Variables aléatoires finies

« Il y a en chacun de nous des calculs que nous nommons espérance. »
Platon

Notations : Dans ce chapitre, Ω représente un ensemble fini, et $\mathbf{P}$ une probabilité sur Ω.

1 Notion de variable aléatoire réelle

A Définition

Exemple *(Exemple suivi...)*

On joue aux dés. Chaque joueur lance deux dés et le gagnant est celui dont la somme des deux dés est la plus grande.
La modélisation naturelle consiste à considérer l'univers constitué de l'ensemble des couples de valeurs possibles et à le munir d'une probabilité uniforme $\mathbf{P}$.

$$\Omega = [\![1,6]\!]^2$$

Cependant, ce n'est pas le couple (dé 1, dé 2) qui nous intéresse, mais simplement sa somme. L'idée est donc de « transformer » l'univers Ω en un autre plus adapté à la question. Pour cela, on construit l'application $X : \Omega \to [\![2,12]\!]$ qui associe à chaque issue possible, la somme des dés correspondante.

$$X : \left\{ \begin{array}{lcl} \Omega = [\![1,6]\!]^2 & \to & [\![2,12]\!] \\ \omega = (i,j) & \mapsto & i+j \end{array} \right.$$

On dit que X est une variable aléatoire, et le nouvel univers de travail, beaucoup plus simple, est $[\![2,12]\!]$.
De tous les événements possibles, seuls nous intéressent, pour $k \in [\![2,12]\!]$,

$$A_k : \text{« la somme des dés vaut } k \text{ ».}$$

Ainsi, parmi les 2^{36} événements possibles, nous n'en retenons que 11.
Pour décrire l'événement A_k dans Ω, il faut chercher les antécédents de k par X.
Par exemple, pour A_3, on cherche tous les couples $(i,j) \in \Omega$ tels que $X(i,j) = 3$.
On trouve l'événement $A_3 = \big\{(1,2),(2,1)\big\}$.
et on en déduit la probabilité de l'événement A_3 :

$$\mathbf{P}(A_3) = \frac{2}{36} = \frac{1}{18}$$

On notera cette probabilité $\mathbf{P}(X = 3)$.

À présent, formalisons cette notion :

Définition 1.1 *(Variable aléatoire réelle finie)*

Une **variable aléatoire** réelle finie discrète X est une application de Ω dans une partie finie $\mathcal{X} \subset \mathbf{R}$.

Remarque : L'univers étant supposé fini discret, on peut écrire $\mathcal{X}$ sous la forme

$$\mathcal{X} = \{x_i,\ i \in I\}$$

avec I une partie finie de $\mathbf{N}$ ou $\mathbf{Z}$. Par exemple $I = [\![1, n]\!]$.

Explications

La variable aléatoire permet de traduire le résultat d'une expérience dans un ensemble sur lequel on sait travailler : on a plus de facilités à manipuler des nombres que des couples de dés ou des paires de chaussettes de différentes couleurs.

⚠ Le vocabulaire consacré par l'usage est assez troublant : on parle de variable aléatoire, bien qu'il s'agisse en fait d'une *application.* Ces difficultés de notation et de langage sont omniprésentes en probabilités, mais elles se justifient par l'usage.

Notation

Soit X une variable aléatoire réelle dans $\mathcal{X} \subset \mathbf{R}$.
Pour $\mathcal{X}'$ une partie de $\mathcal{X}$, on note $[X \in \mathcal{X}']$ l'événement

$$[X \in \mathcal{X}']\,\{\omega, = \text{ tel que } X(\omega) \in \mathcal{X}'\}$$

Ainsi,

$$\mathbf{P}(X \in \mathcal{X}') \ = \ \mathbf{P}(\{\omega,\ \text{tel que } X(\omega) \in \mathcal{X}'\})$$

Lorsque $\mathcal{X}' = \{x\}$ est un singleton, on note plus simplement $[X = x]$:

$$[X = x] = \{\omega,\ \text{tel que } X(\omega) = x\}$$

Ainsi

$$\mathbf{P}(X = x) \ = \ \mathbf{P}(\{\omega,\ \text{tel que } X(\omega) = x\})$$

Explications

Plutôt que d'étudier la réalisation d'un événement $A \subset \Omega$, on préfère chercher les événements qui réalisent certaines valeurs de X.

Exemple *(Exemple suivi)*

Avec l'exemple du lancer de deux dés (équilibrés), on définit la variable aléatoire X qui donne leur somme.
Calculer $\mathbf{P}(X = 2)$, $\mathbf{P}(X \in \{3\})$ et $\mathbf{P}(X \in \{2, 3\})$.
Solution :

⚠ Si $\mathcal{X}' = \{x\}$ est un singleton, il n'y a aucune raison pour que l'événement correspondant dans Ω soit lui-même un singleton.
Rappelez vous l'exemple introductif : A_3 n'est pas un singleton.

Propriété 1.2 *(Loi de probabilité de X)*

Pour X une variable aléatoire $\Omega \to X(\Omega)$,
On appelle **loi de probabilité** de X, l'application

$$f_X \ : \begin{cases} X(\Omega) & \to [0,1] \\ x & \mapsto \mathbf{P}(X = x) \end{cases}$$

Remarque : f_X définit une probabilité $\mathbf{P}_X$ sur $X(\Omega)$ avec

$$\forall x \in X(\Omega),\ \mathbf{P}_X(\{x\}) = f_X(x).$$

Il est aussi possible de travailler sur un ensemble plus gros $\mathcal{X}$ qui contient $X(\Omega)$. Auquel cas, tous les événements supplémentaires seront de probabilité nulle.

Explications

La détermination de la probabilité $\mathbf{P}_X$ permet d'oublier complètement l'univers naturel pour ne considérer que $X(\Omega)$.

Preuve

Il faut prouver que $\mathbf{P}_X$ est bien une probabilité.
Pour cela on peut voir que

$$\forall x \in X(\Omega),\ \mathbf{P}_X(\{x\}) = f_X(x) = \mathbf{P}(X = x) = \mathbf{P}(\{\omega \in \Omega,\ X(\omega) = x\}) \in [0, 1]$$

(car $\{\omega \in \Omega,\ X(\omega) = x\}$ est un événement de Ω et $\mathbf{P}$ est une probabilité sur Ω)

$$\sum_{x \in X(\Omega)} \mathbf{P}_X(\{x\}) = \sum_{x \in X(\Omega)} f_X(x) = \sum_{x \in X(\Omega)} \mathbf{P}(X = x) = 1$$

Donc $\mathbf{P}_X$ est bien une probabilité (définie par ses événements élémentaires).
Par abus de langage, en identifiant $\mathbf{P}_X(\{x\})$ avec f_X, on peut dire que f_X représente une probabilité sur $X(\Omega)$. ■

Exemple *(Exemple suivi)*

Donner la loi de X pour la somme des deux dés.
Solution :

Exemple

Une variable aléatoire sur $[\![1,n]\!]$ suit une **probabilité uniforme** si

$$\forall k \in [\![1,n]\!],\ f_X(k) = \mathbf{P}(X=k) = \frac{1}{n}.$$

On remarque ici que seul l'ensemble $X(\Omega) = [\![1,n]\!]$ est important, on ne s'occupe plus du tout de l'univers Ω sous-jacent.

L'équiprobabilité sur $X(\Omega)$ n'est pas liée à celle sur Ω : si la variable suit une probabilité uniforme sur $X(\Omega)$, cela ne veut pas dire que l'expérience est équiprobable sur Ω. Et réciproquement, si l'expérience est équiprobable sur Ω, cela ne veut pas dire que la variable aura une probabilité uniforme sur $X(\Omega)$. Par exemple, pour la somme des deux dés, la probabilité est uniforme sur Ω, mais ne l'est pas dans $[\![2,12]\!]$.

Propriété 1.3

Soit X une variable aléatoire finie de Ω dans $\mathcal{X}$,
L'ensemble des événements $\{[X=x],\ x\in\mathcal{X}\}$ forme une famille complète d'événements de Ω.

Preuve

- Tout d'abord, tout $\omega\in\Omega$ admet une image par X dans $\mathcal{X}$: $\omega \in [X = X(\omega)]$, donc $\omega \in \bigcup\limits_{x\in\mathcal{X}} [X=x]$.
  Ceci étant vrai pour tous les $\omega\in\Omega$, on a bien $\Omega \subset \bigcup\limits_{x\in\mathcal{X}} [X=x]$.
  L'autre inclusion étant évidente, $\Omega = \bigcup\limits_{x\in\mathcal{X}} [X=x]$.
- Soient $(x,y) \in (X(\Omega))^2$.
  Si on suppose que $\omega \in [X=x]\cap[X=y]$, alors $X(\omega)=x$ et $X(\omega)=y$, donc $x=y$.
  Ainsi, pour $x\neq y$, $[X=x]\cap[X=y]=\emptyset$.

L'ensemble $\{[X=x],\ x\in\mathcal{X}\}$ forme donc bien une famille complète d'événements de Ω. ■

Propriété 1.4

Soit X une variable aléatoire finie de Ω dans $\mathcal{X}$, et $\mathcal{X}'$ une partie de $\mathcal{X}$.

$$\mathbf{P}(X\in\mathcal{X}') = \sum_{x\in\mathcal{X}'} \mathbf{P}(X=x) = \sum_{x\in\mathcal{X}'} f_X(x)$$

Explications

On écrit B comme union d'événements élémentaires.

B Fonction de répartition d'une variable aléatoire réelle

Cette notion est surtout utile pour les variables aléatoires continues qu'on n'étudiera pas cette année, mais elle peut parfois simplifier les raisonnements avec les variables aléatoires discrètes.
La fonction de répartition correspond tout simplement aux « probabilités cumulées croissantes » :

Définition 1.5 *(Fonction de répartition)*

Soit X une variable aléatoire réelle sur Ω.
On définit la **fonction de répartition** F_X sur $\mathbf{R}$ par

$$F_X : \left\{ \begin{array}{ccl} \mathbf{R} & \to & \mathbf{R} \\ x & \mapsto & \mathbf{P}(X\leq x) = \sum\limits_{x_i\leq x} \mathbf{P}(X=x_i) \end{array} \right\}$$

Remarque : Avec des probabilités continues, on prendrait l'intégrale qui est l'équivalent continu de la somme.

Exemple *(Exemple suivi)*

Donner la fonction de répartition pour la variable X correspondant à la somme des deux dés.

Solution :

Propriété 1.6

Une fonction de répartition sur $\mathbf{R}$ est positive, croissante et

$$\lim_{x\to-\infty} F_X(x) = 0 \qquad \text{et} \qquad \lim_{x\to+\infty} F_X(x) = 1$$

Preuve

Soit $x \leq y$, alors $[X \leq x] \subset [X \leq y]$, et par croissance de la probabilité $\mathbf{P}(X \leq x) \leq \mathbf{P}(X \leq y)$, donc $F_X(x) \leq F_X(y)$: la fonction F_X est croissante.
Si $X(\Omega)$ est fini, alors il admet un minimum x_m et un maximum x_M.
$\forall x < x_m,\ F_X(x) = 0$, donc $\lim\limits_{x\to-\infty} F_X(x) = 0$. $\forall x \geq x_M,\ F_X(x) = 1$, donc $\lim\limits_{x\to+\infty} F_X(x) = 1$. ■

Propriété 1.7

F_X décrit parfaitement la loi de X : la donnée de F_X permet de retrouver la probabilité.
Lorsque les probabilités sont à valeurs entières,

$$\forall k \in \mathbf{Z},\ \mathbf{P}(X = k) = \mathbf{P}(X \leq k) - \mathbf{P}(X \leq k-1) = F_X(k) - F_X(k-1)$$

Définition 1.8 *(Quantiles)*

Soit X une variable aléatoire et F_X sa fonction de répartition.
Pour $r \in]0,1[$, on appelle **quantile d'ordre** r l'unique valeur $q_r \in \mathcal{X}$ définie par

$$q_r = \max\left(x_i \in \mathcal{X},\ \text{tel que } F(x_i) \leq r\right)$$

Le plus souvent, on définit les quantiles pour des valeurs rationnelles

Définition 1.9

- Le quantile d'ordre $\frac{1}{2}$, $q_{1/2}$ est appelé la **médiane** de X.
- Le quantile d'ordre $\frac{1}{4}$, $q_{1/4}$ est appelé le premier **quartile** de X.
- Le quantile d'ordre $\frac{1}{10}$, $q_{1/10}$ est appelé le premier **décile** de X.

On définit de la même façon le troisième quartile $q_{3/4}$ par exemple.
Le second quartile est exactement la médiane. De même que le cinquième décile.

Remarque : La définition des quantiles n'est pas tout à fait unifiée dans la littérature pour les probabilités discrètes.
Prenons l'exemple de la médiane pour comprendre la difficulté.
Comme la fonction de répartition augmente « par sauts », on passe souvent d'une valeur strictement inférieure à $1/2$ à une valeur strictement supérieure.
Par exemple $F_X(x_3) = 0,43$ et $F_X(x_4) = 0,52$.
Le passage se fait entre x_3 et x_4, et on serait tenté de placer la médiane entre les deux. Mais à quelle valeur ? Plusieurs choix sont possibles.
Ce cours adopte la solution la plus simple avec $q_{1/2} = x_3$: c'est la dernière valeur avant le passage des 50%. C'est le choix le plus fréquent dans notre situation et il permet d'avoir un quantile qui fait partie de $\mathcal{X}$.
Cependant, il faut savoir que la plupart des logiciels (plutôt pour les statistiques) réalise une interpolation linéaire entre x_3 et x_4 pour placer la médiane. Mais à nouveau, cette règle n'est pas unifiée et tous les logiciels ne donneront donc pas la même valeur.
Dans les situations concrètes, ces choix sont de faible importance car les quantiles ne sont utiles que si l'univers image $X(\Omega)$ contient suffisamment de valeurs pour que l'écart entre les différents choix proposés soit négligeable.

Explications

La médiane permet de partager l'univers en deux événements dont les probabilités sont les plus proches possibles de $\frac{1}{2}$; tout en préservant l'ordre des issues élémentaires.

C Image d'une variable aléatoire par une fonction

Propriété 1.10

Si X est une variable aléatoire de Ω dans $\mathcal{X}$ et $u : \mathcal{X} \to \mathcal{X}'$ une application, alors $u(X)$ définit une nouvelle variable aléatoire sur Ω.

Exemple

Dans le cas d'un double lancer de dés, on peut par exemple associer un gain à chaque résultat.
On note X la variable aléatoire correspondant à somme des valeurs des deux dés, et u l'application qui définit le gain en fonction du résultat.

$$u : \left\{ \begin{array}{ccc} [\![2,12]\!] & \to & \mathbf{R} \\ x & \mapsto & \left\{ \begin{array}{ll} 10 & \text{si } x = 12 \\ 1 & \text{si } x \in [\![9,11]\!] \\ -1 & \text{sinon} \end{array} \right. \end{array} \right.$$

On trouve donc $\mathbf{P}(u(X) = 10) = \mathbf{P}(X = 12) = \dfrac{1}{36}$

$\mathbf{P}(u(X) = 1) = \mathbf{P}(X \in [\![9,11]\!]) = \dfrac{4}{36} + \dfrac{3}{36} + \dfrac{2}{36} = \dfrac{9}{36} = \dfrac{1}{4}$

$\mathbf{P}(u(X) = -1) = 1 - \mathbf{P}(u(X) \neq -1) = 1 - \dfrac{1}{36} - \dfrac{9}{36} = \dfrac{26}{36} = \dfrac{13}{18}$

On peut alors construire le tableau avec les nouvelles probabilités.

y	-1	1	10
$f_{u(X)}(y) = \mathbf{P}(u(X) = y)$	$\dfrac{26}{36}$	$\dfrac{9}{36}$	$\dfrac{1}{36}$

Explications

Cela consiste à refaire exactement ce que nous avions fait en début de chapitre avec X, mais au lieu de passer de l'espace probabilisé « naturel » $(\Omega, \mathbf{P})$ dans $(\mathcal{X}, \mathbf{P}_X)$, on passe cette fois-ci de l'univers $(\mathcal{X}, \mathbf{P}_X)$ dans l'univers $(\mathcal{X}', \mathbf{P}_{u \circ X})$.
Cela revient à considérer la variable aléatoire $u \circ X$ au lieu de X.

2 Espérance et écart type

A Indicateur de position

Définition 2.1 *(Espérance d'une variable aléatoire et théorème de transfert)*

Soit X une variable aléatoire finie de Ω dans $\mathbf{R}$.
L'espérance de X est définie par

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x)$$

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

L'espérance est parfois aussi appelée la *moyenne*, elle représente simplement la valeur moyenne de X sur un très grand nombre d'expériences.
Le terme d'**espérance** fait référence à la théorie des jeux. Il représente le gain moyen que je peux espérer après un grand nombre d'essais.

Dans la pratique, pour calculer l'espérance, on utilise souvent la deuxième expression.
En effet, elle a l'avantage de ne pas faire intervenir l'univers initial explicitement, mais uniquement l'univers image.

Preuve *(Égalité des deux formulations)*

$$\begin{aligned} \sum_{\omega \in \Omega} X(\omega)\mathbf{P}(\{\omega\}) &= \sum_{x \in X(\Omega)} \sum_{\omega \in X^{-1}(x)} X(\omega)\mathbf{P}(\{\omega\}) \\ &\text{car les événements } [X = x] \text{ forment une famille complète d'événements.} \\ &= \sum_{x \in X(\Omega)} \sum_{\omega \in X^{-1}(x)} x\mathbf{P}(\{\omega\}) \\ &= \sum_{x \in X(\Omega)} x\mathbf{P}(\omega \in X^{-1}(x)) \quad \text{car les événements sont disjoints} \\ &= \sum_{x \in X(\Omega)} x\mathbf{P}(X = x) \end{aligned}$$

■

Exemple

Calculer la valeur moyenne de la somme lorsqu'on lance deux dés équilibrés.
Calculer l'espérance du gain (voir exercice page 5).
Solution :

Exemple *(La roulette)*

La roulette est composée de 37 numéros : $[\![0, 36]\!]$.
Faire un pari simple consiste à parier sur la moitié des numéros sauf le 0.
Par exemple, vous pariez sur les rouges.
Si un rouge sort, vous gagnez une fois votre mise, si un noir sort, vous perdez votre mise, et si le zéro sort, vous perdez la moitié de la mise (pour simplifier).
Calculer l'espérance de gain.

Solution :

Exemple

Donner l'espérance d'une variable uniforme sur $[\![1, n]\!]$.

Solution :

Propriété 2.2 *(Propriétés élémentaires)*

X et Y désignent deux variables aléatoires réelles.

1. L'espérance est **linéaire** : $\forall (a,b) \in \mathbf{R}^2,\ \mathbf{E}(aX+bY) = a\mathbf{E}(X) + b\mathbf{E}(Y)$
2. Si $\exists a \in \mathbf{R},\ \forall \omega \in \Omega,\ X(\omega) = a,$ alors $\mathbf{E}(X) = a$.
   L'espérance d'une variable constante est égale à cette constante.
3. $\left|\mathbf{E}(X)\right| \leq \mathbf{E}\left(\left|X\right|\right)$ (inégalité triangulaire)
4. L'espérance est **positive** : si $\forall \omega \in \Omega,\ X(\omega) \geq 0,$ alors $\mathbf{E}(X) \geq 0$.
5. L'espérance est **croissante** : si $\forall \omega \in \Omega,\ X(\omega) \leq Y(\omega),$ alors $\mathbf{E}(X) \leq \mathbf{E}(Y)$.

⚠ La positivité de l'espérance ne veut pas dire que $\mathbf{E}(X)$ est toujours positive. C'est à comprendre dans le même sens que la positivité de l'intégrale par exemple.

Preuve

1. **Linéarité :**
$$\begin{aligned}\mathbf{E}(aX+bY) &= \sum_{\omega\in\Omega} \Big(aX(\omega) + bY(\omega)\Big)\mathbf{P}(\{\omega\}) \\ &= \sum_{\omega\in\Omega} \Big(aX(\omega)\mathbf{P}(\{\omega\}) + bY(\omega)\mathbf{P}(\{\omega\})\Big) \\ &= a\sum_{\omega\in\Omega} X(\omega)\mathbf{P}(\{\omega\}) + b\sum_{\omega\in\Omega} Y(\omega)\mathbf{P}(\{\omega\}) \quad \text{(linéarité de la somme)} \\ &= a\mathbf{E}(X) + b\mathbf{E}(Y)\end{aligned}$$
2. **Espérance d'une constante :**
   Si $\forall \omega \in \Omega, X(\omega) = a$, alors $\mathbf{E}(X) : \sum_{x=a} a\mathbf{P}(X=a) = \sum_{x=a} a1 = a$.
3. **Inégalité triangulaire :**
$$\begin{aligned}\left|\mathbf{E}(X)\right| &= \left|\sum_{x\in X(\Omega)} x\mathbf{P}(X=x)\right| \\ &\leq \sum_{x\in X(\Omega)} \left|x\mathbf{P}(X=x)\right| \quad \text{(inégalité triangulaire)} \\ &\leq \sum_{x\in X(\Omega)} |x|\,\mathbf{P}(X=x) \quad \text{(probabilité positive)} \\ &\leq \mathbf{E}\left(\left|X\right|\right) \quad \text{(théorème de transfert 2.3 vu plus loin.)}\end{aligned}$$
4. **Positivité :**
   $\forall x \in X(\Omega)$, $x \geq 0$ par hypothèse, et $\mathbf{P}(X=x) \geq 0$ par définition d'une probabilité.
   Donc $\forall x \in X(\Omega),\ x\mathbf{P}(X=x) \geq 0$, et par somme de termes positifs, $\mathbf{E}(X) \geq 0$.
5. **Croissance :**
   C'est une conséquence de la linéarité et de la positivité avec pour $X \geq Y$,
   $\mathbf{E}(X) - \mathbf{E}(Y) = \mathbf{E}(X-Y) \geq 0$.

■

Exemple

Si X est une variable aléatoire, alors sa variable aléatoire centrée est $X - \mathbf{E}(X)$. La linéarité de l'espérance montre bien que cette nouvelle variable aléatoire est d'espérance nulle.

Exemple *(Indicatrice)*

Quelle est l'espérance de $\mathbf{1}_A$ pour A une partie de Ω ?

Solution :

Exemple

Soit X une variable aléatoire réelle, quelle est l'espérance de $\mathbf{1}_{[a,b]}(X)$?

Solution :

Théorème 2.3 *(Théorème de transfert)*

Soit X une variable aléatoire de Ω sur E et $u : E \to E'$,

$$\mathbf{E}\left(u(X)\right) = \sum_{x \in X(\Omega)} u(x)\mathbf{P}(X = x)$$

On dit la même chose que la première expression de la définition. On remplace simplement l'univers Ω par $X(\Omega)$, et P par P_X.
La variable aléatoire considérée n'est plus X sur Ω, mais u sur $X(\Omega)$.

Preuve

Admis ■

Exemple

Nous avions calculé directement l'espérance de gain avec le lancer des deux dés. On peut aussi faire le calcul avec le théorème de transfert (même si, dans cet exemple c'est plus plus compliqué puisque nous avions déjà calculé la loi de $u(X)$).

$$\begin{aligned}
\mathbf{E}(u(X)) &= \sum_{x \in X(\Omega)} u(x)\mathbf{P}(X = x) \\
&= \sum_{x=1}^{12} u(x)\mathbf{P}(X = x) \\
&= u(1)\cdot 0 + u(2)\frac{1}{36} + u(3)\frac{2}{36} + u(4)\frac{3}{36} + u(5)\frac{4}{36} + u(6)\frac{5}{36} + ... \\
&\quad ... + u(7)\frac{6}{36} + u(8)\frac{5}{36} + u(9)\frac{4}{36} + u(10)\frac{3}{36} + u(11)\frac{2}{36} + u(12)\frac{1}{36} \\
&= 0 - \frac{1}{36} - \frac{2}{36} - \frac{3}{36} - \frac{4}{36} - \frac{5}{36} - \frac{6}{36} - \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + 10\frac{1}{36} \\
&= -\frac{7}{36} \approx -0,19
\end{aligned}$$

On retrouve le résultat précédent, et on voit bien que dans notre calcul, si on avait factorisé la somme par les valeurs communes de $u(x)$, on aurait retrouvé l'expression calculée sans le théorème de transfert.

⚠ Il en faut pas se tromper sur le théorème du transfert, en général :

$$\mathbf{E}(u(X)) \neq u(\mathbf{E}(X))$$

Par exemple si X suit la loi : $\mathbf{P}(X = 1) = \mathbf{P}(X = -1) = \frac{1}{2}$, on pose $Y = \frac{1}{X}$.
$\mathbf{P}(Y = 1) = \mathbf{P}(\frac{1}{X} = 1) = \mathbf{P}(X = 1) = \frac{1}{2}$ et de même, $\mathbf{P}(Y = -1) = \frac{1}{2}$.
$\mathbf{E}(X) = \mathbf{E}(Y) = 0$ mais $\frac{1}{\mathbf{E}(X)}$ n'existe pas ! (on ne peut pas diviser par 0).

Définition 2.4 *(Moment d'ordre k)*

Pour X une variable aléatoire finie, on désigne le **moment d'ordre** k de X par

$$\mathbf{E}\left(X^k\right) = \sum_{x \in X(\Omega)} x^k \mathbf{P}(X = x)$$

Exemple

Le moment d'ordre 1 est l'espérance.

B Indicateurs de dispersion

Définition 2.5 *(Variance)*

La **variance** d'une variable aléatoire X est définie par

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

C'est le moment d'ordre 2 de la variable centrée $X - \mathbf{E}(X)$.

Explications
Le principe de la variance est de mesurer une dispersion, c'est-à-dire un étalement des valeurs.
Elle mesure donc un étalement moyen autour de l'espérance. Comme l'écart moyen entre X est $\mathbf{E}(X)$ est toujours nul par linéarité de la moyenne, on étudie ici « la moyenne du carré de l'écart à la moyenne ».

⚠ La variance n'est pas *homogène* avec X : lorsque X désigne une longueur, $\mathbf{V}(X)$ désigne une longueur au carré. Il s'agit d'un rapport quadratique.

Propriété 2.6 *(Formule de Kœnig-Huygens)*

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

C'est cette formule que l'on utilise en pratique pour le calcul de la variance.

Preuve
Il suffit de développer le carré et d'utiliser la linéarité de $\mathbf{E}$.

$$\begin{aligned}
\mathbf{V}(X) = \mathbf{E}\left(\left(X - \mathbf{E}(X)\right)^2\right) &= \mathbf{E}\left(X^2 - 2X\mathbf{E}(X) + \mathbf{E}(X)^2\right) \\
&= \mathbf{E}\left(X^2\right) - 2\mathbf{E}\left(X\mathbf{E}(X)\right) + \mathbf{E}\left(\mathbf{E}(X)^2\right) \qquad \text{(par linéarité)} \\
&= \mathbf{E}\left(X^2\right) - 2\mathbf{E}(X)^2 + \mathbf{E}(X)^2 \qquad \text{(car } \mathbf{E}(X) \text{ et } \mathbf{E}(X)^2 \text{ sont des constantes)} \\
&= \mathbf{E}\left(X^2\right) - \mathbf{E}(X)^2
\end{aligned}$$

■

Propriété 2.7

Pour toutes constantes réelles a et b,

$$\mathbf{V}(aX + b) = a^2\,\mathbf{V}(X)$$

Explications
Il est intuitif que V est insensible à l'ajout d'une constante : la translation des valeurs ne modifie pas leur dispersion.
Par contre, l'homothétie par a influence également la dispersion : c'est un changement d'échelle. Comme la variance est quadratique, elle est multipliée par a^2.

Preuve

$$\begin{aligned}
\mathbf{V}(aX + b) &= \mathbf{E}\left((aX + b)^2\right) - \left(\mathbf{E}(aX + b)\right)^2 \\
&= \mathbf{E}\left(a^2X^2 + 2abX + b^2\right) - \left(a\mathbf{E}(X) + b\right)^2 \\
&= a^2\mathbf{E}\left(X^2\right) + 2ab\mathbf{E}(X) + b^2 - a^2\mathbf{E}(X)^2 - 2ab\mathbf{E}(X) - b^2 \\
&= a^2\mathbf{V}(X)
\end{aligned}$$

■

Propriété 2.8

La variance est une grandeur positive : pour toute variable aléatoire X, $\mathbf{V}(X) \geq 0$.

Remarque : ici, on a bien $\mathbf{V}(X) \geq 0$ même si X prend des valeurs négatives. Ce n'est pas la même notion que la positivité de l'espérance qui suppose $X \geq 0$.

Preuve
La variable aléatoire $(X - \mathbf{E}(X))^2$ est à valeurs positives. La positivité de l'espérance permet de conclure. ■

Cette positivité de la variance permet de définir l'écart type :

Définition 2.9 *(Écart type)*

L'**écart type** d'une variable aléatoire X est défini par

$$\sigma_X = \sigma(X) = \sqrt{\mathbf{V}(X)}$$

Une variable est *réduite* lorsque son écart type est égal à 1.

C'est la raison pour laquelle on note souvent la variance $\mathbf{V}(X) = \sigma^2(X) = \sigma_X^2$
L'intérêt de l'écart type est qu'il est *homogène* avec X : il est dans la même unité que X.

Propriété 2.10

Pour toutes constantes réelles a et b,

$$\sigma(aX + b) = |a|\,\sigma(X)$$

⚠ Ne pas oublier la valeur absolue !

Preuve
À faire en exercice. ■

Exemple *(Variable aléatoire centrée réduite)*
Si X une variable aléatoire réelle non certaine, alors $\sigma(X) \neq 0$ et on peut poser

$$Y = \frac{X - \mathbf{E}(X)}{\sigma_X}$$

Montrer que Y est alors une variable aléatoire **centrée réduite**.
Solution :

C Fonction génératrice (hors programme)

Cette notion est hors programme, nous en esquissons simplement le contour car elle est simple, elle facilite les calculs, et elle se retrouve facilement dans les sujets de concours.
Cette partie constitue donc davantage un exercice à savoir refaire, qu'un ensemble de propriétés à connaître par cœur.

Définition 2.11 *(Fonction génératrice)*

Soit $\mathbf{P}$ une probabilité sur un univers Ω et X une variable aléatoire définie sur Ω et à valeurs dans $[\![0, n]\!]$.
On définit la fonction génératrice de X par

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

La fonction génératrice est une application polynômiale. Nous pouvons donc utiliser tous les outils d'analyse et sur les polynômes pour l'étudier. Ses dérivées successives correspondent aux *moments* de la variable. En particulier, on retrouve facilement l'espérance et la variance (qui correspond au moment centré d'ordre 2).

Propriété 2.12

- $G(1) = 1$
- $G'(1) = \mathbf{E}(X)$
- $G''(1) = \mathbf{E}(X(X-1))$ d'où $\mathbf{V}(X) = G''(1) + G'(1) - G'(1)^2$

Preuve

$G(1) = \sum_{k=0}^{n} \mathbf{P}(X = k)1^k = \sum_{k=0}^{n} \mathbf{P}(X = k) = 1$
car $\{[X = k],\ k \in [\![0, n]\!]\}$ forme un système complet d'événements.
$G'(1) = \sum_{k=0}^{n} k\mathbf{P}(X = k)1^{k-1} = \sum_{k=0}^{n} k\mathbf{P}(X = k) = \mathbf{E}(X)$
$G''(1) = \sum_{k=0}^{n} k(k-1)\mathbf{P}(X = k)1^{k-2} = \sum_{k=0}^{n} k(k-1)\mathbf{P}(X = k) = \mathbf{E}(X(X-1))$
(d'après la formule de transfert).

$$\begin{aligned}
\mathbf{V}(X) &= \mathbf{E}\left(X^2\right) - (\mathbf{E}(X))^2 \\
&= \mathbf{E}\left(X(X-1) + X\right) - \left(G'(1)\right)^2 \\
&= \mathbf{E}\left(X(X-1)\right) + \mathbf{E}(X) - \left(G'(1)\right)^2 \\
&= G''(1) + G'(1) - \left(G'(1)\right)^2
\end{aligned}$$

■

3 Lois usuelles

Voici un aperçu rapide des principales lois que l'on rencontre en probabilités finies. Pour l'essentiel, elles ont été vues au lycée.
Lors de l'apprentissage, il est vivement conseillé d'apprendre l'exemple type associé à chaque loi.

A Loi certaine

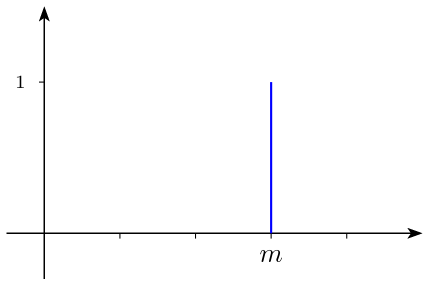
Exemple *(Exemple type : situation déterministe)*
Un processus physique (macroscopique) dont on connait parfaitement les conditions initiales, les conditions de déroulement, et dont on peut donc prévoir l'issue avec certitude. L'événement certain correspond souvent à une constante que l'on ajoute ou soustrait à une autre loi (par exemple une mise initiale dans un jeu de pari : on est totalement sûr de la mettre, quelque soit le tirage qui va suivre).

Définition 3.1 *(Loi certaine)*

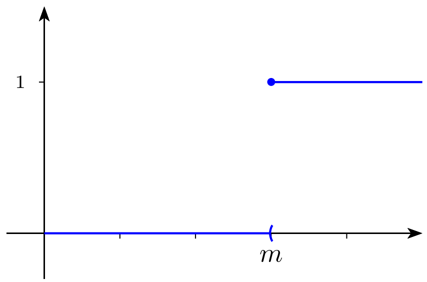
Une variable aléatoire X définie sur Ω suit une loi certaine,
s'il existe $m \in \mathbf{R}$ tel que $\mathbf{P}(X = m) = 1$.
On alors $X(\Omega) = \{m\}$.

Propriété 3.2

Pour une loi certaine égale à m.



Loi de probabilité



Fonction de répartition

Propriété 3.3

Si X est de loi certaine égale à m, alors

$$\mathbf{E}(X) = m \qquad \text{et} \qquad \mathbf{V}(X) = 0$$

Remarque : On a vu que la réciproque est vraie.

B Loi uniforme : situation équiprobable

Exemple *(Exemple type : lancer d'un dé)*
On lance un dé (équilibré) à 6 faces, on note X la variable aléatoire correspondant au numéro obtenu.

$$\forall k \in [\![1;6]\!],\ \mathbf{P}(X = k) = \frac{1}{6}$$

Définition 3.4 *(Loi uniforme)*

Soit X une variable aléatoire définie sur Ω à valeurs dans $[\![1, n]\!]$, on dit que X suit une loi uniforme si

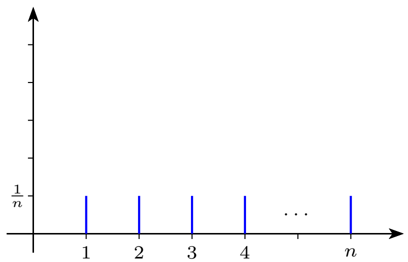
$$\forall k \in [\![1, n]\!],\ \mathbf{P}(X = k) = \frac{1}{n}$$

On note $X \sim \mathcal{U}([\![1, n]\!])$ ou $X \hookrightarrow \mathcal{U}([\![1, n]\!])$.

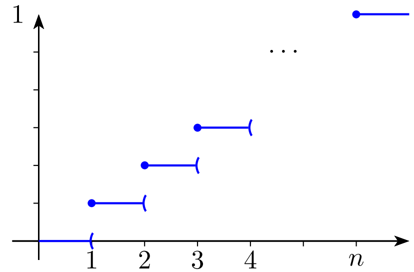
Remarque : Le signe $\sim$ n'a rien à voir avec les équivalents en analyse.

Propriété 3.5

Si X suit une loi uniforme dans $[\![1, n]\!]$,



Loi de probabilité



Fonction de répartition

Propriété 3.6

Si X une loi uniforme sur $[\![1, n]\!]$, alors

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

Preuve

$$\mathbf{E}(X) = \sum_{k=1}^{n} k\mathbf{P}(X = k) = \frac{n(n+1)}{2}\frac{1}{n} = \frac{n+1}{2}.$$

■

Exemple *(Variance de la loi uniforme)*

Calculer $\mathbf{V}(X)$.

Solution :

C Loi de Bernoulli : tirage à deux issues

Exemple *(Exemple type : lancer à deux issues)*

Une pièce truquée possède p chances de tomber sur pile (succès) et $q = 1 - p$ chances de tomber sur face (échec). On définit la variable aléatoire X qui teste si pile est sorti : X vaut 1 si on obtient pile, et 0 sinon.

$$\mathbf{P}(X = 1) = p \quad \text{et} \quad \mathbf{P}(X = 0) = 1 - p$$

Définition 3.7 *(Loi de Bernoulli)*

Soit X une variable aléatoire définie sur Ω à valeurs dans $\{0, 1\}$.
Pour $p \in [0, 1]$, on dit que X suit une loi de Bernoulli de paramètre p si

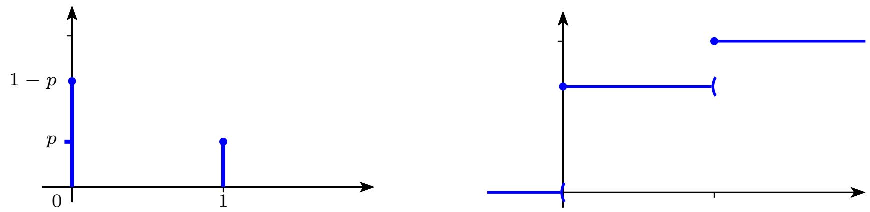
$$\mathbf{P}(X = 1) = p \quad \text{et} \quad \mathbf{P}(X = 0) = 1 - p = q$$

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

Remarque : En général $X(\Omega) = \{0, 1\}$, mais si $p = 0$, alors $X(\Omega) = \{0\}$ ou si $p = 1$, alors $X(\Omega) = 1$. Ce sont des cas critiques qui correspondent à la loi certaine.
Certains ouvrages excluent ces situations de la définition de la loi de Bernoulli.

Propriété 3.8

Pour une loi de Bernoulli de paramètre p,



Loi de probabilité — Fonction de répartition

Propriété 3.9

Soit X une variable aléatoire de loi $\mathcal{B}(p)$,

$$\mathbf{E}(X) = p, \qquad \mathbf{V}(X) = p(1 - p) = pq, \qquad \sigma_X = \sqrt{p(1 - p)} = \sqrt{pq}$$

Preuve

$\mathbf{E}(X) = 0 \times \mathbf{P}(X = 0) + 1 \times \mathbf{P}(X = 1) = p$

$\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = 0^2 \times \mathbf{P}(X = 0) + 1^2 \times \mathbf{P}(X = 1) - p^2 = p(1 - p)$ ■

Exemple

La loi $\mathbf{1}_A$ suit une loi de Bernoulli de paramètre $p = \mathbf{P}(A)$.

D Loi binomiale : tirages avec remise

Exemple *(Exemple type 1 : tirages avec remise)*
On dispose d'une urne avec N boules, dont R sont rouges et $N-R$ sont jaunes.
On effectue n tirages successifs avec remise, la probabilité d'obtenir exactement k boules rouges parmi les n tirages est

$$\mathbf{P}(X=k)=\binom{n}{k}\frac{R^k(N-R)^{n-k}}{N^n}$$

Exemple *(Exemple type 2 : répétitions indépendantes d'une épreuve de Bernoulli)*
On reprend la pièce de l'épreuve de Bernoulli, et on effectue n lancers indépendants.
On définit la variable aléatoire X qui compte le nombre de fois que l'on a obtenu pile (nombre de succès).

$$\mathbf{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

Définition 3.10 *(Loi binomiale)*

Soit X une variable aléatoire définie sur Ω à valeurs dans $[\![0,n]\!]$,
Pour $p\in[0,1]$, on dit que X suit une loi binomiale de paramètres n et p si

$$\forall k\in[\![0;n]\!],\ \mathbf{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

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

Remarque : D'après la formule du binôme de Newton, on retrouve bien

$$\mathbf{P}(X\in[\![0,n]\!])=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}=1$$

La remarque vue pour la loi de Bernoulli à propos du paramètre p est aussi valable pour la loi binomiale.

Propriété 3.11 *(Interprétation de la loi binomiale)*

Si X est la variable aléatoire égale au nombre de succès obtenus par la répétition indépendante de n expériences de Bernoulli de même paramètre $p\in[0,1]$,
alors, X suit une loi binomiale de paramètres n, p : $X\sim\mathcal{B}(n,p)$.

Autrement dit :
Soit $n\in\mathbf{N}^*$, et $\forall i\in[\![1,n]\!]$, $X_i\sim\mathcal{B}(p)$ (indépendantes entre elles). Si on pose $X=\sum\limits_{i=1}^{n}X_i$, alors

$$X\sim\mathcal{B}(n,p)$$

Preuve

- preuve « **intuitive** » : Soient n et k fixés.
Les tirages successifs s'écrivent comme un n-uplet, composé de succès S et d'échecs $\overline{S}$.
Un événement élémentaire qui vérifie $X=k$, peut donc s'écrire :

$$\underbrace{\underline{S}\ \ \underline{S}\ \ \underline{\overline{S}}\ \ \underline{S}\ \ \underline{\overline{S}}\ \ \underline{\overline{S}}\ \ \underline{S}\ \ \cdots\ \ \underline{\overline{S}}}_{n\text{ tirages dont }k\text{ succès }S\text{ et }n-k\text{ échecs }\overline{S}}$$

On note S_i : « le $i^{\text{ème}}$ tirage donne un succès ».
La probabilité de l'événement donné ci-dessus en exemple s'écrit :

$$\mathbf{P}\left(S_1\cap S_2\cap\overline{S_3}\cap\cdots\cap\overline{S_n}\right)$$

Or les événements sont indépendants entre eux, donc

$$\mathbf{P}\left(S_1\cap S_2\cap\overline{S_3}\cap\cdots\cap\overline{S_n}\right)=\mathbf{P}\left(S_1\right)\mathbf{P}\left(S_2\right)\ \mathbf{P}\left(\overline{S_3}\right)\cdots\mathbf{P}\left(\overline{S_n}\right)$$

Et $\forall i\in[\![1,n]\!]$, $\quad\mathbf{P}(S_i)=p$
$\qquad\qquad\qquad\mathbf{P}\left(\overline{S_i}\right)=1-p$
Dans le produit, il y a k succès et $n-k$ échecs, la probabilité précédente s'écrit donc

$$\mathbf{P}\left(S_1\cap S_2\cap\overline{S_3}\cap\cdots\cap\overline{S_n}\right)=p^k\,(1-p)^{n-k}$$

Or $[X=k]$ correspond à l'union de tous les « événements anagrammes » du précédent.
Ces événements sont incompatibles, donc la probabilité de l'union est égale à la somme des probabilités. Ils ont tous la même probabilité $p^k(1-p)^{n-k}$ et il y en a au total $\binom{n}{k}$.
Donc

$$\mathbf{P}\left(X=k\right)=\binom{n}{k}p^k\,(1-p)^{n-k}$$

- preuve « **formelle** » : Avec les notations précédentes.
On note $\mathcal{P}_k$ l'ensemble des parties à k éléments de $[\![1,n]\!]$.

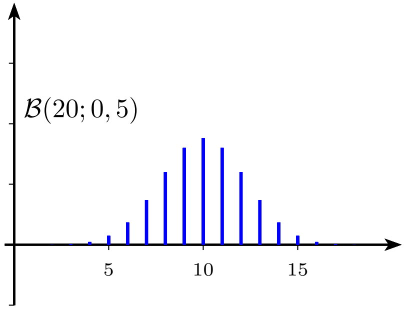
$$[X=k]=\bigcup_{I\in\mathcal{P}_k}\left[\left(\bigcap_{i\in I}S_i\right)\cap\left(\bigcap_{j\in[\![1,n]\!]\setminus I}\overline{S_j}\right)\right]$$

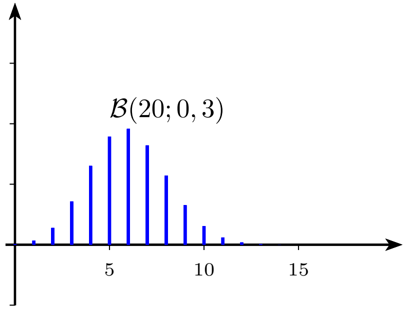
Et cette union est disjointe (événements incompatibles) car les $I\in\mathcal{P}_k$ sont supposés deux à deux différents.
Donc $\mathbf{P}([X=k])=\sum\limits_{I\in\mathcal{P}_k}\mathbf{P}\left[\left(\bigcap\limits_{i\in I}S_i\right)\cap\left(\bigcap\limits_{j\in[\![1,n]\!]\setminus I}\overline{S_j}\right)\right]$
Or les événements $(S_i)_{i\in[\![1,n]\!]}$ sont indépendants, donc

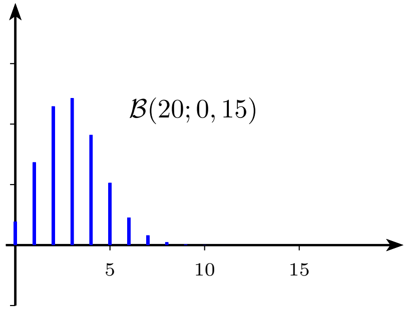
$$\begin{aligned}\mathbf{P}([X=k])&=\sum_{I\in\mathcal{P}_k}\left(\prod_{i\in I}\mathbf{P}(S_i)\right)\ \left(\prod_{j\in[\![1,n]\!]\setminus I}\mathbf{P}\left(\overline{S_j}\right)\right)\\&=\sum_{I\in\mathcal{P}_k}\left(\prod_{i\in I}p\right)\ \left(\prod_{j\in[\![1,n]\!]\setminus I}(1-p)\right)\\&=\sum_{I\in\mathcal{P}_k}p^k(1-p)^{n)-k}\qquad(\text{car Card}\,I=k\text{ et Card}\,([\![1,n]\!]\setminus I)=n-k)\\&=\binom{n}{k}p^k(1-p)^{n)-k}\qquad(\text{car Card}\,\mathcal{P}_k=\tbinom{n}{k})\end{aligned}$$

■

Lois de probabilités en fonction du paramètre p.







On remarque que lorsque $p = 0,5$ les probabilités sont symétriques par rapport à la moyenne, mais pour $p \neq 0,5$, il existe une asymétrie d'autant plus marquée que p est loin de $0,5$.
On remarque également que la distribution ressemble à une gaussienne (courbe en cloche).
Lorsque n tend vers $+\infty$, la loi binomiale tend vers la loi normale (le théorème central limite nous dit que c'est le cas de toutes les lois construites sur ce modèle : la répétitions de n expériences indépendantes qui suivent la même loi).

Propriété 3.12

Soit X une variable aléatoire de loi $\mathcal{B}(n,p)$,

$$\mathbf{E}(X) = np, \qquad \mathbf{V}(X) = np(1-p) = npq, \qquad \sigma_X = \sqrt{np(1-p)} = \sqrt{npq}$$

Preuve

Si on note $\forall i \in [\![1,n]\!]$, X_i une loi de Bernoulli de paramètre p.

On peut alors écrire X sous la forme $X = \sum\limits_{i=0}^{n} X_i$.

Alors $\mathbf{E}(X) = \mathbf{E}\left(\sum\limits_{i=1}^{n} X_i\right) = \sum\limits_{i=1}^{n} \mathbf{E}(X_i) = \sum\limits_{i=1}^{n} p = np$.

Pour la variance, on n'a pas de telle formule (actuellement). On est donc obligé d'aller au charbon.

D'après Kœnig-Huygens, $\mathbf{V}(X) = \mathbf{E}\left(X^2\right) - (\mathbf{E}(X))^2 = \mathbf{E}\left(X^2\right) - (np)^2$.

$$\begin{aligned}
\mathbf{E}\left(X^2\right) &= \sum_{k=0}^{n} k^2 P(X=k) = \sum_{k=\cancel{0}1}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} nk \binom{n-1}{k-1} p^k (1-p)^{n-k} \qquad \text{(formule sans nom)} \\
&= \sum_{k=0}^{n-1} n(k+1) \binom{n-1}{k} p^{k+1} (1-p)^{n-k-1} \qquad \text{(changement d'indice)} \\
&= \sum_{k=0}^{n-1} nk \binom{n-1}{k} p^{k+1} (1-p)^{n-k-1} + n \binom{n-1}{k} p^{k+1} (1-p)^{n-k-1} \\
&= np \sum_{k=0}^{n-1} k \binom{n-1}{k} p^k (1-p)^{n-k-1} + np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-k-1} \\
&= np \sum_{k=1}^{n-1} k \binom{n-1}{k} p^k (1-p)^{n-k-1} + np\,1 \qquad \text{(binôme de Newton)} \\
&= np \sum_{k=1}^{n-1} (n-1) \binom{n-2}{k-1} p^k (1-p)^{n-k-1} + np\,1 \qquad \text{(formule sans nom)} \\
&= n(n-1)p \sum_{k=0}^{n-2} \binom{n-2}{k} p^{k+1} (1-p)^{n-2-k} + np\,1 \qquad \text{(changement d'indice)} \\
&= n(n-1)p^2 \sum_{k=0}^{n-2} \binom{n-2}{k} p^k (1-p)^{n-2-k} + np\,1 \\
&= n(n-1)p^2 + np\,1 \qquad \text{(binôme de Newton)} \\
&= np(np-p+1)
\end{aligned}$$

Donc

$$\mathbf{V}(X) = np(np-p+1) - (np)^2 = np(1-p)$$

autre preuve : avec la fonction génératrice.
La fonction génératrice de la loi binomiale peut s'écrire :

$$G_X(x) = \sum_{k=0}^{n} x^k P(X=k) = \sum_{k=0}^{n} x^k \binom{n}{k} p^k (1-p)^{n-k} = (xp+1-p)^n$$

Or $\mathbf{E}(X) = G'_X(1)$, et $\mathbf{E}(X(X-1)) = G''_X(1)$.

$$\begin{aligned}
\text{Donc } \mathbf{V}(X) &= \mathbf{E}\left(X^2\right) - (\mathbf{E}(X))^2 = \mathbf{E}(X(X-1)+X) - (\mathbf{E}(X))^2 \\
&= \mathbf{E}(X(X-1)) + \mathbf{E}(X) - (\mathbf{E}(X))^2 \\
&= G''_X(1) + G'_X(1)(X) - \left(G'_X(1)\right)^2
\end{aligned}$$

$\forall x \in \mathbf{R}$, $G'_X(x) = np(xp+1-p)^{n-1}$. Donc $\mathbf{E}(X) = G'_X(1) = np$.
$\forall x \in \mathbf{R}$, $G''_X(x) = n(n-1)p^2(xp+1-p)^{n-2}$. Donc $G''_X(1) = n(n-1)p^2$.
Ainsi $\mathbf{V}(X) = n(n-1)p^2 + np - (np)^2 = np(1-p)$. ∎

E Loi hypergéométrique : tirages sans remise

Exemple *(Exemple type : n tirages sans remise)*
On dispose d'une urne avec N boules, dont R sont rouges et $N-R$ sont jaunes. On effectue n tirages successifs **sans** remise, la probabilité d'obtenir exactement k boules rouges parmi les n tirages est

$$\mathbf{P}(X=k)=\frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}.$$

C'est le même résultat que pour un tirage simultané de n boules.

Définition 3.13 *(Loi hypergéométrique)*

Soient $N, n \in \mathbf{N}^*$,
Soit X une variable aléatoire définie sur Ω à valeurs dans $[\![0, n]\!]$,
Pour $p \in [0,1]$[1],
on dit que X suit une loi hypergéometrique de paramètres N, n et p si

$$\forall k \in [\![0;n]\!],\ \mathbf{P}(X=k)=\frac{\binom{pN}{k}\binom{(1-p)N}{n-k}}{\binom{N}{n}}.$$

On note $\mathcal{H}(N,n,p)$ la loi hypergéométrique de paramètres N, n, p.

$$X \sim \mathcal{H}(N,n,p).$$

Par rapport à l'exemple précédent, $p=\frac{R}{N}$ représente la proportion de boules rouges parmi les N boules présentes initialement.

Propriété 3.14

Soit X une variable aléatoire de loi $\mathcal{H}(N,n,p)$,

$$\mathbf{E}(X)=np.$$

Lemme 3.15 *(Rappel : Formule de Vandermonde)*
Soient $n, m, k \in N^*$, avec $k \leq n+m$.

$$\binom{n+m}{k}=\sum_{i=0}^{k}\binom{n}{i}\binom{m}{k-i}.$$

Cette formule n'est pas explicitement au programme, mais peut intervenir facilement dans des exercices : il faut donc savoir la redémontrer.

1. On choisit p tel que $pN \in \mathbf{N}^*$.

Preuve
L'idée (comme souvent pour ce genre de relations) est de repasser aux polynômes.

$$\begin{aligned}
&(1+X)^{m+n}=(1+X)^n(1+X)^m\\
\iff &\sum_{k=0}^{m+n}\binom{n}{k}X^k=\sum_{i=0}^{m}\binom{n}{i}X^i\sum_{j=0}^{m}\binom{m}{j}X^j\\
\iff &\sum_{k=0}^{m+n}\binom{n}{k}X^k=\sum_{i=0}^{n}\sum_{j=0}^{m}\binom{n}{i}\binom{m}{j}X^{j+i}
\end{aligned}$$

On identifie le coefficient de X^k dans les deux expressions :
Dans la deuxième expression, pour $i \geq 0$ fixé, $j=k-i \geq 0$. Donc i varie dans $[\![0,k]\!]$

$$\binom{n}{k}=\sum_{i=0}^{k}\binom{n}{i}\binom{m}{k-i}$$

■

Preuve

$$\begin{aligned}
\mathbf{E}(X)&=\sum_{k=0}^{n}k\mathbf{P}(X=k)=\sum_{k=0}^{n}k\frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}=\frac{1}{\binom{N}{n}}\sum_{k=1}^{n}k\binom{R}{k}\binom{N-R}{n-k}\\
&=\frac{1}{\binom{N}{n}}\sum_{k=1}^{n}R\binom{R-1}{k-1}\binom{N-R}{n-k}=\frac{1}{\binom{N}{n}}\sum_{k=1}^{n}R\binom{R-1}{k-1}\binom{N-R}{n-1-(k-1)}\\
&=\frac{1}{\binom{N}{n}}\sum_{j=0}^{n-1}R\binom{R-1}{j}\binom{N-R}{n-1-j}\\
&=\frac{1}{\binom{N}{n}}R\binom{R-1+N-R}{n-1} \qquad \text{(Formule de Vandermonde)}\\
&=\frac{R}{\binom{N}{n}}\binom{N-1}{n-1}=\frac{Rn}{N} \qquad \text{(Écrire les factorielles et simplifier)}\\
&=np
\end{aligned}$$

■

Propriété 3.16 *(Approximation de la loi hypergéométrique par la loi binomiale)*

Soit $n \in \mathbf{N}$, et $p \in [0,1]$,

$$\mathcal{H}(N,n,p) \xrightarrow[pN\in\mathbf{N},\quad N\to+\infty]{} \mathcal{B}(n,p)$$

C'est-à-dire que $\forall k \in [\![0;n]\!]$,

$$\frac{\binom{pN}{k}\binom{(1-p)N}{n-k}}{\binom{N}{n}} \xrightarrow[pN\in\mathbf{N},\quad N\to+\infty]{} \binom{n}{k}p^k(1-p)^{n-k}$$

Preuve

Soient $X_N \sim \mathcal{H}(N,n,p)$ et $X \sim \mathcal{B}(n,p)$, $\quad \mathbf{P}(X_N = k) = \dfrac{\binom{Np}{k}\binom{(1-p)N}{n-k}}{\binom{N}{n}}$

Or $\dbinom{Np}{k} = \dfrac{(Np)!}{k!(Np-k)!} = \dfrac{(Np)(Np-1)\cdots(Np-k+1)}{k!} \underset{N\to+\infty}{\sim} \dfrac{(Np)^k}{k!}$,

de même $\dbinom{N(1-p)}{n-k} \underset{N\to+\infty}{\sim} \dfrac{(N(1-p))^{n-k}}{(n-k)!}$ et $\dbinom{N}{n} \underset{N\to+\infty}{\sim} \dfrac{N^n}{n!}$

Donc par quotient et produit : $\mathbf{P}(X_N = k) \underset{N\to+\infty}{\sim} \mathbf{P}(X = k)$ ∎

Exemple

Lorsque l'urne est très grande par rapport au nombre de tirages, on voit intuitivement que le fait que le tirage soit avec ou sans remise est sans grand influence : Supposons que je sélectionne 5 individus au hasard parmi la population française. Que je m'autorise ou non de piocher plusieurs fois de suite la même personne ne change rien concrètement car la probabilité de tirer deux fois de suite une personne est négligeable.

4 Inégalité de Bienaymé-Tchebychev

Remarque : Tchebychev est une translittération, on pourra trouver d'autres orthographes comme Chebyshev.

Propriété 4.1 *(Inégalité de Markov)*

Soient X une variable aléatoire réelle **positive** et $a > 0$,

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

Preuve

La fonction $a\mathbf{1}_{[a,+\infty[}$ vaut 0 pour $x < a$ et a pour $x \geq a$.

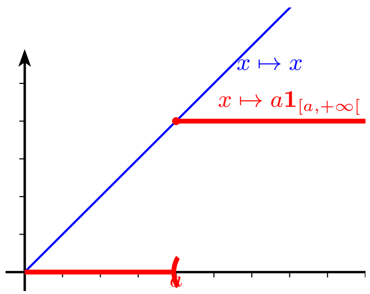
On a donc $\forall x \in \mathbf{R}$, $x \geq a\mathbf{1}_{[a,+\infty[}(x)$. Comme les fonctions précédentes sont croissantes, on peut composer *à droite* par X.

$$\forall \omega \in \Omega, X(\omega) \geq a\mathbf{1}_{[a;+\infty[}(X(\omega))$$

C'est-à-dire $X \geq a\mathbf{1}_{[a,+\infty[}(X)$

On passe à l'espérance qui est croissante :

$$\mathbf{E}(X) \geq a\mathbf{E}\left(\mathbf{1}_{[a;+\infty[}(X)\right) = a\mathbf{P}(X \geq a).$$



∎

Théorème 4.2 *(Inégalité de Bienaymé-Tchebychev)*

Soient X une variable aléatoire réelle et $a > 0$,

$$\mathbf{P}\left(\left|X - \mathbf{E}(X)\right| \geq a\right) \leq \frac{\sigma_X^2}{a^2}.$$

Explications

Cette inégalité offre une façon d'interpréter la dispersion d'une variable aléatoire : les résultats sont d'autant plus concentrés autour de l'espérance que l'écart type est faible : la probabilité d'avoir un résultat éloigné de l'espérance décroît avec l'écart type.

La distance à l'espérance est traduite par $|X - \mathbf{E}(X)| \geq a$. La probabilité d'avoir la valeur de X éloignée de plus de a de l'espérance est majoré par $\left(\frac{\sigma_X}{a}\right)^2$.

Preuve

$$|X - \mathbf{E}(X)| \geq a \iff (X - \mathbf{E}(X))^2 \geq a^2$$

Donc $\mathbf{P}(|X - \mathbf{E}(X)| \geq a) = \mathbf{P}\left((X - \mathbf{E}(X))^2 \geq a^2\right)$.

On applique ensuite l'inégalité de Markov à $X - \mathbf{E}(X)$ (on centre la variable). ∎

Corollaire 4.3

Soit X une variable aléatoire réelle, soit $a > 0$,

$$\mathbf{P}(|X - \mathbf{E}(X)| < a) \geq 1 - \frac{\sigma_X^2}{a^2}$$

Exemple

Donner un majorant de la probabilité de s'éloigner de plus de $10\sigma_X$ de la moyenne pour une épreuve donnée.

Solution :

Exemple

Un tireur a une probabilité de 70% de toucher sa cible à chaque coup. Chaque coup de fusil est indépendant des précédents.

On note X le nombre de fois qu'il a atteint sa cible.

À partir de combien de séances de tir peut-on affirmer que la probabilité d'avoir taux de succès entre 60% et 80% est supérieur à 90% ?

Solution :