

Moyenne empirique – Variance empirique

1. Moyenne empirique

a. Hypothèses et notations

Soit un échantillon (X_1, X_2, \dots, X_n) de la loi de X .

X_1, X_2, \dots, X_n sont donc n variables mutuellement indépendantes et suivant toutes la même loi que X .

Soit

$$E(X_1) = E(X_2) = \dots = E(X_n) = m$$

$$V(X_1) = V(X_2) = \dots = V(X_n) = \sigma^2$$

La variable $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ est appelée moyenne empirique de X_1, X_2, \dots, X_n

On se propose de montrer que \overline{X}_n est un estimateur sans biais de m et convergent

b. Montrons que \overline{X}_n est un estimateur sans biais de m

Pour cela, montrons que $E(\overline{X}_n) = m$

$$\text{En effet : } E(\overline{X}_n) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} nm$$

Finalement, on a bien : $E(\overline{X}_n) = m$. Donc \overline{X}_n est un estimateur sans biais de m

En d'autres termes, $b(\overline{X}_n) = 0$

c. Montrons que \overline{X}_n est un estimateur convergent de m

Pour cela, déterminons d'abord le risque quadratique $r(\overline{X}_n)$ de \overline{X}_n pour ensuite en déduire que $\lim_{n \rightarrow +\infty} r(\overline{X}_n) = 0$

$$r(\overline{X}_n) = b^2(\overline{X}_n) + V(\overline{X}_n)$$

Comme $b(\overline{X}_n) = 0$:

$$r(\overline{X}_n) = V(\overline{X}_n) = V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) \text{ car } X_1, X_2, \dots, X_n \text{ sont mutuellement indépendantes}$$

Ainsi :

$$r(\overline{X}_n) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{1}{n^2} n \sigma^2$$

Finalement : $r(\overline{X_n}) = V(\overline{X_n}) = \frac{\sigma^2}{n}$ et $\lim_{n \rightarrow +\infty} \frac{\sigma^2}{n} = 0$.

Donc $\overline{X_n}$ est un estimateur convergent.

2. Variance empirique

a. Hypothèses et notations

Soit un échantillon (X_1, X_2, \dots, X_n) de la loi de X.

X_1, X_2, \dots, X_n sont donc n variables mutuellement indépendantes et suivant toutes la même loi que X.

Soit

$$E(X_1) = E(X_2) = \dots = E(X_n) = m$$

$$V(X_1) = V(X_2) = \dots = V(X_n) = \sigma^2$$

La variable $\overline{X_n} = \frac{1}{n} \sum_{i=1}^n X_i$ est appelée moyenne empirique de X_1, X_2, \dots, X_n

La variable $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X_n})^2$ est appelée variance empirique de X_1, X_2, \dots, X_n

On se propose de montrer que S_n^2 est un estimateur asymptotiquement sans biais de σ^2

b. Détermination du biais de l'estimateur et de son caractère asymptotiquement sans biais

Par définition : $b_{\sigma^2}(S_n^2) = E(S_n^2) - \sigma^2$

- Détermination de $E(S_n^2)$

$$E(S_n^2) = E \left[\frac{1}{n} \sum_{i=1}^n (X_i - \overline{X_n})^2 \right] = \frac{1}{n} E \left[\sum_{i=1}^n X_i^2 - 2X_i \overline{X_n} + (\overline{X_n})^2 \right]$$

$$E(S_n^2) = \frac{1}{n} E \left[\sum_{i=1}^n X_i^2 - 2\overline{X_n} \sum_{i=1}^n X_i + \sum_{i=1}^n (\overline{X_n})^2 \right]$$

Or :

$$\overline{X_n} = \frac{1}{n} \sum_{i=1}^n X_i$$

Donc :

$$\sum_{i=1}^n X_i = n\overline{X_n}$$

Ainsi :

$$E(S_n^2) = \frac{1}{n} E \left[\sum_{i=1}^n X_i^2 - 2\overline{X_n} n\overline{X_n} + n(\overline{X_n})^2 \right]$$

Finalement :

$$E(S_n^2) = \frac{1}{n} E \left[\sum_{i=1}^n X_i^2 - n(\overline{X_n})^2 \right] = \frac{1}{n} \sum_{i=1}^n E(X_i^2) - E[(\overline{X_n})^2]$$

- Détermination de $\sum_{i=1}^n E(X_i^2)$

Par théorème de Huyghens :

$$V(X_i) = E(X_i^2) - [E(X_i)]^2$$

Donc :

$$E(X_i^2) = V(X_i) + [E(X_i)]^2$$

$$E(X_i^2) = \sigma^2 + m^2$$

- Détermination de $E[(\overline{X_n})^2]$

De même :

$$E[(\overline{X_n})^2] = V(\overline{X_n}) + [E(\overline{X_n})]^2$$

Donc :

$$E[(\overline{X_n})^2] = \frac{\sigma^2}{n} + m^2$$

- Détermination de $E(S_n^2)$

Ainsi :

$$\begin{aligned} E(S_n^2) &= \frac{1}{n} \sum_{i=1}^n (\sigma^2 + m^2) - \left(\frac{\sigma^2}{n} + m^2 \right) = \frac{1}{n} n(\sigma^2 + m^2) - \left(\frac{\sigma^2}{n} + m^2 \right) \\ &= \sigma^2 + m^2 - \frac{\sigma^2}{n} - m^2 \end{aligned}$$

$$E(S_n^2) = \sigma^2 \left(1 - \frac{1}{n} \right)$$

Finalement :

$$E(S_n^2) = \frac{n-1}{n} \sigma^2$$

- Détermination de $b(S_n^2)$ et conclusion

$$b_{\sigma^2}(S_n^2) = \frac{n-1}{n} \sigma^2 - \sigma^2 = -\frac{\sigma^2}{n}$$

$$\lim_{n \rightarrow +\infty} b_{\sigma^2}(S_n^2) = \lim_{n \rightarrow +\infty} -\frac{\sigma^2}{n} = 0$$

Conclusion : S_n^2 est un estimateur asymptotiquement sans biais de σ^2