# 고급수리통계학 기말고사

2020221005

오재권

2021. 6. 13.

1. Given the pdf

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

show that the Rao-Cramer lower bound is $2/n$, where $n$ is the size of a random sample from this Cauchy distribution. What is the asymptotic distribution of $\sqrt{n}(\hat{\theta}-\theta)$ if $\hat{\theta}$ is the mle of $\theta$?

(Solution)

**Fisher Information**

$$\begin{aligned}
I(\theta) &= E\left[\left(\frac{\partial \log f(x;\theta)}{\partial\theta}\right)^2\right] \\
&= E\left[\left(\frac{\partial(-\log\pi - \log(1+(1-\theta)^2))}{\partial\theta}\right)^2\right] \\
&= E\left[\left(\frac{2(x-\theta)}{1+(x-\theta)}\right)^2\right] \\
&= \int_{-\infty}^{\infty}\left(\frac{2(x-\theta)}{1+(x-\theta)^2}\right)^2 \frac{1}{\pi(1+(x-\theta)^2)}dx \\
&= \frac{4}{\pi}\int_{-\infty}^{\infty}\frac{(x-\theta)^2}{(1+(x-\theta)^2)^3}dx, \qquad t = x-\theta,\ dt = dx \\
&= \frac{4}{\pi}\int_{-\infty}^{\infty}\frac{t^2}{(1+t^2)^3}dt \\
&= \frac{8}{\pi}\int_{0}^{\infty}\frac{t^2}{(1+t^2)^3}dt \\
&= \frac{8}{\pi}\int_{0}^{\infty}\frac{t^2}{(1+t^2)}\left(\frac{1}{1+t^2}\right)^2 dt, \qquad u = \frac{1}{1+t^2},\ t = \left(\frac{1}{u}-1\right)^{\frac{1}{2}},\ dt = \frac{1}{2}\left(\frac{1}{u}-1\right)^{-\frac{1}{2}}\left(-\frac{1}{u^2}\right)du \\
&= \frac{8}{\pi}\int_{0}^{1} -(1-u)u^2\frac{1}{2}\left(\frac{1}{u}-1\right)^{-\frac{1}{2}}\left(-\frac{1}{u^2}\right)du \\
&= \frac{4}{\pi}\int_{0}^{1}(1-u)\left(\frac{1}{u}-1\right)^{-\frac{1}{2}}du \\
&= \frac{4}{\pi}\int_{0}^{1}(1-u)\left(\frac{u}{1-u}\right)^{\frac{1}{2}}du \\
&= \frac{4}{\pi}\int_{0}^{1}u^{\frac{1}{2}}(1-u)^{\frac{1}{2}}du \\
&= \frac{4}{\pi}\int_{0}^{1}u^{\frac{3}{2}-1}(1-u)^{\frac{3}{2}-1}du \qquad \text{(Beta integral)} \\
&= \frac{4}{\pi}\frac{\Gamma\left(\frac{3}{2}\right)+\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{3}{2}+\frac{3}{2}\right)} \\
&= \frac{4}{\pi}\frac{(0.5\sqrt{\pi})^2}{2!} \\
&= \frac{1}{2}
\end{aligned}$$

**Rao-Cramer lower bound**

$$RC_{lb} = \frac{1}{nI(\theta)} = \frac{2}{n}$$

**Asymptotic distribution of mle**

$$\begin{aligned}
\sqrt{n}(\hat{\theta} - \theta) &\xrightarrow{d} N\left(0, \frac{1}{I(\theta)}\right) \\
&\xrightarrow{d} N(0, 2)
\end{aligned}$$

2. Let $X$ be $N(0, \theta)$, $0 < \theta < \infty$.

$$\begin{aligned}
f(x;\theta) &= \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{x^2}{2\theta}} \\
\log f(x;\theta) &= -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\theta - \frac{x^2}{2\theta}
\end{aligned}$$

(a) Find the Fisher information $I(\theta)$.

(Solution)

$$\begin{aligned}
I(\theta) &= -E\left[\frac{\partial^2 \log f(x;\theta)}{\partial\theta^2}\right] \\
&= -E\left[\frac{\partial}{\partial\theta}\left(-\frac{1}{2\theta} + \frac{x^2}{2\theta^2}\right)\right] \\
&= -E\left[\frac{1}{2\theta^2} - \frac{x^2}{\theta^3}\right] \\
&= -\frac{1}{2\theta^2} + \frac{1}{\theta^2}E\left(\frac{x^2}{\theta}\right), \qquad Z^2 = \frac{(x-\mu)^2}{\theta} = \frac{x^2}{\theta} \sim \chi^2_{(1)},\ \mu = 0 \\
&= -\frac{1}{2\theta^2} + \frac{1}{\theta^2} \times 1 \\
&= \frac{1}{2\theta^2}
\end{aligned}$$

(b) If $X_1, \ldots, X_n$ is a random sample from this distribution, show that the mle of $\theta$, $\hat{\theta}$, is an efficient estimator of $\theta$.

(Solution)

**mle of $\theta$**

$$\begin{aligned}
L(\theta) &= \Pi_{i=1}^n f(x_i;\theta) = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n e^{-\frac{\sum_{i=1}^n x_i^2}{2\theta}} \\
\ell(\theta) &= \log(L(\theta)) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\theta - \frac{\sum_{i=1}^n x_i^2}{2\theta}
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial \ell(\theta)}{\partial \theta} &= -\frac{n}{2}\cdot\frac{1}{\theta} + \frac{\sum_{i=1}^{n} x_i^2}{2\theta^2} \overset{set}{=} 0 \\
&\Leftrightarrow\ -n\theta + \sum_{i=1}^{n} x_i^2 = 0 \\
&\therefore\ \hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i^2
\end{aligned}
$$

**$\hat{\theta}$의 평균과 분산**

$$
\begin{aligned}
E(\hat{\theta}) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) \\
&= \frac{1}{n}(E(X_1^2) + \cdots + E(X_n^2)) \\
&= \frac{1}{n} n E(X_1^2) \\
&= (Var(X) + (E(X))^2), \qquad E(X) = 0,\ Var(X) = \theta \\
&= \theta
\end{aligned}
$$

$\Rightarrow$ $\hat{\theta}$ 는 $\theta$에 대한 unbiased estimator 이다.

$$
\begin{aligned}
Var(\hat{\theta}) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) \\
&= \frac{1}{n^2}(Var(X_1^2) + \cdots + Var(X_n^2)) \\
&= \frac{1}{n}Var(X_1^2), \qquad \frac{X^2}{\theta} \sim \chi^2_{(1)},\ Var\left(\frac{X^2}{\theta}\right) = 2 \\
&= \frac{1}{n}2\theta^2
\end{aligned}
$$

**Rao-Cramer lower bound**

$$
RC_{lb} = \frac{1}{nI(\theta)} = \frac{2\theta^2}{n}
$$

$\Rightarrow$ $Var(\hat{\theta}) = RC_{lb}$ 이므로 $\hat{\theta}$는 $\theta$에 대한 efficient estimator 이다.

(c) What is the asymptotic distribution of $\sqrt{n}(\hat{\theta} - \theta)$?

(Solution)

**Asymptotic distribution**

$$
\begin{aligned}
\sqrt{n}(\hat{\theta} - \theta) &\xrightarrow{d} N\left(0, \frac{1}{I(\theta)}\right) \\
&\xrightarrow{d} N\left(0, 2\theta^2\right)
\end{aligned}
$$

3. If $X_1, \ldots, X_n$ is a random sample from a distribution with pdf

$$f(x;\theta) = \begin{cases} \frac{3\theta^3}{(x+\theta)^4} & 0 < x < \infty,\ 0 < \theta < \infty \\ 0 & \text{elsewhere,} \end{cases}$$

show that $Y = 2\bar{X}$ is an unbiased estimator of $\theta$ and determine its efficiency.

(Solution)

**$Y$의 평균과 분산**

$$\begin{aligned}
E(Y) &= E(2\bar{X}) = \frac{2}{n}E(\sum_{i=1}^{n} X_i) \\
E(X) &= \int_0^\infty x\frac{3\theta^3}{(x+\theta)^4}dx \\
&= 3\theta^3\int_0^\infty \frac{x}{(x+\theta)^4}dx \qquad t = x+\theta,\ dt = dx \\
&= 3\theta^3\int_\theta^\infty \frac{t-\theta}{t^4}dt \\
&= 3\theta^3\left[\int_\theta^\infty \frac{1}{t^3}dt - \theta\int_\theta^\infty \frac{1}{t^4}dt\right] \\
&= 3\theta^3\left[\left[-\frac{1}{2t^2}\right]_\theta^\infty + \left[\frac{\theta}{3t^3}\right]_\theta^\infty\right] \\
&= 3\theta^3\left(\frac{1}{2\theta^2} - \frac{1}{3\theta^2}\right) \\
&= 3\theta^3\frac{1}{6\theta^2} \\
&= \frac{\theta}{2} \\
E(Y) &= \frac{2}{n}\cdot n\cdot\frac{\theta}{2} = \theta
\end{aligned}$$

$\Rightarrow$ $Y$는 $\theta$에 대한 unbiased estimator 이다.

$$\begin{aligned}
Var(Y) &= Var(2\bar{X}) = \frac{4}{n^2}Var\left(\sum_{i=1}^{n} X_i\right) = \frac{4}{n}Var(X) \\
Var(X) &= E(X^2) - \{E(X)\}^2 \\
E(X^2) &= \int_0^\infty x^2\frac{3\theta^3}{(x+\theta)^4}dx \\
&= 3\theta^3\int_0^\infty \frac{x^2}{(x+\theta)^4}dx \qquad t = x+\theta,\ dt = dx \\
&= 3\theta^3\int_\theta^\infty \frac{(t-\theta)^2}{t^4}dt \\
&= 3\theta^3\left[\int_\theta^\infty \frac{1}{t^2}dt - \int_\theta^\infty \frac{2\theta}{t^3}dt + \int_\theta^\infty \frac{\theta^2}{t^4}dt\right]
\end{aligned}$$

$$
\begin{aligned}
&= 3\theta^3 \left[ \left[ -\frac{1}{t} \right]_\theta^\infty - \left[ -\frac{\theta}{t^2} \right]_\theta^\infty + \left[ -\frac{\theta^2}{3t^3} \right]_\theta^\infty \right] \\
&= 3\theta^3 \left( \frac{1}{\theta} - \frac{1}{\theta} + \frac{1}{3\theta} \right) \\
&= \theta^2 \\
Var(X) &= \theta^2 - \frac{\theta^2}{4} = \frac{3}{4}\theta^2 \\
Var(Y) &= \frac{4}{n} \cdot \frac{3}{4}\theta^2 = \frac{3}{n}\theta^2
\end{aligned}
$$

**Fisher Information**

$$
\begin{aligned}
I(\theta) &= -E\left[ \frac{\partial^2 \log f(x;\theta)}{\partial \theta^2} \right] \\
&= -E\left[ \frac{\partial^2}{\partial \theta^2} (3\log\theta - 4\log(x+\theta)) \right] \\
&= -E\left[ \frac{\partial}{\partial \theta} \left( \frac{3}{\theta} - \frac{4}{x+\theta} \right) \right] \\
&= -E\left[ -\frac{3}{\theta^2} + \frac{4}{(x+\theta)^2} \right] \\
&= \frac{3}{\theta^2} - 4E\left[ \frac{1}{(x+\theta)^2} \right] \\
E\left[ \frac{1}{(x+\theta)^2} \right] &= \int_0^\infty \frac{1}{(x+\theta)^2} \cdot \frac{3\theta^3}{(x+\theta)^4} dx \\
&= \int_0^\infty \frac{3\theta^3}{(x+\theta)^6} dx, \qquad x+\theta = t,\ dt = dx \\
&= \int_\theta^\infty \frac{3\theta^3}{t^6} dt \\
&= \left[ -\frac{3\theta^3}{5t^5} \right]_\theta^\infty \\
&= \frac{3\theta^3}{5\theta^5} \\
&= \frac{3}{5\theta^2} \\
I(\theta) &= \frac{3}{\theta^2} - 4 \cdot \frac{3}{5\theta^2} = \frac{3}{5\theta^2}
\end{aligned}
$$

**Rao-Cramer lower bound**

$$
RC_{lb} = \frac{1}{nI(\theta)} = \frac{5\theta^2}{3n}
$$

**Efficiency**

$$
\text{Efficiency} = \frac{RC_{lb}}{Var(Y)} = \frac{\frac{5\theta^2}{3n}}{\frac{3}{n}\theta^2} = \frac{5}{9} < 1
$$

$\Rightarrow Var(Y) \neq RC_{lb}$ 이므로 $Y$ 는 $\theta$에 대한 efficient estimator가 아니다.

4. Let $\bar{X}$ be the mean of a random sample of size $n$ from a $N(\theta, \sigma^2)$ distribution, $-\infty < \theta < \infty$, $\sigma^2 > 0$. Assume that $\sigma^2$ is known. Show that $\bar{X}^2 - \frac{\sigma^2}{n}$ is an unbiased estimator of $\theta^2$ and find its efficiency.

(Solution)

$$
\begin{aligned}
E\left(\bar{X}^2 - \frac{\sigma^2}{n}\right) &= E(\bar{X}^2) - \frac{\sigma^2}{n} \\
&= Var(\bar{X}) + \left\{E(\bar{X})\right\}^2 - \frac{\sigma^2}{n} \\
&= \frac{\sigma^2}{n} + \theta^2 - \frac{\sigma^2}{n} \\
&= \theta^2
\end{aligned}
$$

$\Rightarrow$ $\bar{X}^2 - \frac{\sigma^2}{n}$ 은 $\theta^2$에 대한 unbiased estimator이다.

**Fisher Information**

$$
\begin{aligned}
I(\theta^2) &= -E\left[\frac{\partial^2 \log f(x;\theta^2)}{\partial(\theta^2)^2}\right] \\
&= -E\left[\frac{\partial^2}{\partial(\theta^2)^2}\left(-\frac{1}{2}\log(2\pi\sigma^2) - \frac{(x-\theta)^2}{2\sigma^2}\right)\right] \\
&= -E\left[\frac{\partial^2}{\partial(\theta^2)^2}\left(-\frac{1}{2}\log(2\pi\sigma^2) - \frac{x^2}{2\sigma^2} + \frac{2x\sqrt{\theta^2}}{2\sigma^2} - \frac{\theta^2}{2\sigma^2}\right)\right] \\
&= -E\left[\frac{\partial}{\partial\theta^2}\left(\frac{x}{2\sigma^2\sqrt{\theta^2}} - \frac{1}{2\sigma^2}\right)\right] \\
&= -E\left[-\frac{x}{4\sigma^2(\theta^2)^{\frac{3}{2}}}\right] \\
&= \frac{1}{4\sigma^2(\theta^2)^{\frac{3}{2}}} \cdot E(X) \\
&= \frac{1}{4\sigma^2(\theta^2)^{\frac{3}{2}}} \cdot \theta \\
&= \frac{1}{4\sigma^2\theta^2}
\end{aligned}
$$

$$
\begin{aligned}
Var\left(\bar{X}^2 - \frac{\sigma^2}{n}\right) &= Var(\bar{X}^2) \\
&= E(\bar{X}^4) - \left\{E(\bar{X}^2)\right\}
\end{aligned}
$$

$\star$ **Moment of normal distribution**

$$
\begin{aligned}
&E(X^2) = \theta^2 + \sigma^2,\ E(X^4) = \theta^4 + 6\theta^2\sigma^2 + 3\sigma^4 \\
&= \theta^4 + 6\theta^2\frac{\sigma^2}{n} + 3\frac{\sigma^4}{n^2} - \left(\theta^2 + \frac{\sigma^2}{n}\right)^2 \\
&= 4\theta^2\frac{\sigma^2}{n} + 2\frac{\sigma^4}{n^2}
\end{aligned}
$$

**Rao-Cramer lower bound**

$$RC_{lb} = \frac{1}{nI(\theta)} = \frac{4\theta^2\sigma^2}{n}$$

**Efficiency**

$$\text{Efficiency} = \frac{RC_{lb}}{Var\left(\bar{X}^2 - \frac{\sigma^2}{n}\right)} = \frac{\frac{4\theta^2\sigma^2}{n}}{4\theta^2\frac{\sigma^2}{n} + 2\frac{\sigma^4}{n^2}} = \frac{4\theta^2}{4\theta^2 + 2\frac{\sigma^2}{n}} < 1, \quad \sigma^2 > 0$$

$\Rightarrow Var\left(\bar{X}^2 - \frac{\sigma^2}{n}\right) \neq RC_{lb}$ 이므로 $\bar{X}^2 - \frac{\sigma^2}{n}$ 은 $\theta^2$에 대한 efficient estimator가 아니다.

5. Let $X_1, \ldots, X_n \overset{iid}{\sim} N(0, \theta),\ 0 < \theta < \infty$. Show that $\sum_{i=1}^n X_i^2$ is a sufficient statistic for $\theta$.

(Solution)

**Factorization**

$$Y = \sum_{i=1}^n X_i^2 = u_1(x_1, \ldots, x_n)$$

**joint pdf**

$$f(x_1;\theta)\cdots f(x_n;\theta) = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n e^{-\frac{\sum_{i=1}^n x_i^2}{2\theta}}$$

$$\left(\frac{1}{\sqrt{\theta}}\right)^n = \theta^{-\frac{n}{2}} = e^{\log\theta^{-\frac{n}{2}}} = e^{-\frac{n}{2}\log\theta}$$

$$f(x_1;\theta)\cdots f(x_n;\theta) = \underbrace{\left(\frac{1}{\sqrt{2\pi}}\right)^n}_{k_2(x_1,\ldots,x_n)} \underbrace{e^{-\frac{\sum_{i=1}^n x_i^2}{2\theta} - \frac{n}{2}\log\theta}}_{k_1(u(x_1,\ldots,x_n;\theta))}$$

$\Rightarrow Y = \sum_{i=1}^n X_i^2$은 $\theta$에 대한 sufficient statistic 이다.

6. What is the sufficient statistic for $\theta$ if the sample arises from a beta distribution in which $\alpha = \beta = \theta > 0$?

(Solution)

$$X_1, \ldots, X_n \overset{iid}{\sim} Beta(\theta, \theta), \quad \theta > 0$$

**pdf**

$$\begin{aligned} f(x;\theta) &= \frac{\Gamma(2\theta)}{\Gamma(\theta)\Gamma(\theta)} x^{\theta-1}(1-x)^{\theta-1} \\ &= \frac{\Gamma(2\theta)}{(\Gamma(\theta))^2}[x(1-x)]^{\theta-1}, \quad x \in [0,1] \end{aligned}$$

**joint pdf**

$$f(x_1;\theta),\ldots,f(x_n;\theta)=\left(\frac{\Gamma(2\theta)}{(\Gamma(\theta))^2}\right)^n\cdot(\Pi_{i=1}^n x_i(1-x_i))^{\theta-1}$$
$$=\underbrace{1}_{k_2(x_1,\ldots,x_n)}\times\underbrace{\left(\frac{\Gamma(2\theta)}{(\Gamma(\theta))^2}\right)^n\cdot(\Pi_{i=1}^n x_i(1-x_i))^{\theta-1}}_{k_1(u_1(x_1,\ldots,x_n);\theta)}$$

$$\therefore\ u(x_1,\ldots,x_n)=\Pi_{i=1}^n x_i(1-x_i)$$

$\Rightarrow \Pi_{i=1}^n x_i(1-x_i)$ 는 $\theta$에 대한 sufficient statistic이다.

7. Let $Y_1<Y_2<Y_3<Y_4<Y_5$ be the order statistics of a random sample of size 5 from the uniform distribution having pdf $f(x;\theta)=1/\theta,\ 0<x<\theta,\ 0<\theta<\infty$, zero elsewhere.

(a) Show that $2Y_3$ is an unbiased estimator of $\theta$.

(Solution)

$$\begin{aligned}f_{Y_k}(y_k)&=\frac{n!}{(k-1)!(n-k)!}f(y_k)[F(y_k)]^{k-1}[1-F(y_k)]^{n-k}\\ f_{Y_3}(y_3)&=\frac{5!}{2!2!}\ \frac{1}{\theta}\ \left(\frac{y_3}{\theta}\right)^{3-1}\left(1-\frac{y_3}{\theta}\right)^{5-3}\\ &=\frac{30}{\theta^3}\ y_3^2\left(1-\frac{y_3}{\theta}\right)^2\end{aligned}$$

$$\begin{aligned}E(2y_3)&=\int_0^\theta 2y_3\cdot\frac{30}{\theta^3}y_3^2\left(1-\frac{y_3}{\theta}\right)^2 dy_3\\ &=60\int_0^\theta\frac{y_3^3}{\theta^3}\left(1-\frac{y_3}{\theta}\right)^2 dy_3 \qquad t=\frac{y_3}{\theta},\ dt=\frac{1}{\theta}dy_3\\ &=60\ \theta\int_0^1 t^3(1-t)^2dt\\ &=60\ \theta\int_0^1 t^{4-1}(1-t)^{3-1}dt \qquad \text{(Beta integral)}\\ &=60\ \theta\ \frac{\Gamma(4)+\Gamma(3)}{\Gamma(4+3)}\\ &=60\ \theta\ \frac{6\times 2}{720}\\ &=\theta\end{aligned}$$

$\Rightarrow 2Y_3$는 $\theta$에 대한 unbiased estimator 이다.

(b) Determine the joint pdf of $Y_3$ and the sufficient statistic $Y_5$ of $\theta$.

(Solution)

$$
\begin{aligned}
f_{Y_3,Y_5}(y_3,y_5) &= \frac{5!}{(3-1)!(5-3-1)!(5-5)!}\left(\frac{y_3}{\theta}\right)^{3-1}\left(\frac{y_5}{\theta}-\frac{y_3}{\theta}\right)^{5-3-1}\left(1-\frac{y_5}{\theta}\right)^{5-5}\frac{1}{\theta}\frac{1}{\theta} \\
&= 60\,\frac{y_3^2(y_5-y_3)}{\theta^5}, \qquad 0<y_3<y_5<\theta
\end{aligned}
$$

(c) Find the conditional expectation $E[2Y_3|y_5]=\rho(y_5)$.

(Solution)

$$
\begin{aligned}
E[2Y_3|Y_5=y_5] &= 2E[Y_3|Y_5=y_5] \\
&= 2\int_0^{y_5} y_3 f_{Y_3|Y_5}(y_3|y_5)dy_3 \\
f_{Y_3|Y_5}(y_3|y_5) &= \frac{f_{Y_3,Y_5}(y_3,y_5)}{f_{Y_5}(y_5)} \\
f_{Y_5}(y_5) &= \frac{5!}{4!}\,\frac{1}{\theta}\left(\frac{y_5}{\theta}\right)^{5-1}\left(1-\frac{y_5}{\theta}\right)^{5-5} \\
&= \frac{5}{\theta^5}\,y_5^4 \\
f_{Y_3|Y_5}(y_3|y_5) &= \frac{60\,\frac{y_3^2(y_5-y_3)}{\theta^5}}{\frac{5}{\theta^5}y_5^4} \\
&= 12\,\frac{y_3^2y_5-y_3^3}{y_5^4}
\end{aligned}
$$

$$
\begin{aligned}
E[2Y_3|Y_5=y_5] &= 2\int_0^{y_5} y_3 f_{Y_3|Y_5}(y_3|y_5)dy_3 \\
&= 2\int_0^{y_5} y_3 12\,\frac{y_3^2y_5-y_3^3}{y_5^4}dy_3 \\
&= \frac{24}{y_5^4}\int_0^{y_5} y_5y_3^3-y_3^4dy_3 \\
&= \frac{24}{y_5^4}\left[\frac{y_5y_3^4}{4}-\frac{y_3^5}{5}\right]_0^{y_5} \\
&= \frac{24}{y_5^4}\left[\frac{y_5^5}{4}-\frac{y_5^5}{5}\right] \\
&= \frac{24}{y_5^4}\frac{y_5^5}{20} \\
&= \frac{6}{5}y_5=\rho(y_5)
\end{aligned}
$$

(d) Compare the variances of $2Y_3$ and $\rho(y_5)$.

(Solution)

**variances of** $2Y_3$

$$
\begin{aligned}
Var(2Y_3) &= E((2Y_3)^2) - (E(2Y_3))^2 \\
E((2Y_3)^2) &= 4\int_0^\theta y_3^2 f_{y_3}(y_3)dy_3 \\
&= 4\int_0^\theta y_3^2 \frac{30}{\theta^3} y_3^2 \left(1-\frac{y_3}{\theta}\right)^2 dy_3 \\
&= 120\int_0^\theta \theta\frac{y_3^4}{\theta^4} \left(1-\frac{y_3}{\theta}\right)^2 dy_3 \qquad t=\frac{y_3}{\theta}, dt=\frac{1}{\theta}dy_3 \\
&= 120\int_0^1 \theta^2 t^4 (1-t)^2 dt \\
&= 120\,\theta^2 \int_0^1 t^{5-1}(1-t)^{3-1}dt \qquad \text{(Beta integral)} \\
&= 120\,\theta^2\ \frac{\Gamma(5)\times\Gamma(3)}{\Gamma(5+3)} = \frac{8}{7}\theta^2 \\
\therefore\ Var(2Y_3) &= \frac{8}{7}\theta^2 - \theta^2 = \frac{1}{7}\theta^2
\end{aligned}
$$

**variances of** $\rho(y_5)$

$$
\begin{aligned}
Var(\rho(y_5)) &= E\left(\left(\frac{6}{5}y_5\right)^2\right) - \left(E\left(\frac{6}{5}y_5\right)\right)^2 \\
E(Y_5) &= \int_0^\theta y_5 \frac{5}{\theta^5} y_5^4 dy_5 = \int_0^\theta \frac{5}{\theta^5} y_5^5 \\
&= \frac{5}{\theta^5}\left[\frac{y_5^6}{6}\right]_0^\theta = \frac{5}{6}\theta
\end{aligned}
$$

$$
\begin{aligned}
E(Y_5^2) &= \int_0^\theta y_5^2 \frac{5}{\theta^5} y_5^4 dy_5 = \int_0^\theta \frac{5}{\theta^5} y_5^6 \\
&= \frac{5}{\theta^5}\left[\frac{y_5^7}{7}\right]_0^\theta = \frac{5}{7}\theta^2 \\
\therefore\ Var(\rho(y_5)) &= \frac{36}{25}\cdot\frac{5}{7}\theta^2 - \theta^2 = \frac{1}{35}\theta^2
\end{aligned}
$$

$$
\therefore\ Var(2Y_3) = \frac{1}{7}\theta^2\ >\ Var(\rho(y_5)) = \frac{1}{35}\theta^2
$$

8. Let $X_1, \ldots, X_n$ be a random sample from a Poisson distribution with mean $\theta$. Find the conditional expectation $E[X_1 + 2X_2 + 3X_3 | \sum_{i=1}^n X_i]$.

(Solution)

$$
\begin{aligned}
E[X_1+2X_2+3X_3|\sum_{i=1}^{n}X_i] &= E(X_1|\sum X_i)+2E(X_2|\sum X_i)+3E(X_3|\sum X_i) \\
&= 6E(X_1|\sum_{i=1}^{n}X_i)
\end{aligned}
$$

$$
\begin{aligned}
E(X_1|\sum X_i=t) &= \sum_{k=0}^{t} kP(X_1=k|\sum_{i=1}^{n}X_i=t) \\
&= \sum_{k=0}^{t} k\frac{P(X_1=k,\sum_{i=1}^{n}X_i=t)}{P(\sum_{i=1}^{n}X_i=t)} \\
&= \sum_{k=0}^{t} k\frac{P(X_1=k,\sum_{k=2}^{n}X_i=t-k)}{P(\sum_{i=1}^{n}X_i=t)} \\
&= \sum_{k=0}^{t} k\frac{\frac{\theta^k}{k!}e^{-\theta}\frac{((n-1)\theta)^{t-k}}{(t-k)!}e^{-(n-1)\theta}}{\frac{(n\theta)^t}{t!}e^{-n\theta}} \\
&= \sum_{k=0}^{t} k\frac{t!}{k!(t-k)!}\frac{\theta^k((n-1)\theta)^{t-k}}{(n\theta)^t} \\
&= \sum_{k=0}^{t} k\binom{t}{k}\frac{(n-1)^t\theta^t\theta^k}{n^t(n-1)^k\theta^t\theta^k} \\
&= \left(1-\frac{1}{n}\right)^t\sum_{k=0}^{t} k\binom{t}{k}\frac{1}{(n-1)^k} \\
E[X_1+2X_2+3X_3|\sum_{i=1}^{n}X_i=t] &= 6E(X_1|\sum_{i=1}^{n}X_i=t) \\
&= 6\left(1-\frac{1}{n}\right)^t\sum_{k=0}^{t} k\binom{t}{k}\frac{1}{(n-1)^k}
\end{aligned}
$$

9. If $az^2+bz+c=0$ for more than two values of $z$, then $a=b=c=0$. Use this result to show that the family $\{Beta(2,\theta):0<\theta<1\}$ is complete.
(Solution)

$$X_1, X_2 \overset{iid}{\sim} Bern(\theta), \quad \theta \in [0.1]$$

Let

$$S = X_1 + X_2 \sim Bin(2,\theta)$$

pmf of $S$

$$P(S=s) = \binom{2}{s}\theta^s(1-\theta)^{2-s}, \quad s \in \{0,1,2\}$$

Let $u$ be any (measurable) function such that $E(u(S)) = 0$. We need to show that $u \equiv 0$ (almost surely, $u \neq 0$ only on a set of points with probability zero).

$$\begin{aligned}
0 = E(u(S)) &= \sum_{k=0}^{2} u(k)P(S=k) = \sum_{k=0}^{2} u(k)\binom{2}{k}\theta^k(1-\theta)^{2-k} \\
&= u(0)(1-\theta)^2 + 2u(1)\theta(1-\theta) + u(2)\theta^2 \\
&= u(0)(\theta^2 - 2\theta + 1) + 2u(1)(\theta - \theta^2) + u(2)\theta^2 \\
&= (u(0) - 2u(1) + u(2))\theta^2 + (-2u(0) + 2u(1))\theta + u(0) \\
&\Rightarrow \begin{cases} u(0) = 0 \\ (-2u(0) + 2u(1)) \ \Rightarrow \ u(1) = 0 \\ (u(0) - 2u(1) + u(2)) \ \Rightarrow \ u(2) = 0 \end{cases} \\
&\therefore \ u(0) = u(1) = u(2) = 0 \ \Rightarrow \ u = 0
\end{aligned}$$

The family $\{Beta(2,\theta) : 0 < \theta < 1\}$ is complete.

10. Let a random sample of size $n$ be taken from a distribution of the discrete type with pmf $f(x;\theta) = 1/\theta,\ x = 1, 2, \ldots, \theta$, zero elsewhere, where $\theta$ is an unknown positive integer.

    (a) Show that the largest observation, say $Y$, of the sample is a complete sufficient statistic for $\theta$.

    (Solution)

    $$Y = u_1(x_1, \ldots, x_n) = \max_{1\le i\le n} x_i$$

    **joint pdf**

    $$\begin{aligned}
    f(x_1, \ldots, x_n; \theta) &= \frac{1}{\theta^n} I_{\{1,2,\ldots,\theta\}}(x_1) \cdots I_{\{1,2,\ldots,\theta\}}(x_n) \\
    &= \frac{1}{\theta^n} I_{\{1,2,\ldots,\theta\}}(\max_{1\le i\le n} x_i) \\
    &= \underbrace{1}_{k_2(x_1,\ldots,x_n)} \cdot \underbrace{\frac{1}{\theta^n} I_{\{1,2,\ldots,\theta\}}(\max_{1\le i\le n} x_i)}_{k_1(u(x_1,\ldots,x_n);\theta)}
    \end{aligned}$$

    $\therefore Y = \max_{1\le i\le n} x_i$ 는 $\theta$에 대한 sufficient statistic 이다.

    **Calculate pmf of $Y$**

    cdf of $X$

    $$F_X(x;\theta) = \sum_{i=1}^{x} P(X=x) = \sum_{i=1}^{x} \frac{1}{\theta} = \frac{x}{\theta}, \quad x \in \{1, 2, \ldots, \theta\}$$

cdf of $Y$

$$
\begin{aligned}
F_Y(k) = P(Y=k) &= P(\max_{1\le i\le n} x_i \le k) \\
&= P(x_1 \le k, x_2 \le k, \ldots, x_n \le k) \\
&= P(x_1 \le k)\cdots P(x_n \le k) \\
&= [F_X(k;\theta)]^n \\
&= \left(\frac{k}{\theta}\right)^n
\end{aligned}
$$

**pmf of $Y$**

$$
f_Y(y) = F_Y(k) - F_Y(k-1) = \left(\frac{k}{\theta}\right)^n - \left(\frac{k-1}{\theta}\right)^n = \frac{k^n - (k-1)^n}{\theta^n}, \quad k \in \{1,2,\ldots,\theta\}
$$

Let $u$ be any (measurable) function such that $E(u(Y)) = 0$, for all $\theta \in \mathbb{N}$.

$$
\begin{aligned}
0 &= E(u(Y)) = \sum_{k=1}^{\theta} u(k)P(Y=k) = \sum_{k=1}^{\theta} u(k)\frac{k^n-(k-1)^n}{\theta^n} \\
0 &= \sum_{k=1}^{\theta} u(k)(k^n-(k-1)^n), \quad \text{for all } \theta \in \mathbb{N}
\end{aligned}
$$

Let's show by induction that $u(k) = 0$, for all $k \in \mathbb{N}$.

- $\theta = 1$
$$0 = u(1)(1-0) = u(1) \;\Rightarrow\; u(1) = 0$$

- Let's assume that $u(k) = 0$, for all $1 \le k \le m$. We prove that $u(m+1) = 0$.
$$0 = \sum_{k=1}^{m+1} u(k)(k^n-(k-1)^n) = u(m+1)\cdot \underbrace{((m+1)^n - m^n)}_{\neq 0}$$
$$\therefore\; u(m+1) = 0$$

- We have shown that $u(k) = 0$, for all $k \in \mathbb{N}$.
$$u = 0$$

- $Y$ is a complete statistic.

(b) Prove that
$$\frac{Y^{n+1}-(Y-1)^{n+1}}{Y^n-(Y-1)^n}$$
is the unique MVUE of $\theta$.
(Solution)

$Y$ is a complete sufficient statistic.

$$U = \frac{Y^{n+1} - (Y-1)^{n+1}}{Y^n - (Y-1)^n}$$

$$\begin{aligned}
E(U) &= \sum_{k=1}^{\theta} \left( \frac{k^{n+1} - (k-1)^{n+1}}{k^n - (k-1)^n} \right) P(Y = k) \\
&= \sum_{k=1}^{\theta} \left( \frac{k^{n+1} - (k-1)^{n+1}}{k^n - (k-1)^n} \right) \frac{k^n - (k-1)^n}{\theta^n} \\
&= \frac{1}{\theta^n} \sum_{k=1}^{\theta} (k^{n+1} - (k-1)^{n+1}) \\
&= \frac{1}{\theta^n} [(1^{n+1} - 0^{n+1}) + (2^{n+1} - 1^{n+1}) + \cdots + ((\theta-1)^{n+1} - (\theta-2)^{n+1}) \\
&\qquad + (\theta^{n+1} - (\theta-1)^{n+1})] \\
&= \frac{1}{\theta^n} \theta^{n+1} \\
&= \theta
\end{aligned}$$

$\therefore$ $U$ 는 $\theta$에 대해 unbiased estimator 이고 complete sufficient statistic $(Y)$ 의 함수이므로 $\theta$ 에 대한 unique MVUE 이다.