

Espérance et variance

I	Espérance	2
	Lois usuelles	
	Propriétés	
	Formule de transfert	
	Espérance d'un produit de variables indépendantes	
II	Variance	6
	Propriétés	
	Lois usuelles	
	Écart type	
III	Covariance — Variance d'une somme	8
	Définition	
	Propriétés de la covariance	
	Variance d'une somme de variables aléatoires	
	Variables aléatoires décorrélées	
IV	Inégalités probabilistes	11



Ici, les VA sont à valeurs dans $\mathbb{K}$ où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

I. Espérance

1

Définition.

Soit X une variable aléatoire définie sur $(\Omega, \mathbb{P})$ à valeurs dans $\mathbb{K}$.

L'espérance de X est le scalaire :

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x \mathbb{P}(X = x) \quad \text{Cette somme est finie}$$

On dispose d'une **formule alternative** pour l'espérance :

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\}) \quad \text{Cette somme est finie}$$

Lorsque $\mathbb{E}(X) = 0$, on dit que la variable aléatoire X est **centrée**.

- Si X est une variable aléatoire **réelle**, alors $\mathbb{E}(X) \in \mathbb{R}$.
- Lorsque $X(\Omega) \subset \{a, b, c\} \subset \mathbb{K}$, on a $\mathbb{E}(X) = \dots\dots\dots$
- Lorsque $X(\Omega) \subset \{x_1, x_2, \dots, x_m\} \subset \mathbb{K}$, on a $\mathbb{E}(X) = \dots\dots\dots$
- Lorsque $X(\Omega) \subset \llbracket 0, n \rrbracket$, on a $\mathbb{E}(X) = \dots\dots\dots$
- L'espérance d'une variable aléatoire ne dépend que de sa loi :
deux variables aléatoires de même loi ont même espérance.
- L'espérance est la moyenne des valeurs prises par X , pondérée par les probabilités que X prenne ces valeurs.
On peut aussi dire que c'est la valeur que l'on attend en moyenne quand on réalise l'expérience aléatoire. Si l'expérience consiste par exemple à tirer des boules dans une urne, on pourra se voir demander : quel est le nombre moyen de boules rouges tirées? La réponse attendue est l'espérance de la variable aléatoire qui représente le nombre de boules rouges tirées.
- L'espérance n'est pas nécessairement une valeur de $X(\Omega)$.
(votre moyenne en maths n'est pas nécessairement égale à une note de pâle).
En revanche, $\mathbb{E}(X)$ est dans l'intervalle $[\min X(\Omega), \max X(\Omega)]$.
(votre moyenne en maths est comprise entre votre plus basse note et votre plus haute note de pâle).
- **Exemple.** On lance deux dés. Soit X la variable aléatoire égale à la somme des chiffres apparus.

$$\text{On a } \mathbb{P}(X = k) = \begin{cases} \frac{k-1}{36} & \text{si } k \in \llbracket 1, 6 \rrbracket \\ \frac{13-k}{36} & \text{si } k \in \llbracket 7, 13 \rrbracket \end{cases}$$

D'où

$$\mathbb{E}(X) = \sum_{k=1}^6 k \frac{k-1}{36} + \sum_{k=7}^{13} k \frac{13-k}{36} \stackrel{\text{WHY}}{=} 7$$

2

Question. Soit X une VA à valeurs dans $\llbracket 0, n \rrbracket$.

Montrer que $\mathbb{E}(X) = \sum_{k=1}^n \mathbb{P}(X \geq k)$.



3 Proposition (espérance et variance des lois usuelles).

— **Uniforme.**

Soit $X \sim \mathcal{U}([1, n])$ où $n \in \mathbb{N}^*$.

$$\mathbb{E}(X) = \frac{n+1}{2}$$

Soit $X \sim \mathcal{U}([a, b])$ où $a \leq b \in \mathbb{Z}$.

$$\mathbb{E}(X) = \frac{a+b}{2}$$

— **Bernoulli.**

Soit $X \sim \mathcal{B}(p)$ où $p \in [0, 1]$.

$$\mathbb{E}(X) = p$$

— **Binomiale.**

Soit $X \sim \mathcal{B}(n, p)$ où $n \in \mathbb{N}^*$ et $p \in [0, 1]$.

$$\mathbb{E}(X) = np$$

- Soit $A \in \mathcal{P}(\Omega)$, on a $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$.
- **Petit moyen mnémotechnique.** On a $\mathcal{B}(1, p) = \mathcal{B}(p)$ et $\mathcal{U}([0, 1]) = \mathcal{B}(\frac{1}{2})$
Autrement dit, on peut vérifier la cohérence des résultats.

4

Proposition (propriétés de l'espérance).

Soit X et Y deux variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathbb{P})$ à valeurs dans $\mathbb{K}$.

— **Linéarité.** Soit $\lambda, \mu \in \mathbb{K}$.

On a

$$\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$$

— **Transformation affine.** Soit $a, b \in \mathbb{K}$.

On a

$$\mathbb{E}(aX + b) = a \mathbb{E}(X) + b$$

— **Inégalité triangulaire.** On a l'inégalité de nombres réels (positifs) :

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$$

- **Remarque.** Soit X une VA. Alors la VA $X - \mathbb{E}(X)$ est centrée (WHY?).
- **Utilité.** La linéarité de l'espérance est très utile pour calculer l'espérance d'une variable dont on ne connaît pas la loi ou dont la loi a une expression compliquée, mais qu'on peut décomposer en somme de variables aléatoires dont on sait calculer l'espérance, qui sont le plus souvent des variables indicatrices.
- **Vocabulaire.** Soit X une VA à valeurs dans E . Soit $a \in E$.
On dit que X est *presque sûrement* égale à $a \in E$ lorsque $\mathbb{P}(X = a) = 1$. On écrit $X \stackrel{\text{p.s.}}{=} a$.
Par exemple, une VA réelle est presque sûrement nulle lorsque $\mathbb{P}(X = 0) = 1$.

5

Proposition (pour une variable aléatoire réelle).

— **Positivité.** Soit X une variable aléatoire réelle.

On a

$$X \succcurlyeq 0 \implies \mathbb{E}(X) \geq 0$$

Si X est positive, alors son espérance est un réel positif.

— **Cas d'égalité.**

On a

$$\begin{cases} X \succcurlyeq 0 \\ \mathbb{E}(X) = 0 \end{cases} \implies X \stackrel{\text{p.s.}}{=} 0$$

On dit, dans ce cas, que la variable X est *presque sûrement nulle*.

Une VA positive d'espérance nulle est presque sûrement nulle.

— **Croissance.** Soit X et Y deux variables aléatoires réelles sur le même espace probabilisé $(\Omega, \mathbb{P})$.

On a

$$X \preccurlyeq Y \implies \mathbb{E}(X) \leq \mathbb{E}(Y)$$

- **Rappel.** La notation de Madame Tête $X \succcurlyeq 0$ signifie

$$\forall \omega \in \Omega, \quad X(\omega) \geq 0$$

- **Remarque pour la Spé.** Si $\mathbb{E}(X) = 0$, on ne peut pas assurer que $X = 0$ (c'est-à-dire que X est constante égale à 0), car un événement peut être différent de $\emptyset$ et avoir une probabilité nulle. Il faudra attendre le contexte de la classe de Spé pour voir un tel exemple (penser au lancer infini d'une pièce, et considérer l'év^t « obtenir que des PILE »).



6
sol → 13

Question. Une urne contient n boules numérotées de 1 à n . On les extrait successivement sans remise. On dit qu'il y a rencontre au $i^{\text{ème}}$ tirage si la boule tirée porte le numéro i . Déterminer le nombre moyen de rencontres.

Formule de transfert

7
preuve

Théorème de transfert.

Soit X une variable aléatoire définie sur $(\Omega, \mathbb{P})$.

Soit g une fonction définie sur $X(\Omega)$ à valeurs dans $\mathbb{K}$.

La variable aléatoire $g(X)$ admet une espérance qui est

$$\mathbb{E}(g(X)) = \sum_{x \in X(\Omega)} g(x) \mathbb{P}(X = x)$$

- **Cas particuliers.** Pour $g : t \mapsto t$, quelle formule trouve-t-on? Et pour $g : t \mapsto at + b$?
- **Utilité.** De manière générale, pour une VA Y , si on veut déterminer son espérance, il nous faut la loi de Y , c'est-à-dire les $\mathbb{P}(Y = y)$ pour tout y .

Le théorème de transfert dit que l'on peut déterminer l'espérance de Y sans connaître la loi de Y .

Voilà ce qu'il faut retenir du théorème de transfert :

« Soit X une VA avec sa loi et g une fonction.

Alors on connaît l'espérance de $Y = g(X)$ sans nécessairement connaître sa loi. »

- La formule s'applique en particulier quand X est un couple ou un n -uplet de variables aléatoires.

8

Question. Soit $X \hookrightarrow \mathcal{B}(n, p)$. On pose $Y = \frac{1}{X+1}$. Déterminer $\mathbb{E}(Y)$.

Espérance d'un produit de variables indépendantes

9
preuve

Théorème. Soit X et Y deux variables aléatoires définies sur $(\Omega, \mathbb{P})$ à valeurs dans $\mathbb{K}$.

Si X et Y sont indépendantes, alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

- **Refrain!** Sous couvert d'indépendance, l'espérance d'un produit est le produit des espérances.
- **Généralisation.** Évidemment, on a le même résultat pour n variables mutuellement indépendantes.



II. Variance

On considère ici des VA à valeurs réelles.

- **Définition hors-programme, mais usuelle.**

Le moment d'ordre $r \in \mathbb{N}^*$ de X est l'espérance de la variable aléatoire X^r .

Autrement dit :

$$\text{moment}_r(X) \stackrel{\text{déf}}{=} \mathbb{E}(X^r) \stackrel{\text{transfert}}{=} \sum_{x \in X(\Omega)} x^r \mathbb{P}(X = x)$$

En particulier, le moment d'ordre 1 est l'espérance !

10

Définition. Soit X une variable aléatoire réelle définie sur $(\Omega, \mathbb{P})$.

La variance de X est le moment d'ordre 2 de la variable aléatoire *centrée* $X - \mathbb{E}(X)$:

$$\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$$

La variance, c'est la moyenne des carrés des écarts à la moyenne

Lorsque $\mathbb{V}(X) = 0$, on dit que la variable aléatoire X est *réduite*.

- **Refrain.** La variance est le moment *centré* d'ordre 2.
- L'expression $(X - \mathbb{E}(X))^2$ est bien une variable aléatoire, car $\mathbb{E}(X)$ désigne ici la variable aléatoire constante égale à $\mathbb{E}(X)$.
- La variance est l'espérance du carré de la distance entre les valeurs de X et l'espérance de X . La variance mesure donc la **dispersion** de X autour de $\mathbb{E}(X)$.
- On aurait pu songer à prendre $\mathbb{E}(|X - \mathbb{E}(X)|)$ comme paramètre de dispersion, mais la valeur absolue donne des calculs beaucoup moins agréables que le carré.
- Lorsque l'on calcule une variance par un calcul direct en connaissant la loi, on utilise presque toujours la formule de Kœnig-Huygens et non la définition qui entraîne des calculs compliqués.

Propriétés

11

Formule de Kœnig-Huygens. Soit X une variable aléatoire réelle définie sur $(\Omega, \mathbb{P})$.

On a la formule suivante :

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

La variance, c'est la moyenne des carrés moins le carré de la moyenne

et ses cousines germaines :

$$\mathbb{V}(X) = \mathbb{E}(X(X - 1)) + \mathbb{E}(X) - \mathbb{E}(X)^2$$

$$\mathbb{V}(X) = \mathbb{E}(X(X + 1)) - \mathbb{E}(X) - \mathbb{E}(X)^2$$

12

Proposition (propriétés de la variance).

Soit X une variable aléatoire réelle définie sur $(\Omega, \mathbb{P})$.

- **Réel positif.** On a toujours $\mathbb{V}(X) \geq 0$.

Et on a

$$\mathbb{V}(X) = 0 \iff X \text{ est presque sûrement constante.}$$

- **Transformation affine.** Soit $a, b \in \mathbb{R}$. On a $\mathbb{V}(aX + b) = a^2 \mathbb{V}(X)$.



13

Proposition (espérance et variance des lois usuelles).

— Uniforme.

Soit $X \sim \mathcal{U}([1, n])$ où $n \in \mathbb{N}^*$.

$$\mathbb{E}(X) = \frac{n+1}{2}$$

$$\mathbb{V}(X) = \frac{n^2-1}{12}$$

Soit $X \sim \mathcal{U}([a, b])$ où $a \leq b \in \mathbb{Z}$.

$$\mathbb{E}(X) = \frac{a+b}{2}$$

$$\mathbb{V}(X) = \frac{(b-a+1)^2-1}{12}$$

— Bernoulli.

Soit $X \sim \mathcal{B}(p)$ où $p \in [0, 1]$.

$$\mathbb{E}(X) = p$$

$$\mathbb{V}(X) = pq \text{ où } q = 1 - p$$

— Binomiale.

Soit $X \sim \mathcal{B}(n, p)$ où $n \in \mathbb{N}^*$ et $p \in [0, 1]$.

$$\mathbb{E}(X) = np$$

$$\mathbb{V}(X) = npq \text{ où } q = 1 - p$$

Écart type

14

Définition. L'écart-type d'une variable aléatoire réelle X est le réel $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

- On prend la racine carrée de la variance pour des raisons d'homogénéité; ainsi, l'écart-type s'exprime dans les mêmes unités que les valeurs prises par la variable aléatoire.
- Une variable aléatoire réelle X telle que $\mathbb{E}(X) = 0$ et $\sigma(X) = 1$ est dite *centrée réduite*.

15

Proposition (VA centrée réduite).

Soit X une variable aléatoire telle que $\mathbb{V}(X) \neq 0$.

La variable aléatoire

$$X^* = \frac{1}{\sqrt{\mathbb{V}(X)}}(X - \mathbb{E}(X))$$

est une variable aléatoire centrée réduite, appelée *variable aléatoire centrée réduite associée à X* .



III. Covariance — Variance d'une somme

Définition

16

Définition.

Soit X et Y deux variables aléatoires réelles définies sur $(\Omega, \mathbb{P})$.

On appelle *covariance* de X et Y , le réel noté $\text{cov}(X, Y)$ défini par :

$$\text{cov}(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right).$$

- Cette définition a un sens car $(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))$ est une variable aléatoire réelle. C'est la variable aléatoire $g(X, Y)$, où $g : (x, y) \mapsto (x - \mathbb{E}(X))(y - \mathbb{E}(Y))$.

- **Refrain!** La covariance est l'espérance du produit des variables centrées.

La covariance de deux VA est le moment *centré* d'ordre 1 de leur produit.

La variance d'une VA est le moment *centré* d'ordre 1 de son carré.

17

Formule de Kœnig Huygens pour la covariance.

Soit X et Y deux variables aléatoires réelles définies sur $(\Omega, \mathbb{P})$.

La covariance de X et Y est donnée par la formule :

$$\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

- **Remarque.** Si la variable X est constante égale à a , on a $\text{cov}(X, Y) = 0$ puisqu'alors :

$$\mathbb{E}(XY) = \mathbb{E}(aY) = a\mathbb{E}(Y) = \mathbb{E}(X)\mathbb{E}(Y).$$

- Comment calculer $\mathbb{E}(XY)$?

Propriétés de la covariance

18

Proposition.

Soit X, X', Y et Y' des variables aléatoires réelles définies sur $(\Omega, \mathbb{P})$ et λ un réel. On a :

(i) $\text{cov}(X, X) = \mathbb{V}(X) \geq 0$

(ii) $\text{cov}(X, Y) = \text{cov}(Y, X)$

(iii) $\text{cov}(X + X', Y) = \text{cov}(X, Y) + \text{cov}(X', Y) \quad \text{cov}(\lambda X, Y) = \lambda \text{cov}(X, Y)$

(iv) $\text{cov}(X, Y + Y') = \text{cov}(X, Y) + \text{cov}(X, Y') \quad \text{cov}(X, \lambda Y) = \lambda \text{cov}(X, Y)$

- L'application $(X, Y) \mapsto \text{cov}(X, Y)$ est une forme bilinéaire, symétrique et positive sur l'espace vectoriel des variables aléatoires réelles sur $(\Omega, \mathbb{P})$.
- L'application $(X, Y) \mapsto \text{cov}(X, Y)$ n'est pas un produit scalaire, car $\mathbb{V}(X) = 0$ n'implique pas $X = 0$.
En effet, $\mathbb{V}(X) = 0$ équivaut à X est presque sûrement constante.
- L'utilisation de la bilinéarité peut éviter des calculs de covariance souvent pénibles.
- On peut retenir que :

$$\text{cov}(X + a, Y + b) = \text{cov}(X, Y) + \text{cov}(a, Y) + \text{cov}(X, b) + \text{cov}(a, b) = \text{cov}(X, Y)$$

car la covariance est nulle si l'une des variables aléatoires est constante.



Variance d'une somme de variables aléatoires

19

Proposition. Soit X et Y des variables aléatoires réelles sur $(\Omega, \mathbb{P})$.

On a :

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + 2\text{cov}(X, Y) + \mathbb{V}(Y)$$

- Soit $\lambda, \mu \in \mathbb{R}$. On a

$$\mathbb{V}(\lambda X + \mu Y) = \lambda^2 \mathbb{V}(X) + 2\lambda\mu \text{cov}(X, Y) + \mu^2 \mathbb{V}(Y)$$

- La formule permet de calculer la covariance de X et Y à partir des variances de X , Y et $X + Y$.
- Par récurrence, on généralise la proposition précédente au cas de n variables aléatoires.

Soit $(X_1, \dots, X_n)$ une famille de n variables aléatoires réelles sur $(\Omega, \mathbb{P})$.

La variance de la somme $X_1 + \dots + X_n$ est donnée par :

$$\mathbb{V}(X_1 + \dots + X_n) = \sum_{k=1}^n \mathbb{V}(X_k) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j).$$

- On peut démontrer (en suivant les idées de la preuve de l'inégalité de Cauchy-Schwarz) que :

$$|\text{cov}(X, Y)| \leq \sigma(X)\sigma(Y).$$

20

Question. Reprenons le problème des rencontres. Déterminer $\mathbb{V}(X)$.

sol → 14

Variables aléatoires décorrélées

Soit X et Y deux variables aléatoires réelles sur $(\Omega, \mathbb{P})$.

21

Proposition. Si X et Y sont indépendantes, alors

$$\text{cov}(X, Y) = 0 \quad \text{et} \quad \mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y).$$

- On peut calculer la covariance de variables aléatoires s'exprimant en fonctions de variables aléatoires indépendantes, en utilisant la bilinéarité de la covariance.

22

Question. Soit trois variables aléatoires mutuellement indépendantes X , Y , Z suivant la même loi de Bernoulli de paramètre p . Déterminer $\text{cov}(X + Y, Y + Z)$.

sol → 14

23

Définition. On dit que X et Y sont *décorrélées* lorsque $\text{cov}(X, Y) = 0$,

24

Proposition. Si X et Y sont indépendantes, alors elles sont décorrélées.

- **Attention.** La réciproque est fausse.

Soit X suivant une loi uniforme sur $\{-1, 0, 1\}$ et $Y = \mathbb{1}_{\{X=0\}}$. Alors :

- les variables aléatoires X et Y ne sont pas indépendantes, car :

$$\mathbb{P}(\{X = 1\} \cap \{Y = 1\}) = \mathbb{P}(\{X = 1\} \cap \{X = 0\}) = 0$$

alors que $\mathbb{P}(X = 1)\mathbb{P}(Y = 1) \neq 0$ (WHY?)

- les variables X et Y sont décorrélées;

en effet, comme $XY = 0$, on a $\mathbb{E}(XY) = 0$; de plus, $\mathbb{E}(X) = 0$, donc $\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0$.



- **Cas des Bernoulli.**

Les variables de Bernoulli constituent une exception agréable : deux variables aléatoires décorréées suivant une loi de Bernoulli sont indépendantes.

En effet, si X et Y sont deux variables de Bernoulli, on a :

$$\begin{aligned}\text{cov}(X, Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{P}(XY = 1) - \mathbb{P}(X = 1)\mathbb{P}(Y = 1) \\ &= \mathbb{P}(\{X = 1\} \cap \{Y = 1\}) - \mathbb{P}(X = 1)\mathbb{P}(Y = 1).\end{aligned}$$

Ainsi, si $\text{cov}(X, Y) = 0$, alors on a $\mathbb{P}(\{X = 1\} \cap \{Y = 1\}) = \mathbb{P}(X = 1)\mathbb{P}(Y = 1)$.

D'après ..., cela entraîne l'indépendance de X et Y .

25
preuve

Proposition. Soit $(X_1, \dots, X_n)$ une famille de variables aléatoires réelles sur $(\Omega, \mathbb{P})$.

Si $X_1, \dots, X_n$ sont deux à deux décorréées, alors on a :

$$\mathbb{V}(X_1 + \dots + X_n) = \sum_{k=1}^n \mathbb{V}(X_k).$$

- Ce résultat s'applique en particulier si les variables aléatoires $X_1, \dots, X_n$ deux à deux indépendantes et *a fortiori* si elles sont mutuellement indépendantes. Cette formule montre donc l'intérêt de décomposer une variable aléatoire en somme de variables aléatoires indépendantes plus simples.



IV. Inégalités probabilistes

Soit X une variable aléatoire réelle définie sur $(\Omega, \mathbb{P})$.

26

Proposition (inégalité de Markov).

Si X est positive, alors :

$$\forall a > 0, \quad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

27

Proposition (inégalité de Bienaymé-Tchebychev).

On a

$$\forall \varepsilon > 0, \quad \mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}$$

- On a l'inclusion $\{Y > \text{truc}\} \subset \{Y \geq \text{truc}\}$. Donc on a aussi :

$$\forall \varepsilon > 0, \quad \mathbb{P}(|X - \mathbb{E}(X)| > \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}$$

- L'inégalité de Bienaymé-Tchebychev est une *inégalité de concentration* : elle fournit une majoration de la probabilité que X s'écarte de son espérance.
- Comme $\{|X - \mathbb{E}(X)| \geq \varepsilon\}$ est l'événement contraire de $\{|X - \mathbb{E}(X)| < \varepsilon\}$, on a aussi :

$$\forall \varepsilon > 0, \quad \mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon) \geq 1 - \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

- Mes refrains.**

Markov, c'est du positif et du moment d'ordre 1.

Bienaymé-Tchebychev, c'est du moment d'ordre 2.

28

Question. Soit $X_1, \dots, X_n$ des variables aléatoires réelles indépendantes.

Pour tout k , on suppose que $X_k \hookrightarrow \mathcal{B}(p)$ et on pose $S_n = X_1 + \dots + X_n$.

Montrer que

$$(\spadesuit) \quad \forall \varepsilon > 0, \quad \mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow +\infty} 0$$

- Commentaires sur cette question.**

Considérons une suite de n épreuves indépendantes telle que, pour chaque épreuve, la probabilité qu'un certain événement A soit réalisé est p .

Pour tout k , on note X_k la variable indicatrice de l'événement « A est réalisé à la $k^{\text{ème}}$ épreuve ».

On obtient des variables aléatoires indépendantes qui suivent la loi $\mathcal{B}(p)$.

Alors la variable aléatoire $\frac{S_n}{n}$ est égale à la fréquence avec laquelle l'événement A s'est réalisé parmi les n épreuves.

La limite $(\spadesuit)$ dit que cette fréquence est proche de p pour les grandes valeurs de n .

Ce résultat conforte l'intuition qui voit dans une probabilité une fréquence.

Si, par exemple, on lance un dé un grand nombre de fois, 3 apparaîtra à peu près dans $\frac{1}{6}$ des lancers.



Espérance et variance

preuve et éléments de correction

6

Il s'agit de calculer l'espérance de la variable X égale au nombre de rencontres.

Pour tout i , on note X_i la variable qui prend la valeur 1 s'il y a une rencontre au $i^{\text{ème}}$ tirage, et 0 sinon.

Ainsi, X_i suit une loi de Bernoulli de paramètre :

$$\mathbb{P}(X_i = 1) = \frac{(n-1)!}{n!} = \frac{1}{n}$$

car il y a $n!$ tirages possibles et $(n-1)!$ tirages pour lesquels la boule i est tirée au $i^{\text{ème}}$ tirage.

Donc

$$\mathbb{E}(X_i) = \frac{1}{n}$$

Or, on a $X = \sum_{i=1}^n X_i$.

Par linéarité, on a $\mathbb{E}(X) = \sum_{i=1}^n \mathbb{E}(X_i) = n \frac{1}{n} = 1$.

7

Pour alléger les notations, posons $Y = g(X)$.

Partons de la définition alternative de l'espérance :

$$\mathbb{E}(Y) = \sum_{\omega \in \Omega} Y(\omega) \mathbb{P}(\{\omega\}) = \sum_{\omega \in \Omega} g(X(\omega)) \mathbb{P}(\{\omega\})$$

Maintenant, utilisons le fait que

$$\Omega = \bigsqcup_{x \in X(\Omega)} [X = x]$$

Ainsi, la somme simple $\sum_{\omega \in \Omega} \text{truc}_\omega$ vaut la somme double $\sum_{x \in X(\Omega)} \sum_{\omega \in [X=x]} \text{truc}_\omega$.

Reprenons le calcul avec cette information.

$$\begin{aligned} \mathbb{E}(Y) &= \sum_{x \in X(\Omega)} \sum_{\omega \in [X=x]} g(X(\omega)) \mathbb{P}(\{\omega\}) \\ &= \sum_{x \in X(\Omega)} \sum_{\omega \in [X=x]} g(x) \mathbb{P}(\{\omega\}) \\ &= \sum_{x \in X(\Omega)} g(x) \underbrace{\sum_{\omega \in [X=x]} \mathbb{P}(\{\omega\})}_{\mathbb{P}(X=x)} \\ &= \sum_{x \in X(\Omega)} g(x) \mathbb{P}(X=x) \end{aligned}$$

9

Pour calculer $\mathbb{E}(XY)$, On remarque que $XY = g(Z)$ où $Z = (X, Y)$ et $g : (x, y) \mapsto xy$.

En appliquant la formule de transfert, on obtient :

$$\mathbb{E}(XY) = \mathbb{E}(g(Z)) = \sum_{z \in Z(\Omega)} g(z) \mathbb{P}(Z=z) = \sum_{(x,y) \in Z(\Omega)} g(x,y) \mathbb{P}((X,Y) = (x,y)).$$

On peut même changer l'ensemble d'indexation de la somme en $X(\Omega) \times Y(\Omega)$.

En effet, si $(x, y) \in (X(\Omega) \times Y(\Omega)) \setminus Z(\Omega)$, alors $\mathbb{P}((X, Y) = (x, y)) = 0$.



D'où

$$\mathbb{E}(XY) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xy \mathbb{P}(X=x, Y=y).$$

Utilisons à présent l'indépendance de X et Y .

On obtient

$$\mathbb{E}(XY) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xy \mathbb{P}(X=x) \mathbb{P}(Y=y)$$

Cette somme double, devient le produit de deux sommes simples car « les variables sont bien séparées » :

$$\mathbb{E}(XY) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} x \mathbb{P}(X=x) y \mathbb{P}(Y=y) = \left(\sum_{x \in X(\Omega)} x \mathbb{P}(X=x) \right) \left(\sum_{y \in Y(\Omega)} y \mathbb{P}(Y=y) \right)$$

20

On reprend la décomposition $X = X_1 + \dots + X_n$, où X_i est la variable indicatrice de l'événement « il y a une rencontre au $i^{\text{ème}}$ tirage ».

Chaque X_i est une variable de Bernoulli de paramètre $\frac{1}{n}$ et donc de variance :

$$\frac{1}{n} \left(1 - \frac{1}{n} \right) = \frac{n-1}{n^2}.$$

Soit $(i, j) \in \llbracket 1, n \rrbracket^2$ tel que $i \neq j$.

L'événement $\{X_i = 1\} \cap \{X_j = 1\}$ est réalisé s'il y a rencontre aux $i^{\text{ème}}$ et $j^{\text{ème}}$ tirages.

Les résultats des autres tirages sont quelconques. On obtient donc :

$$\mathbb{P}(\{X_i = 1\} \cap \{X_j = 1\}) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}.$$

Mais $X_i X_j$ est aussi une variable de Bernoulli donc :

$$\mathbb{E}(X_i X_j) = \mathbb{P}(X_i X_j = 1) = \mathbb{P}(\{X_i = 1\} \cap \{X_j = 1\}) = \frac{1}{n(n-1)}.$$

On en déduit :

$$\text{cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i) \mathbb{E}(X_j) = \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{1}{n^2(n-1)}.$$

Appliquons alors ..., en remarquant que toutes les variances intervenant sont égales, ainsi que toutes

les covariances, et qu'il y a $\binom{n}{2} = \frac{n(n-1)}{2}$ termes dans la seconde somme :

$$\mathbb{V}(X) = \sum_{i=1}^n \mathbb{V}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j) = n \mathbb{V}(X_1) + 2 \frac{n(n-1)}{2} \text{cov}(X_1, X_2)$$

et, après calculs et simplifications, $\mathbb{V}(X) = 1$.

22

Par bilinéarité de la covariance, on a :

$$\text{cov}(X+Y, Y+Z) = \text{cov}(X, Y) + \text{cov}(X, Z) + \text{cov}(Y, Y) + \text{cov}(Y, Z).$$

On a $\text{cov}(X, Y) = \text{cov}(X, Z) = \text{cov}(Y, Z) = 0$ par indépendance deux à deux des variables aléatoires.

De plus, $\text{cov}(Y, Y) = \mathbb{V}(Y) = p(1-p)$.

D'où

$$\text{cov}(X+Y, Y+Z) = p(1-p).$$



25

On a

$$\mathbb{V}(X_1 + \cdots + X_n) = \sum_{k=1}^n \mathbb{V}(X_k) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j).$$

Pour $i \neq j$, les variables X_i et X_j sont indépendantes donc $\text{cov}(X_i, X_j) = 0$, d'où :

$$\mathbb{V}(X_1 + \cdots + X_n) = \sum_{k=1}^n \mathbb{V}(X_k).$$

