Concentration et loi des grands nombres

I) Inégalité de Bienaymé-Tchebychev

1) Inégalité et interprétation

X est une variable aléatoire d'espérance μ et de variance V.
Pour tout réel $\delta > 0$, $P(|X - \mu| \geq \delta) \leq \frac{V}{\delta^2}$

Remarques:

Pour $\delta \leq \sigma$ où σ est l'écart type de X, $\frac{V}{\delta^2} \geq 1$ et l'inégalité est évidente.

Exemple et interprétation :

On lance une pièce équilibrée 100 fois de suite, soit X la variable aléatoire qui donne le nombre de Pile obtenus.

X suit la loi binomiale de paramètres $n = 100$ et $p = \frac{1}{2}$

L'espérance de X est $\mu = 100 \times \frac{1}{2} = 50$ et sa variance est V $= 100 \times \frac{1}{2} \times \frac{1}{2} = \frac{100}{4} = 25$

D'après l'inégalité de Bienaymé-Tchebychev :

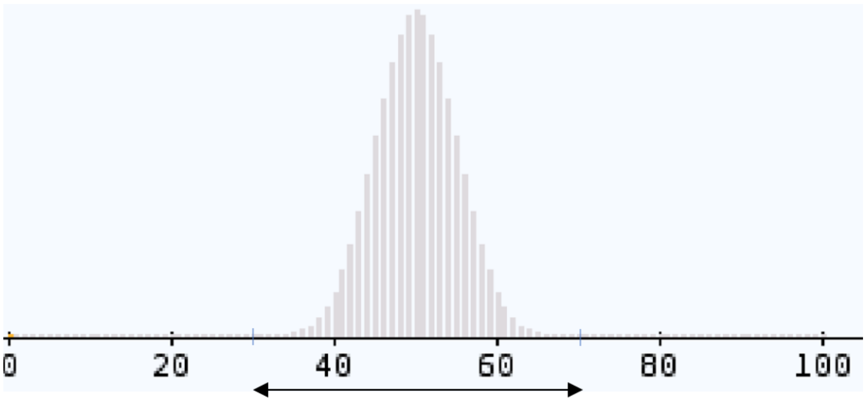
Pour tout réel $\delta > 0$, $P(|X - 50| \geq \delta) \leq \frac{25}{\delta^2}$

Prenons par exemple $\delta = 20$ alors l'inégalité devient : $P(|X - 50| \geq 20) \leq \frac{25}{20^2}$

La probabilité pour que l'écart de X à $\mu = 50$ soit supérieure ou égale à 20 est inférieur ou égale à 0,0625.

Or, $P(|X - 50| < 20) = 1 - P(|X - 50| \geq 20)$ donc $P(|X - 50| < 20) \geq 0{,}9375$ $\quad (1 - 0{,}0625 = 0{,}9375)$

La probabilité que X prenne une valeur dans l'intervalle $]50 - 20\,;\,50 + 20[$ est supérieure à $0{,}9375$c'est-à-dire la probabilité que X prenne une valeur dans l'intervalle $]30\,;\,70[$ est supérieure à $\mathbf{0{,}9375}$



[30 ;70] est un intervalle de fluctuation de X au seuil de 93,75%.

L'inégalité de Bienaymé-Tchebychev explicite le rôle de la variance comme indicateur de dispersion.

2) Cas particuliers

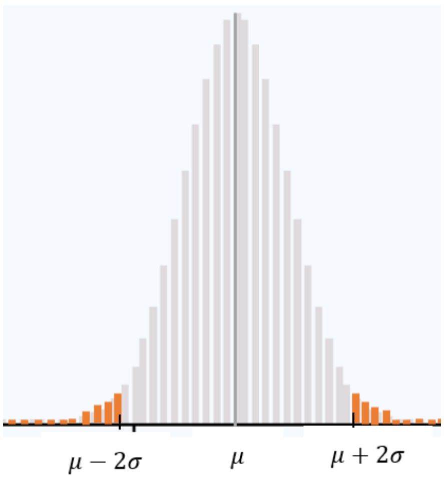
Soit X une variable aléatoire d'espérance μ et d'écart type σ.
L'inégalité de Bienaymé-Tchebychev avec $\delta = k\sigma$ où k est un nombre entier naturel $k \geq 2$ devient $P(|X-\mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2}$ soit

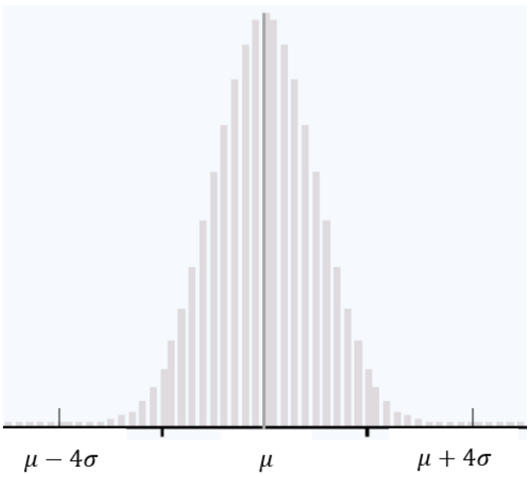
$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

Par exemple :

$k = 2$
$P(|X-\mu| \geq 2\sigma) \leq \frac{1}{2^2}$
$P(|X-\mu| \geq 2\sigma) \leq \frac{1}{2^2}$
$P(|X-\mu| \geq 2\sigma) \leq 0{,}25$
La probabilité que l'écart de X à μ soit supérieur ou égale à 2σ est inférieure ou égale à 0,25

$k = 4$
$P(|X-\mu| \geq 4\sigma) \leq \frac{1}{4^2}$
$P(|X-\mu| \geq 4\sigma) \leq \frac{1}{4^2}$
$P(|X-\mu| \geq 4\sigma) \leq 0{,}0625$
La probabilité que l'écart de X à μ soit supérieur ou égale à 4σ un évènement quasi-improbable puisque c'est une probabilité inférieure ou égale à 0,0625





II) Loi des grands nombres

1) Variable aléatoire moyenne

Soit $(X_1, X_2, \ldots, X_n)$ un échantillon de taille n de variables aléatoires indépendantes, suivant une même loi de probabilité.

La variable aléatoire moyenne est : $M_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$

2) Inégalité de concentration

Pour tout réel $\delta > 0$, $P(|M_n - \mu| \geq \delta) \leq \frac{V}{n\delta^2}$

Démonstration : L'espérance de M_n est E$(M_n) = \frac{1}{n} \times n \times \mu = \mu$ et sa variance est V$(M_n) = \frac{1}{n^2} \times n \times V = \frac{V}{n}$. D'après l'inégalité de Bienaymé-Tchebychev, pour tout réel $\delta > 0$, $P(|M_n - \mu| \geq \delta) \leq \frac{V(M_n)}{\delta^2}$ or V$(M_n) = \frac{V}{n}$ donc $P(|M_n - \mu| \geq \delta) \leq \frac{V}{n\delta^2}$

Exemple :

Dans une ville moyenne de 20 000 habitants, lors d'une consultation portant sur la rénovation du théâtre municipal, 75% des personnes consultées ont émis un avis positif.

1. On interroge n personnes. Pour $1 \leq k \leq n$, la variable aléatoire X_k donne 1 si la k-ième personne interrogée est favorable au projet et 0 sinon.
Donner la loi de probabilité de X_k, son espérance μ et sa variance V.

2. On note M_n la variable aléatoire moyenne de X_1, X_2,….., X_n

Déterminer une taille n d'échantillon afin d'obtenir pour M_n une précision de 0,05 et un risque de 0,1, c'est-à-dire $P(|M_n - \mu| \geq 0{,}05) \leq 0{,}1$

Réponse :
1. La loi de probabilité de la variable X_k est donnée par :

a	0	1
P$(X_k = a)$	0,25	0,75

Son espérance est : $\mu = 0 \times 0{,}25 + 1 \times 0{,}75 = 0{,}75$
Sa variance est : $V = 0{,}25 \times (0 - 0{,}75)^2 + 0{,}75 \times (1 - O,75)^2 = 0{,}1875$

2. D'après l'inégalité de concentration, pour tout réel $\delta > 0$

$P(|M_n - 0{,}75| \geq \delta°) \leq \frac{0{,}1875}{n\delta^2}$ Comme $\delta = 0{,}05$ on obtient

$P(|M_n - 0{,}75| \geq 0{,}01) \leq \frac{75}{n}$ il suffit de choisir n telle que $\frac{75}{n} \leq 0{,}1$ soit $n \geq \frac{75}{0{,}1}$ soit $n \geq 750$

L'échantillon doit être composé de 750 personnes.

3) Loi des grands nombres

Pour tout réel $t > 0$, $\lim\limits_{n \to +\infty} P(|M_n - \mu| \geq t) = 0$

Démonstration :

Pour tout réel $t > 0$ donné, l'inégalité de concentration, permet d'écrire :
$0 \leq P(|M_n - \mu| \geq t) \leq \frac{V}{nt^2}$
Or $\lim\limits_{n \to +\infty} \frac{V}{nt^2} = 0$ donc d'après le théorème des gendarmes $\lim\limits_{n \to +\infty} P(|M_n - \mu| \geq t) = 0$