ST210A - Homework 2

Hoang Duong

September 19, 2014

Problem 1. Sample of two-dimensional normal distribution

Proof. **(a)** We have:

$$W = \begin{bmatrix} X_1 \\ \ldots \\ X_n \\ Y_1 \\ \ldots \\ Y_n \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \ldots \\ 0 \\ 0 \\ \ldots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & & 0 & \theta & & 0 \\ & \ddots & & & \ddots & \\ 0 & & 1 & 0 & & \theta \\ \theta & & 0 & 1 & & 0 \\ & \ddots & & & \ddots & \\ 0 & & \theta & 0 & & 1 \end{bmatrix} \right)$$

Denote $X = [X_1, X_2, ..., X_n]^T, Y = [Y_1, Y_2, ..., Y_n]^T, W = [X^T, Y^T]$. For simpification, we write the covariance matrix as $\Sigma = \begin{bmatrix} I_n & \theta I_n \\ \theta I_n & I_n \end{bmatrix}$. For a block matrix $S = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, we have $\det(S) = \det(A)\det(D - CA^{-1}B)$

$$\begin{aligned}
\Rightarrow \det(\Sigma) =& \det(I_n)\det(I_n - \theta I_n I_n \theta I_n) \\
=& \det((1-\theta^2)I_n) \\
=& (1-\theta^2)^n \\
\Sigma^{-1} =& \frac{1}{(1-\theta^2)^n} \begin{bmatrix} I_n & -\theta I_n \\ -\theta I_n & I_n \end{bmatrix} \\
\Rightarrow f_W(w) =& \frac{1}{\sqrt{(2\pi)^{2n}(1-\theta^2)^n}} \exp\left\{ -\frac{1}{2} w^T \Sigma^{-1} w \right\} \\
=& \frac{1}{\sqrt{(2\pi)^{2n}(1-\theta^2)^n}} \exp\left\{ -\frac{1}{2(1-\theta^2)^n} \left[x^T, y^T\right] \begin{bmatrix} I_n & -\theta I_n \\ -\theta I_n & I_n \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\} \\
=& \frac{1}{\sqrt{(2\pi)^{2n}(1-\theta^2)^n}} \exp\left\{ -\frac{1}{2(1-\theta^2)^n} \left[x^T - \theta y^T, -\theta x^T + y^T\right] \begin{bmatrix} x \\ y \end{bmatrix} \right\} \\
=& \frac{1}{\sqrt{(2\pi)^{2n}(1-\theta^2)^n}} \exp\left\{ -\frac{1}{2(1-\theta^2)^n} \left(x^T x + y^T y - 2\theta x^T y\right) \right\}
\end{aligned}$$

Let $T = \left[X^T X + Y^T Y, X^T Y\right]^T := [T_1, T_2]^T$, then $f_T(t) = g_\theta(T)$, for:

$$g_\theta(T) = \frac{1}{\sqrt{(2\pi)^{2n}(1-\theta^2)^n}} \exp\left\{ -\frac{1}{2(1-\theta^2)^n} [1, -2\theta]\, T \right\}.$$

Then by the factorization theorem, T is a sufficient statistics.
Now assume that $f_W(w) = k f_W(w'), k > 0$, we have $\forall\theta$:

$$\log f_W(w) = \log f_W(w') + \log k$$
$$\Rightarrow -\frac{1}{2(1-\theta^2)^n}(T_1 - 2\theta T_2) = -\frac{1}{2(1-\theta^2)^n}(T_1' - 2\theta T_2') + \log k$$
$$\Rightarrow T_1 - 2\theta T_2 = T_1' - 2\theta T_2' - 2(1-\theta^2)^n \log k$$

Since the LHS and RHS of the last expression are polynomial of θ, and they are equal $\forall \theta \in (-1, 1)$, we must have all the coefficient for each degree of the two polynomial are equal. For degree $2n$, we have $-2\log k = 0 \Rightarrow k = 1$. For degree 1, we have $T_2 = T_2'$, and for degree zero we have $T_1 = T_1'$. Thus $T = T'$. By theorem 3.11 in Keener, we have T is minimal sufficient.

(b) Let $h(T) = [1, 0]\, T = X^T X + Y^T Y$, then $h(T)$ is not a constant function, but

$$\begin{aligned} \mathbb{E}h(T) =& \mathbb{E}\left[\sum_{i=1}^n X_i^2 + \sum_{i=1}^n Y_i^2\right] \\ =& 2n \end{aligned}$$

is a constant. Thus T is not complete.

(c) We have $[X_1, X_2, ..., X_n]^T \sim N(\mathbf{0}, I)$, as such $Z_1 = \sum_{i=1}^n X_i^2$ has distribution independent of θ. In fact it has a χ^2 distribution with k degree of freedom. Thus Z_1 is ancillary. Similarly Z_2 is ancillary.

Now consider $Z = [Z_1, Z_2]^T$. For $X_i, Y_i, Y_i = \theta X_i + Y_i - \theta X_i := \theta X_i + U_i$. Then:

$$\begin{aligned} \theta = \mathrm{Cov}(X_i, Y_i) &= \mathrm{Cov}(X_i, \theta X_i) + \mathrm{Cov}(X_i, U_i) \\ &= \theta + \mathrm{Cov}(X_i, U_i) \\ \Rightarrow \mathrm{Cov}(X_i, U_i) &= 0 \end{aligned}$$

Thus X_i, U_i are independent since they are normal and uncorrelated. Let $U_i = \sqrt{1-\theta^2}V_i$ then V_i has variance 1. At the end we have: $Y_i = \theta X_i + \sqrt{1-\theta^2}V_i$. Now:

$$\begin{aligned} X_i^2 Y_i^2 =& \theta^2 X_i^4 + 2\theta\sqrt{1-\theta^2}X_i^3 V + (1-\theta^2)X_i^2 V_i^2 \\ \Rightarrow \mathbb{E}\left[X_i^2 Y_i^2\right] =& 3\theta^2 + 1 - \theta^2 = 2\theta^2 + 1. \\ \Rightarrow \mathrm{Cov}(X_i^2, Y_i^2) =& \mathbb{E}\left[X_i^2 Y_i^2\right] - \mathbb{E}\left[X_i^2\right]\mathbb{E}\left[Y_i^2\right] \\ =& 2\theta^2 \\ \Rightarrow \mathrm{Cov}(Z_1, Z_2) =& \sum_{i=1}^n \sum_{j=1}^n \mathrm{Cov}(X_i^2, Y_j^2) \\ =& \sum_{i=1}^n \mathrm{Cov}(X_i^2, Y_i^2) \\ =& 2n\theta^2 \end{aligned}$$

Since the $\mathrm{Cov}(Z_1, Z_2)$ depends on θ, the joint distribution of $[Z_1, Z_2]$ must also depend on θ. Thus $[Z_1, Z_2]$ is not ancillary. □

Problem 2. Uniform on $(-\theta, \theta)$.

Proof. **(a)** For each $X_i, f_{X_i}(x_i) = \frac{1}{2\theta}\mathbb{I}[-\theta < x_i < \theta]$, since $X_i's$ are independent,

$$
\begin{aligned}
f_X(x) =& \frac{1}{(2\theta)^n} \prod_{i=1}^{n} \mathbb{I}\left[-\theta < x_i < \theta\right] \\
=& \frac{1}{(2\theta)^n} \mathbb{I}\left[x_i \in (-\theta, \theta), \forall i = 1, ..., n\right] \\
=& \frac{1}{(2\theta)^n} \mathbb{I}\left[\max_{i\in\{1,\dots,n\}} \mid x_i \mid < \theta\right]
\end{aligned}
$$

Thus $T = \max_{i\in\{1,\dots,n\}} \mid x_i \mid$ is a sufficient statistic. Now if $\forall\theta$, we have

$$
\begin{aligned}
f_X(x) = k f_X(x') \\
\Rightarrow \mathbb{I}\left[T < \theta\right] = k\mathbb{I}\left[T' < \theta\right]
\end{aligned}
$$

Then $k = 1$, and $T = T'$. Thus T is minimal sufficient.
(b) We will show that V is ancillary.

$$
\begin{aligned}
V =& \frac{\bar{X}}{X_{(n)} - X_{(1)}} \\
=& \frac{\frac{1}{n}\sum_{i=1}^{n} X_i}{X_{(n)} - X_{(1)}} \\
=& \frac{1}{n} \frac{\sum_{i=1}^{n} \frac{X_i}{\theta}}{\frac{X_{(n)}}{\theta} - \frac{X_{(1)}}{\theta}}
\end{aligned}
$$

Let Y_i be i.i.d uniform on $(-1, 1)$, then $V = \frac{1}{n}\frac{\sum_{i=1}^{n} Y_i}{Y_{(n)} - Y_{(1)}}$, as such the distribution of V does not depend on θ. Thus V is ancillary. By the Basu theorem, T is independent with V. □

Problem 3. Uniform on $\left[\theta - \frac{1}{2}, \theta + \frac{1}{2}\right]$.

Proof. **(a)** $T = \left(X_{(1)}, X_{(2)}\right)$. $X = [X_1, X_2, ..., X_n]$

$$
\begin{aligned}
f_{X_i}(x_i) =& \mathbb{I}\left[\theta - \frac{1}{2} \le x_i \le \theta + \frac{1}{2}\right] \\
\Rightarrow f_X(x) =& \prod_{i=1}^{n} \mathbb{I}\left[\theta - \frac{1}{2} \le x_i \le \theta + \frac{1}{2}\right] \\
=& \mathbb{I}\left[x_i \in \left[\theta - \frac{1}{2}, \theta + \frac{1}{2}\right], \forall i \in \{1, 2, ..., n\}\right] \\
=& \mathbb{I}\left[x_{(n)} < \theta + \frac{1}{2} \wedge x_{(1)} > \theta - \frac{1}{2}\right].
\end{aligned}
$$

So T is sufficient.
(b) By the Rao-Blackwell theorem for the sufficient statistics T, we have

$$
\begin{aligned}
R(\theta, \delta(X_1, ..., X_n)) \le& R(\theta, \bar{X}) \\
\Leftrightarrow \mathbb{E}\left[(\theta - \delta)^2\right] \le& \mathbb{E}\left[(\theta - \bar{X})^2\right] \\
\Leftrightarrow \text{MSE}(\delta) \le& \text{MSE}(\bar{X}).
\end{aligned}
$$

Since the loss function is not linear, the equality hold iff $\bar{X} \mid \min\{X_i\}, \max\{X_i\}$ is equal to a constant with probability 1, which is not true when $n > 2$. So MSE for δ is strictly better than MSE for $\bar{X}$ when $n > 2$. And we can easily see that they are equal when $n = 2$ or $n = 1$.

Now we will calculate δ.

First for A, B independent random variables, we have:

$$\mathbb{E}\left[X\right] = \mathbb{E}\left[\mathbb{E}\left[X \mid A\right]\right]$$

$$\begin{aligned}
\mathbb{E}\left[\bar{X} \mid \min\{X_i\} = a, \max\{X_i\} = b\right] =& \mathbb{E}\left[\bar{X} \mid X_i \in [a,b], \forall i = 1, ..., n\right] \\
=& \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i \mid X_i \in [a,b], \forall i = 1, ..., n\right] \\
=& \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[X_i \mid X_j \in [a,b], \forall i = 1, ..., n\right] \\
=& \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[X_i \mid X_i \in [a,b]\right] \\
=& \frac{1}{n}\sum_{i=1}^{n} \frac{b+a}{2} \\
=& \frac{b+a}{2} = \frac{\min\{X_i\} + \max\{X_i\}}{2}
\end{aligned}$$

□

Problem 4. Location-scale family

Proof. Write $X_i = a + bZ_i$ then Z_i has known cumulative distribution function $F(x)$ does not depend on neither a nor b.

(a) Now $(X_1 - X_i)/b = Z_1 - Z_i$ has distribution not depending on a. Thus $(X_1 - X_i)/b$ are ancillary if b is known.

(b) $(X_1 - a)/(X_i - a) = Z_1/Z_i$ has distribution not depending on b. Thus it is ancillary

(c) $(X_1 - X_i)/(X_2 - X_i) = b(Z_1 - Z_i)/(b(Z_2 - Z_i)) = (Z_1 - Z_i)/(Z_2 - Z_i)$ has distribution not depending on neither a nor b. Thus the statistics is ancillary. □

Problem 5. Unbiased estimator

Proof. **(a)** $\mathbb{E}S_1 = \mathbb{P}[X_1 = 0] \times 1 + \mathbb{P}[X_1 \neq 0] \times 0 = \mathbb{P}[X_1 = 0] = \exp(-\lambda)$. Thus S_1 is an unbiased estimator.

$\mathbb{E}S_2 = \frac{1}{n}\mathbb{E}\mathbb{I}\left[X_i = 0\right] = \frac{1}{n} n \exp(-\lambda) = \exp(-\lambda)$. Thus S_2 is also unbiased. Note that we can see that S_2 has a smaller risk than S_1 as it has smaller variance (due to averaging independent observation).

(b) Consider the density function

$$\begin{aligned}
f_X(x) =& \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \\
=& e^{-n\lambda} \lambda^{\sum_{i=1}^{n} x_i} \frac{1}{\prod_{i=1}^{n} x_i!} \\
=& g_\lambda(\sum_{i=1}^{n} x_i) h(x)
\end{aligned}$$

Thus by the factorization theorem, $T(x) = \sum_{i=1}^{n} x_i$ is sufficient.

(c) From the lemma we proved in the last homework, sum of independent Poisson is Poisson with mean parameter equal to sum of mean parameter.

$$
\begin{aligned}
\mathbb{E}[S_1 \mid T=t] =& \mathbb{E}[\mathbb{I}[X_1=0] \mid T=t] \\
=& \mathbb{P}[X_1=0 \mid T=t] \\
=& \frac{\mathbb{P}[X_1=0 \wedge \sum_{i=1}^n X_i = t]}{\mathbb{P}[\sum_{i=1}^n X_i = t]} \\
=& \frac{\mathbb{P}[X_1=0 \wedge \sum_{i=2}^n X_i = t]}{\mathbb{P}[\sum_{i=1}^n X_i = t]} \\
=& \frac{\mathbb{P}[X_1=0]\mathbb{P}[\sum_{i=2}^n X_i = t]}{\mathbb{P}[\sum_{i=1}^n X_i = t]} \\
=& \frac{\exp(-\lambda)((n-1)\lambda)^t \exp(-(n-1)\lambda)/t!}{(n\lambda)^t \exp(-n\lambda)/t!} \\
=& \frac{(n-1)^t}{n^t} = \left(1 - \frac{1}{n}\right)^t
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[S_2 \mid T=t] =& \mathbb{E}\left[\frac{1}{n}\sum_{j=1}^n \mathbb{I}[X_j=0] \mid \sum_{i=1}^n X_i = t\right] \\
=& \frac{1}{n}\sum_{j=1}^n \mathbb{E}\left[\mathbb{I}[X_j=0] \mid \sum_{i=1}^n X_i = t\right] \\
=& \frac{1}{n} n \mathbb{E}[S_1 \mid T=t] \\
=& \left(1 - \frac{1}{n}\right)^t
\end{aligned}
$$

So apply Rao-Blackwell to two estimators result in the same estimator. Now since T is complete as it comes from a full-rank exponential distribution family. It is expected that the two Rao-Blackwellized of S_1 and S_2 are the same as they are both the UMVU. ☐

Problem 6. Determine UMUV for $e^{-2\lambda}$ for single *Poisson* distribution mean λ.

Proof. Let an unbiased estimator be $g(X)$, we need:

$$
\begin{aligned}
\mathbb{E}[g(X)] &= e^{-2\lambda} \\
\Rightarrow \sum_{i=0}^{\infty} g(i) e^{-\lambda} \frac{\lambda^i}{i!} &= e^{-2\lambda} \\
\Rightarrow \sum_{i=0}^{\infty} g(i) \frac{\lambda^i}{i!} &= e^{-\lambda} \\
\Rightarrow \sum_{i=0}^{\infty} g(i) \frac{\lambda^i}{i!} &= \sum_{i=0}^{\infty} (-1)^{i+1} \frac{\lambda^i}{i!} \\
\Rightarrow g(i) &= (-1)^{i+1}, \forall i
\end{aligned}
$$

So $g(X) = (-1)^{X+1}$. Since there is only one observation, this $g(X)$ is the only unbiased estimator, thus it is UMUV. ☐