Mathematical Statistics I
Recitation Session 2

Definition. Let $X_1, X_2, \ldots, X_n$ denote a sample of size n from a distribution that has pdf or pmf $f(x;\theta),\ \theta \in \Omega$. Let $Y = u\left(X_1, X_2, \ldots, X_n\right)$ be a statistic whose pdf or pmf is $f_Y\left(y;\theta\right)$. Then Y is a sufficient statistic for θ if and only if

$$\frac{f\left(x_1, x_2, \ldots, x_n;\theta\right)}{f_Y\left[u\left(x_1, x_2, \ldots, x_n\right);\theta\right)]} = H\left(x_1, x_2, \ldots, x_n\right)$$

does not depend upon $\theta \in \Omega$, where $f\left(x_1, x_2, \ldots, x_n;\theta\right)$ is the joint pdf or pmf of $X_1, X_2, \ldots, X_n$.

Exercise. Let $X_1, X_2, \ldots, X_n$ be a random sample from Poisson distribution $P(\lambda), \lambda > 0$. Show that $\sum_{i=1}^n X_i$ is a sufficient statistic for λ.

Exercise. Let $X_1, X_2, \ldots, X_n$ be a random sample from normal distribution $N(0,\theta), \theta > 0$. Show that $\sum_{i=1}^n X_i^2$ is a sufficient statistic for θ.

Theorem. *(Neyman). Let $X_1, X_2, \ldots, X_n$ denote a sample from a distribution that has pdf or pmf $f(x;\theta),\ \theta \in \Omega$. The statistic $Y = u\left(X_1, \ldots, X_n\right)$ is a sufficient statistic for θ if and only if we can find two nonnegative functions, k_1 and k_2, such that*

$$f\left(x_1, x_2, \ldots, x_n;\theta\right) = k_1\left[u\left(x_1, x_2, \ldots, x_n\right);\theta\right] k_2\left(x_1, x_2, \ldots, x_n\right),$$

where $k_2\left(x_1, x_2, \ldots, x_n\right)$ does not depend upon θ.

Exercise. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a geometric distribution that has pmf $f(x;\theta) = (1-\theta)^x\theta, x = 0, 1, 2, \ldots, 0 < \theta < 1$, zero elsewhere. Show that $\sum_{i=1}^n X_i$ is a sufficient statistic for θ.

Exercise. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution with pdf

$$f(x;\theta) = \begin{cases} \theta x^{\theta-1} & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

where $0 < \theta$. Show that $\prod_{i=1}^n X_i$ is a sufficient statistic for θ.

Theorem. *Let $X_1, X_2, \ldots, X_n$ denote a sample from a distribution that has pdf or pmf $f(x;\theta),\ \theta \in \Omega$. If a sufficient statistic $Y = u\left(X_1, X_2, \ldots, X_n\right)$ for θ exists and if a maximum likelihood estimator $\hat{\theta}$ of θ also exists uniquely, then $\hat{\theta}$ is a function of $Y = u\left(X_1, X_2, \ldots, X_n\right)$.*

Exercise. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution with pdf

$$f(x;\theta) = \begin{cases} \theta e^{-\theta x} & 0 < x < \infty,\ \theta > 0 \\ 0 & \text{elsewhere} \end{cases}$$

(a) Show that $Y = \sum_{i=1}^n X_i$ is a sufficient statistic for θ.
(b) Show that the maximum likelihood estimator $\hat{\theta}$, is a function of Y.