# MATH243 Statistical Inference

**Exercise 3** **(Parameter Estimation II: Rao-Blackwell Theorem)**

Cramer-Rao lower bound:

1. Let $\underset{\sim}{X}$ be a random sample of size $n$ from $\text{Bi}(1, \theta)$. Find the maximum likelihood estimator for $\theta$ and verify in this case the Cramer-Rao Inequality. Comment on what you have found.

2. Let $\underset{\sim}{X}$ be a random sample of size $n$ from an exponential distribution with parameter $\dfrac{1}{\theta}$. The p.d.f. is
$$f(x; \theta) = \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right), \quad x \in (0, \infty)$$
Find the Cramer-Rao lower bound and deduce whether or not $\bar{X}$ is a fully efficient estimator for $\theta$.

Exponential family, Sufficient Statistic, Rao-Blackwell Theorem

3. Show that the following p.d.f.'s belong to the exponential family (discrete case)
   (a) Poisson;
   (b) Binomial;
   (c) Negative Binomial.

4. Show that the following p.d.f.'s belong to the exponential family (continuous cases).
   (a) Gamma $(\theta; k)$, where $k > 0$ is known;
   (b) $N(\theta; 1)$;
   (c) $N(0; \theta)$.

5. Let $X_1, X_2, ..., X_n$ be a random sample of size $n$ from a geometric distribution that has p.d.f. $f(x; \theta) = (1-\theta)^x \theta,\ x = 0, 1, 2, ...,\ 0 < \theta < 1$, zero elsewhere. Show that $\sum_1^n X_i$ is a sufficient statistic for $\theta$.

6. Show that the sum of the items of a random sample of size $n$ from a gamma distribution which has p.d.f. $f(x; \theta) = (1/\theta)e^{-x/\theta},\ 0 < x < \infty,\ 0 < \theta < \infty$, zero elsewhere, is a sufficient statistic for $\theta$.

7. Suppose we have a random sample $\underset{\sim}{X}$ of size $n$ from a normal distribution $N(0, \theta),\ \theta \in R^+$. Show that
$$T = t(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i^2$$
is a sufficient statistic for $\theta$.

8. Let $\underset{\sim}{X} = (X_1, ..., X_n)$ be a random sample from a uniform distribution over the interval $[0, \theta]$. Show that $Y = \max(X_1, X_2, ..., X_n)$ is sufficient for $\theta$.

9. Prove that the sum of the items of a random sample of size $n$ from a Poisson distribution having parameter $\theta, 0 < \theta < \infty$, is a sufficient statistic for $\theta$.

10. Consider a random sample of size $n$ drawn from a Bernoulli distribution. Show that the single statistic, $X_k, k \in \{1, 2, ..., n\}$, is an unbiased estimator for $\theta$, the probability of success, but not sufficient for $\theta$. Show that $Y = \sum_{i=1}^{n} X_i$ is sufficient for $\theta$. Find an improved unbiased estimator provided by the Rao-Blackwell Theorem.

11. Let the random variable $X$ denote the time to failure of television tubes and assume that $X$ is exponentially distributed with unknown mean $\theta$. Let $\underset{\sim}{X} = (X_1, X_2)$ be a random sample from this distribution

    (a) Show that $X_1, X_2$ are unbiased estimator for $\theta$ but not sufficient for $\theta$.

    (b) Show that $Y = X_1 + X_2$ is sufficient for $\theta$.

    (c) Hence obtain an unbiased estimator for $\theta$ which has variance which is not greater than that of $X_1$ or $X_2$ for all $\theta (0 < \theta < \infty)$.

12. If $X_1, X_2$ denote a random sample of size 2 from a distribution having p.d.f.

$$f(x; \theta) = (1/\theta) e^{-x/\theta}, \quad 0 < x < \infty, \quad 0 < \theta < \infty.$$

    (a) Find the joint p.d.f. of $Y_1 = X_1 + X_2$, and $Y_2 = X_2$.

    (b) Show that $Y_2$ is an unbiased estimator for $\theta$ having variance $\theta^2$, and apply Rao-Blackwell Theorem to find an improved estimator. Does the variance of the improved estimator attain the Cramer-Rao lower bound?

13. Write the p.d.f.

$$f(x; \theta) = \frac{1}{6\theta^4} x^3 e^{-x/\theta}, \qquad 0 < x < \infty, \ 0 < \theta < \infty,$$

zero elsewhere, in the exponential form. If $X_1, X_2, ..., X_n$ is a random sample from this distribution, find a complete sufficient statistic $Y_1$ for $\theta$ and the unique continuous function $\varphi(Y_1)$ of this statistic which is the best statistic for $\theta$. Is $\varphi(Y_1)$ itself a complete sufficient statistic ?

14. Let $X_1, X_2, ..., X_n$ denote a random sample of size $n > 2$ from a distribution with p.d.f. $f(x; \theta) = \theta e^{-\theta x}$, $0 < x < \infty$, zero elsewhere, and $\theta > 0$. Then $Y = \sum_{1}^{n} X_i$ is a sufficient statistic for $\theta$. Prove that $(n-1)/Y$ is the best statistic for $\theta$.

15. Let $X_1, X_2, ..., X_n$ denote a random sample from a distribution which is $b(1, \theta)$. Find the best statistic for the variance $n\theta(1-\theta)$ of $Y = \Sigma X_i$.

16. Let $X_1, X_2, ..., X_n$ denote a random sample from a distribution which is $n(0, \theta)$. Then $Y = \Sigma X_i^2$ is a sufficient statistic for $\theta$. Find the best statistic for $\theta^2$.

17. Let $X_1, X_2, \ldots, X_n$ denote a random sample of size $n > 2$ from a distribution with p.d.f. $f(x; \theta) = \theta e^{-\theta x}, 0 < x < \infty$, zero elsewhere, and $\theta > 0$. Find the best estimator for $\theta$.

18. Let $(X_1, ..., X_n)$ be a sample from $P_0(\lambda)$, find the UMVUE of $\tau(\lambda) = \dfrac{\lambda^k e^{-\lambda}}{k!}$ for any $k = 0, 1, 2, ...$ .

19. Let $X$ be a r.v. having the Negative Binomial distribution with known $k$ and parameter $\theta \in \Omega = (0, 1)$. Find the UMVU estimator of $g(\theta) = 1/\theta$ and determine its variance.

20. Let $X_1, \ldots X_n$ be i.i.d. r.v.'s from the Gamma distribution with $\alpha$ known and $\theta \in \Omega = (0, \infty)$ unknown, i.e.

$$f(x, \theta) = \frac{x^{\alpha-1} e^{-x/\theta}}{\theta^\alpha \Gamma(\alpha)}.$$

Then show that the UMVU estimator of $\theta$ is

$$U(X_1, \ldots X_n) = \frac{1}{n\alpha} \sum_{j=1}^{n} X_j$$

and its variance attains the Cramer-Rao bound.

21. Let $X$ be a r.v. denoting the life length of an equipment. Then the reliability of the equipment at time $x$, $R(x)$, is defined as the probability that $X > x$. If $X$ has the exponential distribution with parameter $\theta \in \Omega = (0, \infty)$, find the UMVU estimator of the reliability $R(x, \theta)$ on the basis of $n$ observations on $X$.

Difficult questions:

22. 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 p.d.f. $f(x; \theta) = 1/\theta,\ 0 < x < \theta,\ 0 < \theta < \infty$, zero elsewhere. Show that $2Y_3$ is an unbiased statistic for $\theta$. Determine the joint p.d.f. of $Y_3$ and the sufficient statistic $Y_5$ for $\theta$. Find the conditional expectation $E(2Y_3|y_5) = \varphi(y_5)$. Compare the variances of $2Y_3$ and $\varphi(Y_5)$.

23. Let the random variables $X$ and $Y$ have the joint p.d.f. $f(x, y) = (2/\theta^2)e^{-(x+y)/\theta}$, $0 < x < y < \infty$, zero elsewhere.

    (a) Show that the mean and the variance of $Y$ are, respectively, $3\theta/2$ and $5\theta^2/4$.

    (b) Show that $E(Y|x) = x + \theta$. In accordance with the Rao-Blackwell theorem, the expected value of $X + \theta$ is that of $Y$, namely, $3\theta/2$, and the variance of $X + \theta$ is less than that of $Y$. Show that the variance of $X + \theta$ is in fact $\theta^2/4$.

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

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

    (b) Prove that

    $$[Y^{n+1} - (Y-1)^{n+1}]/[Y^n - (Y-1)^n]$$

    is the unique best statistic for $\theta$.

25. Suppose $X_1, \ldots, X_n$ are independent random variable with distribution $B(1, p)$.

    (a) Find the maximum likelihood estimator of $\theta = (1-p)^2$.

    (b) Show that $\hat{\theta}$ is an unbiased estimator of $\theta$, where $\hat{\theta} = 1$ if $X_1 + X_2 = 0$ and $\hat{\theta} = 0$ otherwise.

    (c) Find the best estimator for $\theta$.

26. Let $X_1, \ldots, X_n$ be i.i.d. r.v.'s from the $U(\theta, 2\theta)$, $\theta \in \Omega = (0, \infty)$ distribution and set

    $$U_1 = \frac{n+1}{2n+1} Y_n \quad \text{and} \quad U_2 = \frac{n+1}{5n+4}[2Y_n + Y_1].$$

    Then show that both $U_1$ and $U_2$ are unbiased estimators of $\theta$ and the $U_2$ is better than $U_1$.

27. Let $X_1, \ldots X_n$ be independent r.v.'s distributed as $N(\theta, 1)$. Show that $\bar{X}^2 - (1/n)$ is the UMVU estimator of $g(\theta) = \theta^2$. Also show that the Cramer-Rao bound is not attained.

28. Let $X_1, \ldots X_n$ be independent r.v.'s distributed as $N(\mu, \sigma^2)$, where both $\mu$ and $\sigma^2$ are unknown. Find the UMVU estimator of $\mu/\sigma$.

29. Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be two independent random samples with the same mean $\theta$ and known variances $\sigma_1^2$ and $\sigma_2^2$, respectively. Then show that for every $c \in [0, 1]$, $U = c\bar{X} + (1-c)\bar{Y}$ is an unbiased estimator of $\theta$. Also find the value of $c$ for which the variance of $U$ is minimum.