MTH 418a: Inference-I
Assignment No. 3: Minimal Sufficiency and Completeness

1. Show that every function of a complete statistic is complete.

2. Let $X_1, \ldots, X_n$ be a random sample from a population with p.d.f./p.m.f. $f_\theta(\cdot), \theta \in \Theta$. In each of the following cases, show that $T(\underline{X})$ is a minimal sufficient statistic. Also verify if it is complete.

   (a) For known $\mu_0 \in \mathbb{R}, X_1 \sim \mathrm{N}(\mu_0, \theta^2), \Theta = (0, \infty)$, $T(\underline{X}) = \sum_{i=1}^n (X_i - \mu_0)^2$. Is $\underline{U}(\underline{X}) = (\overline{X},\ \sum_{i=1}^n (X_i - \overline{X})^2)$ sufficient and complete?;
   (b) For known $\sigma_0 > 0$, $X_1 \sim \mathrm{N}(\theta, \sigma_0^2), \Theta = \mathbb{R}$, $T(\underline{X}) = \sum_{i=1}^n X_i$;
   (c) $X_1 \sim \mathrm{N}(\mu, \sigma^2), \underline{\theta} = (\mu, \sigma) \in \mathbb{R} \times (0, \infty)$, $\underline{T}(\underline{X}) = (\overline{X},\ \sum_{i=1}^n (X_i - \overline{X})^2)$. Is $U(\underline{X}) = \overline{X} + S^2$ sufficient and complete?;
   (d) For known $\mu_0 \in \mathbb{R}, X_1 \sim \mathrm{Exp}(\mu_0, \theta), \Theta = (0, \infty)$, $T(\underline{X}) = \sum_{i=1}^n (X_i - \mu_0)$. Is $U(\underline{X}) = \overline{X}$ complete and sufficient?;
   (e) For known $\sigma_0 > 0$, $X_1 \sim \mathrm{Exp}(\theta, \sigma_0), \Theta = \mathbb{R}$, $T(\underline{X}) = X_{(1)}$. Is $\underline{U}(\underline{X}) = (X_{(1)},\ \sum_{i=1}^n (X_i - X_{(1)}))$ sufficient and complete?;
   (f) $X_1 \sim \mathrm{Exp}(\mu, \sigma), \underline{\theta} = (\mu, \sigma) \in \mathbb{R} \times (0, \infty)$, $\underline{T}(\underline{X}) = (X_{(1)},\ \sum_{i=1}^n (X_i - X_{(1)}))$;
   (g) For known $\alpha_0 > 0, X_1 \sim \mathrm{Gamma}(\alpha_0, \theta), \Theta = (0, \infty)$, $T(\underline{X}) = \sum_{i=1}^n X_i$.

3. Let $X_1, \ldots, X_n$ be a random sample from a population with p.d.f./p.m.f. $f_\theta(\cdot), \theta \in \Theta$. In each of the following cases, show that $T(\underline{X})$ is a minimal sufficient statistic. Also verify if it is complete.

   (a) For known $\mu_0 \in \mathbb{R}, X_1 \sim \mathrm{N}(\mu_0, \theta^2), \Theta = (0, \infty)$, $T(\underline{X}) = \sum_{i=1}^n (X_i - \mu_0)^2$;
   (b) $X_1 \sim \mathrm{Beta}(\theta_1, \theta_2), \underline{\theta} = (\theta_1, \theta_2), \Theta = (0, \infty) \times (0, \infty)$, $T(\underline{X}) = (\prod_{i=1}^n X_i,\ \prod_{i=1}^n (1 - X_i)$);
   (c) $X_1 \sim \mathrm{N}(\theta, \theta), \Theta = (0, \infty)$, $T(\underline{X}) = \sum_{i=1}^n X_i^2$;
   (d) $X_1 \sim \mathrm{N}(\theta, \theta^2), \Theta = (0, \infty)$, $T(\underline{X}) = (\overline{X}, \sum_{i=1}^n (X_i - \overline{X})^2)$;
   (e) $X_1 \sim \mathrm{U}(\theta_1, \theta_2), \underline{\theta} = (\theta_1, \theta_2), \Theta = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 < x_2\}$, $T(\underline{X}) = (X_{(1)}, X_{(n)})$;
   (f) $X_1 \sim \mathrm{U}(\theta - \frac{1}{2}, \theta + \frac{1}{2}), \Theta = \mathbb{R}$, $T(\underline{X}) = (X_{(1)}, X_{(n)})$;
   (g) $f_\theta(x) = \frac{2(\theta - x)}{\theta^2} I_{(0,\theta)}(x), \Theta = (0, \infty)$, $T(\underline{X}) = (X_{(1)}, \ldots, X_{(n)})$;
   (h) $X_1 \sim \mathrm{Poisson}(\theta), \Theta = (0, \infty)$, $T(\underline{X}) = \overline{X}$;
   (i) For known positive integer m, $X_1 \sim \mathrm{Bin}(m, \theta), \Theta = (0, 1)$, $T(\underline{X}) = \sum_{i=1}^n X_i$;
   (j) For normalizing constant $c(\theta)$, $f_\theta(x) = \frac{c(\theta)}{x^2}$, $x = \theta + 1,\ \theta + 2, \ldots,\ \Theta = \{0, 1, 2, \ldots\}$, $T(\underline{X}) = X_{(1)}$.

4. Each of the parts (a)-(j) below correspond to the corresponding parts of Problem 3. For each of the parts (a)-(j), determine whether the statistic $U(\underline{X})$, as defined below, is minimal sufficient for θ. Also verify if it is complete.

   (a) $U(\underline{X}) = (\overline{X},\ \sum_{i=1}^n (X_i - \overline{X})^2)$;
   (b) $U(\underline{X}) = (X_{(1)}, \ldots, X_{(n)})$;
   (c) $U(\underline{X}) = 15T^2(\underline{X}) + 8T(\underline{X}) - 23$;
   (d) $U(\underline{X}) = (\overline{X},\ \sum_{i=1}^n (X_i - 1)^2)$;

(e) $U(\underline{X}) = (2X_{(n)} - X_{(1)},\ X_{(n)} + 3X_{(1)})$;
(f) $U(\underline{X}) = (X_{(1)}, X_{(2)}, \ldots, X_{(n)})$;
(g) $U(\underline{X}) = (\prod_{i=1}^{n} X_i, X_{(2)} + X_{(3)},\ 2X_{2)} + 3X_{(3)},\ X_{(4)}, \ldots, X_{(n)})$;
(h) $U(\underline{X}) = 2\overline{X}^2 + 3\overline{X} - 100$;
(i) $U(\underline{X}) = \overline{X}^5$;
(j) $U(\underline{X}) = 9e^{2X_{(1)}} - 17e^{X_{(1)}} - 1083$.

5. Let X_1, X_2 be a random sample from a discrete uniform distribution on the set $\{\theta, \theta+1, \theta+2\}$, where $\theta \in \Theta = \{0, 1, 2, \ldots\}$. Show that $\underline{T} = (\frac{X_{(1)}+X_{(2)}}{2}, X_{(2)} - X_{(1)})$ is a minimal sufficient statistic. Is it complete?

6. Let X be a r.v. with distributional support $\chi = \{-1, 0, 1, 2, \ldots\}$ and p.m.f. $f_\theta, \theta \in (0,1) = \Theta$, where $f_\theta(-1) = \theta$ and $f_\theta(x) = (1-\theta)^2\theta^x,\ x = 0, 1, 2, \ldots$. Show that the family $\mathcal{P} = \{f_\theta : \theta \in \Theta\}$ is boundedly complete but not complete.

7. Let $X_1, X_2, \ldots, X_n$ be a random sample from a discrete uniform distribution on the set $\{1, 2, \ldots, \theta\}$, where $\theta \in \{1, 2, \ldots\} = \Theta$.

   (a) Show that that the family $\mathcal{P} = \{f_\theta : \theta \in \Theta\}$ is complete;
   (b) Let $m_0 \in \Theta$. Show that the family $\mathcal{P}_1 = \{f_\theta : \theta \in \Theta - \{m_0\}\}$ is not complete;
   (c) Show that $T(\underline{X}) = X_{(n)}$ is complete.

Honors Problems

8. Let $X_1, \ldots, X_n$ $(n \geq 2)$ be a random sample from $\text{Exp}(\mu, \sigma)$ distribution, where $\mu \in \mathbb{R}, \sigma > 0$. Show that $T_1(\underline{X}) = X_{(1)} \sim \text{Exp}(\mu, \frac{\sigma}{n})$ and $T_2(\underline{X}) = \sum_{i=1}^{n}(X_i - X_{(1)}) \sim \text{Gamma}(n-1, \sigma)$. Also, show that $T_1(\underline{X})$ and $T_2(\underline{X})$ are independently distributed. (**Hint:** Note that $T_2(\underline{X}) = \sum_{i=2}^{n}(X_{(i)} - X_{(1)})$)

9. Let $\mathcal{P}$ be a family of densities with common support, and let $\mathcal{P}_0 \subseteq \mathcal{P}$. If a statistic T is minimal sufficient for $\mathcal{P}_0$ and sufficient for $\mathcal{P}$, show that T is minimal sufficient for $\mathcal{P}$.

10. (a) Let $\mathcal{P} = \{f_0, f_1, f_2, \ldots, f_k\}$ be a family of densities, each having the same support. Prove that the statistic $\underline{T}(X) = (\frac{f_1(X)}{f_0(X)}, \frac{f_2(X)}{f_0(X)}, \ldots, \frac{f_k(X)}{f_0(X)})$ is minimal sufficient;
    (b) Let $\mathcal{P} = \{f_0, f_1, f_2\}$, where $f_0(x) = I_{(-1,0)}(x),\ f_1(x) = I_{(0,1)}(x)$ and $f_2(x) = 2xI_{(0,1)}(x)$. Show that $\underline{T}(X) = (\frac{f_1(X)}{f_0(X)}, \frac{f_2(X)}{f_0(X)})$ is not minimal sufficient. Is it sufficient? Find a minimal sufficient statistic.

11. (a) Let X be a random sample from a population having p.m.f. $f_\theta,\ \theta \in (0,1) = \Theta$, where $f_\theta(x) = \frac{1}{4}$, if $x = 1, 2$; $= \frac{1+\theta}{4}$, if $x = 3$; $= \frac{1-\theta}{4}$, if $x = 4$; $= 0$, otherwise; $\theta \in \Theta$. Find a minimal sufficient statistic for θ. Is it complete?
    (b) Let X be a discrete r.v. with distributional support $\chi = \{-1, 0, 1\}$ and p.m.f. $f_\theta,\ \theta \in \Theta = \{1, 2, 3\}$. Suppose that $f_1(-1) = 0.4, f_1(0) = 0.2, f_1(1) = 0.4, f_2(-1) = 0.6, f_2(0) = 0.3, f_2(1) = 0.1, f_3(-1) = 0.2, f_3(0) = 0.1, f_3(1) = 0.7$. Find a minimal sufficient statistic for θ. Is it complete?

12. Let $X_1, \ldots, X_n$ be a random sample from a population with p.d.f. $f_\theta(\cdot), \theta \in \Theta$. In each of the following cases, find a minimal sufficient statistic $T(\underline{X})$. Verify if $T(\underline{X})$ is complete.

    (a) $f_\theta(x) = \frac{1}{\pi} \cdot \frac{1}{1+(x-\theta)^2},\ -\infty < x < \infty,\ \theta \in \mathbb{R} = \Theta;$
    (b) $f_\theta(x) = \frac{1}{2} \cdot e^{-|x-\theta|},\ -\infty < x < \infty,\ \theta \in \mathbb{R} = \Theta.$

13. (a) Let $\mathcal{P}_0$ and $\mathcal{P}_1$ be two families of distributions such that every null set of $\mathcal{P}_0$ is also a null set of $\mathcal{P}_1$, and $\mathcal{P}_0 \subseteq \mathcal{P}_1$. Show that a sufficient statistic that is complete for $\mathcal{P}_0$ is also complete for $\mathcal{P}_1$.
    (b) Let $X_1, \ldots, X_n$ be a random sample from from a population having the p.d.f. $f \in \mathcal{P}$, where $\mathcal{P}$ is the family of all the Lebesgue p.d.f.s. Show that $\underline{T}(\underline{X}) = (X_{(1)}, \ldots, X_{(n)})$ is complete. (**Hint:** In (a), take $\mathcal{P}_0$ to be the exponential family with p.d.f. $f_{\underline{\theta}} = c(\underline{\theta})e^{-\sum_{i=1}^n \theta_i \sum_{j=1}^n x_j^i - \sum_{i=1}^n x_i^{2n}}, \underline{x} \in \mathbb{R}^n, \underline{\theta} \in \mathbb{R}^n$).