MOOC Statistique pour ingénieur
Thème 2 : échantillonnage, estimation
Vidéo 2 : Distributions d'échantillonnage

F. Delacroix M. Lecomte

Institut Mines-Télécom
École Nationale Supérieure des Mines de Douai

Sommaire

1. Moyenne empirique
2. Variance empirique

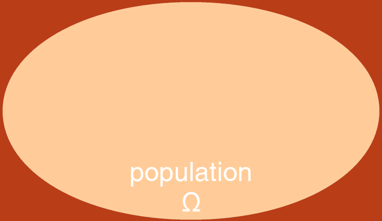
Moyenne empirique de l'échantillon

X=longueur

Moyenne empirique de l'échantillon

X caractère étudié

$\mathbb{E}(X) = \mu$, $\mathbb{V}(X) = \sigma^2$.

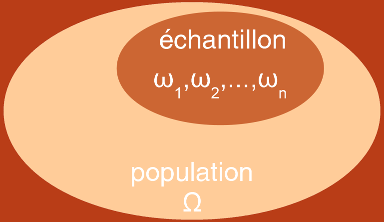


Moyenne empirique de l'échantillon

X caractère étudié

$\mathbb{E}(X) = \mu$, $\mathbb{V}(X) = \sigma^2$.

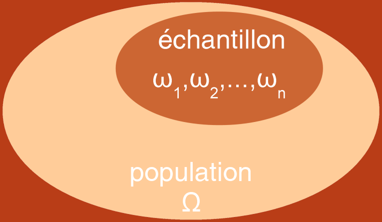
$$\begin{array}{ccc} & X & \\ \omega_1 & \longrightarrow & x_1 \\ \omega_2 & \longrightarrow & x_2 \\ \vdots & & \vdots \\ \omega_n & \longrightarrow & x_n \end{array}$$



Moyenne empirique de l'échantillon

X caractère étudié
$\mathbb{E}(X) = \mu$, $\mathbb{V}(X) = \sigma^2$.

$$\begin{array}{ccccc} x_1, & x_2, & \ldots, & x_n & \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \\ \uparrow & \uparrow & & \uparrow & \uparrow \\ X_1, & X_2, & \ldots, & X_n & \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \end{array}$$



Moyenne empirique de l'échantillon

X=longueur

$\mu = \mathbb{E}(X)$ relative à la population

$\bar{x}$ est la moyenne de l'échantillon

$\bar{x}$ est une estimation ponctuelle de μ

Espérance et variance de $\overline{X}$

$$\mathbb{E}\left(\overline{X}\right) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\mu = \mu$$

$$\mathbb{V}\left(\overline{X}\right) = \mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{V}\left(X_i\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n}$$

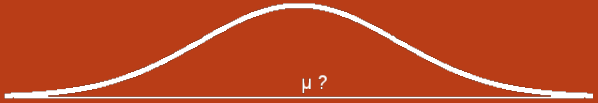
Espérance et variance de $\overline{X}$

Théorème

$$\mathbb{E}\left(\overline{X}\right) = \mu$$

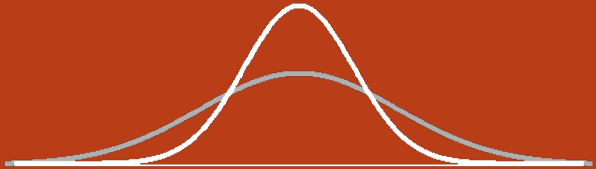
$$\mathbb{V}\left(\overline{X}\right) = \frac{\sigma^2}{n}$$

Cas des échantillons gaussiens



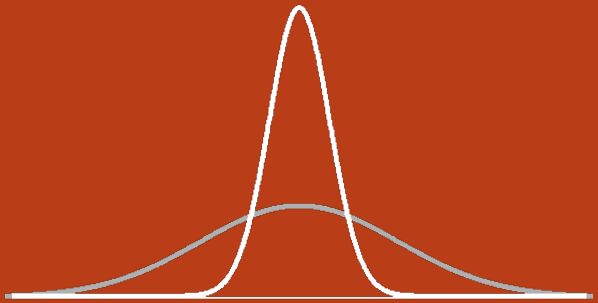
$n = 1$ (loi de X)

Cas des échantillons gaussiens



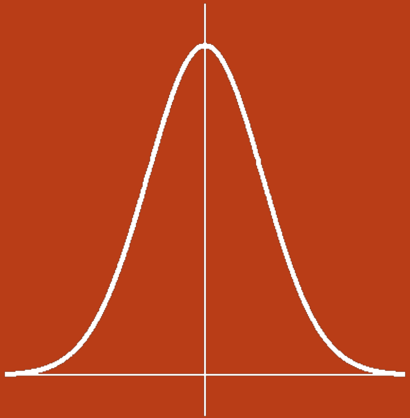
$\overline{X}$ avec $n = 3$

Cas des échantillons gaussiens



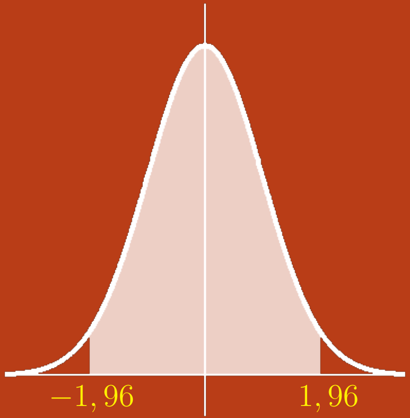
$\overline{X}$ avec $n = 10$

Cas des échantillons gaussiens



$$U = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \sim \mathcal{N}(0,1)$$

Cas des échantillons gaussiens



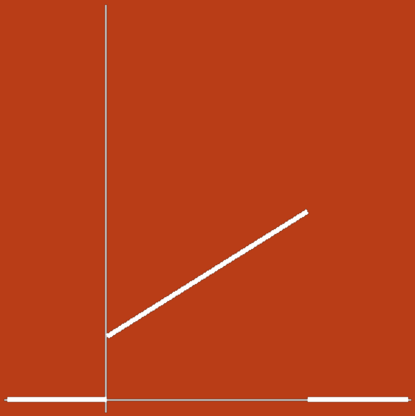
$$U = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

$$\mathbb{P}\left(-1,96 \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq 1,96\right) = 0,95$$

$$\text{soit } \mathbb{P}\left(\bar{X} - 1,96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1,96\frac{\sigma}{\sqrt{n}}\right) = 0,95$$

$$IC_{0,95}(\mu) = \left[\bar{X} - 1,96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1,96\frac{\sigma}{\sqrt{n}}\right]$$

Et si X ne suit pas une loi normale ?



$n = 1$ (densité de X)

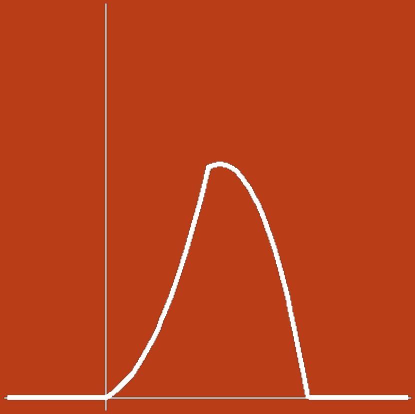
Exemple

Soit une v.a. X de densité

$$f(x) = \begin{cases} \dfrac{1+x}{4} & \text{si } x \in [0, 2] \\ 0 & \text{sinon.} \end{cases}$$

Quelle est la loi de $\overline{X}$?

Et si X ne suit pas une loi normale ?



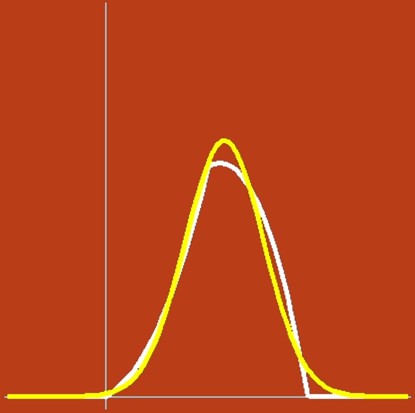
densité de $\overline{X}$ pour $n = 2$

Exemple

Soit une v.a. X de densité

$$f(x) = \begin{cases} \frac{1+x}{4} & si\ x \in [0,2] \\ 0 & sinon. \end{cases}$$

Et si X ne suit pas une loi normale ?



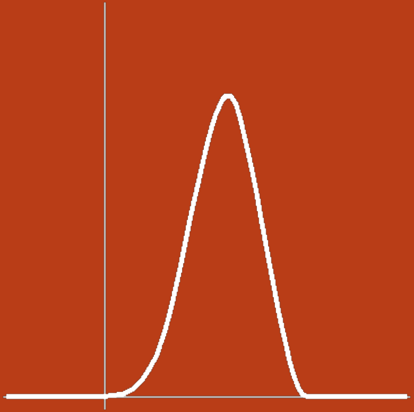
densité $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ pour $n = 2$

Exemple

Soit une v.a. X de densité

$$f(x) = \begin{cases} \frac{1+x}{4} & \text{si } x \in [0,2] \\ 0 & \text{sinon.} \end{cases}$$

Et si X ne suit pas une loi normale ?



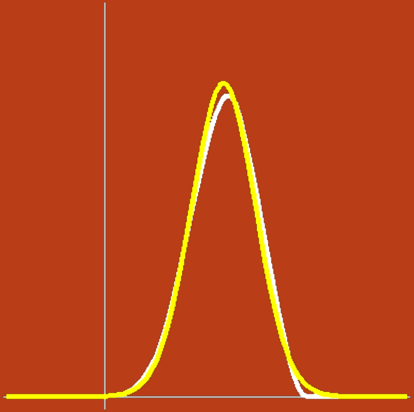
densité de $\overline{X}$ pour $n = 3$

Exemple

Soit une v.a. X de densité

$$f(x) = \begin{cases} \dfrac{1+x}{4} & si\ x \in [0,2] \\ 0 & sinon. \end{cases}$$

Et si X ne suit pas une loi normale ?



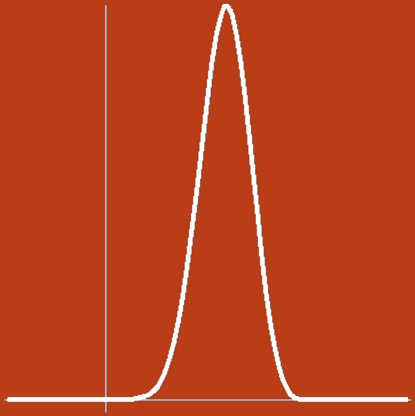
densité $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ pour $n = 3$

Exemple

Soit une v.a. X de densité

$$f(x) = \begin{cases} \frac{1+x}{4} & \text{si } x \in [0,2] \\ 0 & \text{sinon.} \end{cases}$$

Et si X ne suit pas une loi normale ?



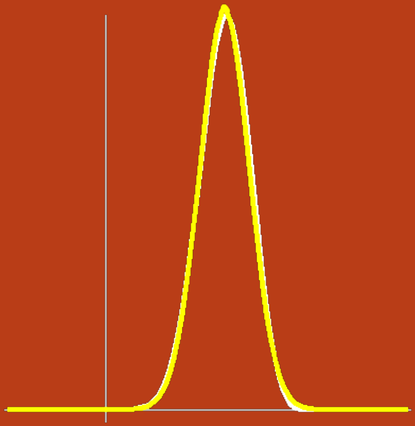
densité de $\overline{X}$ pour $n = 5$

Exemple

Soit une v.a. X de densité

$$f(x) = \begin{cases} \dfrac{1+x}{4} & \text{si } x \in [0,2] \\ 0 & \text{sinon.} \end{cases}$$

Et si X ne suit pas une loi normale ?



densité $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ pour $n = 5$

Exemple

Soit une v.a. X de densité

$$f(x) = \begin{cases} \frac{1+x}{4} & \text{si } x \in [0,2] \\ 0 & \text{sinon.} \end{cases}$$

Et si X ne suit pas une loi normale ?

Théorème (Central-limite)

Sous les hypothèses de la statistique classique, la variable aléatoire

$$U = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

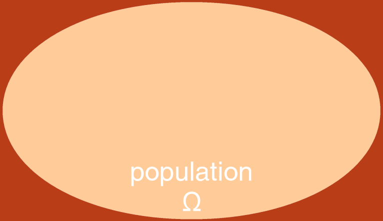
suit approximativement $\mathcal{N}(0, 1)$.

Autrement dit, $\overline{X}$ suit approximativement $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$.

Sommaire

Variance empirique de l'échantillon

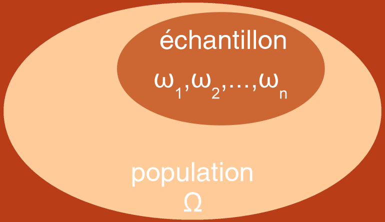
X caractère étudié
$\mathbb{E}(X) = \mu, \mathbb{V}(X) = \sigma^2.$



Variance empirique de l'échantillon

X caractère étudié
$\mathbb{E}(X) = \mu$, $\mathbb{V}(X) = \sigma^2$.

$x_1, \quad x_2, \quad \dots, \quad x_n \qquad s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$

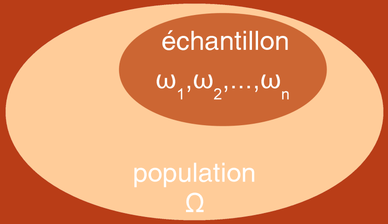


Variance empirique de l'échantillon

X caractère étudié

$\mathbb{E}(X) = \mu$, $\mathbb{V}(X) = \sigma^2$.

$$
\begin{array}{cccccc}
x_1, & x_2, & \ldots, & x_n & & s^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})^2 \\
\uparrow & \uparrow & & \uparrow & & \uparrow \\
X_1, & X_2, & \ldots, & X_n & & S^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2
\end{array}
$$



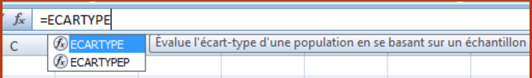
Estimateurs de la variance

$$\mathbb{E}\left(S^2\right) = \frac{n-1}{n}\sigma^2 \neq \sigma^2 \qquad S^2 = \text{Variance empirique}$$

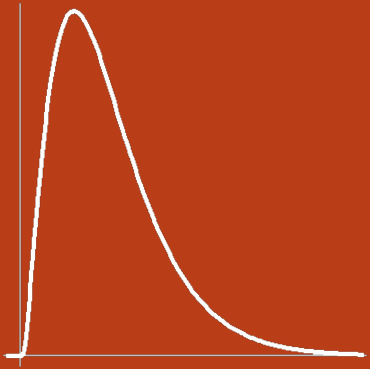
S^2 est une statistique biaisée pour σ^2.

$$S^{*2} = \frac{n}{n-1}S^2 \quad \mathbb{E}\left(S^{*2}\right) = \sigma^2 \qquad S^{*2} = \text{Variance corrigée}$$

S^{*2} est une statistique non biaisée pour σ^2.



Cas des échantillons gaussiens



$\nu = 5$

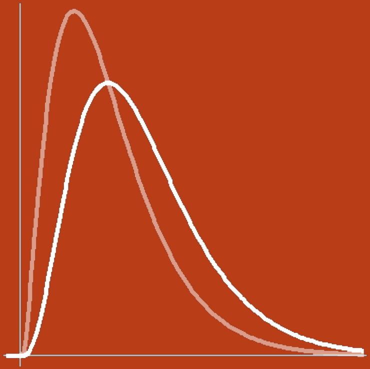
Théorème

Si $X \sim \mathcal{N}(\mu, \sigma^2)$ *alors la v.a.*

$$Z = \frac{n S^2}{\sigma^2}$$

suit la loi du χ^2 *à* $\nu = n - 1$ *degrés de liberté.*

Cas des échantillons gaussiens



$\nu = 7$

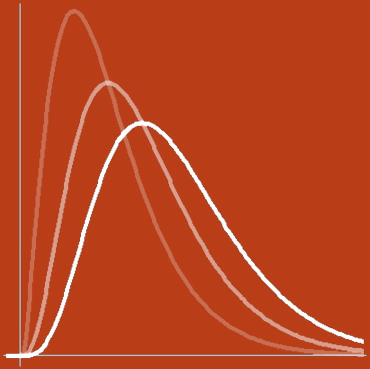
Théorème

Si $X \sim \mathcal{N}(\mu, \sigma^2)$ *alors la v.a.*

$$Z = \frac{n\,S^2}{\sigma^2}$$

suit la loi du χ^2 *à* $\nu = n - 1$ *degrés de liberté.*

Cas des échantillons gaussiens



$\nu = 9$

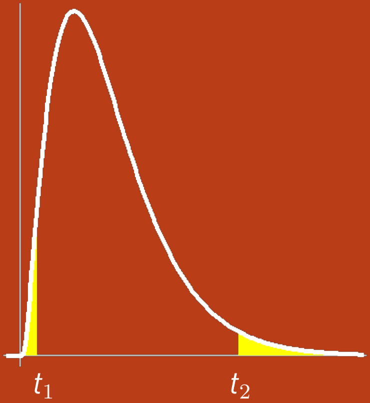
Théorème

Si $X \sim \mathcal{N}(\mu, \sigma^2)$ *alors la v.a.*

$$Z = \frac{n\,S^2}{\sigma^2}$$

suit la loi du χ^2 *à* $\nu = n - 1$ *degrés de liberté.*

Intervalle de confiance pour σ^2



$$\mathbb{P}(Z < t_1) = 2,5\%$$
$$\mathbb{P}(Z > t_2) = 2,5\%$$
$$\text{Alors } \mathbb{P}\left(t_1 \leq \frac{n\,S^2}{\sigma^2} \leq t_2\right) = 95\%$$
$$IC_{0,95}(\sigma^2) = \left[\frac{n\,S^2}{t_2}, \frac{n\,S^2}{t_1}\right]$$

INSTITUT
Mines-Télécom