Espérance et variance

Cours de É. Bouchet – PCSI

4 juin 2024

Table des matières

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

1 Espérance d'une variable aléatoire

1.1 Définition et premières propriétés

Définition 1.1 (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. Soit X une variable aléatoire qui vérifie $X(\Omega) = [\![-1, 1]\!]$, $P(X = -1) = \frac{1}{4}$, $P(X = 0) = \frac{1}{2}$ 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. Soit A un ensemble fini non vide et X une variable aléatoire telle que $X \sim \mathcal{U}(A)$. On a :

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

Proposition 1.2 (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 1.3 (Linéarité)

Soit X et Y deux variables aléatoires réelles ou complexes et λ, μ 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.

Exercice 1. On lance deux dés, et on note S la somme de leurs résultats. Déterminer l'espérance de S.
Solution : 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 $S = X_1 + X_2$, $X_1 \sim \mathcal{U}([\![1,6]\!])$ et $X_2 \sim \mathcal{U}([\![1,6]\!])$. 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 1.4 (Inégalité triangulaire)

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

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

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

□

Proposition 1.5 (Positivité)

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

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

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

□

Proposition 1.6 (Croissance)

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

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

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

1.2 Espérance des lois usuelles

Proposition 1.7 (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 1.8 (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 $E(X) = 1P(X = 1) + 0P(X = 0) = p$. □

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 1.9 (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 :

$$\begin{aligned}
E(X) &= \sum_{k=0}^{n} kP(X=k) \\
&= 0 + \sum_{k=1}^{n} k\binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k\frac{n}{k}\binom{n-1}{k-1} p^k (1-p)^{n-k} \\
&= n\sum_{i=0}^{n-1} \binom{n-1}{i} p^{i+1} (1-p)^{n-1-i} \qquad \text{en posant } i = k-1 \\
&= np\sum_{i=0}^{n-1} \binom{n-1}{i} p^{i} (1-p)^{n-1-i} \\
E(X) &= np \qquad \text{par formule du binôme de Newton,}
\end{aligned}$$

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

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

□

1.3 Formule de transfert et espérance du produit

Proposition 1.10 (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 :

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

□

Exercice 2. Soit X une variable aléatoire de loi $\mathcal{U}(\llbracket 1,6\rrbracket)$. Déterminer $E(X^2)$.
Solution : D'après la formule de transfert,

$$E(X^2) = \sum_{k=1}^{6} k^2 P(X=k) = \frac{1}{6}\sum_{k=1}^{6} k^2 = \frac{1}{6}\frac{6\times 7\times 13}{6} = \frac{91}{6}.$$

Remarque. La formule de transfert s'applique en particulier aux couples (ou 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 1.11 (Cas de variables aléatoires indépendantes)

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,

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

□

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

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

Exercice 3. 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.
Solution : 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 :

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

où $p^3\neq p^4$ car $p\neq 0$ et $p\neq 1$. 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. 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 2.1 (Variance, écart-type)

Soit X une variable aléatoire réelle. On appelle **variance** de X la valeur $V(X) = E\left((X - E(X))^2\right)$ 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 2.2 (Variance d'une combinaison linéaire)

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 :

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

□

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 $Z = \dfrac{X - E(X)}{\sigma(X)}$, 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 2.3 (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,

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

□

Proposition 2.4 (Cas de la variance nulle)

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

Remarque. Dans le cas d'un univers Ω fini, $P(X = E(X)) = 1$ signifie que X est une variable aléatoire constante, qui vaut $E(X)$.

Démonstration. Par formule de transfert et définition de la variance, $V(X) = \sum_{x \in X(\Omega)} (x - E(X))^2 P(X = x)$.

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 2.5 (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 2.6 (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 :

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

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

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. □

2.3 Covariance et variables aléatoires décorrélées

Définition 2.7 (Covariance)

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

$$\mathrm{Cov}(X, Y) = E\Big((X - E(X))(Y - E(Y))\Big).$$

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

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

Proposition 2.8 (Formule de König-Huygens)

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

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

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

□

Remarque. Si X et Y sont deux variables aléatoires indépendantes, on a déjà montré que $E(XY) = E(X)E(Y)$, et donc $\mathrm{Cov}(X, Y) = 0$. Deux variables aléatoires indépendantes sont donc toujours décorrélées.
La réciproque est fausse, 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 2.9 (Variance d'une somme de variables aléatoires)

Soit X et Y deux variables aléatoires réelles. Alors : $V(X + Y) = V(X) + V(Y) + 2\,\mathrm{Cov}(X, Y)$.
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 :

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

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, ..., X_n de loi $\mathcal{B}(p)$. Alors $X_1 + \cdots + X_n$ 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 3.1 (Inégalité de Markov)

Soit X une variable aléatoire réelle positive et $\lambda \in \mathbb{R}_+^*$, alors $P\left(X \geqslant \lambda\right) \leqslant \dfrac{E\left(X\right)}{\lambda}$.

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

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

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

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

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

Exercice 4. Une entreprise compte 300 employés. Chacun d'eux téléphone en moyenne 6 minutes par heure, indépendamment des autres. 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 $\frac{1}{4}$.
Solution : 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 :

$$P(X \geqslant N) \leqslant \frac{E(X)}{N} = \frac{30}{N}.$$

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

3.2 Inégalité de Bienaymé-Tchebychev

Proposition 3.2 (Inégalité de Bienaymé-Tchebychev)

Soit X une variable aléatoire réelle et $\varepsilon \in \mathbb{R}_+^*$, alors $P\left(\left|X - E\left(X\right)\right| \geqslant \varepsilon\right) \leqslant \dfrac{V(X)}{\varepsilon^2}$.

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

$$P\left((X - E(X))^2 \geqslant \varepsilon^2\right) \leqslant \frac{E\left((X - E(X))^2\right)}{\varepsilon^2} = \frac{V(X)}{\varepsilon^2}.$$

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

$$P\left(|X - E(X)| \geqslant \varepsilon\right) = P\left((X - E(X))^2 \geqslant \varepsilon^2\right) \leqslant \frac{V(X)}{\varepsilon^2}.$$

□

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

Exercice 5. On dispose d'une pièce truquée dont la probabilité d'obtention de pile est notée $p \in [0, 1]$. 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 $0,9$?
<u>Solution :</u> 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 :

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

On cherche une valeur de n telle que $P(|F - p| < 10^{-2}) \geqslant \frac{9}{10}$, ce qui équivaut à $1 - P(|F - p| \geqslant 10^{-2}) \geqslant \frac{9}{10}$ puis à $P(|F - p| \geqslant 10^{-2}) \leqslant \frac{1}{10}$ (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 $\frac{p(1-p)10^4}{n} \leqslant \frac{1}{10}$, ce qui équivaut à $n \geqslant 10^5 p(1 - p)$. Sauf qu'on ne connaît pas la valeur de p... On sait par contre que $p \in [0, 1]$ et une étude de $p \to p(1 - p)$ sur cet intervalle donne que $p(1 - p) \leqslant \frac{1}{4}$. Choisir $n \geqslant \frac{10^5}{4} = 25\ 000$ convient donc.

Remarque. Soit $n \in \mathbb{N}^*$ et X_1, ..., X_n des variables aléatoires indépendantes de même loi. On note m leur espérance et σ 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 $\varepsilon > 0$,

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

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.