Homework 5

蒋翌坤 20307100013

Problem 1

(a)

We can write Bernoulli distribution as $f(x|p) = p^x(1-p)^{1-x}$ for $x \in \{0,1\}$. So, log likelihood function is

$$l(p) = \log\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = \log p\sum_{i=1}^{n} x_i + \log(1-p)(n - \sum_{i=1}^{n} x_i)$$
$$\frac{dl}{dp} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n-\sum_{i=1}^{n} x_i}{1-p} = 0 \Rightarrow \hat{p} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{X}$$

So, $\bar{X}$ is the MLE of p. $\mathrm{E}(\bar{X}) = \mathrm{E}(X) = p$ $\mathrm{Var}(\bar{X}) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) = \frac{p(1-p)}{n}$

$f(x|p) = p^x(1-p)^{1-x} = \exp\Big\{x\log p + (1-x)\log(1-p)\Big\}$, so f is from exponential family.

$$I(p) = \mathrm{E}\Big[\Big(\frac{\partial}{\partial p}\log\prod_{i=1}^{n} f(x_i|p)\Big)^2\Big] = -n\mathrm{E}\Big[\frac{\partial^2}{\partial p^2}\log f(x|p)\Big] = -n\mathrm{E}\Big[\frac{\partial}{\partial p}\frac{x}{p} - \frac{1-x}{1-p}\Big]$$
$$= -n\mathrm{E}\Big[-\frac{x}{p^2} - \frac{1-x}{(1-p)^2}\Big] = -n\Big(-\frac{p}{p^2} - \frac{1-p}{(1-p)^2}\Big) = \frac{n}{p(1-p)}$$

So the Cramer-Rao lower bound is $\frac{1}{I(p)} = \frac{p(1-p)}{n}$. So the variance of the MLE of p attains the Cramer-Rao lower bound.

□

(b)

Let $Y = X_1X_2X_3X_4$, $P(Y=1) = p^4, P(Y=0) = 1-p^4$, so $\mathrm{E}(Y) = p^4$, Y is an unbiased estimator of p^4.

We next show $T = \sum_{i=1}^{n} \sim \mathrm{Bin}(n,p)$ is complete sufficient. Sample's joint distribution is $f(\mathbf{x}, p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$, so T is sufficient.

$$0 = \mathrm{E}(g(T)) = \sum_{t=0}^{n} g(t)\binom{n}{t}p^t(1-p)^{n-t} = (1-p)^n\sum_{t=0}^{n} g(t)\binom{n}{t}\Big(\frac{p}{1-p}\Big)^t$$

The right side $\sum_{t=0}^{n} g(t)\binom{n}{t}\Big(\frac{p}{1-p}\Big)^t$ is a n-th polynomial about $\frac{p}{1-p}$. If $g(t) \neq 0$, it has at most n roots. However, $\frac{p}{1-p}$ can be any positive value, so $g(t) = 0$ for all t. So T is complete.

$$\phi(T) = \mathrm{E}(Y|T) = P(X_1=1, X_2=1, X_3=1, X_4=1|T=t) = \frac{P(X_1=1, X_2=1, X_3=1, X_4=1, T=t)}{P(T=t)}$$
$$= \frac{P(X_1=1, X_2=1, X_3=1, X_4=1, \sum_{i=5}^{n} X_i = t-4)}{P(T=t)}$$
$$= \frac{p^4\cdot\binom{n-4}{t-4}p^{t-4}(1-p)^{n-t}}{\binom{n}{t}p^t(1-p)^t} = \frac{\binom{n-4}{t-4}}{\binom{n}{t}} = \frac{t(t-1)(t-2)(t-3)}{n(n-1)(n-2)(n-3)}$$

Since T is complete sufficient, we know $\frac{T(T-1)(T-2)(T-3)}{n(n-1)(n-2)(n-3)}$ is UMVUE of p^4.

Problem 2

(a)

sample's joint distribution is $f(\mathbf{x}, \lambda) = \prod_{i=1}^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \frac{1}{\prod_{i=1}^n x_i!}$. Let $T = \sum_{i=1}^n x_i$, T is sufficient.

Poisson distribution is from exponential family: $f(x, \lambda) = \frac{1}{x!} e^{-\lambda} e^{x \log \lambda}$, so $T = \sum_{i=1}^n x_i$ is complete. $\bar{X} = \frac{1}{n} T$, since T is complete sufficient, $\bar{X}$ is the UMVUE of λ. Therefore, $\bar{X}$ is the best unbiased estimator of λ.

□

(b)

For any $i = 1, \ldots, n$, we have

$$P(X_i = x_i | \bar{X} = \frac{x}{n}) = \frac{P(X_i = x_i, \sum_{j=1}^n X_j = x)}{P(\sum_{j=1}^n X_j = x)} = \frac{P(X_i = x_i, \sum_{j=1, j \neq i}^n X_j = x - x_i)}{P(\sum_{j=1}^n X_j = x)}$$

$$= \frac{\frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \cdot \frac{((n-1)\lambda)^{x - x_i} e^{-(n-1)\lambda}}{(x - x_i)!}}{\frac{(n\lambda)^x e^{-n\lambda}}{x!}} = \binom{x}{x_i} \left(\frac{1}{n}\right)^x \left(1 - \frac{1}{n}\right)^{x - x_i}$$

So, $X_i | \bar{X} \sim \text{Bin}(n\bar{X}, \frac{1}{n})$, variance is $\mathrm{E}((X_i - \bar{X})^2 | \bar{X}) = (1 - \frac{1}{n}) \bar{X}$. Therefore. $\mathrm{E}(\sum_{i=1}^n (X_i - \bar{X})^2 | \bar{X}) = (n-1)\bar{X} \Rightarrow \mathrm{E}(S^2 | \bar{X}) = \bar{X}$.

We know that $\bar{X}$ is sufficient because sample's joint distribution is $f(\mathbf{x}, p) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} = p^{n\bar{X}} (1-p)^{n - n\bar{X}}$. From Rao-Blackwell, we know $\mathrm{Var}(\bar{X}) \leq \mathrm{Var}(S^2)$, equality is achieved when $\mathrm{E}(\mathrm{Var}(S^2 | \bar{X})) = 0$, i.e., $S^2 = \bar{X}$, but $S^2 \neq \bar{X}$, so $\mathrm{Var}(\bar{X}) < \mathrm{Var}(S^2)$.

□

Problem 3

Let $T = \sum_{i=1}^n X_i \sim \text{Bin}(n, p)$, in problem 1 we already show that T is a complete sufficient statistic for p.

We know $\mathrm{E}(T) = np$, $\mathrm{E}(T^2) = np + n(n-1)p^2$, let $\phi(T) = \frac{nT - T^2}{n(n-1)}$ $\mathrm{E}\Big(\phi(T)\Big) = \frac{n^2 p - np - n(n-1)p^2}{n(n-1)} = p(1-p)$. Since T is complete and sufficient, so $\phi(T)$ is the UMVUE of $p(1-p)$.

Problem 4

Let $\mathrm{E}(\phi(T)) = \tau(\theta)$. Suppose there is another estimator based on T, $\phi'(T)$, such that $\phi'(T) = \tau(\theta)$, then $\mathrm{E}(\phi(T) - \phi'(T)) = 0$, since T is complete, so $P(g(T) = \phi(T) - \phi'(T) = 0) = 1 \Rightarrow \phi(T) = \phi'(T)$. So $\phi(T)$ is the unique unbiased estimator of its expected value.

□