MTH 418a: Inference-I
Assignment No. 2: Sufficiency

1. 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, using the factorization theorem, show that $T(\underline{X})$ is a sufficient statistic:

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

2. Each of the parts (a)-(k) below correspond to the corresponding part of Problem 1. For each of the parts (a)-(k), show that the statistic $U(\underline{X})$, as defined below, is sufficient for θ:

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

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, using the conditional distribution of the sample $X_1, \ldots, X_n$ given the statistic $T(\underline{X})$, show that the statistic $T(\underline{X})$ is sufficient for θ.

   (a) $X_1 \sim \mathrm{Poisson}(\theta), \Theta = (0, \infty)$, $T(\underline{X}) = \sum_{i=1}^n X_i$;
   (b) For known positive integer m, $X_1 \sim \mathrm{Bin}(m, \theta), \Theta = (0, 1)$, $T(\underline{X}) = \sum_{i=1}^n X_i$;
   (c) $n = 2$, $X_1 \sim \mathrm{U}(0, \theta), \Theta = (0, \infty)$, $T(\underline{X}) = \max\{X_1, X_2\}$;
   (d) $n = 2$, $X_1 \sim \mathrm{N}(0, \theta), \Theta = (0, \infty)$, $T(\underline{X}) = X_1^2 + X_2^2$.

4. Let $X_1, \ldots, X_n$ be a random sample from a population having p.d.f. $f_\theta(x) = \theta x^{\theta-1}, 0 < x < 1,\ \theta \in \Theta = (0, \infty)$.

(a) Show that $T(\underline{X}) = -\sum_{i=1}^{n} \ln X_i$ is a sufficient statistic for θ;
(b) Find the p.d.f. and $T(\underline{X})$;
(c) Are $T_1(\underline{X}) = (X_1, \ldots, X_n)$, $T_2(\underline{X}) = (X_{(1)}, \ldots, X_{(n)})$, $T_3(\underline{X}) = (X_1, X_1 + X_2, \prod_{i=3}^{n} X_i)$, $T_4(\underline{X}) = T^2(\underline{X}) + 3T(\underline{X}) + 2$, $T_5(\underline{X}) = e^{T(\underline{X})} - T(\underline{X}) + \frac{1}{2}$ sufficient statistics for θ?.