☙ Variables aléatoires et loi des grands nombres

1 Sommes de variables aléatoires

Exemple 1 *On considère trois variables aléatoires X, Y et S.*
La variable X peut prendre les valeurs 0 et 1 de manière équiprobable.
La variable Y peut également prendre les valeurs 0 et 1 de manière équiprobable.
La variable S est la somme des deux variables précédentes : $S = X + Y$.

1. *Déterminer la loi de X et la loi de Y ; autrement dit, il s'agit de remplir les tableaux suivants :*

$X =$	0	1
$P(X =)$		

$Y =$	0	1
$P(Y =)$		

2. *Calculer l'espérance et la variance des lois X et Y.*
On rappelle que si Z est une variable aléatoire discrète prenant un nombre fini de valeurs $\{z_1; ..; z_n\}$ avec $P(Z = z_i) = p_i$, alors :

$$\mathbb{E}(Z) = \sum_{i=1}^{n} z_i \times p_i$$

$$V(Z) = \sum_{i=1}^{n} p_i \times (z_i - \mathbb{E}(Z))^2$$

3. *Déterminer la loi de S.*
4. *Calculer l'espérance et la variance de S.*
Comparer les résultats avec $\mathbb{E}(X) + \mathbb{E}(Y)$ et $V(X) + V(Y)$.
5. *Est ce que les variables X et Y semblent indépendantes ?*

Exemple 2 *On considère trois variables aléatoires X, Y et S.*
La variable X peut prendre les valeurs 0 et 1 de manière équiprobable.
La variable Y peut également prendre les valeurs 0 et 1 mais :

- ✏ *si $X = 0$ alors $Y = 1$.*
- ✏ *si $X = 1$ alors $Y = 0$.*

La variable S est la somme des deux variables précédentes : $S = X + Y$.

1. *Déterminer la loi de X et la loi de Y ; autrement dit, il s'agit de remplir les tableaux suivants :*
2. *Calculer l'espérance et la variance des lois X et Y.*
3. *Déterminer la loi de S.*
4. *Calculer l'espérance et la variance de S.*
Comparer les résultats avec $\mathbb{E}(X) + \mathbb{E}(Y)$ et $V(X) + V(Y)$.
5. *Est ce que les variables X et Y semblent indépendantes ?*

Définition 1 *Soient X et Y deux variables aléatoires.*
La loi de probabilité de la variable aléatoire $S = X + Y$ est donnée par :

$$P(S = k) = \sum_{i+j=k} P\big((X = i) \cap (Y = j)\big)$$

De plus, si les variables X et Y sont indépendantes, alors :

$$P\big((X = i) \cap (Y = j)\big) = P(X = i) \times P\big(Y = j\big)$$
$$\text{donc } P(S = k) = \sum_{i+j=k} P(X = i) \times P\big(Y = j\big)$$

Propriétés 1 *Soit X et Y deux variables aléatoires ainsi que a et b deux nombres réels :*

$$E(aX + b) = aE(X) + b$$
$$E(X + Y) = E(X) + E(Y)$$
$$V(aX + b) = a^2 V(X)$$

En général, on a $V(X + Y) \neq V(X) + V(Y)$ ***mais si X et Y sont indépendants*** *:*

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

Exercice 1 *On considère le jeu suivant qui se déroule en deux étapes :*

- ✏ *dans un premier temps, on lance une pièce de monnaie.*
  Pile fait gagner un euro et Face fait gagner deux euros.
- ✏ *dans un second temps, on lance un dé à 6 faces. Si on tombe sur un chiffre pair, on ne gagne rien, si on tombe sur le « 3 » ou le « 5 », on gagne 2 €. Si on tombe sur le « 1 », on perd 6 €.*
  La variable aléatoire X désigne les gains à la 1ère partie, la variable aléatoire Y désigne les gains à la 2e partie.
  On considère que les variables aléatoires X et Y sont indépendantes.
  Établir la loi de probabilité de la variable aléatoire somme $S = X + Y$ donnant le gain total cumulé à la fin des deux parties.

Exercice 2 *Une entreprise qui fabrique des roulements à bille fait une étude sur une gamme de billes produites. Le diamètre théorique doit être égal à 1,3 cm mais cette mesure peut être légèrement erronée. L'expérience consiste à tirer au hasard une bille d'un lot de la production et à mesurer son diamètre. On considère la variable aléatoire X qui, à une bille choisie au hasard, associe son diamètre. La loi de probabilité de X est résumée dans le tableau suivant :*

x_i	1.298	1.299	1.3	1.301	1.302
$P(X = x_i)$	0.2	0.1	0.2	0.4	0.1

Calculer l'espérance et l'écart-type de la loi de probabilité de X. Pour simplifier les calculs, on étudiera la variable aléatoire $Y = 1000X - 1300$.

Définition 2 *Un échantillon de taille n d'une loi de probabilité est une liste de n variables aléatoires indépendantes suivant cette loi.*

Propriétés 2 *Soit $(X_1, .., X_n)$ un échantillon de taille n de variables aléatoires indépendantes suivant une même loi.*
On pose $S = X_1 + .. + X_n$; on a alors :

$$E(S) = E(X_1) + .. + E(X_n)$$
$$V(S) = V(X_1) + .. + V(X_n)$$

Exercice 3 *Sur un axe gradué, on part de l'origine et on marche aléatoirement d'une unité vers la gauche ou d'une unité vers la droite :*

- ✏ *on se déplace d'une unité vers la gauche avec la probabilité $\frac{2}{5}$.*
- ✏ *on se déplace d'une unité vers la droite avec la probabilité $\frac{3}{5}$*

On suppose que chacun des déplacements sont indépendants les uns des autres.
Pour k un entier naturel plus grand que 1, on appelle X_k la variable qui représente le $k$ième déplacement et que l'on modélise ainsi : $X_k = 1$ pour un déplacement vers la droite et $X_k = -1$ pour un déplacement vers la gauche.
La variable $S_n = X_1 + .. + X_n$ représente le déplacement global sur la droite graduée.

1. *Si on se déplace suffisamment longtemps, vers quel côté peut on espérer se trouver ?*
2. *Calculer $E(X_k)$ et $V(X_k)$.*
3. *En déduire $E(S_n)$ et $V(S_n)$.*
4. *Au bout de combien de déplacements peut on espérer se trouver à 10 unité vers la droite ?*
   Calculer $\sigma(S_n)$.

Propriétés 3 *Soit $(X_1, .., X_n)$ un échantillon de taille n de variables aléatoires indépendantes suivant une loi de Bernouilli de paramètre p.*
On pose $S = X_1 + .. + X_n$; S suit une loi binomiale de paramètres n et p avec :

$$E(S) = E(X_1) + .. + E(X_n) = np$$
$$V(S) = V(X_1) + .. + V(X_n) = np(1-p)$$
$$\sigma(S) = \sqrt{V(S)} = \sqrt{np(1-p)}$$

Exercice 4 *On lance 5 fois de suite un dé à six faces.*
On considère comme succès le fait d'obtenir 5 ou 6.
On considère la variable aléatoire S donnant le nombre de succès.
Calculer $E(S)$, $V(S)$ et $\sigma(S)$.

Exercice 5 *un garagiste dispose de deux voitures de location. Chacune est utilisable en moyenne* 4 *jours sur* 5.
Il loue les voitures avec une marge brute de 300 *euros par jour et par voiture.*
On considère X la variable aléatoire égale au nombre de clients se présentant chaque jour pour louer une voiture. On suppose $X(\Omega) = \{0;1;2;3\}$ avec :

$$P(X=0)=0,1 \quad P(X=1)=0,3 \quad P(X=2)=0,4 \quad P(X=3)=0,2$$

1. *On note Z le nombre de voitures disponibles par jour. Déterminer la loi de Z.*
   Par la suite, on va considérer que Z et X sont indépendantes.
2. *On note Y la variable aléatoire donnant le nombre de clients satisfaits par jour. Déterminer la loi de Y.*
3. *Calculer la marge brute moyenne par jour.*

Exercice 6 *Les vaches laitières sont atteintes par une maladie M avec la probabilité* $p = 0.15$.
Pour dépister la maladie M dans une étable de n vaches, on fait procéder à une analyse de lait. On dispose de deux méthodes :

- ✏ *la première consiste à faire une analyse sur un échantillon de lait de chaque vache.*
- ✏ *la seconde s'effectue en faisant d'abord une analyse sur un échantillon de lait provenant du mélange des n vaches. Si le résultat est positif, on effectue une nouvelle analyse, cette fois sur chaque vache.*

On veut connaître la méthode la plus économique (celle qui nécessite le moins d'analyse).
Pour cela, on note X_n la variable aléatoire du nombre d'analyses réalisées dans la deuxième méthode.
On pose $Y_n = \frac{X_n}{n}$.

1. *Déterminer la loi de Y_n et calculer son espérance.*
2. *Étudier la fonction $f(x) = ax + \ln(x)$, pour $a = \ln(0.85)$. Donner la liste des entiers tels que $f(n) > 0$.*
3. *Montrer que $f(n) > 0 \Leftrightarrow E(Y_n) < 1$. En déduire la réponse à la question posée.*

2 Moyenne d'un échantillon

Définition

Soit $(X_1, X_2, ..., X_n)$ un échantillon de taille n de variables aléatoires indépendantes suivant une même loi.
La variable aléatoire moyenne M_n de l'échantillon est donnée par :

$$M_n = \frac{1}{n}(X_1 + X_2 + ... + X_n)$$

Propriétés

Soit $(X_1, X_2, ..., X_n)$ un échantillon de taille n de variables aléatoires indépendantes suivant la même loi que X :

$$E(M_n) = E(X)$$
$$V(M_n) = \frac{1}{n}V(X)$$
$$\sigma(M_n) = \frac{1}{\sqrt{n}}\sigma(X)$$

Exemple 3 *On considère la variable aléatoire X qui prend, de façon équiprobable, les valeurs –4, 0, 1, 3 et 6.*
M_{50} est la variable aléatoire moyenne d'un échantillon de taille 50 de la loi de X.
Calculer l'espérance, la variance et l'écart type de M_{50}.

3 Inégalité de Bienaymé-Tchebychev

Propriétés

Soit une variable aléatoire X. Pour tout réel strictement positif δ, on a :

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

Exemple 4 *Soit une variable aléatoire X qui suit la loi binomiale de paramètres $n = 20$ et $p = 0,1$.*

1. *Appliquer l'inégalité de Bienaymé-Tchebychev avec $\delta = 2\sigma(X)$. Interpréter.*
2. *Recommencer avec $\delta = 3\sigma(X)$, puis $\delta = 4\sigma(X)$. Que constate-t-on ?*
3. **a.** *Simuler N valeurs de la variable aléatoire X par une fonction en Python dans le but d'estimer la probabilité $P(|X - 2| \geq 2\sigma(X))$.*
   On testera le programme pour différentes valeurs de N.

   b. *Au regard des résultats obtenus par le programme, peut-on penser que l'inégalité de Bienaymé-Tchebychev a un caractère optimal ?*

4 Inégalité de concentration

Propriétés

Soit la variable aléatoire moyenne M d'un échantillon de taille n de la variable aléatoire X. Pour tout réel strictement positif δ, on a :

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

Exemple 5 *Soit une variable aléatoire X qui suit la loi de Bernouilli de paramètre 0.3.*
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 tel que la probabilité que la moyenne M_n appartienne à l'intervalle $]0.03;0.37[$ soit supérieure à 0.95.

5 Lois des grands nombres

Propriétés

Soit la variable aléatoire moyenne M d'un échantillon de taille n de la variable aléatoire X. Pour tout réel strictement positif δ, on a :

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

On peut traduire cette propriété de la manière suivante : plus n est grand, plus la valeur de M_n tend à être proche de $E(X)$.

Exemple 6 *On considère la variable aléatoire X qui prend ses valeurs de manière équiprobable parmi les entiers 1 à 5 et on nomme M_n la variable aléatoire moyenne d'un échantillon de taille n de la variable aléatoire X*

1. *Simuler 500 valeurs de la variable aléatoire M par une fonction en Python dans le but d'estimer la probabilité $P(|M_n - E(X)| \geq \delta)$.*
   Tester le programme pour différentes valeurs de δ et des valeurs de n de plus en plus grande.
2. *Que constate-t-on ?*

Exercice 7 *On considère une variable aléatoire X.*

1. *Exprimer les événements suivants sous la forme $(|X - a| \geq t)$.*
   a. $(X - 2 \geq 2)$ *ou* $(X - 2 \leq -2)$
   b. $(X - a \geq 3)$ *ou* $(X - a \leq -3)$

c. *La distance entre* X *et* 5 *est supérieure ou égale à* 4.

d. $(X \geq 6)$ *ou* $(X \leq 0)$

e. $(X \geq 5)$ *ou* $(X \leq 1)$

f. $((X-2)^2 \geq 9)$

2. *Exprimer l'événement complémentaire à l'aide d'une valeur absolue.*

a. $(-2 < X - 1 < 2)$.

b. $(1 < X < 5)$.

c. $(1 + X > 2 > -1 + X)$

Exercice 8 *Soit* X *une varibale aléatoire d'espérance* 10 *et de variance* 16.

1. *Donner une majoration de la probabilité que* X *s'écarte de son espérance d'au moins* 4.

2. *Donner une majoration de la probabilité que* X *s'écarte de* 10 *d'au moins* 10.

Exercice 9 *Soit* X *une variable aléatoire.*
Donner une majoration de chaque probabilité suivante :

1. $P(|X - E(X)| \geq \sigma(X))$

2. $P(|X - E(X)| \geq 2\sigma(X))$

3. $P(|X - E(X)| \geq 3\sigma(X))$

Exercice 10 *Chaque jour un train subit au départ un retard aléatoire, noté* X, *mesuré en minutes.*
On suppose que $E(X) = 10$ *et* $V(X) = 100$.
Un passager porte réclamation en assurant qu'un jour sur deux, le train a plus de 30 *minutes de retard.*
Justifier que cette affirmation est sans doute très exagérée.

Exercice 11 *On considère une marche aléatoire sur* $\mathbb{Z}$ *définie de la façon suivante : on part de* 0 *et, à chaque étape, indépendamment de la précédente, on a un probabilité* $\frac{1}{2}$ *de faire une pas vers la droite et une probabilité* $\frac{1}{2}$ *de faire un pas vers la gauche.*
On note, pour i *un entier naturel non nul,* X_i, *la variable aléatoire dont la loi est donnée par :*

$$P(X_i = 1) = \frac{1}{2}$$
$$P(X_i = -1) = \frac{1}{2}$$

Pour n *un entier strictement positif, on pose :*

$$S_n = X_1 + X_2 + .. + X_n$$

1. *Que représente* S_1 *?* S_2 *?*

2. *Que représente* S_n *?*

3. *Déterminer l'espérance et la variance des variables aléatoires X_i pour $i \geq 1$.*
4. *En déduire l'espérance et la variance de S_n.*
5. *Montrer que pour $t > 0$:*

$$P(|S_n| \geq t) \leq \frac{n}{t^2}$$

6. *En déduire, pour tout x réel positif, une majoration de $P(|M_n| \geq x)$.*
7. *Calculer, pour $x > 0$:*

$$\lim_{n\to+\infty} P(|M_n| \geq x)$$

et interpréter le résultat.

Aurait-on pu faire autrement pour la question précédente ?

Exercice 12 *On lance cinq fois une pièce non truquée.*
Pour $1 \leq i \leq 5$, on note X_i la variable aléatoire égale à 1 si on obtient Pile au i-ième lancer, et 0 sinon.
On note $S_5 = X_1 + X_2 + X_3 + X_4 + X_5$.

1. *Quelle est la loi suivie par chacun des X_i ?*
2. *Calculer $E(X_i)$ et $V(X_i)$.*
3. *En déduire $E(S_5)$ et $V(S_5)$.*
4. *Que représente la variable aléatoire S_5 ?*
5. *Justifier que S_5 suit une loi binomiale dont on précisera les paramètres.*
6. *Montrer que $P(|S_5 - 2.5| \geq 2) \leq 0.3125$.*
   Interpréter ce résultat.

Exercice 13 *Combien de fois faut-il lancer une pièce équilibrée pour être sûr à 99% que la proportion de Pile soit comprise entre 0.45 et 0.55 ?*

Exercice 14 *En France métropolitaine en 2019, il y a eu 343003 naissances de garçons et 329909 naissances de filles.*

1. *On note X la variable aléatoire qui prend pour valeurs le nombre de garçons qui naissent dans une année.*
   En supposant que les événements "Avoir un garçon" et "Avoir une fille" sont équiprobables, justifier que le nombre de garçons qui sont nés en 2019 suivait une loi binomiale $\mathscr{B}(n; 0.5)$, où n est un entier naturel à déterminer.
2. *Calculer le nombre de naissances de garçons qui étaient espérés en 2019.*
3. *Quel est le pourcentage de naissances supplémentaires de garçons par rapport au nombre espéré pour l'année 2019 ?*

4. *Donner une majoration de :*

$$P\left(\left|\frac{X}{n}-\frac{1}{2}\right| \geq 0.019\right)$$

et une minoration de :

$$P\left(\left|\frac{X}{n}-\frac{1}{2}\right| < 0.019\right)$$

Comment interpréter ce résultat ?

Exercice 15 *Quatre jetons indiscernables au toucher, numérotés de* 1 *à* 4*, sont placés dans un sac opaque.*

1. *On choisit de manière aléatoire un jeton dans le sac. On ne le remet pas et on en choisit un deuxième. On appelle* X *la variable aléatoire qui associe au premier tirage le numéro du jeton choisi et* Y *celle qui associe au second tirage le numéro du jeton choisi.*
   a. *Déterminer la loi de la variable aléatoire* $X \times Y$ *et son espérance.*
   b. *A-t-on* $E(X \times Y) = E(X) \times E(Y)$ *?*
2. *On considère maintenant qu'avant de tirer le second jeton, on remet le premier dans le sac.*
   a. *Peut-on considérer comme indépendants les deux tirages ?*
   b. *Déterminer la loi de* $X \times Y$ *et son espérance.*
   c. *Comparer* $E(X \times Y)$ *et* $E(X) \times E(Y)$.

Exercice 16 *Une association organise un jeu d'argent pour récolter des fonds.*
Chaque participant doit verser 9 *euros de participation avant de lancer trois dés équilibrés.*
Si le résultat est un 421*, il gagne* 99 *euros.*

1. *Calculer la probabilité de faire un* 421.
   On note G_i *le gain algébrique du* i*-ième participant.*
2. *Donner la loi de chaque* G_i.
3. *Calculer son espérance et sa variance.*
   On pose :

$$M_n = \frac{\sum_{i=1}^{n} G_i}{n}$$

4. *Déterminer :*

$$\lim_{n\to+\infty} P(|M_n - E(M_n)| \geq t)$$

5. *On note* X_n *la variable aléatoire égale au gain algébrique de l'association après la participation de* n *personnes.*
   a. *Exprimer* X_n*, à l'aide des variables aléatoires* G_i.
   b. *Calculer* $E(X_n)$ *et* $V(X_n)$.
6. *On suppose que* 100 *personnes vont participer.*
   Donner une minoration de la probabilité que le gain algébrique de l'association soit strictement compris entre 600 *euros et* 650 *euros.*