2009

1.

a) Since $E(X_i|T) = f(T)$ a.s. $\forall i$

$$\Rightarrow T = E(T|T) = E\left(\sum_{i=1}^{n} X_i | T\right)$$
$$= \sum_{i=1}^{n} E(X_i|T) \quad \text{(by linearity)}$$
$$= n E(X_i|T) \quad \text{(by sym.)}$$
$$= n f(T)$$

$\therefore$ $E(X_i|T) = T/n$ $\forall i$.

b)

$$T = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S = \frac{1}{n}\sum X_i^2.$$

$$T^2 = \frac{1}{n^2}\left(\sum_{i=1}^{n} X_i\right)^2 = \frac{1}{n^2}\left[\sum X_i^2 + \sum_{k\neq l} X_k X_l\right] = n^{-1}S + \frac{1}{n^2}\sum_{k\neq l} X_k X_l$$

$$\Rightarrow E(T^2|T,S) = n^{-2}\left[nE(S|T,S) + E\left(\sum_{k\neq l} X_k X_l | T,S\right)\right]$$

$$\Rightarrow T^2 = n^{-1}S + n^{-2}\binom{n}{2} E(X_i X_j | T,S)$$

Note that $T^2 - n^{-1}S = \frac{1}{n^2}\sum_{k\neq l} X_k X_l$

$$2\frac{1}{n^2}\binom{n}{2} E(X_i X_j|T,S) = \frac{1}{n^2}\sum_{k\neq l} X_k X_l$$

$$\Rightarrow E(X_i X_j|T,S) = \frac{1}{n(n-1)}\sum_{k\neq l} X_k X_l.$$

Scanned with CamScanner

c) by (b)

$$E(X_j X_j | T, S) = \frac{1}{n(n-1)} \sum_{k \neq l} X_k X_l$$

is indeed U-statistic with kernel $h(x,y) = xy$

$$\Rightarrow E(X_i X_j | T, S) \xrightarrow{P_r} E(X_1 X_2) = \mu^2, \text{ where } \mu = E(X_1)$$

By (a).

$$E(X_i | T) = T/n$$

$$E(X_i | T, S) = E(E(X_i | T) | S) = \frac{1}{n} E(T | S) \quad (\text{Prop } 1.10 - (v))$$

$$\xrightarrow{P_r} \mu. \quad \left(\begin{array}{c} \text{as } \frac{1}{n}\sum X_i, \ \frac{1}{n}\sum (X_i - \bar{X})^2 \\ \text{are asy. indep.} \\ \text{-- } (*) \end{array} \ \begin{array}{c} \text{and } E(T|S) \\ \text{is function of } T, S \end{array} \right)$$

By Slutsky, the result follows.

"Proof of (*)": By CLT,

$$\sqrt{n} \left(\begin{bmatrix} \bar{X} \\ S^2 \end{bmatrix} - \begin{bmatrix} E(X_1) \\ \underset{\sigma^2}{\overset{\shortparallel}{Var(X_1)}} \end{bmatrix} \right) \xrightarrow{d} N\left(0, \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^2 \end{bmatrix}\right)$$

$\Rightarrow T$ is asy. indep to S. (by continuous mapping.)

2.

a)

$$E(X_1) = \frac{1}{\theta^2}\int_a^\infty x(x-a)\, e^{-(x-a)/\theta}\, dx. \quad , \text{let } \mu_j = E(X_i^j)$$

$$\hat{\mu}_j = \frac{1}{n}\sum_{i=1}^{n} X_i^j$$

Let $y = \frac{x-a}{\theta}$, $x = \theta y + a$

$dy = \frac{1}{\theta}\, dx$

$$= \int_0^\infty (\theta y + a) y\, e^{-y}\, dy$$

$$= \theta\int_0^\infty \underbrace{y^2 e^{-y}\, dy}_{\Gamma(3)=2!} + a\int_0^\infty \underbrace{y e^{-y}\, dy}_{\Gamma(2)=1!}$$

$$= 2\theta + a$$

$$E(X_1^2) = \frac{1}{\theta^2}\int_a^\infty x^2(x-a)\, e^{-(x-a)/\theta}\, dx$$

$$= \int_0^\infty (\theta y + a)^2 y\, e^{-y}\, dy$$

$$= \theta^2\int_0^\infty \underbrace{y^3 e^{-y}\, dy}_{\Gamma(4)=3!} + 2a\theta\int_0^\infty \underbrace{y^2 e^{-y}\, dy}_{\Gamma(3)=2!} + a^2\int_0^\infty \underbrace{y e^{-y}\, dy}_{\Gamma(2)=1!}$$

$$= 6\theta^2 + 4a\theta + a^2. =$$

$$\Rightarrow \begin{cases} \mu_1 = 2\theta + a \\ \mu_2 = 6\theta^2 + 4a\theta + a^2. \end{cases}$$

$① : \quad a = \mu_1 - 2\theta$

$$\Rightarrow \quad \mu_2 = 6\theta^2 + 4(\mu_1 - 2\theta)\theta + (\mu_1 - 2\theta)^2$$

$$= 6\theta^2 + \cancel{4\mu_1\theta} - 8\theta^2 + 4\theta^2 - \cancel{4\mu_1\theta} + \mu_1^2$$

$$= 2\theta^2 + \mu_1^2$$

$$\therefore \quad \hat{\theta} = \sqrt{\frac{\hat{\mu}_2 - \hat{\mu}_1^2}{2}}$$

and $\hat{a} = \hat{\mu}_1 - \sqrt{2(\hat{\mu}_2 - \hat{\mu}_1^2)}$ //

b) Consider $Y_i = (X_i, X_i^2)^T$ are iid. by CLT, we have

$$\sqrt{n}(\bar{Y} - E(Y_1)) \xrightarrow{d} N(0, \Sigma), \text{ where}$$

$$\Sigma = \begin{bmatrix} Var(X_1) & Cov(X_1, X_1^2) \\ Cov(X_1, X_1^2) & Var(X_1^2) \end{bmatrix}$$

$$= \begin{bmatrix} E(X_1^2) - E(X_1)^2 & E(X_1^3) - E(X_1)E(X_1^2) \\ E(X_1^3) - E(X_1)E(X_1^2) & E(X_1^4) - E(X_1^2)^2 \end{bmatrix}$$

$$= \begin{bmatrix} \mu_2 - \mu_1^2 & \mu_3 - \mu_1\mu_2 \\ \mu_3 - \mu_1\mu_2 & \mu_4 - \mu_2^2 \end{bmatrix}$$

$\bar{Y} = \frac{1}{n}\left(\sum x_i, \sum x_i^2\right)^T$ and $E(Y_1) = [\mu_1, \mu_2]$

Let $g(s,t) = \frac{1}{\sqrt{2}}(t-s^2)^{\frac{1}{2}}$, $h(s,t) = s - \sqrt{2(t-s^2)}$

$$\nabla g(s,t) = \frac{1}{\sqrt{2}}\begin{bmatrix} \frac{1}{\cancel{2}}(t-s^2)^{-\frac{1}{2}}(-\cancel{2}s) \\ \frac{1}{2}(t-s^2)^{-\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} -\frac{s}{\sqrt{2(t-s^2)}} \\ \frac{1}{2\sqrt{2(t-s^2)}} \end{bmatrix}$$

$$\nabla h(s,t) = \begin{bmatrix} 1-\sqrt{2}(t-s^2)^{-\frac{1}{2}}(-2s) \\ -\sqrt{2}(t-s^2)^{-\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} 1+\frac{2^{\frac{3}{2}}s}{\sqrt{t-s^2}} \\ -\frac{\sqrt{2}}{\sqrt{t-s^2}} \end{bmatrix}.$$

By Detla's method, let $G = \nabla g(\mu_1,\mu_2)$, $H = \nabla h(\mu_1,\mu_2)$

$$\sqrt{n}(\hat{\theta}-\theta) \xrightarrow{d} N(0, G\Sigma G^T)$$

$$\sqrt{n}(\hat{a}-a) \xrightarrow{d} N(0, H\Sigma H^T).$$

c) $f(\underset{\sim}{x}) = \prod_{i=1}^{n} \frac{x_i-a}{\theta^2} e^{-(x_i-a)/\theta} I_{(a,\infty)}(x_i)$

$= \exp\left\{\sum_{i=1}^{n}\log\left(\frac{x_i-a}{\theta^2}\right) - \frac{1}{\theta}\sum(x_i-a)\right\} I_{(a,\infty)}(x_{(1)})$

$= \exp\left\{\sum_{i=1}^{n}\log(x_i-a) - \frac{1}{\theta}\sum_{i=1}^{n}x_i + \frac{na}{\theta} - 2n\log\theta\right\} I_{(a,\infty)}(x_{(1)})$

$T = (X_{(1)}, \dots, X_{(n)})$ is sufficient by Factorization.

(Indeed is minimal sufficient by $\frac{f(x)}{f(y)} = \phi(a,y)$)
$\Rightarrow (X_{(1)}, \dots X_{(n)}) = (Y_{(1)} \dots Y_{(n)})$.

d) When $n=2$,

$$f(x) = \exp\left\{ \log(x_1-a)+\log(x_2-a) + \frac{1}{\theta}(x_1+x_2) + \frac{2a}{\theta} - 4\log\theta \right\} I_{(a,\infty)}(x_{(1)}).$$

$$= \exp\left\{ \log(x_1x_2 - a(x_1+x_2)-a^2) + \frac{1}{\theta}(x_1+x_2) + \frac{2a}{\theta} - 4\log\theta \right\} I_{(a,\infty)}(x_{(1)}).$$

$\Rightarrow$ $T=(X_1+X_2, \min(X_1,X_2))$ is sufficient.

(T is eq. to $(X_{(1)}, X_{(2)})$)

for

$$\frac{f(\underline{x})}{f(\underline{y})} = \phi(x,y) \Rightarrow \phi(x,y) = \frac{I_{(a,\infty)}(x_{(1)})}{I_{(a,\infty)}(y_{(1)})} \times$$

$$\exp\left\{ \log\left(\frac{x_1x_2-a(x_1+x_2)-a^2}{y_1y_2-a(y_1+y_2)-a^2}\right) + \frac{1}{\theta}\left[(x_1+x_2)-(y_1+y_2)\right] \right\}$$

$\Rightarrow$ $(x_1+x_2, \min(x_1,x_2)) = (y_1+y_2, \min(y_1, y_2))$

$\Rightarrow$ T is minimal sufficient.

e) When $\theta = 1$, known.

$$F_a(x) = \int_a^x (y-a) e^{-(y-a)} dy \quad , \text{ let } t = y-a ,$$
$$dt = dy$$

$$= \int_0^{x-a} t e^{-t} dt$$

$$\Rightarrow f_{X_{(1)}}(x) = n\left[1 - F_a(x)\right]^{n-1} (x-a) e^{-(x-a)} I_{(a,\infty)}(x).$$

$$= n\left(\int_{x-a}^{\infty} t e^{-t} dt\right)^{n-1} (x-a) e^{-(x-a)} I_{(a,\infty)}(x).$$

$$E(X_{(1)}) = n \int_a^{\infty} x \left(\int_{x-a}^{\infty} t e^{-t} dt\right)^{n-1} (x-a) e^{-(x-a)} dx$$

$$= n \int_0^{\infty} (y+a) \left(\int_y^{\infty} t e^{-t} dt\right)^{n-1} y e^{-y} dy \quad \text{let } y = x-a$$
$$dy = dx$$

$$= n \int_0^{\infty} y^2 e^{-y} \left(\int_y^{\infty} t e^{-t} dt\right)^{n-1} dy \quad \cdots ①$$

$$+ n a \int_0^{\infty} y e^{-y} \left(\int_y^{\infty} t e^{-t} dt\right)^{n-1} dy \quad \cdots ②$$

Note that

$$\int_y^{\infty} t e^{-t} dt = -\left[t e^{-t}\right]_y^{\infty} + \int_y^{\infty} e^{-t} dt$$
$$= y e^{-y} + e^{-y}$$

$$② = na\int_0^\infty ye^{-y}\left(ye^{-y}+e^{-y}\right)^{n-1}dy$$

let $u = ye^{-y} + e^{-y}$ when $y=0 \Rightarrow u=1$, $y\to\infty \Rightarrow u\to 0$

$$du = \left(-ye^{-y} + \cancel{e^{-y}} - \cancel{e^{-y}}\right)dy$$

$$= -ye^{-y}dy$$

$$= -na\int_1^0 u^{n-1}du$$

$$= na\int_0^1 u^{n-1}du$$

$$= \frac{n}{n}a = a$$

$\Rightarrow$ Bias of $X_{(1)}$ to a

is $① : n\int_0^\infty\left(\int_x^\infty te^{-t}dt\right)^{n-1}x^2e^{-x}dx$ //

3.

a) $h(x_1, x_2) = \frac{1}{2}[x_1 g(x_2) + x_2 g(x_1)]$

is sym. to argument.

$E[h(X_1, X_2)] = E[X_1 g(X_2)] = \mu\sigma^{-1}$

$\Rightarrow$ the u-statistic is

$$u_n = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} h(X_i, X_j) \quad //$$

b) $h_1(x) = E[h(x, X_2)]$

$$= \frac{1}{2}[x\sigma^{-1} + \mu g(x)]$$

c) $\zeta_1 = Var(h_1(X))$ $, > 0$ as h_1 is not constant.

$$= \frac{1}{4}\left[Var(X\sigma^{-1}) + 2\,Cov(X\sigma^{-1}, \mu g(X)) + Var(\mu g(X))\right]$$

$$= \frac{1}{4}\left[\sigma^{-2} Var(X) + 2\mu\sigma^{-1} Cov(X, g(X)) + \mu^2 Var(g(X))\right]$$

By Thm. 3.5 (i.)

$$\Rightarrow \sqrt{n}(u_n - \mu\sigma^{-1}) \xrightarrow{d} N(0, \sigma^{-2}Var(X_1) + 2\mu\sigma^{-1}Cov(X_1, g(X_1)) + \mu^2 Var(g(X_1))) \quad //$$

4. a)

$$Z = \begin{bmatrix} 1 & 0 & 1 \\ 1 & \frac{1}{2} & \frac{1}{2} \\ 1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \xrightarrow[R_3 - R_1]{R_2 - R_1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\therefore \mathcal{R}(Z) = \left\{ c_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \;\middle|\; c_1, c_2 \in \mathbb{R} \right\}$$

if $l = \begin{bmatrix} c_1 \\ c_2 \\ c_1 - c_2 \end{bmatrix}$ $\forall c_1, c_2 \in \mathbb{R}$ then $l^T\beta$ is estimable.

b) by Prop. 3.4 with condition (e) (as (c) $\Leftrightarrow$ (e))

LSE is BLUE iff

$Z(Z^TZ)^{-1}Z^T \, \mathrm{Var}(\varepsilon)$ is sym.

$$\Leftrightarrow \begin{bmatrix} a & u^T \\ u & C \end{bmatrix} \begin{bmatrix} r^2 & 0 \\ 0 & \sigma^2 I \end{bmatrix} \text{ is sym.}$$

$$= \begin{bmatrix} ar^2 & \sigma^2 u^T \\ r^2 u & \sigma^2 C \end{bmatrix} \text{ is sym} \quad -(*)$$

since $r^2 \neq \sigma^2$, $(*)$ is sym $\Leftrightarrow$ $u = 0$ and C is sym.

c)

$$Z^T Z = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & -1 \\ 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

$$\Rightarrow (Z^T Z)^{-1} = \frac{1}{8-4} \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

$$\Rightarrow \quad \hat{\beta} = (Z^T Z)^{-1} Z^T X$$

$$= \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & -1 & -1 \end{bmatrix} X$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} \sum_{i=1}^{4} x_i \\ \frac{1}{2}(x_3 - x_4) \end{bmatrix}$$ //