Department of Statistics, The Chinese University of Hong Kong
STAT5010 Advanced Statistical Inference (Term 1, 2022–23)

Assignment 2 · due on 24 October 2021

Please submit your answers in `.pdf` *format via Blackboard.*

1. Let $\{X_i\}_{i=1,\ldots,n}$ be a random sample i.i.d. from F,
   (a) If F is $N(\mu, \sigma^2)$ with μ and σ^2 unknown, find a sufficient statistic for (μ, σ^2).
   (b) If F is Uniform$[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$, find the sufficient statistic for θ.

2. Let $\{X_i\}_{i=1,\ldots,n}$ be a random sample i.i.d. from F, where $f = F'$ is continuous. For $\tau \in (0,1)$, denote ξ_τ as the τ-th quantile of the distribution (i.e., $F(\xi_\tau) = \tau$), and $f(\xi_\tau) > 0$, then show that $X_{(k)} \xrightarrow{P} \xi_\tau$ where $X_{(k)}$ is the k-th order statistic of the sample and $k = [n\tau]$.

3. Let X be one observation from a $N\left(0, \sigma^2\right)$ population. Is $|X|$ a sufficient statistic?

4. Let $X_1, \cdots, X_n$ be a random sample from the pdf

$$f\left(x \mid \mu, \sigma\right) = \frac{1}{\sigma} e^{-(x-\mu)/\sigma}, \mu < x < \infty, 0 < \sigma < \infty.$$

   Find a two-dimensional sufficient statistic for (μ, σ).

5. Let $X_1, \cdots, X_n$ be iid observations from a pdf or pmf $f\left(x \mid \theta\right)$ that belongs to an exponential family given by

$$f\left(x \mid \boldsymbol{\theta}\right) = h\left(x\right) c\left(\boldsymbol{\theta}\right) \exp\left(\sum_{i=1}^{k} w_i\left(\boldsymbol{\theta}\right) t_i\left(x\right)\right),$$

   where $\boldsymbol{\theta} = \left(\theta_1, \cdots, \theta_d\right), d \leq k$. Prove that $T\left(\boldsymbol{X}\right) = \left(\sum_{j=1}^{n} t_1\left(\boldsymbol{X}_j\right), \cdots, \sum_{j=1}^{n} t_k\left(\boldsymbol{X}_j\right)\right)^\top$ is a sufficient statistic for $\boldsymbol{\theta}$.

6. Let $X_1, \cdots, X_n$ be independent random variables with pdfs

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

   where $\theta > 0$. Find a two-dimensional sufficient statistic for θ.

7. Let $f(x, y \mid \theta_1, \theta_2, \theta_3, \theta_4)$ be a bivariate pdf for the uniform distribution on the rectangle with lower left corner (θ_1, θ_2) and upper right corner (θ_3, θ_4) in $\mathbb{R}^2$. The parameters satisfy $\theta_1 < \theta_3$ and $\theta_2 < \theta_4$. Let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be a random sample from this pdf. Find a four-dimensional sufficient statistic for $(\theta_1, \theta_2, \theta_3, \theta_4)$.

8. Consider an exponential family whose density is given by

$$p(x \mid \eta) = \exp\left\{\sum_{i=1}^{s} \eta_i T_i(x) - A(\eta)\right\} h(x)$$

with nature parameter space Θ. Show that Θ is convex. *(Hint: See Lemma 2.7.1 of Lehmann and Romano (2005).)*

9. [Rao-Blackwell Theorem]. Let X be a random observable with distribution $P_\theta \in \mathcal{P} = \{P_{\theta'} : \theta' \in \Theta\}$, and let T be sufficient for $\mathcal{P}$. Let δ be an estimator of an estimand $g(\theta)$, and let the loss function $L(\theta, d)$ be a strictly convex function of d. If δ has finite expectation and risk,

$$R(\theta, \delta) = E\left(L(\theta, \delta(X))\right) < \infty\ ,$$

and if

$$\eta(t) = E\left(\delta(X) \mid t\right)\ ,$$

then the risk of the estimator $\eta(T)$ satisfies

$$R(\theta, \eta) < R(\theta, \delta)$$

unless $\delta(X) = \eta(T)$ with probability 1.

10. Let $X_1, \ldots, X_n$ be i.i.d according to the exponential distribution $E(a, b)$, i.e., X_i has density

$$f_X(x) = \frac{1}{b} e^{-(x-a)/b} \cdot I(x \geq a)\ , \quad a \in \mathbb{R}\ , b > 0\ .$$

Now let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ be the corresponding order statistic of the sample and let $T_1 = X_{(1)}$, $T_2 = \sum_{i=1}^{n}\{X_i - X_{(1)}\}$. Show that (T_1, T_2) are independently distributed as $E(a, b/n)$ and $\frac{1}{2} b \chi^2_{2n-2}$ respectively, and there are jointly sufficient and complete.

11. Let $X_1, X_2, \ldots, X_n$ be i.i.d according to the logistic distribution $L(\theta, 1)$, i.e., X_i has density

$$f_X(x) = \frac{e^{-(x-\theta)}}{\{1 + e^{-(x-\theta)}\}^2}\ , \quad \theta \in \mathbb{R}\ . \tag{1}$$

Consider a subfamily $\mathcal{P}_0$ consisting of the distribution (1) with $\theta_0 = 0$ and $\theta_1, \ldots, \theta_{n+1}$. Show that the order statistic $T(X) = (X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ is minimal sufficient for $\mathcal{P}_0$.

12. [Problem 3.27 of Keener (2010)]. Let $X_1, \ldots, X_n$ be i.i.d. from a uniform distribution on $(-\theta, \theta)$, where $\theta > 0$ is an unknown parameter.

    (a) Find a minimal sufficient statistic T.

    (b) Define

    $$V = \frac{\bar{X}_n}{\max_{1 \le i \le n} X_i - \min_{1 \le i \le n} X_i},$$

    where $\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i$ denotes the sample average. Show that T and V are independent.