^2 + 1}{2\theta}.
\]

Let \(\bar{X}\) be the sample mean. Setting \(\bar{X} = (\theta^2 + 1)/(2\theta)\), we obtain that \(\theta^2 - 2\bar{X}\theta + 1 = 0\), which has solutions \(\bar{X} = \sqrt{\theta^2 - 1}\). Since \(\bar{X} \geq 1\), \(\bar{X} = \sqrt{\theta^2 - 1} < 1\). Since \(\theta \geq 1\), the moment estimator of \(\theta\) is \(\hat{\theta} = \bar{X} + \sqrt{\theta^2 - 1}\).

From the central limit theorem,

\[
\sqrt{n}(\bar{X} - \mu_1) \rightarrow_d N\left(0, \frac{\theta^3 + 2}{3\theta} - \frac{(\theta^2 + 2)^2}{4\theta^2}\right).
\]

By the \(\delta\)-method with \(g(x) = x + \sqrt{x^2 - 1}\),

\[
\sqrt{n}(\hat{\theta} - \theta) \rightarrow_d N\left(0, \left(1 + \frac{\theta}{\sqrt{\theta^2 - 1}}\right)^2 \left[\frac{\theta^3 + 2}{3\theta} - \frac{(\theta^2 + 2)^2}{4\theta^2}\right]\right).
\]

(ii) From Exercise 11(ii),

\[
\mu_1 = EX_1 = \frac{1}{\theta} \int_0^1 x dx + P(Y_1 > 1) = \frac{1}{2\theta} + 1 - \frac{1}{\theta} = 1 - \frac{1}{2\theta}.
\]
Hence the moment estimator of \( \theta^{-1} \) is \( 2(1 - \bar{X}) \). From the central limit theorem,
\[
\sqrt{n}(\bar{X} - \mu_1) \rightarrow_d N \left( 0, \frac{1}{3\theta^2} - \frac{1}{4\theta^2} \right).
\]
By the \( \delta \)-method with \( g(x) = 2(1 - x) \),
\[
\sqrt{n}[2(1 - \bar{X}) - \theta^{-1}] \rightarrow_d N \left( 0, \frac{4}{3\theta^2} - \frac{1}{\theta^2} \right).
\]
Let \( R_i = 0 \) if \( X_i = 1 \) and \( R_i = 1 \) if \( X_i \neq 1 \). From the solution of Exercise 11(ii), the UMVUE of \( \theta^{-1} \) is \( \bar{R} = n^{-1} \sum_{i=1}^{n} R_i \). By the central limit theorem,
\[
\sqrt{n}(\bar{R} - \theta^{-1}) \rightarrow_d N \left( 0, \frac{1}{\theta} - \frac{1}{\theta^2} \right).
\]
Hence, the asymptotic relative efficiency of \( 2(1 - \bar{X}) \) with respect to \( \bar{R} \) is equal to \( (\theta^{-1}) / (\frac{4}{3} \theta - 1) \).

**Exercise 50 (#3.107).** Let \((X_1, ..., X_n)\) be a random sample from a population having the Lebesgue density \( f_{\alpha, \beta}(x) = \alpha \beta x^{\alpha - 1} I_{(0, \beta)}(x) \), where \( \alpha > 0 \) and \( \beta > 0 \) are unknown. Obtain a moment estimator of \( \theta = (\alpha, \beta) \) and its asymptotic distribution.

**Solution.** Let \( \mu_j = E X_1^j \). Note that
\[
\mu_1 = \frac{\alpha}{\beta^\alpha} \int_0^\beta x^\alpha dx = \frac{\alpha \beta}{\alpha + 1}
\]
and
\[
\mu_2 = \frac{\alpha}{\beta^\alpha} \int_0^\beta x^{\alpha + 1} dx = \frac{\alpha \beta^2}{\alpha + 2}.
\]
Then \( \beta = (1 + \frac{1}{\alpha}) \mu_1 \) and
\[
\left( 1 + \frac{1}{\alpha} \right)^2 \mu_1^2 = \left( 1 + \frac{2}{\alpha} \right) \mu_2,
\]
which leads to
\[
\frac{1}{\alpha} = \frac{\mu_2 - \mu_1^2}{\mu_1^2 - \mu_1 \mu_2} \pm \sqrt{\frac{\mu_2^2 - \mu_1^2}{\mu_1^2} - \mu_1 \mu_2}
\]
Since \( \alpha > 0 \), we obtain the moment estimators
\[
\hat{\alpha} = \frac{\hat{\mu}_1^2}{\hat{\mu}_2 - \hat{\mu}_1^2 + \sqrt{\hat{\mu}_2^2 - \hat{\mu}_1 \hat{\mu}_2}}
\]
and
\[
\hat{\beta} = \frac{\hat{\mu}_2 + \sqrt{\hat{\mu}_2^2 - \hat{\mu}_1 \hat{\mu}_2}}{\hat{\mu}_1},
\]
where \( \mu_j = \frac{1}{n} \sum_{i=1}^{n} X_i^j \). Let \( \gamma = (\mu_1, \mu_2) \) and \( \hat{\gamma} = (\hat{\mu}_1, \hat{\mu}_2) \). From the central limit theorem,

\[
\sqrt{n}(\hat{\gamma} - \gamma) \rightarrow_d N(0, \Sigma),
\]

where

\[
\Sigma = \begin{pmatrix}
\mu_2 - \mu_1^2 & \mu_3 - \mu_1 \mu_2 \\
\mu_3 - \mu_1 \mu_2 & \mu_4 - \mu_2^2
\end{pmatrix}.
\]

Let \( \alpha(x, y) = \frac{x^2}{y - x^2 + \sqrt{y^2 - xy}} \) and \( \beta(x, y) = \frac{y + \sqrt{y^2 - xy}}{x} \).

Then

\[
\frac{\partial (\alpha, \beta)}{\partial (x, y)} = \begin{pmatrix}
-\frac{2x}{y - x^2 + \sqrt{y^2 - xy}} + \frac{x^2(y + x/\sqrt{y^2 - xy})}{2(y - x^2 + \sqrt{y^2 - xy})^2} & -\frac{x^2(1 + (y - x/2)/\sqrt{y^2 - xy})}{(y - x^2 + \sqrt{y^2 - xy})^2} \\
-\frac{y}{2x\sqrt{y^2 - xy}} - \frac{y + \sqrt{y^2 - xy}}{x^2} & \frac{1}{x} + \frac{2y - x}{2x\sqrt{y^2 - xy}}
\end{pmatrix}.
\]

Let \( \hat{\theta} = (\hat{\alpha}, \hat{\beta}) \) and \( \Lambda = \frac{\partial (\alpha, \beta)}{\partial (x, y)}|_{x=\mu_1, y=\mu_2} \). Then, by the \( \delta \)-method,

\[
\sqrt{n}(\hat{\theta} - \theta) \rightarrow_d N(0, \Lambda \Sigma \Lambda^T).
\]

**Exercise 51 (#3.108).** Let \((X_1, ..., X_n)\) be a random sample from the following discrete distribution:

\[
P(X_1 = 1) = \frac{2(1 - \theta)}{2 - \theta}, \quad P(X_1 = 2) = \frac{\theta}{2 - \theta},
\]

where \( \theta \in (0, 1) \) is unknown. Obtain a moment estimator of \( \theta \) and its asymptotic distribution.

**Solution.** Note that

\[
EX_1 = \frac{2(1 - \theta)}{2 - \theta} + \frac{2\theta}{2 - \theta} = \frac{2}{2 - \theta}.
\]

Hence, a moment estimator of \( \theta \) is \( \hat{\theta} = 2(1 - \bar{X}^{-1}) \), where \( \bar{X} \) is the sample mean. Note that

\[
\text{Var}(X_1) = \frac{2(1 - \theta)}{2 - \theta} + \frac{4\theta}{2 - \theta} - \frac{4}{(2 - \theta)^2} = \frac{4\theta - 2\theta^2 - 4}{(2 - \theta)^2}.
\]

By the central limit theorem and \( \delta \)-method,

\[
\sqrt{n}(\hat{\theta} - \theta) \rightarrow_d N \left( 0, \frac{(2 - \theta)^2(2\theta - \theta^2 - 2)}{2} \right).
\]
Exercise 52 (#3.110). Let \((X_1, ..., X_n)\) be a random sample from a population having the Lebesgue density
\[
f_{\theta_1, \theta_2}(x) = \begin{cases} 
(\theta_1 + \theta_2)^{-1}e^{-x/\theta_1} & x > 0 \\
(\theta_1 + \theta_2)^{-1}e^{x/\theta_2} & x \leq 0,
\end{cases}
\]
where \(\theta_1 > 0\) and \(\theta_2 > 0\) are unknown. Obtain a moment estimator of \((\theta_1, \theta_2)\) and its asymptotic distribution.

Solution. Let \(\mu_j = E X_j^j\) and \(\hat{\mu}_j = \sum_{i=1}^n X_i^j\). Note that
\[
\mu_1 = \frac{1}{\theta_1 + \theta_2} \left( \int_{-\infty}^0 x e^{x/\theta_2} dx + \int_0^\infty x e^{-x/\theta_1} dx \right) = \theta_1 - \theta_2
\]
and
\[
\mu_2 = \frac{1}{\theta_1 + \theta_2} \left( \int_{-\infty}^0 x^2 e^{x/\theta_2} dx + \int_0^\infty x^2 e^{-x/\theta_1} dx \right) = 2(\theta_1^2 + \theta_2^2 - \theta_1 \theta_2).
\]

Then, \(\mu_2 - \mu_1^2 = \theta_1^2 + \theta_2^2\). Since \(\theta_1 = \mu_1 + \theta_2\), we obtain that
\[
2\theta_2^2 + 2\mu_1 \theta_2 + 2\mu_2^2 - \mu_2 = 0,
\]
which has solutions
\[
-\mu_1 \pm \sqrt{2\mu_2 - 3\mu_1^2}.
\]
Since \(\theta_2 > 0\), the moment estimators are
\[
\hat{\theta}_2 = -\hat{\mu}_1 + \sqrt{2\hat{\mu}_2 - 3\hat{\mu}_1^2}
\]
and
\[
\hat{\theta}_1 = \frac{\hat{\mu}_1 + \sqrt{2\hat{\mu}_2 - 3\hat{\mu}_1^2}}{2}.
\]

Let \(g(x, y) = (\sqrt{2y - 3x} - x)/2\) and \(h(x, y) = (\sqrt{2y - 3x} + x)/2\). Then
\[
\frac{\partial(g, h)}{\partial(x, y)} = \begin{pmatrix} -\frac{1}{2} - \frac{3}{4\sqrt{2y - 3x}} & \frac{1}{2\sqrt{2y - 3x}} \\
\frac{3}{4\sqrt{2y - 3x}} & \frac{1}{2}\end{pmatrix}.
\]

Let \(\gamma = (\mu_1, \mu_2)\), \(\hat{\gamma} = (\hat{\mu}_1, \hat{\mu}_2)\), \(\theta = (\theta_1, \theta_2)\), and \(\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)\). From the central limit theorem,
\[
\sqrt{n}(\hat{\gamma} - \gamma) \to_d N(0, \Sigma),
\]
where \(\Sigma\) is as defined in the solution of Exercise 50. By the \(\delta\) method,
\[
\sqrt{n}(\hat{\theta} - \theta) \to_d N(0, \Lambda\Sigma\Lambda^\tau),
\]
Chapter 3. Unbiased Estimation

where \( \Lambda = \frac{\partial(g,h)}{\partial(x,y)}|_{x=\mu_1, y=\mu_2}. \)

Exercise 53 (#3.111). Let \((X_1, ..., X_n)\) be a random sample from \(P\) with discrete probability density \(f_{\theta,j}\), where \(\theta \in (0,1), j = 1, 2, f_{\theta,1}\) is the Poisson distribution with mean \(\theta\), and \(f_{\theta,2}\) is the binomial distribution with size 1 and probability \(\theta\). Let \(h_k(\theta, j) = E_{\theta,j}(X_i^k), k = 1, 2,\) where \(E_{\theta,j}\) is the expectation is with respect to \(f_{\theta,j}\). Show that
\[
\lim_{n} P(\hat{\mu}_k = h_k(\theta, j) \text{ has a solution}) = 0
\]
when \(X_i\)'s are from the Poisson distribution, where \(\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^{n} X_i^k, k = 1, 2.\)

Solution. Note that \(h_1(\theta, 1) = h_1(\theta, 2) = \theta\). Hence \(h_1(\theta, j) = \hat{\mu}_1\) has a solution \(\theta = \hat{\mu}_1\). Assume that \(X_i\)'s are from the Poisson distribution with mean \(\theta\). Then \(\hat{\mu}_2 \rightarrow_p \theta + \theta^2.\) Since \(h_2(\theta, 1) = \theta - \theta^2,\)
\[
\lim_{n} P(\hat{\mu}_2 = h_2(\theta, 1)) = 0.
\]
It remains to show that
\[
\lim_{n} P(\hat{\mu}_2 = h_2(\theta, 2)) = 0.
\]
Since \(h_2(\theta, 2) = \theta + \theta^2\) and \(\theta = \hat{\mu}_1\) is a solution to the equation \(h_1(\theta, 1) = h_1(\theta, 2) = \theta,\) it suffices to show that
\[
\lim_{n} P(\hat{\mu}_2 = \hat{\mu}_1^2) = 0.
\]
Let \(\gamma = (\mu_1, \mu_2)\) and \(\hat{\gamma} = (\hat{\mu}_1, \hat{\mu}_2).\) From the central limit theorem,
\[
\sqrt{n} (\hat{\gamma} - \gamma) \to_d N(0, \Sigma),
\]
where \(\Sigma\) is as defined in the solution of Exercise 50. Then, we only need to show that \(\Sigma\) is not singular. When \(X_1\) has the Poisson distribution with mean \(\theta,\) a direct calculation shows that \(\mu_1 = \theta, \mu_2 = \theta + \theta^2, \mu_3 = \theta + 3\theta^2 + \theta^3,\)
and \(\mu_4 = \theta + 7\theta^2 + 6\theta^3 + \theta^4.\) Hence,
\[
\Sigma = \begin{pmatrix}
\theta & \theta + 2\theta^2 \\
\theta + 2\theta^2 & \theta + 6\theta^2 + 4\theta^3
\end{pmatrix}.
\]
The determinant of \(\Sigma\) is equal to
\[
\theta^2 + 6\theta^3 + 4\theta^4 - (\theta + 2\theta^2)^2 = 2\theta^3 > 0.
\]
Hence \(\Sigma\) is not singular.

Exercise 54 (#3.115). Let \((X_1, ..., X_n)\) be a random sample from a population on \(\mathcal{R}\) having a finite sixth moment. Consider the estimation of \(\mu^3,\)
where $\mu = EX_1$. Let $\bar{X}$ be the sample mean. When $\mu = 0$, find the asymptotic relative efficiency of the V-statistic $\bar{X}^3$ with respect to the U-statistic $U_n = \frac{1}{\binom{n}{3}} \sum_{1 \leq i < j < k \leq n} X_i X_j X_k$.

**Solution.** We adopt the notation in Exercise 25. Note that $U_n$ is a U-statistic with $\zeta_1 = \zeta_2 = 0$, since $\mu = 0$. The order of the kernel of $U_n$ is 3. Hence, by Exercise 25(iii),

$$\text{Var}(U_n) = \frac{6\zeta_3}{n^3} + O\left(\frac{1}{n^4}\right),$$

where $\zeta_3 = \text{Var}(X_1 X_2 X_3) = E(X_1^2 X_2^2 X_3^2) = \sigma^6$ and $\sigma^2 = EX_1^2 = \text{Var}(X_1)$.

The asymptotic mean squared error of $U_n$ is then $6\sigma^6/n^3$.

From the central limit theorem and $\mu = 0$, $\sqrt{n} \bar{X} \rightarrow_d N(0, \sigma^2)$. Then $n^{3/2} \bar{X}^3 / \sigma^3 \rightarrow_d Z^3$, where $Z$ is a random variable having distribution $N(0, 1)$. Then the asymptotic mean square error of $\bar{X}^3$ is $\sigma^6EZ^6/n^3$. Note that $EZ^6 = 15$. Hence, the asymptotic relative efficiency of $\bar{X}^3$ with respect to $U_n$ is $6/15 = 2/5$. $\blacksquare$