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

Assignment 2 · due on 17 October 2023
Please submit your answers in `.pdf` *format via Blackboard.*

1. Let $\{X_i\}_{i=1,\dots,n}$ 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 $\{X_i\}_{i=1,\dots,n}$ 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_1 and X_2 be independent discrete random varibles with common mass function
$$\Pr(X_i = x) = -\frac{\theta^x}{x \log(1-\theta)}, \quad x = 1, 2, \dots,$$
where $\theta \in (0,1)$.
   (a) Find the mean and variance of X_1.
   (b) Find the UMVU of $\theta / \log(1-\theta)$.

4. Let $\theta = (\alpha, \lambda)$ and let P_θ denote the gamma distribution with shape parameter α and scale $1/\lambda$. So P_θ has density
$$p_\theta(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-x\lambda}}{\Gamma(\alpha)} I(x > 0).$$
   (a) Find the Fisher information matrix $I(\theta)$, expressed using $\psi \equiv \Gamma'/\Gamma$ and its derivatives.
   (b) What is the Cramér-Rao lower bound for the variance of an unbiased estimator of $\alpha + \lambda$?
   (c) Find the mean μ and variance σ^2 for P_θ. Show that there is a one-to-one correspondence between θ and (μ, σ^2).
   (d) Find the Fisher information matrix if the family of gamma distributions is parameterised by (μ, σ^2) instead of θ.

5. Let $X_1, \cdots, X_n$ be iid observations from a pdf or pmf $f(x \mid \theta)$ that belongs to an exponential family given by
$$f(x \mid \boldsymbol{\theta}) = h(x)\, c(\boldsymbol{\theta}) \exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})\, t_i(x)\right),$$

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

6. 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).)*

7. [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.

8. 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.

9. 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

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

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 $\mathcal{P}_0$.

10. 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.