MOOC Statistique pour ingénieur
Thème 2 : échantillonnage, estimation
Vidéo 3 : Estimateurs

F. Delacroix M. Lecomte

Institut Mines-Télécom
École Nationale Supérieure des Mines de Douai

Sommaire

1. Estimateur d'une proportion
2. Notion d'estimateur, exemple de $\overline{X}$
3. Qualité d'un estimateur
4. Estimateurs d'une variance

Estimateur d'une proportion

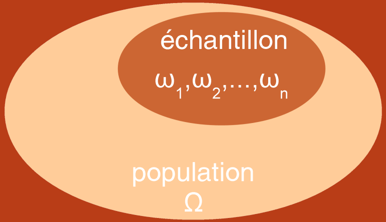
Exemple

Proportion p de pièces défectueuses au sein de la production ?

$$K \sim \mathcal{B}(n,p) \qquad F = \frac{K}{n}$$

$$\mathbb{E}(F) = p \qquad \mathbb{V}(F) = \frac{p(1-p)}{n} \to 0$$

F est un estimateur de p



K=nombre de pièces défectueuses dans l'échantillon

Sommaire

1. Estimateur d’une proportion
2. Notion d’estimateur, exemple de $\overline{X}$
3. Qualité d’un estimateur
4. Estimateurs d’une variance

Notion d'estimateur

X variable aléatoire sur Ω
θ paramètre de la loi de X dont la valeur exacte est inconnue

Définition

$\widehat{\Theta}_n$ *est un estimateur de θ si :*

1. $\widehat{\Theta}_n = f(X_1, \ldots, X_n)$
2. $\widehat{\Theta}_n \xrightarrow{\mathbb{P}} \theta$, *c'est-à-dire :*

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\left|\widehat{\Theta}_n - \theta\right| \geq \varepsilon\right) \xrightarrow[n \to +\infty]{} 0$$

Exemple : moyenne empirique

X variable aléatoire

$\mu = \mathbb{E}(X)$

Théorème (Loi faible des grands nombres)

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\mathbb{P}} \mu$$

$\overline{X}$ est un estimateur de μ.

Condition suffisante

En pratique on utilise souvent les conditions suffisantes suivantes.

Théorème

Si $\widehat{\Theta}_n$, fonction de l'échantillon $X_1, \ldots, X_n$ est tel que

1. $\mathbb{E}\left(\widehat{\Theta}_n\right) \xrightarrow[n \to +\infty]{} \theta$
2. $\mathbb{V}\left(\widehat{\Theta}_n\right) \xrightarrow[n \to +\infty]{} 0$

alors $\widehat{\Theta}_n$ est un estimateur de θ.

Exemple de l'estimateur F d'une proportion p :

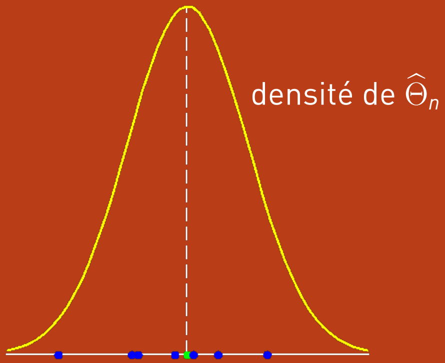
$$\mathbb{E}(F) = p \qquad \text{et} \qquad \mathbb{V}(F) = \frac{p(1-p)}{n} \longrightarrow 0.$$

Sommaire

1. Estimateur d'une proportion
2. Notion d'estimateur, exemple de $\overline{X}$
3. **Qualité d'un estimateur**
4. Estimateurs d'une variance

Biais d'un estimateur

Un estimateur peut être biaisé ou sans biais.



$$\theta = \mathbb{E}\left(\widehat{\Theta}_n\right)$$

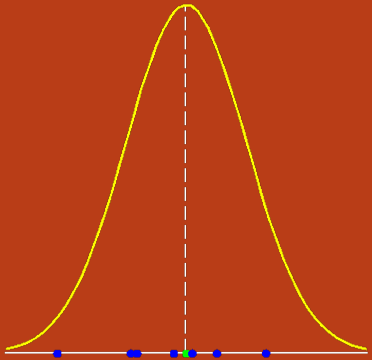
Définition

Un estimateur $\widehat{\Theta}_n$ de θ est sans biais si

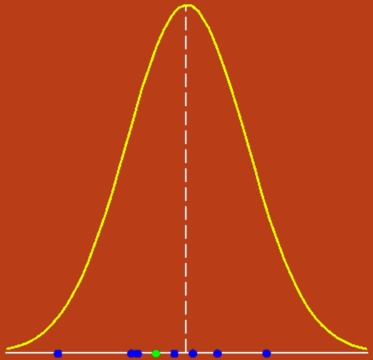
$$\theta = \mathbb{E}\left(\widehat{\Theta}_n\right).$$

Biais d'un estimateur

Un estimateur peut être biaisé ou sans biais.



$$\theta = \mathbb{E}\left(\widehat{\Theta}_n\right)$$



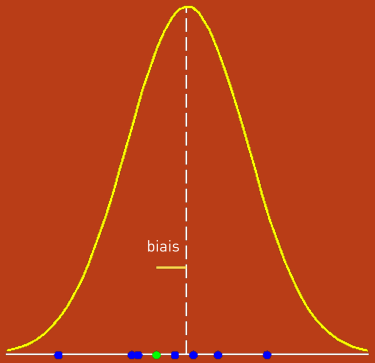
$$\theta \neq \mathbb{E}\left(\widehat{\Theta}_n\right)$$

Biais d'un estimateur

Un estimateur peut être biaisé ou sans biais.

Définition

$$biais\left(\widehat{\Theta}_n\right) = \mathbb{E}\left(\widehat{\Theta}_n\right) - \theta$$



$$\theta \neq \mathbb{E}\left(\widehat{\Theta}_n\right)$$

Exemple

Exemple

$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ *est tel que*

$$\mathbb{E}\left(\overline{X}\right) = \mu \quad et \quad \mathbb{V}\left(\overline{X}\right) = \frac{\sigma^2}{n}$$

$\overline{X}$ est un estimateur sans biais de $\mu = \mathbb{E}\left(X\right)$.

Erreur quadratique moyenne

Définition

Erreur quadratique moyenne d'un estimateur $\widehat{\Theta}_n$ de θ :

$$EQM\left(\widehat{\Theta}_n\right) = \mathbb{E}\left(\left(\widehat{\Theta}_n - \theta\right)^2\right)$$

$$= biais\left(\widehat{\Theta}_n\right)^2 + \mathbb{V}\left(\widehat{\Theta}_n\right).$$

Sommaire

1. Estimateur d’une proportion
2. Notion d’estimateur, exemple de $\overline{X}$
3. Qualité d’un estimateur
4. Estimateurs d’une variance

Variances empirique et corrigée

On a défini dans la vidéo 2

- $S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$
- $S^{*2} = \frac{n}{n-1}S^2$

Théorème

$$\mathbb{E}\left(S^2\right) = \frac{n-1}{n}\sigma^2 \quad et \quad \mathbb{V}\left(S^2\right) \to 0$$

S^2 et S^{*2} sont des estimateurs de σ^2.
S^2 est biaisé tandis que S^{*2} est sans biais.

En pratique

Estimations ponctuelles de σ^2 :

- s^2 obtenue par l'estimateur S^2
- s^{*2} obtenue par l'estimateur S^{*2}.

Laquelle utiliser ?

- Si $n \geq 30$, $s^2 \simeq s^{*2}$
- si $n < 30$, S^2 a tendance à sous-estimer σ^2, on utilise plutôt S^{*2}.

Exemple

Pour $s^2 = 3$ et $n = 20$,

$$s^{*2} = \frac{20}{19} s^2 \simeq 3,16.$$

INSTITUT
Mines-Télécom