

Variables aléatoires réelles

ECE1 2018–2019

Sommaire

1	Définition et généralités	1
1.1	Définition et exemples	1
1.2	Fonction de répartition et loi d'une VAR	3
2	Variables aléatoires réelles discrètes (VARD)	4
2.1	Définition et propriétés	4
2.2	VARD obtenues par compositions	5
2.3	Moments d'une VARD	5
2.3.1	Espérance d'une VARD	5
2.3.2	Moments, variance et écart-type d'une VARD	8
3	Lois usuelles	9
3.1	Lois discrètes finies	9
3.1.1	Loi certaine	9
3.1.2	Loi uniforme	9
3.1.3	Loi de Bernoulli	10
3.1.4	Loi binomiale	11
3.2	Lois discrètes infinies	14
3.2.1	Loi géométrique	14
3.2.2	Loi de Poisson	15
4	Variables à densité	16
4.1	Définition et propriétés	16
4.2	Espérance d'une variable à densité	17
4.3	Lois à densité usuelles	18
4.3.1	Loi uniforme	18
4.3.2	Loi exponentielle	20
4.3.3	Loi normale	22

Une variable aléatoire réelle (VAR) est grosso modo une fonction qui envoie chaque résultat d'une expérience aléatoire sur un nombre réel (plus certaines conditions), par exemple le nombre de fois qu'on a obtenu "pile" en jouant à pile ou face un certain nombre de fois, la somme des numéros affichés par deux dés qu'on a jetés, etc.

1 Définition et généralités

1.1 Définition et exemples

Définition

Variable aléatoire réelle (VAR)

Soit $(\Omega, \mathcal{A})$ un espace probablisable. On dit qu'une fonction $X : \Omega \rightarrow \mathbb{R}$ est une *variable aléatoire (réelle)* sur $(\Omega, \mathcal{A})$ si pour tout $x \in \mathbb{R}$, l'ensemble $\{\omega \in \Omega : X(\omega) \leq x\}$ est un évènement (c'est-à-dire, pour rappel, si cet ensemble appartient à la tribu $\mathcal{A}$).

Remarque

On peut alors montrer, en revenant aux propriétés d'une tribu d'évènements, que les ensembles

$$\begin{aligned} &\{\omega \in \Omega : X(\omega) > x\}, \\ &\{\omega \in \Omega : X(\omega) \in]a, b]\}, \\ &\{\omega \in \Omega : X(\omega) = x\}, \\ &\{\omega \in \Omega : X(\omega) < x\}, \\ &\{\omega \in \Omega : X(\omega) \geq x\}, \\ &\{\omega \in \Omega : X(\omega) \in [a, b]\}, \\ &\{\omega \in \Omega : X(\omega) \in [a, b[\}, \\ &\{\omega \in \Omega : X(\omega) \in]a, b]\} \end{aligned}$$

sont également des évènements pour tous réels x, a et b .

Remarque

Toute fonction $f : \Omega \rightarrow \mathbb{R}$ est une variable aléatoire sur l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$, puisque les évènements de $(\Omega, \mathcal{P}(\Omega))$ sont simplement les sous-ensembles de Ω .

Note

Démontrer qu'une fonction est une variable aléatoire n'est pas exigible, on se contente de l'admettre dans les exercices.

Notation

Etant donnée une variable aléatoire X sur un espace probabilisable $(\Omega, \mathcal{A})$, on note

$$[X \in I] := \{\omega \in \Omega : X(\omega) \in I\}$$

pour tout ensemble $I \subset \mathbb{R}$, en particulier

$$[X \leq x] = \{\omega \in \Omega : X(\omega) \leq x\},$$

$$[X = x] = \{\omega \in \Omega : X(\omega) = x\},$$

etc...

Si l'on dispose d'une probabilité P sur $(\Omega, \mathcal{A})$, on note également

$$P(X \leq x) = P([X \leq x]),$$

$$P(X = x) = P([X = x]),$$

etc...

Important

Il est toujours utile, voire demandé, de préciser l'image d'une variable aléatoire $X : \Omega \rightarrow \mathbb{R}$, c'est-à-dire l'ensemble $X(\Omega) = \{X(\omega) : \omega \in \Omega\}$.

Exemple/Exercice 1

On joue à pile ou face n fois d'affilée, $(\Omega, \mathcal{A})$ est l'espace probabilisé fini où $\Omega = \{\text{pile, face}\}^n$ et $\mathcal{A} = \mathcal{P}(\Omega)$. Pour tout tirage $\omega \in \Omega$, on définit

$$X(\omega) = \text{nombre de fois où on a obtenu pile.}$$

La fonction $X : \Omega \rightarrow \mathbb{R}$ est une variable aléatoire sur $(\Omega, \mathcal{A})$.

Notons que $X(\Omega)$ (l'ensemble des valeurs prises par X) est plus précisément $\{0, 1, 2, \dots, n\} \subset \mathbb{R}$.

Révisions chapitre 5 : montrer que si la pièce est équilibrée, alors $P(X = k) = \frac{\binom{n}{k}}{2^n}$ pour tout $k \in \{0, 1, \dots, n\}$.

Note : autrement dit, X suit la loi binomiale $\mathcal{B}(n, \frac{1}{2})$ (voir Section 3.1.4).

Exemple/Exercice 2

On lance un dé équilibré une infinité de fois (l'espace probabilisable est ici infini), la fonction X qui associe à chaque tirage le nombre maximum que l'on obtient sur tous les lancers est une variable aléatoire, d'image $X(\Omega) = \{1, 2, 3, 4, 5, 6\} \subset \mathbb{R}$.

Exercice : montrer que $P(X = 6) = 1$ et $P(X = i) = 0$ pour tout $i \in \{1, 2, 3, 4, 5\}$.

Exemple/Exercice 3

On lance simultanément deux dés une seule fois (espace probabilisable fini), la fonction X qui associe chaque lancer à la somme des dés est une variable aléatoire, d'image $X(\Omega) = \{2, 3, 4, \dots, 12\} \subset \mathbb{R}$.

Exercice : calculer $P(X = k)$ pour tout $k \in \{2, 3, 4, \dots, 12\}$.

Exemple A

On choisit un réel x "au hasard" (voir note ci-dessous) dans un segment $[a, b]$, la fonction X qui donne ce réel aléatoire x est une variable aléatoire d'image $X(\Omega) = [a, b]$ (ici en fait $\Omega = [a, b]$ et X est la fonction identité sur Ω).

Note : choisir un réel au hasard entre a et b signifie ici que X suit la loi uniforme continue $\mathcal{U}([a, b])$ (voir Section 4.3.1).

Exemple B

On patiente à la caisse d'un supermarché, la fonction X qui donne le temps $t \geq 0$ qu'on aura attendu avant de payer est une variable aléatoire d'image $X(\Omega) = [0, +\infty[$.

Note : il est raisonnable de modéliser X par une loi exponentielle $\mathcal{E}(\lambda)$ où $\frac{1}{\lambda}$ est le temps moyen qu'on attend avant de payer (voir Section 4.3.2).

1.2 Fonction de répartition et loi d'une VAR

Définition

Fonction de répartition d'une VAR

Soit X une variable aléatoire réelle sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, on nomme *fonction de répartition de X* (notée F_X) la fonction

$$\begin{aligned} F_X : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto P(X \leq x). \end{aligned}$$

Propriété

Pour toute variable aléatoire X , la fonction de répartition F_X de X est croissante et continue à droite en tout point (c'est-à-dire $\lim_{x \rightarrow x_0^+} F_X(x) = F_X(x_0)$ pour tout $x_0 \in \mathbb{R}$), et

$$\begin{aligned} \lim_{x \rightarrow -\infty} F_X(x) &= 0 \\ \lim_{x \rightarrow +\infty} F_X(x) &= 1. \end{aligned}$$

Définition

Loi d'une VAR

Soit X une variable aléatoire réelle sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, on nomme *loi de X* la donnée de $X(\Omega)$ et des réels $\{P(X = x) : x \in X(\Omega)\}$.

Propriété

La fonction de répartition caractérise la loi d'une variable aléatoire, c'est-à-dire que l'on peut déduire $X(\Omega)$ et $P(X = x)$ pour tout $x \in X(\Omega)$ seulement à partir de F_X .

Exemple/Exercice 4

On lance un dé équilibré, on note X la variable aléatoire donnant le numéro du dé.

La loi de X est donnée par $X(\Omega) = \{1, 2, 3, 4, 5, 6\}$ et $P(X = i) = \frac{1}{6}$ pour tout $i \in [1, 6]$ (**autrement dit, X suit la loi uniforme discrète $\mathcal{U}([1; 6])$, voir Section 3.1.2).**

Tracer le graphe de F_X .

Propriété

- Pour tout $x \in \mathbb{R}$,

$$P(X > x) = 1 - F_X(x).$$

- Pour tous réels $a \leq b$,

$$F_X(b) - F_X(a) = P(X \in]a, b]).$$

Démonstration

- Le complémentaire de l'évènement $[X \leq x]$ est $[X > x]$ donc $P(X > x) = 1 - P(X \leq x) = 1 - F_X(x)$.
- Si $a \leq b$ alors $[X \leq a] \subset [X \leq b]$ et

$$\begin{cases} [X \leq b] = [X \leq a] \cup [X \in]a, b] \\ [X \leq a] \cap [X \in]a, b] = \emptyset \end{cases}$$

par conséquent

$$P(X \leq b) = P(X \leq a) + P(X \in]a, b]) \Rightarrow P(X \in]a, b]) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a).$$

2 Variables aléatoires réelles discrètes (VARD)

2.1 Définition et propriétés

Définition

VARD

Une variable aléatoire X sur un espace probabilisable est dite *discrète* si $X(\Omega)$ est indexé par un sous-ensemble fini ou infini de $\mathbb{Z}$, c'est-à-dire qu'il existe des entiers $i_1 < i_2 < i_3 \dots$ tels que $X(\Omega) = \{x_{i_1}, x_{i_2}, x_{i_3}, \dots\}$.

Exemples

Les Exemples/Exercices de 1 à 4 décrivent des VARD (où $X(\Omega)$ est fini), mais pas les Exemples A et B.

Exemple/exercice

On lance un dé équilibré une infinité de fois. Soit X le nombre de lancers nécessaires pour obtenir le premier 6 (défini comme valant 0 si on n'obtient jamais 6), X est une VARD d'image $X(\Omega) = \mathbb{N}$.

Exercice : montrer que $P(X = n) = \left(\frac{5}{6}\right)^{n-1} \times \frac{1}{6}$ pour tout $n \geq 1$ et que $P(X = 0) = 0$.

(Autrement dit, X suit la loi géométrique $\mathcal{G}\left(\frac{1}{6}\right)$, voir Section 3.2.1.)

Exercices

Exercices 1,2 et 3 du TD14.

Propriété

Si une VARD X est à valeurs dans $\mathbb{Z}$ (c'est-à-dire, si les éléments de $X(\Omega)$ sont des entiers), la fonction de répartition F_X de X caractérise la loi de X via l'équation

$$P(X = k) = F_X(k) - F_X(k-1)$$

pour tout $k \in \mathbb{Z}$.

Démonstration

Pour tout $k \in \mathbb{Z}$,

$$F_X(k) - F_X(k-1) = P(X \leq k) - P(X \leq k-1) = P(X \in]k-1, k]) = \underbrace{P(X \in]k-1, k])}_0 + P(X = k) = P(X = k)$$

puisque les valeurs prises par X sont entières.

Exercice

Reprenons l'exemple précédent du dé équilibré lancé une infinité de fois, et X la variable aléatoire affichant le nombre de lancers nécessaires pour obtenir 6.

Calculer $F_X(k)$ pour tout $k \in \mathbb{Z}$.

Propriété

SCE associé à une VARD

Soit X une VARD, notons $X(\Omega) = \{x_i : i \in I\}$ où I est un sous-ensemble de $\mathbb{Z}$, alors la famille d'évènements

$$([X = x_i])_{i \in I}$$

est un système complet d'évènements.

Démonstration

Il s'agit de montrer que ces évènements sont incompatibles deux à deux et que leur union est Ω tout entier.

- Si $i \neq j$, on a $[X = x_i] \cap [X = x_j] = [x_i = X = x_j] = \emptyset$.
- On a $\bigcup_{i \in I} [X = x_i] = [\text{il existe } i \in I \text{ tel que } X = x_i] = [X \in X(\Omega)] = \Omega$.

Remarque

En corollaire de la propriété précédente, on a

$$\sum_{i \in I} P(X = x_i) = P(\Omega) = 1.$$

Exercices

Exercice 4 du TD14.

2.2 VARD obtenues par compositions

Proposition/définition

VARD composées

Soit X une VARD et g une fonction à valeurs réelles définie sur $X(\Omega)$.
Alors : la fonction $Y = g(X)$ définie par

$$Y(\omega) = g(X(\omega))$$

pour tout $\omega \in \Omega$ (autrement dit, Y est la composition $g \circ X$), est aussi une variable aléatoire (réelle discrète).

Exemple

On lance un dé équilibré, on pose X = le nombre affiché par le dé. La loi de X est donc donnée par $X(\Omega) = \{1, 2, 3, 4, 5, 6\}$ et $P(X = i) = \frac{1}{6}$ pour tout $i \in \{1, 2, 3, 4, 5, 6\}$ (**autrement dit** $X \hookrightarrow \mathcal{U}(\{1; 6\})$), voir Section 3.1.2).
On considère la VARD $Y = 2X - 1$, c'est-à-dire $Y = g(X)$ où $g(x) = 2x - 1$ pour tout $x \in \mathbb{R}$.
La loi de Y est donnée par $Y(\Omega) = \{1, 3, 5, 7, 9, 11\}$ et

$$P(Y = i) = P(2X - 1 = i) = P(2X = i + 1) = P\left(X = \frac{i + 1}{2}\right) = \frac{1}{6}$$

pour tout $i \in \{1, 3, 5, 7, 9, 11\}$.

Exemple

Considérons une VARD X dont la loi est donnée par $X(\Omega) = \{-2, -1, 0, 1, 2\}$ et $P(X = i) = \frac{1}{5}$ pour tout $i \in \{-2, -1, 0, 1, 2\}$.
Étudions la loi de la VARD $Y = X^2$ (c'est-à-dire $Y = g(X)$ où $g(x) = x^2$ pour tout $x \in \mathbb{R}$).
 Y a pour image $Y(\Omega) = g(\{-2, -1, 0, 1, 2\}) = \{0, 1, 4\}$ et

$$\begin{aligned} P(Y = 0) &= P(X^2 = 0) = P(X = 0) = \frac{1}{5}, \\ P(Y = 1) &= P(X^2 = 1) = P(X = -1 \text{ ou } X = 1) = P(X = -1) + P(X = 1) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}, \\ P(Y = 4) &= P(X^2 = 2) = P(X = -2 \text{ ou } X = 2) = P(X = -2) + P(X = 2) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}. \end{aligned}$$

2.3 Moments d'une VARD

2.3.1 Espérance d'une VARD

L'espérance d'une VARD, si elle existe, est la "valeur moyenne" que peut prendre cette variable aléatoire.

Définition

Espérance d'une VARD lorsque $X(\Omega)$ est fini

Soit X une variable aléatoire réelle discrète telle que $X(\Omega)$ soit fini, notons $X = \{x_1, x_2, \dots, x_n\}$.
On définit l'espérance de X , notée $E(X)$, comme le réel

$$E(X) = \sum_{k=1}^n x_k \times P(X = x_k).$$

Remarque

Si X est une variable "aléatoire" constante (c'est-à-dire, s'il existe $x \in \mathbb{R}$ tel que $X(\Omega) = \{x\}$, ainsi la loi de X est donnée par $P(X = x) = 1$), alors X a pour espérance $E(X) = x$.

Exemple

On lance un dé équilibré, X = numéro affiché par le dé (**ainsi** $X \hookrightarrow \mathcal{U}([1;6])$), **voir Section 3.1.2**).

On a $X(\Omega) = [1;6]$ et $P(X = i) = \frac{1}{6}$ pour tout $i \in [1,6]$.

L'espérance de X est alors

$$E(X) = \sum_{i=1}^6 i \times \frac{1}{6} = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = \frac{1+2+3+4+5+6}{6} = \frac{\frac{6 \times 7}{2}}{6} = \frac{7}{2} = 3,5.$$

Exemple

On lance deux fois successivement une pièce truquée ayant trois chances sur quatre de tomber sur "pile".

On note X la variable aléatoire donnant le nombre de fois où l'on est tombé sur "pile" lors de ces deux lancers.

La loi de X est donnée par $X(\Omega) = \{0, 1, 2\}$ et

$$\begin{aligned} P(X = 0) &= \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}, \\ P(X = 1) &= \frac{3}{4} \times \frac{1}{4} + \frac{1}{4} \times \frac{3}{4} = \frac{6}{16}, \\ P(X = 2) &= \frac{3}{4} \times \frac{3}{4} = \frac{9}{16}. \end{aligned}$$

Ainsi l'espérance de X est

$$E(X) = 0 \times \frac{1}{16} + 1 \times \frac{6}{16} + 2 \times \frac{9}{16} = \frac{24}{16} = \frac{3}{2} = 1,5.$$

Remarque : X suit en fait la loi binomiale $\mathcal{B}(n, p)$ où $n = 2$ et $p = \frac{3}{4}$ (**voir Section 3.1.4**).

On a alors la formule $E(X) = np = 2 \times \frac{3}{4} = \frac{3}{2}$.

Définition

Espérance d'une VARD lorsque $X(\Omega)$ est infini

Soit X une variable aléatoire réelle discrète telle que $X(\Omega)$ soit infini, notons $X(\Omega) = \{x_n : n \geq n_0\}$.

On dit que X *admet une espérance* si la série $\sum_{n \geq n_0} x_n \times P(X = x_n)$ converge absolument (donc simplement), et l'on définit alors *l'espérance de X* , notée $E(X)$, comme le réel

$$E(X) = \sum_{n=n_0}^{+\infty} x_n \times P(X = x_n).$$

Exemple

On lance un dé équilibré jusqu'à ce qu'on obtienne 6, on note X la variable aléatoire affichant le nombre de lancers nécessaires pour obtenir 6 (défini comme 0 si on n'obtient jamais 6), on a vu que $X(\Omega) = \mathbb{N}$ et

$$\begin{aligned} P(X = 0) &= 0, \\ P(X = n) &= \left(\frac{5}{6}\right)^{n-1} \times \frac{1}{6} \quad \text{pour tout } n \geq 1. \end{aligned}$$

Pour déterminer si X admet une espérance, il faut étudier l'absolue convergence de la série $\sum_{n \geq 1} n \times \left(\frac{5}{6}\right)^{n-1} \times \frac{1}{6}$, c'est-à-dire sa convergence tout court car c'est une série à termes positifs.

Or on sait que la série $\sum_{n \geq 1} nq^{n-1}$ converge vers $\frac{1}{(1-q)^2}$ pour tout $q \in]-1, 1[$, donc X admet effectivement l'espérance

$$E(X) = \frac{1}{6} \times \sum_{n=1}^{+\infty} n \left(\frac{5}{6}\right)^{n-1} = \frac{1}{6} \frac{1}{\left(1 - \frac{5}{6}\right)^2} = 6.$$

Remarque : X suit en fait la loi géométrique $\mathcal{G}(p)$ où $p = \frac{1}{6}$ (**voir Section 3.2.1**).

On a alors la formule $E(X) = \frac{1}{p} = 6$.

Exercices

Exercice 5 du TD14.

Propriété

Linéarité de l'espérance

Si X et Y sont deux VARD admettant une espérance sur un même espace probabilisé, alors les variables aléatoires $X + Y$ et λX (pour tout $\lambda \in \mathbb{R}$) admettent également une espérance, donnée par

$$\begin{aligned} E(X + Y) &= E(X) + E(Y), \\ E(\lambda X) &= \lambda E(X) \quad \text{pour tout } \lambda \in \mathbb{R}. \end{aligned}$$

En particulier, si X admet une espérance alors pour tous réels a et b , la VARD $aX + b$ admet l'espérance

$$E(aX + b) = aE(X) + b.$$

Exemple

On lance n dés équilibrés.

Pour tout $i \in [1; n]$, on note X_i le numéro du $i^{\text{ème}}$ dé, on a déjà calculé que $E(X_i) = \frac{7}{2}$.

Notons ensuite S la variable aléatoire affichant la somme des numéros des dés, autrement dit

$$S = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i.$$

Par linéarité de l'espérance, on a donc

$$E(S) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n \frac{7}{2} = \frac{7n}{2}.$$

Remarque : on verra plus tard (voir Section 4.3.3) que plus n est un entier élevé, plus la loi de S ressemble à la loi normale $\mathcal{N}\left(\frac{7n}{2}, \frac{35n}{12}\right)$.

Propriété

Positivité

Si X est une VARD à valeurs positives (c'est-à-dire $X(\Omega) \subset [0, +\infty[)$, alors son espérance, si elle existe, est aussi positive.

Exercices

Exercice 6 du TD14.

Définition

Variables centrées

Soit X une VARD admettant une espérance. La *variable centrée* de X est la variable aléatoire $Y = X - E(X)$

Propriété

Si X est une variable centrée, alors elle admet une espérance nulle.

Propriété

Théorème de transfert

Si X est une VARD admettant une espérance et g une fonction à valeurs réelles définies sur $X(\Omega)$, notons $Y = g(X)$.

- Si $X(\Omega)$ est un ensemble fini, disons $X(\Omega) = \{x_1; x_2; \dots; x_n\}$, **alors Y admet l'espérance**

$$E(Y) = \sum_{i=1}^n g(x_i)P(X = x_i).$$

- Si $X(\Omega)$ est infini, disons $X(\Omega) = \{x_i : i \geq 1\} = \{x_1; x_2; \dots\}$, **alors Y admet une espérance si et seulement si la série $\sum_{i \geq 1} g(x_i)P(X = x_i)$ converge absolument, et dans ce cas**

$$E(Y) = \sum_{i=1}^{+\infty} g(x_i)P(X = x_i).$$

Exercice

On considère une VARD X de loi donnée par $X(\Omega) = \{n \in \mathbb{N} : n \geq 1\} = \{1; 2; \dots\}$ et $P(X = n) = \left(\frac{5}{6}\right)^{n-1} \times \frac{1}{6}$ pour tout $n \geq 1$.
On pose $Y = \frac{1}{2^X}$.
Montrer que Y admet une espérance et calculer $E(Y)$.

2.3.2 Moments, variance et écart-type d'une VARD

Définition

Moments d'ordre r d'une VARD

Pour tout entier $r \geq 1$, on dit qu'une VARD X admet un moment d'ordre r si la variable X^r admet une espérance, et l'on définit alors le moment d'ordre r de X comme le réel

$$m_r(X) = E(X^r).$$

Remarque

Les notions d'espérance et de moment d'ordre 1 sont confondues.

Exercice

Soit X = numéro d'un dé équilibré lancé au hasard.
Calculer le moment d'ordre 2 de X .

Exercice

Même exercice pour X = nombre de lancers nécessaires pour obtenir un 6 avec un dé équilibré.

Définition

Variance et écart-type d'une VARD

Soit X une VARD. On dit que X admet une variance si la variable centrée $X - E(X)$ admet un moment d'ordre 2, c'est-à-dire si le réel

$$E((X - E(X))^2)$$

existe, auquel cas on nomme *variance* de X ce réel $V(X) = E((X - E(X))^2)$.

On définit alors l'écart-type de X comme le réel

$$\sigma(X) = \sqrt{V(X)}.$$

Remarque

La variance et l'écart-type d'une variable aléatoire X sont des quantités positives qui traduisent la "dispersion" de X , c'est-à-dire sa susceptibilité à prendre des valeurs éloignées de son espérance (sa valeur moyenne), qui est d'autant plus élevée que sa variance est élevée.

Une variable aléatoire ayant une variance nulle ne prend ainsi qu'une seule valeur (son espérance, et n'a rien donc finalement rien d'aléatoire littéralement parlant), et réciproquement une variable aléatoire constante a bien sûr une variance nulle.

En pratique

Pour calculer la variance d'une variable aléatoire, on utilisera toujours le résultat suivant.

Propriété

Formule de Kœnig-Huygens

Si une variable aléatoire X admet un moment d'ordre 2, alors elle admet une variance qui vaut

$$V(X) = E(X^2) - (E(X))^2.$$

Exercice

Soit X = numéro d'un dé équilibré lancé au hasard.
Calculer la variance et l'écart-type de X .

Exercice

Même exercice pour X = nombre de lancers nécessaires pour obtenir un 6 avec un dé équilibré.

Propriété

Soit X une VARD admettant une variance. Alors, pour tous réels a et b , la variable aléatoire $aX + b$ admet la variance

$$V(aX + b) = a^2 V(X),$$

et donc l'écart-type

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

Exercices

Exercices 9 et 10 du TD14.

Définition

Variable aléatoire réduite

Une variable aléatoire X dotée d'une variance est dite *réduite* si $\sigma(X) = 1$.

Définition/Proposition

Soit X une variable aléatoire dotée d'une espérance $E(X) = m$ et d'un écart-type $\sigma(X) = \sigma$.

Alors : la variable aléatoire $X^* = \frac{X - m}{\sigma}$, nommée la *variable centrée réduite de X* , est une variable aléatoire centrée réduite.

3 Lois usuelles

3.1 Lois discrètes finies

3.1.1 Loi certaine

Définition

On dit qu'une variable aléatoire X suit la loi certaine si X prend une seule valeur avec une probabilité 1, autrement dit si X est une fonction constante.

Remarque

Ainsi une variable aléatoire X suit la loi certaine si et seulement si elle admet la variance $V(X) = 0$.

3.1.2 Loi uniforme

Définition

Soient a et b deux entiers vérifiant $a \leq b$.

On dit qu'une variable aléatoire X suit **la loi uniforme sur $[[a; b]]$** , ce que l'on note $X \hookrightarrow \mathcal{U}([a; b])$, si

$$X(\Omega) = [[a; b]] = \{a; a+1; \dots; b-1; b\},$$
$$P(X = k) = \frac{1}{b - a + 1} \quad \text{pour tout } k \in [[a; b]].$$

Cas particulier

$X \hookrightarrow \mathcal{U}([1; n])$ si X prend pour valeurs les entiers de 1 à n , chacun avec une probabilité $\frac{1}{n}$.

Exemple

On lance un dé **équilibré**, X = numéro du dé.
Alors : $X \hookrightarrow \mathcal{U}([1; 6])$.

Exemple

Une urne contient n boules numérotées de 1 à n . On tire une boule.
 X = numéro de la boule tirée.
Alors : $X \hookrightarrow \mathcal{U}([1; n])$.

Propriété

Si $X \hookrightarrow \mathcal{U}([1; n])$, alors

$$E(X) = \frac{n+1}{2},$$
$$V(X) = \frac{n^2-1}{12}.$$

Corollaire

Plus généralement, si $X \hookrightarrow \mathcal{U}([a; b])$, alors

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

Démonstration

Posons $Y = X - a + 1$ et $n = b - a + 1$.

On a $Y(\Omega) = \{a - a + 1; a + 1 - a + 1; \dots; b - a + 1\} = \{1; 2; \dots; b - a + 1\} = [1; n]$ et, pour tout $k \in [1; n]$,

$$P(Y = k) = P(X - a + 1 = k) = P(X = k + a - 1) = \frac{1}{b - a + 1} = \frac{1}{n},$$

autrement dit $Y \hookrightarrow \mathcal{U}([1; n])$ donc $E(Y) = \frac{n+1}{2}$ et $V(Y) = \frac{n^2-1}{12}$.

Ainsi, par linéarité de l'espérance,

$$E(X) = E(Y + a - 1) = E(Y) + E(a - 1) = \frac{n+1}{2} + a - 1 = \frac{b-a+2}{2} + a - 1 = \frac{a+b}{2}.$$

Ensuite, on sait que $V(\alpha Z + \beta) = \alpha^2 V(Z)$ pour toute variable aléatoire Z (admettant une espérance), donc

$$V(X) = V(1 \times Y + (a - 1)) = 1^2 V(Y) = V(Y) = \frac{n^2-1}{12} = \frac{(b-a+1)^2-1}{12}.$$

Scilab

On peut simuler une variable aléatoire suivant la loi uniforme sur $[a; b]$ via la commande
`grand(1, 1, "uin", a, b)`

Exercice(s)

Exercice 11 du TD14.

3.1.3 Loi de Bernoulli

Définition

Soit X une VARD et $p \in]0, 1[$. On dit que X suit la loi de Bernoulli de paramètre p , ce que l'on note $X \hookrightarrow \mathcal{B}(p)$, si la loi de X est donnée par $X(\Omega) = \{0; 1\}$ et $P(X = 1) = p$ (et donc $P(X = 0) = 1 - p$).

Archétype : épreuve de Bernoulli

On appelle épreuve de Bernoulli de probabilité de succès $p \in]0; 1[$ toute expérience aléatoire ayant deux résultats possibles : le succès, obtenu avec la probabilité p , et l'échec, obtenu avec une probabilité $1 - p$ (par exemple, jouer à pile ou face, obtenir ou non le numéro 6 en lançant un dé, etc).
Soit X la variable aléatoire valant 1 si cette expérience réussit et 0 sinon, par définition $X \hookrightarrow \mathcal{B}(p)$.

Exemple

On joue à pile ou face une fois avec une pièce ayant une probabilité $p \in]0, 1[$ de tomber sur pile, X est la variable aléatoire valant 1 si on obtient pile, 0 si on obtient face.
Alors : $X \hookrightarrow \mathcal{B}(p)$.

Exemple

On lance un dé équilibré, X est la variable aléatoire valant 1 si on obtient 6, 0 si on obtient un autre numéro.
Alors : $X \hookrightarrow \mathcal{B}\left(\frac{1}{6}\right)$.

Exemple

Une urne contient a boules bleues et b boules rouges, X est la variable aléatoire valant 1 si on tire une boule bleue, 0 si on tire une boule rouge.
Alors : $X \hookrightarrow \mathcal{B}\left(\frac{a}{a+b}\right)$.

Propriété

Si $X \hookrightarrow \mathcal{B}(p)$, alors

$$E(X) = p,$$
$$V(X) = p - p^2 = p(1 - p).$$

Scilab

On peut implanter une variable aléatoire suivant la loi de Bernoulli de paramètre $p \in]0, 1[$ de la façon suivante :

```
function X = Bernoulli(p)
    a = rand()
    X = 1*(a < p)
endfunction
```

3.1.4 Loi binomiale

Définition

On dit qu'une variable aléatoire X **suit la loi binomiale de taille $n \geq 0$ et de paramètre $p \in]0, 1[$** , ce que l'on note

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

si la loi de X est donnée par

$$X(\Omega) = \llbracket 0; n \rrbracket,$$
$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{pour tout } k \in \llbracket 0; n \rrbracket.$$

Remarque

Cette définition est cohérente avec la propriété $\sum_{x \in X(\Omega)} P(X = x) = 1$ vérifiée par toute VARD X : en effet, par l'identité du binôme de Newton,

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

Remarque

Si $n = 0$, X suit une loi certaine.
Si $n = 1$, X suit la loi de Bernoulli de paramètre p , autrement dit $\mathcal{B}(1, p) = \mathcal{B}(p)$.

Propriété/archétype : succession d'épreuves de Bernoulli

On répète n fois une épreuve de Bernoulli de probabilité de succès $p \in]0; 1[$ (par exemple, on joue à pile ou face n fois avec une pièce ayant une probabilité p de tomber sur pile).
Soit X le nombre de succès de cette expérience (dans l'exemple, X = nombre de piles obtenus).
Alors : $X \hookrightarrow \mathcal{B}(n, p)$.

Démonstration

On illustre cette preuve avec l'exemple donné dans la propriété.
Il est clair que $X(\Omega) = \llbracket 0; n \rrbracket$ (après n lancers, on obtient entre 0 et n piles).
Calculons $P(X = k)$ pour tout $k \in \llbracket 0; n \rrbracket$. Soit $P_i =$ "on obtient pile au $i^{\text{ème}}$ lancer". On a

$$(X = k) = \bigcup_{\substack{S \subseteq \llbracket 1; n \rrbracket, \\ |S| = k}} \left(\bigcap_{i \in S} P_i \cap \bigcap_{i \notin S} \overline{P_i} \right).$$

Les évènements regroupés par le signe union étant clairement incompatibles deux à deux, on obtient

$$\begin{aligned} P(X = k) &= \sum_{\substack{S \subseteq \llbracket 1; n \rrbracket, \\ |S| = k}} P \left(\underbrace{\bigcap_{i \in S} P_i \cap \bigcap_{i \notin S} \overline{P_i}}_{p^k (1-p)^{n-k}} \right) = \sum_{\substack{S \subseteq \llbracket 1; n \rrbracket, \\ |S| = k}} p^k (1-p)^{n-k} \\ &= p^k (1-p)^{n-k} \times \underbrace{\left(\sum_{\substack{S \subseteq \llbracket 1; n \rrbracket, \\ |S| = k}} 1 \right)}_{\binom{n}{k}} \\ &= \binom{n}{k} p^k (1-p)^{n-k}. \end{aligned}$$

Rappelons en effet que $\binom{n}{k}$ est le nombre de sous-ensembles S de $\llbracket 1; n \rrbracket$ possédant exactement k éléments.

Autres exemples

- X = nombre de 6 obtenus en lançant simultanément n dés équilibrés (ce qui revient à faire n lancers indépendants d'un même dé).
Alors : $X \hookrightarrow \mathcal{B}\left(n, \frac{1}{6}\right)$.
- Une urne contient a boules bleues et b boules rouges. On tire n boules avec remise. X = nombre de boules rouges tirées.
Alors : $X \hookrightarrow \mathcal{B}\left(n, \frac{a}{a+b}\right)$.

Remarque

Si $X \hookrightarrow \mathcal{B}(n, p)$, alors X est une somme de variables aléatoires suivant la loi de Bernoulli : si par exemple X = nombre de piles obtenus en jouant à pile ou face n fois avec une pièce ayant une probabilité p de tomber sur pile, notons $N_i = 1$ si on obtient pile au $i^{\text{ème}}$ lancer, 0 sinon (ainsi $N_i \hookrightarrow \mathcal{B}(p)$) pour tout $i \in \llbracket 1; n \rrbracket$.
Alors : $X = N_1 + N_2 + \dots + N_n$.

Propriété

Si $X \hookrightarrow \mathcal{B}(n, p)$, alors

$$\begin{aligned} E(X) &= np, \\ V(X) &= np(1-p). \end{aligned}$$

Démonstration

Il y a deux façons de démontrer ces résultats. Pour simplifier, on suppose que X = nombre de piles obtenus en jouant à pile ou face n fois avec une pièce ayant une probabilité p de tomber sur pile.

La première façon est de considérer X comme la somme $N_1 + N_2 + \dots + N_n$ où, pour tout $i \in \llbracket 1; n \rrbracket$, $N_i = 1$ si on obtient pile au $i^{\text{ème}}$ lancer, 0 sinon. On sait que N_i suit la loi de Bernoulli de paramètre p pour tout i , ainsi, par linéarité de l'espérance,

$$E(X) = E(N_1 + N_2 + \dots + N_n) = E(N_1) + E(N_2) + \dots + E(N_n) = p + p + \dots + p = np.$$

En deuxième année, on voit de plus que la variance est linéaire **pour les variables aléatoires indépendantes** (notion définie à ce moment-là), ce que sont les variables $N_1, N_2, \dots, N_n$, ainsi

$$V(X) = V(N_1) + V(N_2) + \dots + V(N_n) = p(1-p) + p(1-p) + \dots + p(1-p) = np(1-p).$$



Démonstration (suite)

L'autre façon est purement calculatoire. Revenons à la définition de l'espérance. On a

$$\begin{aligned} E(X) &= \sum_{k=0}^n kP(X=k) = 0P(X=0) + \sum_{k=1}^n \underbrace{k \binom{n}{k}}_{n \binom{n-1}{k-1}} p^k (1-p)^{n-k} = np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} \\ &= np \sum_{l=0}^{n-1} \binom{n-1}{l} p^l (1-p)^{n-1-l} \\ &= np(p + (1-p))^{n-1} \quad \text{par l'identité du binôme de Newton,} \\ &= np. \end{aligned}$$

Ensuite,

$$\begin{aligned} E(X^2) &= \sum_{k=0}^n k^2 P(X=k) = 0^2 P(X=0) + \sum_{k=1}^n \underbrace{k^2 \binom{n}{k}}_{nk \binom{n-1}{k-1}} p^k (1-p)^{n-k} = np \sum_{k=1}^n k \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} \\ &= np \sum_{l=0}^{n-1} (l+1) \binom{n-1}{l} p^l (1-p)^{n-1-l} = np \left(\underbrace{\sum_{l=0}^{n-1} l \binom{n-1}{l} p^l (1-p)^{n-1-l}}_{(n-1)p \text{ d'après ce qui précède}} + \underbrace{\sum_{l=0}^{n-1} \binom{n-1}{l} p^l (1-p)^{n-1-l}}_{(p+(1-p))^{n-1}} \right) \\ &= np((n-1)p + 1) = np(np - p + 1) = (np)^2 + np(1-p), \end{aligned}$$

donc, par la formule de Koenig-Huygens,

$$V(X) = E(X^2) - (E(X))^2 = (np)^2 + np(1-p) - (np)^2 = np(1-p).$$

Scilab

On peut simuler une variable aléatoire $X \hookrightarrow \mathcal{B}(n, p)$ comme suit :

grand(1,1,"bin",n,p)

Exercices

Exercices 12,13,14 du TD14.

3.2 Lois discrètes infinies

3.2.1 Loi géométrique

Définition

On dit qu'une variable aléatoire X suit la loi géométrique de paramètre $p \in]0; 1[$, ce que l'on note $X \hookrightarrow \mathcal{G}(p)$, si sa loi est donnée par

$$X(\Omega) = \mathbb{N} \setminus 0 = \{1; 2; \dots\},$$
$$P(X = n) = (1 - p)^{n-1} p \quad \text{pour tout } n \geq 1.$$

Archétype

C'est la loi de toute variable aléatoire de la forme $X =$ premier succès d'une épreuve de Bernoulli répétée soit à l'infini, soit tant que l'épreuve échoue.

Exemple

$X =$ nombre de lancers nécessaires pour obtenir pile en lançant une pièce ayant une probabilité $p \in]0; 1[$ de tomber sur pile.
Alors : $X \hookrightarrow \mathcal{G}(p)$.

Exemple

$X =$ nombre de lancers nécessaires pour obtenir 6 en lançant un dé équilibré.
Alors : $X \hookrightarrow \mathcal{G}(\frac{1}{6})$.

Propriété

Si $X \hookrightarrow \mathcal{G}(p)$, alors X admet l'espérance et la variance suivantes :

$$E(X) = \frac{1}{p},$$
$$V(X) = \frac{1 - p}{p^2}.$$

Scilab

On peut simuler une variable aléatoire $X \hookrightarrow \mathcal{G}(p)$ via la commande

grand(1, 1, "geom", p)

Remarque

Si $X \hookrightarrow \mathcal{G}(p)$, alors $P(X > k)$ est la probabilité que les k premières épreuves de Bernoulli de paramètre p se soldent par un échec, c'est-à-dire

$$P(X > k) = (1 - p)^k.$$

Ainsi, pour tous entiers $k \geq 1$ et $l \geq 1$,

$$\begin{aligned} P_{(X > k)}(X = k + l) &= \frac{P((X > k) \cap (X = k + l))}{P(X > k)} = \frac{P(X = k + l)}{P(X > k)} \\ &= \frac{(1 - p)^{k+l-1} p}{(1 - p)^k} = (1 - p)^{l-1} p \\ &= P(X = l). \end{aligned}$$

Cette propriété (l'égalité $P_{(X > k)}(X = k + l) = P(X = l)$) est appelée **absence de mémoire** de la variable aléatoire X : si k épreuves de Bernoulli successives échouent (c'est l'évènement $(X > k)$), la probabilité que le premier succès survienne au bout de l autres épreuves est la même que si les k premières épreuves n'avaient pas eu lieu.

Exercices

Exercice 15 du TD14.

3.2.2 Loi de Poisson

Définition

On dit qu'une variable aléatoire X suit la loi de Poisson de paramètre $\lambda > 0$ si sa loi est donnée par

$$X(\Omega) = \mathbb{N} = \{0; 1; 2; \dots\},$$
$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{pour tout } k \geq 0.$$

Remarque

Cette définition est cohérente avec l'égalité $\sum_{x \in X(\Omega)} P(X = x) = 1$ vérifiée par toute VARD : en effet, on sait que la série exponentielle $\sum_{k \geq 0} \frac{x^k}{k!}$ converge vers e^x pour tout $x \in \mathbb{R}$, ainsi

$$\sum_{k=0}^{+\infty} P(X = k) = \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \times \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \times e^{\lambda} = 1.$$

Propriété

Soit λ un réel strictement positif.

Pour tout entier $n \geq 1$, on considère une variable aléatoire $X_n \hookrightarrow \mathcal{B}(n, \frac{\lambda}{n})$.

Alors : la suite de variables aléatoires $(X_n)_{n \geq 1}$ converge "en loi" (plus de détails en deuxième année) vers une variable aléatoire $X \hookrightarrow \mathcal{P}(\lambda)$, c'est-à-dire : pour tout entier $k \geq 0$, on a

$$\lim_{n \rightarrow +\infty} \underbrace{\binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}}_{P(X_n = k)} = \underbrace{\frac{\lambda^k}{k!} e^{-\lambda}}_{P(X = k)}.$$

Exemple

Un pêcheur pêche en moyenne $\lambda > 0$ poisson(s) en une heure. On cherche à modéliser l'expérience consistant à pêcher pendant une heure, pour déterminer la loi de probabilité de la variable aléatoire donnant le nombre de poissons pêchés.

Pour tout $n \geq 1$, on peut subdiviser l'heure en n intervalles de temps égaux, et considérer que le pêcheur peut pêcher un poisson sur chaque intervalle de temps avec une probabilité constante p_n . Ainsi, si X_n = nombre de poissons pêchés et une heure, alors $X_n \hookrightarrow \mathcal{B}(n, p_n)$ (d'espérance np_n), et puisque le pêcheur pêche en moyenne λ poisson(s) (autrement dit on veut avoir $E(X_n) = \lambda$), alors $p_n = \frac{\lambda}{n}$ (et cette modélisation n'a de sens que si $n > \lambda$).

Le modèle ci-dessus supposant que le pêcheur ne peut pêcher qu'un seul poisson sur chacun de ces intervalles de temps (et au total un maximum de n poissons), il est naturel de considérer la limite de ce modèle quand n tend vers l'infini, où cette limite disparaît. Par la propriété précédente, la variable aléatoire X = nombre de poissons pêchés en une heure suit alors la loi de Poisson de paramètre λ .

Propriété

Si $X \hookrightarrow \mathcal{P}(\lambda)$, alors X admet l'espérance et la variance suivante :

$$E(X) = \lambda,$$
$$V(X) = \lambda.$$

Scilab

On peut simuler une variable aléatoire $X \hookrightarrow \mathcal{P}(\lambda)$ via la commande

grand(1,1,"poi",lambda)

Exercices

Exercice 16 du TD14.

4 Variables à densité

4.1 Définition et propriétés

Définition

On dit qu'une variable aléatoire réelle X est à densité si sa fonction de répartition F_X vérifie les conditions suivantes :

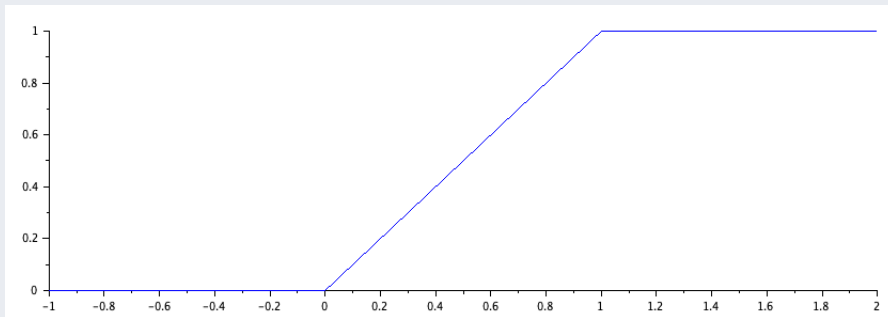
- F_X est continue sur $\mathbb{R}$;
- F_X est de classe C^1 sur $\mathbb{R}$, ou bien sur $\mathbb{R}$ privé d'un nombre **fini** de points.

Exemple

On pioche un réel X au hasard dans le segment $[0; 1]$, de manière uniforme (voir loi uniforme plus loin pour une définition générale), c'est-à-dire que $P(a \leq X \leq b) = b - a$ pour tous réels $0 \leq a \leq b \leq 1$.

On a ainsi

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ x & \text{si } x \in [0; 1], \\ 1 & \text{si } x > 1. \end{cases}$$



F_X est ainsi continue sur $\mathbb{R}$ et de classe C^1 sur $\mathbb{R} \setminus \{0; 1\}$ donc X est une variable à densité.

Définition

Soit X est une variable à densité. On considère la dérivée F'_X de F_X sur l'ensemble des points où elle est dérivable (c'est-à-dire tout $\mathbb{R}$ sauf éventuellement un nombre fini de points).

On nomme *densité* de X toute fonction f_X définie sur $\mathbb{R}$ vérifiant :

- f_X est à valeurs positives;
- $f_X = F'_X$ sauf éventuellement en un nombre **fini** de points.

Exemple

Reprenons l'exemple précédent.

Soit f_X la fonction définie sur $\mathbb{R}$ par

$$f_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 & \text{si } x \in [0; 1], \\ 0 & \text{si } x > 1. \end{cases}$$

Ainsi f_X est à valeurs positives et $F'_X = f_X$ sur $\mathbb{R} \setminus \{0; 1\}$ donc f_X est une densité de X .

Remarque

Si f_X est une densité de X , puisque $f_X = F'_X$ (sauf éventuellement en un nombre fini de points), et comme F_X est de classe C^1 sur $\mathbb{R}$ (sauf éventuellement en un nombre fini de points), alors f est continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points, en particulier f est continue par morceaux et l'intégrale $\int_a^b f(t)dt$ est définie pour tous réels a et b .

Propriété

Si f_X est une densité de X , alors

$$\int_{-\infty}^x f_X(t)dt = F_X(x)$$

pour tout $x \in \mathbb{R}$.

Remarque

Ainsi, si f_X est une densité de X , alors

$$\int_{-\infty}^{+\infty} f_X(t) dt = \lim_{x \rightarrow +\infty} F_X(x) = 1.$$

Corollaire

Si X est une variable à densité de densité f_X , alors f_X détermine totalement la loi de X .

Propriété

Si f est une fonction définie sur $\mathbb{R}$ vérifiant les propriétés suivantes :

- f est à valeurs positives ;
- f est continue sur $\mathbb{R}$ (sauf éventuellement en un nombre fini de points) ;
- $\int_{-\infty}^{+\infty} f(t) dt = 1$;

alors il existe une variable aléatoire X dont f est une densité.

Exercice

Soit f la fonction définie sur $\mathbb{R}$ par

$$f(x) = \begin{cases} 0 & \text{si } x < 1, \\ \frac{1}{x^2} & \text{si } x \geq 1. \end{cases}$$

Montrer que f est une densité d'une variable aléatoire X .

Propriété

Transformation affine d'une variable à densité

Soit X une variable aléatoire de densité f_X .

Alors : pour tous réels a et b , la variable aléatoire $Y = aX + b$ est également à densité.

Exercice

Soit X une variable aléatoire ayant pour densité la fonction f définie dans l'exercice précédent.
Déterminer une densité de la variable aléatoire $Y = -2X + 3$.

4.2 Espérance d'une variable à densité

Définition

On dit qu'une variable à densité X , de densité f_X , admet une espérance si l'intégrale $\int_{-\infty}^{+\infty} t f_X(t) dt$ **converge absolument**, auquel cas l'espérance de X est définie comme

$$E(X) = \int_{-\infty}^{+\infty} t f_X(t) dt.$$

Exercice

1 Montrer que si X est une variable à densité, de densité la fonction f des exercices précédents, alors X n'admet **pas** d'espérance.

2 Soit g la fonction définie sur $\mathbb{R}$ par

$$g(x) = \begin{cases} 0 & \text{si } x < 1, \\ \frac{2}{x^3} & \text{si } x \geq 1. \end{cases}$$

- Montrer que g est la densité d'une variable aléatoire Y (qu'on ne cherchera pas à déterminer).
- Montrer que Y admet une espérance et la calculer.

Définition/propriété

Variable centrée

Comme pour les variables aléatoires discrètes, on dit qu'une variable à densité X est *centrée* si elle admet l'espérance

$$E(X) = 0,$$

et si Y est une variable à densité admettant une espérance quelconque $E(Y)$, alors la variable aléatoire $Y - E(Y)$ **est une variable à densité centrée**.

Exemple

Soit Y une variable de densité g comme dans l'exercice précédent, montrer que la variable à densité $Z = Y - 2$ est centrée.

4.3 Lois à densité usuelles

4.3.1 Loi uniforme

Définition

On dit qu'une variable à densité X *suit la loi uniforme sur le segment* $[a, b]$ si $X(\Omega) = [a; b]$ et si une densité de X est donnée par la fonction f_X définie sur $\mathbb{R}$ par

$$f_X(x) = \begin{cases} 0 & \text{si } x < a, \\ \frac{1}{b-a} & \text{si } a \leq x \leq b, \\ 0 & \text{si } x > b, \end{cases}$$

autrement dit si la fonction de répartition de X est la fonction F_X donnée par

$$F_X(x) = \begin{cases} 0 & \text{si } x < a, \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b, \\ 1 & \text{si } x > b. \end{cases}$$

On note alors

$$X \hookrightarrow \mathcal{U}([a; b]).$$

Archétype

C'est la loi de la variable X = réel pioché aléatoirement et uniformément dans le segment $[a; b]$.

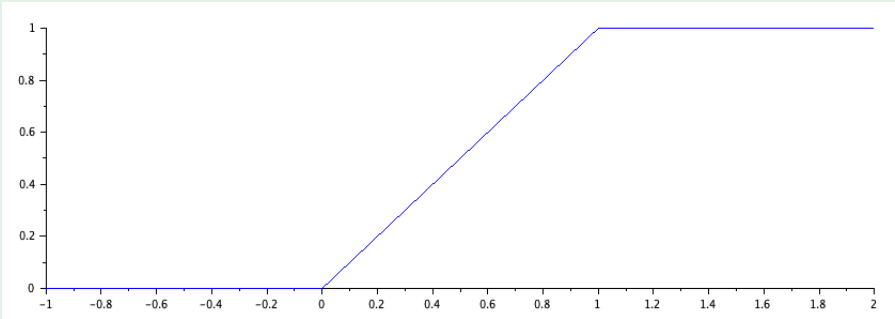
Remarque

Si $X \hookrightarrow \mathcal{U}([a; b])$, alors

$$\begin{cases} P(X \leq a) = F_X(a) = 0, \\ P(X \in [c; d]) = F_X(d) - F_X(c) = \frac{d-c}{b-a} \quad \text{pour tous réels } a \leq c \leq d \leq b, \\ P(X > b) = 1 - F_X(b) = 0. \end{cases}$$

Cas particulier

$X \hookrightarrow \mathcal{U}([0; 1])$: exemple précédemment traité, F_X a pour graphe



et

$$\begin{cases} P(X \leq 0) = 0, \\ P(X \in [c; d]) = d - c \quad \text{pour tous réels } 0 \leq c \leq d \leq 1, \\ P(X > 1) = 0. \end{cases}$$

Remarque

Soit $Y \hookrightarrow \mathcal{U}([a; b])$.

Pour tous réels A et $B > 0$,

$$A + BY \hookrightarrow \mathcal{U}([A + Ba; A + Bb]).$$

En particulier, si X est une variable aléatoire et $Y = a + (b - a)X \left(\Leftrightarrow X = \frac{Y - a}{b - a} \right)$, alors

$$X \hookrightarrow \mathcal{U}([0; 1]) \Leftrightarrow Y \hookrightarrow \mathcal{U}([a; b]).$$

Propriété

Si $X \hookrightarrow \mathcal{U}([a; b])$, alors comme les variables aléatoires uniformes discrètes, X admet l'espérance

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

Scilab

La commande

rand()

simule une variable aléatoire $X \hookrightarrow \mathcal{U}([0; 1])$.

Ensuite, soit $Y \hookrightarrow \mathcal{U}([a; b])$. Puisque la variable aléatoire $X = \frac{Y - a}{b - a}$ vérifie $X \hookrightarrow \mathcal{U}([0; 1])$ et $Y = a + (b - a)X$, alors une façon de simuler Y est la commande

a+(b-a)*rand()

4.3.2 Loi exponentielle

Définition

Soit $\lambda > 0$. On dit qu'une variable à densité X suit la loi exponentielle de paramètre λ , ce que l'on note

$$X \hookrightarrow \mathcal{E}(\lambda),$$

si $X(\Omega) = [0; +\infty[$ et si X a pour densité la fonction f_X définie sur $\mathbb{R}$ par

$$f_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ \lambda e^{-\lambda x} & \text{si } x \geq 0, \end{cases}$$

autrement dit si la fonction de répartition F_X de X est donnée par

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - e^{-\lambda x} & \text{si } x \geq 0, \end{cases}$$

Remarque

Si $X \hookrightarrow \mathcal{E}(\lambda)$, alors $P(X > x) = 1 - F_X(x) = e^{-\lambda x}$ pour tout $x \geq 0$, ainsi

$$\begin{aligned} P_{(X>x)}(X > x + y) &= \frac{P((X > x) \cap (X > x + y))}{P(X > x)} = \frac{P(X > x + y)}{P(X > x)} \\ &= \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}} = e^{-\lambda y} \\ &= P(X > y) \end{aligned}$$

pour tous réels positifs x et y . Cette propriété (l'égalité $P_{(X>x)}(X > x + y) = P(X > y)$) est appelée **absence de mémoire** de la variable aléatoire X .

La proposition suivante donne la réciproque de cette propriété.

Propriété

Si une variable à densité X vérifiant $X(\Omega) = [0; +\infty[$ est sans mémoire (c'est-à-dire $P_{(X>x)}(X > x + y) = P(X > y)$), autrement dit si $P(X > x + y) = P(X > x)P(X > y)$), alors il existe $\lambda > 0$ tel que

$$X \hookrightarrow \mathcal{E}(\lambda).$$

Démonstration

Notons $G = 1 - F_X$, donnée par $G(x) = 1 - F_X(x) = P(X > x)$ pour tout $x \in \mathbb{R}$.

Introduisons une densité f_X de X . Par définition d'une variable à densité, on sait que f_X est continue sur $\mathbb{R}$ éventuellement privé d'un nombre fini de points, et même en ces points f_X admet une limite à gauche et à droite par continuité de f_X ailleurs qu'en ces points (sur un voisinage ouvert assez petit).

Autrement dit, G définie par $G(x) = \int_x^{+\infty} f_X(t) dt$ est continue sur $\mathbb{R}$, et de classe C^1 sur $\mathbb{R}$ éventuellement privé d'un nombre fini de points (mais même en ces points les dérivées à gauche et à droite existent, car ce sont les limites à gauche et à droite de $-f_X$, en particulier $G'(0^+) = \lim_{h \rightarrow 0^+} \frac{G(h) - G(0)}{h - 0}$ existe).

Puisque $X(\Omega) = [0; +\infty[$, on a $G(x) = 1 - P(X < x) = 1$ pour tout $x < 0$ (et donc $G(0) = 1$ également par continuité de G).

On sait aussi que $G(x) = P(X > x)$ est positif pour tout $x \in \mathbb{R}$ et décroît de $G(0) = 1$ vers 0 quand x tend vers $+\infty$, par propriété d'une fonction de répartition.

Soit $x \geq 0$ tel que G soit dérivable en x . On a

$$\begin{aligned} G'(x) &= \lim_{h \rightarrow 0^+} \frac{G(x+h) - G(x)}{h} = \lim_{h \rightarrow 0^+} \frac{G(x)G(h) - G(x)}{h} = G(x) \times \lim_{h \rightarrow 0^+} \frac{G(h) - 1}{h} \\ &= G(x) \times \lim_{h \rightarrow 0^+} \frac{G(h) - G(0)}{h - 0} \\ &= G(x) \times G'(0^+). \end{aligned}$$

Posons $\lambda = -G'(0^+)$. Puisque G est décroissante sur $\mathbb{R}$, on a donc $\lambda \geq 0$. Pour tout $x \geq 0$ tel que $G(x) \neq 0$ (autrement dit $G(t) > 0$ pour tout $t \in [0; x]$, puisque G décroît de $G(0) = 1$ à $\lim_{x \rightarrow +\infty} G(x) = 0$), on a donc

$$\begin{aligned} \int_0^x \frac{G'(t)}{G(t)} dt &= [\ln(G(t))]_0^x = \ln(G(x)) - \ln(G(0)) = \ln(G(x)) - \ln(1) = \ln(G(x)) \\ &= \int_0^x (-\lambda) dt = [-\lambda t]_0^x = -\lambda x, \end{aligned}$$

ainsi

$$G(x) = e^{\ln(G(x))} = e^{-\lambda x}.$$

Puisque G est continue, elle ne peut donc s'annuler sur $[0; +\infty[$ (l'exponentielle est toujours strictement positive), et l'on a toujours $G(x) = e^{-\lambda x}$ en tout point x où G n'est pas dérivable (ce qui implique finalement que ces points n'existent pas puisque la fonction $e^{-\lambda x}$ est dérivable sur $\mathbb{R}$ tout entier).

Finalement, comme $\lim_{x \rightarrow +\infty} e^{-\lambda x} = \lim_{x \rightarrow +\infty} G(x) = 0$, on a forcément $\lambda > 0$ (sinon $\lambda = 0$ et la fonction G est constante valant 1 sur tout $\mathbb{R}$, ce qui est absurde).

En conclusion, G est donc définie par

$$G(x) = \begin{cases} 1 & \text{si } x < 0, \\ e^{-\lambda x} & \text{si } x \geq 0, \end{cases}$$

donc la fonction de répartition $F_X = 1 - G$ est définie par

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - e^{-\lambda x} & \text{si } x \geq 0, \end{cases}$$

autrement dit $X \hookrightarrow \mathcal{E}(\lambda)$.

Propriété

Si $X \hookrightarrow \mathcal{E}(\lambda)$, alors X admet l'espérance

$$E(X) = \frac{1}{\lambda}.$$

Exemple

Soit X le temps séparant notre dernier appel téléphone reçu du prochain (ainsi $X(\Omega) = [0; +\infty[$).

On suppose que :

- X admette une espérance de 12 heures (en moyenne, il faut attendre 12 heures avant de recevoir un autre appel),
- X est une variable aléatoire sans mémoire (les chances de recevoir un appel ne changent pas avec le temps).

Alors $X \hookrightarrow \mathcal{E}(\lambda)$ où $\lambda = \frac{1}{E(X)} = \frac{1}{12}$.

Calculer la probabilité de recevoir un appel d'ici une heure.

Propriété

Lien entre lois uniforme et exponentielle

Soit Y une variable aléatoire et $\lambda > 0$. On considère la variable aléatoire $X = 1 - e^{-\lambda Y} \left(\Leftrightarrow Y = -\frac{\ln(1-X)}{\lambda} \right)$.

Alors :

$$Y \hookrightarrow \mathcal{E}(\lambda) \Leftrightarrow X \hookrightarrow \mathcal{U}([0; 1]).$$

Scilab

Pour simuler une variable aléatoire $Y \hookrightarrow \mathcal{E}(\lambda)$, on peut utiliser la propriété précédente : puisque $Y = -\frac{\ln(1-X)}{\lambda}$ où $X \hookrightarrow \mathcal{U}([0; 1])$ est simulé par la commande **rand()**, alors on peut simuler Y via

`Y = -log(1-rand())/lambda`

Sinon, on peut utiliser **grand** via

`Y = grand(1,1,'exp',1/lambda)`

4.3.3 Loi normale

Propriété

$$\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi}.$$

Définition

Loi normale centrée réduite

On dit qu'une variable aléatoire suit la *loi normale centrée réduite*, ce que l'on note

$$X \hookrightarrow \mathcal{N}(0, 1),$$

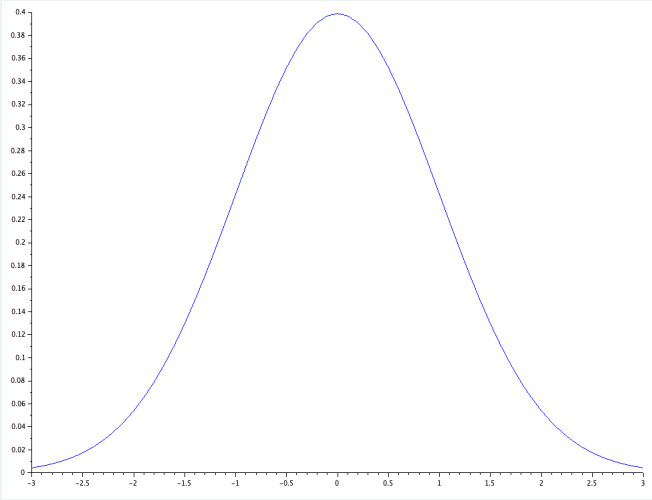
si la fonction de répartition F_X de X est donnée par

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt,$$

autrement dit si X est la variable de densité f_X donnée par

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

dont le graphe est tracé ci-dessous.



Propriété

Si $X \hookrightarrow \mathcal{N}(0, 1)$, alors X admet l'espérance et la variance suivantes :

$$E(X) = 0,$$

$$V(X) = 1,$$

autrement dit X est une variable centrée (son espérance est nulle) et réduite (sa variance vaut 1), d'où le nom de la loi $\mathcal{N}(0, 1)$.

Définition

Loi normale (définition générale)

Soient $\mu \in \mathbb{R}$ et $\sigma > 0$.

On dit qu'une variable aléatoire suit *la loi normale d'espérance m et de variance σ^2* , ce que l'on note

$$X \hookrightarrow \mathcal{N}(m, \sigma^2),$$

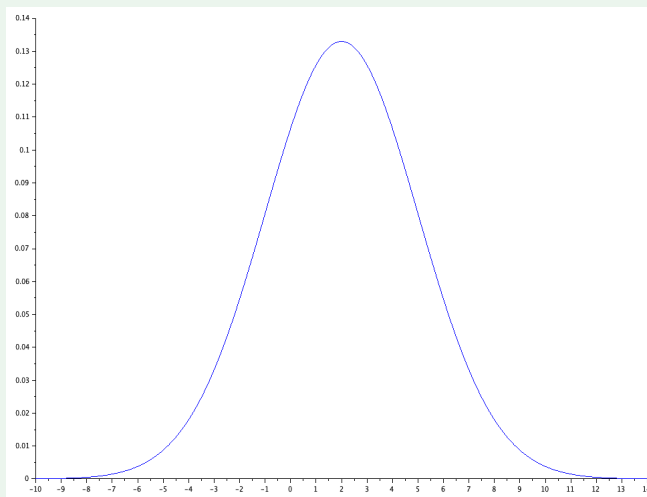
si la fonction de répartition F_X de X est donnée par

$$F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt,$$

autrement dit si X est la variable de densité f_X donnée par

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

dont on donne un exemple de graphe ci-dessous (avec $\mu = 2$ et $\sigma = 3$) :



Propriété

Soit X une variable aléatoire admettant l'espérance $E(X) = \mu$ et la variance $V(X) = \sigma^2$.

Notons $X^* = \frac{X - \mu}{\sigma}$ la variable centrée réduite de X .

Alors :

$$X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \Leftrightarrow X^* \hookrightarrow \mathcal{N}(0, 1).$$

Démonstration

Pour tout $x \in \mathbb{R}$, posons $y = \frac{x - \mu}{\sigma}$, ainsi $x = \sigma y + \mu$. On a

$$F_X(x) = P(X \leq x) = P(X \leq \sigma y + \mu) = P\left(\frac{X - \mu}{\sigma} \leq y\right) = P(X^* \leq y) = F_{X^*}(y)$$

donc $X^* \hookrightarrow \mathcal{N}(0, 1)$ si et seulement si

$$F_X(x) = F_{X^*}(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{t^2}{2}} dt = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(u-\mu)^2}{2\sigma^2}} du$$

(en faisant le changement de variable $u = \sigma t + \mu$, ainsi $t = \frac{u - \mu}{\sigma}$),
c'est-à-dire si et seulement si $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$.

Corollaire

Si $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$, alors X admet l'espérance et la variance suivantes :

$$\begin{aligned} E(X) &= \mu, \\ V(X) &= \sigma^2 \text{ (donc } \sigma(X) = \sigma). \end{aligned}$$

Archétype

Soit X n'importe quelle variable aléatoire (discrète ou non) d'espérance $E(X) = \mu$ et variance $V(X) = \sigma^2$. D'après le théorème central limite (voir deuxième année), $\mathcal{N}(\mu, \sigma^2)$ est la loi approximativement suivie par :

- la somme des valeurs prises par la variable aléatoire X en répétant son expérience un très grand nombre de fois de manière indépendante ;
- la moyenne arithmétique des valeurs prises par X toujours en répétant son expérience un très grand nombre de fois de manière indépendante.

Exemple

On lance n dés équilibrés. Soit X_i le numéro du $i^{\text{ème}}$ dé pour tout $i \in \llbracket 1; n \rrbracket$, ainsi $X_i \hookrightarrow \mathcal{U}(\llbracket 1; 6 \rrbracket)$ pour tout i (donc $E(X_i) = \frac{7}{2}$ et $V(X_i) = \frac{35}{12}$). Notons $S_n = X_1 + X_2 + \dots + X_n$, autrement dit S_n est la somme des numéros des dés. On sait que

$$E(S_n) = E(X_1) + E(X_2) + \dots + E(X_n) = \frac{7n}{2}.$$

Admettons que

$$V(S_n) = V(X_1) + V(X_2) + \dots + V(X_n) = \frac{35n}{12}.$$

Alors : plus n est grand, plus la loi de S_n ressemble à la loi normale $\mathcal{N}\left(\frac{7n}{2}, \frac{35n}{12}\right)$, c'est-à-dire

$$P(S_n \leq x) \simeq \frac{1}{\sqrt{\frac{35n}{12} \times 2\pi}} \int_{-\infty}^x e^{-\frac{(t - \frac{7n}{2})^2}{2 \times \frac{35n}{12}}} dt$$

pour tout $x \in \mathbb{R}$.

Exemple

Soit X la variable aléatoire donnant la note sur 20 d'un élève à un devoir quelconque. Lors des 5 derniers devoirs, X a pris les valeurs 7, 9, 12, 6 et 9.

La moyenne empirique (arithmétique) de ces valeurs est $\frac{7+9+12+6+9}{5} = \frac{43}{5} = 8.6$, et leur variance empirique est donc

$$\frac{(7-8.6)^2 + (9-8.6)^2 + (12-8.6)^2 + (9-8.6)^2}{5} = 4.24.$$

La loi des grands nombres (voir deuxième année) dit que $E(X) \simeq 8.6$ et $V(X) \simeq 4.24$.

L'élève passe n devoirs. Soit X_i sa note au $i^{\text{ème}}$ devoir pour tout $i \in \llbracket 1; n \rrbracket$ (X_i suit donc la même loi que X).

Notons enfin $\overline{X}_n$ la moyenne arithmétique de l'élève à ces n devoirs, c'est-à-dire

$$\overline{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

On a donc

$$E\left(\overline{X}_n\right) = \frac{E(X_1) + E(X_2) + \dots + E(X_n)}{n} \simeq \frac{8.6 + 8.6 + \dots + 8.6}{n} = 8.6.$$

Admettons de plus (voir deuxième année) que

$$V\left(\overline{X}_n\right) = \frac{V(X_1) + V(X_2) + \dots + V(X_n)}{n^2} \simeq \frac{4.24 + 4.24 + \dots + 4.24}{n^2} = \frac{4.24}{n}.$$

Alors : plus n est grand, plus la loi de $\overline{X}_n$ ressemble à la loi normale $\mathcal{N}\left(8.6, \frac{4.24}{n}\right)$, c'est-à-dire

$$P(\overline{X}_n \leq x) \simeq \frac{1}{\sqrt{\frac{4.24}{n}} \sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-8.6)^2}{2 \times \frac{4.24}{n}}} dt$$

pour tout $x \in \mathbb{R}$.

Remarque

On peut retenir que si $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$, alors

$$\begin{aligned} P(\mu - \sigma \leq X \leq \mu + \sigma) &\simeq 0.68, \\ P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) &\simeq 0.95, \\ P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) &\simeq 0.99, \end{aligned}$$

autrement dit **la loi normale se concentre essentiellement autour de son espérance**.

Par exemple, pour reprendre l'étude de S_n = somme des numéros de n dés équilibrés traitée plus haut, prenons $n = 12$, ainsi S_{12} peut valoir de 12 à 72, pour une valeur moyenne $E(S_{12}) = \frac{7 \times 12}{2} = 42$ et une variance $V(S_{12}) = \frac{35 \times 12}{12} = 35$ (donc $\sigma(S_{12}) = \sqrt{35} \simeq 5.9$).

On sait que S_{12} suit approximativement la loi normale $\mathcal{N}(42, 35)$, ainsi la probabilité que la somme de nos 12 dés soit comprise entre $42 - 2\sqrt{35} \simeq 30.16$ et $42 + 2\sqrt{35} \simeq 53.83$ (ce qui revient à dire que cette somme est comprise entre 31 et 53) est

$$P(31 \leq S_{12} \leq 53) \simeq 95\%.$$

Exemple

Légitimité des sondages

Une population contient $N = 30$ millions de votants pour un référendum à venir. Notons D le nombre de personnes qui voteront "oui", puis $p = \frac{D}{N} \in [0; 1]$.

On aimerait estimer la valeur de p avant le vote au moyen d'un sondage auprès d'un échantillon de 1000 personnes $V_1, V_2, \dots, V_{1000}$ de cette population.

Pour tout $i \in [1; 1000]$, notons $X_i = \begin{cases} 1 & \text{si } V_i \text{ déclare qu'elle votera "oui",} \\ 0 & \text{si } V_i \text{ déclare qu'elle votera "non".} \end{cases}$

Si l'échantillon de votants est représentatif de l'ensemble de la population, il est raisonnable de supposer que $X_i \hookrightarrow \mathcal{B}(p)$ pour tout $i \in [1; 1000]$, ainsi $E(X_i) = p$ et $V(X_i) = p(1 - p)$.

Notons maintenant $S = \sum_{i=1}^{1000} X_i$ = nombre de personnes déclarant qu'elles voteront "oui" au référendum parmi les 1000 personnes sondées, ainsi $S \hookrightarrow \mathcal{B}(1000, p)$ et donc

$$E(S) = 1000p, \quad V(S) = 1000p(1 - p).$$

Notons finalement $M = \frac{S}{1000}$ = pourcentage de personnes déclarant qu'elles voteront "oui" au référendum parmi les 1000 personnes sondées, ainsi

$$\begin{aligned} E(M) &= E\left(\frac{M}{1000}\right) = \frac{E(M)}{1000} = \frac{1000p}{1000} = p, \\ V(M) &= V\left(\frac{M}{1000}\right) = \frac{V(M)}{1000^2} = \frac{1000p(1 - p)}{1000^2} = \frac{p(1 - p)}{1000}. \end{aligned}$$

Or, d'après ce qui précède, on sait que M suit approximativement la loi normale $\mathcal{N}\left(p, \frac{p(1-p)}{1000}\right)$, donc d'après la remarque précédente,

$$P\left(p - 2\sqrt{\frac{p(1-p)}{1000}} \leq M \leq p + 2\sqrt{\frac{p(1-p)}{1000}}\right) \simeq 95\%,$$

c'est-à-dire que le pourcentage M que l'on obtient à l'issue du sondage est, avec une probabilité de 95%, une valeur approchée du pourcentage réel p que l'on obtiendra le jour du référendum, à $2\sqrt{\frac{p(1-p)}{1000}}$ près.

Puisque $p(1 - p) \leq \frac{1}{4}$ pour tout $p \in [0; 1]$ (on peut le montrer facilement par exemple en étudiant les variations de la fonction $f(x) = x(1 - x)$ sur $[0; 1]$), alors

$$2\sqrt{\frac{p(1-p)}{1000}} \leq 2\sqrt{\frac{1}{4000}} \simeq 0.0316 = 3.16\%,$$

donc il y a une probabilité supérieure à 95% que le pourcentage M obtenu auprès des 1000 votants soit une valeur approchée du résultat réel p à 3.16% près, ce qui fait du sondage un estimateur très fiable (pour peu que l'échantillon de population soit bien choisi).