1. Let $X_1, ..., X_n$ be i.i.d. from $U(a,b)$, $a, b \in R$. Prove that $T = (X_{(1)}, X_{(n)})$ is a complete sufficient statistics.

The joint density is $(b-a)^{-n}I(a < X_1, ..., X_n < b) = (b-a)^{-n}I(a < X_{(1)} \leq X_{(n)} < b)$. By Factorization Criterion $T = (X_{(1)}, X_{(n)})$ is sufficient for $(a,b)$. To prove completeness, we first obtain the joint density of $T$:

$$P(X_{(1)} \leq t_1, X_{(n)} \leq t_2) = F(t_2)^n - (F(t_2) - F(t_1))^n, \quad (1)$$

$$f_T(t_1, t_2) = n(n-1)(F(t_2) - F(t_1))^{n-2}f(t_1)f(t_2) = n(n-1)(t_2 - t_1)^{(n-2)}(b-a)^{-n}$$

for $a < t_1 \leq t_2 < b$. Secondly, consider a function $g(t_1, t_2)$ such that for all $-\infty < a < b < \infty$

$$\int_a^b \int_{t_1}^b g(t_1, t_2)n(n-1)(t_2 - t_1)^{(n-2)}(b-a)^{-n}dt_2dt_1 = 0.$$

We represent $g = g^+ - g^-$, where $g^+, g^- \geq 0$ everywhere. Then for all $-\infty < a < b < \infty$

$$\int_a^b \int_{t_1}^b g^+(t_1, t_2)(t_2 - t_1)^{(n-2)}dt_2dt_1 = \int_a^b \int_{t_1}^b g^-(t_1, t_2)(t_2 - t_1)^{(n-2)}dt_2dt_1.$$

Now let $u = t_2 - t_1$ then for all $-\infty < a < b < \infty$

$$\int_a^b \int_0^{b-a} g^+(t_1, t_1 + u)u^{(n-2)}dudt_1 = \int_a^b \int_0^{b-a} g^-(t_1, t_1 + u)u^{(n-2)}dudt_1.$$

And for all Borel sets $A \subset R, B \subset R^+$

$$\int_A \int_B g^+(t_1, t_1 + u)u^{(n-2)}dudt_1 = \int_A \int_B g^-(t_1, t_1 + u)u^{(n-2)}dudt_1,$$

which implies $g^+(t, t + u) = g^-(t, t + u)$ a.e. on $(t, u) \in R \times R^+$, and finally $g = 0$ a.e. on $[a, b]^2$. Thus, $T$ is a complete sufficient statistics.

2. Let $(X_1, Y_1), ..., (X_n, Y_n)$ be i.i.d. from $U([a_1, a_2] \times [b_1, b_2]), a_1 < a_2, b_1 < b_2$. Let $T = (X_{(1)}, X_{(n)}, Y_{(1)}, Y_{(n)})$. Prove that $T$ is a complete statistics when $a_1, a_2, b_1, b_2$ are arbitrary, but not complete when $a_1, b_1, \Delta > 0$ are arbitrary and $a_2 = a_1 + \Delta, b_2 = b_1 + \Delta$.

When $a_1, a_2, b_1, b_2$ are arbitrary, $T$ is complete because $(X_{(1)}, X_{(n)})$ and $Y_{(1)}, Y_{(n)})$ are independent and complete for $U([a_1, a_2])$ and $U([b_1, b_2])$, respectively by the previous problem.

When $a_1, b_1, \Delta > 0$ are arbitrary and $a_2 = a_1 + \Delta, b_2 = b_1 + \Delta$, we have $EX_{(n)} = a_1 + \frac{n}{n+1}\Delta$, $EX_{(1)} = a_1 + \frac{1}{n+1}\Delta$, $EY_{(n)} = b_1 + \frac{n}{n+1}\Delta$, $EY_{(1)} = b_1 + \frac{1}{n+1}\Delta$. So

$$E[(X_{(n)} - X_{(1)}) - (Y_{(n)} - Y_{(1)})] = 0 \quad \text{but} \quad (X_{(n)} - X_{(1)}) - (Y_{(n)} - Y_{(1)}) \neq 0,$$

which implies that $T$ is not complete.
3. Let $X$ take on the values $-1, 0, 1, 2, 3$ with probabilities $P(X = -1) = 2p(1 - p)$ and $P(X = k) = p^k(1 - p)^{(3-k)}$ for $k = 0, 1, 2, 3$.

(a) Check that it is a probability distribution.

$$2p(1-p)+(1-p)^3+p(1-p)^2+p^2(1-p)+p^3 = 2p-2p^2+1-3p+3p^2-p^3+p-2p^2+p^3+p^2-p^3+p^2 = 1$$

(b) Determine the LMVU estimator at $p_0$ of $p(1 - p)$. Is it UMVU?

Let $\delta(X)$ be unbiased estimator of $p(1 - p)$. Then for all $0 < p < 1$

$$p(1 - p) = \delta(-1)2p(1 - p) + \delta(0)(1 - p)^3 + \delta(1)p(1 - p)^2 + \delta(2)p^2(1 - p) + \delta(3)p^3$$

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

$$+p^3[-\delta(0) + \delta(1) - \delta(2) + \delta(3)].$$

Thus the coefficients of both polynomials of $p$ should be the same

$$\delta(0) = 0, 2\delta(-1) + \delta(1) = 1, -2\delta(-1) - 2\delta(1) + \delta(2) = -1, \delta(1) - \delta(2) + \delta(3) = 0.$$ Simplifying

$$\delta(0) = 0, \delta(1) = 1 - 2\delta(-1), \delta(2) = 1 - 2\delta(-1), \delta(3) = 0.$$

Let $\delta(-1) = k$, then the variance of $\delta(X)$ is

$$Var[\delta(X)] = (k - p(1 - p))^2p(1 - p) + (0 - p(1 - p))^2(1 - p)^3$$

$$+ (1 - 2k - p(1 - p))^2p(1 - p)^2 + (1 - 2k - p(1 - p))^2p^2(1 - p) + (0 - p(1 - p))^2p^3.$$ We need to minimize it with respect to $k$ at $p_0$, so take derivative, set it equal to 0 and check it is minimum not maximum.

$$2(k - p_0(1 - p_0))2p_0(1 - p_0) + 2(1 - 2k - p_0(1 - p_0)(-2)[p_0(1 - p_0)^2 + p_0^2(1 - p_0)] = 0$$

$$k - p_0(1 - p_0) - (1 - 2k - p_0(1 - p_0)) = 0, \text{ so } k = 1/3.$$ It is a point of minimum because the variance is a quadratic function in $k$ with a positive leading coefficient. Then the LMVU at $p_0$ is given by

$$\delta(-1) = 1/3, \delta(0) = 0, \delta(1) = 1/3, \delta(2) = 1/3, \delta(3) = 0.$$ Now we need to check if it is the UMVU estimator. We will use Theorem that unbiased $\delta$ is UMVU iff $E(\delta U) = 0$ for all $U$ with $EU = 0$ and all $0 < p < 1$. By (2) $U$ is given by

$$0 = U(0) + p[2U(-1) - 3U(0) + U(1)] + p^2[-2U(-1) + 3(U(0) - 2U(1) + U(2)]$$

$$+ p^3[-U(0) + U(1) - U(2) + U(3)],$$

$$U(0) = 0, [2U(-1) - 3U(0) + U(1)] = 0, [-2U(-1) + 3U(0) - 2U(1) + U(2)] = 0,$$

$$[-U(0) + U(1) - U(2) + U(3)] = 0,$$

$$U(0) = U(3) = 0, U(1) = U(2) = -2U(-1).$$ (3)
And moreover, $\delta$ is UMVU iff for all $U$ satisfying (3)
\[
\delta(0)U(0) = \delta(3)U(3) = 0, \delta(1)U(1) = \delta(2)U(2) = -2\delta(-1)U(-1)
\]
or equivalently, $\delta$ is UMVU iff
\[
\delta(1) = \delta(2) = \delta(-1).
\]
Then the estimator
\[
\delta(-1) = 1/3, \delta(0) = 0, \delta(1) = 1/3, \delta(2) = 1/3, \delta(3) = 0.
\]
is the UMVU of $p(1-p)$.

(c) Determine the LMVU estimator at $p_0$ of $p^2(1-p)$. Is it UMVU? By (2) $\delta$ is given by
\[
p^2(1-p) = \delta(0) + p[2\delta(-1) - 3\delta(0) + \delta(1)] + p^2[-2\delta(-1) + 3\delta(0) - 2\delta(1) + \delta(2)]
\]
\[
+ p^3[-\delta(0) + \delta(1) - \delta(2) + \delta(3)],
\]
\[
\delta(0) = 0, [2\delta(-1) - 3\delta(0) + \delta(1)] = 0, [-2\delta(-1) + 3\delta(0) - 2\delta(1) + \delta(2)] = 1,
\]
\[
[-\delta(0) + \delta(1) - \delta(2) + \delta(3)] = -1.
\]
Then
\[
\delta(-1) = k, \delta(0) = 0, \delta(1) = -2k, \delta(2) = 1 - 2k, \delta(3) = 0.
\]
Again we need to minimize the variance with respect to $k$ at $p_0$.
\[
Var[\delta(X)] = (k - p^2(1-p))^22p(1-p) + (0 - p^2(1-p))^2(1-p)^3
\]
\[
+(-2k - p^2(1-p))^2p(1-p)^2 + (1 - 2k - p^2(1-p))^2p^2(1-p) + (0 - p^2(1-p))^2p^3.
\]
So the derivative is
\[
2(k - p_0^2(1-p))(2p_0(1-p) + 2(-2k - p_0^2(1-p))(1-2k) = 0
\]
\[
k - p_0^2(1-p) + (2k + p_0^2(1-p))(1-p) - (1 - 2k - p_0^2(1-p))p_0 = 0, \quad \text{so} \quad k = p_0/3.
\]
It is a point of minimum because the variance is a quadratic function in $k$ with a positive leading coefficient. Then the LMVU at $p_0$ is given by
\[
\delta(-1) = p_0/3, \delta(0) = 0, \delta(1) = -2p_0/3, \delta(2) = 1 - 2p_0/3, \delta(3) = 0.
\]
In order for $\delta$ to be UMVU by part (a) we should have $\delta(-1) = \delta(1) = \delta(2)$, which is not true here. So this estimator is not UMVU.

4. If $\delta_1, \delta_2$ have finite second moments, and are UMVU estimators of $g_1(\theta), g_2(\theta)$, respectively, then $a_1\delta_1 + a_2\delta_2$ also has finite second moment and is UMVU for estimating $a_1g_1(\theta) + a_2g_2(\theta)$, for any real $a_1, a_2$. 
We have $E\delta_i(X) = g_i(\theta)$, $E\delta^2(X) < \infty$, $i = 1, 2$, for all $\theta$ and by Theorem $E[\delta_i(X)U] = 0$, $i = 1, 2$, for all $\theta$ and all $U \in \mathcal{U}$. Then for any real $a_1, a_2$ we have $E[a_1\delta_1(X) + a_2\delta_2(X)] = a_1g_1(\theta) + a_2g_2(\theta)$, $E[a_1\delta_1(X) + a_2\delta_2(X)]^2 < \infty$ for all $\theta$. Moreover, for all $U \in \mathcal{U}$ and for all $\theta$ we have $E[(a_1\delta_1(X) + a_2\delta_2(X))U] = 0$, so by Theorem $a_1\delta_1 + a_2\delta_2$ is the UMVU of $a_1g_1(\theta) + a_2g_2(\theta)$, for any real $a_1, a_2$.

5. Let $X_1, \ldots, X_n$ be i.i.d. from Poisson distribution with parameter $\lambda > 0$. Use both methods to find UMVU of $g(\lambda) = e^{-\lambda}$. Hint: consider $\delta = 1$ if $X_1 = 0$ and 0 otherwise.

First, note that $T = \sum_{i=1}^n X_i$ is a complete sufficient statistics for this exponential family of full rank. Also $T$ has Poisson distribution with parameter $n\lambda$.

Method 1: find unbiased estimator $\delta(T)$, such that $E\delta(T) = e^{-\lambda}$. To this end

$$
\sum_{k=0}^{\infty} \frac{(n\lambda)^k}{k!} e^{-n\lambda} \delta(k) = e^{-\lambda} \quad \forall \lambda > 0,
$$

$$
\sum_{k=0}^{\infty} \frac{n^k}{k!} \delta(k) \lambda^k = e^{(n-1)\lambda} = \sum_{k=0}^{\infty} \frac{(n-1)^k}{k!} \lambda^k \quad \forall \lambda > 0,
$$

$$
\frac{n^k}{k!} \delta(k) = \frac{(n-1)^k}{k!}, \quad \text{so} \quad \delta(k) = \left( \frac{n-1}{n} \right)^k, k = 0, 1, 2, \ldots \quad (4)
$$

Thus, $\delta(T) = \left( \frac{n-1}{n} \right)^T$.

Method 2: consider $\delta = 1$ if $X_1 = 0$ and 0 otherwise. Then $E\delta = P(X_1 = 0) = e^{-\lambda}$, so $\delta = I(X_1 = 0)$ is unbiased estimator of $e^{-\lambda}$. Now we need to find $E(\delta|T)$, which is the UMVU estimator of $e^{-\lambda}$. We have $E(\delta|T) = P(X_1 = 0|T)$ and for $t = 0, 1, 2, \ldots$

$$
E(\delta|T = t) = P(X_1 = 0|T = t) = \frac{P(X_1 = 0, T = t)}{P(T = t)} = \frac{P(X_1 = 0, X_2 + \ldots + X_n = t)}{P(T = t)}
$$

$$
= \frac{e^{-\lambda}((n-1)\lambda)^t e^{-(n-1)\lambda}}{(\lambda)^t e^{-\lambda}} = \left( \frac{n-1}{n} \right)^t.
$$

Thus, $\delta(T) = \left( \frac{n-1}{n} \right)^T$.

6. Let $T$ be a binomial random variable with $n > 3$ and $0 < p < 1$. Use both methods to find the UMVU of $p^3$. Hint: consider $\delta = 1$ if $X_1 = X_2 = X_3 = 1$ and 0 otherwise.

First, note that $X_1, \ldots, X_n$ are i.i.d. Bernoulli random variables with parameter $0 < p < 1$ and $T = \sum_{i=1}^n X_i$ is a complete sufficient statistics for this exponential family of full rank. And $T$ is a binomial random variable with $n$ and $0 < p < 1$.

Method 1: find unbiased estimator $\delta(T)$, such that $E\delta(T) = p^3$. To this end for all $0 < p < 1$

$$
\sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \delta(k) = p^3, \quad \text{set} \quad \rho = p/(1-p), \ p = \rho/(1+\rho), \ 1-p = 1/(1+\rho),
$$
Thus, \( \delta(T) = \frac{T(T-1)(T-2)}{n(n-1)(n-2)} \).

Method 2: consider \( \delta = 1 \) if \( X_1 = X_2 = X_3 = 1 \) and 0 otherwise. Then \( E\delta = P(X_1 = X_2 = X_3 = 1) = p^3 \), so \( \delta = I(X_1 = X_2 = X_3 = 1) \) is unbiased estimator of \( p^3 \). Now we need to find \( E(\delta|T) \), which is the UMVU estimator of \( p^3 \). We have \( E(\delta|T) = P(X_1 = X_2 = X_3 = 1|T) \) and for \( t = 0, 1, 2, ..., n \)

\[
E(\delta|T = t) = P(X_1 = X_2 = X_3 = 1|T = t) = \frac{P(X_1 = X_2 = X_3 = 1, T = t)}{P(T = t)} = \frac{p^3 (n-3)!}{(t-3)!} \frac{p^3 (1-p)^{n-3-t}}{l(t-3)!} P(T = t) = \frac{t(t-1)(t-2)}{n(n-1)(n-2)}.
\]

Thus, \( \delta(T) = \frac{T(T-1)(T-2)}{n(n-1)(n-2)} \).

7. Let \( X_1, ..., X_m \) and \( Y_1, ..., Y_n \) be independently distributed as \( E(a, b) \) and \( E(c, d) \), respectively.

(a) If \( a, b, c, d \) are unknown, then show \( T = (X_{(1)}, Y_{(1)}, \sum [X_i - X_{(1)}], \sum [Y_k - Y_{(1)}]) \) is a complete sufficient statistics. The joint density is

\[
b^{-m} \exp \left\{ - \sum_{i=1}^{m} (X_i - a)/b \right\} I(X_{(1)} > a) d^{-n} \exp \left\{ - \sum_{j=1}^{n} (Y_j - c)/d \right\} I(Y_{(1)} > c).
\]

Then by Factorization Criterion \( T \) is a sufficient statistics. Since \( X_1, ..., X_m \) and \( Y_1, ..., Y_n \) are independent samples, then \( (X_{(1)}, \sum [X_i - X_{(1)}]) \) and \( (Y_{(1)}, \sum [Y_k - Y_{(1)}]) \) are independent and they are complete for \( E(a, b) \) and \( E(c, d) \), respectively by the argument on page 43 in the textbook. Then \( T \) is a complete statistics.

(b) Find the UMVU estimators of \( b \) and \( (c-a)/b \).

In homework 1 we showed that

\[
\frac{m}{b} (X_{(1)} - a) \sim E(0, 1), \quad \frac{2}{b} \sum [X_i - X_{(1)}] \sim \chi^2_{2(m-1)}, \quad \frac{n}{d} (Y_{(1)} - c) \sim E(0, 1), \quad \frac{2}{d} \sum [Y_j - Y_{(1)}] \sim \chi^2_{2(n-1)},
\]

and they are jointly independent. Note \( EC = 2(m-1), EC^{-1} = 1/(2(m-1)) \) for \( C \sim \chi^2_{2(m-1)} \) and \( E\xi = 1 \) for \( \xi \sim E(0, 1) \). Then

\[
EX_{(1)} = a+b/m, \quad EY_{(1)} = c+d/n, \quad E \sum_{i=1}^{m} [X_i - X_{(1)}] = (m-1)b, \quad E \sum_{j=1}^{n} [Y_j - Y_{(1)}] = (n-1)d,
\]
\[
E(\sum_{i=1}^{m} [X_i - X_{(1)}])^{-1} = \frac{1}{b(m-2)}, \quad E(\sum_{j=1}^{n} [Y_j - Y_{(1)}])^{-1} = \frac{1}{d(n-2)}.
\]

Thus, the UMVU estimator of \(b\) is

\[
\sum_{i=1}^{m} [X_i - X_{(1)}]/(m-1)
\]

and the UMVU estimator of \((c - a)/b\) is

\[
\frac{Y_{(1)}(m-2)}{\sum_{i=1}^{m} [X_i - X_{(1)}]} - \frac{(m-2) \sum_{j=1}^{n} [Y_j - Y_{(1)}]}{n(n-1) \sum_{i=1}^{m} [X_i - X_{(1)}]} - \frac{X_{(1)}(m-2)}{\sum_{i=1}^{m} [X_i - X_{(1)}]} + \frac{1}{m}.
\]