18 (d)

$$f(\vec{x};\alpha)=\prod_{i=1}^{n} f(x_i;\alpha)=\prod_{i=1}^{n}\frac{\Gamma(3\alpha)}{\Gamma(\alpha)\Gamma(2\alpha)}x_i^{\alpha-1}(1-x_i)^{2\alpha-1}$$

$$=\left(\frac{\Gamma(3\alpha)}{\Gamma(\alpha)\Gamma(2\alpha)}\right)^{n}\left(\prod x_i\right)^{\alpha-1}\left(\prod(1-x_i)\right)^{2\alpha-1}$$

$$=\prod[x_i(1-x_i)]^{-1}\exp\left\{\alpha\sum\ln x_i+2\alpha\sum\ln(1-x_i)\right\}$$

$$=\underbrace{\prod[x_i(1-x_i)]^{-1}}_{h(\vec{x})}\exp\Big\{\underbrace{\alpha}_{h(\alpha)}\underbrace{\left[\sum(\ln x_i+2\ln(1-x_i))\right]}_{T(\vec{x})}\Big\}$$

$\Rightarrow f(\vec{x};\alpha)$ is exp. family

since $\{h(\alpha)\}$ is linear indep. and can form a 1-dim. retangle space

so $T(\vec{x})=\sum[\ln x_i+2\ln(1-x_i)]$ is complete and suff. for α

21. (c)

$$f(\vec{x};\theta)=\prod f(x_i;\theta)=\prod e^{-(x_i-\theta)}\,I(\theta\le x_i)$$

$$=\underbrace{e^{-\sum x_i}}_{h(\vec{x})}\cdot\underbrace{e^{n\theta}\,I(\theta\le x_{(1)})}_{g(\theta;x_{(1)})}$$

By fact. thm $X_{(1)}=\min(X_1,\cdots,X_n)$ is suff. for θ

The p.d.f. of $X_{(1)}$ is:

$$f(x_{(1)})=n(1-F_x(x_{(1)}))^{n-1}f_x(x_{(1)})=n\cdot(e^{-(x_{(1)}-\theta)})^{n-1}\cdot e^{-(x_{(1)}-\theta)}$$

$$=ne^{-n(x_{(1)}-\theta)},\quad x_{(1)}>\theta$$

Let $E(g(x_{(1)}))=0,\ \forall\theta$

$$\Rightarrow\int_\theta^\infty g(u)\cdot ne^{-n(u-\theta)}\,du=0,\ \forall\theta$$

$\Rightarrow \frac{d}{d\theta}\int_\theta^\infty g(u)\cdot ne^{-n(u-\theta)}du = 0 \quad , \forall \theta$

$\Rightarrow -g(\theta)\cdot ne^0 + \int_\theta^\infty g(u)\frac{d}{d\theta}(ne^{-n(u-\theta)})du = 0 \quad , \forall \theta$

$\Rightarrow -g(\theta)\cdot n + n\cdot \underbrace{\int_\theta^\infty g(u)\, ne^{-n(u-\theta)}du}_{=0} = 0 \quad , \forall \theta$

$\Rightarrow g(\theta) = 0 \Rightarrow X_{(1)}$ is complete for θ by definition

$f(x;\theta) = e^{-(x-\theta)} \quad , x > \theta$

since the range of random variable X depends on unknown parameter θ, so $f(x;\theta)$ don't belong to exp. family

49 (a) Let $Y_i = \begin{cases} 1 & , \text{if } X_i \in (0,1] \\ 0 & , \text{if } X_i \in [-1,0] \end{cases}$, $i = 1, \cdots, n$

then $Y_i \overset{iid}{\sim} Ber(p)$, w.l. $p = P(X_i \in (0,1]) = \int_0^1 \frac{1+\alpha x}{2}dx$

$= \frac{1}{2} + \frac{\alpha}{4}$

$\Rightarrow \alpha = 4p - 2$

since $\bar{Y} = \frac{\sum Y_i}{n}$ is MLE for p (check Ex 6.14 in LN p.3-23)

so $\hat{\alpha} = 4\bar{Y} - 2$ is MLE for α (by the invariance propty of MLE)

We can use $4\bar{Y} - 2$ to estimate α

(b) (i) $Var(\hat{\alpha}) = 16\, Var(\bar{Y}) = \frac{16\,p(1-p)}{n} = \frac{4-\alpha^2}{n}$

(ii) $\mu_1 = E(X) = \int_{-1}^1 x\cdot\frac{1+\alpha x}{2}dx = \frac{\alpha}{3} \Rightarrow 3\mu_1 = \alpha$

Use $\frac{1}{n}\sum X_i = \bar{X}$ to estimate $\mu_1 \Rightarrow \hat{\alpha}_{MoME} = 3\bar{X}$ (Recall: LN p.3-27)

$Var(3\bar{X}) = \frac{9}{n}Var(X) = \frac{9}{n}\cdot\frac{3-\alpha^2}{9} = \frac{3-\alpha^2}{n}$ (Recall: LN p.3-61)

(iii) 因 $\hat{\alpha}_{MLE}$ 無法使用 finite-sample variance

故使用 asymptotic variance：

Fisher Information

$$I(\alpha) = E\left[-\frac{d^2}{d\alpha^2}\ln L(\alpha;\vec{x})\right]$$

$$= E\left[\sum X_i^2(1-\alpha X_i)^{-2}\right]$$

$$= n\int_{-1}^{1}\frac{x^2}{1-\alpha x}\cdot\frac{1}{2}\,dx$$

$$= \begin{cases} \frac{n[\ln(1+\alpha)-\ln(1-\alpha)-2\alpha]}{2\alpha^3}, & \alpha\in[-1,1),\ \alpha\neq 0 \\ \frac{n}{3}, & \alpha=0 \end{cases}$$

(Recall: LN p.3-61)

$$Var(\hat{\alpha}_{MLE}) = \frac{1}{I(\alpha)}$$

(iv)

$$eff(\hat{\alpha},\hat{\alpha}_{MME}) = \frac{Var(\hat{\alpha}_{MME})}{Var(\hat{\alpha})} = \frac{3-\alpha^2}{4-\alpha^2}$$

$$eff(\hat{\alpha},\hat{\alpha}_{MLE}) = \frac{Var(\hat{\alpha}_{MLE})}{Var(\hat{\alpha})} = \begin{cases} \frac{2\alpha^3}{(4-\alpha^2)[\ln(1+\alpha)-\ln(1-\alpha)-2\alpha]}, & \alpha\in[-1,1],\ \alpha\neq 0 \\ \frac{3}{4}, & \alpha=0 \end{cases}$$

(v)

α	0	.1	.2	.3	.4	.5	.6	.7	.8	.9
$eff(\hat{\alpha},\hat{\alpha}_{MME})$	0.75	0.7449	0.7448	0.7444	0.7440	0.7433	0.7425	0.7415	0.7402	0.6817
$eff(\hat{\alpha},\hat{\alpha}_{MLE})$	0.75	0.7447	0.7439	0.7425	0.7405	0.7376	0.6317	0.584	0.510	0.399

$\Rightarrow$ ∵ eff < 1 for all α, $\hat{\alpha}_{MLE}$ and $\hat{\alpha}_{MME}$ will be the better estimator for α than $\hat{\alpha}$

#69.

$X_1, \dots, X_n$ are an iid sample from a geometric distribution with parameter p.

and $f(x|p) = p(1-p)^{x-1}$, $x = 1, 2, \dots$, zero ow.

$L(p) = \prod_{i=1}^n f(x_i|p) = p^n (1-p)^{\sum_{i=1}^n x_i - n}$.

Let $T = \sum_{i=1}^n x_i$, $L(p) = p^n (1-p)^{T-n} \cdot 1$, where $g(T,p) = p^n (1-p)^{T-n}$ and $h(x) = 1$

By factorization thm, $T = \sum_{i=1}^n x_i$ is sufficient statistic for p.

#71.

1° $L(\theta) = \prod_{i=1}^n \frac{\theta}{(1+x_i)^{\theta+1}} = \theta^n \prod_{i=1}^n (1+x_i)^{-(\theta+1)} = \theta^n \left(\prod_{i=1}^n (1+x_i)\right)^{-1} \left(\prod_{i=1}^n (1+x_i)\right)^{-\theta}$

Let $T = \prod_{i=1}^n (1+x_i)$, then $L(\theta) = g(T,\theta) \cdot h(x)$

Then by factorization thm, $T = \prod_{i=1}^n (1+x_i)$ is a sufficient statistic for θ.

2° $f(x|\theta) = \theta (1+x)^{-(\theta+1)} = \underbrace{\theta}_{c(\theta)} \cdot \underbrace{(1+x)^{-1}}_{h(x)} \exp\{ \underbrace{-\theta}_{w(\theta)} \underbrace{\ln(1+x)}_{t(x)} \}$

This is a 1-parameter exponential family.

3° For an exponential family in canonical form,

$S = \sum_{i=1}^n t(x_i)$ is complete if the range of $w(\theta)$ contains a open interval

$S = \sum_{i=1}^n \ln(1+x_i) = \ln\left(\prod_{i=1}^n (1+x_i)\right) = \ln(T)$

Since T is a one-to-one function of S

Then $T = \prod_{i=1}^n (1+x_i)$ is a sufficient and complete statistic for θ.

#72.

$f(x|\alpha,\beta) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\Gamma(\alpha)\beta^\alpha}$, $0 < x < \infty$, zero ow.

$= \underbrace{\frac{1}{\Gamma(\alpha)\beta^\alpha}}_{c(\alpha,\beta)} \exp\{ \underbrace{(\alpha-1)}_{w_1(\alpha)} \underbrace{\ln x}_{t_1(x)} - \underbrace{\frac{1}{\beta}}_{w_2(\beta)} \underbrace{x}_{t_2(x)} \}$, This is a 2-parameter exponential family.

For an iid sample, $L(\alpha,\beta) = \frac{1}{(\Gamma(\alpha)\beta^\alpha)^n} \exp\{ (\alpha-1) \sum_{i=1}^n \ln x_i - \frac{1}{\beta} \sum_{i=1}^n x_i \}$

Then $T_1 = \sum_{i=1}^{n} \ln x_i$, $T_2 = \sum_{i=1}^{n} x_i$

since $\sum_{i=1}^{n} \ln x_i = \ln \prod_{i=1}^{n} x_i$, the set $(\prod_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i)$ provides the same information

$\Rightarrow (\prod_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i)$ are sufficient statistics for (α, β) by factorization thm.

Also, since it's an exponential family and the parameter space for (α, β) is $\{(\alpha, \beta): \alpha > 0, \beta > 0\}$, which contains 2D open set in $\mathbb{R}^2$, then $\prod_{i=1}^{n} x_i$ and $\sum_{i=1}^{n} x_i$ are sufficient and complete.