Concentration, lois des grands nombres.

I. L'inégalité de Bienaymé-Tchebychev.

1. L'inégalité de Markov.

Définition : Une variable aléatoire est dite positive ou nulle dans un univers Ω lorsque toutes les valeurs prises par celle-ci sont des réels positifs ou nuls.

Exemple :une variable aléatoire suivant une loi binomiale est une variable aléatoire positive.

Propriété : Inégalité de Markov.
Soit X une variable aléatoire positive ou nulle d'espérance E(X).
Pour tout réel *a* strictement positif, $\mathrm{P}(\mathrm{X}>a)\leqslant\frac{\mathrm{E}(\mathrm{X})}{a}$
Cette inégalité est appelée l'inégalité de Markov.

Démonstration : Soit X une variable aléatoire positive ou nulle.
Soit $x_1,..,x_n$ les valeurs prises par X.
X étant positive, pour tout $i \in \mathbb{N}$, $x_i\geqslant 0$.
Par définition de l'espérance, $\mathrm{E}(\mathrm{X})=\sum_{i=1}^{n} x_i\mathrm{P}(\mathrm{X}=x_i)$.
Soit a un réel strictement positif. Séparons les x_i en deux parties : ceux supérieurs ou égaux à a et ceux inférieurs ou égaux à a .
Nous pouvons donc séparer E(X) en deux sommes :

$$\mathrm{E}(\mathrm{X})=\sum_{x_i<a} x_i\mathrm{P}(\mathrm{X}=x_i)+\sum_{x_i\geq a} x_i\mathrm{P}(\mathrm{X}=x_i) .$$

Pour tout i , $x_i\mathrm{P}(\mathrm{X}=x_i)\geqslant 0$, donc en particulier, $\sum_{x_i<a} x_i\mathrm{P}(\mathrm{X}=x_i)\geq 0$.

Par suite, nous obtenons que $\mathrm{E}(\mathrm{X})\geq\sum_{x_i\geq a} x_i\mathrm{P}(\mathrm{X}=x_i)$.

Or $x_i\geqslant a$ => $\sum_{x_i\geq a} x_i\mathrm{P}(\mathrm{X}=x_i)\geq\sum_{x_i\geq a} a\,\mathrm{P}(\mathrm{X}=x_i)$

Par conséquent, $\mathrm{E}(\mathrm{X})\geq\sum_{x_i\geq a} a\,\mathrm{P}(\mathrm{X}=x_i)$

Or $\sum_{x_i\geq a} a\,\mathrm{P}(\mathrm{X}=x_i)=a\sum_{x_i\geq a}\mathrm{P}(\mathrm{X}=x_i)=a\,P(x\geq a)$

Ainsi, $\mathrm{E}(\mathrm{X})\geqslant a\mathrm{P}(\mathrm{X}\geqslant a)$ et donc $\mathrm{P}(\mathrm{X}\geqslant a)\leqslant\frac{\mathrm{E}(\mathrm{X})}{a}$.

Exemple 1: Une entreprise produit en moyenne 35 pièces par semaine.
On note X la variable aléatoire donnant le nombre de pièces produites par semaine.
a. Que peut-on dire de la probabilité que l'usine produise plus de 70 pièces par semaine ?
b. Que peut-on dire de la probabilité que l'usine produise moins de 5 pièces par semaine ?

2. Inégalité de Bienaymé-Tchebychev.

Théorème : Soit X une variable aléatoire d'espérance E(X) et de variance V(X).
Alors, pour tout réel a strictement positif, $P(|X-E(X)|\geqslant a)\leqslant\frac{V(x)}{a^2}$

Cette inégalité est appelée l'inégalité de Bienaymé-Tchebychev.

Démonstration : Soit $a>0$.
$|X-E(X)|\geqslant a \Leftrightarrow |X-E(X)^2|\geqslant a^2$
Donc $P(|X-E(X)|\geqslant a)=P(|X-E(X)|^2\geqslant a^2)$
De plus, nous admettrons que E(X-E(X))=V(X).
Nous allons donc appliquer l'inégalité de Markov à la variable aléatoire $(X-E(X))^2$.
Celle-ci est positive.
Nous obtenons donc $P(|X-E(X)|^2\geqslant a^2)\leqslant\frac{V(X)}{a^2}$

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

Exemple 2: Le taux moyen de glycémie dans une population est de 1 g/L avec une variance de 0,1. Une personne présente un taux X critique si son taux ne se situe pas dans l'intervalle]0,5;1,5[.
a. Quelle est la probabilité qu'une personne ne présente pas de taux critique ?
b. En déduire la probabilité qu'une personne présente un taux critique.

II. Loi des grands nombres.

1. Inégalité de concentration.

Théorème :
Soient X_1 , X_2, ..., X_n n variables aléatoires indépendantes, suivant la même loi de probabilité qu'une variable aléatoire X. Soit E(X) l'espérance de X et V(X) la variance de X.
On pose M_n la variable aléatoire moyenne d'un échantillon de taille n. On a donc $M_n=\frac{1}{n}\sum_{i=1}^{n}X_i$ **.**
Pour tout réel a strictement positif, $P(|M_n-E(X)|\geqslant a)\leqslant\frac{V(X)}{na^2}$

Démonstration :
Nous allons appliquer l'inégalité de Bienaymé-Tchebychev à la variable aléatoire M_n .
Cela nous donne :
$P(|M_n-E(M_n)|\geqslant a)\leqslant\frac{V(M_n)}{a^2}$ (I)
Déterminons $E(M_n)$.

$$E(M_n)=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X)=\frac{1}{n}\times nE(X)=E(X)$$

Déterminons $V(M_n)$.

$V(\mathrm{M}_n)=V(\frac{1}{n}\sum_{\mathrm{i}=1}^{n}\mathrm{X}_\mathrm{i})=\frac{1}{n^2}V(\sum_{\mathrm{i}=1}^{n}\mathrm{X}_\mathrm{i})$.

Or X_1 , X_2 , ... X_n étant indépendantes, $V(\sum_{\mathrm{i}=1}^{n}\mathrm{X}_\mathrm{i})=\sum_{\mathrm{i}=1}^{n}V(\mathrm{X}_\mathrm{i})=nV(X)$

Par conséquent, (I) devient : $\mathrm{P}(|\mathrm{M}_n-\mathrm{E}(X)|\geqslant a)\leqslant\frac{V(X)}{na^2}$

Exemple 3 : Soit X une variable aléatoire suivant une loi de binomiale de paramètre n=0,02 et p=10.
a. Déterminer E(X) et V(X).
b. On considère un échantillon de n variables aléatoires suivant la loi de X.
On appelle M_n la variable aléatoire moyenne associée à cet échantillon.
Déterminer la taille n de l'échantillon afin que la probabilité que la moyenne M_n appartienne à l'intervalle]0,03,0,37[soit supérieure à 0,95.

b. Loi (faible) des grands nombres.

Propriété : Soit $\mathrm{X}_{1,}\mathrm{X}_{2,\ldots,}\mathrm{X}_n$ un échantillon de taille n d'une variable aléatoire X.
On pose $\mathrm{M}_n=\frac{1}{n}\sum_{\mathrm{i}=1}^{n}\mathrm{X}_\mathrm{i}$.
Alors, pour tout réel a strictement positif, $\lim_{n\to+\infty}\mathrm{P}(\mathrm{M}_n-\mathrm{E}(\mathrm{X})\geqslant a)=0$

Démonstration.
Appliquons l'inégalité de concentration à la variable aléatoire M_n .
On sait que pour tout réel a , strictement positif, $0\leqslant\mathrm{P}(|\mathrm{M}_n-\mathrm{E}(\mathrm{X})|\geqslant a)\leqslant\frac{\mathrm{V}(\mathrm{X})}{na^2}$.
Or $\lim_{n\to+\infty}\frac{\mathrm{V}(\mathrm{X})}{na^2}=0$, donc, d'après le théorème des gendarmes, on a :

$$\lim_{n\to+\infty}\mathrm{P}(\mathrm{M}_n-\mathrm{E}(\mathrm{X})\geqslant a)=0$$

Remarque : La loi des grands nombres traduit le fait que plus la taille de l'échantillon d'une variable aléatoire X est grande, plus l'écart entre la moyenne de cet échantillon et l'espérance de la variable aléatoire X est faible.