ISyE 6412A **Theoretical Statistics** **Fall 2023**

HW #6 (due at Canvas midnight on Friday, Oct 20, ET)

(There are 6 questions. After you spent at least 30 minutes per question, please look at the hints on the last two pages.)

1. Suppose $X_1, X_2, \ldots, X_n$ ($n \geq 2$) are independent and identically distributed (iid) with a Uniform$[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$ distribution for some unknown $-\infty < \theta < \infty$, i.e., the X_i's have density

$$f_\theta(x) = \begin{cases} 1, & \text{if } \theta - \frac{1}{2} \leq x \leq \theta + \frac{1}{2}; \\ 0, & \text{otherwise.} \end{cases}$$

It is desired to guess the value of θ under the loss function $L(\theta, d) = (\theta - d)^2$ based on the observed data $\mathbf{X} = (X_1, \ldots, X_n)$. The purpose of this question is to show that the sample mean is inadmissible.

(a) Specify S, Ω, D, and L (i.e., the sample space, the set of all possible distribution functions, the decision space, and the loss function).

(b) Find the risk function of the procedure $\delta_0(\mathbf{X}) = \bar{X}_n = (X_1 + \ldots + X_n)/n$, which is the so-called method of moment estimator.

(c) **Prove that** $T = (X_{(1)}, X_{(n)})$ is a sufficient statistic for θ, where $X_{(1)} = \min(X_1, \ldots, X_n)$ and $X_{(n)} = \max(X_1, \ldots, X_n)$ are the sample minimum and maximum.

(d) While $T = (X_{(1)}, X_{(n)})$ gives all the information about θ, the T itself is not a statistical procedure for estimating θ, since a point estimator of θ must take on real values. To produce point estimators from sufficient statistic T, let us consider a family of procedures of the form

$$\delta_{a,b}(\mathbf{X}) = aX_{(1)} + (1-a)X_{(n)} + b$$

for some real-valued constants a, b. Show that the risk function of $\delta_{a,b}(\mathbf{X})$ is given by

$$R_{\delta_{a,b}}(\theta) = \Big[a\frac{1}{n+1} + (1-a)\frac{n}{n+1} + b - \frac{1}{2}\Big]^2 + \frac{a^2 n + (1-a)^2 n + 2a(1-a)}{(n+1)^2(n+2)},$$

which is minimized at $b = \frac{1}{2} - \frac{a + n(1-a)}{n+1}$ for any given constant a.

(e) *Among all procedures* $\delta_{a,b}(\mathbf{X})$ in part (d), **show that** the choice $a = \frac{1}{2}$ and $b = \frac{1}{2} - \frac{a+n(1-a)}{n+1} = 0$, i.e., $\delta^*(\mathbf{X}) = (X_{(1)} + X_{(n)})/2$, gives uniformly smallest risk function.

(f) **Prove that** when $n \geq 3$, the procedure $\delta^*(\mathbf{X}) = (X_{(1)} + X_{(n)})/2$ in part (c)(iv) is better than $\delta_0(\mathbf{X}) = \bar{X}_n$, and conclude that $\delta_0(\mathbf{X}) = \bar{X}_n$ is inadmissible when $n \geq 3$.

Answer: (a) $S = \mathcal{R}^n$, $\Omega = \{-\infty < \theta < \infty\}, D = (-\infty, \infty)$ and the loss function is $L(\theta, d) = (\theta - d)^2$.
(b) Note that $\mathbf{E}_\theta(\bar{X}_n) = \mathbf{E}_\theta(X_i) = \theta$ and $Var_\theta(\bar{X}_n) = \frac{1}{n} Var_\theta(X_i) = \frac{1}{12n}$. Thus the risk function of $\delta_0(\mathbf{X}) = \bar{X}_n$ is

$$R_{\delta_0}(\theta) = \mathbf{E}_\theta(\theta - \bar{X}_n)^2 = [\theta - \mathbf{E}_\theta(\bar{X}_n)]^2 + Var_\theta(\bar{X}_n) = \frac{1}{12n}.$$

(c) The joint pdf of $\mathbf{X} = (X_1, \ldots, X_n)$ is

$$\begin{aligned} f_\theta(\mathbf{x}) &= \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n I\Big(\theta - \frac{1}{2} < x_i < \theta + \frac{1}{2}\Big) \\ &= I\Big(\theta < x_i + \frac{1}{2} \text{ and } \theta > x_i - \frac{1}{2} \quad \text{for all } i = 1, \ldots, n\Big) \\ &= I\Big(\theta < x_i + \frac{1}{2} \quad \text{for all } i = 1, \ldots, n\Big) \times I\Big(\theta > x_i - \frac{1}{2} \quad \text{for all } i = 1, \ldots, n\Big) \\ &= I\Big(\theta < \min_{1 \leq i \leq n} x_i + \frac{1}{2}\Big) \times I\Big(\theta > \max_{1 \leq i \leq n} x_i - \frac{1}{2}\Big) \\ &= I\Big(\max_{1 \leq i \leq n} x_i - \frac{1}{2} < \theta < \min_{1 \leq i \leq n} x_i + \frac{1}{2}\Big) \end{aligned}$$

Let $T_1(\mathbf{x}) = \min(x_i) = x_{(1)}$ and $T_2(\mathbf{x}) = \max(x_i) = x_{(n)}$. Define $h(\mathbf{x}) = 1$ and

$$g(\theta, \mathbf{t}) = g(\theta, t_1, t_2) = I\Big(t_2 - \frac{1}{2} < \theta < t_1 + \frac{1}{2}\Big).$$

Then $f_\theta(\mathbf{x}) = g(\theta, T_1(\mathbf{x}), T_2(\mathbf{x}))h(\mathbf{x})$. Thus, by the Factorization Theorem, $T(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X}) = (X_{(1)}, X_{(n)})$ is a (two-dimensional) sufficient statistic for θ.
(d)
(i) By hints, for $u < v$,

$$\begin{aligned}
F_{U_{(1)},U_{(n)}}(u,v) &= \mathbf{P}(U_{(1)} \le u, U_{(n)} \le v) = \mathbf{P}(U_{(n)} \le v) - \mathbf{P}(u < U_{(1)} \le U_{(n)} \le v) \\
&= \mathbf{P}(U_i \le v \text{ for all } i = 1, \dots, n) - \mathbf{P}(u < U_i \le v \text{ for all } i = 1, \dots, n) \\
&= \prod_{i=1}^n \mathbf{P}(U_i \le v) - \prod_{i=1}^n \mathbf{P}(u \le U_i \le v) \\
&= \begin{cases} v^n - (v-u)^n, & \text{if } 0 \le u \le v \le 1 \text{ (\textbf{the most interesting case})}; \\ 0, & \text{if } u \le v \le 0; \\ 1^n - 0 = 1, & \text{if } 1 \le u \le v; \\ v^n - v^n = 0, & \text{if } u < 0 \le v \le 1; \\ 1 - (1-u)^n, & \text{if } 0 \le u \le 1 \le v; \\ 1^n - 1^n = 0, & \text{if } u < 0 \le 1 \le v; \end{cases}
\end{aligned}$$

Thus,

$$f_{U_{(1)},U_{(n)}}(u,v) = \frac{\partial^2}{\partial u \partial v} F_{U_{(1)},U_{(n)}}(u,v) = \begin{cases} n(n-1)(v-u)^{n-2}, & \text{if } 0 \le u \le v \le 1; \\ 0, & \text{otherwise.} \end{cases}$$

For the marginal densities of $U_{(1)}$, we can either compute from

$$\begin{aligned}
f_{U_{(1)}}(u) &= \int_\infty^\infty f_{U_{(1)},U_{(n)}}(u,v)dv = \begin{cases} \int_u^1 [n(n-1)(v-u)^{n-2}]dv = n(v-u)^{n-1}|_{v=u}^{v=1}, & \text{if } 0 \le u \le 1; \\ 0, & \text{otherwise.} \end{cases} \\
&= \begin{cases} n(1-u)^{n-1}, & \text{if } 0 \le u \le 1; \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

or directly from

$$\begin{aligned}
F_{U_{(1)}}(u) &= \mathbf{P}(U_{(1)} \le u) = 1 - \mathbf{P}(U_{(1)} > u) = 1 - \mathbf{P}(U_i \ge u \text{ for all } i = 1, \dots, n) \\
&= 1 - \prod_{i=1}^n \mathbf{P}(U_i \ge u) \\
&= \begin{cases} 1 - 1^n = 0, & \text{if } u < 0; \\ 1 - (1-u)^n, & \text{if } 0 \le u \le 1; \\ 1 - 0^n = 1, & \text{if } u > 1. \end{cases} \\
f_{U_{(1)}}(u) &= \frac{\partial}{\partial u} F_{U_{(1)}}(u) = \begin{cases} n(1-u)^{n-1}, & \text{if } 0 \le u \le 1; \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

Similar arguments go for the marginal densities of $U_{(n)}$.
(ii) By definition, we have

$$\mathbf{E}(U_{(1)}) = \int_0^1 n(1-u)^{n-1} u du = \int_0^1 n t^{n-1}(1-t)dt \qquad \text{letting } t = 1-u$$

$$
\begin{aligned}
&= n(\int_0^1 t^{n-1}dt - \int_0^1 t^n dt) = n(\frac{t^n|_{t=0}^{t=1}}{n} - \frac{t^{n+1}|_{t=0}^{t=1}}{n+1}) \\
&= n(\frac{1}{n} - \frac{1}{n+1}) \\
&= \frac{1}{n+1}; \\
\mathbf{E}(U_{(n)}) &= \int_0^1 nv^{n-1}vdv = n\int_0^1 v^n dv = n\frac{v^{n+1}|_{v=0}^{v=1}}{n+1} \\
&= \frac{n}{n+1}; \\
\mathbf{E}(U_{(1)}^2) &= \int_0^1 n(1-u)^{n-1}u^2du = \int_0^1 nt^{n-1}(1-t)^2dt \qquad \text{letting } t = 1-u \\
&= n(\int_0^1 t^{n-1}dt - 2\int_0^1 t^n dt + \int_0^1 t^{n+1}dt) \\
&= n(\frac{1}{n} - \frac{2}{n+1} + \frac{1}{n+2}) \\
&= \frac{2}{(n+1)(n+2)}; \\
\mathbf{E}(U_{(n)}^2) &= \int_0^1 nv^{n-1}v^2dv = n\int_0^1 v^{n+1}dv \\
&= \frac{n}{n+2}; \\
\mathbf{E}(U_{(1)}U_{(n)}) &= \int_0^1 [\int_0^v n(n-1)(v-u)^{n-2}uvdu]dv \\
&= n(n-1)\int_0^1 v[\int_0^v t^{n-2}(v-t)dt]dv \qquad (\text{letting } u = v-t) \\
&= n(n-1)\int_0^1 v[v\frac{t^{n-1}}{n-1} - \frac{t^n}{n}\Big|_{t=0}^{t=v}]dv \\
&= n(n-1)\int_0^1 v[\frac{v^n}{n-1} - \frac{v^n}{n}]dv \\
&= \int_0^1 v^{n+1}dv \\
&= \frac{1}{n+2}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\mathbf{E}(U_{(1)}) &= \frac{1}{n+1}; \\
\mathbf{E}(U_{(n)}) &= \frac{n}{n+1}; \\
Var(U_{(1)}) &= \mathbf{E}(U_{(1)}^2) - (\mathbf{E}(U_{(1)})^2 = \frac{2}{(n+1)(n+2)} - (\frac{1}{n+1})^2 \\
&= \frac{n}{(n+1)^2(n+2)} \\
Var(U_{(n)}) &= \mathbf{E}(U_{(n)}^2) - (\mathbf{E}(U_{(n)})^2 = \frac{n}{n+2} - (\frac{n}{n+1})^2 \\
&= \frac{n}{(n+1)^2(n+2)}
\end{aligned}
$$

$$\begin{aligned} Cov(U_{(1)}, U_{(n)}) &= \mathbf{E}(U_{(1)}U_{(n)}) - \mathbf{E}(U_{(1)})\mathbf{E}(U_{(n)}) = \frac{1}{n+2} - \frac{1}{n+1}\frac{n}{n+1} \\ &= \frac{1}{(n+1)^2(n+2)}. \end{aligned}$$

(iii) Note that

$$R_{\delta_{a,b}}(\theta) = \mathbf{E}_\theta(\delta_{a,b} - \theta)^2 = \mathbf{E}(Y^2) = [\mathbf{E}(Y)]^2 + Var(Y)$$

where

$$\begin{aligned} Y &= \delta_{a,b} - \theta = aX_{(1)} + (1-a)X_{(n)} + b - \theta \\ &= a(U_{(1)} + \theta - \frac{1}{2}) + (1-a)[U_{(n)} + \theta - \frac{1}{2}) + b - \theta \qquad \textbf{(since under } \mathbf{E}_\theta\textbf{)} \\ &= aU_{(1)} + (1-a)U_{(n)} + b - \frac{1}{2} \end{aligned}$$

Now by (ii),

$$\mathbf{E}(Y) = a\mathbf{E}(U_{(1)}) + (1-a)\mathbf{E}(U_{(n)}) + b - \frac{1}{2} = a\frac{1}{n+1} + (1-a)\frac{n}{n+1} + b - \frac{1}{2}$$

and

$$\begin{aligned} Var(Y) &= Var(aU_{(1)} + (1-a)U_{(n)} + b - \frac{1}{2}) = Var(aU_{(1)} + (1-a)U_{(n)}) \\ &= a^2Var(U_{(1)}) + (1-a)^2Var(U_{(n)}) + 2a(1-a)Cov(U_{(1)}, U_{(n)}) \\ &= \frac{a^2n + (1-a)^2n + 2a(1-a)}{(n+1)^2(n+2)} \end{aligned}$$

Hence, the risk function of $\delta_{a,b}(\mathbf{X})$ is

$$R_{\delta_{a,b}}(\theta) = [\mathbf{E}(Y)]^2 + Var(Y) = \Big[a\frac{1}{n+1} + (1-a)\frac{n}{n+1} + b - \frac{1}{2}\Big]^2 + \frac{a^2n + (1-a)^2n + 2a(1-a)}{(n+1)^2(n+2)}.$$

Clearly, for a given a, this is minimized at $b = \frac{1}{2} - \frac{a+n(1-a)}{n+1}$, in which the squared term is zero.
(iv) For any b, by (c)(iii)

$$R_{\delta_{a,b}}(\theta) \geq \frac{a^2n + (1-a)^2n + 2a(1-a)}{(n+1)^2(n+2)},$$

where the numerator of the right-hand side

$$a^2n + (1-a)^2n + 2a(1-a) = 2(n-1)a^2 - 2(n-1)a + n = 2(n-1)(a - \frac{1}{2})^2 + \frac{n+1}{2}$$

is minimized at $a = \frac{1}{2}$ with the minimum value $(n+1)/2$. Hence, for all a, b

$$R_{\delta_{a,b}}(\theta) \geq \frac{a^2n + (1-a)^2n + 2a(1-a)}{(n+1)^2(n+2)} \geq \frac{(n+1)/2}{(n+1)^2(n+2)} = \frac{1}{2(n+1)(n+2)}.$$

(e) Now for $a = \frac{1}{2}$ and $b = \frac{1}{2} - \frac{a+n(1-a)}{n+1} = 0$, this leads to the procedure $\delta^*(\mathbf{X}) = (X_{(1)} + X_{(n)})/2$, with

$$R_{\delta^*}(\theta) = \frac{(n+1)/2}{(n+1)^2(n+2)} = \frac{1}{2(n+1)(n+2)} \leq R_{\delta_{a,b}}(\theta) \text{ for all } \theta \text{ and for any } a \text{ and } b.$$

This shows that *among all procedures* $\delta_{a,b}(\mathbf{X})$, the choice $a = \frac{1}{2}$ and $b = \frac{1}{2} - \frac{a+n(1-a)}{n+1} = 0$, i.e., $\delta^*(\mathbf{X}) = (X_{(1)} + X_{(n)})/2$, gives uniformly smallest risk function.
(f) The above computation shows that

$$R_{\delta^*}(\theta) = \frac{1}{2(n+1)(n+2)} \qquad \text{and} \qquad R_{\delta_0}(\theta) = \frac{1}{12n}$$

Thus for all θ,

$$\frac{R_{\delta_0}(\theta)}{R_{\delta^*}(\theta)} - 1 = \frac{(n+1)(n+2)}{6n} - 1 = \frac{(n-1)(n-2)}{6n},$$

which is positive for $n \geq 3$ (and goes to ∞ as $n \to \infty$). That is, for $n \geq 3$, $R_{\delta_0}(\theta) > R_{\delta^*}(\theta)$ for all θ and thus δ^* is better than $\delta_0(\mathbf{X})$ when $n \geq 3$ (and much better for large n).
It is interesting to note that when $n = 1$ or 2, $\delta^*(\mathbf{X}) = (X_{(1)} + X_{(n)})/2$ is equivalent to $\delta_0(\mathbf{X}) = \bar{X}_n$. □

2. **(Modified from 7.19(a))** Suppose that the random variables $Y_1, \ldots, Y_n (n \geq 2)$ satisfy

$$Y_i = \beta x_i + \epsilon_i, \qquad i = 1, 2, \ldots, n,$$

where $\epsilon_1, \ldots, \epsilon_n$ are iid $N(0, \sigma^2)$, and both β and σ^2 are unknown.
(a) Assume $x_1, \ldots, x_n$ are fixed known constants, and we observe $Y_1 = y_1, \cdots, Y_n = y_n$, e.g., the observed data $\mathbf{Y} = (y_1, \ldots, y_n)$. **Find** a two-dim sufficient statistic of $\mathbf{Y} = (Y_1, \cdots, Y_n)$ for (β, σ^2).
(b) Assume now that $x_1, \ldots, x_n$ are random variables with a known joint distribution $m(x_1, \ldots, x_n)$, and the x_i's are independent of ϵ_i's (it is traditional in the linear regression to use lower case for independent variables x_i's). In this case, the observed data $(\mathbf{Y}, \mathbf{x}) = \{(Y_i, x_i)\}_{i=1,\ldots,n}$. **Find** a three-dimensional sufficient statistic of $(\mathbf{Y}, \mathbf{x})$ for (β, σ^2).

Answer: Let $\theta = (\beta, \sigma^2)$.
(a) As mentioned in the hints, the sample is $\mathbf{Y} = (Y_1, \ldots, Y_n)$, and the joint density function of $\mathbf{Y}$ is

$$\begin{aligned} f_\theta(\mathbf{y}) &= \prod_{i=1}^n f(y_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(- \frac{(y_i - \beta x_i)^2}{2\sigma^2} \Big) \\ &= \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\Big(- \frac{\sum_{i=1}^n (y_i - \beta x_i)^2}{2\sigma^2} \Big) \\ &= \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\Big(- \frac{\sum_{i=1}^n y_i^2 - 2\beta \sum_{i=1}^n x_i y_i + \beta^2 \sum_{i=1}^n x_i^2}{2\sigma^2} \Big). \end{aligned}$$

Let $T_1(\mathbf{y}) = \sum_{i=1}^n y_i^2$ and $T_2(\mathbf{y}) = \sum_{i=1}^n (x_i y_i)$. Define $h(\mathbf{y}) = 1$ and

$$g(\theta, \mathbf{t}) = g((\beta, \sigma), (t_1, t_2)) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\Big(- \frac{t_1 - 2\beta t_2 + \beta^2 \sum_{i=1}^n x_i^2}{2\sigma^2} \Big).$$

Then $f_\theta(\mathbf{y}) = g(\theta, (T_1(\mathbf{y}), T_2(\mathbf{y}))) \times h(\mathbf{y})$. Thus, by the Factorization Theorem, $T(\mathbf{Y}) = (T_1(\mathbf{Y}), T_2(\mathbf{Y})) = (\sum_{i=1}^n Y_i^2, \sum_{i=1}^n (x_i Y_i))$ is sufficient for (β, σ^2).
(b) As mentioned in the hints, when $x_1, \ldots, x_n$ are random variables with a known joint distribution $m(x_1, \ldots, x_n)$, and the x_i's are independent of ϵ_i's, the joint density of the data $(\mathbf{y}, \mathbf{x}) = \{(y_i, x_i)\}_{i=1,\ldots,n}$ is

$$\begin{aligned} f_\theta(\mathbf{y}, \mathbf{x}) &= m(\mathbf{x}) f_\theta(\mathbf{y}|\mathbf{x}) \\ &= m(x_1, \ldots, x_n) \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(- \frac{(y_i - \beta x_i)^2}{2\sigma^2} \Big) \\ &= m(x_1, \ldots, x_n) \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\Big(- \frac{\sum_{i=1}^n y_i^2 - 2\beta \sum_{i=1}^n x_i y_i + \beta^2 \sum_{i=1}^n x_i^2}{2\sigma^2} \Big). \end{aligned}$$

Let $T_1(\mathbf{y},\mathbf{x}) = \sum_{i=1}^n y_i^2$, $T_2(\mathbf{y},\mathbf{x}) = \sum_{i=1}^n (x_i y_i)$, and $T_3(\mathbf{y},\mathbf{x}) = \sum_{i=1}^n x_i^2$, Define $h(\mathbf{y},\mathbf{x}) = m(x_1,\ldots,x_n)$ and

$$g(\theta,\mathbf{t}) = g((\beta,\sigma),(t_1,t_2,t_3)) = \frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\Big(-\frac{t_1 - 2\beta t_2 + \beta^2 t_3}{2\sigma^2}\Big).$$

Then $f_\theta((\mathbf{y},\mathbf{x})) = g(\theta,(T_1(\mathbf{y},\mathbf{x}),T_2(\mathbf{y},\mathbf{x}),T_3(\mathbf{y},\mathbf{x}))) \times h(\mathbf{y},\mathbf{x})$. Thus, by the Factorization Theorem, $T(\mathbf{Y},\mathbf{X}) = (T_1(\mathbf{Y},\mathbf{X}),T_2(\mathbf{Y},\mathbf{X}),T_3(\mathbf{Y},\mathbf{X})) = (\sum_{i=1}^n Y_i^2, \sum_{i=1}^n (x_i Y_i), \sum_{i=1}^n x_i^2)$ is sufficient for (β,σ^2). □

3. **(Modified from 6.5)**. Let $X_1,\cdots,X_n (n \geq 2)$ be independent random variables with pdfs

$$f(x_i|\theta) = \begin{cases} \frac{1}{3i\theta}, & \text{if } -i(\theta-1) < x_i < i(2\theta+1); \\ 0, & \text{otherwise,} \end{cases}$$

for $i = 1,2,\cdots,n$, where $\theta > 0$.

(a) Show that $T_a(\mathbf{X}) = (\min_{1\leq i\leq n}(X_i/i), \max_{1\leq i\leq n}(X_i/i))$ is a two-dim sufficient statistic for θ.

(b) Find a minimal sufficient statistic for θ. Hints: the minimal sufficient statistic is one-dimensional.

Answer: (a) The joint pdf of $\mathbf{X} = (X_1,\ldots,X_n)$ is

$$\begin{aligned}
f_\theta(\mathbf{x}) &= \prod_{i=1}^n f_{X_i}(x_i|\theta) = \prod_{i=1}^n \Big[\frac{1}{3i\theta} I(-i(\theta-1) < x_i < i(2\theta+1)\Big] \\
&= \frac{1}{3^n n! \theta^n} I\Big(-(\theta-1) < \frac{x_i}{i} < 2\theta+1 \quad \text{for all } i = 1,\ldots,n\Big) \\
&= \frac{1}{3^n n! \theta^n} \times I\Big(-(\theta-1) < \frac{x_i}{i} \quad \text{for all } i = 1,\ldots,n\Big) \times I\Big(\frac{x_i}{i} < 2\theta+1 \quad \text{for all } i = 1,\ldots,n\Big) \\
&= \frac{1}{3^n n! \theta^n} \times I\Big(-(\theta-1) < \min_{1\leq i\leq n}\frac{x_i}{i}\Big) \times I\Big(\max_{1\leq i\leq n}\frac{x_i}{i} < 2\theta+1\Big) \\
&= \frac{1}{3^n n! \theta^n} \times I\Big(\theta > 1 - (\min_{1\leq i\leq n}\frac{x_i}{i})\Big) \times I\Big(\theta > \frac{1}{2}[(\max_{1\leq i\leq n}\frac{x_i}{i}) - 1]\Big)
\end{aligned}$$

Let $T_1(\mathbf{x}) = \min_{1\leq i\leq n}(x_i/i)$ and $T_2(\mathbf{x}) = \max_{1\leq i\leq n}(x_i/i)$. Define $h(\mathbf{x}) = 1$ and

$$g(\theta,\mathbf{t}) = g(\theta,t_1,t_2) = \frac{1}{3^n n! \theta^n} \times I\Big(\theta > 1 - t_1\Big) \times I\Big(\theta > \frac{1}{2}(t_2 - 1)\Big).$$

Then $f_\theta(\mathbf{x}) = g(\theta,T_1(\mathbf{x}),T_2(\mathbf{x}))h(\mathbf{x})$. Thus, by the Factorization Theorem, $T_a(\mathbf{X}) = (T_1(\mathbf{X}),T_2(\mathbf{X}) = (\min_{1\leq i\leq n}(X_i/i), \max_{1\leq i\leq n}(X_i/i))$ is a two-dimensional sufficient statistic for θ.
Note that the sufficient statistic is not unique! In fact, this two-dimensional sufficient statistic is not minimal sufficient.
Remark: a typical mistake is to think that

$$T_1(\mathbf{X}) = \min_{1\leq i\leq n}(X_i/i) = \min(\frac{X_1}{1},\frac{X_2}{2},\frac{X_3}{3},\cdots,\frac{X_n}{n})$$

which be written in the form something like $\frac{1}{i}\min_{1\leq i\leq n}(X_i)$ or similar. This is incorrect because the index i is not a constant. In this case, it is fine to keep the form of T or T_2 as is.
(b) **The key observation is** to note that $I(\theta > u)I(\theta > v) = I(\theta > \max(u,v))$, and thus the joint density function $f(\mathbf{x}|\theta)$ in part (a) can be further rewritten as

$$\begin{aligned}
f(\mathbf{x}|\theta) &= \frac{1}{3^n n! \theta^n} \times I\Big(\theta > 1 - (\min_{1\leq i\leq n}\frac{x_i}{i})\Big) \times I\Big(\theta > \frac{1}{2}[(\max_{1\leq i\leq n}\frac{x_i}{i}) - 1]\Big) \\
&= \frac{1}{3^n n! \theta^n} \times I\Big(\theta > \max\Big\{1 - \min_{1\leq i\leq n}\frac{x_i}{i}, \quad \frac{1}{2}(\max_{1\leq i\leq n}\frac{x_i}{i} - 1)\Big\}\Big).
\end{aligned}$$

If we define

$$T(\mathbf{X}) = \max\Big\{1 - \min_{1\le i\le n} \frac{x_i}{i}, \quad \frac{1}{2}(\max_{1\le i\le n} \frac{x_i}{i} - 1)\Big\},$$

(it is okay that we do not need to simplify T here), then

$$f(\mathbf{x}|\theta) = \frac{1}{3^n n! \theta^n} \times I\Big(\theta > T(\mathbf{X})\Big) = g(\theta, T(\mathbf{X}))h(\mathbf{X}),$$

where $h(\mathbf{X}) = 1$ and $g(\theta, t) = \frac{1}{2^n n! \theta^n} \times I\Big(\theta > t\Big)$. Hence, by the Factorization theorem, this $T(\mathbf{X})$ is also sufficient for θ, and it is a one-dimensional sufficient statistic!
Given two sample points, $\mathbf{x}$ and $\mathbf{y}$,

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\frac{1}{3^n n!\theta^n} \times I\Big(\theta > T(\mathbf{x})\Big)}{\frac{1}{3^n n!\theta^n} \times I\Big(\theta > T(\mathbf{y})\Big)} = \frac{I\Big(\theta > T(\mathbf{x})\Big)}{I\Big(\theta > T(\mathbf{y})\Big)}$$

This will be constant as a function of θ if and only if $T(\mathbf{x}) = T(\mathbf{y})$. Thus the statistic

$$T(\mathbf{X}) = \max\Big\{1 - \min_{1\le i\le n} \frac{x_i}{i}, \quad \frac{1}{2}(\max_{1\le i\le n} \frac{x_i}{i} - 1)\Big\},$$

is minimal sufficient for θ.
Remark: this minimal sufficient statistic is unique (in the sense that there is a one-to-one map with any other minimal sufficient statistic). Moreover, while this minimal sufficient is defined as the maximum of two other statistics, itself is one-dimensional, since for a given observed data set, we can compute a single numerical value of this statistic! □

4. **(6.25) (b) and (d)**. We have seen a number of theorems concerning sufficiency and related concepts for exponential families. Let $X_1, \dots, X_n (n \ge 2)$ be a random sample for each of the following distribution families, and establish the following results.

   (b) The statistic $T(\mathbf{X}) = \sum_{i=1}^n X_i^2$ is minimal sufficient in the $N(\mu, \mu)$ family.

   (d) The statistic $T(\mathbf{X}) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ is minimal sufficient for $\theta = (\mu, \sigma^2)$ in the $N(\mu, \sigma^2)$ family.

Answer: (b) Given two sample points, $\mathbf{x}$ and $\mathbf{y}$,

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\mu}} e^{-(x_i-\mu)^2/2\mu}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\mu}} e^{-(y_i-\mu)^2/2\mu}} = \exp\Big\{\frac{\sum_{i=1}^n y_i^2 - \sum_{i=1}^n x_i^2}{2\mu} + \Big(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\Big)\Big\}.$$

This will be constant as a function of μ if and only if $\sum x_i^2 = \sum y_i^2$. Thus the statistic $T(\mathbf{X}) = \sum_{i=1}^n X_i^2$ is a minimal sufficient statistic.
(d) Given two sample points, $\mathbf{x}$ and $\mathbf{y}$,

$$\frac{f(\mathbf{x}|\mu, \sigma^2)}{f(\mathbf{y}|\mu, \sigma^2)} = \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_i-\mu)^2/2\sigma^2}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y_i-\mu)^2/2\sigma^2}} = \exp\Big\{\frac{\sum_{i=1}^n y_i^2 - \sum_{i=1}^n x_i^2}{2\sigma^2} + \frac{\mu}{\sigma^2}\Big(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\Big)\Big\}.$$

This will be constant as a function of μ and σ^2 if and only if $\sum x_i = \sum y_i$ and $\sum x_i^2 = \sum y_i^2$. Thus the statistic $T(\mathbf{X}) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ is a minimal sufficient statistic. □

5. **(6.9)(a)(b)(d)(e).** For each of the following distribution let $X_1, \ldots, X_n (n \geq 2)$ be a random sample. **Find** a minimal sufficient statistic for θ.

   (a) $f(x|\theta) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}, \quad -\infty < x < \infty, -\infty < \theta < \infty \qquad$ (normal)

   (b) $f(x|\theta) = e^{-(x-\theta)}, \quad \theta < x < \infty, -\infty < \theta < \infty \qquad$ (location exponential)

   (d) $f(x|\theta) = \frac{1}{\pi[1+(x-\theta)^2]}, \quad -\infty < x < \infty, -\infty < \theta < \infty \qquad$ (Cauchy)

   (e) $f(x|\theta) = \frac{1}{2} e^{-|x-\theta|}, \quad -\infty < x < \infty, -\infty < \theta < \infty \qquad$ (double exponential)

   [In class we will discuss part (c) $f(x|\theta) = \frac{e^{-(x-\theta)}}{(1+e^{-(x-\theta)})^2}, -\infty < x < \infty, -\infty < \theta < \infty$(logistic).]

   Answer: The key idea is to use Theorem 6.2.13 to find a minimal sufficient statistic. Given two sample points, $\mathbf{x}$ and $\mathbf{y}$, and calculate the ratio of densities:
   (a)

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-(x_i-\theta)^2/2}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-(y_i-\theta)^2/2}} = \exp\Big\{ \frac{\sum_{i=1}^n y_i^2 - \sum_{i=1}^n x_i^2}{2} + \theta\Big(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\Big)\Big\}.$$

   This will be constant as a function of θ if and only if $\sum x_i = \sum y_i$. Thus the statistic $T(\mathbf{X}) = \sum_{i=1}^n X_i$ is a minimal sufficient statistic.
   (b) It is important to note that the range of X depends on θ. Now

$$\begin{aligned} \frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} &= \frac{\prod_{i=1}^n [e^{-(x_i-\theta)} I(\theta < x_i)]}{\prod_{i=1}^n [e^{-(x_i-\theta)} I(\theta < x_i)]} = \frac{\exp(n\theta - \sum_{i=1}^n x_i) I(\theta < x_i \text{ for all } i)}{\exp(n\theta - \sum_{i=1}^n y_i) I(\theta < y_i \text{ for all } i)} \\ &= \exp(\sum_{i=1}^n y_i - \sum_{i=1}^n x_i) \times \frac{I(\theta < \min x_i)}{I(\theta < \min y_i)}. \end{aligned}$$

   This ratio is independent of θ if and only if $\min x_i = \min y_i$. So $T(\mathbf{X}) = \min(X_1, \ldots, X_n) = X_{(1)}$ is a minimal sufficient statistic.
   (c) This has been discussed in class, and the order statistic is a minimal sufficient statistic.
   (d) It is easy to see that

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\prod_{k=1}^n \frac{1}{\pi(1+(x_k-\theta)^2)}}{\prod_{k=1}^n \frac{1}{\pi(1+(y_k-\theta)^2)}} = \frac{\prod_{k=1}^n (1+(y_k-\theta)^2)}{\prod_{k=1}^n (1+(x_k-\theta)^2)}.$$

   Now

$$\begin{aligned} \frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} \text{ is constant in } \theta \quad &\Longleftrightarrow \quad \frac{f(\mathbf{x}|\theta)}{f(\mathbf{x}|\theta=0)} = \frac{f(\mathbf{y}|\theta)}{f(\mathbf{y}|\theta=0)} \quad \text{for all } \theta \\ &\Longleftrightarrow \quad \prod_{k=1}^n \frac{1+(x_k-\theta)^2}{1+x_k^2} = \prod_{k=1}^n \frac{1+(y_k-\theta)^2}{1+y_k^2} \quad \text{for all } \theta. \end{aligned}$$

   Now both sides are polynomial of θ of degree $2n$, comparing the coefficient of θ^{2n} yields that $\prod(1+x_k^2) = \prod(1+y_k^2)$, and thus

$$\prod_{k=1}^n [1+(x_k-\theta)^2] = \prod_{k=1}^n [1+(y_k-\theta)^2].$$

   Setting these two polynomials to 0 and solving the complex root for θ, the left-hand side polynomial has $2n$ complex roots, $\widehat{\theta} = x_k \pm \sqrt{-1}$, for $k = 1, \ldots, n$, whereas the right-hand polynomial leads to another set of $2n$ complex roots, $\widehat{\theta} = y_k \pm \sqrt{-1}$, for $k = 1, \ldots, n$. Of course these two polynomials in θ will have the same (complex) roots, and thus $x_{(k)} = y_{(k)}$ for $k = 1, \ldots, n$. Hence, the order statistic is

a minimal sufficient statistic here.
(e) In this case, the order statistic is also a minimal sufficient statistic. The proof is tedious but straightforward. To see this, note that

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\prod_{i=1}^n e^{-|x_i-\theta|}}{\prod_{i=1}^n e^{-|y_i-\theta|}} = \frac{\prod_{i=1}^n e^{-|x_{(i)}-\theta|}}{\prod_{i=1}^n e^{-|y_{(i)}-\theta|}} = \exp\Big\{\sum_{i=1}^n |y_{(i)}-\theta| - \sum_{i=1}^n |x_{(i)}-\theta|\Big\}.$$

Clearly, if the x's and y's have the same order statistic, then $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = 1$ does not depend on θ. On the other hand, if $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)}$ does not depend on θ, we will prove that $x_{(i)} = y_{(i)}$ for all $i = 1, \ldots, n$.
First, let us prove $x_{(1)} = y_{(1)}$. Assume $x_{(1)} \neq y_{(1)}$, and without loss of generality, assume $x_{(1)} < y_{(1)}$. For convenience of notation, define $x_{(0)} = y_{(0)} = -\infty$ and define $x_{(n+1)} = y_{(n+1)} = \infty$. Now let r be the largest $i \geq 1$ such that $x_{(i)} < y_{(1)}$. In other words, $x_{(1)} \leq x_{(r)} < y_{(1)} \leq x_{(r+1)}$ for some $1 \leq r \leq n$. Consider the interval $x_{(r)} < \theta < y_{(1)}$, we have (as conventional $\sum_{i=n+1}^n = 0$ below)

$$\begin{aligned}
\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} &= \exp\Big\{\big[\sum_{i=1}^r (y_{(i)}-\theta) + \sum_{i=r+1}^n (y_{(i)}-\theta)\big] - \big[\sum_{i=1}^r (\theta - x_{(i)}) + \sum_{i=r+1}^n (x_{(i)}-\theta)\big]\Big\} \\
&= \exp\Big\{\sum_{i=1}^r (y_{(i)} + x_{(i)}) - 2r\theta + \sum_{i=r+1}^n (y_{(i)} - x_{(i)})\Big\},
\end{aligned}$$

which depends on $\theta \in (x_{(r)}, y_{(1)})$ since $1 \leq r \leq n$. This is a contradiction that $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)}$ is a constant of θ. Thus the assumption that $x_{(1)} \neq y_{(1)}$ is wrong, and hence we must have $x_{(1)} = y_{(1)}$.
The above arguments can be easily extended to show that $x_{(i)} = y_{(i)}$ for all $i = 1, \ldots, n$. Assume this is not true, and let k be the smallest i such that $x_{(i)} \neq y_{(i)}$, say $x_{(k)} < y_{(k)}$. As above, let r be the largest $i \geq k$ such that $x_{(i)} < y_{(k)}$. Then

$$x_{(1)} = y_{(1)} \leq x_{(2)} = y_{(2)} \leq \cdots \leq x_{(k-1)} = y_{(k-1)} \leq x_{(k)} \leq x_{(r)} < y_{(k)}$$

for some $k \leq r \leq n$. Consider the interval $x_{(r)} < \theta < y_{(k)}$,

$$\begin{aligned}
\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} &= \exp\Big\{\sum_{i=k}^n |y_{(i)}-\theta| - \sum_{i=k}^n |x_{(i)}-\theta|\Big\} \qquad (\text{since } x_{(i)} = y_{(i)} \text{ for } i \leq k-1) \\
&= \exp\Big\{\big[\sum_{i=k}^r (y_{(i)}-\theta) + \sum_{i=r+1}^n (y_{(i)}-\theta)\big] - \big[\sum_{i=k}^r (\theta - x_{(i)}) + \sum_{i=r+1}^n (x_{(i)}-\theta)\big]\Big\} \\
&= \exp\Big\{\sum_{i=k}^r (y_{(i)} + x_{(i)}) - 2(r-k+1)\theta + \sum_{i=r+1}^n (y_{(i)} - x_{(i)})\Big\},
\end{aligned}$$

which clearly depends on θ since $r - k + 1 \geq 1$ as $r \geq k$. Thus, such k does not exist, and hence $x_{(i)} = y_{(i)}$ for all i. □

6. **(6.12).** A natural ancillary statistic in most problems in the *sample size*. For example, let N be an integer-valued random variable taking values $1, 2, \cdots$ with known probabilities $p_1, p_2, \cdots$, where $\sum_{i=1}^\infty p_i = 1$. Having observed $N = n$, perform n Bernoulli trials with success probability θ, getting X successes.

   (a) Prove that the pair (X, N) is minimal sufficient and N is ancillary for θ.
   (Note that the similarity to some of the hierarchical models discussed in Section 4.4.)

   Answer: By definition,

$$P(N = n) = P(N = n|\theta) = p_n$$

does not depend on θ, and thus N is ancillary for θ. To find the sufficient statistic, note that

$$P(X=x,N=n|\theta)=P(N=n)P(X=x|N=n,\theta)=p_n\binom{n}{x}\theta^x(1-\theta)^{n-x}.$$

To show that (X,N) is minimal sufficient, note that for any (x,n) and (y,m),

$$\begin{aligned}
&\frac{P(X=x,N=n|\theta)}{P(X=y,N=m|\theta)} \quad \text{is a constant for } \theta\in(0,1)\\
\iff\quad & \frac{P(X=x,N=n|\theta)}{P(X=x,N=n|\theta=1/2)}=\frac{P(X=y,N=m|\theta)}{P(X=y,N=m|\theta=1/2)} \qquad \text{for all } \theta\in(0,1)\\
\iff\quad & (2\theta)^x\Big(2(1-\theta)\Big)^{n-x}=(2\theta)^y\Big(2(1-\theta)\Big)^{m-y} \qquad \text{for all } \theta\in(0,1)\\
\iff\quad & (2\theta)^{x-y}=\Big(2(1-\theta)\Big)^{(m-y)-(n-x)} \qquad \text{for all } \theta\in(0,1)\\
\iff\quad & x-y=0 \text{ and } (m-y)-(n-x)=0 \qquad \text{(Why? see what happens if } \theta\to 0 \text{ or } 1)\\
\iff\quad & x=y \text{ and } n=m.
\end{aligned}$$

This implies that (X,N) is a minimal sufficient statistic for θ. □

(b) Prove that the estimator X/N is unbiased for θ and has variance $\theta(1-\theta)\mathbf{E}(1/N)$. In other words, prove that $\mathbf{E}_\theta(X/N)=\theta$ and $Var_\theta(X/N)=\theta(1-\theta)\mathbf{E}(1/N)$.

Answer: Using the hints,

$$\begin{aligned}
\mathbf{E}(\frac{X}{N}) &= \mathbf{E}\Big(\mathbf{E}(\frac{X}{N}|N)\Big)=\mathbf{E}\Big(\frac{1}{N}\mathbf{E}(X|N)\Big)\\
&= \mathbf{E}(\frac{1}{N}\times\theta N)=\theta. \qquad \text{So } X/N \text{ is unbiased for } \theta.\\
Var(\frac{X}{N}) &= \mathbf{E}\Big(Var(\frac{X}{N}|N)\Big)+Var\Big(\mathbf{E}(\frac{X}{N}|N)\Big)\\
&= \mathbf{E}\Big(\frac{\theta(1-\theta)}{N}|N\Big)+Var\Big(\theta\Big)\\
&= \mathbf{E}(\frac{\theta(1-\theta)}{N})=\theta(1-\theta)\mathbf{E}(\frac{1}{N}).
\end{aligned}$$

□

Hints of Problem 1 (d): To compute its risk function, it is useful to split in the following steps.

(i) Note that if we let $U_i = X_i - \theta + 1/2$, then $X_{(1)} = U_{(1)} + \theta - 1/2$ and $X_{(n)} = U_{(n)} + \theta - 1/2$. Hence we first need to investigate the properties of $U_{(1)} = \min(U_1, \ldots, U_n)$ and $U_{(n)} = \max(U_1, \ldots, U_n)$ when $U_1, \ldots, U_n$ are iid with Uniform$[0, 1]$. Using the fact $\mathbf{P}(u \leq U_{(1)} \leq U_{(n)} \leq v) = \mathbf{P}(u \leq U_i \leq v \text{ for all } i = 1, \ldots, n) = \prod_{i=1}^{n} \mathbf{P}(u \leq U_i \leq v)$ for any u and v, show that the joint density of $U_{(1)}$ and $U_{(n)}$ is

$$f_{U_{(1)},U_{(n)}}(u, v) = \begin{cases} n(n-1)(v-u)^{n-2}, & \text{if } 0 \leq u \leq v \leq 1; \\ 0, & \text{otherwise.} \end{cases}$$

whereas the respective (marginal) densities of $U_{(1)}$ and $U_{(n)}$ are

$$f_{U_{(1)}}(u) = \begin{cases} n(1-u)^{n-1}, & \text{if } 0 \leq u \leq 1; \\ 0, & \text{otherwise.} \end{cases} \quad \text{and} \quad f_{U_{(n)}}(v) = \begin{cases} nv^{n-1}, & \text{if } 0 \leq v \leq 1; \\ 0, & \text{otherwise.} \end{cases}$$

(ii) Show that $\mathbf{E}(U_{(1)}) = \frac{1}{n+1}$, $\mathbf{E}(U_{(n)}) = \frac{n}{n+1}$, $Var(U_{(1)}) = Var(U_{(n)}) = \frac{n}{(n+1)^2(n+2)}$ and $Cov(U_{(1)}, U_{(n)}) = \frac{1}{(n+1)^2(n+2)}$.

(iii) Use the fact of $\mathbf{E}(Y^2) = [\mathbf{E}(Y)]^2 + Var(Y)$ to show that the risk function of $\delta_{a,b}(\mathbf{X})$ is

$$R_{\delta_{a,b}}(\theta) \quad = \quad \mathbf{E}\Big(aU_{(1)} + (1-a)U_{(n)} + b - 1/2\Big)^2.$$

Hints of Problem 2: Let $\theta = (\beta, \sigma^2)$.

(a) The sample is $\mathbf{Y} = (Y_1, \ldots, Y_n)$, and the joint density function of $\mathbf{Y}$ is

$$f_\theta(\mathbf{Y}) = \prod_{i=1}^{n} f(Y_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(-\frac{(Y_i - \beta x_i)^2}{2\sigma^2}\Big).$$

How to factor this joint pdf into two parts? The part that depends on $\theta = (\beta, \sigma^2)$ depends on the sample $\mathbf{Y} = (Y_1, \ldots, Y_n)$ only through which kind of two-dimensional function $T(\mathbf{Y})$? Note that the x_i's are treated as known constants here.

(b) When $x_1, \ldots, x_n$ are random variables with a known joint distribution $m(x_1, \ldots, x_n)$, and the x_i's are independent of ϵ_i's, the joint density of the data $(\mathbf{Y}, \mathbf{X}) = \{(Y_i, x_i)\}_{i=1,\ldots,n}$ is

$$f_\theta(\mathbf{Y}, \mathbf{X}) = m(\mathbf{x}) f_\theta(\mathbf{Y}|\mathbf{X}) = m(x_1, \ldots, x_n) \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(-\frac{(Y_i - \beta x_i)^2}{2\sigma^2}\Big).$$

Can you factor this joint pdf into two parts? The part that depends on $\theta = (\beta, \sigma^2)$ depends on the sample $(\mathbf{Y}, \mathbf{X}) = \{(Y_i, x_i)\}_{i=1,\ldots,n}$ only through which kind of three-dimensional function $T(\mathbf{Y}, \mathbf{X})$?

Hints of Problem 3: It is important to focus on the domain of θ in the joint pdf of $\mathbf{X} = (X_1, \ldots, X_n)$. You can write $a(\theta) < x_i < b(\theta)$ for $i = 1, \cdots, n$, into two separate inequalities: $a(\theta) < x_i$ for all i and $x_i < b(\theta)$ for all i. From this, we can conclude that $a(\theta) < \min_i x_i$ and $\max_i x_i < b(\theta)$, and then solve for θ, respectively. To be more specific, the joint density is

$$\begin{aligned} f_\theta(\mathbf{x}) &= \prod_{i=1}^{n} f_{X_i}(x_i|\theta) = \prod_{i=1}^{n} \Big[\frac{1}{3i\theta} I(-i(\theta-1) < x_i < i(2\theta+1)\Big] \\ &= \frac{1}{3^n n! \theta^n} I\Big(-(\theta-1) < \frac{x_i}{i} < 2\theta+1 \quad \text{for all } i = 1, \ldots, n\Big) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{3^n n! \theta^n} \times I\Big(-(\theta-1) < \frac{x_i}{i} \quad \text{for all } i=1,\ldots,n \Big) \times I\Big(\frac{x_i}{i} < 2\theta+1 \quad \text{for all } i=1,\ldots,n \Big) \\
&= \frac{1}{3^n n! \theta^n} \times I\Big(-(\theta-1) < \min_{1\le i\le n} \frac{x_i}{i} \Big) \times I\Big(\max_{1\le i\le n} \frac{x_i}{i} < 2\theta+1 \Big) \\
&= \frac{1}{3^n n! \theta^n} \times I\Big(\theta > 1 - (\min_{1\le i\le n} \frac{x_i}{i}) \Big) \times I\Big(\theta > \frac{1}{2}[(\max_{1\le i\le n} \frac{x_i}{i}) - 1] \Big)
\end{aligned}
$$

Part (a) follows from this immediately. To find the minimal sufficient statistic in part (b), using the fact that $I(\theta > u)I(\theta > v) = I(\theta > \max(u,v))$, you can further simplify the above density function as a function of one-dimensional statistic. Hint: how about us defining

$$T(\mathbf{X}) = \max\Big\{ 1 - \min_{1\le i\le n} \frac{x_i}{i}, \quad \frac{1}{2}(\max_{1\le i\le n} \frac{x_i}{i} - 1) \Big\}.$$

Also we do not need to simplify T here and it is okay to leave it as is.

Hints of Problem 5(d): The key observation is that

$$
\begin{aligned}
\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} \text{ is constant in } \theta \quad &\Longleftrightarrow \quad \frac{f(\mathbf{x}|\theta)}{f(\mathbf{x}|\theta=0)} = \frac{f(\mathbf{y}|\theta)}{f(\mathbf{y}|\theta=0)} \quad \text{for all } \theta \\
&\Longleftrightarrow \quad \prod_{k=1}^{n} \frac{1+(x_k-\theta)^2}{1+x_k^2} = \prod_{k=1}^{n} \frac{1+(y_k-\theta)^2}{1+y_k^2} \quad \text{for all } \theta.
\end{aligned}
$$

Now both sides are polynomial of θ of degree $2n$, comparing the coefficient of θ^{2n} yields that $\prod(1+x_k^2) = \prod(1+y_k^2)$, and thus

$$\prod_{k=1}^{n} [1+(x_k-\theta)^2] = \prod_{k=1}^{n} [1+(y_k-\theta)^2].$$

Setting these two polynomials to 0 and solving the complex root for θ, the left-hand side polynomial has $2n$ complex roots, $\widehat{\theta} = x_k \pm \sqrt{-1}$, for $k=1,\ldots,n$, whereas the right-hand polynomial leads to another set of $2n$ complex roots, $\widehat{\theta} = y_k \pm \sqrt{-1}$, for $k=1,\ldots,n$. Of course these two polynomials in θ will have the same (complex) roots, and thus $x_{(k)} = y_{(k)}$ for $k=1,\ldots,n$. What does this mean?

Hints of Problem 5(e): In this case, the order statistic is also a minimal sufficient statistic. the main difficulty is to show that if $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)}$ does not depend on θ, then $x_{(i)} = y_{(i)}$ for all $i=1,\ldots,n$.
First, let us prove $x_{(1)} = y_{(1)}$. Assume $x_{(1)} \neq y_{(1)}$, and without loss of generality, assume $x_{(1)} < y_{(1)}$. For convenience of notation, define $x_{(0)} = y_{(0)} = -\infty$ and define $x_{(n+1)} = y_{(n+1)} = \infty$. Now let r be the largest $i \ge 1$ such that $x_{(i)} < y_{(1)}$. In other words, $x_{(1)} \le x_{(r)} < y_{(1)} \le x_{(r+1)}$ for some $1 \le r \le n$. Consider the interval $x_{(r)} < \theta < y_{(1)}$, and show that $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)}$ depends on $\theta \in (x_{(r)}, y_{(1)})$ since $1 \le r \le n$. This is a contradiction that $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)}$ is a constant of θ. Thus the assumption that $x_{(1)} \neq y_{(1)}$ is wrong, and hence we must have $x_{(1)} = y_{(1)}$.
The above arguments can be easily extended to show that $x_{(i)} = y_{(i)}$ for all $i=1,\ldots,n$. Assume this is not true, and let k be the smallest i such that $x_{(i)} \neq y_{(i)}$, say $x_{(k)} < y_{(k)}$. As above, let r be the largest $i \ge k$ such that $x_{(i)} < y_{(k)}$. Then

$$x_{(1)} = y_{(1)} \le x_{(2)} = y_{(2)} \le \cdots \le x_{(k-1)} = y_{(k-1)} \le x_{(k)} \le x_{(r)} < y_{(k)}$$

for some $k \le r \le n$. Then consider the interval $x_{(r)} < \theta < y_{(k)}$, and see what happens?

Hints of Problem 6(b): Use the facts that $\mathbf{E}(U) = \mathbf{E}(\mathbf{E}(U|V))$ and $Var(U) = \mathbf{E}(Var(U|V)) + Var(\mathbf{E}(U|V))$ for $U = X/N$ and $V = N$. See Theorems 4.4.3 and 4.4.7 on page 164-167 of our text for the proofs of these two useful facts which will be used later.