Probabilités
Inégalités et grands nombres

Terminale générale
spécialité maths

Table des matières

I - Somme de variables aléatoires

1) Espérance

Propriété *Linéarité de l'espérance*

Soit X et Y deux variables aléatoires, et a un réel, alors

$$E(X+Y) = E(X) + E(Y) \quad \text{et} \quad E(aX) = aE(X)$$

Exercice 1 On joue au jeu suivant :
— on lance un dé équilibré à 6 faces numérotées de 1 à 6 et on gagne autant de points que le chiffre obtenu
— on lance ensuite une pièce équilibrée : si on obtient pile on gagne 4 points, si c'est face on perd 8 points

1. Soit X la variable aléatoire égale au nombre total de points obtenus. Calculer $E(X)$. Ce jeu est-il équitable ?
2. On multiplie par 2 tous les points obtenus, ou perdus, dans les deux étapes. Quelle est l'espérance avec cette nouvelle règle ?

Exercice 2 On lance 2 dés équilibrés à 6 faces numérotées de 1 à 6. Le score obtenu est la somme des chiffres sur les 2 dés. On note X la variable aléatoire égale à ce score.
Donner la loi de probabilité de X puis son espérance.

2) Variance

Propriété *Soit X une variable aléatoire et a un réel alors*

$$V(aX) = a^2 V(X) \quad \text{et donc} \quad \sigma(aX) = |a|\,\sigma(X)$$

Propriété *Soit X et Y deux variables aléatoires* ***indépendantes****, alors*

$$V(X+Y) = V(X) + V(Y)$$

Exercice 3 Calculer les variances dans les exercices précédents.

Exercice 4 On lance 5 dés équilibrés. On note X_k la variable aléatoire égale au résultat du k-ième dé et X la variable aléatoire égale à la somme des résultats obtenus.
Calculer $E(X_k)$ et $V(X_k)$ pour tout $1 \leqslant k \leqslant 5$, et en déduire $E(X)$, $V(X)$ et $\sigma(X)$.
Solution : Pour tout k, on a $E(X_k) = 1 \times \dfrac{1}{6} + 2 \times \dfrac{1}{6} + \cdots + 6 \times \dfrac{1}{6} = \dfrac{21}{6} = 3,5$
et $V(X_k) = (1-3,5)^2 \times \dfrac{1}{6} + (2-3,5)^2 \times \dfrac{1}{6} + \cdots + (6-3,5)^2 \times \dfrac{1}{6} = \dfrac{35}{12} \simeq 2,92$.

On a alors $X = X_1 + X_2 + X_3 + X_4 + X_5 = \displaystyle\sum_{k=1}^{5} X_k$, d'où $E(X) = 5E(X_k) = 17,5$ et, comme les X_k sont indépendants $V(X) = 5V(X_k) = \dfrac{175}{12} \simeq 14,6$. Enfin, $\sigma(X) = \sqrt{V(X)} \simeq 3,82$.

Exercice 5 On lance n fois une pièce équilibrée et on joue au jeu suivant : si on obtient pile au $k^{\text{ème}}$ lancer, on gagne k euros, sinon on ne gagne, ni ne perd, rien.
On note X_k la variable aléatoire correspondant au gain obtenu lors du $k^{\text{ème}}$ tirage et Y_n le total des gains obtenus à l'issue de n lancers.

1. Donner la loi de probabilité de X_k.
   En déduire $E(X_k)$ puis montrer que $V(X_k) = \dfrac{k^2}{4}$.

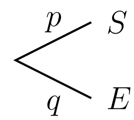
2. a) Exprimer Y_n en fonction des variables aléatoires X_k.
   b) Déterminer l'espérance de Y_n.
   c) Déterminer le nombre théorique de lancers nécessaires pour que le gain total dépasse en moyenne 280 euros.
   d) Montrer, par récurrence, que pour $n \geqslant 1$ on a $\displaystyle\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

      En déduire $V(Y_n)$.

3) Application à la loi binomiale

La loi binomiale est la loi de probabilité d'une variable aléatoire X égale au nombre de succès dans un schéma de Bernoulli : la répétition de n épreuves, identiques et indépendantes, de Bernoulli.

Si on note X_k la variable aléatoire associée à chacune de ces épreuves,



dont la loi de probabilité est

x_i	1	0
$P\left(X_k = x_i\right)$	p	$1-p$

on a alors $E(X_k) = p$ et $V(X_k) = pq = p(1-p)$.

On a aussi, $X = \displaystyle\sum_{k=1}^{n} X_k = X_1 + X_2 + \cdots + X_n$ et donc,

Corollaire $E(X) = \displaystyle\sum_{k=1}^{n} E(X_k) = np$

et, par indépendance des X_k, $V(X) = \displaystyle\sum_{k=1}^{n} V(X_k) = npq$, *d'où* $\sigma(X) = \sqrt{npq}$.

Exemple : X suit la loi $\mathcal{B}(30; 0,2)$,
alors $E(X) = 30 \times 0,2 = 6$
et $V(X) = 30 \times 0,2 \times (1-0,2) = 4,8$ et donc $\sigma(X) = \sqrt{V(X)} = \sqrt{4,8} \simeq 2,19$.

4) Estimation de la moyenne

On considère n variables aléatoires indépendantes et suivant la même loi de probabilité que X.
On note la somme $S_n = X_1 + X_2 + \cdots + X_n$ et la moyenne $M_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$.

Propriété
— *$E(S_n) = nE(X)$ et $V(S_n) = nV(X)$ donc $\sigma(S_n) = \sqrt{n}\sigma(X)$.*
— *$E(M_n) = E(X)$ et $V(M_n) = \dfrac{V(X)}{n}$ donc $\sigma(M_n) = \dfrac{\sigma(X)}{\sqrt{n}}$.*

Exercice 6 Un appareil de mesure permet de réaliser des mesures avec une certaine précision.
Plus précisément, les mesures données par cet appareil peuvent être modélisée par une variable aléatoire X dont l'espérance est la valeur exacte $E(X) = m$ et l'écart type $\sigma(X) = 1$.
On suppose les mesures réalisées avec cet appareil indépendantes les unes des autres.

1. On réalise n mesures. Que peut-on dire de l'espérance de la moyenne de ces mesures ?
2. On sait que la probabilité de l'événement $M_n \in [m - 2\sigma\,;\, m + 2\sigma]$ est d'environ 95%.

   a) Quelle est la précision de la mesure avec 10 mesures ?

b) Combien faut-il faire de mesures pour que la moyenne de celles-ci donne la valeur exacte avec une précision d'au moins 10^{-2} ?

Répéter un grand nombre de fois une mesure, et utiliser la moyenne des résultats obtenus donne tout simplement une estimation de la valeur recherchée.
En augmentant le nombre de mesures, on réduit ***l'incertitude*** *sur la valeur trouvée.*
La ***loi des grands nombres*** *permet aussi de préciser cela.*

II - Inégalités de Markov et de Bienaymé-Tchebychev

1) Inégalité de Markov

Théorème *Soit $a > 0$ un réel et X une variable aléatoire réelle positive ou nulle, alors, on a l'**inégalité de Markov***

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

<u>Démonstration:</u> On sépare dans la loi de probabilité de X les valeurs inférieures ou supérieures ou égales à a : soit k le plus petit entier tel que $x_k \geqslant a$

a

x_i	x_1	x_2	$\dots$	x_k	x_{k+1}	$\dots$	x_n
$P\left(X = x_i\right)$	$\dots$	$\dots$					

On a alors :

$$\begin{aligned}
E(X) &= \sum_{i=1}^{n} x_i P\left(X = x_i\right) \\
&= \underbrace{\sum_{i=1}^{k-1} x_i P\left(X = x_i\right)}_{>0} + \sum_{i=k}^{n} \underbrace{x_i}_{\geqslant a} P\left(X = x_i\right) \\
&\geqslant 0 + a \underbrace{\sum_{i=k}^{n} P\left(X = x_i\right)}_{P(X \geqslant a)}
\end{aligned}$$

D'où l'inégalité de Markov en divisant par $a > 0$. □

Exercice 7 Majorer la probabilité $P(X \geqslant 1)$ d'une variable aléatoire X d'espérance $E(X) = 0,5$.

Exercice 8 Un phénomène naturel se produit, au hasard semble-t-il, en moyenne tous les 75 ans.
Majorer la probabilité qu'il se produise dans plus de 300 ans ?
Que peut-on dire de la probabilité qu'il se produise dans moins de 300 ans ?

Exercice 9 Une usine produit en moyenne 100 pièces par semaine. Que peut-on dire de la probabilité qu'une semaine elle en produise le double ?

Exercice 10 Montrer que, pour une variable aléatoire positive X et un entier $k > 0$ quelconque, on a $P\left(X \geqslant kE(X)\right) \leqslant \frac{1}{k}$.
En déduire la loi : "moins de 10% des salariés gagnent plus de 10 fois le salaire moyen".

Exercice 11 La température moyenne sur une île est de 25 degrés.

a) Majorer la probabilité que la température soit supérieure à 40 degrés un jour donné.

b) Minorer la probabilité que la température soit inférieure à 30 degrés un jour donné.

Exercice 12 Sur mon compte j'ai en moyenne 400 euros. Ma banque ne m'autorise aucun découvert.
Minorer la probabilité que je finisse un mois avec moins de 500 euros sur mon compte.

2) Inégalité de Bienaymé-Tchebychev

Théorème *Pour X une variable aléatoire d'espérance $E(X)$ et de variance $V(X)$, et pour tout réel $a > 0$, on a l'**inégalité de Bienaymé-Tchebychev***

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

<u>Démonstration:</u> On applique l'inégalité de Markov à la variable aléatoire $|X - E(X)|^2$, qui est bien positive, et avec le réel $a^2 > 0$:

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

qui est l'inégalité voulue car :

— $\big|X - E(X)\big|^2 \geqslant a^2 \iff \big|X - E(X)\big| \geqslant a$,
d'où $P\Big(\big|X - E(X)\big|^2 \geqslant a^2\Big) \iff P\Big(\big|X - E(X)\big| \geqslant a\Big)$,

— $V(X) = E\Big(\big|X - E(X)\big|^2\Big)$ *(moyenne des carrés des écarts à la moyenne)*

□

Exercice 13 Soit X la variable aléatoire dont la loi de probabilité est donnée ci-dessous.

1. Calculer l'espérance et la variance de X.
2. Calculer $P(|X - E(X)| \geqslant 2)$.
3. Appliquer l'inégalité de Bienaymé-Tchebychev pour $a = 2$ et comparer avec le résultat précédent.

x_i	1	4	10
$P(X = x_i)$	0,6	0,3	0,1

Exercice 14 Dans un grand magasin, on a relevé que le temps d'attente moyen des clients est de 12 minutes. La probabilité qu'un client attende entre 9 et 15 minutes est de 0,55.
Que peut-on dire de l'écart type du temps d'attente ?

Exercice 15 Une équipe de rugby a marqué 60 essais sur les 20 derniers matchs. On considère que le nombre d'essais marqués à chaque match par l'équipe est une variable aléatoire X.

1. Que vaut l'espérance de X ?
   Majorer la probabilité que l'équipe marque plus de 5 essais au prochain match.
2. On a estimé, statistiquement, que la variance de X est $0,6$.
   a) Majorer la probabilité qu'au cours du prochain match, l'écart entre le nombre d'essais marqués et la moyenne soit supérieur ou égal à 1.
   En déduire une information sur la probabilité que l'équipe marque exactement 3 essais au prochain match.
   b) Minorer la probabilité que l'équipe marque 2, 3 ou 4 essais.

III - Loi des grands nombres

1) Inégalité de concentration

Théorème *Soit X une variable aléatoire d'espérance $E(X)$ et de variance $V(X)$, et M_n la variable aléatoire moyenne d'un échantillon de taille n de X.*
On a alors, pour tout réel $a > 0$,

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

*Cette inégalité s'appelle **inégalité de concentration**.*

Exemple : On lance n fois successivement une pièce équilibrée. À chaque lancer, on note X la variable aléatoire égale à 0 si on obtient face et 1 si on obtient pile, alors $E(X) = 1/2$ et $V(X) = 1/4$.
Pour tout $a > 0$, on a alors

$$P\left(\left|M_n - E(X)\right| \geqslant a\right) \leqslant \frac{V(X)}{na^2}$$
$$\iff \quad P\left(\left|M_n - \frac{1}{2}\right| \geqslant a\right) \leqslant \frac{1}{4na^2}$$

Par exemple,

— pour $a = 0,1$ et $n = 100$,

$$P\left(\left|M_n - \frac{1}{2}\right| \geqslant 0,1\right) \leqslant \frac{1}{4}$$

en d'autres termes, sur $n = 100$ lancers, la probabilité que la proportion de pile obtenue s'écarte de plus d'un dixième de 1/2, donc d'avoir moins de 40 ou plus de 60 pile, est inférieure à 1/4.

— pour $a = 0,01$ et $n = 10\,000$,

$$P\left(\left|M_n - \frac{1}{2}\right| \geqslant 0,01\right) \leqslant \frac{1}{4}$$

en d'autres termes, sur $n = 10\,000$ lancers, la probabilité que la proportion de pile obtenue s'écarte de plus d'un centième de 1/2, donc d'avoir moins de 4900 ou plus de 5100 pile, est inférieure à 1/4.

— pour $a = 0,01$ et $n = 100\,000$,

$$P\left(\left|M_n - \frac{1}{2}\right| \geqslant 0,01\right) \leqslant \frac{1}{40}$$

en d'autres termes, sur $n = 100\,000$ lancers, la probabilité que la proportion de pile obtenue s'écarte de plus d'un centième de 1/2, donc d'avoir moins de 49 000 ou plus de 51 000 pile, est inférieure à 1/40.

Exercice 16 On lance n fois un dé équilibré et on note le nombre de fois où on obtient 1.
Combien de fois, au minimum, faut-il lancer ce dé pour que la probabilité de s'écarter de la moyenne de plus de 0,1 soit inférieure à 5% ?

On peut simuler numériquement cette expérience pour observer ce résutat, voir cette page.

Exercice 17 On reprend l'exercice 6 sur les mesures physiques. On rappelle qu'avec l'appareil utilisé, une mesure est modélisée par une variable aléatoire X d'espérance m, qui est aussi la valeur exacte recherchée (et inconnue donc), et d'écart type $\sigma = 1$.

1. On réalise $n = 10$ mesures. Majorer la probabilité que la moyenne des mesures réalisées s'écarte de plus de $0,1$ de la valeur exacte recherchée.
2. Combien de mesures faut-il réaliser pour que la probabilité que l'imprécision sur la mesure soit inférieure à 10^{-2} soit supérieure à 99% ?

2) Loi des grands nombres

En passant à la limite $n \to +\infty$ dans l'inégalité de concentration, comme une probabilité est positive et grâce au théorème des gendarmes, on obtient

Corollaire ***Loi faible des grands nombres***

Soit (X_n) un échantillon d'une variable aléatoire X et sa moyenne $M_n = \dfrac{1}{n}\sum_{k=1}^{n} X_k$, alors, pour tout réel $a > 0$,

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

Pour une illustration numérique simulée de cette propriété, voir cette page.

Démonstration: Il s'agit d'une conséquence (un corollaire donc !) de l'inégalité de concentration précédente :

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

comme cette probabilité est (comme toutes les probabilités) positives et que

$$\lim_{n\to+\infty} \frac{V(X)}{na^2} = 0$$

le théorème des gendarmes donne alors la limite énoncée dans le corollaire. □

Cette loi des grands nombres est largement utilisée en pratique. Elle nous dit qu'on peut estimer la probabilité d'un événement aléatoire à partir de la moyenne d'un (grand) nombre de réalisations.
*En théorie ce **grand** nombre signifie **la limite** lorsque le nombre de réalisations tend vers l'**infini**.*
*En pratique, on prend un (grand) nombre fini. Comment préciser si un nombre n est **grand ??***
*Cela dépend du contexte et de la précision souhaitée, et cela est alors donné par l'**inégalité de concentration** qui est un résultat plus précis.*