Statistical Inference

B. Statistical Data Science 2nd Year   Indian Statistical Institute

Teacher: Soham Sarkar

Exercise Series 3

Exercise 1. Recall that a sufficient statistic $T = T(\mathbf{X})$ for θ is said to be *minimal sufficient* for θ if, for any other sufficient statistic $S = S(\mathbf{X})$ of θ, we can write T as a function of S, i.e., there exists some function g such that $T(\mathbf{x}) = g(S(\mathbf{x}))$ for all $\mathbf{x} \in \mathcal{X}$. We have also seen a condition which ensures that a statistic T is *minimal sufficient*.

Consider the previous exercise set (Series 2). Verify whether the sufficient statistics that you obtained in that exercise set are minimal sufficient. If not, then try to identify a minimal sufficient statistic.

Exercise 2. Find minimal sufficient statistics for the unknown parameter(s) for a random sample $X_1, \ldots, X_n$ from the following distributions.

(a) $\mathsf{Beta}(\alpha, \beta)$, $\theta = (\alpha, \beta)$.

(b) $\Gamma(\alpha, \lambda)$, $\theta = (\alpha, \lambda)$.

(c) $\mathsf{Pareto}(\mu, \alpha)$, $\mu > 0, \alpha > 0$, $\theta = (\mu, \alpha)$.

Hint: $\mathsf{Pareto}(\mu, \alpha)$ *is a continuous distribution with pdf*

$$f_{\mu,\alpha}(x) = \begin{cases} \dfrac{\alpha\mu^\alpha}{x^{\alpha+1}} & \text{if } x \geq \mu, \\ 0 & \text{otherwise.} \end{cases}$$

This is used to model data which are heavy-tailed.

(d) $\mathsf{Weibull}(\lambda, k)$, $\lambda > 0, k > 0$, $\theta = (\lambda, k)$.

Hint: $\mathsf{Weibull}(\lambda, k)$ *is a continuous distribution with pdf*

$$f_{\lambda,k}(x) = \begin{cases} \dfrac{\lambda}{k}\left(\dfrac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Exercise 3. Let $(X_1, Y_1), \ldots, (X_n, Y_n) \overset{\text{i.i.d.}}{\sim} \mathsf{BVN}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ be a random sample from the bivariate normal distribution. Here, $\mu_x, \mu_y \in \mathbb{R}, \sigma_x, \sigma_y > 0$ and $\rho \in (-1, 1)$ are the unknown parameters (i.e., $\boldsymbol{\theta} = (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$. Find minimal sufficient statistics for the unknown parameters under the following setups.

(a) All the parameters are unknown.

(b) $\sigma_x = \sigma_y = 1$.

(c) $\mu_x = \mu_y = 0$.

(d) $\rho = 0$.

(e) $\mu_x = \mu_y = \mu$, $\mu \in \mathbb{R}$ unknown.

(f) $\mu_x = \mu_y = \mu$, $\sigma_x^2 = \sigma_y^2 = \sigma^2$, $\mu \in \mathbb{R}, \sigma > 0$ unknown.

(g) $\mu_x = \mu_y = \theta$, $\sigma_x = \sigma_y = \theta^2$, $\theta \in \mathbb{R}$ unknown.

(h) $\mu_x = \mu_y = \theta$, $\sigma_x^2 = \sigma_y^2 = \theta$, $\theta > 0$ unknown.