Variables aléatoires

BCPST 1, 27/03/2018

Notations du chapitre — $(\Omega, \mathscr{T}, P)$ *est un espace probabilisé fini.*

L'idée d'une variable aléatoire est qu'on s'attache à des « mesures » sur les évènements aléatoires.

Exemple A Une maladie a 2 chances sur 5 de toucher un animal chaque année. On considère un élevage de 4 animaux et on s'intéresse :

1. au nombre d'animaux malades une année donnée (nombre noté X);
2. à la première année ou aucun animal est malade (T).

Ainsi X et T ne sont pas des événements. En effet, ici, l'ensemble des événements est l'ensemble des suites de 4 lettres M ou S (malade ou sain) :

$$\Omega = \left\{(L_1, \ldots, L_5) \ / \ \forall i \in [\![1\,;4]\!], L_i = M \text{ ou } S\right\}.$$

Plus concrètement les évènements élémentaires ressemblent à $\{(M,S,M,S)\}$ ou $\{(S,M,M,S)\}$. Les valeurs de X et de T peuvent s'exprimer pour un résultat $\omega \in \Omega$ de l'expérience. Par exemple $X(MSMS) = 2$ ou encore $T(SSSM) = 4$. On peut voir que

$$X(\omega) \in \{0,1,2,3,4\} \qquad \text{et} \qquad T(\omega) \in \mathbb{N}^*$$

On peut donc dire que X et T sont des « fonctions » des événements.

I — Définitions

Définition 1.1 — Variable aléatoire réelle
On appelle variable aléatoire réelle (en abrégé v.a.r.) toute application de Ω dans $\mathbb{R}$.

On note couramment $(X = x)$ l'évènement $X^{-1}(\{x\})$: c'est l'ensemble des évènements élémentaires pours lesquelles X vaut x, ou tout simplement l'évènement « X est égal à x ».
Plus généralement I un sous-ensemble de $\mathbb{R}$. L'ensemble $X^{-1}(I) = \{\omega \in \Omega / X(\omega) \in I\}$, qui est un sous-ensemble de Ω, est un évènement aléatoire : on le note $(X \in I)$. On note aussi $(X \geqslant 2)$, $(1 < X \leqslant 2)$, etc.
Dans le cadre de ce cours, Ω peut s'écrire $\left\{\omega_1, \omega_2, \cdots, \omega_p\right\}$. L'ensemble des valeurs de X s'écrit donc $X(\Omega) = \left\{X(\omega_1), X(\omega_2), \cdots, X(\omega_p)\right\} = \{x_1, x_2, \cdots, x_n\}$ (avec $n \geqslant p$).

Définition 1.2 — Support d'une variable aléatoire
Le support d'une variable aléatoire X est l'ensemble $X(\Omega)$ des valeurs atteintes par X.
La variable aléatoire X est discrète si $X(\Omega)$ est dénombrable.
La variable aléatoire X est discrète finie si $X(\Omega)$ est fini.
La variable aléatoire X est continue si $X(\Omega)$ est indénombrable.

Évidemment, dans le cadre de ce chapitre, les variable aléatoire sont discrètes finies.

Exemple — A
On trouve facilement $X(\Omega) = \{0, 1, 2, 3, 4\}$ et $T(\Omega) = \mathbb{N}^*$.
Dans le cas de T la variable est discrète infinie. En fait l'univers utilisé pour étudier T n'est pas Ω. C'est l'ensemble des suites infinies de S et M. C'est un ensemble infini ; ce cas sera étudié en Spé.

Cette année nous n'étudions que le cas des variable aléatoire finie.

Exemple — B
On lance 2d6 parfaitement équilibrés et on note S la somme des résultats des deux dés.
Alors $S(\Omega) = [\![2\,;12]\!]$.

Propriété 1.5 — Système complet d'évènements associé
Soit $(\Omega, \mathscr{T})$ un espace probabilisable et X un v.a.r. finie de support $X(\Omega) = \left\{x_1, \dots, x_p\right\}$.
Alors $(X = x_1)$, $(X = x_2)$, …, $(X = x_p)$ est un s.c.e. : c'est le système complet d'évènements associé à la variable aléatoire X

DÉM. D'une part si $(i, j) \in [\![0\,; n]\!]^2$ avec $i \neq j$ alors $(X = x_i) \cap (X = x_j) = (X = x_i \cap X = x_j) = \varnothing$.

D'autre part
$$\bigcup_{i=1}^{n} (X = x_i) = (X \in X(\Omega)) = \Omega$$
□

Propriété 1.6 — Fonction d'une variable aléatoire
Soit X une variable aléatoire sur Ω et g une fonction de $\mathbb{R}$ dans $\mathbb{R}$
Alors $g \circ X$ est une variable aléatoire, noté $g(X)$.

Par exemple X^2, $\sin(X)$, etc. sont des variable aléatoire si X en est une.

II — Loi de probabilité

II.1 — Définition

Définition 2.1 — Loi de probabilité d'une variable aléatoire
On appelle loi de probabilité d'une variable aléatoire X la donnée de

1) *l'ensemble $X(\Omega)$;*
2) *la valeur de $P(X = x)$ pour tout réel x dans $X(\Omega)$.*

Exemple B — A
Dans le cas de l'expérience A, on a vu que l'univers le plus adapté est l'ensemble des mots de 4 lettres écrits avec M et S muni de l'équiprobabilité.
On a déjà vu que $X(\Omega) = [\![0 ; 4]\!]$.
Pour $k \in X(\Omega)$, l'évènement $(X = k)$ se produit lorsque un des évènements avec k lettres M et $4 - k$ lettres S se produit. Il y a $\binom{4}{k}$ tels évènements. Ils sont tous de probabilités $\left(\frac{2}{5}\right)^k \left(\frac{3}{5}\right)^{4-k}$. Ainsi

$$\forall k \in X(\Omega), \quad P(X = k) = \binom{4}{k} \left(\frac{2}{5}\right)^k \left(\frac{3}{5}\right)^{4-k}$$

On peut représenter cette loi à l'aide d'un diagramme en bâtons.

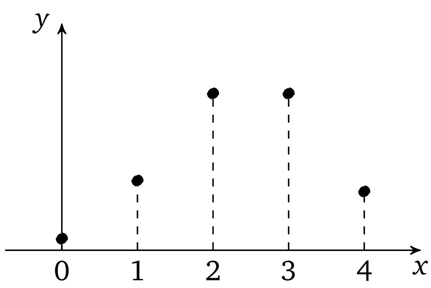


FIGURE I.1 — *Loi binomiale* $\mathscr{B}(4, 3/5)$.

En pratique Déterminer la loi de probabilité d'une variable aléatoire consiste à donner d'une part son support et d'autre part les probabilités de chacun des évènement $(X = x_i)$.
Cela se fait en général sous forme d'une expression algébrique ou, dans les cas les plus simples, sous forme d'un tableau.

L'intérêt de la loi de X est qu'elle permet, en théorie, de calculer toutes les probabilités d'évènements associés à X : $X \leqslant 2$, (X pair), etc. Dans le cas de l'expérience A, on peut s'intéresser à l'évènement C « au plus un animal est malade ». Il s'écrit $(X=0)\cup(X=1)$. En utilisant le système complet d'évènements associé à X, on s'assure que l'union est disjointe et donc

$$P(X) = P(X=0) + P(X=1) = \left(\frac{3}{5}\right)^4 + \left(\frac{2}{5}\right)\left(\frac{3}{5}\right)^3 = \frac{35}{125}$$

Théorème 2.2 — Définition de la loi d'une variable aléatoire
Soit $(p_1, \dots, p_n) \in \mathbb{R}^n$. Ces réels définissent la loi de probabilité d'une variable aléatoire si et seulement si

1) $\forall i \in [\![1 ; n]\!], \quad p_i > 0;$

2) $\displaystyle\sum_{i=1}^{n} p_i = 1.$

DÉM. Ce résultat est admis. □

En pratique Ce théorème permet de définir une loi de probabilité sans avoir à décrire une expérience aléatoire sous-jacente. On va même souvent plus loin : une variable aléatoire X étant décrite *on fait l'hypothèse* qu'elle suit telle ou telle loi. C'est le cas par exemple des problèmes de files d'attente, de durée de vie, etc.

Exemple C Soit $n \in \mathbb{N}$ et $(p_i)_{1\leqslant i\leqslant n}$ la suite définie par

$$\forall i \in [\![1 ; n]\!], \quad p_i = \frac{1}{n}$$

Ces réels sont positifs et leur somme vaut 1. Ils définissent donc une loi de probabilité : la loi uniforme sur $[\![1 ; n]\!]$ (on parle de loi uniforme car toutes les valeurs ont la même probabilité de survenir).

Exercice 1 Soit $n \in \mathbb{N}$ et $(p_i)_{1\leqslant i\leqslant 2n+1}$ la suite définie par

$$\forall i \in [\![1 ; 2n+1]\!], \quad p_i = \frac{i}{(n+1)^2} \quad \text{si} \quad i \leqslant n+1 \qquad \text{et} \qquad p_i = \frac{2n+2-i}{(n+1)^2} \quad \text{sinon}$$

Justifier que l'on définit bien ainsi une loi de probabilité et la représenter.
Cette loi porte le nom de loi triangulaire. Je vous laisse deviner pourquoi.

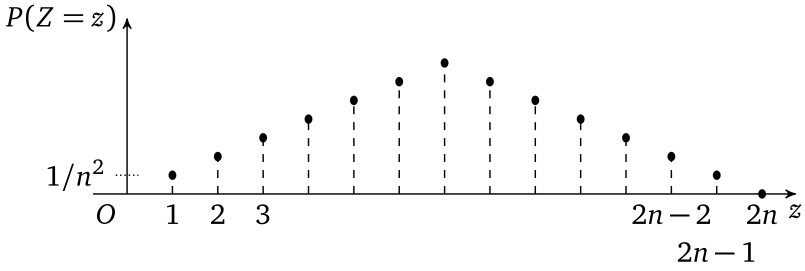


FIGURE I.2 — *Loi triangulaire.*

II.2 — Loi d'une fonction d'une variable aléatoire

Exemple Considérons la variable aléatoire X dont la loi est donnée par :

- $X(\Omega) = [\![-2\,;1]\!]$;
- $\forall k \in X(\Omega),\quad P(X=k) = \dfrac{(k-2)^2}{30}$

On a vu que $Y = X^2$ est aussi une variable aléatoire. Déterminons sa loi de probabilité.

- $Y(\Omega)$ est l'image de $[\![-2\,;2]\!]$ par la fonction $x \mapsto x^2$. Ainsi $Y(\Omega) = \{0,1,4\}$.
- Pour $y=1$: $P(Y=y) = P(X^2=1) = P(X=1 \cup X=-1) = P(X=1)+P(X=-1) = \dfrac{10}{30}$.

  On trouve de même $P(Y=0) = \dfrac{4}{30}$ et $P(Y=4) = \dfrac{16}{30}$.

En pratique Comme on peut le constater, déterminer la loi d'une fonction d'une variable aléatoire n'est pas, en soi, difficile. Mais c'est pénible... Nous essaierons de nous en dispenser et d'obtenir les informations utiles sans calculer complètement cette loi.

II.3 — Fonction de répartition

Définition 2.4 — Fonction de répartition d'une variable aléatoire
Soit X une variable aléatoire. La fonction de répartition de X est la fonction F_X définie par

$$\forall x \in \mathbb{R},\quad F_X(x) \overset{\text{déf.}}{=} P(X \leqslant x)$$

Considérons la variable aléatoire X de loi définie par

$$X(\Omega) = [\![1\,;4]\!] \qquad \text{et} \qquad \forall x \in X(\Omega),\quad P(X=x) = \frac{1}{4}$$

Cette loi s'appelle la loi uniforme sur $[\![1\,;7]\!]$. On la note $\mathscr{U}\big([\![1\,;n]\!]\big)$.

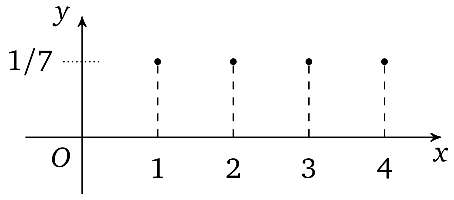


FIGURE I.3 — *Loi uniforme* $\mathscr{U}(\llbracket 1\,;4\rrbracket)$.

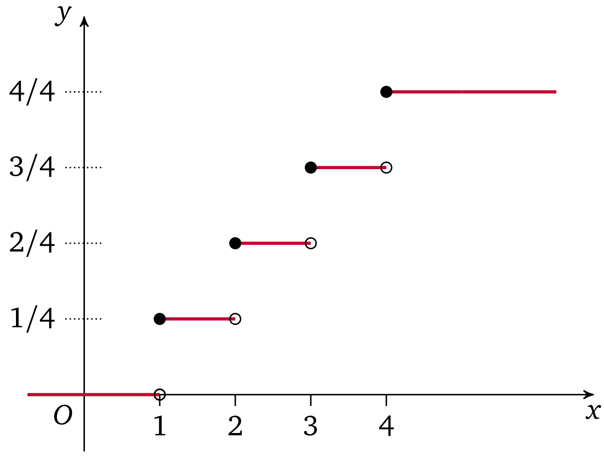


FIGURE I.4 — *Fonction de répartition de la loi uniforme* $\mathscr{U}(\llbracket 1\,;4\rrbracket)$ *(le repère n'est pas normé).*

Propriété 2.5 — Propriété de la fonction de répartition
La fonction de répartition d'une variable aléatoire est une fonction croissante, de limites 0 *en* $-\infty$ *et* 1 *en* $+\infty$.

Dans le cas d'une variable aléatoire discrète finie, la fonction de répartition est une fonction en escalier.

En pratique Elle est peut utile dans le cas des variables finies, mais indispensable dans le cas des variables à densité (vues en Spé). Nous verrons en exercice les quelques cas d'applications utiles.

Théorème 2.6 — Passage de la fonction de répartition à la loi de probabilité
Soit X *une variable aléatoire finie. On note* $X(\Omega)=\left\{x_1,\dots,x_p\right\}$ *avec* $x_1<x_2<\cdots<x_p$.
Passage de la loi de probabilité à la fonction de répartition :

1) $\forall x\in\,]-\infty\,;x_1[\,,\quad F_X(x)=0;$

2) $\forall i\in\llbracket 1\,;p-1\rrbracket,\ \ \forall x\in[x_i\,;x_{i+1}[\,,\quad F_X(x)=\sum_{k=1}^{i}P(X=x_k).$

3) $\forall x \in \left[x_p\,;\,+\infty\right[,\quad F_X(x) = 1;$

Passage de la fonction de répartition à la loi de probabilité :

1) $P(X = x_1) = F_X(x_1);$

2) $\forall i \in [\![2\,;\,p]\!],\quad P(X = x_i) = F_X(x_i) - F_X(x_{i-1}).$

III — Espérance mathématique, valeurs typiques

III.1 — Espérance

Reprenons l'expérience *A*. On cherche à définir la valeur moyenne de X. On peut penser simplement à faire la moyenne arithmétique des valeurs de X, sur toutes les situations possibles. Il y a en tout 32 situations, et donc

$$\text{moyenne}(X) = \frac{1}{32}\Big(0 + \underbrace{1+\cdots+1}_{\times 5} + \underbrace{2+\cdots+2}_{\times 10} + + \underbrace{3+\cdots+3}_{\times 10} + \underbrace{4+\cdots+4}_{\times 5} + 5\Big)$$

$$\text{moyenne}(X) = \frac{1}{32}(0 + 5\times 1 + 10\times 2 + 10\times 3 + 5\times 4 + 5)$$

Ce qui peut aussi s'écrire

$$\begin{aligned} &= P(X=0)\times 0 + P(X=1)\times 1 + P(X=2)\times 2 \\ &\quad + P(X=3)\times 3 + P(X=4)\times 4 + P(X=5) \end{aligned}$$

Cette dernière écriture mène à la définition théorique de la « moyenne » d'une variable aléatoire.

Définition 3.1 — Espérance d'une variable aléatoire

Soit X une variable aléatoire dont l'ensemble des valeurs s'écrit $\{x_1, \cdots, x_n\}$. L'espérance mathématique de X est le réel

$$\mathrm{E}(X) \overset{\text{déf.}}{=} \sum_{k=1}^{n} x_k\, P(X = x_k)$$

On peut aussi écrire cette somme sous la forme

$$\mathrm{E}(X) \overset{\text{déf.}}{=} \sum_{x\in X(\Omega)} x\, P(X = x)$$

L'espérance est donc la moyenne des valeurs de X, moyenne pondérée par les probabilités.

On parle d'*espérance* par opposition à *moyenne* : la moyenne est une valeur effectivement calculée dans un expérience, alors que l'espérance est une valeur limite, une valeur idéale vers laquelle la moyenne va tendre, en répétant l'expérience un nombre croissante de fois.

Notez que l'espérance n'est pas une valeur typique de la variable aléatoire X : ce n'est pas nécessairement une valeur prise par X, elle n'est pas proche d'une valeur très probable de X, etc.
Pensez par exemple à une classe de 20 élèves dont 15 élèves ont 5/20 et 5 élèves ont 20/20. La moyenne de la classe est de 8,75... et ne représente pas grand chose de la répartition des notes. En fait la moyenne n'a de sens que si la loi de probabilité est relativement symétrique et/ou « régulière ».

Propriété 3.2 — *Soit X une variable aléatoire. Si $m = \min(X(\Omega))$ et $M = \max(X(\Omega))$, alors $\mathrm{E}(X) \in [\, m \,;\, M \,]$.*

DÉM. En effet

$$\sum_{k=1}^{n} mP(X = x_k) \leqslant \sum_{k=1}^{n} x_k P(X = x_k) \leqslant \sum_{k=1}^{n} MP(X = x_k)$$

puisque $m \leqslant x_k \leqslant M$ pour tout k

$$m\left[\sum_{k=1}^{n} P(X = x_k)\right] \leqslant \mathrm{E}(X) \leqslant M\left[\sum_{k=1}^{n} MP(X = x_k)\right]$$

$$m \leqslant \mathrm{E}(X) \leqslant M \qquad \text{avec } \sum_{k=1}^{n} P(X = x_k) = 1 \quad \square$$

Corollaire 3.3 — *En particulier, si X est une variable aléatoire positive, alors $\mathrm{E}(X) \geqslant 0$.*

Exemple Calculons l'espérance de X lorsque X suit la loi uniforme $\mathscr{U}\big(()\llbracket 1 \,;\, n \rrbracket)$.

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

III.2 — Formule de transfert & conséquences

On a vu que la loi d'une variable aléatoire est pénibles à déterminer. Heureusement on dispose, pour calculer son espérance, du puissant résultat suivant.

Théorème 3.5 — **Théorème de transfert**
Soit X une variable aléatoire et g une fonction de $\mathbb{R}$ dans $\mathbb{R}$. Alors

$$\mathrm{E}\big(g(X)\big) = \sum_{k=1}^{n} g(x_k)\, P(X = x_k)$$

DÉM. Notons $Y = g(X)$, et $Y(\Omega) = \{y_1, y_2, \dots, y_p\}$.
Écrivons les valeurs de X en les regroupant selon leur image par g

$$X(\Omega) \qquad \underbrace{x_1 \;\; x_2 \;\; \dots x_{k_1}}_{g(x_1)=g(x_2)=\dots=g(x_{k_1})=y_1} \quad \underbrace{x_{k_1+1} \;\; \dots x_{k_2}}_{g(x_{k_1+1})=\dots=g(x_{k_2})=y_2} \quad \cdots$$

Pour calculer $\mathrm{E}\big(g(X)\big)$ on part de la définitions de l'espérance

$$\begin{aligned}\mathrm{E}\big(g(X)\big) = \mathrm{E}(Y) &= \sum_{i=1}^{p} y_p \, P(Y = y_p) \\ &= y_1 \, P(Y = y_1) + y_2 \, P(Y = y_2) + \dots + y_p \, P(Y = y_p)\end{aligned}$$

Pour calculer $P(Y = y_1)$, on peut se ramener à une somme sur certains coefficients de la loi de X. En effet

$$\begin{aligned}P(Y = y_1) &= P\big((X = x_1) \cup (X = x_2) \cup \dots \cup (X = x_{k_1})\big) \\ P(Y = y_1) &= P(X = x_1) + \dots + P(X = x_{k_1}) && \text{év. disjoints} \\ y_1 P(Y = y_1) &= y_1 P(X = x_1) + \dots + y_1 P(X = x_{k_1}) \\ \text{ou aussi} \qquad y_1 P(Y = y_1) &= g(x_1) P(X = x_1) + \dots + g(x_{k_1}) P(X = x_{k_1})\end{aligned}$$

En procédant ainsi pour tous les y_i et en sommant, on trouve

$$\begin{aligned}\mathrm{E}\big(g(X)\big) &= g(x_1)P(X = x_1) + \dots + g(x_{k_1})P(X = x_{k_1}) + \\ &\quad + g(x_{k_1+1})P(X = x_{k_1+1}) + \dots + g(x_n)P(X = x_n) \\ \mathrm{E}\big(g(X)\big) &= \sum_{k=1}^{n} g(x_k) P(X = x_k)\end{aligned}$$

□

Ce résultat important permet d'établir de nouvelles propriétés de l'espérance

Propriété 3.6 — Linéarité de l'espérance – I
Soit X une variable aléatoire et $(a, b) \in \mathbb{R}^2$. On a $\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$.

DÉM. Utilisons la formule de transfert

$$\begin{aligned}\mathrm{E}(aX + b) &= \sum_{k=1}^{n} (a x_k + b) P(X = x_k) \\ &= a \sum_{k=1}^{n} x_k P(X = x_k) + b \sum_{k=1}^{n} P(X = x_k) \\ \mathrm{E}(aX + b) &= a\,\mathrm{E}(X) + b\end{aligned}$$

□

Vocabulaire On dit que X est une variable centrée si et seulement si $\mathrm{E}(X) = 0$

Si X est une variable aléatoire quelconque, on appelle variable centrée associée à X la variable aléatoire $X^0 = X - \mathrm{E}(X)$. On a bien $\mathrm{E}(X^0) = \mathrm{E}(X - \mathrm{E}(X)) = \mathrm{E}(X) - \mathrm{E}(X) = 0$...
Nous utiliserons tout de suite le théorème suivant, qui ne sera démontré que plus tard, mais qui est tellement utile !

Propriété 3.7 — Linéarité de l'espérance – II
Soit X et Y deux variable aléatoire alors $\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$.

DÉM. Ce résultat sera démontré dans le chapitre sur les couples de variables aléatoires. □

III.3 — Moments d'une variable aléatoire — Variance

Définition 3.8 — *Soit $p \in \mathbb{N}$. Le moment d'ordre p de X est le réel* $\mathrm{E}(X^p)$.

Les moments servent à évaluer la dispersion d'une variable aléatoire. En particulier le moment suivant

Définition 3.9 — Variance d'une v.a.r.
Soit X une variable aléatoire. La variance de X est définie par

$$\mathrm{V}(X) \overset{\text{déf.}}{=} \mathrm{E}\left((X - \mathrm{E}(X))^2\right)$$

On appelle écart-type de X le réel $\sigma(X) = \sqrt{\mathrm{V}(X)}$.

Remarquez qu'en tant qu'espérance d'une variable aléatoire positive, la variance est toujours positive. L'écart-type est donc correctement défini.

Interprétation de la variance Il est difficile de donner un sens très concret à la variance. C'est une mesure des écarts de la loi de X par rapport à l'espérance. D'une façon assez grossière, plus la variance est elevée, plus la probabilité de mesurer des valeurs de X éloignées de la moyenne est elevé.
Par exemple, je note deux groupes et je trouve comme moyenne pour les deux groupes 10. Pour un groupe je mesure un écart-type de 1 et pour l'autre un écart-type de 4. Et bien je peux en déduire que le niveau du premier groupe est très homogène, au contraire du second.
La propriété suivante traite d'un cas extrême.

Propriété 3.10 — Cas de la variance nulle
Soit X une variable aléatoire.
On a $\mathrm{V}(X) = 0$ si et seulement si X est une v.a.r. certaine, égale dans ce cas à $\mathrm{E}(X)$.

DÉM. Une variable aléatoire certaine est une variable qui ne prend qu'une seule valeur, avec une probabilité égale à 1.
Supposons $\mathrm{V}(X) = 0$. En calculant cette variance par la formule de transfert

$$\mathrm{V}(X) = \mathrm{E}\big((X - \mathrm{E}(X))^2\big) = \sum_{k=1}^{n} (x_k - \mathrm{E}(X))^2 P(X = x_k) = 0$$

Or cette somme ne porte que sur des réels positifs. Si elle est nulle, c'est que chacun des terme est nul. Ainsi

$$\forall k \in [\![0\,;n]\!], \quad (x_k - \mathrm{E}(X))^2 P(X = x_k) = 0 \quad \Longrightarrow \quad \forall k \in [\![0\,;n]\!], \quad (x_k - \mathrm{E}(X))^2 = 0$$

puisque par hypothèse $P(X = x_k) \neq 0$ (sinon $x_k \neq X(\Omega)$...). On a donc bien prouvé que

$$\forall k \in [\![0\,;n]\!], \quad x_k = \mathrm{E}(X)$$ □

Nous verrons une autre « explication » de la variance dans le cadre du cours de statistiques.

Théorème 3.11 — Théorème de König-Huygens
Soit X une variable aléatoire.

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

DÉM. Aisée avec la linéarité de l'espérance

$$\begin{aligned} \mathrm{V}(X) &= \mathrm{E}\left(X^2 - 2X\,\mathrm{E}(X) + \mathrm{E}(X)^2\right) = \mathrm{E}(X^2) - 2\,\mathrm{E}(X)^2 + \mathrm{E}(X)^2 \\ &= \mathrm{E}(X^2) - \mathrm{E}(X)^2 \end{aligned}$$ □

En pratique On utilise toujours la formule de König-Huygens pour calculer une espérance. Toute la difficulté du calcul est donc reportée sur le calcul de deux espérance.

Exemple Soit X un variable aléatoire suivant la loi uniforme $\mathscr{U}(1, n)$. Nous avons vu que $\mathrm{E}(X) = \dfrac{n+1}{2}$. Calculons $\mathrm{E}(X^2)$, à l'aide de la formule de transfert

$$\mathrm{E}(X^2) = \sum_{k=1}^{n} k^2\, P(X = k) = \sum_{k=1}^{n} \frac{k^2}{n} = \frac{1}{n} \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6n} = \frac{(n+1)(2n+1)}{6}$$

D'où la variance

$$V(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}$$

Propriété 3.13 — *Soit X une variable aléatoire et a et b deux réels.*

- $\mathrm{V}(X+b) = \mathrm{V}(X)$ *et* $\sigma(X+b) = \sigma(X)$ *(invariance par translation) ;*
- $\mathrm{V}(aX) = a^2\,\mathrm{V}(X)$ *et* $\sigma(aX) = |a|\,\sigma(X)$*;*

DÉM. Sans difficultés avec la formule de transfert. □

III.4 — Inégalité de Bienaymé-Tchebychev

L'inégalité de Bienaymé-Tchebychev [1] rend compte également, quoique de façon très approximative, de cette notion d'étalement.

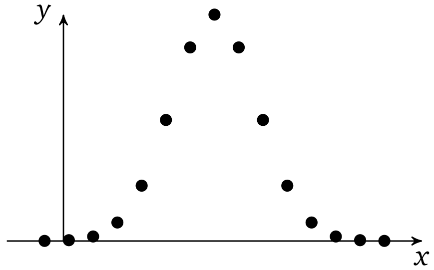
Théorème 3.14 — Inégalité de Bienaymé-Tchebychev
Soit X une variable aléatoire d'espérance $\mathrm{E}(X)$ et de variance $\mathrm{V}(X)$. Alors, pour tout $\varepsilon > 0$ on a

$$P\big(\,|X-\mathrm{E}(X)| \geqslant \varepsilon\big) \leqslant \frac{\mathrm{V}(X)}{\varepsilon^2}$$

DÉM. D'après la définition de la variance et la formule de transfert,

$$\begin{aligned}
V(X) &= \sum_{x\in X(\Omega)} (x-\mathrm{E}(X))^2 P(X=x)\\
V(X) &= \sum_{x\in X(\Omega), |x-\mathrm{E}(X)|\geqslant\varepsilon} (x-\mathrm{E}(X))^2 P(X=x)\\
&\qquad + \sum_{x\in X(\Omega), |x-\mathrm{E}(X)|<\varepsilon} (x-\mathrm{E}(X))^2 P(X=x)\\
V(X) &\geqslant \sum_{x\in X(\Omega), |x-\mathrm{E}(X)|\geqslant\varepsilon} (x-\mathrm{E}(X))^2 P(X=x)\\
V(X) &\geqslant \varepsilon^2 \sum_{x\in X(\Omega), |x-\mathrm{E}(X)|\geqslant\varepsilon} P(X=x)\\
\frac{V(X)}{\varepsilon^2} &\geqslant P\left(|X-\mathrm{E}(X)| \geqslant \varepsilon\right)
\end{aligned}$$

□



1. Irénée-Jules Bienaymé, (1796–1878), probabiliste et statisticien français et Pafnouti Lvovitch Tchebychev (1821–1894), mathématicien russe.

Cette inégalité permet d'évaluer des probabilités, comme nous le verrons dans plusieurs exercices. En voici un exemple

Exemple La moyenne de la classe est de 10,5 avec un écart-type de 1. Vous avez 12,5. Est-ce un résultat significativement au-dessus de la moyenne ?
Notons X la variable aléatoire égale au résultat d'un élève. La note 12,5 réalise l'évènement $(X - \mathrm{E}(X)) \geqslant 2$. D'après l'inégalité de Bienaymé-Tchebychev cet évènement à une probabilité inférieure à $(1^2)/(2.5)^2$ soit 16%. C'est assez rare, mais ça ne semble pas exceptionnel.
En réalité, il faut savoir que l'inégalité de Bienaymé-Tchebychev donne une approximation exécrable de la probabilité.
Si un paquet de copie est « bien noté », sa répartition de note suit une gaussienne. Dans ce cas, 95% des notes sont dans l'intervalle $[\,10 - 1.96\sigma\,;\,10 + 1.96\sigma\,]$. Un écart-type de 1 dans un paquet de copies signifie, en réalité, que quasiment tout le monde a entre 8 et 12. La note 12,5 est en fait excellente !
Toutefois, la seule réponse rigoureuse possible viendrait d'une analyse de l'histogramme des notes...

IV — Lois usuelles

Nous commençons maintenant l'étude de lois usuelles. Lorsqu'une expérience permet de reconnaître l'une de ces lois... et bien on saute sur l'occasion !

IV.1 — Variable certaine

Cette variable est un cas limite. On a vu que X suit la loi certaine si et seulement si

$$X(\Omega) = \{x_0\} \qquad \text{et} \qquad P(X = x_0) = 1$$

Dans ce cas

$$\mathrm{E}(X) = x_0 \qquad \text{et} \qquad V(X) = 0$$

IV.2 — Loi uniforme

Définition 4.1 — Loi uniforme
Soit n un entier naturel non nul.
On dit qu'une variable aléatoire X suit la loi uniforme sur $[\![\,1\,;\,n\,]\!]$ si et seulement si

1) $X(\Omega) = [\![\,1\,;\,n\,]\!]$*;*

2) $\forall k \in [\![1 ; n]\!], \quad P(X=k) = \dfrac{1}{n}$

On note alors $X \hookrightarrow \mathscr{U}(1, n)$.

Situation type On tire un nombre au hasard entre 1 et n sans favoriser un résultat par rapport à un autre. La variable aléatoire X est alors le résultat obtenu.

Espérance, variance On a vu que

$$\mathrm{E}(X) = \frac{n+1}{2} \qquad \text{et} \qquad V(X) = \frac{n^2-1}{12}$$

IV.3 — Loi de Bernoulli

Paramètre Un réel $p \in]0 ; 1[$.

Définition 4.2 — Loi de Bernoulli
Soit $p \in]0;1[$. Une variable aléatoire X suit la loi de Bernoulli de paramètre p si et seulement si
– $X(\Omega) = \{0, 1\}$;
– $P(X=1) = p$ *(et donc* $P(X=0) = 1-p = q$*).*
On note alors $X \hookrightarrow \mathscr{B}(1, p)$. *La variable aléatoire X est qualifiée de variable aléatoire de Bernoulli.*

Situation type Toute expérience a deux issues : on attribue arbitrairement la valeur 1 à un « succès » et la valeur 0 à un « échec ». Par exemple le lancer d'une pièce de monnaie ($1=$ obtention d'un pile), naissance d'un enfant ($1=$ naissance d'une fille), etc.

Espérance, variance Par un calcul simple : $\mathrm{E}(X) = p$; $V(X) = pq$.
Notez que pour $p=1$ (ou $p=0$) la variable X est certaine, égale à 1 (ou à 0).

Définition 4.3 — Variable indicatrice
Soit A un évènement. On associe à A la variable X_A définie par

$$\forall \omega \in \Omega, \qquad X_A(\omega) = 1 \quad \textit{si} \quad \omega \in A \qquad \textit{et} \qquad X_A(\omega) = 0 \quad \textit{sinon}$$

La variable aléatoire X_A est la variable indicatrice de l'évènement A.

En terme plus simple : l'évènement A est considéré comme un succès alors que $\overline{A}$ est un échec. La variable aléatoire X_A compte alors le nombre de succès. On verra l'intérêt de ce type de variable en exercice. Retenons pour le moment le résultat intuitivement évident

Propriété 4.4 — *Soit A un évènement et X_A sa variable indicatrice. On a*

$$\mathrm{E}(X_A) = P(X_A = 1) = P(A)$$

IV.4 — Loi binomiale

Paramètres Un réel $p \in \,]0\,;1[$ et un entier $n \in \mathbb{N}^*$.

Situation type On considère une succession de n expérience de type succès-échec *indépendantes*. On suppose qu'à chaque expérience la probabilité de succès est p.
On compte le nombre X de succès ainsi obtenu.
D'une part $X(\Omega) = [\![0\,;n]\!]$.
D'autre part, soit $k \in X(\Omega)$. Notons S_i l'évènement « on a obtenu un succès à la i—ième expérience » et E_i l'évènement « on a obtenu un échec à la i—ième expérience ». On a

$$\begin{aligned} P(X=k) = P(S_1 \cap S_2 \cap \cdots \cap S_k \cap E_{k+1} \cap E_{k+2} \cap \cdots \cap E_n) + \\ P(S_1 \cap \cdots S_{k-1} \cap E_k \cap S_{k+1} \cap E_{k+2} \cap \cdots \cap E_n) + \cdots \\ P(E_1 \cap \cdots E_{n-k} \cap S_{n-k+1} \cap \cdots \cap S_n) + \cdots \end{aligned}$$

où la somme porte sur toutes les permutations de k succès et $n-k$ échecs parmi les n expériences. Il y a donx $\binom{n}{k}$ termes dans cette somme. Par indépendance des expériences succès-échec, on a

$$P(S_1 \cap S_2 \cap \cdots \cap S_k \cap E_{k+1} \cap E_{k+2} \cap \cdots \cap E_n) = p^k q^{n-k}$$

En permutant les succès et les échecs, on ne modifie pas cette valeur. Donc tous les termes de la somme ci-dessous sont égaux, et finalement

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

Définition 4.5 — Loi binomiale
Soit $n \in \mathbb{N}^*$ et $p \in \,]0\,;1[$. Une variable aléatoire X suit la loi binomiale de paramètres n et p si et seulement si
- $X(\Omega) = [\![\,0\,;n\,]\!]$;
- $\forall k \in [\![\,0\,;n\,]\!], \quad P(X=k) = \binom{n}{k} p^k q^{n-k}$.

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

Situation type Une succession de n expériences de type succès-échec *indépendantes*. On suppose qu'à chaque expérience la probabilité de succès est *constante* égale à p.
La variable aléatoire X est le nombre de succès observés.

Espérance et variance $\mathrm{E}(X) = np$ et $V(X) = npq$.

Espérance Le calcul est classique et utilise la formule $k\binom{n}{k} = n\binom{n-1}{k-1}$:

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{k=0}^{n} kP(X=k) = \sum_{k=0}^{n} k\binom{n}{k}p^k q^{n-k} \\
&= \sum_{k=1}^{n} n\binom{n-1}{k-1}p^k q^{n-k} && \text{petite formule} \\
&= n\sum_{j=0}^{n-1}\binom{n-1}{j}p^{j+1}q^{n-1-j} = np\left(\sum_{j=0}^{n-1}\binom{n-1}{j}p^j q^{n-1-j}\right) \\
\mathrm{E}(X) &= np(p+q)^{n-1} = np
\end{aligned}$$

Variance Le calcul est classique. Une petite astuce consiste à calculer $\mathrm{E}(X(X-1))$ en utilisant deux fois la formule $k\binom{n}{k} = n\binom{n-1}{k-1}$. D'après la formule de transfert :

$$\begin{aligned}
\mathrm{E}(X(X-1)) &= \sum_{k=0}^{n} k(k-1)P(X=k) = \sum_{k=2}^{n} k(k-1)\binom{n}{k}p^k q^{n-k} \\
&= \sum_{k=2}^{n} n(k-1)\binom{n-1}{k-1}p^k q^{n-k} = \sum_{k=2}^{n} n(n-1)\binom{n-2}{k-2}p^k q^{n-k} \\
&= n(n-1)\sum_{j=0}^{n-2}\binom{n-2}{j}p^{j+2}q^{n-2-k} \\
\mathrm{E}(X(X-1)) &= n(n-1)p^2
\end{aligned}$$

Par linéarité de l'espérance, on a (c'est l'astuce !)

$$\mathrm{E}(X(X-1)) = \mathrm{E}(X^2 - X) = \mathrm{E}(X^2) - \mathrm{E}(X)$$

donc $$\mathrm{E}(X^2)s = \mathrm{E}(X(X-1)) + \mathrm{E}(X) = n(n-1)p^2 + np$$

On trouve ainsi $V(X) = npq$.

IV.5 — Loi hypergéométrique

Situation type
Une urne contient a boules blanches et b boules noires. On tire une poignée de n boules dans l'urne. On appelle X la variable aléatoire égale au nombre de boules blanches dans la poignée.
On peut aussi tirer les boules une à une et sans remise et trouver le même résultat.
D'une part $X(\Omega) = [\![\max(0, n-b)\ ;\ \min(n, a)]\!]$. En effet, si $n > a$ (si on tire plus de boules qu'il n'y a de boules blanches dans l'urne) la valeur maximale de X est a. Sinon c'est n. Etc.

D'autre part, on peut tirer $\binom{N}{n}$ poignées dans l'urne, en notant $N = a + b$ le nombre total de boules dans l'urne. Toutes ces poignées sont équiprobables.
Soit $k \in X(\Omega)$. Parmi ces poignées, $\binom{a}{k}\binom{b}{n-k}$ contiennent k boules blanches et $n-k$ boules noires. Finalement

$$P(X = k) = \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{N}{n}}$$

En sommant ces coefficients, on obtient un résultat utile, la formule de Vandermonde (qui est hors programme) :

$$\sum_{x \in X(\Omega)} P(X = x) = \sum_{k=0}^{n} \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{N}{n}} = 1 \qquad \text{donc} \qquad \sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} = \binom{N}{n}$$

Notez que la seconde somme a été prise de 0 à n, c'est-à-dire au-delà de l'ensemble $X(\Omega)$. Étant donné les conventions sur les coefficients binomiaux, cela ne pose pas de problème.

Définition 4.6 — Loi hypergéométrique
Soit $N \in \mathbb{N}^*$, $n \in [\![0 \,;\, N]\!]$ et $p \in \,]0\,;1[$. Une variable aléatoire X suit la loi hypergéométrique de paramètres N, n et p si et seulement si
- $X(\Omega) = [\![\max(0, n - N(1-p)) \,;\, \min(n, Np)]\!]$;
- $\forall k \in X(\Omega), \quad P(X = k) = \dfrac{\binom{Np}{k}\binom{N(1-p)}{n-k}}{\binom{N}{n}}$.

On note alors $X \hookrightarrow \mathscr{H}(N, n, p)$.

Situation type Une urne contient a boules blanches et b boules noires. On tire une poignée de n boules dans l'urne. On appelle X la variable aléatoire égale au nombre de boules blanches dans la poignée.

Espérance et variance $\mathrm{E}(X) = \dfrac{an}{N} = np$ et $V(X) = \dfrac{npq(N-n)}{N-1}$.
On retrouve la formule précédente en interprétant $p = a/N$ comme la proportion de boules blanches dans l'urne, au départ.

Espérance Le calcul est classique et utilise la formule $k\binom{n}{k} = n\binom{n-1}{k-1}$:

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{k=0}^{n} kP(X = k) = \sum_{k=0}^{n} k \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{N}{n}} \\
&= \sum_{k=1}^{n} a \frac{\binom{a-1}{k-1}\binom{b}{n-k}}{\binom{N}{n}} \\
&= \frac{a}{\binom{N}{n}} \sum_{j=0}^{n-1} \binom{a-1}{j}\binom{b}{(n-1)-j} = \frac{a}{\binom{N}{n}} \binom{N-1}{n-1} && \text{formule de Vandermonde}
\end{aligned}$$

$$\mathrm{E}(X) = \frac{an}{N} = np$$

Variance On utilise la même astuce que précedemment. D'après la formule de transfert :

$$\begin{aligned}
\mathrm{E}(X(X-1)) &= \sum_{k=0}^{n} k(k-1)P(X=k) \\
&= \sum_{k=2}^{n} k(k-1)\frac{\binom{a}{k}\binom{b}{n-k}}{\binom{N}{n}} \\
&= \sum_{k=2}^{n} a(a-1)\frac{\binom{a-2}{k-2}\binom{b}{n-k}}{\binom{N}{n}} \\
&= \frac{a(a-1)}{\binom{N}{n}} \sum_{k=2}^{n} \binom{a-2}{j}\binom{b}{n-2-j} \\
&= \frac{a(a-1)}{\binom{N}{n}} \binom{N-2}{n-2} \\
\mathrm{E}(X(X-1)) &= \frac{a(a-1)n(n-1)}{N(N-1)}
\end{aligned}$$

On en déduit $\mathrm{E}(X^2) - \mathrm{E}(X) = \dfrac{a(a-1)N(N-1)}{n(n-1)}$, donc

$$\begin{aligned}
V(X) &= \frac{a(a-1)n(n-1)}{N(N-1)} + \frac{an}{N} - \frac{a^2n^2}{N^2} \\
&= \frac{a(a-1)n(n-1)}{N(N-1)} + \frac{an}{N} - \frac{a^2n^2}{N^2} \\
&= \frac{Na(a-1)n(n-1) + anN(N-1) - a^2n^2(N-1)}{N^2(N-1)} \\
&= \frac{na(N(a-1)(n-1) + N(N-1) - an(N-1)}{N^2(N-1)} \\
&= \frac{na(Nan - Na - Nn + N + N^2 - N - anN + an)}{N^2(N-1)} \\
&= \frac{na(-Na - Nn + N^2 + an)}{N^2(N-1)} \\
&= \frac{na(N-a)(N-n)}{N^2(N-1)} \\
V(X) &= \frac{npq(N-n)}{N-1}
\end{aligned}$$

Approximation de la loi hypergéométrique par la loi binomiale
Soit $k \in \mathbb{N}$. D'après l'expression de la loi hypergéométrique

$$P(X=k) = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}$$

On a noté $q = 1 - p$. Comme Np et Nq sont proportionnels à N, pour N suffisament grand on a $k \leqslant Np$, $k \leqslant Nq$ et $n \leqslant N$ (on se rappelle que n et k sont fixés dans ce calcul). Il est donc possible d'utiliser l'expression factorielle des combinaisons. Ainsi

$$\begin{aligned}
P(X=k) &= \frac{(Np)!\,(Nq)!\,(N-n)!\,n!}{N!\,k!\,(Np-k)!\,(n-k)!\,(Nq-n+k)!} \\
&= \frac{(Np)!}{(Np-k)!}\frac{(Nq)!}{(Nq-n+k)!}\frac{(N-n)!}{N!}\frac{n!}{k!(n-k)!} \qquad \text{en réordonnant} \\
P(X=k) &= \binom{n}{k} \times \frac{Np}{N}\frac{Np-1}{N-1}\cdots\frac{Np-k+1}{N-k+1} \times \frac{Nq}{N-k}\frac{Nq-1}{N-k-1}\cdots\frac{Nq-n+k+1}{N-n+1}
\end{aligned}$$

Les k premières fractions tendent toutes vers p lorsque N tend vers $+\infty$. Les $n-k$ suivantes tendent quant à elles vers q. Ainsi, en passant à la limite sur N

$$P(X=k) \xrightarrow[N\to+\infty]{} \binom{n}{k} p^k q^{n-k}$$

D'ailleurs $\mathrm{E}(X) = np$ est aussi l'espérance de la loi binomiale $\mathscr{B}(n,p)$.
Quand à la variance on a $V(X) = \dfrac{N-n}{N-1} npq \xrightarrow[N\to+\infty]{} npq$.