Autumn Semester 2021 **December 31, 2021**

IE 605: Engineering Statistics

Tutorial 8

Exercise 1. *Let $X_1, X_2, \ldots, X_n$ be independent random variables from pdfs*

$$f(x_i|\theta) = \begin{cases} \frac{1}{2i\theta}, & -i(\theta - 1) < x_i < i(\theta + 1) \\ 0, & \text{otherwise} \end{cases}$$

where $\theta > 0$. Find the sufficient statistics of θ.

Exercise 2. *Let X_1, X_2 be iid $Poisson(\lambda)$ RVs.*

1. *Is $T_1 = X_1 + X_2$ sufficient for λ?*
2. *Is $T_2 = X_1 + 2X_2$ sufficient for λ?*

Exercise 3. *Let $X_1, X_2, ..., X_n$ be a sequence of independent Bernoulli trials with $P(X_i = 1) = \theta$. Prove that $T = \sum_{i=1}^{n} X_i$ is sufficient for θ. Show that T is also complete.*

Exercise 4. *Suppose that $X_1, X_2, ..., X_n$ form a random sample from a normal distribution for which the mean μ is unknown but the variance σ^2 is known. Find a sufficient statistic for μ.*

Exercise 5. *Suppose that $X_1, X_2, ..., X_n$ form a random sample from a beta distribution with parameters α and β, where the value of α is known and the value of β is unknown ($\beta > 0$). Show that the following statistic T is a sufficient statistic for the parameter β:*

$$T = \frac{1}{n}\left(\sum_{i=1}^{n} \log \frac{1}{1 - X_i}\right)^3$$

Exercise 6. *Let $X_1, X_2, \ldots, X_n$ be a random sample drawn from the pdf given as,*

$$f_X(x) = \frac{e^{-(x-\theta)}}{(1 + e^{-(x-\theta)})^2} \qquad -\infty < x < \infty, -\infty < \theta < \infty$$

Find the minimal sufficient statistic of θ.

Exercise 7. *Show that the minimal sufficient statistic of $U(\theta, \theta + 1)$ is not complete.*

Exercise 8. *Let $X_1, X_2, \ldots, X_n$ be a sample from $N(\theta, \theta^2)$ where $\theta > 0$. Show that $T = (\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2)$ is sufficient for θ but T is not complete.*

Exercise 9. *Let $X_1, X_2, \ldots, X_n$ be a sample from $N(\theta, \alpha\theta^2)$, where α is known constant and $\theta > 0$. Show that $T = (\bar{X}, S^2)$ is sufficient statistics for θ but T is not complete.*

Exercise 10. *Suppose X_1 and X_2 are i.i.d. observations from the pdf $f(x|\alpha) = \alpha x^{\alpha-1}e^{-x^{\alpha}}, x > 0, \alpha > 0$. Show that $\frac{\log X_1}{\log X_2}$ is an ancillary statistic.*

Exercise 11. *Samples are drawn from $Unif(\theta, \theta+1)$, where θ is unknown. Joint PDF of sample x is then*

$$f(x|\theta) = \begin{cases} 1 & \textit{if} \quad \theta < x_i < \theta + 1 \quad i = 1, 2, \ldots, n \\ 0 & \textit{otherwise} \end{cases}$$

Find the minimal sufficient statistic for θ. Show that $T(X) = ((X_{(n)} - X_{(1)}), (X_{(n)} + X_{(1)})/2)$ is also a minimum sufficient statistic.

Exercise 12. *In Exercise 11, show that the range statistic, i.e., $R = X_{(n)} - X_{(1)}$ is an ancillary statistic.*

Exercise 13. *A natural ancillary statistic in most problems is the sample size. For example, let N be a random variable taking values $1, 2, \ldots$ with known probabilities $p_1, p_2, \ldots$ where $\sum_i p_i = 1$. Having observed $N = n$, perform n Bernoulli trials with success probability θ, getting X successes.*

(a) Prove that the pair (X, N) is minimal sufficient and N is ancillary for θ.

(b) Prove that the estimator X/N is unbiased for θ and has variance $\theta(1-\theta)\mathbb{E}[1/N]$.

Exercise 14. *Consider the model in Exercise 13 above. Show that the Formal Likelihood Principle implies that any conclusions about θ should not depend on the fact that the sample size n was chosen randomly. That is, the likelihood for (n, x), a sample point from Exercise 13, is proportional to the likelihood for the sample point x, a sample point from a fixed-sample-size binomial (n, θ) experiment.*

Exercise 15. *A risky experiment treatment is to be given to at most three patients. The treatment will be given to one patient. If it is a success, then it will be given to a second. If it is a success, it will be given to a third patient. Model the outcomes for the patients as independent Bernoulli (p) r.v.s. Identify the four sample points in this model and show that, according to the Formal Likelihood Principle, the inference about p should not depend on the fact that the sample size was determined by the data.*

Exercise 16. *Consider a Negative Binomial Distribution with $r = 3$ and probability of success be p. If $X = 3$ (X represents no of success) is observed, then find its likelihood in terms of p. Also generalize the result for $X = x$*

Exercise 17. *Let $X_1, X_2, \ldots, X_n$ be a random sample from the Inverse Gaussian distribution with pdf*

$$f(x|\mu, \lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} e^{\frac{-\lambda(x-\mu)^2}{2\mu^2 x}}, 0 < x < \infty$$

(a) Show that the statistics

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } T = \frac{n}{\sum\limits_{i=1}^{n} \frac{1}{X_i} - \frac{1}{\bar{X}}}$$

are sufficient and complete.

(b) For $n = 2$, show that $\bar{X}$ has an inverse Gaussian distribution, $n\lambda/T$ has a χ^2_{n-1} distribution, and they are independent.

Exercise 18. *Suppose that $(X_1, Y_1), ..., (X_n, Y_n)$ are independent and identically distributed random 2-vectors having the normal distribution with $\mathbb{E}[X_1] = \mathbb{E}[Y_1] = 0, Var(X_1) = Var(Y_1) = 1$, and $Cov(X_1, Y_1) = \theta \in (-1, 1)$.*

1. *Find a minimal sufficient statistic for θ.*
2. *Show whether the minimal sufficient statistic in (i) is complete or not.*
3. *Prove that $T_1 = \sum\limits_{i=1}^{n} X_i^2$ and $T_2 = \sum\limits_{i=1}^{n} Y_i^2$ are both ancillary but (T_1, T_2) is not ancillary.*

Exercise 19. *One advantage of using a minimal sufficient statistic is that unbiased estimators will have smaller variance, as the following exercise will show. Suppose that T_1 is sufficient and T_2 is minimal sufficient, U is an unbiased estimator of θ, and define $U_1 = \mathbb{E}[U|T_1]$ and $U_2 = \mathbb{E}[U|T_2]$.*

(a) Show that $U_2 = \mathbb{E}[U_1|T_2]$.

(b) Now use conditional variance formula to show that $Var(U_2) \leq Var(U_1)$.

Exercise 20. *Let $(X_1, ..., X_n)$ be a random sample of random variables having the Cauchy distribution with location parameter μ and scale parameter σ, where $\mu \in \mathbb{R}$ and $\sigma > 0$ are unknown parameters. Show that the vector of order statistics is minimal sufficient for (μ, σ).*