



On étudie les variables aléatoires à valeurs finies. L'espérance et la variance, en cas d'existence, apportent des renseignements qualitatifs sur les expériences aléatoires étudiées. On termine par un bref aperçu des résultats asymptotique (lois des grands nombres et théorème limit-centrale).



La roue de la fortune, Sébastien Brant

2	Variables aléatoires finies	1
1	Se concentrer sur l'information importante	2
2	Loi d'une variable aléatoire discrète	3
2.1	Loi image	3
2.2	Lois usuelles : uniforme, Bernoulli et binomiale	4
3	Couples de variables aléatoires, vecteurs aléatoires	6
4	Variables indépendantes	7
5	Espérance et variance d'une variable aléatoire	9
5.1	Espérance	9
5.2	Variance, écart-type et covariance	11
5.3	Inégalités de Markov et de Tchebychev	12
6	Les grands résultats asymptotiques	13
6.1	La loi faible des grands nombres	13
6.2	Le théorème « limit-centrale »	14
7	Passage du discret au continu	15
8	Énoncés des tests	16
9	Solutions des tests	17

1. Se concentrer sur l'information importante

En théorie moderne des probabilités, on préfère prendre un point de vue fonctionnel plutôt qu'ensembliste, et utiliser les variables aléatoires plutôt que les événements. Ce point de vue sera développé dans la suite du cours.

Variables aléatoires, heuristique

Une variable aléatoire est une grandeur qui dépend du résultat de l'expérience. Par exemple :

- ⇒ le nombre de 6 obtenus dans un lancé de 3 dés;
- ⇒ le nombre d'appels dans un central téléphonique pendant une heure;
- ⇒ la distance du point d'atteinte d'une flèche au centre de la cible;
- ⇒ la valeur maximale d'un prix d'actif sur un intervalle de temps donné.

Plus formellement,

Définition 2.0. Variable aléatoire sur un espace probabilisé fini

Soit (Ω, \mathbf{P}) un espace probabilisé fini.

- ⇒ On appelle variable aléatoire discrète sur (Ω, \mathbf{P}) toute application X définie sur Ω à valeurs dans un ensemble E .
- ⇒ Lorsque $E \subset \mathbb{R}$, la variable aléatoire est dite réelle.

La terminologie, consacrée par l'usage, est très malencontreuse et engendre une difficulté liée au vocabulaire employé. Une variable aléatoire, malgré son nom, n'est pas une variable (au sens de l'analyse), mais une fonction de la variable $\omega \in \Omega$.

Intérêt fondamental des variables aléatoires

Comme l'ensemble E est connu dans la pratique, nous allons préférer nous intéresser aux chances de réalisation des valeurs de X plutôt qu'aux chances de réalisation des résultats de l'expérience.

✘ Soit (Ω, \mathbf{P}) un espace probabilisé fini. Toute application $X : \Omega \rightarrow E$ est une variable aléatoire. Par exemple, pour le lancé simultané de trois dés, $S : \Omega \rightarrow \mathbb{N}$, définie par $S(\omega)$ est la somme des trois numéros obtenue, est une variable aléatoire.

Notation 2.1. Événements usuels

Si X est une variable aléatoire et si A est une partie de E , notation $\{X \in A\}$ ou $(X \in A)$ pour l'événement $X^{-1}(A)$. Notations $\mathbf{P}(X \in A)$, $\mathbf{P}(X = x)$, $\mathbf{P}(X \leq x)$.

2. Loi d'une variable aléatoire discrète

Soit $X : \Omega \rightarrow E$ une variable aléatoire. Nous allons voir comment X permet de définir sur l'espace probabilisable $\text{image } X(\Omega)$ une probabilité, appelée *loi de la variable* X . Nous étudierons ensuite quelques lois classiques.

2.1. Loi image

On commence par définir la loi d'une variable aléatoire X , appelée loi image par X de la loi de probabilité \mathbf{P} et notée \mathbf{P}_X .

Définition 2.2. Loi d'une variable aléatoire discrète

Soit (Ω, \mathbf{P}) un espace probabilisé fini et $X : \Omega \rightarrow E$ une variable aléatoire.

⇒ La loi \mathbf{P}_X de la variable aléatoire X est la probabilité définie sur $X(\Omega)$ par :

$$\forall A \in \mathcal{P}(X(\Omega)), \mathbf{P}_X(A) = \mathbf{P}(X \in A)$$

⇒ L'application \mathbf{P}_X est entièrement définie par la donnée des $\mathbf{P}(X = x)$ pour x dans $X(\Omega)$, ie par la distribution $\{(x, \mathbf{P}(X = x)), x \in X(\Omega)\}$.

✘ Loi uniforme sur un ensemble fini E . Si Ω est fini et non vide. Une variable aléatoire X sur la loi uniforme sur E si $X(\Omega) = E$ et si la loi de X est donnée par

$$\left\{ \left(e, \frac{1}{|E|} \right); e \in E \right\}$$

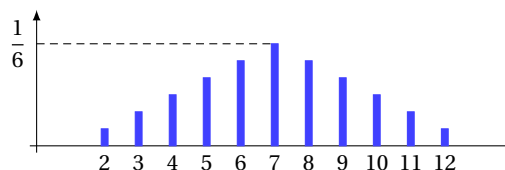
✘ Somme des numéros obtenus lors du lancé de deux dés équilibrés. L'univers vaut $\Omega = \llbracket 1, 6 \rrbracket^2$, et on considère naturellement l'équiprobabilité sur $(\Omega, \mathcal{P}(\Omega))$. L'application $X : \Omega \rightarrow \llbracket 1, 12 \rrbracket$ définie par $X(i, j) = i + j$ est la variable aléatoire « *somme des résultats des deux dés* ». Elle a pour loi

$$\mathbf{P}_X(B) = \frac{\text{nombre de couples } (i, j) \text{ tels que } i + j \in B}{36}$$

On trouve que X suit la loi suivante :

n	2	3	4	5	6	7	8	9	10	11	12
$\mathbf{P}(X = n)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

On représente une distribution de probabilité discrète sous la forme d'un histogramme. Par exemple, pour la loi de X , on met en abscisse les valeurs x_n prises par Ω et en ordonnée les probabilités $\mathbf{P}(X = x_n)$: cf. ci-contre la distribution de la variable X de l'exemple précédent.



✘ Deux variables aléatoires peuvent avoir la même loi sans être égales. Considérons par exemple le jet de deux dés, l'un bleu et l'autre rouge. Notons X le nombre de points indiqué par le dé bleu et Y celui du rouge. Les variables aléatoires X et Y sont définies sur le même espace probabilisé $\Omega = \llbracket 1, 6 \rrbracket^2$ muni de l'équiprobabilité. Les variables suivant la loi uniforme sur $X(\Omega) = Y(\Omega) = \llbracket 1, 6 \rrbracket$ mais ne sont pas égales!

Il sera très courant de manipuler des *fonctions d'une variable aléatoire*. Ici encore, les notations probabilistes diffèrent des usages de la théorie de ensembles.

Notation 2.3. Image d'une variable aléatoire par une fonction.

Si $f : X(\Omega) \rightarrow F$, alors $f \circ X$ est notée $f(X)$.

Il n'est pas toujours facile de calculer la loi de $Y = f(X)$ connaissant celle de X : pour tout $v \in f(X(\Omega))$, $(Y = v) = (X \in f^{-1}(\{v\}))$. Lorsque f est injective, on est ramené à $(Y = v) = (X = f^{-1}(v))$ (en notant f la corestriction – bijective – de f à son image).

✘ Loi de $Y = (-1)^X$ où X suit la loi uniforme sur $\llbracket 1, 2n \rrbracket$ pour $n \in \mathbb{N}^*$. On a clairement $Y(\Omega) = \{-1, 1\}$. De plus,

$$(Y = -1) = (X \text{ impair}) = \bigsqcup_{k=0}^{n-1} (X = 2k + 1) \text{ donc } \mathbf{P}(Y = -1) = \frac{n}{2n} = \frac{1}{2} \text{ et } \mathbf{P}(Y = 1) = 1 - \mathbf{P}(Y = -1) = \frac{1}{2}$$

ainsi Y est une variable aléatoire et Y suit donc la loi uniforme sur $\{-1, 1\}$.

On retiendra la démarche suivante :

Calculer la loi de $f(X)$

On suppose connue la loi de X et on pose $Y = f(X)$.

⇒ On calcule d'abord $Y(\Omega)$, ensemble des valeurs prises par Y .

⇒ Puis, pour tout $v \in Y(\Omega)$, on décrit l'événement $(f(X) = v)$ au moyen de X .

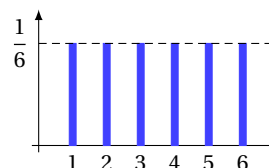
⇒ En en déduit $\mathbf{P}(Y = v)$.

2.2. Lois usuelles : uniforme, Bernoulli et binomiale

Définition 2.4. Loi uniforme

X suit une loi uniforme si \mathbf{P}_X est l'équiprobabilité sur $X(\Omega)$. On note $X \hookrightarrow \mathcal{U}(E)$ pour signifier que X suit la loi uniforme sur un ensemble E .

✘ Soit X le chiffre indiqué lors du lancé d'un dé équilibré. On a $X \hookrightarrow \mathcal{U}(\llbracket 1, 6 \rrbracket)$. On trouvera ci-contre l'histogramme de cette loi.



Type d'expériences aléatoires modélisées par la loi uniforme

Lors d'un tirage au hasard d'un objet parmi n , la variable égale au numéro de l'objet tiré suit la loi $\mathcal{U}(\llbracket 1, n \rrbracket)$.

Définition 2.4. Loi de Bernoulli

On dit qu'une variable suit la loi de Bernoulli de paramètre p dans $[0, 1]$ si

$$X(\Omega) = \{0, 1\}, \quad \mathbf{P}(X = 0) = 1 - p, \quad \mathbf{P}(X = 1) = p$$

On note alors $X \hookrightarrow \mathcal{B}(p)$.

Type d'expériences aléatoires modélisées par la loi de Bernoulli

Cette loi permet de modéliser une expérience aléatoire à deux issues : succès (1) et échec (0), le paramètre p étant la probabilité de succès.

- ✘ Lien entre variable aléatoire de Bernoulli et indicatrice d'un événement. Pour tout $A \in \mathcal{T}$, on a $\mathbb{1}_A \hookrightarrow \mathcal{B}(\mathbf{P}(A))$.

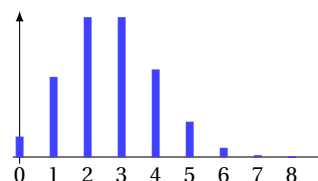
Définition 2.5. Loi binomiale

On dit qu'une variable X suit la loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0, 1]$ si

$$X(\Omega) = \llbracket 0, n \rrbracket, \quad \forall k \in \llbracket 0, n \rrbracket, \quad \mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

On note alors $X \hookrightarrow \mathcal{B}(n, p)$.

- ✘ Un modèle d'urne. Le nombre de boules noires obtenues lors d'un tirage avec remise à 8 tours dans une urne contenant 6 boules dont exactement deux noires suit la loi $\mathcal{B}(8, 1/3)$.



Plus généralement,

Type d'expériences aléatoires modélisées par la loi binomiale

Le nombre de succès lors de la répétition de n expériences de Bernoulli suivant la loi $\mathcal{B}(p)$ et indépendantes suit la loi $\mathcal{B}(n, p)$.

3. Couples de variables aléatoires, vecteurs aléatoires

Soit X et Y deux variables aléatoires définies sur le même espace probabilisé fini (Ω, \mathbf{P}) et à valeurs dans E et F . La connaissance des deux lois images \mathbf{P}_X et \mathbf{P}_Y ne donne pas d'information sur les liens qui peuvent unir les comportements aléatoires de X et de Y . Il est plus intéressant de considérer le couple $Z = (X, Y)$ comme une variable aléatoire à valeurs dans le produit cartésien $F \times G$.

Définition 2.6. Lois marginales d'un couple

Soit (Ω, \mathbf{P}) un espace probabilisé fini.

⇒ On appelle couple de variables aléatoires toute variable aléatoire à valeur dans un produit.

⇒ Soit Z une variable aléatoire à valeurs dans $E \times F$. Les variables aléatoires X et Y telles que $Z = (X, Y)$ ont des lois \mathbf{P}_X et \mathbf{P}_Y appelées lois marginales du couple (X, Y) .

Notation 2.7. Événements sur un couple

Pour deux variables aléatoires X et Y , on note souvent $(X \in A, Y \in B)$ au lieu de $(X \in A) \cap (Y \in B)$ afin d'alléger les expressions.

✘ On lance deux dés équilibrés numérotés 1 et 2; on note X_1 et X_2 les chiffres obtenus.

✓ Le couple (X_1, X_2) suit la loi uniforme sur $\llbracket 1, 6 \rrbracket^2$.

✓ On peut donner la loi d'un couple (X, Y) au moyen d'un tableau :

(X, Y)	$Y = y_1$	$Y = y_2$	$Y = y_3$...
$X = x_1$	$\mathbf{P}(X = x_1, Y = y_1)$	$\mathbf{P}(X = x_1, Y = y_2)$	$\mathbf{P}(X = x_1, Y = y_3)$...
$X = x_2$	$\mathbf{P}(X = x_2, Y = y_1)$	$\mathbf{P}(X = x_2, Y = y_2)$	$\mathbf{P}(X = x_2, Y = y_3)$...
$X = x_3$	$\mathbf{P}(X = x_3, Y = y_1)$	$\mathbf{P}(X = x_3, Y = y_2)$	$\mathbf{P}(X = x_3, Y = y_3)$...
\vdots	\vdots	\vdots	\vdots	\ddots

Il contient toute l'information sur la loi du couple (X, Y) et nous allons voir qu'il permet de reconstruire facilement les lois des composantes X et Y .

✘ Attention, en général, $(X, Y)(\Omega) \neq X(\Omega) \times Y(\Omega)$ On reprend l'exemple du lancé de deux dés développé ci-dessus. On a $(X_1, X_2)(\Omega) = X_1(\Omega) \times X_2(\Omega)$ mais, en notant $Y = X_1 + X_2$, on a $Y(\Omega) = \llbracket 2, 12 \rrbracket$ mais $(X_1, Y)(\Omega) \neq \llbracket 1, 6 \rrbracket \times \llbracket 2, 12 \rrbracket$ car $(6, 2) \notin (X_1, Y)(\Omega)$.

La connaissance de la loi du couple (X, Y) (on parle de la loi conjointe de X et Y), permet de trouver les lois marginales \mathbf{P}_X et \mathbf{P}_Y par simple sommation.

Proposition 2.8. Calcul des marginales à partir de la loi conjointe

Sous les hypothèses précédentes, on a $\forall x \in X(\Omega), \mathbf{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbf{P}(X = x, Y = y)$.

Le même type de formule vaut bien-sûr pour la loi de X .

✘ Attention, la réciproque de ce théorème est fautive : les lois marginales ne déterminent pas la loi conjointe. Considérons les deux couples Z et Z' dont les lois sont données par les tables ci-dessous :

(X, Y)	$Y = 0$	$Y = 1$
$X = 0$	1/4	1/4
$X = 1$	1/4	1/4

(X', Y')	$Y' = 0$	$Y' = 1$
$X' = 0$	1/2	0
$X' = 1$	0	1/2

Il est clair que Z et Z' ont les mêmes lois marginales mais pourtant $\mathbf{P}_Z \neq \mathbf{P}_{Z'}$.

La notion de couple se généralise sans peine à des n -listes.

Définition 2.9. Vecteurs aléatoires

On appelle vecteur aléatoire sur (Ω, \mathbf{P}) toute variable aléatoire à valeurs dans un produit. Un vecteur aléatoire admet donc des composantes X_1, \dots, X_n qui sont-elles mêmes des variables aléatoires sur Ω .

Les notions précédentes (lois marginales, etc.) se généralisent sans peine à ce cadre.

4. Variables indépendantes

Il est très intéressant d'étudier le cas où l'information que l'on possède sur une variable aléatoire X ne change rien à la loi d'une autre variable aléatoire Y , généralisant ainsi la notion d'indépendance introduite pour les événements aléatoires.

Définition 2.10. Couples de variables aléatoires indépendantes

Les variables aléatoires X et Y définies sur Ω sont dites indépendantes si, pour tout x de $X(\Omega)$ et tout y de $Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants. Cette propriété est notée $X \perp\!\!\!\perp Y$.

Il s'agit d'une notation hors programme mais que nous utiliserons à loisir.

✘ Loi de $X + Y$ avec X et Y indépendantes à valeurs dans \mathbb{N} . Notons $S = X + Y$. Pour tout $n \in \mathbb{N}$, on a

$$\mathbf{P}(S = n) = \sum_{k=0}^n \mathbf{P}(X = k, Y = n - k) = \sum_{k=0}^n \mathbf{P}(X = k) \mathbf{P}(Y = n - k)$$

On obtient ce qu'on appelle *un produit de convolution*.

✘ Soit X et Y deux variables aléatoires indépendantes à valeurs dans $\llbracket 1, n \rrbracket$. On suppose que les variables X et Y suivent la loi uniforme. Déterminer la loi de $Z := \max(X, Y)$. Pour $k \in \llbracket 0, n \rrbracket$. On a

$$\mathbf{P}(Z \leq k) = \mathbf{P}(X \leq k, Y \leq k) = \mathbf{P}(X \leq k) \mathbf{P}(Y \leq k) = \mathbf{P}(X \leq k)^2 = \frac{k^2}{n^2}$$

$$\text{d'où } \forall k \in \llbracket 1, n \rrbracket, \mathbf{P}(Z = k) = \mathbf{P}((Z \leq k) \setminus (Z \leq k - 1)) = \mathbf{P}(Z \leq k) - \mathbf{P}(Z \leq k - 1) = \frac{k^2 - (k - 1)^2}{n^2} = \frac{2k - 1}{n^2}.$$

✘ Le lecteur poursuivra par le test (**2.1**).

On étend la propriété de la définition à des événements plus généraux de la forme $(X \in A)$ et $(Y \in B)$.

Proposition 2.11. Caractérisation de l'indépendance

Soit X et Y des variables aléatoires définies sur le même espace probabilisé (Ω, \mathbf{P}) fini. Les propriétés suivantes sont équivalentes :

$\Rightarrow X \perp\!\!\!\perp Y$;

\Rightarrow Pour tout $A \subset X(\Omega)$ et tout $B \subset Y(\Omega)$, on a $\mathbf{P}((X, Y) \in A \times B) = \mathbf{P}(X \in A) \mathbf{P}(Y \in B)$.

La seconde propriété est équivalente à $(X \in A) \perp\!\!\!\perp (Y \in B)$ pour tout $A \subset X(\Omega)$ et $B \subset Y(\Omega)$.

Définition 2.12. Variables aléatoires mutuellement indépendantes

Les variables aléatoires X_1, \dots, X_n sont dites mutuellement indépendantes si, pour toute famille (x_1, \dots, x_n) de $X_1(\Omega) \times \dots \times X_n(\Omega)$, les événements $(X_i = x_i)$ sont mutuellement indépendants. On note de même $X_1 \perp\!\!\!\perp \dots \perp\!\!\!\perp X_n$.

Les prescriptions de vocabulaire et de notation sont les mêmes que pour les événements.

On ne précise pas toujours *mutuellement*, en revanche il faudra toujours préciser *deux à deux* si c'est le cas. Comme dans le cas des événements, l'indépendance entraîne l'indépendance deux à deux mais la réciproque est fautive (on peut reprendre le contre-exemple du cours sur les espaces probabilités et le reformulé en termes de variables aléatoires).

On généralise la caractérisation de l'indépendance à ce cadre.

Proposition 2.13. Caractérisation de l'indépendance

Des variables aléatoires X_1, \dots, X_n sont mutuellement indépendantes si et seulement si quel que soit $(A_1, \dots, A_n) \in \mathcal{P}(X_1(\Omega)) \times \dots \times \mathcal{P}(X_n(\Omega))$, les événements $(X_i \in A_i)$ sont mutuellement indépendants.

La démonstration de ce résultat est en tout point identique à celle du 4, les sommes doubles étant remplacées par des sommes multiples.

On modélise n expériences aléatoires indépendantes par une suite finie $(X_i)_{1 \leq i \leq n}$ de variables aléatoires indépendantes. Un cas particulier important est celui où on répète une même expérience.

Définition 2.14. Variables iid

Des variables indépendantes de même loi sont dites iid, *indépendantes et identiquement distribuées*.

Proposition 2.15. Sommes finies de variables de Bernoulli indépendantes

Soit X_1, \dots, X_n des var iid de loi $\mathcal{B}(p)$ où $n \in \mathbb{N}^*$. La variable $X_1 + \dots + X_n$ suit la loi $\mathcal{B}(n, p)$.

De l'existence de X_1, \dots, X_n va iid de loi $\mathcal{B}(p)$ où $n \in \mathbb{N}^*$.

Considérons le jeu fini de pile ou face (cf. le cours PB 1) à n lancers. Pour tout $i \in \llbracket 1, n \rrbracket$, variable X_i égale à 1 si le lancer $n^o i$ a donné pile, et 0 sinon. Les variables X_1, \dots, X_n sont indépendantes car pour tout $u \in \{0, 1\}^n$, les événements $(X_i = u_i)$ sont indépendants (cf. les numéros 1.30. et 1.31. du cours PB 1).

De nombreuses va se construisent comme fonction des termes d'une famille $(X_k)_{k \in \llbracket 1, n \rrbracket}$ de variables indépendantes (cf. la somme $X_1 + \dots + X_n$ de l'exemple précédent). Dans ce contexte, le théorème suivant sera fort commode.

Proposition 2.16. Lemme des coalitions

- Soit X et Y deux va indépendantes, f et g des applications définies respectivement sur $X(\Omega)$ et $Y(\Omega)$. Les va $f(X)$ et $g(Y)$ sont indépendantes.
- Soit $X_1, \dots, X_n, Y_1, \dots, Y_m$ des variables aléatoires indépendantes, f et g des applications définies respectivement sur $X_1(\Omega) \times \dots \times X_n(\Omega)$ et $Y_1(\Omega) \times \dots \times Y_m(\Omega)$. Les variables aléatoires $f(X_1, \dots, X_n)$ et $g(Y_1, \dots, Y_m)$ sont indépendantes.

Nous recommandons de finir ce paragraphe par le test (**2.2**).

5. Espérance et variance d'une variable aléatoire

Répétons n fois une expérience aléatoire d'univers fini, et notons X_1, \dots, X_n les valeurs successives prises par X . Pour avoir une idée du comportement de la variable X , il est naturel de considérer la moyenne arithmétique

$$M_n = \frac{X_1 + \dots + X_n}{n}$$

En regroupant suivant les différents valeurs possibles x_1, \dots, x_m de X , nous obtenons

$$M_n = \sum_{i=1}^m f_n(X = x_i) x_i$$

où $f_n(X = x_i)$ est la fréquence de réalisation de l'événement $(X = x_i)$ au cours des n expériences. Nous voulons faire tendre n vers l'infini. D'un point de vue intuitif, $f_n(X = x_i) \xrightarrow{n \rightarrow +\infty} \mathbf{P}(X = x_i)$ et donc

$$M_n \xrightarrow{n \rightarrow +\infty} \sum_{i=1}^m \mathbf{P}(X = x_i) x_i$$

Ce nombre, appelé *espérance mathématique* de X , est la limite des moyennes arithmétiques lorsque le nombre d'expériences tend vers l'infini. Nous justifierons cette assertion plus loin, dans l'un des théorèmes les plus importants de la théorie des probabilités, appelé *la loi des grands nombres*.

5.1. Espérance

L'espérance est une moyenne pondérée, on emploie parfois le terme *valeur moyenne* de X .

Définition 2.17. Espérance

Soit X une va réelle. On appelle espérance de X , notée $\mathbf{E}(X)$, le réel $\mathbf{E}(X) := \sum_{x \in X(\Omega)} x \mathbf{P}(X = x)$

On remarque que l'espérance d'une va X ne dépend que de la loi de X . Les résultats suivants sont à connaître sans aucune hésitation.

Proposition 2.18. Espérances des lois usuelles

On considère des variables aléatoires sur (Ω, \mathbf{P}) .

- a. Pour une X constante de valeur a , $\mathbf{E}(X) = a$.
- b. Pour tout événement A , on a $\mathbf{E}(\mathbb{1}_A) = \mathbf{P}(A)$.
- c. Pour $X \hookrightarrow \mathcal{B}(p)$, on a $\mathbf{E}(X) = p$.
- d. Pour $X \hookrightarrow \mathcal{B}(n, p)$, on a $\mathbf{E}(X) = np$.

Nous avons vu que la loi de $Y = f(X)$ n'est pas toujours simple à calculer. Le théorème suivant, dit *du transfert*, nous permet de faire l'économie de la loi de Y pour calculer $\mathbf{E}(Y)$ en cas d'existence.

Proposition 2.19. Théorème du transfert

Pour une va X et $f : X(\Omega) \rightarrow \mathbb{R}$, on a $\mathbf{E}(f(X)) = \sum_{x \in X(\Omega)} f(x) \mathbf{P}(X = x)$.

✕ Calculons $\mathbf{E}(Y)$ où $Y := \frac{1}{X+1}$ et $X \hookrightarrow \mathcal{B}(n, p)$ avec $p \in]0, 1[$. Par la formule du transfert :

$$\begin{aligned} \mathbf{E}(Y) &= \sum_{k=0}^n \frac{\binom{n}{k}}{k+1} p^k (1-p)^{n-k} = \frac{1}{n+1} \sum_{k=0}^n \binom{n+1}{k+1} p^k (1-p)^{n-k} \\ &= \frac{1}{(n+1)p} \sum_{k=0}^n \binom{n+1}{k+1} p^{k+1} (1-p)^{n+1-(k+1)} = \frac{1}{(n+1)p} \sum_{k=1}^{n+1} \binom{n+1}{k} p^k (1-p)^{n+1-k} \\ &= \frac{(p+1-p)^{n+1} - (1-p)^{n+1}}{p(n+1)} = \frac{1 - (1-p)^{n+1}}{p(n+1)} \end{aligned}$$

Comme nous allons le voir, l'espérance a les propriétés d'une intégrale¹ :

Proposition 2.20. Propriétés de l'espérance

Soit X et Y deux variables aléatoires et $\lambda \in \mathbb{R}$.

- a. Linéarité : $\mathbf{E}(\lambda X + Y) = \lambda \mathbf{E}(X) + \mathbf{E}(Y)$.
- b. Positivité : si $X \geq 0$, alors $\mathbf{E}(X) \geq 0$.
- c. Croissance : si $X \leq Y$, alors $\mathbf{E}(X) \leq \mathbf{E}(Y)$.
- d. Inégalité triangulaire : $|\mathbf{E}(X)| \leq \mathbf{E}(|X|)$.

1. Ce n'est pas un hasard et on peut pousser la comparaison plus loin qu'une simple analogie : il existe un cadre plus général permettant d'unifier ces notions *a priori* différentes.

Proposition 2.21. Produit de deux variables indépendantes

Si X et Y sont deux variables aléatoires indépendantes, alors $\mathbf{E}(XY) = \mathbf{E}(X) \mathbf{E}(Y)$.

- ✘ La réciproque est fautive. Soit X une variable vérifiant $X(\Omega) = \{-1, 0, 1\}$ et suivant la loi uniforme sur cet ensemble. On pose $Y = |X|$. On a bien $\mathbf{E}(XY) = 0 = \mathbf{E}(X) \mathbf{E}(Y)$ mais

$$\mathbf{P}(X = 0, Y = 0) = \mathbf{P}(X = 0) = \frac{1}{3} \neq \mathbf{P}(X = 0) \mathbf{P}(Y = 0) = \frac{1}{9}$$

ainsi X et Y ne sont pas indépendantes.

- ✘ Démonstration de la formule de Poincaré : pour des événements A_1, \dots, A_n , on a

$$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{\substack{\Lambda \subset [1, n] \\ |\Lambda|=k}} \mathbf{P}\left(\bigcap_{\ell \in \Lambda} A_\ell\right)$$

Notons $B := \bigcup_{i=1}^n A_i$. Comme $\bar{B} = \bigcap_{i=1}^n \bar{A}_i$, on a

$$\mathbb{1}_{\bar{B}} = \prod_{i=1}^n \mathbb{1}_{\bar{A}_i} = \prod_{i=1}^n (1 - \mathbb{1}_{A_i}) = \sum_{\Lambda \subset [1, n]} \prod_{\ell \in \Lambda} -\mathbb{1}_{A_\ell} = \sum_{k=0}^n \sum_{\substack{\Lambda \subset [1, n] \\ |\Lambda|=k}} (-1)^k \mathbb{1}_{\bigcap_{\ell \in \Lambda} A_\ell}$$

Ainsi, par linéarité de l'espérance :

$$\mathbf{P}(B) = 1 - \mathbf{P}(\bar{B}) = 1 - \mathbf{E}(\mathbb{1}_{\bar{B}}) = 1 - \sum_{k=0}^n \sum_{\substack{\Lambda \subset [1, n] \\ |\Lambda|=k}} (-1)^k \underbrace{\mathbf{E}(\mathbb{1}_{\bigcap_{\ell \in \Lambda} A_\ell})}_{=\mathbf{P}(\bigcap_{\ell \in \Lambda} A_\ell)} = \sum_{k=1}^n (-1)^{k-1} \sum_{\substack{\Lambda \subset [1, n] \\ |\Lambda|=k}} \mathbf{P}\left(\bigcap_{\ell \in \Lambda} A_\ell\right)$$

- ✘ Le lecteur abordera le test (❗ 2.3) avec profit.

5.2. Variance, écart-type et covariance

L'écart-type est une grandeur qui mesure la moyenne (en un certain sens) de l'écart des valeurs de X à sa moyenne.

Définition 2.22. Variance, covariance et écart type

Soit X et Y deux va.

⇒ On appelle variance de X le réel $\mathbf{V}(X) := \mathbf{E}((X - \mathbf{E}(X))^2)$ et écart type de X , le réel $\sigma(X) = \sqrt{\mathbf{V}(X)}$.

⇒ On appelle covariance de X et Y le réel $\mathbf{Cov}(X, Y) := \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))$.

Ces nombres sont indicateurs de dispersion de X par rapport à sa moyenne.

Proposition 2.23. Propriétés de la variance

Soit X une variable aléatoire réelle.

- a. Relation de Huygens : $V(X) = E(X^2) - E(X)^2$.
- b. $\forall (a, b) \in \mathbb{R}^2, V(aX + b) = a^2 V(X)$.

Définition 2.24. Moments d'une variable aléatoire.

Le moment d'ordre r de X est par définition égal à $E(X^r)$.

Nous retrouvons dans ce contexte l'inégalité de Cauchy-Schwarz.

Proposition 2.25. Inégalité de Cauchy-Schwarz

Pour X et Y variables aléatoires réelles, on a $E(XY)^2 \leq E(X^2) E(Y^2)$.

Proposition 2.26. Propriétés de la covariance

Soient X et Y deux variables aléatoires réelles.

- a. Relation de Huygens : $\text{Cov}(X, Y) = E(XY) - E(X) E(Y)$.
- b. Inégalité de Cauchy-Schwarz : $\text{Cov}(X, Y)^2 \leq V(X) V(Y)$.
- c. Pour des variables aléatoires X_1, \dots, X_n , on a $V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$.
- d. Si de plus X_1, \dots, X_n sont deux à deux indépendantes, alors $V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i)$.

Proposition 2.27. Variance des lois usuelles

- a. Pour X constante, $V(X) = 0$.
- b. Pour $X \hookrightarrow \mathcal{B}(p)$, on a $V(X) = p(1 - p)$.
- c. Pour $X \hookrightarrow \mathcal{B}(n, p)$, on a $V(X) = np(1 - p)$.

On pourra aborder le test (2.4).

5.3. Inégalités de Markov et de Tchebychev

Les inégalités probabilistes sont nombreuses et particulièrement importantes pour démontrer des propriétés asymptotiques. Nous nous contenterons ici des deux plus élémentaires.

Proposition 2.28. Inégalité de Markov

Pour toute variable aléatoire Y , on a $\forall a > 0$, $\mathbf{P}(|Y| \geq a) \leq \frac{\mathbf{E}(|Y|)}{a}$.

Il faut savoir appliquer avec une certaine souplesse cette inégalité.

✕ Variations sur l'inégalité de Markov. Soit X une variable aléatoire et $a > 0$. Si $f : \mathbb{R} \rightarrow \mathbb{R}_+^*$ est croissante, on a alors l'inclusion $(X \geq a) \subset (f(X) \geq f(a))$ et donc

$$\mathbf{P}(X \geq a) \leq \mathbf{P}(f(X) \geq f(a)) \leq \frac{\mathbf{E}(f(X))}{f(a)}$$

Ainsi, par exemple, $\mathbf{P}(X \geq a) \leq \frac{\mathbf{E}(\exp(X))}{\exp(a)}$.

Proposition 2.29. Inégalité de Bienaymé-Tchebychev

Pour toute variable aléatoire X , on a $\forall a > 0$, $\mathbf{P}(|X - \mathbf{E}(X)| \geq a) \leq \frac{\mathbf{V}(X)}{a^2}$.

On poursuivra avec le test (**2.5**).

6. Les grands résultats asymptotiques

Les théorèmes asymptotiques sont les sommets de la théorie des probabilités, tout particulièrement *les lois des grands nombres et le théorème central-limite*.

6.1. La loi faible des grands nombres

Nous avons à de nombreuses reprises motivé les définitions et les modèles choisis par des considérations heuristiques. *La loi des grands nombres* montre rigoureusement que, quand le nombre de répétitions de l'expérience tend vers l'infini, la fréquence de réalisations d'un événement converge vers la probabilité de réalisation de cet événement. Ainsi, notre modèle est bien cohérent l'intuition.

Ce résultat a également un intérêt en analyse numérique; il est à l'origine de méthodes de calcul numérique appelées *Méthodes de Monte-Carlo*, qui sont extrêmement puissantes et robustes. Elles sont par exemple très utilisées en Physique.

Proposition 2.30. Loi faible des grands nombres

Si $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires indépendantes admettant une espérance μ et une variance σ^2 alors, pour tout $\varepsilon > 0$:

$$\mathbf{P}\left(\left|\frac{1}{n} \sum_{k=1}^n X_k - \mu\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow +\infty} 0$$

La loi forte des grands nombres

L'existence d'une suite de n indépendantes impose que l'univers Ω est infini, la démonstration qui précède est donc une anticipation du cours de spé (d'où les hypothèses d'existence des deux moments). Sous les mêmes hypothèses que la loi faible, on peut démontrer un résultat beaucoup plus précis, *la loi forte des grands nombres* :

$$\mathbf{P}\left(\left\{\omega \in \Omega; \frac{1}{n} \sum_{k=1}^n X_k(\omega) \xrightarrow{n \rightarrow +\infty} \mu\right\}\right) = 1$$

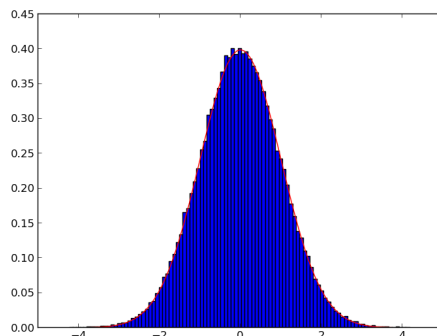
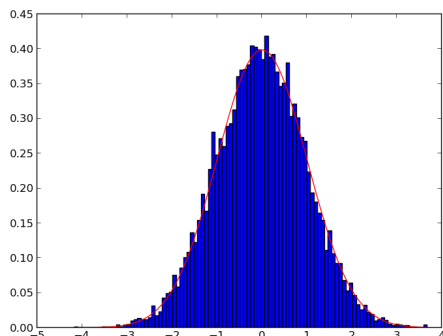
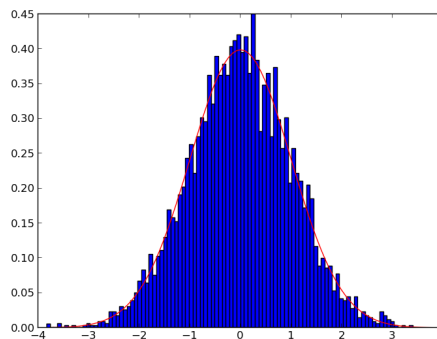
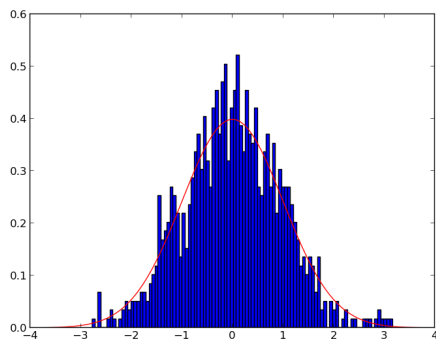
6.2. Le théorème « limit-centrale »

Proposition 2.31. Théorème limit-centrale

Soit (X_n) une suite de variables aléatoires discrètes définies sur le même espace probabilisé, indépendantes et identiquement distribuées telle que X_1 admet un moment d'ordre deux. En notant σ l'écart-type de X_1 et μ l'espérance de X_1 , on a

$$\forall (a, b) \in \mathbb{R}^2, \mathbf{P}\left(a \leq \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq b\right) \xrightarrow{n \rightarrow +\infty} \int_a^b \exp\left(-\frac{u^2}{2}\right) \frac{du}{\sqrt{2\pi}} \quad \text{où } \bar{X}_n := \frac{X_1 + \dots + X_n}{n}$$

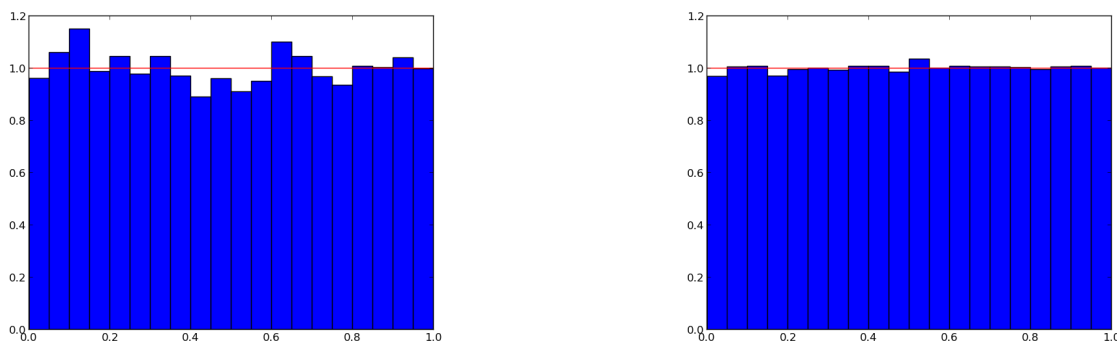
Dans les figures ci-dessous, le lecteur trouvera une illustration de ce théorème au moyen de simulations effectuées sous Python.



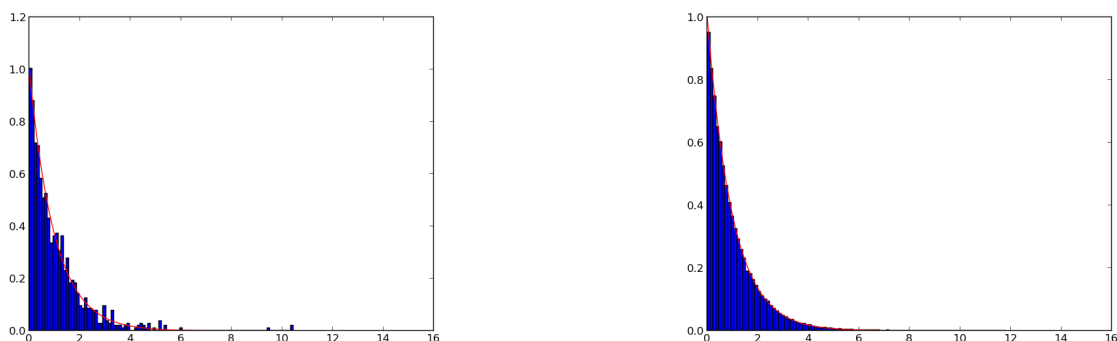
7. Passage du discret au continu

Nous terminerons par un bref exposé sur le cas des *variables aléatoires continues*. Essayons d'abord d'appréhender ce que peut être la loi d'une variable non discrète.

Considérons une variable aléatoire X à valeurs dans $[0, 1]$, comme par exemple le temps d'attente (en heure) d'un bus à une station. Pour tout $n \in \mathbb{N}$ et tout $\omega \in \Omega$, on note $X_n(\omega)$ la variable égale à l'unique indice $k \in \llbracket 0, n-1 \rrbracket$ tel que $X(\omega) \in [k/n, (k+1)/n[$. La variable X_n est discrète et on peut se faire une idée de la loi de X en traçant les histogrammes de la distribution de X_n . En effectuant quelques relevés pour des valeurs de n croissantes, on observe ceci :



On semble observer une convergence vers une distribution uniforme. En essayant de reproduisant cette expérience avec une variable Y , on observe le même phénomène de convergence.



En notant f cette fonction, on s'achemine donc vers un modèle tel que

$$\mathbf{P}(a \leq X \leq b) = \int_a^b f(t) dt$$

car en sommant les aires des histogrammes on obtient une approximation de l'aire sous la courbe de f sur $[a, b]$ mais également une approximation de la probabilité de l'événement $X \in [a, b]$. Cette fonction f est appelée *densité de X* et caractérise la loi de X . La théorie dépasse cependant de loin les outils de classes préparatoires (il faut une théorie de l'intégrale un peu plus maniable que la nôtre pour manipuler rigoureusement ce type d'objets).

8. Énoncés des tests

2.1.

Soit X et Y deux variables aléatoires réelles indépendantes. Les variables aléatoires $X + Y$ et $X - Y$ sont-elles indépendantes ?

2.2.

Soit X et Y deux variables aléatoires prenant pour valeurs a_1, \dots, a_n avec $\mathbf{P}(X = a_i) = \mathbf{P}(Y = a_i) = p_i$. On suppose que les variables X et Y sont indépendantes. Montrer que

$$\mathbf{P}(X \neq Y) = \sum_{i=1}^n p_i(1 - p_i)$$

2.3.

Soit $p \in]0, 1[$ et (X_k) une suite de variables aléatoires mutuellement indépendantes vérifiant $\mathbf{P}(X_k = 1) = p$ et $\mathbf{P}(X_k = -1) = 1 - p$.

- a. Calculer l'espérance de X_k .
- b. On pose $Y_n = \prod_{k=1}^n X_k$. En calculant de deux façons l'espérance de Y_n , déterminer $p_n = \mathbf{P}(Y_n = 1)$.
- c. Quelle est la limite de p_n quand $n \rightarrow +\infty$?

2.4.

Soit X une variable aléatoire à valeurs dans $\llbracket 0, n \rrbracket$ telle qu'il existe $a \in \mathbb{R}$ vérifiant $\mathbf{P}(X = k) = a \binom{n}{k}$. Calculer l'espérance et la variance de X .

2.5.

Soit X une variable aléatoire d'espérance μ et de variance σ^2 . En introduisant la variable aléatoire

$$Y = (\alpha(X - \mu) + \sigma)^2$$

Montrer que pour tout $\alpha > 0$, $\mathbf{P}(X \geq \mu + \alpha\sigma) \leq \frac{1}{1 + \alpha^2}$.

9. Solutions des tests

2.1. _____

Non en général. Supposons que X et Y suivent la loi $\mathcal{B}(1/2)$. On a

$$\mathbf{P}(X+Y=2) = \mathbf{P}(X=1)\mathbf{P}(Y=1) = 1/4$$

et

$$\begin{aligned} \mathbf{P}(X-Y=0) &= \mathbf{P}(X=0)\mathbf{P}(Y=0) + \mathbf{P}(X=1)\mathbf{P}(Y=1) \\ &= 1/2 \end{aligned}$$

Comme $(X+Y=2) \subset (X-Y=0)$, on a

$$\mathbf{P}(X+Y=2, X-Y=0) = \mathbf{P}(X+Y=2)$$

et l'on constate que

$$\mathbf{P}(X+Y=2, X-Y=0) \neq \mathbf{P}(X+Y=2)\mathbf{P}(X-Y=0)$$

2.2. _____

Comme

$$(X=Y) = \bigsqcup_{i=1}^n (X=a_i) \cap (Y=a_i)$$

donc, par indépendance de X et Y ,

$$\mathbf{P}(X=Y) = \sum_{i=1}^n p_i^2$$

ainsi

$$\mathbf{P}(X \neq Y) = 1 - \sum_{i=1}^n p_i^2$$

On conclut sachant que $1 = \sum_{i=1}^n p_i$.

2.3. _____

a. $\mathbf{E}(X_k) = 1 \times p + (-1) \times (1-p) = 2p-1$.

b. Par l'indépendance des variables

$$\mathbf{E}(Y_n) = \prod_{k=1}^n \mathbf{E}(X_k) = (2p-1)^n$$

Comme $Y_n \in \{1, -1\}$ et

$$\mathbf{E}(Y_n) = 1 \times p_n + (-1) \times (1-p_n) = 2p_n - 1$$

On en déduit que

$$p_n = \frac{1 + (2p-1)^n}{2}$$

c. Puisque $|p| < 1$, $p_n \xrightarrow[n \rightarrow +\infty]{} 1/2$.

2.4. _____

En écrivant que

$$\sum_{k=0}^n \mathbf{P}(X=k) = 1$$

on obtient $a = 1/2^n$. Ainsi, $X \hookrightarrow \mathcal{B}(n, 1/2)$ et

$$\mathbf{E}(X) = \frac{n}{2}, \quad \mathbf{V}(X) = \frac{n}{4}$$

2.5. _____

On a

$$\mathbf{E}(Y) = (\alpha^2 + 1)\sigma^2$$

L'inégalité de Markov appliquée à la variable positive Y donne

$$\mathbf{P}(Y \geq a) \leq \frac{\mathbf{E}(Y)}{a}$$

Pour $a = \sigma^2(\alpha^2 + 1)2$,

$$\mathbf{P}(Y \geq a) \leq \frac{1}{1 + \alpha^2}$$

Or $(X \geq \mu + \alpha\sigma) = (\alpha(X - \mu) + \sigma \geq (\alpha^2 + 1)\sigma)$ et donc $(X \geq \mu + \alpha\sigma) \subset (Y \geq a)$ puis

$$\mathbf{P}(X \geq \mu + \alpha\sigma) \leq \frac{1}{1 + \alpha^2}$$