

# Espérance et variance

Cours de É. Bouchet – PCSI

24 mai 2022

## Table des matières

<b>1</b>	<b>Espérance d'une variable aléatoire</b>	<b>2</b>
1.1	Définition et premières propriétés . . . . .	2
1.2	Espérance des lois usuelles . . . . .	3
1.3	Formule de transfert et espérance du produit . . . . .	4
<b>2</b>	<b>Variance, écart-type et covariance</b>	<b>6</b>
2.1	Variance, écart-type et premières propriétés . . . . .	6
2.2	Variance des lois usuelles . . . . .	7
2.3	Covariance et variables aléatoires décorréliées . . . . .	8
<b>3</b>	<b>Inégalités probabilistes</b>	<b>9</b>
3.1	Inégalité de Markov . . . . .	9
3.2	Inégalité de Bienaymé-Tchebychev . . . . .	10

Dans tout le chapitre, on se place dans un espace probabilisé fini  $(\Omega, P)$ .

# 1 Espérance d'une variable aléatoire

## 1.1 Définition et premières propriétés

### Définition (Espérance).

Soit  $X$  une variable aléatoire réelle ou complexe. On appelle **espérance** de  $X$  la valeur

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$

Lorsque  $E(X) = 0$ , on dit que la variable est **centrée**.

**Remarque.** L'espérance donne une moyenne des valeurs atteintes par  $X$ , c'est donc un **indicateur de position**. Si on décale toutes les valeurs de  $X$ , l'espérance sera décalée en conséquence.

**Exemple 1.** Soit  $X$  une variable aléatoire telle que  $X(\Omega) = \llbracket -1, 1 \rrbracket$ ,  $P(X = -1) = \frac{1}{4}$ ,  $P(X = 0) = \frac{1}{2}$  et  $P(X = 1) = \frac{1}{4}$ . Alors  $E(X) = -\frac{1}{4} + 0 + \frac{1}{4} = 0$ , donc  $X$  est une variable centrée.

**Exemple 2.** Soit  $A$  un ensemble fini non vide et  $X$  une variable aléatoire telle que  $X \sim \mathcal{U}(A)$ . On a :

$$E(X) = \sum_{x \in A} xP(X = x) = \frac{1}{\text{Card}(A)} \sum_{x \in A} x.$$

### Proposition (Expression alternative).

Soit  $X$  une variable aléatoire réelle ou complexe. Alors  $E(X) = \sum_{\omega \in \Omega} X(\omega)P(\{\omega\})$ .

*Démonstration.*  $(\{X = x\})_{x \in X(\Omega)}$  forme un système complet d'événements, donc :

$$\sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) = \sum_{x \in X(\Omega)} \sum_{\omega \in \{X=x\}} X(\omega)P(\{\omega\}) = \sum_{x \in X(\Omega)} \sum_{\omega \in \{X=x\}} xP(\{\omega\}) = \sum_{x \in X(\Omega)} xP(\{X = x\}) = E(X).$$

□

**Remarque.** Cette forme servira peu en exercices, mais sera très utile pour démontrer les différents résultats du chapitre.

### Proposition (Linéarité).

Soit  $X$  et  $Y$  deux variables aléatoires réelles ou complexes et  $\lambda, \mu$  deux constantes, alors

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$

*Démonstration.* L'expression alternative donne directement :

$$E(\lambda X + \mu Y) = \sum_{\omega \in \Omega} (\lambda X + \mu Y)(\omega)P(\{\omega\}) = \lambda \sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) + \mu \sum_{\omega \in \Omega} Y(\omega)P(\{\omega\}) = \lambda E(X) + \mu E(Y).$$

□

**Remarque.** En particulier, l'espérance de la somme est égale à la somme des espérances.

**Exemple 3.** On lance deux dés, et on note  $S$  la somme de leurs résultats. Déterminer l'espérance de  $S$ . On ne cherche surtout pas à calculer la loi de  $S$  (qui n'est pas demandée).

On note  $X_1$  et  $X_2$  les résultats des dés, alors  $S = X_1 + X_2$ ,  $X_1 \sim \mathcal{U}(\llbracket 1, 6 \rrbracket)$  et  $X_2 \sim \mathcal{U}(\llbracket 1, 6 \rrbracket)$ . Cela donne :

$$E(X_1) = \sum_{k=1}^6 kP(X_1 = k) = \sum_{k=1}^6 \frac{k}{6} = \frac{1}{6} \times \frac{6 \times 7}{2} = \frac{7}{2},$$

de même pour  $E(X_2)$ . Donc  $E(S) = E(X_1) + E(X_2) = \frac{7}{2} + \frac{7}{2} = 7$ .

**Proposition** (Inégalité triangulaire).

Soit  $X$  une variable aléatoire réelle ou complexe, alors  $|E(X)| \leq E(|X|)$ .

*Démonstration.* Il suffit d'utiliser l'expression alternative, l'inégalité triangulaire sur les sommes et la positivité de la probabilité :

$$|E(X)| = \left| \sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) \right| \leq \sum_{\omega \in \Omega} |X(\omega)P(\{\omega\})| = \sum_{\omega \in \Omega} |X(\omega)| P(\{\omega\}) = E(|X|).$$

□

**Proposition** (Positivité).

Soit  $X$  une variable aléatoire réelle positive. Alors  $E(X) \geq 0$ .

*Démonstration.* Puisque  $X$  est positive,  $\forall \omega \in \Omega, X(\omega)P(\{\omega\}) \geq 0$ , donc par passage à la somme :

$$E(X) = \sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) \geq 0.$$

□

**Proposition** (Croissance).

Soit  $X$  et  $Y$  deux variables aléatoires réelles. Si  $X \leq Y$ , alors  $E(X) \leq E(Y)$ .

*Démonstration.* Comme  $Y - X$  est une variable aléatoire positive, alors  $E(Y - X) \geq 0$ . La linéarité donne alors  $E(Y) - E(X) \geq 0$ , donc  $E(Y) \geq E(X)$ . □

**Remarque.** En particulier, si  $X$  est à valeurs dans un intervalle  $I$ ,  $E(X) \in I$ .

## 1.2 Espérance des lois usuelles

**Proposition** (Espérance d'une variable aléatoire constante).

Soit  $X$  une variable aléatoire constante de valeur  $m$ , alors  $E(X) = m$ .

*Démonstration.* Un retour à la définition donne directement  $E(X) = m \times P(X = m) = m \times 1 = m$ . □

**Proposition** (Espérance d'une variable aléatoire de Bernoulli).

Soit  $p \in [0, 1]$ . Si  $X$  est une variable aléatoire telle que  $X \sim \mathcal{B}(p)$ , alors  $E(X) = p$ .

*Démonstration.* Un retour à la définition donne directement  $E(X) = 1P(X = 1) + 0P(X = 0) = p$ .  $\square$

**Remarque.** Soit  $A \in \mathcal{P}(\Omega)$ , le cas particulier de la variable indicatrice donne  $E(\mathbb{1}_A) = P(A)$ , puisqu'on sait que  $\mathbb{1}_A \sim \mathcal{B}(P(A))$ .

**Proposition** (Espérance d'une variable aléatoire binomiale).

Soit  $p \in [0, 1]$  et  $n \in \mathbb{N}^*$ . Si  $X$  est une variable aléatoire telle que  $X \sim \mathcal{B}(n, p)$ , alors  $E(X) = np$ .

*Démonstration.* On peut aborder la preuve de deux manières différentes.

— Un retour à la définition donne :

$$\begin{aligned} E(X) &= \sum_{k=0}^n kP(X = k) \\ &= 0 + \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n k \frac{n}{k} \binom{n-1}{k-1} p^k (1-p)^{n-k} \\ &= n \sum_{i=0}^{n-1} \binom{n-1}{i} p^{i+1} (1-p)^{n-1-i} \quad \text{en posant } i = k - 1 \\ &= np \sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{n-1-i} \end{aligned}$$

$$E(X) = np \quad \text{par formule du binôme de Newton,}$$

— On peut aussi introduire des variables aléatoires indépendantes  $X_1, \dots, X_n$  de loi  $\mathcal{B}(p)$ . Alors  $X_1 + \dots + X_n$  et  $X$  ont même loi, ce qui donne :

$$E(X) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n p = np.$$

$\square$

### 1.3 Formule de transfert et espérance du produit

**Proposition** (Formule de transfert).

Soit  $X$  une variable aléatoire réelle ou complexe et  $g$  une fonction définie sur  $X(\Omega)$ . Alors :

$$E(g(X)) = \sum_{x \in X(\Omega)} g(x)P(X = x).$$

*Démonstration.*  $(\{X = x\})_{x \in X(\Omega)}$  forme un système complet d'événements, donc :

$$\begin{aligned} E(g(X)) &= \sum_{\omega \in \Omega} g(X(\omega))P(\{\omega\}) \\ &= \sum_{x \in X(\Omega)} \sum_{\omega \in \{X=x\}} g(X(\omega))P(\{\omega\}) \\ &= \sum_{x \in X(\Omega)} \sum_{\omega \in \{X=x\}} g(x)P(\{\omega\}) \\ E(g(X)) &= \sum_{x \in X(\Omega)} g(x)P(X = x). \end{aligned}$$

□

**Remarque.** Cette formule s'applique en particulier aux couples et  $n$ -uplets de variables aléatoires. Si  $X$  et  $Y$  sont des variables aléatoires et  $g$  est une fonction définie sur  $X(\Omega) \times Y(\Omega)$ , on a par exemple :

$$E(g(X, Y)) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} g(x, y)P(X = x, Y = y).$$

**Proposition.**

Soit  $X$  et  $Y$  deux variables aléatoires réelles ou complexes. Si elles sont indépendantes, alors

$$E(XY) = E(X)E(Y).$$

*Démonstration.* D'après la formule de transfert, puis par indépendance des variables aléatoires,

$$\begin{aligned} E(XY) &= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} xyP(X = x, Y = y) \\ &= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} xyP(X = x)P(Y = y) \\ &= \left( \sum_{x \in X(\Omega)} xP(X = x) \right) \left( \sum_{y \in Y(\Omega)} yP(Y = y) \right) \\ E(XY) &= E(X)E(Y). \end{aligned}$$

□

**Remarque.** On peut utiliser ce résultat pour montrer que deux variables aléatoires ne sont pas indépendantes : il suffit de vérifier que  $E(XY) \neq E(X)E(Y)$ .

**Remarque.** Ce résultat s'étend sans difficulté au cas de  $n$  variables aléatoires indépendantes.

**Exemple 4.** Soient  $X_1, X_2$  et  $X_3$  indépendantes, suivant des lois de Bernoulli de paramètre  $p \in ]0, 1[$ . Montrer que les variables aléatoires  $Y_1 = X_1X_2$  et  $Y_2 = X_2X_3$  ne sont pas indépendantes.

L'indépendance des  $X_i$  donne  $E(Y_1) = E(X_1)E(X_2) = p^2$  et  $E(Y_2) = E(X_2)E(X_3) = p^2$ . Par ailleurs,  $X_2^2 = X_2$  puisque  $X_2$  ne prend que les valeurs 0 et 1, donc :

$$E(Y_1Y_2) = E(X_1X_2X_2X_3) = E(X_1X_2X_3) = E(X_1)E(X_2)E(X_3) = p^3 \neq p^4 = E(Y_1)E(Y_2).$$

Donc  $Y_1$  et  $Y_2$  ne sont pas indépendantes.

**Remarque.** Attention, ce résultat n'est pas une équivalence : la relation  $E(XY) = E(X)E(Y)$  ne suffit pas à garantir l'indépendance des variables.

**Exemple 5.** Soit  $Z \sim \mathcal{U}(\{-1, 1\})$  et  $Y$  une variable aléatoire indépendante de  $Z$ . Alors,  $Y$  et  $X = ZY$  ne sont clairement pas indépendantes (elles ont même module), pourtant le lemme des coalitions la relation  $E(Z) = 0$  et l'indépendance de  $Z$  et  $Y$  donnent :

$$E(XY) = E(ZY^2) = E(Z)E(Y^2) = 0 = E(Y)E(Z)E(Y) = E(Y)E(X).$$

## 2 Variance, écart-type et covariance

### 2.1 Variance, écart-type et premières propriétés

**Définition** (Variance, écart-type).

Soit  $X$  une variable aléatoire réelle. On appelle **variance** de  $X$  la valeur

$$V(X) = E((X - E(X))^2)$$

et **écart-type** de  $X$  la valeur  $\sigma(X) = \sqrt{V(X)}$ . Lorsque  $V(X) = 1$ , on dit que la variable est **réduite**.

**Remarque.** Contrairement au cas de l'espérance, le programme ne s'intéresse pas à la variance de variables aléatoires complexes. Cela permet de garantir que la variance sera toujours positive (puisque c'est l'espérance d'une fonction positive).

**Remarque.** La variance (ou l'écart-type) donnent une idée d'à quel point les valeurs de  $X$  s'écartent de leur valeur moyenne, c'est donc un **indicateur de dispersion**. Si on décale toutes les valeurs de  $X$ , la variance ne sera pas modifiée.

**Proposition.**

Soient  $(a, b) \in \mathbb{R}^2$  et  $X$  une variable aléatoire réelle, alors  $V(aX + b) = a^2V(X)$ .

*Démonstration.* Il suffit d'utiliser la définition de la variance et la linéarité de l'espérance :

$$\begin{aligned} V(aX + b) &= E((aX + b - E(aX + b))^2) \\ &= E((aX + b - aE(X) - b)^2) \\ &= E((a(X - E(X)))^2) \\ &= a^2E((X - E(X))^2) \\ V(aX + b) &= a^2V(X) \end{aligned}$$

□

**Remarque.** On en déduit que  $\sigma(aX + b) = \sqrt{a^2V(X)} = |a|\sigma(X)$ .

**Remarque.** Soit  $X$  une variable aléatoire réelle telle que  $\sigma(X) > 0$ . On pose  $Z = \frac{X - E(X)}{\sigma(X)}$ , les propriétés de l'espérance et la variance donnent alors :

$$E(Z) = \frac{E(X) - E(X)}{\sigma(X)} = 0 \quad \text{et} \quad V(Z) = \frac{V(X)}{\sigma(X)^2} = \frac{V(X)}{V(X)} = 1.$$

$Z$  est donc une variable aléatoire centrée réduite.

**Proposition** (Formule de König-Huygens).

Soit  $X$  une variable aléatoire réelle. Alors :  $V(X) = E(X^2) - E(X)^2$ .

*Démonstration.* Par linéarité de l'espérance,

$$V(X) = E((X - E(X))^2)$$

$$\begin{aligned}
&= E(X^2 + E(X)^2 - 2E(X)X) \\
&= E(X^2) + E(X)^2 - (2E(X))E(X) \\
&= E(X^2) + E(X)^2 - 2E(X)^2 \\
V(X) &= E(X^2) - E(X)^2
\end{aligned}$$

□

**Proposition** (Cas de la variance nulle).

Soit  $X$  une variable aléatoire réelle. Alors :  $V(X) = 0 \iff P(X = E(X)) = 1$ .

*Démonstration.* La formule de transfert et la définition de la variance donnent :

$$V(X) = \sum_{x \in X(\Omega)} (x - E(X))^2 P(X = x).$$

Les termes de cette somme sont tous à valeurs positives, on a donc :

$$\begin{aligned}
V(X) = 0 &\iff \sum_{x \in X(\Omega)} (x - E(X))^2 P(X = x) = 0 \\
&\iff \forall x \in X(\Omega), (x - E(X))^2 P(X = x) = 0 \\
&\iff \forall x \in X(\Omega), (x - E(X))^2 = 0 \text{ ou } P(X = x) = 0 \\
&\iff \forall x \in X(\Omega) \setminus \{E(X)\}, P(X = x) = 0 \\
V(X) = 0 &\iff P(X = E(X)) = 1,
\end{aligned}$$

où la dernière équivalence fonctionne car  $\sum_{x \in X(\Omega)} P(X = x) = 1$ .

□

## 2.2 Variance des lois usuelles

**Proposition** (Variance d'une variable aléatoire de Bernoulli).

Soit  $p \in [0, 1]$ . Si  $X$  est une variable aléatoire telle que  $X \sim \mathcal{B}(p)$ , alors  $V(X) = p(1 - p)$ .

*Démonstration.* Un retour à la définition et la formule de transfert donnent :

$$\begin{aligned}
V(X) &= (1 - E(X))^2 P(X = 1) + (0 - E(X))^2 P(X = 0) \\
&= (1 - p)^2 p + p^2 (1 - p) \\
&= p(1 - p)(1 - p + p) \\
V(X) &= p(1 - p).
\end{aligned}$$

□

**Proposition** (Variance d'une variable aléatoire binomiale).

Soit  $p \in [0, 1]$  et  $n \in \mathbb{N}^*$ . Si  $X$  est une variable aléatoire telle que  $X \sim \mathcal{B}(n, p)$ , alors  $V(X) = np(1 - p)$ .

*Démonstration.* La formule de transfert donne :

$$\begin{aligned}
 E(X^2) &= \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} \\
 &= 0 + n \sum_{i=0}^{n-1} (i+1) \binom{n-1}{i} p^{i+1} (1-p)^{n-1-i} \quad \text{en simplifiant } \binom{n}{k} \text{ puis en posant } i = k-1 \\
 &= np \left( \sum_{i=0}^{n-1} i \binom{n-1}{i} p^i (1-p)^{n-1-i} + (p+1-p)^{n-1} \right) \quad \text{par linéarité et binôme de Newton} \\
 &= np \left( 0 + \sum_{i=1}^{n-1} i \frac{n-1}{i} \binom{n-2}{i-1} p^i (1-p)^{n-1-i} \right) + np \\
 &= n(n-1)p^2 \left( \sum_{k=0}^{n-2} \binom{n-2}{k} p^k (1-p)^{n-2-k} \right) + np \quad \text{en posant } k = i-1 \\
 &= n(n-1)p^2 + np \quad \text{par binôme de Newton} \\
 E(X^2) &= np(pn - p + 1).
 \end{aligned}$$

Par formule de König-Huygens, on trouve alors  $V(X) = E(X^2) - E(X)^2 = np(pn - p + 1) - (np)^2 = np(1 - p)$ . Une preuve alternative basée sur des sommes de variables aléatoires indépendantes sera vue plus tard dans le chapitre.  $\square$

### 2.3 Covariance et variables aléatoires décorrélées

#### Définition (Covariance).

Soit  $X$  et  $Y$  deux variables aléatoires réelles. On appelle **covariance** de  $X$  et  $Y$  la valeur :

$$\text{Cov}(X, Y) = E\left((X - E(X))(Y - E(Y))\right).$$

Lorsque  $\text{Cov}(X, Y) = 0$ , on dit que les variables  $X$  et  $Y$  sont **décorrélées**.

**Remarque.** En particulier,  $\text{Cov}(X, X) = V(X)$  et  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ .

#### Proposition (Formule de König-Huygens).

Soient  $X$  et  $Y$  deux variables aléatoires réelles. Alors :  $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$ .

*Démonstration.* Un retour à la définition et la linéarité de l'espérance donnent :

$$\begin{aligned}
 \text{Cov}(X, Y) &= E((X - E(X))(Y - E(Y))) \\
 &= E(XY - E(Y)X - E(X)Y + E(X)E(Y)) \\
 &= E(XY) - E(Y)E(X) - E(X)E(Y) + E(X)E(Y) \\
 \text{Cov}(X, Y) &= E(XY) - E(X)E(Y)
 \end{aligned}$$

$\square$

**Remarque.** Si  $X$  et  $Y$  sont deux variables aléatoires indépendantes, on a déjà montré que  $E(XY) = E(X)E(Y)$ , et donc  $\text{Cov}(X, Y) = 0$ . Deux variables aléatoires indépendantes sont donc toujours décorrélées.

La réciproque est fautive, par contre : des variables peuvent être décorrélées sans être indépendantes, pour peu que  $E(XY) = E(X)E(Y)$  (cf contre-exemple étudié plus haut).



**Proposition.**

Soit  $X$  et  $Y$  deux variables aléatoires réelles. Alors :

$$V(X + Y) = V(X) + V(Y) + 2 \operatorname{Cov}(X, Y).$$

En particulier, si  $X$  et  $Y$  sont décorrélées,  $V(X + Y) = V(X) + V(Y)$ .

*Démonstration.* La définition de la variance et la linéarité de l'espérance donnent :

$$\begin{aligned} V(X + Y) &= E((X + Y - E(X + Y))^2) \\ &= E((X - E(X) + Y - E(Y))^2) \\ &= E((X - E(X))^2 + (Y - E(Y))^2 + 2(X - E(X))(Y - E(Y))) \\ V(X + Y) &= V(X) + V(Y) + 2 \operatorname{Cov}(X, Y). \end{aligned}$$

En particulier, si  $X$  et  $Y$  sont décorrélées,  $V(X + Y) = V(X) + V(Y)$ . □

**Remarque.** Dans le cas de variables indépendantes, on a donc  $V(X + Y) = V(X) + V(Y)$ . Ce résultat se généralise naturellement à une somme de  $n$  variables aléatoires mutuellement indépendantes.

**Remarque.** Ce résultat donne une nouvelle manière de calculer la variance de la loi binomiale.

Soit  $p \in [0, 1]$ ,  $n \in \mathbb{N}^*$  et  $X$  une variable aléatoire réelle telle que  $X \sim \mathcal{B}(n, p)$ . On introduit des variables aléatoires indépendantes  $X_1, \dots, X_n$  de loi  $\mathcal{B}(p)$ . Alors  $X_1 + \dots + X_n$  et  $X$  ont même loi, ce qui donne :

$$V(X) = V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) = \sum_{i=1}^n p(1-p) = np(1-p).$$

### 3 Inégalités probabilistes

#### 3.1 Inégalité de Markov

**Proposition (Inégalité de Markov).**

Soit  $X$  une variable aléatoire réelle positive et  $\lambda \in \mathbb{R}_+^*$  :

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

*Démonstration.* On a  $P(X \geq \lambda) = \sum_{\substack{x \in X(\Omega) \\ x \geq \lambda}} P(X = x)$ . La définition de  $E(X)$  et la positivité de  $X$  donnent donc :

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x) \geq \sum_{\substack{x \in X(\Omega) \\ x \geq \lambda}} xP(X = x) \geq \sum_{\substack{x \in X(\Omega) \\ x \geq \lambda}} \lambda P(X = x) = \lambda \sum_{\substack{x \in X(\Omega) \\ x \geq \lambda}} P(X = x).$$

Donc  $E(X) \geq \lambda P(X \geq \lambda)$ . Diviser par  $\lambda \in \mathbb{R}_+^*$  fournit alors le résultat souhaité.

Variante : on pouvait aussi remarquer que  $\mathbb{1}_{\{X \geq \lambda\}} \lambda \leq \mathbb{1}_{\{X \geq \lambda\}} X \leq X$ . La croissance de l'espérance donne alors directement  $\lambda P(X \geq \lambda) \leq E(X)$ . □

**Remarque.** Les majorations fournies par l'inégalité de Markov sont assez brutales : pour de petites valeurs de  $\lambda$ ,  $\frac{E(X)}{\lambda}$  peut même dépasser 1 ! C'est cependant une première approche intéressante, en attendant de savoir comment obtenir des bornes plus précises.

**Exemple 6.** Une entreprise compte 300 employés. Chacun d'eux téléphone en moyenne 6 minutes par heures. Déterminer un entier  $N$  tel que si l'entreprise a  $N$  lignes de téléphone, la probabilité que toutes les lignes soient utilisées au même instant est au plus égale à  $\frac{1}{4}$ .

On pose  $X$  la variable aléatoire égale au nombre d'employés au téléphone à l'instant étudié.  $X$  compte le nombre de succès (« employé au téléphone », de probabilité  $\frac{6}{60} = \frac{1}{10}$ ) dans une succession de 300 expériences indépendantes, donc  $X \sim \mathcal{B}(300, \frac{1}{10})$ . De plus,  $X$  est à valeurs positives et  $E(X) = 300 \times \frac{1}{10} = 30$ . Soit  $N \in \mathbb{N}$ , l'inégalité de Markov donne alors :

$$P(X \geq N) \leq \frac{E(X)}{N} = \frac{30}{N}.$$

On cherche donc  $N$  tel que  $\frac{30}{N} \leq \frac{1}{4}$ , ce qui équivaut à  $120 \leq N$ . Disposer de 120 lignes de téléphones suffirait donc.

### 3.2 Inégalité de Bienaymé-Tchebychev

**Proposition** (Inégalité de Bienaymé-Tchebychev).

Soit  $X$  une variable aléatoire réelle et  $\varepsilon \in \mathbb{R}_+^*$  :

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

*Démonstration.* La variable  $(X - E(X))^2$  est positive, on peut donc lui appliquer l'inégalité de Markov, avec  $\lambda = \varepsilon^2 \in \mathbb{R}_+^*$  :

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

Or,  $(X - E(X))^2 \geq \varepsilon^2 \iff |X - E(X)| \geq \varepsilon$  car  $\varepsilon > 0$  et par stricte croissance de la fonction racine carrée sur  $\mathbb{R}_+$ . Donc

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

□

**Remarque.** Ici encore, les majorations obtenues sont brutales, mais l'approche reste intéressante.

**Exemple 7.** On dispose d'une pièce truquée dont la probabilité d'obtention de pile est notée  $p$ . Pour déterminer  $p$ , on lance cette pièce  $n \in \mathbb{N}^*$  fois et on note  $F$  la fréquence d'apparition de pile ainsi obtenue. À partir de quelle valeur de  $n$  la probabilité pour que  $F$  soit une approximation de  $p$  à  $10^{-2}$  près est-elle supérieure à 0,9 ?

On note  $N$  la variable aléatoire égale au nombre de pile obtenu en  $n$  lancers.  $N$  compte le nombre de succès (« obtenir pile », de probabilité  $p$ ) dans une succession de  $n$  expériences indépendantes, donc  $N \sim \mathcal{B}(n, p)$ .

Par ailleurs,  $F = \frac{N}{n}$ , donc  $E(F) = \frac{E(N)}{n} = \frac{np}{n} = p$  et  $V(F) = \frac{V(N)}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$ . L'inégalité de Bienaymé-Tchebychev donne alors :

$$P(|F - p| \geq 10^{-2}) = P(|F - E(F)| \geq 10^{-2}) \leq \frac{V(F)}{10^{-4}} = \frac{p(1-p)10^4}{n}.$$

On cherche une valeur de  $n$  telle que  $P(|F - p| < 10^{-2}) \geq \frac{9}{10}$ , ce qui équivaut à  $1 - P(|F - p| \geq 10^{-2}) \geq \frac{9}{10}$  puis à  $P(|F - p| \geq 10^{-2}) \leq \frac{1}{10}$  (choisir ici quelles égalités sont strictes ou larges n'a pas d'importance tant qu'on garde la cohérence du passage au complémentaire).

Il suffit donc de déterminer  $n$  tel que  $\frac{p(1-p)10^4}{n} \leq \frac{1}{10}$ , ce qui équivaut à  $n \geq 10^5 p(1-p)$ . Sauf qu'on ne connaît pas la valeur de  $p$ ... On sait par contre que  $p \in [0, 1]$  et une étude de  $p \rightarrow p(1-p)$  sur cet intervalle donne que  $p(1-p) \leq \frac{1}{4}$ . Choisir  $n \geq \frac{10^5}{4} = 25\,000$  convient donc.

**Remarque.** Soit  $n \in \mathbb{N}^*$  et  $X_1, \dots, X_n$  des variables aléatoires indépendantes de même loi. On note  $m$  leur espérance et  $\sigma$  leur écart-type. Posons  $S_n = \frac{1}{n} \sum_{i=1}^n X_i$  leur valeur moyenne. Alors par indépendance et propriétés de l'espérance et la variance,

$$E(S_n) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} nm = m \quad \text{et} \quad V(S_n) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

Appliquer l'inégalité de Bienaymé-Tchebychev à  $X_1$  et à  $S_n$  donne alors respectivement : soit  $\varepsilon > 0$ ,

$$P(|X_1 - m| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \quad \text{et} \quad P(|S - m| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Le  $n$  au dénominateur dans cette deuxième inégalité la rend bien plus intéressante : si on part du principe que la valeur de  $m$  est inconnue, plus  $n$  augmente, plus  $S_n$  représente une approximation fiable de  $m$ . Pour estimer la valeur du paramètre  $m$ , utiliser une moyenne de variables indépendantes de même loi est donc très efficace.