EX 3.1 The following problems are selected from the book (problem)

Proof: $T(x)$ is sufficient for θ in an experiment E iff the model $P_\theta(x)$ can be factorized in the form $p_\theta(x) = g(t(x), \theta)\, h(x)$ where $h(x)$ is free of θ

Sufficiency $\Rightarrow$

If $T(x)$ is sufficient for θ $\stackrel{\text{Def}}{\Longrightarrow}$ $P(X=x \mid T=t)$ is free of θ

$\Rightarrow P(X=x \mid T=t) \stackrel{\Delta}{=} h(x)$

$P(T=t)$ is a function involves T and θ

$\Rightarrow P(T=t) \stackrel{\Delta}{=} g(t(x), \theta)$

Thus,

$$P(X=x) = P(X=x \mid T=t)\, P(T=t)$$
$$= h(x) \cdot g(t(x), \theta)$$

Necessity $\Leftarrow$

If $p(x) = g(t(x), \theta) \cdot h(x)$

$$P(X=x \mid T=t) = \begin{cases} 0 & \text{when } T(x) \neq t \\ \dfrac{P(X=x, T=t)}{P(T(x)=t)} = \dfrac{P(X=x)}{P(T(x)=t)} & (*) \end{cases}$$

$$(*) = \frac{P(X=x)}{\sum_{T(x)=t} P(X=x)} = \frac{g(t,\theta)\cdot h(x)}{g(t,\theta)\sum_{T(x)=t} h(x)} = \frac{h(x)}{\sum_{T(x)=t} h(x)}$$

So. $P(X=x \mid T=t) = \begin{cases} 0 & \text{when } T(x) \neq t \\ \dfrac{h(x)}{\sum_{T(x)=t} h(x)} \end{cases}$ is free of θ.

QED

Ex3.2. problem2

(a) $X=(X_1, X_2, \cdots X_n) \qquad x=(x_1, x_2 \cdots x_n)$

$$f(x) = \prod_{i=1}^{n} f(x_i) = \left(\frac{1}{\theta}\right)^n \cdot \prod_{i=1}^{n} \mathbb{I}\{0<x_i<\theta\}.$$

$$= \left(\frac{1}{\theta}\right)^n \cdot \mathbb{I}\{0<x_{(1)}<x_{(2)}<\cdots<x_{(n)}<\theta\}$$

$$= \left(\frac{1}{\theta}\right)^n \cdot \mathbb{I}\{0<x_{(1)}\} \underbrace{\mathbb{I}\{x_{(n)}<\theta\}}_{g(x_{(n)},\, \theta).}$$

By Factorization Thm. $\quad T(X) = X^{(n)}$

(b). $$f(x) = \prod_{i=1}^{n} f(x_i) = \left(\frac{1}{2}\right)^n \prod_{i=1}^{n} \mathbb{I}\{\theta-1<x_i<\theta+1\}$$

$$= \frac{1}{2^n} \prod_{i=1}^{n} \mathbb{I}\{\theta-1<x_{(1)}<x_{(2)}<\cdots<x_{(n)}<\theta+1\}$$

$$= \frac{1}{2^n} \mathbb{I}\{\theta-1<x_{(1)}<x_{(n)}<\theta+1\}$$

$$= \frac{1}{2^n} \cdot \underbrace{\mathbb{I}\{x_{(n)}-1<\theta<x_{(1)}+1\}}_{g(T,\theta)} \qquad T \triangleq (x_{(1)}, x_{(n)})$$

By Factorization Thm. $\quad T(X) = (X^{(1)}, X^{(n)})$

(c) $$f(x) = \prod_{i=1}^{n} f(x_i) = \left(\frac{1}{\theta_2-\theta_1}\right)^n \cdot \prod_{i=1}^{n} \mathbb{I}\{\theta_1 \le x_i \le \theta_2\}$$

$$= \left(\frac{1}{\theta_2-\theta_1}\right)^n \cdot \mathbb{I}\{\theta_1 \le x_{(1)} \le \cdots \le x_{(n)} \le \theta_2\}$$

$$= \left(\frac{1}{\theta_2-\theta_1}\right)^n \underbrace{\mathbb{I}\{\theta_1 \le x_{(1)}\}\, \mathbb{I}\{x_{(n)} \le \theta_2\}}_{g(T,\, \binom{\theta_1}{\theta_2})} \qquad T \triangleq (x_{(1)}, x_{(n)})$$

By Factorization Thm. $\quad T(X) = (X^{(1)}, X^{(n)})$

(d) $$f(x) = \prod_{i=1}^{n} f(x_i) = \left(\frac{1}{2\theta}\right)^n \prod_{i=1}^{n} \mathbb{I}\{-\theta<x_i<\theta\}$$

$$= \left(\frac{1}{2\theta}\right)^n \mathbb{I}\{x_{(1)}<\theta\}\, \mathbb{I}\{x_{(n)}<\theta\}$$

$$= \left(\frac{1}{2\theta}\right)^n \underbrace{\mathbb{I}\{\max\{x_{(1)}, x_{(n)}\}<\theta\}} \qquad T \triangleq \max\{x_{(1)}, x_{(n)}\}$$

By Factorization Thm. $\quad T(X) = \max\{X_{(1)}, X_{(n)}\}$

(e) $f(x) = \prod_{i=1}^{n} f(x_i) = \left(\frac{1}{\theta}\right)^n \prod_{i=1}^{n} \mathbb{I}\{\theta < x_i < 2\theta\}$

$$= \left(\frac{1}{\theta}\right)^n \underbrace{\mathbb{I}\{\theta < x_{(1)}\} \mathbb{I}\{\frac{x_{(n)}}{2} < \theta\}}_{g(T,\theta)} \quad T \triangleq (x_{(1)}, \frac{x_{(n)}}{2})$$

Actually $T = (x^{(1)}, x^{(n)})$ is also right

By Factorization Thm. $T(X) = (x^{(1)}, \frac{x^{(n)}}{2})$

(f) $f(x, y) = \left(\frac{1}{\pi\theta^2}\right)^n \cdot \prod_{i=1}^{n} \mathbb{I}\{x_i^2 + y_i^2 \le \theta\}$

$z_i \triangleq x_i^2 + y_i^2$

$$= \left(\frac{1}{\pi\theta^2}\right)^n \underbrace{\mathbb{I}\{z^{(n)} \le \theta\}}_{g(T,\theta)} \quad T \triangleq \max\{x_i^2 + y_i^2\}_{i=1}^{n}$$

By Factorization Thm. $T(X) = \max\{x_i^2 + y_i^2\}_{i=1}^{n}$

(g) $f(X, Y) = \left(\frac{1}{4}\right)^n \cdot \prod_{i=1}^{n} \left[\mathbb{I}\{\theta - 1 < x_i < \theta + 1\} \mathbb{I}\{\theta - 1 < y_i < \theta + 1\}\right] \quad T \triangleq \{x_{(1)}, x_{(n)}, y_{(1)}, y_{(n)}\}$

$$= \left(\frac{1}{4}\right)^n \cdot \prod_{i=1}^{n} \underbrace{\left[\mathbb{I}\{x_{(n)} - 1 < \theta < x_{(1)} + 1\} \mathbb{I}\{y_{(n)} - 1 < \theta < y_{(1)} + 1\}\right]}_{g(T,\theta)}$$

By Factorization Thm. $T(X) = \{x_{(1)}, x_{(n)}, y_{(1)}, y_{(n)}\}$

(h) $f(x) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \frac{1}{\theta} \cdot e^{-\frac{1}{\theta} x_i} = \underbrace{\frac{1}{\theta^n} \exp\{-\frac{1}{\theta} \sum_{i=1}^{n} x_i\}}_{g(T,\theta)} \quad T \triangleq \sum_{i=1}^{n} x_i$

By Factorization Thm. $T(X) = \sum_{i=1}^{n} x_i$

(i) $f(x) = \prod_{i=1}^{n} f(x_i) = \left(\frac{1}{\Gamma(\alpha)}\right)^n (\lambda^\alpha)^n \cdot \underbrace{\left(\prod_{i=1}^{n} x_i\right)^{\alpha - 1} \cdot \exp\{-\lambda \cdot \sum_{i=1}^{n} x_i\}}_{g(T, (\alpha, \lambda))} \quad T \triangleq \left(\prod_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i\right)$

By Factorization Thm. $T(X)=(\prod_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i)$ if both α, λ are unknown.

$T(x)=(\prod_{i=1}^{n} x_i)$ if λ is known.

$T(x)=(\sum_{i=1}^{n} x_i)$ if α is known.

(j) $F(x)=1-e^{-(\lambda x)^{\alpha}}$

$$f(x)=\frac{dF(x)}{dx}=\left(-e^{-(\lambda x)^{\alpha}}\right)\left(-\alpha\cdot(\lambda x)^{\alpha-1}\right)\cdot\lambda=\alpha\cdot\lambda^{\alpha}\cdot x^{\alpha-1}e^{-(\lambda x)^{\alpha}}$$

$$f(x)=\prod_{i=1}^{n} f(x_i)=\prod_{i=1}^{n}\left[\alpha\cdot\lambda^{\alpha}\cdot x^{\alpha-1}e^{-(\lambda x)^{\alpha}}\right]$$

$$=\underbrace{\alpha^{n}\cdot\lambda^{n\alpha}\cdot\left(\prod_{i=1}^{n} x_i\right)^{\alpha-1}\cdot e^{-\lambda^{\alpha}\sum_{i=1}^{n} x_i^{\alpha}}}_{g(T,\binom{\alpha}{\lambda})} \qquad T\triangleq\left(\prod_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i^{\alpha}\right)$$

By Factorization Thm.

$T=(\prod_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i^{\alpha})$ if α is known.

$T=(x_1\cdots x_n)$ or $(x_{(1)}\cdots x_{(n)})$ otherwise.

(k) $$f(x)=\prod_{i=1}^{n} f(x_i)=\left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right)^{n}\underbrace{\left(\prod_{i=1}^{n} x_i\right)^{\alpha-1}\left(\prod_{i=1}^{n}(1-x_i)\right)^{\beta-1}}_{g(T,\binom{\alpha}{\beta})} \qquad 0<x_i<1$$

$T\triangleq\left(\prod_{i=1}^{n} x_i, \prod_{i=1}^{n}(1-x_i)\right)$

By Factorization Thm. $T=\left(\prod_{i=1}^{n} x_i, \prod_{i=1}^{n}(1-x_i)\right)$ if both α, β are unknown

$T=(\prod_{i=1}^{n} x_i)$ if β is known.

$T=(\prod_{i=1}^{n}(1-x_i))$ if α is known.

Ex 3.3. $x_1, x_2\cdots x_n \overset{iid}{\sim} N(\theta,1)$ $\qquad X\overset{p.d.f.}{\sim} f_X(x)=\frac{1}{\sqrt{2\pi}}\exp\{-\frac{(x-\theta)^2}{2}\}$

$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i \sim N(\theta,\frac{1}{n})$ $\qquad Y=\bar{x}\overset{p.d.f.}{\sim} f_Y(y)=\frac{1}{\sqrt{2\pi\frac{1}{n}}}\cdot\exp\{\frac{n(y-\theta)^2}{2}\}$

$p(X|\bar{x})=\frac{p(X,\bar{x})}{p(\bar{x})}$ (*) $\qquad$ problems.

Let $X=(x_1,x_2\cdots x_n)$ $\bar{x}=t$

if $\frac{1}{n}\sum_{i=1}^{n} x_i\neq t \Rightarrow (*)=0$

if $\frac{1}{n}\sum_{i=1}^{n} x_i=t \Rightarrow (*)=\frac{p(X)}{(p(\bar{x}))^{n}}=\frac{\prod_{i=1}^{n} f_X(x_i)}{\left(f_Y(\frac{1}{n}\sum_{i=1}^{n} x_i)\right)^{n}}=\frac{(\frac{1}{\sqrt{2\pi}})^{n}\exp\{-\frac{1}{2}\overset{A}{\sum_{i=1}^{n}(x_i-\theta)^2}\}}{\left(\frac{\sqrt{n}}{\sqrt{2\pi}}\right)^{n}\exp\{-\frac{1}{2}\underset{B}{n(\frac{1}{n}\sum_{i=1}^{n} x_i-\theta)^2}\}}$

$$A \triangleq \sum_{i=1}^{n}(x_i-\theta)^2$$
$$= \sum_{i=1}^{n}(x_i^2-2x_i\theta+\theta^2)$$
$$= \sum_{i=1}^{n}x_i^2-2\cdot\Big(\sum_{i=1}^{n}x_i\Big)\cdot\theta+n\theta^2$$

$$B \triangleq \Big(\big(\tfrac{1}{n}\sum_{i=1}^{n}x_i-\theta\big)^2\Big)n$$
$$= \Big(\big(\tfrac{1}{n}\sum_{i=1}^{n}x_i\big)^2-2\big(\tfrac{1}{n}\sum_{i=1}^{n}x_i\big)\theta+\theta^2\Big)n$$
$$= \tfrac{1}{n}\Big(\sum_{i=1}^{n}x_i\Big)^2-2\cdot\Big(\sum_{i=1}^{n}x_i\Big)\theta+n\theta^2$$

$$A-B = \sum_{i=1}^{n}x_i^2-\frac{1}{n}\Big(\sum_{i=1}^{n}x_i\Big)^2$$

Then. $$p(X\mid\bar{X}) = \frac{1}{(\sqrt{n})^n}\exp\Big\{-\frac{1}{2}\Big(\sum_{i=1}^{n}x_i^2-\frac{(\sum_{i=1}^{n}x_i)^2}{n}\Big)\Big\}$$

I will use bootstrap to sample a small population. $Z^{(i)}=\{X_1,\cdots X_k\}$ from the original population. $X=(x_1\cdots x_n)$ for the i-th iteration.

Ex 3.4 Since P_θ is continuous

problem 4

Let $T=(X_{(1)},X_{(2)}\cdots X_{(n)})$

$$P(X=x\mid T=t) = P\big((X_1\cdots X_n)=(x_1\cdots x_n)\,\big|\,(X_{(1)}\cdots X_{(n)})=(t_1\cdots t_n)\big) \quad (*)$$

when $\{x_1,x_2,\cdots x_n\}\neq\{t_1,t_2,\cdots t_n\}$ we have $(*)=0$

when $\{x_1,x_2,\cdots x_n\}=\{t_1,t_2,\cdots t_n\}$ we have $(*)=\frac{1}{n!}$

Thus,

$$p(X=x\mid T=t) = \frac{1}{n!}\,\mathbb{I}\big\{\{x_1\cdots x_n\}=\{t_1\cdots t_n\}\big\}$$

does not depend on θ

Thus, T is sufficient for continuous P_θ for a non-parametric model.

Ex. 3.7. problem 5

(a)

$T(x)=x_{(n)}$

$$\frac{f_\theta(x)}{f_\theta(y)} = \frac{\mathbb{I}\{0<x_{(1)}\}\,\mathbb{I}\{x_{(n)}<\theta\}}{\mathbb{I}\{0<y_{(1)}\}\,\mathbb{I}\{y_{(n)}<\theta\}}$$ is free of $\theta \Rightarrow x_{(n)}=y_{(n)}$,

$\therefore$ $T(x)=X_{(n)}$ is minimal sufficient for θ

(b). $T(x)=\{X_{(1)},X_{(n)}\}$

$$\frac{f_\theta(x)}{f_\theta(y)} = \frac{\mathbb{I}\{x_{(n)}-1<\theta<x_{(1)}+1\}}{\mathbb{I}\{y_{(n)}-1<\theta<y_{(1)}+1\}}$$ is free of $\theta \Rightarrow \begin{cases}x_{(1)}=y_{(1)}\\ x_{(n)}=y_{(n)},\end{cases}$

$\therefore$ $T(x)=\{x_{(1)}\ x_{(n)}\}$ is minimal sufficient for θ.

(c). $T(x)=\{x_{(1)},x_{(n)}\}$

$$\frac{f_\theta(x)}{f_\theta(y)} = \frac{\mathbb{I}\{\theta_1\le x_{(1)}\}\,\mathbb{I}\{x_{(n)}\le\theta_2\}}{\mathbb{I}\{\theta_1\le y_{(1)}\}\,\mathbb{I}\{y_{(n)}\le\theta_2\}}$$ is free of $\theta \Rightarrow \begin{cases}x_{(1)}=y_{(1)}\\ x_{(n)}=y_{(n)}\end{cases}$

$\therefore T(x) = \{x_{(1)}, x_{(n)}\}$ is minimal sufficient for θ

(d) $T(x) = \{\max\{x_{(1)}, x_{(n)}\}\}$

$$\frac{f_\theta(x)}{f_\theta(y)} = \frac{\mathbb{I}\{\max\{x_{(1)}, x_{(n)}\} < \theta\}}{\mathbb{I}\{\max\{y_{(1)}, y_{(n)}\} < \theta\}} \text{ is free of } \theta$$

$\Rightarrow \max\{x_{(1)}, x_{(n)}\} = \max\{y_{(1)}, y_{(n)}\}$

$\therefore T(x) = \{\max\{x_{(1)}, x_{(n)}\}\}$ is minimal sufficient

(e) $T(x) = \{x_{(1)}, \frac{x_{(n)}}{2}\}$

$$\frac{f_\theta(x)}{f_\theta(y)} = \frac{\mathbb{I}\{\theta < x_{(1)}\}\,\mathbb{I}\{\frac{x_{(n)}}{2} < \theta\}}{\mathbb{I}\{\theta < y_{(1)}\}\,\mathbb{I}\{\frac{y_{(n)}}{2} < \theta\}} \text{ is free of } \theta.$$

$$\Rightarrow \begin{cases} x_{(1)} = y_{(1)} \\ \frac{x_{(n)}}{2} = \frac{y_{(n)}}{2} \end{cases}$$

$T(x) = \{x_{(1)}, \frac{x_{(n)}}{2}\}$ is minimal sufficient.

(f) $T(x) = \max\{x_i^2 + y_i^2\}_{i=1}^n$

$$\frac{f_\theta(x,y)}{f_\theta(x',y')} = \frac{\mathbb{I}\{\max\{x_i^2+y_i^2\}_{i=1}^n \le \theta\}}{\mathbb{I}\{\max\{x_i'^2+y_i'^2\}_{i=1}^n \le \theta\}} \text{ is free of } \theta$$

$\Rightarrow \max\{x_i^2+y_i^2\}_{i=1}^n = \max\{x_i'^2+y_i'^2\}$

$T(x) = \{\max\{x_i^2+y_i^2\}_{i=1}^n\}$ is thus minimal sufficient

(g) $$\frac{f_\theta(x,y)}{f_\theta(x',y')} = \frac{\mathbb{I}\{x_{(n)}-1<\theta<x_{(1)}+1\}\,\mathbb{I}\{y_{(n)}-1<\theta<y_{(1)}+1\}}{\mathbb{I}\{x'_{(n)}-1<\theta<x'_{(1)}+1\}\,\mathbb{I}\{y'_{(n)}-1<\theta<y'_{(1)}+1\}}$$

is free of θ.

$$\Rightarrow \begin{cases} x_{(1)} = x_{(1)}' \\ x_{(n)} = x_{(n)}' \\ y_{(1)} = y_{(1)}' \\ y_{(n)} = y_{(n)}' \end{cases}$$

$\therefore T(x) = \{x_{(1)}, x_{(n)}, y_{(1)}, y_{(n)}\}$ is minimal sufficient

(h) $T(x) = \{\sum_{i=1}^n x_i\}$

$$\frac{f_\theta(x)}{f_\theta(y)} = \exp\{-\frac{1}{\theta}(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i)\} \text{ is free of } \theta$$

$$\Rightarrow \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

$\therefore T(x) = \sum_{i=1}^n x_i$ is minimal sufficient

(i). $$\frac{f_\theta(x)}{f_\theta(y)} = \frac{(\prod_{i=1}^n x_i)^{\alpha-1} \exp\{-\lambda \cdot \sum_{i=1}^n x_i\}}{(\prod_{i=1}^n y_i)^{\alpha-1} \exp\{-\lambda \cdot \sum_{i=1}^n y_i\}} \text{ is free of } \theta = \binom{\alpha}{\lambda}$$

$$\Rightarrow \begin{cases} \prod_{i=1}^n x_i = \prod_{i=1}^n y_i \\ \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \end{cases}$$

$\therefore T(x) = (\prod_{i=1}^n x_i, \sum_{i=1}^n x_i)$ is minimal sufficient

(j). $$\frac{f_\theta(x)}{f_\theta(y)} = \frac{(\prod_{i=1}^n x_i)^{\alpha-1} e^{-\lambda^\alpha \cdot \sum_{i=1}^n x_i^\alpha}}{(\prod_{i=1}^n y_i)^{\alpha-1} e^{-\lambda^\alpha \cdot \sum_{i=1}^n y_i^\alpha}} \text{ is free of } \theta$$

$T(x) = (\prod_{i=1}^n x_i, \sum_{i=1}^n x_i)$ is minimal sufficient

(k) $$\frac{f_\theta(x)}{f_\theta(y)} = \frac{(\prod_{i=1}^n x_i)^{\alpha-1} (\prod_{i=1}^n (1-x_i))^{\beta-1}}{(\prod_{i=1}^n y_i)^{\alpha-1} (\prod_{i=1}^n (1-y_i))^{\beta-1}} \text{ is free of } \theta = \binom{\alpha}{\beta}$$

$$\Rightarrow \begin{cases} \prod_{i=1}^n x_i = \prod_{i=1}^n y_i \\ \prod_{i=1}^n (1-x_i) = \prod_{i=1}^n (1-y_i) \end{cases}$$

$T = (\prod_{i=1}^n x_i, \prod_{i=1}^n (1-x_i))$ is minimal sufficient

The following five problem are selected from the note.

1. let $X_1, X_2 \cdots X_n \overset{iid}{\sim} F_\theta$ Find a non-trivial sufficient statistics for θ

let $\mathbb{X} = (X_1, X_2, \cdots X_n)$ $\quad X = (x_1, x_2, \cdots x_n)$

problem 1.

(a). $f(x) = \dfrac{1}{\Gamma(\alpha)\cdot\beta^{\alpha}} x^{\alpha-1}\cdot e^{-x/\beta} \qquad x>0 \qquad \theta = (\alpha, \beta)^T$

$$f(X) \overset{iid}{=} \prod_{i=1}^{n} f(x_i) = \left(\frac{1}{\Gamma(\alpha)\cdot\beta^{\alpha}}\right)^n \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1}\cdot e^{-\frac{\sum_{i=1}^{n} x_i}{\beta}}$$

$$T_1(X) \triangleq \prod_{i=1}^{n} x_i \qquad T_2(X) \triangleq \sum_{i=1}^{n} x_i \qquad T = \binom{T_1}{T_2} \qquad h(X) \triangleq 1$$

$$g(\theta, T) = g\left(\binom{\alpha}{\beta}, \binom{T_1}{T_2}\right) \triangleq \frac{(T_1)^{\alpha-1}}{\Gamma(\alpha)^n} \times \frac{\exp\{-T_2/\beta\}}{(\beta^{\alpha})^n}$$

Then. $f(X) = g(\theta, T)\cdot h(X) \xRightarrow[\text{Thm.}]{\text{Factorization}} T = \left(\prod_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i\right)$ is sufficient.

Actually, the p.d.f. here belongs to r.v. Gamma $(\alpha, \frac{1}{\beta})$

(b) $f(x) = \frac{1}{6}\cdot \exp\{-(x-\mu)/6\}$ for $x>\mu$. $\quad \theta = \binom{\mu}{6}$

$$f(X) = \prod_{i=1}^{n} f(x_i) = \frac{1}{6^n}\cdot\exp\left\{-\left(\sum_{i=1}^{n} x_i - n\mu\right)/6\right\}$$

$$T \triangleq \sum_{i=1}^{n} x_i \qquad h(x) \triangleq 1 \qquad g(T, \theta) \triangleq \frac{1}{6^n}\exp\{-(T-n\mu)/6\}$$

Then. $f(X) = g(\theta, T)\, h(X)$

$\xRightarrow[\text{Thm.}]{\text{Factorization}} T = \sum_{i=1}^{n} x_i$ is sufficient.

(c). $X \sim \text{Uniform}(\theta - \frac{1}{2}, \theta + \frac{1}{2})$

$$f(x) = \mathbb{I}.\{x \in (\theta-\tfrac{1}{2}, \theta+\tfrac{1}{2})\} = \mathbb{I}\{\theta-\tfrac{1}{2} < x < \theta+\tfrac{1}{2}\}$$

$$f(X) = \prod_{i=1}^{n} \mathbb{I}\{\theta-\tfrac{1}{2}<x<\theta+\tfrac{1}{2}\} = \mathbb{I}.\{\theta-\tfrac{1}{2} < x_{(1)} < \cdots < x_{(n)} < \theta+\tfrac{1}{2}\}$$

$$= \mathbb{I}\{x_{(n)} - \tfrac{1}{2} < \theta < x_{(1)} + \tfrac{1}{2}\}$$

$T = \{x_{(1)}, x_{(n)}\}$ is the suffient statistics by factorization Thm.

2. Let $X_1, X_2, \cdots X_n$ be independent r.v. with densities. $f_{X_i}(x) = e^{i\theta - x}$

for $x > i\theta$ prove that $T = \min_i \{X_i / i\}$ is a sufficient statistic for θ

$$f(x) = \prod_{i=1}^n f(x_i) = \exp\left\{\sum_{i=1}^n (i\theta - x_i)\right\} \prod_{i=1}^n \mathbb{I}\{x_i > i\theta\}$$

problem 2

$$= \exp\left\{\frac{(1+n)n}{2}\theta - \sum_{i=1}^n x_i\right\} \cdot \prod_{i=1}^n \mathbb{I}\left\{\frac{x_i}{i} > \theta\right\}$$

$$= \exp\left\{\frac{n(n+1)}{2}\theta\right\} \times \exp\left\{-\sum_{i=1}^n x_i\right\} \cdot \mathbb{I}\left\{\min_i\left\{\frac{x_i}{i}\right\} > \theta\right\}.$$

$$\begin{cases} T(x) = \min_i \left\{\frac{x_i}{i}\right\} \\ h(x) \triangleq \exp\left\{-\sum_{i=1}^n x_i\right\} \\ g(\theta, T) \triangleq \exp\left\{\frac{n(n+1)}{2}\theta\right\} \cdot \mathbb{I}\{T > \theta\}. \end{cases} \xRightarrow[\text{Thm.}]{\text{Factorization.}} T(x) = \min_i \left\{\frac{x_i}{i}\right\} \text{ is the sufficient statistic for } \theta$$

4. Let. $X_i \overset{iid.}{\sim} X \sim N(\mu, \sigma^2)$. where σ^2 is known $\theta = \mu$.

Then $T = (X_1, \cdots X_{n-1})$ is not sufficient for θ.

problem 3.

Claim:

$S = \sum_{i=1}^n x_i$ is the minimal. sufficient statistic for μ.

① proof: $S = \sum_{i=1}^n x_i$ is the sufficient statistic for μ.

$$f(x) = \prod_{i=1}^n f(x_i) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \cdot \exp\left\{-\frac{1}{2\sigma^2} \cdot \left(\left(\sum_{i=1}^n x_i\right)^2 - 2\mu \cdot \sum_{i=1}^n x_i + \mu^2\right)\right\}$$

S is sufficient statistic for μ. by Factorization Thm.

② proof.: S is actually a minimal one.

for $\forall X, Y$

$$r(\mu) \equiv \frac{f(X)}{f(Y)} = \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2 - 2\mu\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)\right)\right\}$$

$$\equiv \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2 - 2\mu(T(x) - T(y)\right)\right\}$$

Then. $r(\mu)$ is ind of μ iff $T(x) = T(y)$

Thus, by Thm 2.3. S is minimal sufficient for μ.

we finished the proof of the claim: S is minimal sufficient for μ.

Suppose. $T = (X_1, X_2, \cdots X_{n-1})$ is a sufficient statistic

By the Def of minimal sufficient S is a function of T

However we cannot find a map from $(X_1 \cdots X_{n-1})$ to $\mathbb{1}^T \cdot (X_1 \cdots X_n)$

where $\mathbb{1}$ is a vector full of 1 of dim $= n$

$\Rightarrow$ Contradictory.

Therefore., $T = (X_1, X_2 \cdots X_{n-1})$ cannot be sufficient for μ.

5. Let. $X_i \overset{i.i.d.}{\sim} X \sim U(\theta, \theta+1)$ Then T is sufficient but not complete

$$f(\vec{x}) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \mathbb{1}(\theta < x_i < \theta+1) = \mathbb{1}\{\theta < x_{(1)} < \cdots < x_{(n)} < \theta+1\}$$

$$= \mathbb{1}\{x_{(n)} - 1 < \theta < x_{(1)}\}$$

By factorization Thm. $T = \{x_{(1)}, x_{(n)}\}$ is the sufficient statistic for θ

Next we prof T is not complete.

① $X_1, X_2 \cdots X_n \overset{iid}{\sim} X \overset{pdf}{\sim} f(x). \overset{cdf}{\sim} F(x)$. problem 4

Induct the p.d.f. of order statistic. $X^{(k)}$

$$P(X^{(k)} \in [x, x+\varepsilon)) = P(\{\text{one of the } X \in [x, x+\varepsilon) \text{ and exactly } k-1 \text{ others} < x\})$$

$$= P\left(\bigcup_{i=1}^{n} \{X_i \in [x, x+\varepsilon) \text{ and exactly } k-1 \text{ others} < x\}\right)$$

$$\overset{\text{disjoint events}}{=} \sum_{i=1}^{n} P(\{X_i \in [x, x+\varepsilon) \text{ and exactly } k-1 \text{ others} < x\})$$

$$= n \cdot P(X \in [x, x+\varepsilon)) \cdot \binom{n-1}{k-1} \cdot P(X < x)^{k-1} \, P(X > x)^{n-k}$$

$$= n \cdot \binom{n-1}{k-1} f_X(x) \cdot F_X(x)^{k-1} (1 - F_X(x))^{n-k}.$$

Thus. $f_{X^{(k)}}(x) = n \cdot \binom{n-1}{k-1} f_X(x) \cdot F_X(x)^{k-1} (1 - F_X(x))^{n-k}.$

$X_i \overset{iid}{\sim} U(\theta, \theta+1)$

$$f_X(x) = \mathbb{1}\{x \in (\theta, \theta+1)\} \qquad F_X(x) = \int_{-\infty}^{x} \mathbb{1}\{x \in \theta, \theta+1\} \cdot dx = \begin{cases} 0 & x < \theta \\ x - \theta & \theta < x < \theta+1 \\ 1 & x > \theta+1 \end{cases}$$

when $x < \theta$ $\quad f_{X^{(k)}}(x) = 0$

$\theta < x < \theta + 1$ $\quad f_{X^{(k)}}(x) = n \cdot \binom{n-1}{k-1} (x-\theta)^{k-1} \cdot (1-(x-\theta))^{n-k}$

$x > \theta + 1$ $\quad f_{X^{(k)}}(x) = 0$

let $\alpha - 1 = k - 1$ $\quad \beta - 1 = n - k$. $\quad \begin{cases} \alpha = k \\ \beta = n - k + 1. \end{cases}$

Then. $\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} = \dfrac{(\alpha+\beta)!}{\alpha!\ \beta!} = \dfrac{(n+1)!}{(k-1)!\,(n-k)!}$

$n \cdot \binom{n-1}{k-1} = n \cdot \dfrac{(n-1)!}{(n-k)!(k-1)!}$

$\therefore n \cdot \binom{n-1}{k-1} = \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$.

$\therefore f_{X^{(k)}}(x) = \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot (x-\theta)^{\alpha-1} (1-(x-\theta))^{\beta-1}$

Let $Y_i = X_i - \theta$ $\quad Y_i \sim U(0,1)$ $\quad$ for $Y^{(k)}$

$f_{Y_{(k)}}(y) = \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1} (1-y)^{\beta-1} \sim Beta(\alpha,\beta)$

$E(Y_{(k)}) = \dfrac{k}{n+1}$ $\quad Var(Y_{(k)}) = \dfrac{k(n-k+1)}{(n+1)^2(n+2)}$

$\therefore E(X_{(1)} - \theta) = E(Y_{(1)}) = \dfrac{1}{n+1}$

$E(X_{(n)} - \theta) = E(Y_{(n)}) = \dfrac{n}{n+1}$

Let $c = \dfrac{n-1}{n+1}$

$g(T) \triangleq X_{(1)} - X_{(n)} + c$

$E(g(T)) = E[X_{(1)} - X_{(n)} + c] = E[(X_{(1)} - \theta)] - E[X_{(n)} - \theta] + c = 0$

However $\quad g(T) = X_{(1)} - X_{(n)} + c \neq 0$

$\therefore$ $T=(X_{(1)}, X_{(n)})$ is not complete.

7. $X_1 \cdots X_n$ are i.i.d. r.v.s. with p.d.f. $f(x) = \theta^{-1} \cdot e^{-(x-a)/\theta} \, \mathbb{I}_{(a,\infty)}(x)$.

Show that: (1) $\sum_{i=1}^{n} (X_i - X_{(1)})$ and $X_{(1)}$ are independent. problem5.

(2) $Z_i =: (X_{(n)} - X_{(i)}) / (X_{(n)} - X_{(n-1)})$ $i=1,2,\cdots n-2$ are independent of $(X_{(1)}, \sum_{i=1}^{n} (X_i - X_{(1)}))$

(1). Constrain the case to $\forall\ \theta$ which is fixed.

$$f(x) = \prod_{i=1}^{n} f(x_i) = \theta^{-n} \exp\{-\tfrac{1}{\theta}(\textstyle\sum_{i=1}^{n} x_i - na)\} \prod_{i=1}^{n} \mathbb{I}_{(a,\infty)}(x_i)$$

$$= \theta^{-n} \exp\{-\tfrac{1}{\theta}(\textstyle\sum_{i=1}^{n} x_i - na)\} \, \mathbb{I}\{x_{(1)} > a\}$$

$$= \theta^{-n} \cdot \exp\{\tfrac{na}{\theta}\} \cdot \exp\{-\tfrac{1}{\theta}(\textstyle\sum_{i=1}^{n} x_i)\} \, \mathbb{I}\{x_{(1)} > a\}$$

By factorization. Thm. $X_{(1)}$ is sufficient for a

For $x^{(1)}$.

$$p(x^{(1)} \in [x, x+\varepsilon)) = P(\text{some } X \in [x, x+\varepsilon) \text{ and. others.} > x)$$

$$= n \cdot P(X_1 \in [x, x+\varepsilon) \text{ and others} > x)$$

$$= n \cdot f(x) \cdot (1 - F(x))^{n-1}.$$

$$f_{X^{(1)}}(x) = n f(x) \cdot (1 - F(x))^{n-1}$$

For $F(x)$.

$$Y = X - a \overset{p.d.f}{\sim} \theta^{-1} \cdot e^{-\theta^{-1} \cdot y} \, \mathbb{I}_{(0,\infty)}(y). \sim \mathcal{E}(\tfrac{1}{\theta}).$$

$$F_Y(y) = \begin{cases} 1 - e^{-\frac{1}{\theta} y} & y \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

$X = Y + a$

$$F_X(x) = F_Y(x-a) = \begin{cases} 1 - e^{-\frac{1}{\theta}(x-a)} & x \geq a \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $f_{X_{(1)}}(x) = n \cdot \theta^{-1} \cdot e^{-(x-a)/\theta} \mathbb{I}_{(a,\infty)}(x) \cdot \left(e^{-(x-a)/\theta}\right)^{n-1}$

$= \frac{n}{\theta} \cdot e^{-\frac{n(x-a)}{\theta}} \cdot \mathbb{I}_{(a,\infty)}(x).$

For any g.

if $\mathbb{E}_T(g(T)) = \mathbb{E}_{X_{(1)}}[g(X^{(1)})] = \int_x g(x) \cdot \frac{n}{\theta} \cdot e^{-\frac{(x-a)}{\theta}} \mathbb{I}_{(a,\infty)}(x) \cdot dx$

$= \int_a^\infty g(x) \cdot \frac{n}{\theta} e^{-\frac{(x-a)}{\theta}} dx = \frac{n}{\theta} \int_a^\infty g(x) \cdot e^{-\frac{x-a}{\theta}} dx = 0$

Then $g(x) = 0$ a.s. for any a

So $X^{(1)}$ is complete.

Since the p.d.f of $Y = X - a$ doesn't depend on a.

Thus the p.d.f of $\sum_{i=1}^n (X_i - X_{(1)}) = \sum_{i=1}^n [(X_i - a) - (X_{(1)} - a)]$

does not depend on a

Thus $\sum_{i=1}^n (X_i - X_{(1)})$ is ancillary.

By Basu's Thm.

$\sum_{i=1}^n (X_i - X_{(1)}) \perp X_{(1)}$. $\forall\, \theta$ is fixed

Therefore for all θ we have $\sum_{i=1}^n (X_i - X_{(1)}) \perp X_{(1)}$

$\Rightarrow \sum_{i=1}^n (X_i - X_{(1)}) \perp X_{(1)}$.

QED.

(2). Actually $\sum_{i=1}^n (X_i - X_{(1)})$ is the sufficient and complete statistic for θ

sufficient: obvious by Factorization Thm. (See top on this solution)

Complete: The proof is analogous to question (1).

Thus. $\left(\sum_{i=1}^n (X_i - X_{(1)}), X_{(1)}\right)$ is the sufficient and complete statistic for (a, θ).

Since. $H_i = \frac{X_i - a}{\theta}$ has p.d.f. $e^{-h} \mathbb{I}_{(0,\infty)}(h)$.

which does not depend on. a and θ

H_i are ancillary. $H_{(i)}$ are ancillary.

Thus.

$$z_i = \frac{X_{(i)} - X_{(1)}}{X_{(n)} - X_{(1)}} = \frac{\frac{X_{(i)} - a}{\theta} - \frac{X_{(1)} - a}{\theta}}{\frac{X_{(n)} - a}{\theta} - \frac{X_{(1)} - a}{\theta}} = \frac{H_{(i)} - H_{(1)}}{H_{(n)} - H_{(1)}}$$

are ancillary.

By Basu's Theorem $(z_1, z_2 \cdots z_n) \perp \left(X_{(1)}, \sum_{i=1}^{n} (X_i - X_{(1)}) \right)$.

QED.