Le cours – T° spé : Loi des grands nombres

Prérequis : Variables aléatoires

Inégalité de Markov:

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

La probabilité que X prenne des valeurs plus grandes que a est d'autant plus petite que a est grand.

Inégalité de Bienaymé-Tchebychev :

$$P(|X-E(X)| \geqslant \delta) \leqslant \frac{V(X)}{\delta^2}$$

La probabilité que les valeurs prises par X s'écartent d'au moins δ de l'espérance $E(X)$ est d'autant plus petite que δ est grand.

Loi des grands nombres :

Soit M_n la variable aléatoire moyenne d'un échantillon de taille n d'une variable aléatoire d'espérance $E(X)$ et de variance $V(X)$. Pour tout réel strictement positif δ, on a :

$$P(|M_n-E(X)| \geqslant \delta) \leqslant \frac{V(X)}{n\delta^2}$$

(inégalité de concentration)

Et donc :

$$\lim_{n \to +\infty} P(|M_n-E(X)| \geqslant \delta) = 0$$

On dit que M_n converge en probabilité vers $E(X)$ lorsque n tend vers $+\infty$.
Ce résultat justifie la possibilité de **définir des probabilités en prenant pour valeurs approchées les fréquences** obtenues pour un grand nombre d'essais.

Les méthodes – T° spé : Loi des grands nombres

Appliquer l'inégalité de Bienaymé-Tchebychev :

On lance 3600 fois une pièce de monnaie non truquée. Soit X la variable aléatoire qui associe à cette expérience le nombre de Pile obtenus.
a) Écrire l'inégalité de Bienaymé-Tchébychev relative à la variable X.
b) Minorer la probabilité que le nombre d'apparitions de Pile soit strictement compris entre 1600 et 2000.

Solution :

a) Nous avons besoin de $E(X)$ et de $V(X)$ pour écrire l'inégalité de Bienaymé-Tchébychev.
X suit une loi binomiale de paramètres 3600 et 0,5. On a donc : $E(X) = 3600 \times 0{,}5 = 1800$ et $V(X) = 3600 \times 0{,}5 \times 0{,}5 = 900$.

On écrit alors l'inégalité de Bienaymé-Tchébychev : $P(|X-1800| \geqslant \delta) \leqslant \frac{900}{\delta^2}$ avec $\delta > 0$.

b) L'intervalle $]1600\,;2000[$ est un intervalle centré sur 1800.
$X \in]1600\,;2000[$ équivaut à $|X-1800| < 200$.
Nous cherchons donc à minorer $P(|X-1800| < 200) = 1 - P(|X-1800| \geqslant 200)$.

D'après la question 1, avec $\delta = 200$, nous avons :
$P(|X-1800| \geqslant 200) \leqslant \frac{900}{200^2}$
$P(|X-1800| \geqslant 200) \leqslant 0{,}0225$
$-P(|X-1800| \geqslant 200) \geqslant -0{,}0225$
$1-P(|X-1800| \geqslant 200) \geqslant 1-0{,}0225$
$1-P(|X-1800| \geqslant 200) \geqslant 0{,}9775$
$P(|X-1800| < 200) \geqslant 0{,}9775$
La probabilité que le nombre d'apparitions de Pile soit strictement compris entre 1600 et 2000 est au moins de 0,9775.

Utiliser l'inégalité de concentration pour définir une taille d'échantillon :

On effectue n tirages successifs, avec remise, d'une boule dans une urne contenant 2 boules rouges et 3 boules noires. On note X la variable aléatoire qui, à un tirage donné, associe 1 si la boule tirée est rouge, et 0 sinon, et M_n la variable aléatoire moyenne d'un échantillon de taille n de X.

a) Déterminer $E(X)$ et $V(X)$, puis écrire l'inégalité de concentration relative à M_n.
b) A partir de quel nombre de tirages peut-on garantir à au moins 95 % que la proportion de boules rouges obtenues restera strictement comprise entre 0,35 et 0,45 ?

Solution :

a) X est une variable de Bernoulli de paramètre $\frac{2}{5} = 0{,}4$ donc $E(X) = 0{,}4$ et $V(X) = 0{,}4 \times 0{,}6 = 0{,}24$.

L'inégalité de concentration s'écrit : $P(|M_n-0{,}4| \geqslant \delta) \leqslant \frac{0{,}24}{n\delta^2}$

b) L'intervalle $]0{,}35\,;0{,}45[$ est un intervalle centré sur 0,4.
$M_n \in]0{,}35\,;0{,}45[$ équivaut à $|M_n-0{,}4| < 0{,}05$.

Nous cherchons n tel que :
$P(|M_n-0{,}4| < 0{,}05) \geqslant 0{,}95$
$-P(|M_n-0{,}4| < 0{,}05) \leqslant -0{,}95$
$1-P(|M_n-0{,}4| < 0{,}05) \leqslant 1-0{,}95$
$1-P(|M_n-0{,}4| < 0{,}05) \leqslant 0{,}05$
$P(|M_n-0{,}4| \geqslant 0{,}05) \leqslant 0{,}05$
Il faut donc, d'après l'inégalité de concentration, que $\frac{0{,}24}{n\,0{,}05^2} \leqslant 0{,}05$
$\frac{n\,0{,}05^2}{0{,}24} \geqslant \frac{1}{0{,}05}$
$n \geqslant \frac{0{,}24}{0{,}05^3}$
$n \geqslant 1920$

Le nombre minimal de tirages doit être de 1920.