# STATISTICS

$X$ a population characteristic, $X_1, X_2, ..., X_n$ a sample of size $n$, i.e. independent and identically distributed, with the same pdf as $X$; $\theta$ target parameter, $\overline{\theta} = \overline{\theta}(X_1, X_2, ..., X_n)$ point estimator for $\theta$.

**Sample Functions**:

**Sample Mean**: $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$,

**Sample Moment**: $\overline{\nu_k} = \frac{1}{n}\sum_{i=1}^{n} X_i^k$,

**Sample Absolute Moment**: $\overline{\mu_k} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k$,

**Sample Variance**: $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$.

**Sample Distribution Function**: $\overline{F}(x) = \frac{\operatorname{card}\{X_i | X_i \leq x\}}{n}$.

**Likelihood Function of a Sample**: $L(X_1, ..., X_n; \theta) = \prod_{i=1}^{n} f(X_i; \theta)$.

**Fisher Information**: $I_n(\theta) = E\left[\left(\frac{\partial \ln L(X_1, ..., X_n; \theta)}{\partial \theta}\right)^2\right]$;

- if the range of $X$ does not depend on $\theta$, then $I_n(\theta) = -E\left[\frac{\partial^2 \ln L(X_1, ..., X_n; \theta)}{\partial \theta^2}\right]$ and $I_n(\theta) = nI_1(\theta)$.

**Efficiency of an Absolutely Correct Estimator**: $e(\overline{\theta}) = \frac{1}{I_n(\theta)V(\overline{\theta})}$.

**Estimator $\overline{\theta}$ is**

**- unbiased**: $E(\overline{\theta}) = \theta$;

**- MVUE (minimum variance unbiased estimator)**: $E(\overline{\theta}) = \theta$ and $V(\overline{\theta}) \leq V(\hat{\theta})$, $\forall\ \hat{\theta}$ unbiased estimator;

**- consistent**: $\lim_{n\to\infty} P\left(|\overline{\theta} - \theta| \leq \varepsilon\right) = 1,\ \forall \varepsilon > 0$;

**- absolutely correct**: $E(\overline{\theta}) = \theta$ and $\lim_{n\to\infty} V(\overline{\theta}) = 0$;

**- efficient**: absolutely correct and $e(\overline{\theta}) = 1$.

**Statistic** $S = S(X_1, \ldots, X_n)$ **with value** $s = S(x_1, \ldots, x_n)$ **is**

**- sufficient**: - conditional joint pdf $f(x_1, \ldots, x_n; \theta | s)$ does not depend on $\theta$; OR
- $L(x_1, ..., x_n; \theta) = g(x_1, \ldots, x_n) h(s; \theta)$, for some measurable functions $g, h$.

**- complete for the family of distributions** $f(x; \theta), \theta \in A$: if $E\Big(\varphi(S)\Big) = 0, \ \forall \theta \in A$, then $\varphi = 0$ a. s.

**Rao - Cramer Inequality**: If $\overline{\theta}$ is an absolutely coorect estimator for $\theta$, then $V\left(\overline{\theta}\right) \geq \dfrac{1}{I_n(\theta)}$.

**Rao - Blackwell Theorem**: If $\hat{\theta}$ is an unbiased estimator for $\theta$ and $S$ is a sufficient statistic for $\theta$, then $\overline{\theta} = E\Big(\hat{\theta}|S\Big)$ is also an unbiased estimator for $\theta$ and $V\left(\overline{\theta}\right) \leq V\left(\hat{\theta}\right)$.

**Lehmann - Scheffé Theorem**: If $\hat{\theta}$ is an unbiased estimator for $\theta$ and $S$ is a sufficient and complete statistic for $\theta$, then $\overline{\theta} = E\Big(\hat{\theta}|S\Big)$ is a MVUE.

**Method of Moments**:
Solve the system $\nu_k \ (= E(X^k)) = \overline{\nu}_k$, for as many parameters as needed ($k = 1 \ldots$ nr. of unknown parameters).

**Method of Maximum Likelihood**:

Solve the system $\dfrac{\partial \ln L(X_1, ..., X_n | \theta_1, \ldots, \theta_m)}{\partial \theta_j} = 0, \ j = \overline{1, m}$ for the unknown parameters $\theta_1, ..., \theta_m$.

**Hypothesis Testing**: $H_0 : \theta = \theta_0$ with one of the alternatives $H_1 : \begin{cases} \theta < \theta_0 & \text{(left-tailed test)}, \\ \theta > \theta_0 & \text{(right-tailed test)}, \\ \theta \neq \theta_0 & \text{(two-tailed test)}, \end{cases}$

$TS$ is the test statistic, $RR$ is the rejection region.

**Significance Level**: $\alpha = P(\text{type I error}) = P(\text{reject } \ H_0 \mid H_0) = P(TS \in RR \mid \theta = \theta_0)$.

**Type II Error**: $\beta(\theta^*) = P(\text{type II error}) = P(\text{not reject } \ H_0 \mid H_1) = P(TS \notin RR \mid \theta = \theta^*)$.

**Power of a Test**: $\pi(\theta^*) = P(\text{reject } \ H_0 \mid \theta = \theta^*) = P(TS \in RR \mid \theta = \theta^*)$.

**Neyman-Pearson Lemma (NPL)**: Suppose we test two simple hypotheses $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$. Let $L(\theta^*)$ denote the likelihood function of the sample, when $\theta = \theta^*$. Then for every $\alpha \in (0, 1)$, a most powerful test (a test that maximizes the power $\pi(\theta_1)$) is the test with $RR = \left\{ \dfrac{L(\theta_1)}{L(\theta_0)} \geq k_\alpha \right\}$, for some constant $k_\alpha > 0$.