Lycée Lucie Aubrac - Terminale

Inégalités classiques en probabilités et loi des grands nombres

Table des matières

1 Inégalités de Markov et de Bienaymé-Tchebychev

1.1 Inégalité de Markov

Définition 1

Soit $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire. On dit que X est positive si pour tout $\omega \in \Omega$, $X(\omega) \geqslant 0$.

Proposition 1 – Inégalité de Markov

Soit $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire positive. Pour tout $a \in \mathbb{R}_+^*$,

$$\mathrm{P}(X \geqslant a) \leqslant \frac{\mathrm{E}(X)}{a}.$$

Démonstration.
Soit $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire positive. Soit $a \in \mathbb{R}_+^*$. On note $x_1, \dots, x_n$ les valeurs prises par X et $p_i = \mathrm{P}(X = x_i)$. Par définition de l'espérance, $\mathrm{E}(X) = \sum_{i=0}^{n} p_i x_i$.
On sépare cette somme en deux. Plus précisément,

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{\{i, x_i \geqslant a\}} p_i x_i + \sum_{\{i, x_i < a\}} p_i x_i \\
&\geqslant \sum_{\{i, x_i \geqslant a\}} p_i x_i \quad \text{(car pour tout i, } x_i \geqslant 0) \\
&\geqslant \sum_{\{i, x_i \geqslant a\}} p_i a \\
&\geqslant a \left(\sum_{\{i, x_i \geqslant a\}} p_i \right) \\
&\geqslant a\mathrm{P}(X \geqslant a)
\end{aligned}$$

On en déduit l'inégalité souhaitée ($a \neq 0$) :

$$\mathrm{P}(X \geqslant a) \leqslant \frac{\mathrm{E}(X)}{a}.$$

□

Exemple.
Le salaire moyen en 2022 en France s'élève à 2 424 euros. On choisit un salarié au hasard et on note X son salaire. Déterminer un majorant de $\mathrm{P}(X \geqslant 5\,000)$ à l'aide de l'inégalité de Markov.

Solution :
X est une variable aléatoire positive et on peut appliquer l'inégalité de Markov pour $a = 5\,000$:

$$\begin{aligned}
\mathrm{P}(X \geqslant 5\,000) &\leqslant \frac{\mathrm{E}(X)}{5\,000} \\
&\leqslant \frac{2\,424}{5\,000} \\
&\leqslant 0{,}485
\end{aligned}$$

Cela signifie qu'il y a au maximum 48,5% de salariés gagnant 5 000 euros, cela correspondant au cas limite où tous les autres salariés gagnent 0 euros.

1.2 Inégalité de Bienaymé-Tchebychev

Proposition 2 – Inégalité de Bienaymé-Tchebychev

Soit $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire. Pour tout $a \in \mathbb{R}_+^*$,

$$\mathrm{P}\Big(|X - \mathrm{E}(X)| \geqslant a\Big) \leqslant \frac{V(X)}{a^2}.$$

Démonstration.
Pour tout $a \in \mathbb{R}_+^*$,

$$|X - \mathrm{E}(X)| \geqslant a \Longleftrightarrow (X - \mathrm{E}(X))^2 \geqslant a^2.$$

Par conséquent, $\mathrm{P}\Big(|X - \mathrm{E}(X)| \geqslant a\Big) = \mathrm{P}\Big(\,(X - \mathrm{E}(X))^2 \geqslant a^2\Big)$.
On applique ensuite l'inégalité de Markov à la variable aléatoire $(X - \mathrm{E}(X))^2$:

$$\begin{aligned}\mathrm{P}\Big(|X - \mathrm{E}(X)| \geqslant a\Big) &= \mathrm{P}\Big(\,(X - \mathrm{E}(X))^2 \geqslant a^2\Big) \\ &\leqslant \frac{\mathrm{E}\left((X - \mathrm{E}(X))^2\right)}{a^2} \\ &\leqslant \frac{V(X)}{a^2}\end{aligned}$$

□

Proposition 3 – (Corollaire)

Soit $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire. Pour tout $a \in \mathbb{R}_+^*$,

$$\mathrm{P}\Big(|X - \mathrm{E}(X)| < a\Big) \geqslant 1 - \frac{V(X)}{a^2}.$$

Démonstration.
Il suffit d'appliquer l'inégalité de Bienaymé-Tchebychev puis de remarquer que l'événement $\{|X - \mathrm{E}(X)| < a\}$ est exactement l'événement contraire de l'événement $\{|X - \mathrm{E}(X)| \geqslant a\}$.

□

Remarque.
Si on note $p = 1 - \dfrac{V(X)}{a}$, on peut interpréter la proposition 3 de la manière suivante : la probabilité que X soit proche de $E(X)$ (à une distance de a près) est au moins égale à p.

Exemple.
On lance 600 fois un dé équilibré à 6 faces. On note X le nombre de fois où l'on obtient le résultat 6. Déterminer un minorant de $\mathrm{P}(90 \leqslant X \leqslant 110)$.

Solution :
X correspond à la répétitions d'épreuves indépendantes à deux issues possibles de même probabilité de succès. La variable X suit donc une loi binomiale de paramètres $n = 600$ et $p = \dfrac{1}{6}$

et on a donc $\mathrm{E}(X) = np = 100$ et $V(X) = np(1-p) \simeq 83{,}3$. De plus, d'après l'inégalité de Bienaymé-Tchebychev appliquée pour $a = 11$:

$$\begin{aligned}
& \mathrm{P}\Big(|X - \mathrm{E}(X)| \geqslant a\Big) \leqslant \frac{V(X)}{a^2} \\
\iff & \mathrm{P}\Big(|X - 100| \geqslant 11\Big) \leqslant \frac{83{,}3}{11^2} \\
\iff & \mathrm{P}\Big(|X - 100| \geqslant 11\Big) \leqslant 0{,}68 \\
\iff & \mathrm{P}\Big(|X - 100| < 11\Big) \geqslant 0{,}68 \\
\iff & \mathrm{P}\Big(90 \leqslant X \leqslant 110\Big) \geqslant 0{,}68
\end{aligned}$$

Ainsi, cela signifie que le nombre de 6 sera compris entre 90 et 110 avec une probabilité d'au moins 68%.

Remarque.
Si on applique l'inégalité de Bienaymé-Tchebychev pour $a = 2\sigma$, on obtient

$$\mathrm{P}\Big(|X - \mathrm{E}(X)| \geqslant 2\sigma\Big) \leqslant \frac{\sigma^2}{(2\sigma)^2} \leqslant \frac{1}{4},$$

soit, de manière équivalente,

$$\mathrm{P}\Big(|X - \mathrm{E}(X)| < 2\sigma\Big) \geqslant \frac{3}{4}.$$

Cela donne une majoration intéressante mais qui est loin d'être optimale. On sait par exemple que si X suit une loi binomiale, l'intervalle $\Big[\mathrm{E}(X) - 2\sigma\,;\mathrm{E}(X) + 2\sigma\Big]$ contient, en pratique, environ 95% des valeurs. Cela est bien supérieur à 0,75, valeur obtenu par l'inégalité de Bienaymé-Tchebychev.

2 Inégalité de concentration et loi des grands nombres

2.1 Inégalité de concentration

Définition 2 – (Rappel)

Soit X une variable aléatoire. Un échantillon de taille n de X est la donnée de n variables aléatoires indépendantes $X_1, \ldots, X_n$ et qui sont toutes identiquement distribuées, suivant la même loi que X.

Proposition 4 – Inégalité de concentration

Soit $(X_k)_{1 \leqslant k \leqslant n}$ un échantillon d'une variable aléatoire X. On note $M_n = \dfrac{X_1 + \ldots + X_n}{n}$. Alors, pour tout $a \in \mathbb{R}_+^*$,

$$\mathrm{P}\Big(|M_n - \mathrm{E}(X)| \geqslant a\Big) \leqslant \frac{V(X)}{na^2}.$$

Démonstration.
Pour tout $a \in \mathbb{R}_+^*$,

$$|M_n - \mathrm{E}(X)| \geqslant a \iff \left|\frac{X_1 + \ldots + X_n}{n} - \mathrm{E}(X)\right| \geqslant a \iff |X_1 + \ldots X_n - n\mathrm{E}(X)| \geqslant na.$$

Par conséquent, $\mathrm{P}\Big(|M_n - \mathrm{E}(X)| \geqslant a\Big) = \mathrm{P}\Big(|X_1 + \ldots X_n - n\mathrm{E}(X)| \geqslant na\Big)$.
On applique ensuite l'inégalité de Bienaymé-Tchebychev à la variable aléatoire $X_1 + \ldots + X_n$:

$$\begin{aligned}
\mathrm{P}\Big(|M_n - \mathrm{E}(X)| \geqslant a\Big) &= \mathrm{P}\Big(|X_1 + \ldots X_n - n\mathrm{E}(X)| \geqslant na\Big) \\
&\leqslant \frac{V(X_1 + \ldots X_n)}{(na)^2} \\
&\leqslant \frac{nV(X)}{n^2a^2} \quad \text{(car les } (X_k) \text{ sont indépendantes et identiquement distribuées)} \\
&\leqslant \frac{V(X)}{na^2}
\end{aligned}$$

□

Exemple.
On effectue n lancers d'une pièce équilibrée. Quelle doit-être la valeur de n pour que la proportion de Pile obtenus soit comprise strictement entre 0,4 et 0,5 avec une probabilité supérieure à 95% (soit une marge d'erreur de 5%).

Solution :
On modélise chaque lancer par une variable aléatoire X_i définie par $X_i = 1$ si on obtient Pile et $X_i = 0$ si on obtient Face. Les varaibles X_i sont des variables de Bernoulli indépendantes et identiquement distribuées d'espérance $\mathrm{E}(X) = p = \frac{1}{2}$ et de variance $V(X) = p(1-p) = \frac{1}{4}$.
Ainsi, en appliquant l'inégalité de concentration (avec $a = 0{,}1$) on obtient :

$$\mathrm{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - \frac{1}{2}\right| \geqslant 0{,}1\right) \leqslant \frac{\frac{1}{4}}{n \times 0{,}1^2}$$

$$\mathrm{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - \frac{1}{2}\right| \geqslant 0{,}1\right) \leqslant \frac{25}{n}$$

On cherche donc n tel que $\frac{25}{n} \leqslant 0{,}05$.
On obtient $n \geqslant 500$.
Ainsi, en faisant 500 lancers, la probabilité que la proportion de Pile soit comprise strictement entre 0,49 et 0,51 est supérieure à 95%.

2.2 Loi faible des grands nombres

Proposition 5 – Loi faible des grands nombres

Soit $(X_k)_{1\leqslant k\leqslant n}$ un échantillon d'une variable aléatoire X.On note $M_n = \dfrac{X_1+\ldots+X_n}{n}$.
Alors, pour tout $a \in \mathbb{R}_+^*$,

$$\mathrm{P}\Big(|M_n - \mathrm{E}(X)| \geqslant a\Big) \underset{n\to+\infty}{\longrightarrow} 0.$$

Démonstration.
D'après l'inégalité de concentration,

$$0 \leqslant \mathrm{P}\Big(|M_n - \mathrm{E}(X)| \geqslant a\Big) \leqslant \frac{V(X)}{na^2}.$$

Or, $\lim_{n\to+\infty} \dfrac{V(X)}{na^2} = 0$.
On conclut à l'aide du théorème des gendarmes que $\lim_{n\to+\infty} \mathrm{P}\Big(|M_n - \mathrm{E}(X)| \geqslant a\Big) = 0$. □

Remarque.
Il existe une propriété appelée loi forte des grands nombres qui est néanmoins hors programme.

QCM d'entrainement

Savoir-faire du chapitre

- Appliquer les différentes inégalités du cours.
- Modéliser un problème de probabilité et interpréter les inégalités dans le contexte de l'exercice.