Bias of a point estimator $\hat{\theta}$ of a parameter θ (definition)

$$\boxed{\mathbf{Bias}(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \mathbf{E}\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}}$$

Clarification. θ is a parameter of the (*population*) probability distribution (θ is an unknown *constant*).
$\hat{\theta}$ is a *random variable* called a *point estimator* of the parameter θ. The value of $\hat{\theta}$ is calculated in the process of *sample study*, it is used as an *estimate* for the value of the unknown parameter θ.

Mean squared error of an estimator $\hat{\theta}$ of a parameter θ (definition)

$$\mathrm{MSE}(\hat{\theta}, \theta) = \mathrm{E}\left((\hat{\theta} - \theta)^2\right)$$

Property of the mean squared error

$$\boxed{\mathbf{MSE}(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \mathbf{Var}\,\hat{\boldsymbol{\theta}} + \left(\mathbf{Bias}(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta})\right)^2}$$

Unbiased estimator $\hat{\theta}$ of a parameter θ (definition)

$$\mathrm{Bias}(\hat{\theta}, \theta) = 0$$

Properties of the unbiased estimator

$$\mathrm{E}\hat{\theta} = \theta \qquad\qquad \mathrm{MSE}(\hat{\theta}, \theta) = \mathrm{Var}\,\hat{\theta}$$

Relative efficiency of two point estimators (definition)

$$e(\hat{\theta}_1, \hat{\theta}_2) = \frac{\mathrm{MSE}(\hat{\theta}_2, \theta)}{\mathrm{MSE}(\hat{\theta}_1, \theta)}$$

Comparison of point estimators

- Definition. $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2 \Leftrightarrow \begin{cases} \mathrm{MSE}(\hat{\theta}_1, \theta) \le \mathrm{MSE}(\hat{\theta}_2, \theta) \text{ for all } \theta \\ \mathrm{MSE}(\hat{\theta}_1, \theta) < \mathrm{MSE}(\hat{\theta}_2, \theta) \text{ for some } \theta \end{cases}$
- Property. $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2 \Leftrightarrow \begin{cases} e(\hat{\theta}_1, \hat{\theta}_2) \ge 1 \text{ for all } \theta \\ e(\hat{\theta}_1, \hat{\theta}_2) > 1 \text{ for some } \theta \end{cases}$

Standard estimators and their sample distributions

Estimator	Formula	Properties	Sample distribution
$X \sim Bernoulli(p)$ An estimator of p – *sample proportion*	$\hat{\boldsymbol{p}} = \frac{\boldsymbol{m}}{\boldsymbol{n}}$	$\mathrm{E}\hat{p} = p$ $\mathrm{Var}\,\hat{p} = \frac{p(1-p)}{n}$	$P\left(X = \frac{k}{n}\right) = C_n^k p^k (1-p)^{n-k}$ $np \ge 5$ and $n(p-1) \ge 5 \Rightarrow$ $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$ approximately
An estimator of μ – *sample mean*	$\bar{\boldsymbol{x}} = \frac{\mathbf{1}}{\boldsymbol{n}} \sum_{i=1}^{n} \boldsymbol{x}_i$	$\mathrm{E}\bar{x} = \mu$ $\mathrm{Var}\,\bar{x} = \frac{\sigma^2}{n}$	$n \ge 30 \Rightarrow \bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ approx. $X \sim N(\mu, \sigma^2) \Rightarrow \bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
An estimator of σ^2 when μ is *known*	$\hat{\boldsymbol{\sigma}}^2 = \frac{\mathbf{1}}{\boldsymbol{n}} \sum_{i=1}^{n} (\boldsymbol{x}_i - \boldsymbol{\mu})^2$	$\mathrm{E}\hat{\sigma}^2 = \sigma^2$ $X \sim N(\mu, \sigma^2) \Rightarrow \mathrm{Var}\,\hat{\sigma}^2 = \frac{2\sigma^4}{n}$	$X \sim N(\mu, \sigma^2) \Rightarrow \hat{\sigma}^2 \sim \frac{\sigma^2}{n} \chi^2(n)$
An estimator of σ^2 when μ is *unknown*	$\boldsymbol{s}^2 = \frac{\mathbf{1}}{\boldsymbol{n} - \mathbf{1}} \sum_{i=1}^{n} (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^2$	$\mathrm{E}s^2 = \sigma^2$ $\mathrm{MSE}(\hat{\sigma}^2, \sigma^2) < \mathrm{MSE}(s^2, \sigma^2)$	$X \sim N(\mu, \sigma^2) \Rightarrow$ $\bar{x}$ and s^2 are independent; $s^2 \sim \frac{\sigma^2}{n-1} \chi^2(n-1)$

__Consistency of an estimator__ (definition)

$$\hat{\theta} \text{ is a consistent estimator of } \theta \iff \forall \boldsymbol{\theta} \lim_{\boldsymbol{n} \to \infty} \boldsymbol{\hat{\theta}} = \boldsymbol{\theta}$$

Sufficient condition

$$\hat{\theta} \text{ is a consistent estimator of } \theta \iff \forall \theta \lim_{n \to \infty} \text{MSE}\left(\hat{\theta}, \theta\right) = 0$$

Clarification. $\hat{p}, \bar{x}, \hat{\sigma}^2$ and s^2 are consistent.

__Confidence intervals__

Example (a 2-sided z-interval for the population mean)

$$\mu = \bar{x} \pm z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}$$

Notations

- Confidence level: p
- Critical level: $\alpha = 1 - p$
- Critical value: $z_{\alpha/2}$
  $\boxed{\boldsymbol{P\left(Z > z_{\alpha/2}\right) = \alpha/2}}$
- Standard deviation of the sample mean: $\frac{\sigma}{\sqrt{n}}$
- Standard error of the sample mean: $\frac{s}{\sqrt{n}}$
- Margin of error: $m = z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}$ or $m = z_{\alpha/2} \times \frac{s}{\sqrt{n}}$

General formula

$$Confidence\ interval = (Point\ estimator) \pm (Margin\ of\ error) =$$

$$= (Point\ estimator) \pm (Critical\ value) \times (Standard\ deviation\ of\ the\ estimator)$$

__Hypothesis testing. Notions and notations__

- Null hypothesis: H_0
- Alternative hypothesis: H_a
- Significance of the test: $\alpha = P(type\ I\ error) = P(reject\ H_0 | H_0\ is\ true)$
- $\beta = P(type\ II\ error) = P(not\ reject\ H_0 | H_0\ is\ false)$
- Power of the test: $1 - \beta$
- Test statistic: Z
- p-value: $\boldsymbol{P(|Z| > |z_{\alpha/2}|)}$
  Clarification. This p-value implies that H_a is a 2-sides alternative hypothesis. For 1-sided H_a p-value can be $P(Z > z_\alpha)$ or $P(Z < -z_\alpha)$.
- Rejection rule: reject H_0 if **p-value** $\leq \boldsymbol{\alpha}$.
  Clarifications. 1. The lower p-value is, the stronger statistical evidence to reject H_0 in favor of H_a is.
  2. A value of the test statistic Z is in the *rejection region* if and only if $P\left(|Z| > \left|z_{\alpha/2}\right|\right) \leq \alpha$.