1.5.10: In a Bernoulli sequence of trials with success probability $p$, let $X + m$ be the number of trials required to achieve $m$ successes.

(a) Show that the distribution of $X$, the negative binomial distribution, is as given in Table 5.

**Solution:** $X + m$ be the number of trials required to achieve $m$ successes. The $(X + m)$th is a success, $m - 1$ of the previous $X + m - 1$ trials are successes. Then

$$P(X = x) = p(x + m - 1 \choose m - 1) \cdot p^{m-1}(1-p)^x \cdot I_{\{0,1,\ldots\}}(x)$$

(b) Verify that the negative binomial probabilities add up to 1 by expanding $(\frac{1}{p} - \frac{q}{p})^{-m} = p^m(1-q)^{-m}$

**Solution:**

$$1 = p^m(1-q)^{-m} = p^m \sum_{x=0}^{\infty} \left( -m \atop x \right) (-q)^x$$

$$= p^m \sum_{x=0}^{\infty} \left( m + x - 1 \atop x \right) (-1)^x (-q)^x$$

$$= \sum_{x=0}^{\infty} \left( m + x - 1 \atop x \right) p^m q^x = \sum_{x=0}^{\infty} P(X = x)$$

(c) Show that the distribution of (a) constitute a one-parameter exponential family.

**Solution:**

$$P(X = x) = \left( x + m - 1 \atop m - 1 \right) p^m(1-p)^x I_{\{0,1,\ldots\}}(x)$$

$$= \left( x + m - 1 \atop m - 1 \right) I_{\{0,1,\ldots\}}(x) \exp[x \log(1-p) - (-m \log p)]$$
so the distribution constitute a one-parameter exponential family.

(d) Show that the moment generating function of $X$ is $M_X(u) = p^m/(1 - q e^u)^m$.

**Solution:** From part (c) we know $\eta = \log(1 - p) A(\eta) = -m \log p = -m \log(1 - e^\eta)$. By Theorem 1.5.10,

$$M_X(u) = e^{-m \log(1 - e^{\eta + u})}/e^{-m \log(1 - e^\eta)} = (1 - e^{\eta + u})^{-m}/(1 - e^\eta)^{-m}$$
$$= (1 - e^{\log(1-p)})^m/(1 - e^{\log(1-p)+u})^m$$
$$= p^m/(1 - q e^u)^m$$

(e) Show that $E(X) = mq/p$ and $\text{var}(X) = mq/p^2$

**Solution:** $M'_X(u) = p^m(-m)(1 - q e^u)^{-m-1}(-q e^u)$. $E(X) = M'_X(0) = p^m(-m)(1 - q)^{-m-1}(-q) = mq/p$.

$$M''_X(u) = mq p^m[e^u(1 - q e^u)^{-m-1} + e^u(-m - 1)(1 - q e^u)^{-m-2}(-q e^u)].$$
$$E(X) = M''_X(0) = mq p^m[(1 - q)^{-m-1} + (-m - 1)(1 - q)^{-m-2}(-q)] = \frac{mq}{p^2}(m - mp + 1).$$
$$\text{var}(X) = E(X^2) - (E(X))^2 = \frac{mq}{p^2}(m - mp + 1) - \frac{m^2 q^2}{p^2} = \frac{mq}{p^2}.$$

(f) Show that the first four cumulants of $X$ are $k^1 = mq/p$, $k^2 = mq/p^2$, $k^3 = mq(1 + q)/p^3$, $k^4 = mq(1 + 4q + q^2)/p^4$. 

2
Solution:

\[
K_X(u) = A(\eta + u) - A\eta = -m \log(1 - e^{\eta + u}) + m \log(1 - e^{\eta}) = -m \log(1 - q e^{u}) + m \log p
\]

\[
K_X'(u) = \frac{m}{1 - q e^{u}} = \frac{m q e^{u}}{1 - q e^{u}}
\]

\[
k_1 = K_X'(0) = \frac{m q}{p}
\]

\[
K_X''(u) = \frac{m q e^{u}}{(1 - q e^{u})^2}
\]

\[
k_2 = K_X''(0) = \frac{m q}{p^2}
\]

\[
K_X'''(u) = \frac{-m q e^{u}(1 + q e^{u})}{(q e^{u} - 1)^3}
\]

\[
k_3 = K_X'''(0) = \frac{m q(1 + q)}{p^3}
\]

\[
K_X''''(u) = \frac{m q e^{u}(1 + 4 q e^{u} + q^2 e^{2u})}{(q e^{u} - 1)^4}
\]

\[
k_4 = K_X''''(0) = \frac{m q(1 + 4 q + q^2)}{p^4}
\]
1.5.19: Using lemma 5.15:

(a) Derive the form of the identity for \( X \sim \Gamma(a, b) \) and use it to verify the moments of (5.44).

**Solution:**

\[
f_X(x) = \frac{1}{\Gamma(a)b^a} x^{a-1} e^{-x/b} \mathbb{1}_{x > 0} = x^{-a} \exp \left\{ a \log(x) - x/b - \log(\Gamma(a)b^a) \right\}
\]

Note that this distribution belongs to the exponential family, with \( T_1(x) = \log(x), T_2(x) = x, \eta_1 = a, \eta_2 = -1/b \) and \( h(x) = x^{-1} \mathbb{1}_{x > 0} = x^{-1} \). Also, \( T'_1(x) = x^{-1}, T'_2(x) = 1 \) and \( h'(x) = -x^{-2} \) and note that \( x^{-1} \exp \left\{ a \log(x) - x/b \right\} \to 0 \) as \( x \to 0 \). By Lemma 5.15 (Stein’s identity), for any differentiable function \( g \) with \( E|g'(X)| < \infty \),

\[
E \left\{ \left[ \frac{h'(x)}{h(x)} + \sum_{i=1}^{2} \eta_i T'_i(x) \right] g(X) \right\} = -Eg'(X) \tag{1}
\]

Note that

\[
\frac{h'(X)}{h(X)} = \frac{-X^{-2}}{X^{-1}} = -X^{-1}
\]

so (1) reduces to

\[
E \left\{ \left[ a - 1 - \frac{X}{b} \right] g(X) \right\} = -Eg'(X) \tag{2}
\]

To find \( EX \) take \( g(X) = X \), which is clearly differentiable and has finite expectation, then (2) reduces to

\[
E \left[ a - 1 - \frac{X}{b} \right] = -1 \\
EX = ab
\]

To find \( EX^2 \) take \( g(X) = X^2 \), which is clearly differentiable and has finite expectation, (2) reduces to
\[
E \left[ (a - 1)X - \frac{X^2}{b} \right] = -E(2X)
\]

\[
EX^2 = ab^2 + a^2b^2
\]

Therefore,

\[
VarX = EX^2 - (EX)^2 = ab^2 + a^2b^2 - (ab)^2 = ab^2
\]

To find \(EX^3\) take \(g(X) = X^3\), which is clearly differentiable and has finite expectation, then (2) reduces to

\[
E \left[ (a - 1)X^2 - \frac{X^3}{b} \right] = -E(3X^2)
\]

Solving for \(EX^3\) we get,

\[
EX^3 = b^3a(a + 2) + b^3a^2(a + 2) = 2ab^3 + a^3b^3 + 3a^2b^3
\]

Therefore,

\[
E(X - ab)^3 = E(X^3 - 3X^2ab + 3Xa^2b^2 - a^3b^3)
= (2ab^3 + a^3b^3 + 3a^2b^3) - 3(ab^2 + a^2b^2)ab + 3(ab)a^2b^2 - a^3b^3
= 2ab^3
\]

To find \(EX^4\), take \(g(X) = X^4\), which is clearly differentiable and has finite expectation, then (2) reduces to

\[
E \left[ (a - 1)X^3 - \frac{X^4}{b} \right] = -E(4X^3)
\]

Solving for \(EX^4\) we get,

\[
EX^4 = b(a + 3)(2ab^3 + a^3b^3 + 3a^2b^3)
= 6ab^4 + 11a^2b^4 + 6a^3b^4 + a^4b^4
\]
Therefore,

\[ E(X - ab)^4 = E(X^4 - 4X^3ab + 6X^2a^2b^2 - 4Xa^3b^3 + a^4b^4) \]

\[ = (6ab^4 + 11a^2b^4 + 6a^3b^4 + a^4b^4) - 4(2ab^3 + a^3b^3 + 3a^2b^3)ab \]

\[ + 6(ab^2 + a^2b^2)a^2b^2 - 4(ab)a^3b^3 + a^4b^4 \]

\[ = (3a^2 + 6a)b^4 \]

(b) Derive the form of the identity for \( X \sim \text{Beta}(a, b) \) and use it to verify that \( EX = a/(a + b) \) and \( Var(X) = ab/(a + b)^2(a + b + 1) \)

**Solution:**

\[ f_X(x) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1 - x)^{b-1} I_{(0,1)}(x) \]

\[ = I_{(0,1)}(x)x^{-1}(1 - x)^{-1}exp\{alog(x) + blog(1 - x) + log(\frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)})\} \]

Note that this distribution belongs to the exponential family, with \( T_1(x) = log(x), T_2(x) = log(1 - x), \eta_1 = a, \eta_2 = b \) and \( h(x) = x^{-1}(1 - x)^{-1}I_{(0,1)}(x) \).

Also, \( T_1'(x) = x^{-1}, T_2'(x) = (x - 1)^{-1} \) and \( h'(x) = -(x(1 - x))^{-2}(1 - x)^{-2} \).

Note that \( x^{-1}(1 - x)^{-1}exp\{alog(x) + blog(1 - x) = exp\{(a - 1)log(x) + (b - 1)log(1 - x) \} \to 0 \) as \( x \to 0 \) or \( x \to 1 \), and note that

\[ \frac{h'(X)}{h(X)} = -\frac{1}{X} - \frac{1}{(X - 1)} \]

so by Lemma 5.15 (Stein’s identity), for any differentiable function \( g \) with \( E|g'(X)| < \infty \), (1) holds and reduces to

\[ E \left\{ \left[ \frac{a - 1}{X} + \frac{b - 1}{X - 1} \right] g(X) \right\} = -Eg'(X) \]

(3)

To find \( EX \) take \( g(X) = X(X - 1) \) which is clearly differentiable and has finite expectation, (3) reduces to
\[ E[(a - 1)(X - 1) + (b - 1)X] = -E(2X - 1) \]
\[ EX(a - 1 + b - 1 + 2) = a \]
\[ EX = \frac{a}{a + b} \]

To find \(EX^2\) take \(g(X) = X^2(X - 1)\) which is clearly differentiable and has finite expectation, (3) reduces to

\[ E[(a - 1)(X - 1)X + (b - 1)X^2] = -E(3X^2 - 2X) \]
\[ EX^2 = \frac{a(a + 1)}{(a + b)(a + b + 1)} \]

Therefore,

\[ Var X = EX^2 - (EX)^2 = \frac{a(a + 1)}{(a + b)(a + b + 1)} - \left(\frac{a}{a + b}\right)^2 \]
\[ = \frac{ab}{(a + b)^2(a + b + 1)} \]
1.6.7: Let $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ are independently distributed according to $N(\xi, \sigma^2)$ and $N(\eta, \tau^2)$, respectively. Find the minimal sufficient statistics for these cases:

(a) $\xi, \eta, \sigma, \tau$ are arbitrary: $-\infty < \xi, \eta < +\infty, 0 < \sigma, \tau$.

Solution: From Casella and Berger’s book: A statistics $T(X)$ is minimal sufficient if

$$
\frac{f(X)}{f(X')} = \frac{\prod_{i=1}^m (2\pi \sigma^2)^{-1/2} \exp\left(\frac{-\left(x_i - \xi\right)^2}{2\sigma^2}\right) \prod_{i=1}^n (2\pi \tau^2)^{-1/2} \exp\left(\frac{-\left(y_i - \eta\right)^2}{2\tau^2}\right)}{\prod_{i=1}^m (2\pi \sigma^2)^{-1/2} \exp\left(\frac{-\left(x_i' - \xi\right)^2}{2\sigma^2}\right) \prod_{i=1}^n (2\pi \tau^2)^{-1/2} \exp\left(\frac{-\left(y_i' - \eta\right)^2}{2\tau^2}\right)}
$$

This is constant in $\theta \iff \sum x_i = \sum x_i', \sum y_i = \sum y_i'$.

(b) $\sigma = \tau$ and $\xi, \eta, \sigma$ are arbitrary.

Solution: If $\sigma = \tau$ then

$$
(*) = \exp \left[ \frac{1}{2\sigma^2} \left( \sum x_i'^2 - \sum x_i^2 \right) + \frac{1}{2\tau^2} \left( \sum y_i'^2 - \sum y_i^2 \right) \right] + \frac{\xi}{\sigma^2} \left( \sum x_i - \sum x_i' \right) + \frac{\eta}{\tau^2} \left( \sum y_i - \sum y_i' \right)
$$

This is constant in $\theta \iff \sum x_i + \sum y_i = \sum x_i' + \sum y_i'$.

(c) $\xi = \eta$ and $\xi, \sigma, \tau$ are arbitrary.
Solution: If $\xi = \eta$ then

\[
(*) = \exp \left[ \frac{1}{2\sigma^2} \left( \sum x_i^2 - \sum x_i''^2 \right) + \frac{1}{2\tau^2} \left( \sum y_i^2 - \sum y_i''^2 \right) + \frac{\xi}{\sigma^2} \left( \sum x_i - \sum x_i'' \right) + \frac{\xi}{\tau^2} \left( \sum y_i - \sum y_i'' \right) \right]
\]

This is constant in $\theta \iff \sum x_i^2 = \sum x_i''^2, \sum y_i^2 = \sum y_i''^2, \sum x_i = \sum x_i'', \sum y_i = \sum y_i'$. Therefor $(\sum x_i, \sum x_i^2, \sum y_i, \sum y_i^2)$ is minimal sufficient statistics.
1.6.16: Let $X_1, X_2, \ldots, X_n$ be iid according to a distribution from a family $\mathcal{P}$. Show that $T$ is minimal sufficient in the following cases:

(a) $\mathcal{P} = \{U(0, \theta), \theta > 0\}$; $T = X_{(n)}$

Solution: Let $x$ and $y$ be any two sample points. Now,

$$
p_\theta(x) = \frac{\theta^{-n} \prod_{i=1}^{n} I_{(0, \theta)}(x_i)}{\theta^{-n} \prod_{i=1}^{n} I_{(0, \theta)}(y_i)} = \frac{I_{(x_{(n)}, \infty)}(\theta)}{I_{(y_{(n)}, \infty)}(\theta)}
$$

Note that the numerator and denominator are positive for the same values of $\theta$ and the ratio is a constant function of $\theta$ if and only if $T(x) = x_{(n)} = y_{(n)} = T(y)$. It follows by Theorem 6.1.4 of Casella and Berger’s book that $T = X_{(n)}$ is minimal sufficient for $\theta$.

(b) $\mathcal{P} = \{U(\theta_1, \theta_2), -\infty < \theta_1 < \theta_2 < \infty\}$; $T = (X_{(1)}, X_{(n)})$.

Solution: Let $x$ and $y$ be any two sample points and let $\theta = (\theta_1, \theta_2)$. Now,

$$
p_\theta(x) = \frac{(\theta_2 - \theta_1)^{-n} \prod_{i=1}^{n} I_{(\theta_1, \theta_2)}(x_i)}{(\theta_2 - \theta_1)^{-n} \prod_{i=1}^{n} I_{(\theta_1, \theta_2)}(y_i)} = \frac{I_{(-\infty, x_{(1)})}(\theta_1)I_{(x_{(n)}, \infty)}(\theta_2)}{I_{(-\infty, y_{(1)})}(\theta_1)I_{(y_{(n)}, \infty)}(\theta_2)}
$$

Note that the numerator and denominator are positive for the same values of $\theta$ and the ratio is a constant function of $\theta$ if and only if $x_{(1)} = y_{(1)}$ and $x_{(n)} = y_{(n)}$ that is, if and only if $T(x) = (x_{(1)}, x_{(n)}) = (y_{(1)}, y_{(n)}) = T(y)$. It follows by Theorem 6.1.4 of Casella and Berger’s book that $T = (X_{(1)}, X_{(n)})$ is minimal sufficient for $\theta = (\theta_1, \theta_2)$. 

10
(c) $P = \{U(\theta - \frac{1}{2}, \theta + \frac{1}{2}), -\infty < \theta < \infty\}; T = (X_1, X_n)$.

**Solution:** Let $x$ and $y$ be any two sample points. Now,

$$\frac{p_\theta(x)}{p_\theta(y)} = \frac{\prod_{i=1}^{n} I_{(\theta - \frac{1}{2}, \theta + \frac{1}{2})}(x_i)}{\prod_{i=1}^{n} I_{(\theta - \frac{1}{2}, \theta + \frac{1}{2})}(y_i)} = \frac{I_{(x(1)-\frac{1}{2}, x(1)+\frac{1}{2})}(\theta)}{I_{(y(n)-\frac{1}{2}, y(n)+\frac{1}{2})}(\theta)}$$

Note that the numerator and denominator are positive for the same values of $\theta$ and the ratio is a constant function of $\theta$ if and only if $x(1) + \frac{1}{2} = y(1) + \frac{1}{2}$ and $x(n) - \frac{1}{2} = y(n) - \frac{1}{2}$ if and only if $x(1) = y(1)$ and $x(n) = y(n)$, if and only if $T(x) = (x(1), x(n)) = (y(1), y(n)) = T(y)$. It follows by Theorem 6.1.4 of Casella and Berger's book that $T = (X_1, X_n)$ is minimal sufficient for $\theta$. 
1.6.33: Let $X_1, X_2, \ldots, X_n$ be iid each with density $f(x)$ (with respect to Lebesgue measure) which is unknown. Show that the order statistics are complete.

Solution: Let $\mathcal{P}_1$ be the family of distributions having Lebesgue p.d.f.’s. Let $V = (V_1, \ldots, V_n)$ where $V_k = \sum_{i=1}^n X_i^k, k = 1, \ldots, n.$ and $\mathcal{P}_0$ be the exponential family with density,

$$C(\theta_1, \ldots, \theta_n)\exp\left\{ -\sum_{k=1}^n (\theta_k \sum_{i=1}^n x_i^k) \right\} \exp\left\{ -\sum_{i=1}^n x_i^{2n} \right\} = g_0(V) h(X) \quad (1)$$

Clearly, $\mathcal{P}_0 \subset \mathcal{P}_1$ and $\mathcal{P}_0$ is an exponential family of full rank. By Theorem 6.22, $V(X)$ is complete and sufficient for $\mathcal{P}_0$. Note that every null set of $\mathcal{P}_0$ is also a null set of $\mathcal{P}_1$. To see this let $N$ be null set of $\mathcal{P}_0$, then $P(N) = 0$ for all $P \in \mathcal{P}_0$. Note that since (1) is a positive on $N$ then

$$0 = P(N) = \int_N C(\theta_1, \ldots, \theta_n)\exp\left\{ -\sum_{k=1}^n (\theta_k \sum_{i=1}^n x_i^k) \right\} \exp\left\{ -\sum_{i=1}^n x_i^{2n} \right\} d\mu = \mu(N) = 0$$

Let $P^*$ be any distribution in $\mathcal{P}_1$ then since $\mu(N) = 0$ and $P^*$ is absolutely continuous with respect to $\mu$ with Radon-Nikodym derivative $p$, we get

$$P^*(N) = \int_N p d\mu = 0 \quad \text{for all } P^* \in \mathcal{P}_1$$

Therefore, $N$ is a null set of $\mathcal{P}_1$.

Let $U(X) = (U_1, \ldots, U_n)$ where $U_1 = \sum_{i=1}^n X_i, U_2 = \sum_{i<j} X_i X_j, U_3 = \sum_{i<j<k} X_i X_j X_k, \ldots, U_n = \prod_{i=1}^n X_i$. By Problem 6.9 and 6.10, $T(X) = (X_1, \ldots, X_n)$ is equivalent to $U(X)$ and also $U(X)$ is equivalent to $V(X)$, thus $T(X)$ is equivalent to $V(X)$ (i.e. there exists a one-to-one function $g$ such that $g(T(X)) = V(X)$) thus $T(X)$ is complete and sufficient in $\mathcal{P}_0$. It follows by part (a) for problem 6.32 that $T(X)$ is complete for $\mathcal{P}_1$. 
1.7.5: Prove or disprove by counterexample each of the following statements. If \( \phi \) is convex on \((a, b)\) so is

(i) \( e^{\phi(x)} \)

**Solution:** We know that \( e^x \) is a convex function over \( \mathbb{R} \), and so it is necessarily convex in the interval \((a, b)\). So for all \( \gamma \in (0, 1) \) and \( x, y \in (a, b) \)

\[
e^{\gamma x + (1-\gamma) y} \leq \gamma e^x + (1-\gamma)e^y \tag{4}
\]

Since \( \phi \) is convex. Then for all \( \gamma \in (0, 1) \) and \( x, y \in (a, b) \)

\[
\phi(\gamma x + (1-\gamma) y) \leq \gamma \phi(x) + (1-\gamma)\phi(y) \tag{5}
\]

Exponentiating (5) we get,

\[
e^{\phi(\gamma x + (1-\gamma) y)} \leq e^{\gamma \phi(x) + (1-\gamma)\phi(y)} \\
\leq \gamma e^{\phi(x)} + (1-\gamma)e^{\phi(y)} \text{ using (4)}
\]

Hence, \( e^{\phi(x)} \) is convex.

(i) \( log\phi(x) \)

**Solution:** For a counter example let us take \( \phi(x) = x^2 \) on \((0, 1)\) which is convex for all \( x \in (0, 1) \). (show this!). Now, for all \( x \in (0, 1) \)

\[
\frac{d^2}{dx^2} log\phi(x) = \frac{-2}{x^2} < 0
\]

Hence, \( log\phi(x) \) is not a convex function on \((0, 1)\).
1.7.16: (a) If \( f : \mathbb{R}^p \rightarrow \mathbb{R} \) is superharmonic, then \( \varphi(f(\cdot)) \) is also superharmonic, where \( \varphi : \mathbb{R} \rightarrow \mathbb{R} \) is a twice-differentiable increasing concave function.

**Solution:**

\[
\sum_{i=1}^{p} \frac{\partial^2}{\partial x_i^2} \varphi(f(x)) = \sum_{i=1}^{p} \frac{\partial}{\partial x_i} \left( \frac{\partial \varphi \partial f}{\partial f \partial x_i} \right) \\
= \sum_{i=1}^{p} \left( \frac{\partial \varphi \partial^2 f}{\partial f \partial x_i^2} + \frac{\partial^2 \varphi}{\partial f^2} \left( \frac{\partial f}{\partial x_i} \right)^2 \right) \\
= \frac{\partial \varphi}{\partial f} \sum_{i=1}^{p} \frac{\partial^2 f}{\partial x_i^2} + \frac{\partial^2 \varphi}{\partial f^2} \sum_{i=1}^{p} \left( \frac{\partial f}{\partial x_i} \right)^2 \\
\leq 0
\]

Thus \( \varphi(f(\cdot)) \) is also superharmonic.

(b) If \( h \) is superharmonic, then \( h^*(x) = \int g(x - y)h(y)dy \) is also superharmonic, where \( g(\cdot) \) is a density.

**Solution:** Let \( y = x + t \) then \( h^*(x) = \int g(-t)h(x + t)dt \). We know \( h(x) \) is superharmonic, then \( \frac{\partial^2 h(x)}{\partial x_i^2} \leq 0 \). Also, \( g(\cdot) \) is a density, then \( g(-t) \geq 0 \).

So \( \frac{\partial^2 h^*(x)}{\partial x_i^2} = \int g(-t)\frac{\partial^2 h(x + t)}{\partial x_i^2}dt \leq 0 \). Then \( \sum_{i=1}^{p} \frac{\partial^2 h^*(x)}{\partial x_i^2} \leq 0 \). That means \( h^*(x) = \int g(x - y)h(y)dy \) is also superharmonic.

(c) If \( h_\gamma \) is superharmonic, then so is \( h^*(x) = \int h_\gamma(y)dG(\gamma) \), where \( G(\gamma) \) is a distribution function.

**Solution:**

\[
\sum_{i=1}^{p} \frac{\partial^2}{\partial x_i^2} h^*(x) = \sum_{i=1}^{p} \int \frac{\partial^2}{\partial x_i^2} h_\gamma(x)dG(\gamma) \\
= \int \left( \sum_{i=1}^{p} \frac{\partial^2}{\partial x_i^2} h_\gamma(x) \right)dG(\gamma) \\
\leq 0
\]

So \( h^*(x) = \int h_\gamma(y)dG(\gamma) \) is superharmonic.
1.8.11: Suppose $X_1, \ldots, X_n$ have a common mean $\xi$ and variance $\sigma^2$, and that $\text{cov}(X_i, X_j) = \rho_{j-i}$. For estimating $\xi$, show that:

(b) $\bar{X}$ is consistent if $|\rho_{j-i}| \leq M \gamma^{j-i}$ with $|\gamma| < 1$.

**Solution:** It is easy to check $E\bar{X} = \xi$. If $\text{Var}(\bar{X}) \rightarrow 0$, then $\bar{X}$ is consistent.(by Theorem 1.8.2)

\[
\text{Var}(\bar{X}) = \frac{1}{n^2} \left[ \sum_{i=1}^{2} \text{Var}(X_i) + 2 \sum_{i<j} \rho_{j-i} \right] = \frac{\sigma^2}{n} + \frac{2}{n^2} \sum_{i<j} \rho_{j-i} \\
\leq \frac{\sigma^2}{n} + \frac{2}{n^2} \sum_{i<j} |\rho_{j-i}| \leq \frac{\sigma^2}{n} + \frac{2}{n^2} \sum_{i<j} M \gamma^{j-i} \\
= \frac{\sigma^2}{n} + \frac{2M}{n^2} \sum_{k=1}^{n-1} (n-k) \gamma^k \leq \frac{\sigma^2}{n} + \frac{2M}{n^2} \sum_{k=1}^{n-1} n \gamma^k \\
= \frac{\sigma^2}{n} + \frac{2M}{n} \sum_{k=1}^{\infty} \gamma^k = \frac{\sigma^2}{n} + \frac{2M}{n} \frac{\gamma}{1-\gamma} \rightarrow 0 \text{ as } n \rightarrow \infty.
\]

So $\bar{X}$ is consistent if $|\rho_{j-i}| \leq M \gamma^{j-i}$.
1.8.17: Variance stabilizing transformations are transformations for which the resulting statistics has an asymptotic variance that is independent of the parameters of interest. For each of the following case, find the asymptotic distribution of the transformed statistics and show that it is variance stabilizing.

(a) $T_n = \frac{1}{n} \sum_{i=1}^{n} X_i$, $X_i \sim \text{Poisson}(\lambda)$, $h(T_n) = \sqrt{T_n}$.

Solution: By CLT, $\sqrt{n}(T_n - \lambda) \rightarrow N(0, \lambda)$. Then by Theorem 1.8.12, $\sqrt{n}[h(T_n) - h(\lambda)] \rightarrow N(0, \lambda(h'(\lambda))^2)$, where $h'(\lambda) = \frac{1}{2\sqrt{\lambda}}$. So $\sqrt{n}[h(T_n) - h(\lambda)] \rightarrow N(0, \lambda \frac{1}{4\lambda}) = N(0, 1/4)$ and $1/4$ does not depend on $\lambda$ so the transformation is variance stabilizing.

(b) $T_n = \frac{1}{n} \sum_{i=1}^{n} X_i$, $X_i \sim \text{Bernoulli}(p)$, $h(T_n) = \text{arcsin} \sqrt{T_n}$.

Solution: By CLT, $\sqrt{n}(T_n - p) \rightarrow N(0, p(1-p))$. Then by Theorem 1.8.12, $\sqrt{n}[h(T_n) - h(p)] \rightarrow N(0, p(1-p)(h'(p))^2)$, where $h'(p) = \frac{1}{2\sqrt{p(1-p)}}$. So $\sqrt{n}[h(T_n) - h(p)] \rightarrow N(0, p(1-p)\frac{1}{4p(1-p)}) = N(0, 1/4)$ and $1/4$ does not depend on $\lambda$ so the transformation is variance stabilizing.
2.1.10: Let \( X_{p\times 1} \) satisfy \( E(X) = B\psi \) and \( Var(X) = I \), where \( B_{p\times r} \) is known and \( \psi_{r\times 1} \) is unknown. Let \( D = \{ \delta(X) : \delta(X) = a'X, \text{ for some vector } a \} \).

(a) For a known vector \( c \), show that the estimators in \( D \) that are unbiased estimators of \( c'\psi \) satisfy \( a'B = c' \).

**Solution:** Let \( c_{r\times 1} \) be a known vector. Consider any unbiased estimator \( \delta(X) = a'X \) of \( c'\psi \) that belongs to \( D \). Since \( \delta(X) = a'X \) is an unbiased estimator of \( c'\psi \), then \( Ea'X = c'\psi \) for all \( \psi \). Note that

\[
\begin{align*}
  c'\psi &= Ea'X = a'EX = a'B\psi \\
  \Rightarrow (a'B - c')\psi &= 0 \quad \text{for all } \psi \\
  \Rightarrow a'B - c' &= 0 \Rightarrow a'B = c'
\end{align*}
\]

Thus the estimators in \( D \) that are unbiased estimators of \( c'\psi \) satisfy \( a'B = c' \).

(b) Let \( D_c = \{ \delta(X) : \delta(X) = a'X, a'B = c' \} \) be the class of linear unbiased estimators of \( c'\psi \). Show that the BLUE of \( c'\psi \), the linear unbiased estimator with minimum variance, is \( \delta^* = a^*X \) where \( a^* = a'B(B'B)^{-1}B' \) and \( a^*B = c' \) with variance \( Var(\delta^*) = c'(B'B)^{-1}c \).

**Solution:** Let \( \delta(X) = a'X \) any estimator in \( D_c \). Since \( a \in \mathbb{R}^p \), there exists vectors \( a_1, a_2 \) such that \( a = a_1 + a_2 \) with \( a_1 \in C(B) \) and \( a_2 \in C(B)^\perp \), moreover \( ||a||^2 = ||a_1||^2 + ||a_2||^2 \), and \( a_1 = Pa \) where \( P = B(B'B)^{-1}B' \) is the orthogonal projection onto \( C(B) \). Note that

\[
\begin{align*}
  Var(\delta(X)) &= Var(a'a) = a'VarXa = a'Ia = a'a \\
  &= ||a||^2 = ||a_1||^2 + ||a_2||^2 \geq ||a_1||^2 = ||Pa||^2 = (Pa)'(Pa) = (Pa)'I(Pa) \\
  &= (Pa)'VarX(Pa) = Var((Pa)'X).
\end{align*}
\]

Thus, \( Var((Pa)'X) \leq Var(\delta(X)) \) for all \( \delta(X) \in D_c \). Let \( a^* = Pa \), so \( \delta^*(X) = a''X = (Pa)'X \) has minimum variance with the following three
\[ a^\ast = a'P' = a'P = a'B'(B'B)^{-1}B' \]
\[ a^\ast B = a'B(B'B)^{-1}B' = a'B = c'. \]

\[ \text{Var}(\delta^\ast) = a^\ast a^* = (a'B(B'B)^{-1}B')(B(B'B)^{-1}B'a) = a'B(B'B)^{-1}B'a = c(B'B)^{-1}c \]

(c) Let \( D_0 = \{ \delta(X) : \delta(X) = a'X, a'B = 0 \} \) be the class of linear unbiased estimators of zero. Show that if \( \delta \in D_0, \) then \( \text{Cov}(\delta, \delta^\ast) = 0. \)

**Solution:** Let \( \delta \in D_0, \) so \( \delta(X) = a'X \) with \( a'B = 0 \) and \( E\delta = 0. \) It suffices to show that \( E(\delta\delta^\ast) = 0 \) because \( \text{Cov}(\delta, \delta^\ast) = E(\delta\delta^\ast) - (E\delta)(E\delta^\ast) = E(\delta\delta^\ast) - 0 = E(\delta\delta^\ast). \) Note that

\[ E(\delta\delta^\ast) = E((a'X)(a^\ast X)) = E(a'XX^\ast a^*) = a'E(XX^\ast)a^* \]

\[ = a'(\text{Var}X + (EX)(EX'))a^* = a'(1 + (EX)(EX'))a^* = a'a^* + E(a'X)EX^\ast a^* \]

\[ = a'a^* + (0)EX^\ast a^* = a'a^* = a'(B(B'B)^{-1}B'a) \]

\[ = (a'B)(B'B)^{-1}B'a = (0)(B'B)^{-1}B'a = 0 \]

Thus, \( \text{Cov}(\delta, \delta^\ast) = 0. \)

(d) **Theorem:** An estimator \( \delta^\ast \in D_c \) satisfies \( \text{Var}(\delta^\ast) = \min_{\delta \in D_c} \text{Var}(\delta) \) if and only if \( \text{Cov}(U, \delta^\ast) = 0, \) where \( U \) is any estimator in \( D_0. \)

**Solution:**

Proof. \((\Rightarrow)\) Let \( U \) be any estimator in \( D_0. \) Suppose \( \delta^\ast \in D_c \) satisfies \( \text{Var}(\delta^\ast) = \min_{\delta \in D_c} \text{Var}(\delta), \) then by part (c) \( \text{Cov}(U, \delta^\ast) = 0. \) Thus \( \text{Cov}(U, \delta^\ast) = 0 \) for all \( U \in D_0. \)

\((\Leftarrow)\) Let \( \delta \) be an arbitrary estimator in \( D_c. \) Suppose \( \text{Cov}(U, \delta^\ast) = 0, \) where \( U \) is any estimator in \( D_0. \) Consider \( \delta = \delta^\ast + (\delta - \delta^\ast). \) Note that since \( \delta, \delta^\ast \in D_c, E(\delta - \delta^\ast) = E(\delta) - E(\delta^\ast) = c'\psi - c'\psi = 0. \) So, \( U_0 = \delta - \delta^\ast \in D_0, \) this implies that \( \text{Cov}(U_0, \delta^\ast) = 0. \) Now,

\[ \text{Var}(\delta) = \text{Var}(\delta^\ast + U_0) = \text{Var}(\delta^\ast) + \text{Var}(U_0) + 2\text{Cov}(U_0, \delta^\ast) = \text{Var}(\delta^\ast) + \text{Var}(U_0) \]
Since $\text{Var}(U_0) \geq 0$ then $\text{Var}(\delta^*) \leq \text{Var}(\delta)$ and since $\delta$ was an arbitrary estimator in $D_c$, it follows that $\text{Var}(\delta^*) \leq \text{Var}(\delta)$ for all $\delta \in D_c$, thus $\delta^* \in D_c$ satisfies $\text{Var}(\delta^*) = \min_{\delta \in D_c} \text{Var}(\delta)$.

(e) Show that the results here can be directly extended to the case of $\text{Var}(X) = \Sigma$ where $\Sigma_{p \times p}$ is a known matrix, by considering the transformed problem with $X^* = \Sigma^{-1/2}X$ and $B^* = \Sigma^{-1/2}B$.

**Solution:** Note that $B_{p \times r}^*$ is a known matrix and

\[
EX^* = E(\Sigma^{-1/2}X) = \Sigma^{-1/2}EX = \Sigma^{-1/2}B\psi = B^*\psi
\]

\[
\text{Var}(X^*) = \text{Var}(\Sigma^{-1/2}X) = \Sigma^{-1/2}\text{Var}X\Sigma^{-1/2} = \Sigma^{-1/2}\Sigma\Sigma^{-1/2} = I
\]

So, now we can proceed exactly as we did in parts (a-d), in other order these results can be extended to the case of $\text{Var}X = \Sigma$ and $\Sigma_{p \times p}$ is a known matrix.
2.1.17: If $T$ has the binomial distribution $b(p, n)$ with $n > 3$, use Method 1 to find UMVU estimator of $p^3$.

**Solution:** Let $\delta(T)$ be the UMVU estimator of $p^3$. $E_p \delta(T) = \sum_{t=0}^{n} \binom{n}{t} p^t (1-p)^{n-t} \delta(t) = p^3$. Let $\rho = \frac{p}{1-p}$ then $p = \frac{\rho}{1+\rho}$, $p = \frac{1}{1+\rho}$.

So we have: $\sum_{t=0}^{n} \binom{n}{t} \left(\frac{\rho}{1+\rho}\right)^t \left(\frac{1}{1+\rho}\right)^{n-t} \delta(t) = \left(\frac{\rho}{1+\rho}\right)^3 \sum_{t=0}^{n} \binom{n}{t} \rho^t \delta(t) = \rho^3 (1+\rho)^{n-3} \sum_{t=0}^{n} \binom{n}{t} \rho^t \delta(t) = \sum_{y=0}^{n-3} \binom{n-3}{y} \rho^y \delta(t) = \sum_{t=3}^{n} \binom{n-3}{t-3} \rho^t$. It is clear that both sides of equation are polynomial of $\rho$. So we need $\binom{n}{t} \rho^t \delta(t) = \binom{n-3}{t-3} \rho^t$ for $t = 3, \cdots, n$. That means $\frac{n!}{t!(n-t)!} \delta(t) = \frac{(n-3)!}{(t-3)!(n-t)!}$. So $\delta(t) = \frac{t(t-1)(t-2)}{n(n-1)(n-2)}$ for $t = 3, \cdots, n$ and $\delta(t) = 0$ for $t = 0, 1, 2$. 

20
Extra Problem: To show that the order statistics are not complete in Example 6.15(iv).
Consider the joint density of a sample from \((L(\theta, 1))\),

\[
p_\theta(x) = \frac{\exp[-\sum_{i=1}^{n}(x_i - \theta)]}{\prod_{i=1}^{n} = \{1 + \exp[-(x_i - \theta)]\}^2}
\]

Note that this is a member of the location family. Let \(T(X) = (X^{(1)}, \ldots, X^{(n)})\) be the set of order statistics for the Logistic family. Let \(R(X) = X^{(n)} - X^{(1)} = Z^{(n)} + \theta - (Z^{(1)} + \theta) = Z^{(n)} - Z^{(1)}\) where \(Z_i \sim L(0, 1)\). Since the distribution of \(Z^{(n)} - Z^{(1)}\) does not depend on \(\theta\), it follows that \(R(X)\) is an ancillary statistic. Note that \(R(X)\) is a function of \((T(X))\), in other words \(R(X)\) and \(T(X)\) are dependent. Now, if we assume \(T(X)\) to be complete then from Basu’s Theorem we have \(T(X)\) and \(R(X)\) to be independent, since \(R(X)\) is ancillary, but this is a contradiction. Thus \(T(X)\) is not complete.