

---

# Inégalité de Bienaymé-Tchebychev

---

## Lemme 1 - Inégalité de Markov.

Soit  $X$  une variable aléatoire à valeurs positives ou nulles. Pour tout  $a > 0$ ,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

*Démonstration.*

Notons  $p = \mathbb{P}(X \geq a)$ . Puisque  $\mathbf{1}_{\{X \geq a\}} \hookrightarrow \mathcal{B}(p)$ , on a  $\mathbb{E}(\mathbf{1}_{\{X \geq a\}}) = p$ . Donc

$$a\mathbb{P}(X \geq a) = a\mathbb{E}(\mathbf{1}_{\{X \geq a\}}) = \mathbb{E}(a\mathbf{1}_{\{X \geq a\}}).$$

Soit  $\omega \in \Omega$ . Si  $\omega \in \{X \geq a\}$ , alors  $X(\omega) \geq a$  et ainsi  $a\mathbf{1}_{\{X \geq a\}}(\omega) = a \leq X(\omega)$ . Si  $\omega \notin \{X \geq a\}$ , alors  $a\mathbf{1}_{\{X \geq a\}}(\omega) = 0 \leq X(\omega)$ . Donc  $a\mathbf{1}_{\{X \geq a\}} \leq X$ .

Par croissance de l'espérance, on en déduit que  $\mathbb{E}(a\mathbf{1}_{\{X \geq a\}}) \leq \mathbb{E}(X)$ , donc  $a\mathbb{P}(X \geq a) \leq \mathbb{E}(X)$  puis, puisque  $a > 0$ ,  $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$ .  $\square$

**Remarque 2.** Ce lemme n'est intéressant que si  $a$  est grand et plus précisément lorsque  $a > \mathbb{E}[X]$ .

## Théorème 3 - Inégalité de Bienaymé - Tchebychev.

Soit  $X$  une variable aléatoire réelle d'espérance  $m$  et d'écart-type  $\sigma$  non nul. Pour tout  $d > 0$ ,

$$\mathbb{P}(|X - m| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

*Démonstration.*

Soit  $\varepsilon > 0$ . Appliquons l'inégalité de Markov à la variable aléatoire  $(X - \mathbb{E}(X))^2$  avec  $a = \varepsilon^2$ . Alors

$$\mathbb{P}((X - \mathbb{E}(X))^2 \geq \varepsilon^2) \leq \frac{\mathbb{E}((X - \mathbb{E}(X))^2)}{\varepsilon^2}$$

donc

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}$$

car l'événement  $\{(X - \mathbb{E}(X))^2 \geq \varepsilon^2\}$  est égal à l'événement  $\{|X - \mathbb{E}(X)| \geq \varepsilon\}$ .  $\square$

**Remarque 4.** Cette propriété n'a d'intérêt que si  $\varepsilon > \sigma$ . Elle explique que la probabilité que  $X$  soit « loin » de sa moyenne est petite et décroît au delà de  $\sigma$  en  $O(1/d^2)$  ( $\varepsilon$  étant un minorant de la distance de  $X$  à sa moyenne).