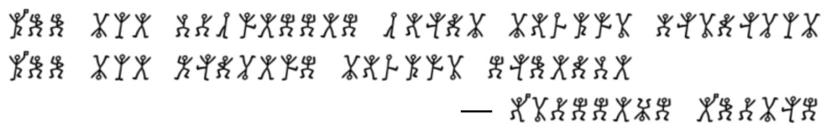
STA3020 Assignment 2.

[1] *School of Data Science, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen)*



1. Sufficient

•*Exercise* 1.1. For two independent samples $X = (X_1, \cdots, X_m)$ and $Y = (Y_1, \cdots, Y_n)$, we have $T = T(X)$ is a sufficient statistic for θ_1 and $S = S(Y)$ is a sufficient statistic for θ_2. Please show that (T, S) is a sufficient statistic for $\theta = (\theta_1, \theta_2)$.

2. Minimal Sufficient

•*Exercise* 2.1. Let $X_1, \cdots, X_n$ be a random sample from $Gamma(\alpha, 1/(\alpha^2+\alpha))$,

$$f(x|\alpha) = \frac{(\alpha^2+\alpha)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\big(-(\alpha^2+\alpha)x\big) \cdot \mathbb{1}(x > 0)$$

with $\Theta = \{\alpha : \alpha > 0\}$. Please show that $T = (\sum_{i=1}^n \ln X_i, \sum_{i=1}^n X_i)$ is minimal sufficient for α.

•*Exercise* 2.2. Let $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ be indenpendently distributed according to $N(\xi, \sigma^2)$ and $N(\eta, \tau^2)$, respectively. Find the minimal sufficient statistics for these cases:

(a) ξ, η, σ, τ are all arbitrary: $-\infty < \xi, \eta < \infty$, $\sigma, \tau > 0$.
(b) $\sigma = \tau$ and ξ, η, σ are all arbitrary: $-\infty < \xi, \eta < \infty$, $\sigma = \tau > 0$.
(c) $\xi = \eta$ and ξ, σ, τ are all arbitrary: $-\infty < \xi = \eta < \infty$, $\sigma, \tau > 0$.

•*Exercise* 2.3. Let $X_1, \cdots, X_n$ be i.i.d random variables from $P_\theta \in \{P_\theta : \theta \in \Theta\}$. In the following cases, find a minimal sufficient statistic for $\theta \in \Theta$.

(a) P_θ is the Poisson(λ) distribution, $\theta = \lambda \in (0, +\infty)$.
(b) P_θ is the binomial(n, p) distribution with a known n and $\theta = p \in (0, 1)$.

(c) P_θ is the Exponential(λ) distribution, $\theta = \lambda \in (0, +\infty)$.
(d) P_θ is the Inverse-Gamma(α, β) distribution, i.e., the density of X_1 is

$$f_\theta(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left(-\frac{\beta}{x}\right) \cdot \mathbb{1}(x > 0)$$

with $\theta = (\alpha, \beta) \in (0, +\infty) \times (0, +\infty)$.
(e) P_θ is the Beta(α, β) distribution, $\theta = (\alpha, \beta) \in (0, 1) \times (0, 1)$.
(f) P_θ is the log-normal(μ, σ^2) distribution, $\theta = (\mu, \sigma^2) \in \mathbb{R} \times (0, +\infty)$.
(g) P_θ is the Weibull(γ, β) distribution, $\theta = (\gamma, \beta) \in (0, +\infty) \times (0, +\infty)$.

3. Complete Statistics

•*Exercise* 3.1. Suppose $X = (X_1, \cdots, X_n)$ is a random sample came from discrete Uniform, i.e., $Uniform\{1, 2, \cdots, \theta\}$, where $\theta \in \mathcal{N}^+$. Please show $T(X) = \max_{\{1 \le i \le n\}} X_i$ is a complete statistic.

Hint. Using mathematical induction. □

4. Basu Theorem

•*Exercise* 4.1. Let $X_1, \cdots, X_n$ be a random sample from the Exponential(a, b), each with density

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

Let $T_1 = X_{(1)}$ and $T_2 = \sum_{i=1}^n \left[X_i - X_{(1)}\right]$. Then (T_1, T_2) are independently distributed as Exponential($a, b/n$) and $\frac{b}{2}\chi^2_{2n-2}$, respectively.

Hint. Using the Basu's Theorem and the characteristic function. □

5. Hierarchical Structure

•*Exercise* 5.1. The time of failure of a super computing machine, recorded as X, has distribution Gamma(α, α^{-1}), i.e.,

$$f_X(x|\alpha) = \frac{\alpha^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\alpha x} \cdot \mathbb{1}(x > 0)$$

with $\alpha \in \mathbb{N}^+$. Meanwhile, within the time of running of this super computing machine, the number of tasks it can solve, denoted as Y, has the conditional Poisson distribution, $Y|X \sim$Poisson(μX), with some $\mu > 0$. Now, in order to test the performance of this machine, an experiment is conducted independently for a total of n products of this machine and Y_i for the $i-$th experiments are recorded, $i = 1, \cdots, n$. Please obtain a statistic that is minimal sufficient and complete for μ and derive its distribution.