MA 413 - Statistics for Data Science

Solutions to Exercise 6

1. (a) To see if S or T is ancillary, we check if their CDFs depend on parameters θ_1 and θ_2. We start to compute the CDF of S,

$$\begin{aligned}
F_S(s) &= \Pr(S \leq s) = \Pr(\min(Y_1, \ldots, Y_n) \leq s) \\
&= 1 - \Pr(\min(Y_1, \ldots, Y_n) > s) \\
&= 1 - \Pr(Y_1 > s, \ldots, Y_n > s) \\
&= 1 - \prod_{i=1}^{n} \Pr(Y_i > s) \ (Y_i \text{ independent}) \\
&= 1 - \prod_{i=1}^{n} (1 - \Pr(Y_i \leq s)),
\end{aligned}$$

and when $s \in (\theta_1, \theta_2)$,

$$\Pr(Y_i \leq s) = \int_{\theta_1}^{s} f_{Y_i}(y_i)\ dy = \int_{\theta_1}^{s} 1/(\theta_2 - \theta_1)\ dy = \frac{s - \theta_1}{\theta_2 - \theta_1}.$$

It follows

$$F_S(s) = 1 - \prod_{i=1}^{n} \left(1 - \frac{s - \theta_1}{\theta_2 - \theta_1}\right) = 1 - \left(1 - \frac{s - \theta_1}{\theta_2 - \theta_1}\right)^n.$$

Similarly for T, we compute its CDF when $t \in (\theta_1, \theta_2)$,

$$\begin{aligned}
F_T(t) &= \Pr(T \leq t) = \Pr(\max(Y_1, \ldots, Y_n) \leq t) \\
&= \Pr(Y_1 \leq t, \ldots, Y_n \leq t) \\
&= \prod_{i=1}^{n} \Pr(Y_i \leq t) \ (Y_i \text{ independent}) \\
&= \left(\frac{t - \theta_1}{\theta_2 - \theta_1}\right)^n.
\end{aligned}$$

Note that neither S nor T is ancillary as their distributions depend on θ_1 and θ_2.

(b) Differentiating $F_S(s)$ and $F_T(t)$ leads to the density function of S and T, respectively,

$$f_S(s) = \frac{n}{\theta_2 - \theta_1} \left(1 - \frac{s - \theta_1}{\theta_2 - \theta_1}\right)^{n-1} \mathbf{1}(\theta_1 < s < \theta_2),$$

$$f_T(t) = \frac{n}{\theta_2 - \theta_1} \left(\frac{t - \theta_1}{\theta_2 - \theta_1}\right)^{n-1} \mathbf{1}(\theta_1 < t < \theta_2).$$

(c) To see if $f_S(s)$ or $f_T(t)$ becomes concentrated, we compute the mean and variance of S and T. Let $\theta_3 := \theta_2 - \theta_1$ and consider change of variable $x = s - \theta_1$. We start with the mean of S,

$$\begin{aligned}
E(S) &= \int_{\theta_1}^{\theta_2} s f_S(s)\ ds = \int_{0}^{\theta_3} (x + \theta_1) \frac{n}{\theta_3} \left(1 - \frac{x}{\theta_3}\right)^{n-1}\ dx \\
&= -(x + \theta_1) \frac{n}{\theta_3} \frac{\theta_3}{n} \left(1 - \frac{x}{\theta_3}\right)^n \Bigg|_0^{\theta_3} + \int_0^{\theta_3} \left(1 - \frac{x}{\theta_3}\right)^n\ dx \\
&= \theta_1 - \frac{\theta_3}{n+1} \left(1 - \frac{x}{\theta_3}\right)^{n+1} \Bigg|_0^{\theta_3} = \theta_1 + \frac{\theta_3}{n+1},
\end{aligned}$$

where we identify the following relationship

$$I(n-1) := \int_0^{\theta_3} (x+\theta_1)\left(1-\frac{x}{\theta_3}\right)^{n-1} dx = \frac{\theta_1\theta_3}{n} + \frac{\theta_3^2}{n(n+1)}.$$

With this, we compute

$$\begin{aligned}
E(S^2) &= \int_{\theta_1}^{\theta_2} s^2 f_S(s)\, ds = \int_0^{\theta_3} (x+\theta_1)^2 \frac{n}{\theta_3}\left(1-\frac{x}{\theta_3}\right)^{n-1} dx \\
&= -(x+\theta_1)^2 \frac{n}{\theta_3}\frac{\theta_3}{n}\left(1-\frac{x}{\theta_3}\right)^{n}\Bigg|_0^{\theta_3} + \int_0^{\theta_3} 2(x+\theta_1)\left(1-\frac{x}{\theta_3}\right)^{n} dx \\
&= \theta_1^2 + 2I(n) = \theta_1^2 + \frac{2\theta_1\theta_3}{n+1} + \frac{2\theta_3^2}{(n+1)(n+2)}.
\end{aligned}$$

The variance of S is

$$\mathrm{Var}(S) = E(S^2) - (E(S))^2 = \theta_1^2 + \frac{2\theta_1\theta_3}{n+1} + \frac{2\theta_3^2}{(n+1)(n+2)} - \left(\theta_1 + \frac{\theta_3}{n+1}\right)^2 = \frac{n\theta_3^2}{(n+1)^2(n+2)}.$$

Therefore for S, its mean approaches to θ_1 and its variance to 0 as n goes to infinity. In other words, $f_S(s)$ becomes concentrated to θ_1 as n increases. Similarly for T, its mean and variance are

$$E(T) = \int_{\theta_1}^{\theta_2} t f_T(t)\, dt = \int_0^{\theta_3} (x+\theta_1)\frac{n}{\theta_3}\left(\frac{x}{\theta_3}\right)^{n-1} dx = \theta_2 - \frac{\theta_3}{n+1},$$

$$E(T^2) = \int_{\theta_1}^{\theta_2} t^2 f_T(t)\, dt = \int_0^{\theta_3} (x+\theta_1)^2\frac{n}{\theta_3}\left(\frac{x}{\theta_3}\right)^{n-1} dx = \theta_2^2 - \frac{2\theta_2\theta_3}{n+1} + \frac{2\theta_3^2}{(n+1)(n+2)},$$

and

$$\mathrm{Var}(T) = E(T^2) - (E(T))^2 = \frac{n\theta_3^2}{(n+1)^2(n+2)}.$$

Therefore for T, its mean approaches to θ_2 and its variance to 0 as n goes to infinity, meaning $f_T(t)$ becomes concentrated to θ_2 as n increases.

(d) To see if U is ancillary, we compute the CDF of U,

$$\begin{aligned}
F_U(u) &= \int\int_{\{(s,t):t-s\le u\}} f_{S,T}(s,t)\, ds\, dt = \int\int_{\{(s,t):t-s\le u\}} f_S(s) f_T(t)\, ds\, dt \ (S,\, T \text{ independent}) \\
&= \int\int_{\{(s,t):t-s\le u\}} n\left(1-(s-\theta)\right)^{n-1} n\left(t-\theta\right)^{n-1}\, ds\, dt \\
&= \int\int_{\{(x,y):y-x\le u\}} n^2\left(1-x\right)^{n-1} y^{n-1}\, dx\, dy,
\end{aligned}$$

where we consider change of variables, $x = s-\theta$ and $y = t-\theta$. U is ancillary as $F_U(u)$ does not depend on θ.

2. (a) Recall that given $Y \sim N(\mu, \sigma^2+\mu^2)$, its density function is

$$f(y;\mu,\sigma^2+\mu^2) = \exp\left\{-\frac{1}{2(\sigma^2+\mu^2)}y^2 + \frac{\mu}{\sigma^2+\mu^2}y - \frac{1}{2}\log(2\pi(\sigma^2+\mu^2)) - \frac{\mu^2}{\sigma^2+\mu^2}\right\},$$

so the density function of $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = \exp\left\{-\frac{1}{2(\sigma^2+\mu^2)}\sum_{i=1}^n y_i^2 + \frac{\mu}{\sigma^2+\mu^2}\sum_{i=1}^n y_i - \frac{n}{2}\log(2\pi(\sigma^2+\mu^2)) - \frac{n\mu^2}{\sigma^2+\mu^2}\right\}.$$

(b) The Fisher-Neyman factorization implies that the statistic $(\sum_{i=1}^n y_i, \sum_{i=1}^n y_i^2)$ is sufficient for (μ, σ^2).

3. (a) As (X_i, Y_i) is iid, the density function of $(\mathbf{X}, \mathbf{Y})$ is

$$f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) = \exp\left\{-\theta\sum_{i=1}^n x_i - \frac{1}{\theta}\sum_{i=1}^n y_i\right\}\mathbf{1}(x_1 > 0, \ldots, x_n > 0, y_1 > 0, \ldots, y_n > 0).$$

(b) To see if T or U is ancillary, we need to compute their density functions. When $x > 0$, $y > 0$, we write

$$f_{X,Y}(x,y) = \{\theta\exp(-\theta x)\}\left\{\frac{1}{\theta}\exp(-\frac{1}{\theta}y)\right\}.$$

Let $f_X(x) := \theta\exp(-\theta x)$ and $f_Y(y) := \theta^{-1}\exp(-\theta^{-1}y)$. We notice that $X \sim \text{Exp}(\theta)$, $Y \sim \text{Exp}(\theta^{-1})$, and they are independent. Recall that given $X_i \overset{\text{iid}}{\sim} \text{Exp}(\theta)$, $P := \sum_i X_i \sim \text{Gamma}(n,\theta)$. Likewise, $Q := \sum_i Y_i \sim \text{Gamma}(n, \theta^{-1})$. As P and Q are independent, their joint distribution is

$$f_{P,Q}(p,q) = f_P(p)f_Q(q) = \left\{\frac{\theta^n}{\Gamma(n)}p^{n-1}e^{-\theta p}\right\}\left\{\frac{\theta^{-n}}{\Gamma(n)}q^{n-1}e^{-\frac{1}{\theta}q}\right\} = \frac{1}{\Gamma(n)^2}(pq)^{n-1}\exp\left(-\theta p - \frac{1}{\theta}q\right).$$

From $T = \sqrt{Q/P}$ and $U = \sqrt{PQ}$, we have $P = U/T$ and $Q = UT$. The transformed density function in terms of U and T is

$$f_{U,T}(u,t) = f_{P,Q}(\frac{u}{t}, ut)\left|\det\begin{bmatrix}\frac{1}{t} & -\frac{u}{t^2} \\ t & u\end{bmatrix}\right| = \frac{2u^{2n-1}}{t\Gamma(n)^2}\exp\left(-\theta\frac{u}{t} - \frac{1}{\theta}ut\right).$$

$f_U(u)$ and $f_T(t)$ are marginal density functions of $f_{U,T}(u,t)$, and in particular,

$$\begin{aligned} f_U(u) &= \frac{2u^{2n-1}}{\Gamma(n)^2}\int_0^\infty \frac{1}{t}\exp\left(-\theta\frac{u}{t} - \frac{1}{\theta}ut\right)\, dt \\ &= \frac{2u^{2n-1}}{\Gamma(n)^2}\int_0^\infty \frac{1}{\theta v}\exp\left(-\frac{u}{v} - uv\right)\, \theta dv \text{ (change of variable } v = \frac{t}{\theta},\ dv = \frac{1}{\theta}dt) \\ &= \frac{2u^{2n-1}}{\Gamma(n)^2}\int_0^\infty \frac{1}{v}\exp\left(-\frac{u}{v} - uv\right)\, dv. \end{aligned}$$

Therefore, U is ancillary, but T is not.

(c) According to $f_{U,T}(u,t)$ and the Fisher-Neyman theorem, $(U/T, UT)$ is sufficient for θ, and so is (U,T) because (U,T) and $(U/T, UT)$ are 1-1. Moreover, (U,T) is minimally sufficient. To see this, we consider

$$\frac{f_{U,T}(u,t)}{f_{U,T}(\tilde{u},\tilde{t})} = 2\frac{\tilde{t}}{t}\left(\frac{u}{\tilde{u}}\right)^{2n-1}\exp\left(-\theta\left(\frac{u}{t} - \frac{\tilde{u}}{\tilde{t}}\right) - \frac{1}{\theta}(ut - \tilde{u}\tilde{t})\right),$$

which does not depend on θ if and only if $u/t = \tilde{u}/\tilde{t}$ and $ut = \tilde{u}\tilde{t}$. This condition is equivalent to $u = \tilde{u}$ and $t = \tilde{t}$ (all variables are greater than 0).

4. As $X_i \overset{\text{iid}}{\sim} \text{Poisson}(\theta)$, the joint mass function is

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^n e^{-\theta}\frac{\theta^{x_i}}{x_i!} = e^{-n\theta}\frac{\theta^{\sum_{i=1}^n x_i}}{x_1!x_2!\cdots x_n!}.$$

Recall that if $Y \sim \text{Poisson}(\lambda)$ and $Z \sim \text{Poisson}(\mu)$, then $Y + Z \sim \text{Poisson}(\lambda + \mu)$. Now as $X_i \overset{\text{iid}}{\sim} \text{Poisson}(\theta)$, $T = \sum_i X_i \sim \text{Poisson}(n\theta)$. The corresponding mass function is

$$f_T(t) = e^{-n\theta}\frac{(n\theta)^t}{t!}.$$

The conditional mass function is

$$f_{\mathbf{X}|T}(\mathbf{x}|t) = \Pr(\mathbf{X} = \mathbf{x}|T = t) = \frac{\Pr(\mathbf{X} = \mathbf{x}, T = t)}{\Pr(T = t)} = \frac{f_{\mathbf{X}}(\mathbf{x})\mathbf{1}(T = t)}{f_T(t)} = \frac{(\sum_{i=1}^n x_i)!}{x_1! \cdots x_n!} \left(\frac{1}{n}\right)^{\sum_{i=1}^n x_i} \mathbf{1}(T = t),$$

which does not depend on θ, so T is sufficient for θ according to definition.

5. To show that F_{X_n} converges in distribution to F_X ($X \sim U(0,1)$), we show that given an arbitrarily small ϵ, one can always find a sufficiently large integer N, such that when $n > N$, $|F_{X_n}(x) - F_X(x)| < \epsilon$. We note that $F_X(x) = x$ ($0 < x < 1$) and $F_{X_n}(x) = k/n$, where k is an integer such that $x \in [k/n, (k+1)/n)$. Indeed, the ϵ-N condition holds when we choose $N > \epsilon^{-1}$. We see this in the following,
$$|F_{X_n}(x) - F_X(x)| = \left|x - \frac{k}{n}\right| < \left|\frac{k+1}{n} - \frac{k}{n}\right| = \frac{1}{n} < \frac{1}{N} < \epsilon.$$

   Therefore, $F_{X_n} \xrightarrow{d} F_X$.

6. Fix $\epsilon > 0$, and let $Y_n \sim N(0, 1/n)$. Then

$$\Pr(|X_n - 1| > \epsilon) = \Pr(|1 + Y_n - 1| > \epsilon) = 1 - \Pr(-\epsilon \le Y_n \le \epsilon).$$

   Note that as n goes to infinity, the density function of Y_n becomes concentrated around 0, meaning $\Pr(-\epsilon \le Y_n \le \epsilon)$ approaches to 1. Therefore, $\Pr(|X_n - 1| > \epsilon) \to 0$ and $X_n \xrightarrow{p} 1$.