Variables aléatoires et lois de probabilité

2017-2018

Introduction

Considérons une expérience aléatoire, et $(\Omega, \mathcal{A}, \mathbb{P})$ l'espace de probabilité associé.

Exemple : expérience = jet de 3 pièces non biaisés
$\Omega = \{P, F\}^3$, $\mathbb{P}(\omega) = \frac{1}{8}$, $\forall \omega \in \Omega$
On s'intéresse au nombre de "Pile" obtenu ; la question doit donc être posée différemment. On définira une variable aléatoire X qui à un événement élémentaire associe le nombre de "Pile" correspondant.

De façon générale, une variable aléatoire sera définie lorsque l'on s'intéresse à une fonction du résultat plutôt qu'au résultat lui-même.

$$
X : \begin{array}{ccc}
\Omega & \to & X(\Omega) \\
(F,F,F) & \to & 0 \\
(F,F,P) & \searrow & \\
(F,P,F) & \to & 1 \\
(P,F,F) & \nearrow & \\
(P,F,P) & \searrow & \\
(P,P,F) & \to & 2 \\
(F,P,P) & \nearrow & \\
(P,P,P) & \to & 3
\end{array}
$$

$X((F,F,F)) = 0,\ X^{-1}(0) = (F,F,F)$

Definition

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé associé à une e.a. *e*. On appelle variable aléatoire toute fonction mesurable de Ω dans un espace E mesurable : $\begin{array}{rcl} X : \Omega & \to & X(\Omega) \subset E \\ \omega & \mapsto & X(\omega) \end{array}$

- Une v.a. est dite discrète si $X(\Omega)$ est un ensemble fini ou infini dénombrable (i.e. $E = \mathbb{N}^n, \mathbb{Z}^n ou\, \mathbb{Q}^n$). Notons que si Ω est dénombrable, toute v.a.r. est discrête.
- Une v.a. est dite continue si $X(\Omega)$ est un ensemble infini non dénombrable, i.e. $\mathbb{R}^n$, un intervalle de $\mathbb{R}^n$, ou une union d'intervalles de $\mathbb{R}^n$, $\mathbb{C}^n$ ...

Exemples de variables aléatoires :

- variable discrête, souvent une variable de comptable, par exemple : nombre de piles dans notre e.a., nombre d'enfants dans une famille, nombre de bactéries dans une boîte de Petri, ...
- autres exemples de variable continue : masse corporelle des individus dans une espèce animale donnée, taux de glucose dans le sang, ...
- Dans une substance radioactive, la désintegration des noyaux se fait de façon spontanée. Le nombre de désintegration sur un intervalle de temps fixé suit une loi de Poisson (discrête). Par contre le temps d'attente entre deux désintégrations est modélisé par une loi exponentielle (continue).

- Si $n = 1$, on parle de variable aléatoire réelle.
- Si $n > 1$, on parle de vecteur aléatoire.
- Notation : X, Y, Z... variables aléatoires ; x, y, z une valeur de X, Y, Z respectivement.
- On est souvent conduit à calculer : $\mathbb{P}(\{\omega; X(\omega) \in S\})$ pour $S \subset X(\Omega)$. Nous allons définir cette mesure image de X. Remarque : en pratique, on connait souvent $(X(\Omega), \mathbb{P}_X)$ plutôt que $(\Omega, \mathbb{P})$.

Definition

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité. Soit $X : \Omega \to X(\Omega)$ une variable aléatoire. La loi de probabilité (ou mesure image) de X est

$$\begin{array}{rl} \mathbb{P}_X : X(\Omega) & \to [0,1] \\ S & \mapsto \mathbb{P}_X(S) = \mathbb{P}(X^{-1}(S)) = \mathbb{P}(\{\omega; X(\omega) \in S\}) \end{array}$$

- C'est une probabilité
- Connaissance de la loi de probabilité $\Rightarrow$ connaissance de X
- Loi de probabilité s'exprime par la fonction de répartition dans le cas réel.

Definition

Soit X une variable aléatoire réelle définie sur $(\Omega, \mathcal{A}, \mathbb{P})$. On appelle fonction de répartition de X la fonction définie par :

$$\begin{aligned} F_X : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}_X(]-\infty, x]) \end{aligned}$$

Propriétés :

- Fonction croissante : soient deux réels $a < b$, alors $F_X(a) \leq F_X(b)$
- Fonction continue à droite : soit $b \in \mathbb{R}$, soit $(b_n)_{n\geq 1}$ suite décroissante, ayant b pour limite lorsque $n \to \infty$, alors $\lim_{n\to\infty} F_X(b_n) = F_X(b)$

- avec des limites à gauche (cadlag) :notons $F_X(a-) = \lim_{x\uparrow a} F_X(x)$
- $\lim\limits_{x\to+\infty} F_X(x) = 1$
- $\lim\limits_{x\to-\infty} F_X(x) = 0$
- On peut utiliser la fonction de répartition pour calculer diverses probabilités :
  - $\mathbb{P}(X < a) = F_X(a-)$
  - $\mathbb{P}(a{<}X{\leq}b) = F_X(b) - F_X(a),\ \forall a < b$ deux réels
  - $\mathbb{P}(a \leq X \leq b) = F_X(b) - F_X(a-)$
  - $\mathbb{P}(a < X < b) = F_X(b-) - F_X(a)$
  - P(X =a) = $\mathrm{F}_X(a) - F_X(a-)$

Vecteur aléatoire

Definition

Soient n variables aléatoires $X_1, \ldots, X_n$ définies sur $(\Omega, \mathcal{A}, \mathbb{P})$; $X_i : \Omega \to E_i$. On appelle X vecteur aléatoire de dimension n : $X(\omega) = (X_1(\omega), \ldots, X_n(\omega))$ pour $\omega \in \Omega$.

Definition

La loi de probabilité conjointe de X est $\{\mathbb{P}(X \in S; S \in E_1 \times \ldots \times E_n)\}$.

Si la loi de probabilité conjointe est connue, alors on peut déterminer les lois marginales, mais la réciproque n'est pas vraie !

Cas particulier dans lequel la réciproque est vraie :

Definition

Les variables aléatoires X_i sont dites indépendantes si pour $S_1, \ldots, S_n \in \mathcal{A}$, les événements $\{X_i \in S_i\}_{i=1,\ldots,n}$ sont indépendants : $\forall n-\text{uplet}(x_1, \ldots x_n) \in E_1 \times \ldots \times E_n$

$$\begin{aligned} \mathbb{P}(\{X_1 \in S_1\} \cap \ldots \cap \{X_n \in S_n\}) &= \mathbb{P}(\{X_1 \in S_1, \ldots, X_n \in S_n\}) \\ &= \prod_{i=1}^{n} \mathbb{P}(\{X_i \in S_i\}) \end{aligned}$$

Si X et Y sont indépendantes, la réciproque est vraie.

Exemples de vecteurs aléatoires :

- Couple aléatoire discrêt : on dispose d'une urne contenant quatre jetons numérotés de 1 à 4, et on tire au sort successivement deux jetons sans remise. On note (X, Y) les résultats des deux tirages. On a : $\mathbb{P}(X = i, Y = i) = 0$ pour tout i entre 1 et 4 et $\mathbb{P}(X = i, Y = j) = 1/12$ si $1 \leq i, j \leq 4$ et $i \neq j$.

- Couple aléatoire continu : Soit $D = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 \leq 1\}$ le disque unité de $\mathbb{R}^2$ et f la fonction définie sur $\mathbb{R}^2$ par : $f(x, y) = \dfrac{1}{\pi} \mathbb{1}_D(x, y)$. f est bien une densité de probabilité sur $\mathbb{R}^2$: c'est la densité de la loi uniforme sur D, c'est-à-dire la loi que l'on obtient en jetant un point au hasard et uniformément sur D.

variable aléatoire discrête - Modélisation

Soit X une v.a. discrête.

- $X(\Omega) = \{x_1, ..., x_n\}$ fini, ou $\{x_i;\ i \geq 1\}$ dénombrable
- Loi de probabilité : distribution des probabilités $(p_i)_{i\geq 1}$ associées aux valeurs x_i

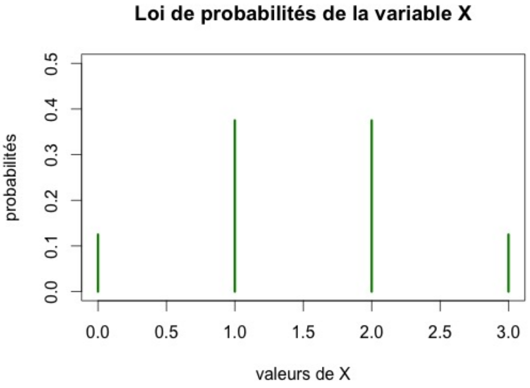
$$p_i = \mathbb{P}(X = x_i),\ i \geq 1 \text{ avec } \begin{cases} 0 \leq p_i \leq 1 \\ \displaystyle\sum_{i=1}^{\infty} p_i = 1 \end{cases}$$

$\hookrightarrow$ $(x_i, p_i)_{i\geq 1}$ loi de probabilité de X

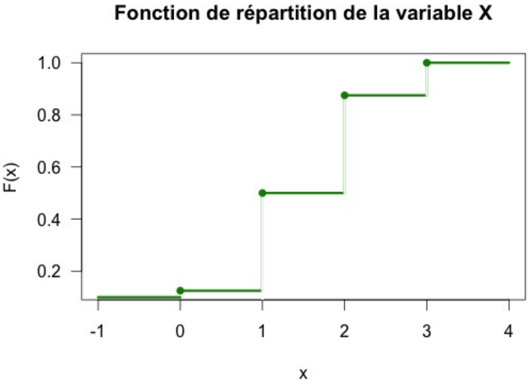
- *Exemple : X nombre de Piles obtenus en jetant trois pièces non biaisées.*

x_i	0	1	2	3
p_i	1/8	3/8	3/8	1/8

- Représentation graphique : diagramme en bâtons



- Fonction de répartition : fonction constante par morceaux
$$F(x) = \sum_{x_i \leq x} p_i$$



Vecteur aléatoire discrêt

- Lorsque les n variables aléatoires $X_i : \Omega \to E_i$ définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ sont discrètes : $X(\omega) = (X_1(\omega), \ldots, X_n(\omega))$ pour $\omega \in \Omega$.
- Loi de probabilité conjointe de X :
  $\{\mathbb{P}(X = (x_1, \ldots x_n); x_i \in E_i \, \forall i\} = \{\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n); x_i \in E_i \, \forall i\}$

Definition

La loi marginale de X_1 est donnée par :

$$\mathbb{P}(X_1 = x_1) = \sum_{(x_2, \ldots, x_n) \in E_2 \times \ldots \times E_n} \mathbb{P}(X = (x_1, \ldots, x_n))$$

Exemple : jet de 3 pièces non biaisées

$Y = (Y_1, Y_2)$; $Y_1 = \mathbb{1}_A$; $Y_2 = \mathbb{1}_B$ *avec les événements :*

A = *"On obtient au plus une fois Pile"*

B= *"On obtient au moins une fois Pile et au moins une fois Face"*

ω	$Y(\omega)$
(P,P,P)	(0,0)
(P,P,F)	(0,1)
(P,F,P)	(0,1)
(P,F,F)	(1,1)
(F,P,P)	(0,1)
(F,P,F)	(1,1)
(F,F,P)	(1,1)
(F,F,F)	(1,0)

Loi conjointe de Y :

Y	0	1	Y_1
0	$1/8$	$3/8$	$1/2$
1	$1/8$	$3/8$	$1/2$
Y_2	$1/4$	$3/4$	1

On vérifie que $\mathbb{P}(Y = (i,j)) = \mathbb{P}(Y_1 = i) \times \mathbb{P}(Y_2 = j)$, $\forall i, j \in \{0, 1\}$, *alors* $Y_1 \perp\!\!\!\perp Y_2$.

- Loi de la somme de deux variables aléatoires X et Y discrêtes indépendantes :
  $\mathbb{P}(X + Y = s) = \sum_{j \leq s} \mathbb{P}(X = j) \times \mathbb{P}(Y = s - j)$
  C'est le produit de convolution des lois de X et Y :
  $\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$.

Suite de l'exemple : loi de $Y_1 + Y_2$

ω	$(Y_1, Y_2)(\omega)$	$Y_1(\omega) + Y_2(\omega)$
(P,P,P)	(0,0)	0
(P,P,F)	(0,1)	1
(P,F,P)	(0,1)	1
(P,F,F)	(1,1)	2
(F,P,P)	(0,1)	1
(F,P,F)	(1,1)	2
(F,F,P)	(1,1)	2
(F,F,F)	(1,0)	1

Alors la loi de $Y_1 + Y_2$ est :

0	1	2
1/8	4/8	3/8

On l'obtient aussi avec le produit de convolution :

$\mathbb{P}(Y_1 + Y_2 = 0) = \mathbb{P}(Y_1 = 0) \times \mathbb{P}(Y_2 = 0) = \frac{1}{2}.\frac{1}{4} = \frac{1}{8}$

$\mathbb{P}(Y_1 + Y_2 = 1) = \mathbb{P}(Y_1 = 0) \times \mathbb{P}(Y_2 = 1) + \mathbb{P}(Y_1 = 1) \times \mathbb{P}(Y_2 = 0) = \frac{1}{2}.\frac{3}{4} + \frac{1}{2}.\frac{1}{4} = \frac{4}{8}$

$\mathbb{P}(Y_1 + Y_2 = 2) = \mathbb{P}(Y_1 = 1) \times \mathbb{P}(Y_2 = 1) = \frac{1}{2}.\frac{3}{4} = \frac{3}{8}$

Lois usuelles discrètes

Loi de Bernouilli

Soit E une épreuve comportant deux résultats possibles S et E (nommés "succès" et "échec") : épreuve ou schéma de Bernouilli (*Mathématicien suisse Jacques Bernouilli 1654-1705, "Ars Conjectandis"*)
On définit la variable aléatoire indicatrice de l'événement "‘succès"=S ($\Omega = \{S, E\}$) :

$$X \mapsto \begin{cases} 1 \text{ si "succès"} & \mathbb{P}(X = 1) = p \\ 0 \text{ si "échec"} & \mathbb{P}(X = 0) = 1 - p \end{cases}$$

Notation : $\mathcal{L}(X) = \mathcal{B}e(p)$.

Exemple : jet d'une pièce de monnaie ; $X = \begin{cases} 1 \textit{ si on obtient Pile} \\ 0 \textit{ sinon} \end{cases}$

$$X \sim \mathcal{B}e\left(\frac{1}{2}\right)$$

Loi binomiale

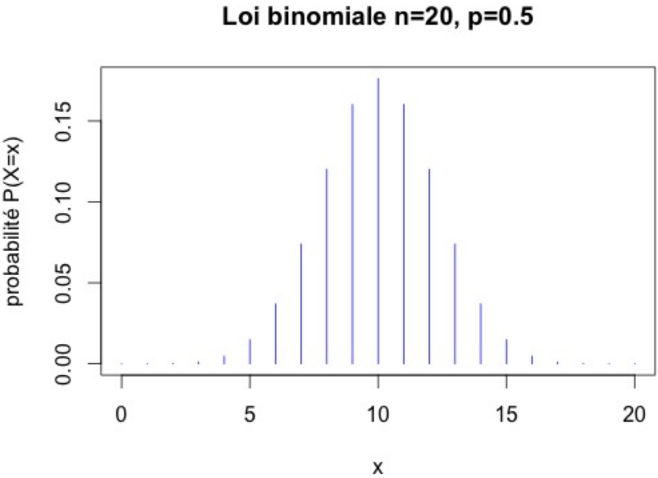
Une variable aléatoire Y est dite binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0,1]$ notée $\mathcal{B}(n,p)$ si $Y(\Omega) = \{0,1,2,...,n\}$ et

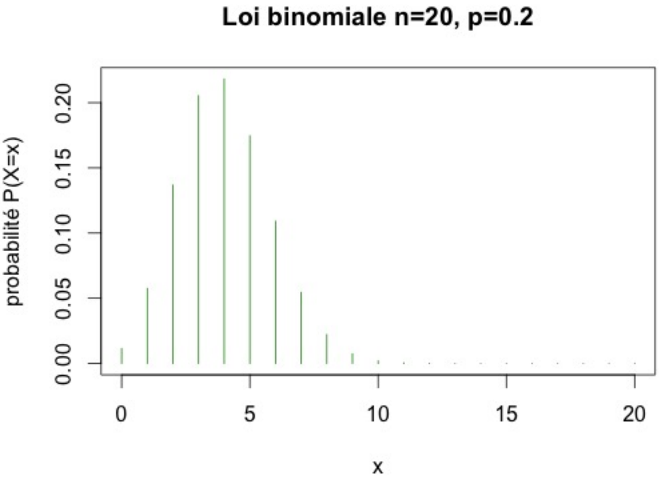
$$\mathbb{P}(Y = k) = C_n^k \, p^k \, (1-p)^{n-k}, \text{ pour } k = 0,1,2,...,n$$

- La loi binomiale modélise le nombre de succès dans n épreuves de Bernouilli $\mathcal{B}e(p)$ indépendantes. $Y = X_1 + ... + X_n$ avec $X_1,...,X_n$ n variables de même loi $\mathcal{B}e(p)$ indépendantes.
- Remarque : Si $(X_i)_{i=1,...,n}$ sont des variables discrètes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$, $X_i : \Omega \to E_i$, on dit qu'elles sont indépendantes si

$$\mathbb{P}(X_1 = x_1, \ldots X_n = x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i)$$

- *Exemple : la variable X de notre exemple précédent (nombre de Piles obtenus lors du jet de 3 pièces) suit une loi Binomiale $\mathcal{B}(3, \frac{1}{2})$.*
- Distribution symétrique si $p = 0,5$; dissymétrique sinon.




Loi binomiale n=20, p=0.2
probabilité P(X=x)
0.00
0.05
0.10
0.15
0.20
0
5
10
15
20
x

- Si $X_1, ..., X_N$ sont des variables mutuellement indépendantes de lois respectives $\mathcal{B}(n_i, p)$ (pour $i = 1, ..., N$), alors $Y = X_1 + ... + X_N$ est une variable de loi $\mathcal{B}(n_1 + ... + n_N, p)$
- $\mathcal{L}(n - Y) = \mathcal{B}(n, 1 - p)$ en choisissant de compter les échecs.
- extension de la loi binomiale : loi multinomiale.

Loi géométrique

Une variable aléatoire X suit une loi géométrique de paramètre $p \in [0,1]$ si $X(\Omega) = \mathbb{N}^*$ et

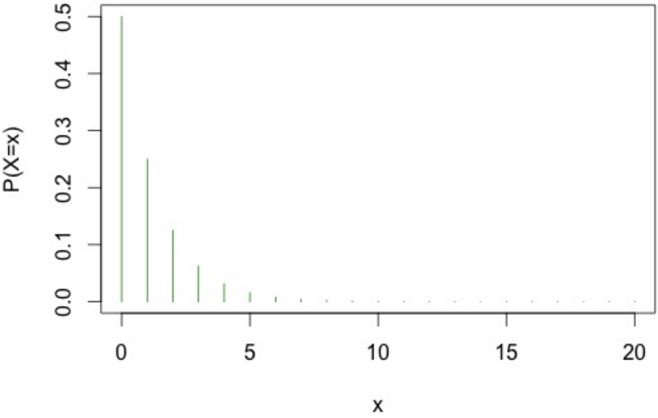
$$\forall k \in \mathbb{N}^*, \mathbb{P}(X = k) = (1-p)^{k-1}\, p$$

Notation : $X \sim \mathcal{G}(p)$

- C'est une loi sans mémoire :

$$\mathbb{P}(X = k + t / X > k) = \mathbb{P}(X = t), \;\forall k \in \mathbb{N}, t \in \mathbb{N}^*$$

- *Exemple : on lance une pièce de monnaie jusqu'à ce qu'elle tombe sur Pile. La variable du nombre de lancés effectués* $X \sim \mathcal{G}(\frac{1}{2})$


P(X=x)
0.0
0.1
0.2
0.3
0.4
0.5
0
5
10
15
20
x

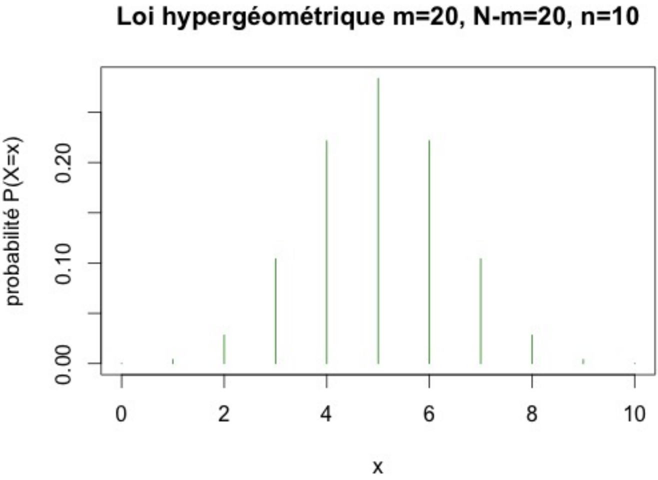
Loi hypergéométrique

Une variable aléatoire X suit une loi hypergéométrique de paramètres N, n, p, avec $N, n \in \mathbb{N}^*$, $p \in [0, 1]$ tels que $Np \in \mathbb{N}$ et $N \geq n$, si $X(\Omega) = \{0, 1, ..., n\}$ et

$$\forall k \in 0, 1, ..., n,\ \mathbb{P}(X = k) = \frac{C_{Np}^{k}\ C_{N(1-p)}^{n-k}}{C_{N}^{n}}$$

Condition : $\max\{0; n - (N - Np)\} \leq k \leq \min\{n; Np\}$
Notation : $X \sim \mathcal{H}(N, n, p)$

$\mathcal{H}(40, 10, 1/2)$



- Une urne contient N boules : m blanches, et $N - m$ noires. On tire n boules sans remise, donc les tirages sont non indépendants. La variable aléatoire X du nombre de boules blanches tirées suit une loi hypergéométrique $\mathcal{H}(N, n, p)$ avec $p = \frac{m}{N}$
- Lorsque $N \to \infty$, la loi hypergéométrique $\mathcal{H}(N, n, p)$ converge vers la loi binomiale $\mathcal{B}(n, p)$. Ceci est admis dès que N est grand devant n : $N \geq 10n$.

Loi de Poisson

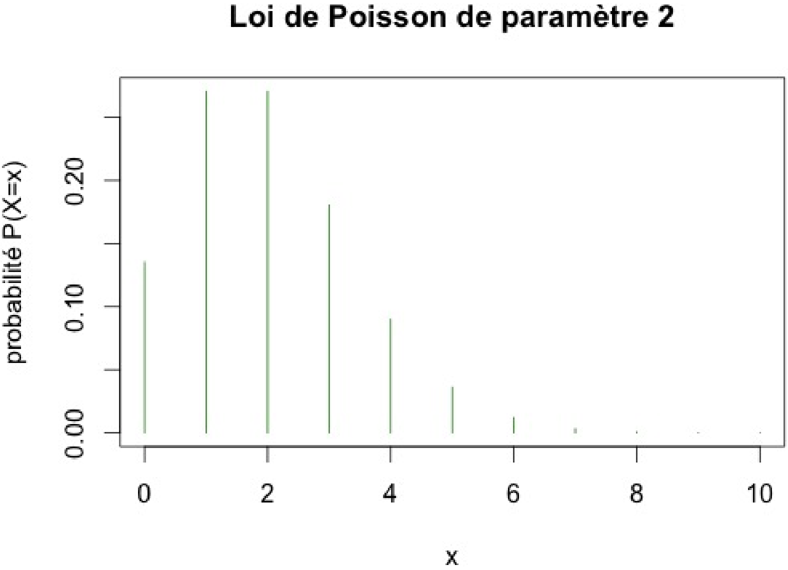
Une v.a. discrête X suit une loi de Poisson de paramètre $\lambda > 0$ lorsque $X(\Omega) = \mathbb{N}$ et

$$\forall k \in \mathbb{N}, \ \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

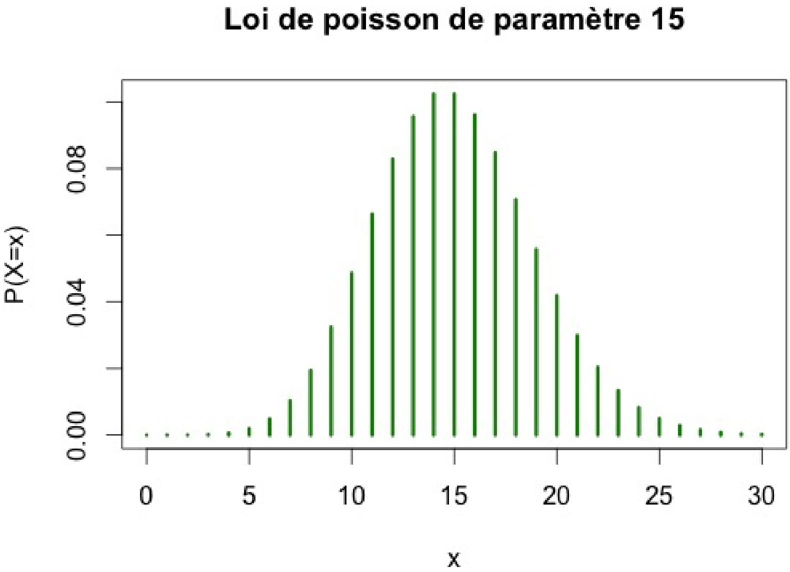
Notation : $\mathcal{L}(X) = \mathcal{P}(\lambda)$.

- introduite par le mathématicien français Simeon Denis Poisson en 1838 (*Recherche sur la probabilités des jugements en matière criminelle et en matière civile*).

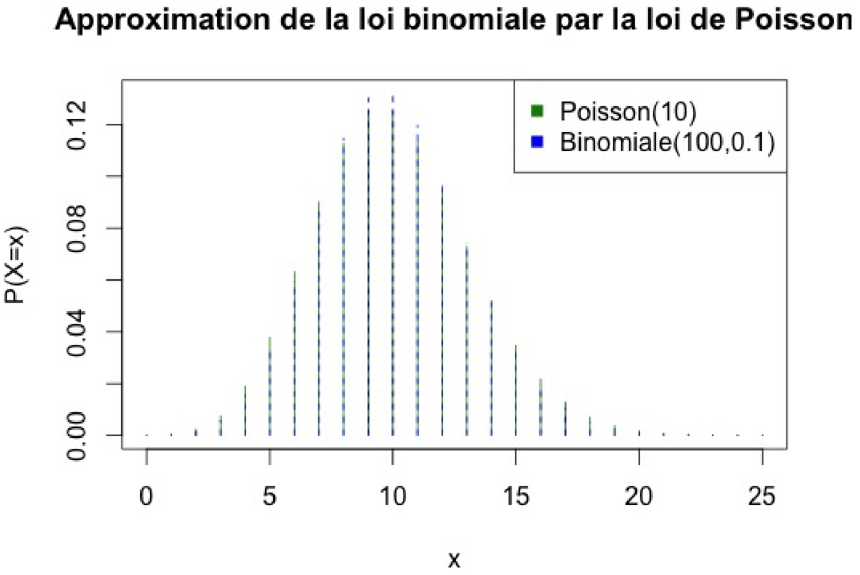
- Distribution dissymétrique étalée à droite



- Distribution tend à devenir symétrique lorsque λ augmente.



- Loi des événements rares ; exemples : nombre de clients arrivant dans une file d'attente dans un intervalle de temps défini, nombre de fautes par pages d'un livre, nombre de particules émises par un matériau radioactif dans un intervalle de temps défini...
- Si X, Y sont deux variables indépendantes de lois respectives $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$, alors $Y = X + Y$ est une variable de loi $\mathcal{P}(\lambda + \mu)$
- Approximation d'une loi binomiale $\mathcal{B}(n, p)$ par une loi de Poisson $\mathcal{P}(n \times p)$ si n est grand et p petit : $n \geq 10$, $p \leq 0,1$.


Approximation de la loi binomiale par la loi de Poisson
Poisson(10)
Binomiale(100,0.1)
P(X=x)
0.00
0.04
0.08
0.12
0
5
10
15
20
25
x

Caractéristiques

Definition

Soit une variable aléatoire discrète $X : \Omega \to E$. Supposons $E = \{x_i; i \geq 1\}$. L'espérance mathématique de X est définie par :
$\mathbb{E}(X) = \sum_{i \geq 1} x_i \, p_i$

- L'espérance représente la valeur moyenne de la variable X.
- Cas particulier : Si A est un événement, alors $\mathbb{E}(\mathbb{I}_A) = \mathbb{P}(A)$.
- Soit $f : E \to \mathbb{R}$. Alors : $\mathbb{E}(f(X)) = \sum_{i \geq 1} f(x_i) \, p_i$ sous réserve que la série soit absolument convergente i.e. $\sum_{i \geq 1} \mid f(x_i) \mid p_i < +\infty$ (on dit que X est intégrable).

- *Retour à notre exemple :* $\mathbb{E}(X) = 1.5$
- Propriétés : soient X, Y des v.a., α, β des réels

$$\left.\begin{aligned} \mathbb{E}(\alpha X + \beta) &= \alpha \mathbb{E}(X) + \beta \\ \mathbb{E}(X + Y) &= \mathbb{E}(X) + \mathbb{E}(Y) \end{aligned}\right\} \text{linéarité}$$

  Généralisation : $\mathbb{E}[\alpha f(X) + \beta g(Y)] = \alpha \mathbb{E}[f(X)] + \beta \mathbb{E}[g(Y)]$
- Indépendance : Si X et Y sont deux variables indépendantes, alors $\mathbb{E}(XY) = \mathbb{E}(X) \times \mathbb{E}(Y)$.

Definition

Une v.a.r. X est dite centrée si $\mathbb{E}(X) = 0$.

$\hookrightarrow$ $X - \mathbb{E}(X)$ est une v.a.r. centrée.

Definition

On appelle variance de la v. a. X : $\mathrm{Var}(X) = s_X^2 = \mathbb{E}[(X - \mathbb{E}(X))^2]$.
Pour une variable discrête :
$\mathrm{Var}(X) = \sum_i (x_i - \mathbb{E}(X))^2 p_i = \sum_i x_i^2 p_i - (\mathbb{E}(X))^2$.

- estimation de la dispersion de la variable autour de la moyenne
- Autre formulation : $s_X^2 = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$

Definition

On appelle écart-type de la variable X : $s_X = \sqrt{s_X^2}$.

Plus clair pour une raison d'unité.
Retour à notre exemple : $s_X^2 = 0.75$; $s_X = 0.866$

Propriétés :

- $\mathrm{Var}(X) \geq 0$
- $\mathrm{Var}(\alpha X + \beta) = \alpha^2 \mathrm{Var}(X)$
- Si $X \perp\!\!\!\perp X$, alors $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.
- $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$

Definition

Soit (X, Y) le couple aléatoire de loi conjointe $\mathbb{P}$. On appelle covariance de X et Y :

$$\mathrm{Cov(X,Y)} = \mathrm{s}^2_{X,Y} \quad = \mathrm{E[(X - E(X))(Y - E(Y))]}$$
$$= \sum_{i,j\geq 1} (x_i - \mathbb{E}(X))\,(y_j - \mathbb{E}(Y))\,\mathbb{P}(X = x_i, Y = y_j)$$

Propriétés :

- $\text{Cov}(X, Y) = \mathbb{E}(X \times Y) - \mathbb{E}(X) \times \mathbb{E}(Y) = \sum_{i,j \geq 1} x_i\, y_j\, \mathbb{P}(X = x_i; Y = y_j) - \mathbb{E}(X) \times \mathbb{E}(Y)$
- X, Y indépendants $\Rightarrow \text{Cov}(X, Y) = 0$ Réciproque fausse !
- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- $\text{Cov}(X, X) = \text{var}(X)$
- $\text{Cov}(\alpha X, Y) = \alpha \text{Cov}(X, Y)$
- $\text{Cov}(X + X', Y + Y') = \text{Cov}(X, Y') + \text{Cov}(X', Y') + \text{Cov}(X', Y) + \text{Cov}(X, Y)$

Definition

Une v.a.r. X est dite réduite si $s_X = 1$.

$\hookrightarrow \dfrac{X - \mathbb{E}(X)}{s_X}$ est une v.a.r. centrée et réduite.

Caractéristiques des lois usuelles vues

Loi de X	$\mathbb{E}(X)$	$\mathrm{Var}(X)$
$\mathcal{B}e(p)$	p	$p(1-p)$
$\mathcal{B}(n,p)$	np	$np(1-p)$
$\mathcal{G}(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$\mathcal{H}(N,n,p)$	np	$np(1-p)\frac{N-n}{N-1}$
$\mathcal{P}(\lambda)$	λ	λ

Definition

Soit (X, Y) un couple aléatoire de loi conjointe $\mathbb{P}$. On appelle coefficient de corrélation linéaire : $\rho(X, Y) = \dfrac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$.

- $-1 \leq \rho(X, Y) \leq 1$
- $X \perp\!\!\!\perp Y \Rightarrow \rho(X, Y) = 0$
- Si $\rho(X, Y) \simeq 1$, alors dépendance linéaire positive entre X et Y, i.e. $Y = aX + b, a > 0$ (idem avec -1, $a < 0$).

Revenons à notre exemple :

$\mathbb{E}(Y_1 \times Y_2) = \dfrac{3}{8}; Cov(Y_1, Y_2) = \dfrac{3}{8} - \dfrac{1}{2} \times \dfrac{3}{4} = 0$

$\rho(Y_1, Y_2) = 0$ *prévisible car* $Y_1 \perp\!\!\!\perp Y_2$.

Definition

Soit une variable aléatoire X, soit $k \in \mathbb{N}^*$. On dit que X admet un **moment d'ordre k** si X^k est intégrable. Le moment d'ordre k est défini par $\mathbb{E}(X^k)$, on le notera m_k.

L'espérance est donc le moment d'ordre 1.

Definition

Soit une variable aléatoire X, soit $k \in \mathbb{N}^*$. Le **moment centré d'ordre k** est défini par $\mathbb{E}((X - \mathbb{E}(X))^k)$, on le notera μ_k.

La variance est donc le moment centré d'ordre 2.

Definition

La fonction génératrice des moments d'une variable X est la fonction $\phi : \mathbb{R} \to \mathbb{R} \cup \infty$ définie par :

$$\phi(t) = \mathbb{E}(e^{tX}) = \sum_i e^{tx_i} p_i$$

lorsque la série est convergente.

Le développement en série entière de ϕ est :

$$\phi(t) = \phi(0) + \phi'(0)\frac{t}{1!} + \phi''(0)\frac{t^2}{2!} + \ldots$$

Or $\phi^{(k)}(t) = \sum_i x_i^k e^{tx_i} p_i$

$\phi^{(k)}(0) = \sum_i x_i^k p_i = \mathbb{E}(X^k) = m_k$, donc :

$$\phi(t) = m_0 + m_1 t + m_2 \frac{t^2}{2!} + \ldots$$

Loi de probabilité conditionnelle

Soient deux v.a.r. X et Y, notons $(x_i, p_i)_{i\in I}$ et $(y_j, p_j)_{j\in J}$ leurs lois de probabilité respectives, et $\mathbb{P}(X = x_i, Y = y_j) = p_{ij}(i \in I, j \in J)$ les probabilités jointes.
Loi de probabilité de Y conditionnelle à $\{X = x_i\}$:

$$b_i(j) = \mathbb{P}(Y = y_j / X = x_i) = \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(X = x_i)} = \frac{p_{ij}}{p_i}$$

Définition simple de l'espérance conditionnelle : si Y admet une espérance finie, alors on définit l'espérance conditionnelle par rapport à $\{X = x_i\}$ par :

$$\mathbb{E}(Y/X = x_i) = \sum_{j\in J} b_i(j) y_j$$

Remarque : $\mathbb{E}(\mathbb{E}(Y/X)) = \mathbb{E}(Y)$.

Soit X une variable aléatoire continue définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans $E = X(\Omega)$. *Rappel : $X(\Omega)$ est un ensemble infini non dénombrable, i.e. $\mathbb{R}^n$, un intervalle de $\mathbb{R}^n$, ou une union d'intervalles de $\mathbb{R}^n$*

On se place dans le cas $n = 1$.

- Loi de probabilité $(\mathbb{P}(X \in A); A \subset E)$ est trop vaste pour être décrite correctement.
- On peut décrire la loi par la fonction de répartition F_X qui est continue
- Pour tout x réel, $\mathbb{P}(X = x) = F_X(x) - F_X(x-) = 0$

- **Cas particulier important :** variable à densité :
  Si en tout réel x, $F_X(x)$ est continue et admet une dérivée notée $f_X(x)$, on dit que X est absolument continue de densité de probabilité f_x
- $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$ pour tout x réel.
- Il existe des variables continues sans densité (hors programme).

- La densité caractérise la loi de la variable X.
- Elle vérifie :
  1. $f_X(x) \geq 0, \forall x \in \mathbb{R}$
  2. $\displaystyle\int_{-\infty}^{+\infty} f_X(x)dx = 1$
- Réciproque : le critère pour qu'une fonction f soit une densité de probabilité est : il faut et il suffit que f soit positive et d'intégrale sur $\mathbb{R}$ égale à 1.

- $\mathbb{P}(X \in S) = \displaystyle\int_S f_X(x)dx$ en particulier si $S = (a; b)$ alors
  $\mathbb{P}(a \leq (<) X \leq (<) b) = \displaystyle\int_a^b f_X(x)dx$
- Moments :
  1. Espérance mathématique : $\mathbb{E}(X) = \displaystyle\int_{-\infty}^{+\infty} x\, f_X(x)\, dx$
  2. Variance : $\mathrm{Var}(X) = \displaystyle\int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2\, f_X(x)\, dx$
- Les propriétés dont similaires au cas discrêt (linéarité de l'intégrale).
- On a : $\mathbb{E}(g(X)) = \displaystyle\int_{-\infty}^{+\infty} g(x)\, f_X(x)\, dx$

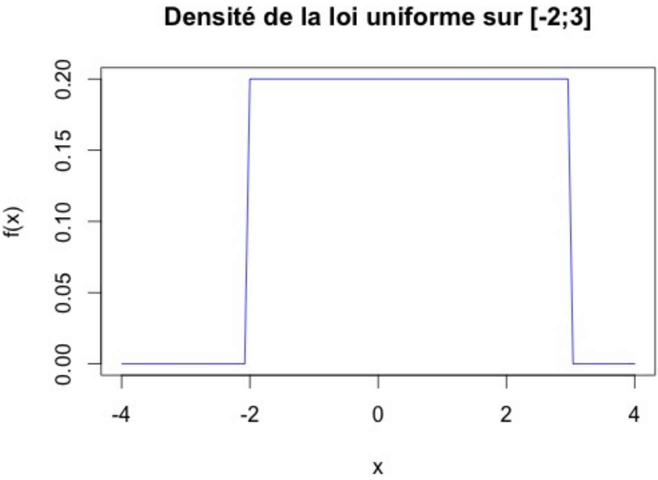
Lois usuelles continues

Loi uniforme

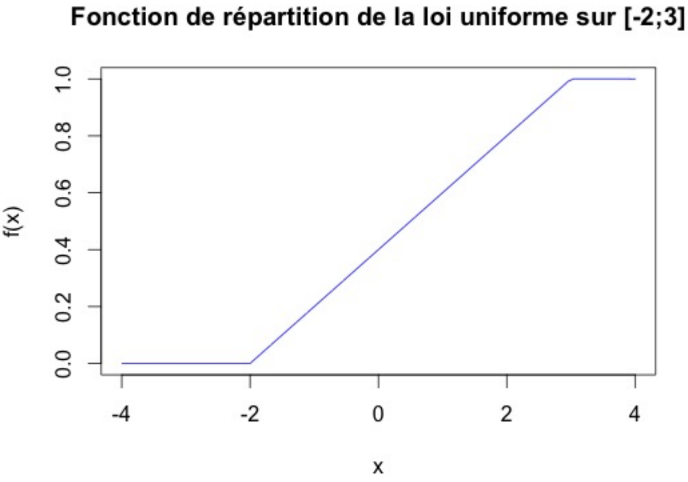
Soient a, b deux réels tels que $-\infty < a < b < +\infty$. La variable X suit la loi uniforme sur l'intervalle $[a, b]$ lorsque la densité de probabilité est : $f_X(x) = \dfrac{1}{b-a}\, \mathbb{I}_{[a,b]}(x)$.

Notation : $\mathcal{L}(X) = \mathcal{U}_{(a,b)}$

- $\mathbb{E}(X) = \dfrac{a+b}{2}$
- $s_X^2 = \dfrac{(b-a)^2}{12}$


Densité de la loi uniforme sur [-2;3]
f(x)
0.00
0.05
0.10
0.15
0.20
-4
-2
0
2
4
x

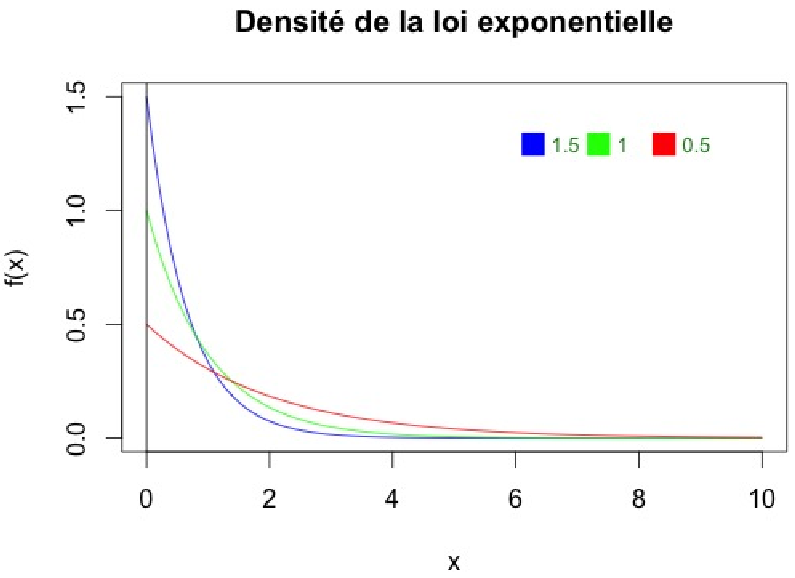
- Fonction de répartition :
$$F(x) = \begin{cases} 0 \text{ si } x \leq a \\ \frac{x-a}{b-a} \text{ si } a < x \leq b \\ 1 \text{ si } x > b \end{cases}$$
- Remarque : c'est la seule variable aléatoire calculée directement sur ordinateur, via la fonction **rand**. On l'utilise pour simuler d'autres lois de probabilités continues.


Fonction de répartition de la loi uniforme sur [-2;3]
f(x)
0.0
0.2
0.4
0.6
0.8
1.0
-4
-2
0
2
4
x

Loi exponentielle

Soit un réel $\lambda > 0$. La variable X suit la loi exponentielle de paramètre λ lorsque $X(\Omega) = \mathbb{R}^+$ et sa densité de probabilité est : $f_X(x) = \lambda\, e^{-\lambda x}\, \mathbb{I}_{\mathbb{R}^+}(x)$.
Notation : $\mathcal{L}(X) = \mathcal{E}xp(\lambda)$

- Exemple : durée de fonctionnement d'un équipement technique, temps séparant les arrivées de deux clients successifs dans une file d'attente...
- $\mathbb{E}(X) = \dfrac{1}{\lambda}$; $\mathrm{Var}(X) = \dfrac{1}{\lambda^2}$
- Loi sans mémoire


Densité de la loi exponentielle
1.5
1
0.5
f(x)
x
0.0
0.5
1.0
1.5
0
2
4
6
8
10

Revenons à une loi uniforme sur $[0;1]$.

Theorem

Soit $F : \mathbb{R} \to [0;1]$ une fonction croissante, continue à droite, vérifiant $\lim_{-\infty} F = 0$ et $\lim_{\infty} F = 1$. On note $F^{-1} :]0;1[\to \mathbb{R}$ sa fonction inverse : $F^{-1}(y) = \inf\{x; F(x) \geq y\}, \forall y \in]0;1[$.
Soit U une variable de loi uniforme sur $[0;1]$. Alors la variable $X = F^{-1}(U)$ admet F comme fonction de répartition.

On pourra ainsi utiliser la loi uniforme sur $[0;1]$ pour simuler d'autres lois de probabilités continues.
Application : Si $X \sim \mathcal{E}xp(\lambda)$, sa fonction de répartition est :
$F(x) = (1 - e^{-\lambda x})\mathbb{1}_{\mathbb{R}^+}(x)$.
$F^{-1}(U)$ permet de simuler des observations de la variable X.

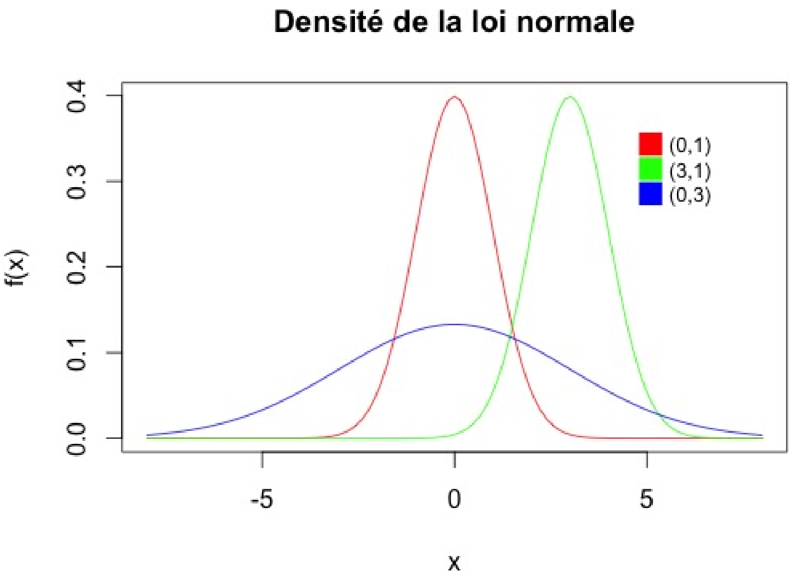
Loi normale (gaussienne/de Laplace-Gauss)

Soit $m \in \mathbb{R}$ et $\sigma > 0$. La variable X suit une loi normale de moyenne m et d'écart-type σ si $X(\Omega) = \mathbb{R}$ et elle admet pour densité de probabilité :

$$f_X(x) = \frac{1}{\sqrt{2\,\pi\,\sigma^2}}\ \exp\left[-\frac{(x-m)^2}{2\,\sigma^2}\right]$$

Notation : $\mathcal{L}(X) = \mathcal{N}(m, \sigma)$

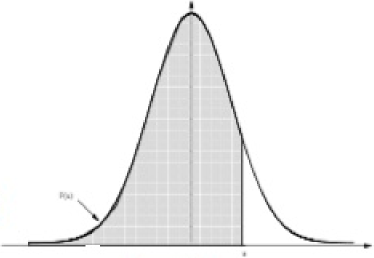
- Loi résultant d'un grand nombre de causes indépendantes dont les effets se cumulent et aucun n'est prépondérant (conditions de Borel), décrite par Gauss (1809) et Laplace (1812).
- Courbe en "cloche", symétrique autour de l'axe $x = m$, de valeur maximale $\frac{1}{\sigma\sqrt{2\pi}}$ en $x = m$.


Densité de la loi normale
(0,1)
(3,1)
(0,3)
f(x)
0.0
0.1
0.2
0.3
0.4
-5
0
5
x

- Stabilité : si X_1 et X_2 sont deux variables. indépendantes de lois respectives $\mathcal{N}(m_1, \sigma_1)$ et $\mathcal{N}(m_2, \sigma_2)$, alors $\mathcal{L}(X_1 + X_2) = \mathcal{N}(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.

- Soit la variable centrée réduite associée à X :
  $T = \dfrac{X - m}{\sigma} \Leftrightarrow X = \sigma T + m.$
- Sa loi est $\mathcal{L}(T) = \mathcal{N}(0, 1)$.
- Sa fonction de densité est : $f_T(x) = \dfrac{1}{\sqrt{2\pi}} \exp\left(-\dfrac{x^2}{2}\right)$
- Notons π sa fonction de répartition. Elle vérifie :
- $\left\{ \begin{array}{rcl} \pi(-x) & = & 1 - \pi(x) \\ \mathbb{P}(|T| \leq t) & = & 2\pi(t) - 1 \end{array} \right.$
- Primitives non directement calculables $\Rightarrow$ tables statistiques de la fonction de répartition.

F(0.24)=0.5948



u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.527
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.567
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.606
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.644
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.680
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.715
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.748
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.779
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.807
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834

Fonction d'une variable aléatoire

Proposition

Soit X v.a. de densité $f_X : \mathbb{R} \to \mathbb{R}+$; soient $a, b, \alpha, \beta \in \mathbb{R} \cup -\infty; +\infty$ tels que $a < b, \alpha < \beta$ et $\mathbb{P}(a < X < b) = 1$. Soit $g :]a, b[\to]\alpha, \beta[$ bijective telle que g, g^{-1} sont C^1.
Posons $Y = g(X)$ et notons $h = g^{-1}$. Alors Y est une v.a. de densité : $f_Y(y) = f_X(h(y)). \mid h'(y) \mid .\mathbb{I}_{y \in]\alpha,\beta[}$.

ou

Théorème de transfert

Pour toute fonction $g : \mathbb{R} \to \mathbb{R}$ continue et bornée, on a :

$$\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x)\, f_X(x)\, dx$$

Application : Soit X variable absolument continue de densité f_X. On pose $Y = h(X)$ avec $h : \mathbb{R} \to \mathbb{R}$. On cherche à déterminer la densité de Y.

Pour g définie dans le théorème, on a :

$$\mathbb{E}(g(Y)) = \int_{-\infty}^{+\infty} g(h(x))\, f_X(x)\, dx$$

Le changement de variable : $y = h(x)$ donne

$$\mathbb{E}(g(Y)) = \int_{-\infty}^{+\infty} g(y)\, f_Y(y)\, dy$$

f_Y sera la densité de la variable Y.

Vecteur aléatoire

- Soient n variables aléatoires $X_1, ..., X_n$ définies sur $(\Omega, \mathcal{F}, \mathbb{P})$. On appelle $X = (X_1, ..., X_n)$ vecteur aléatoire dans $\mathbb{R}^n$.
- Fonction de répartition : $F_X : \mathbb{R}^n \to \mathbb{R}$

$$F_X(x_1, ..., x_n) = \mathbb{P}(X_1 \leq x_1, \ ... \ X_n \leq x_n)$$

- On dit que X admet une densité $f_X : \mathbb{R}^n \to \mathbb{R}_+$ si

$$F_X(x_1, ..., x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f_X(t_1, ..., t_n)\, dt_1 ... dt_n$$

- La loi marginale de X_i est donnée par :
$$f_{X_i}(x_i) = \int_{\mathbb{R}^{n-1}} f_X(t_1, ..., t_{i-1}, t_{i+1}, ..., t_n)\, dt_1 ... dt_{i-1} dt_{i+1} ... dt_n$$

- Les variables $X_1, ..., X_n$ sont mutuellement indépendantes si

$$f_X(x_1, ..., x_n) = \prod_{i=1}^{n} f_{X_i}(x_i),\ \forall (x_1, ..., x_n) \in \mathbb{R}^n$$

- Covariance d'un couple aléatoire (X, Y) :
$$\int_{\mathbb{R}} \int_{\mathbb{R}} (x - \mathbb{E}(X))\,(y - \mathbb{E}(Y))\, f(x, y)\, dx\, dy = \int_{\mathbb{R}} \int_{\mathbb{R}} xyf(x, y)dxdy - \mathbb{E}(X)\mathbb{E}(Y)$$
- Densité conditionnelle sachant $X_n = x_n$:
$$f_{X_1,...,X_{n-1}}(x/X_n = x_n) = \frac{f_X(x_1, ..., x_n)}{f_{X_n}(x_n)}$$

Fonction d'un vecteur aléatoire

Proposition

Soit $X = (X_1, ..., X_m)$ vecteur aléatoire de densité $f_X : \mathbb{R}^m \to \mathbb{R}+$; soit $U \subset \mathbb{R}^m$ ouvert tel que $\mathbb{P}(X \in U) = 1$. Soient $V \subset \mathbb{R}^m$ un autre sous-ensemble ouvert et $g : U \to V$ bijective :
$g(x_1, ..., x_m) = (g_1(x_1, ..., x_m), ..., g_m(x_1, ..., x_m))$.
Posons $Y = g(X)$ et notons $h = g^{-1}$. Si h est C^1, alors Y est une v.a. continue dont la fonction de densité conjointe est :
$f_Y(y) = f_X(h(y)). \mid J_h(y) \mid .\mathbb{I}_{y \in V}$, où J_h est le Jacobien de h :

$$J_h(y) = \det \begin{pmatrix} \frac{\partial h_1}{\partial y_1} & \cdots & \frac{\partial h_m}{\partial y_1} \\ \vdots & & \vdots \\ \frac{\partial h_1}{\partial y_m} & \cdots & \frac{\partial h_m}{\partial y_m} \end{pmatrix}$$

Soit une suite de variables $X_1, ..., X_n$ et X toutes définies sur $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition

La suite $(X_i)_i$ converge vers la variable X de fonction de répartition F_X quand $n \to \infty$

- en loi si la suite des fonctions de répartition $F_{X_1}, ..., F_{X_n}$ tend vers F_X pour tout x pour lequel F_X est continue. On note : $X_n \overset{\text{Loi}}{\to} X$
- en probabilité si pour tout $\epsilon > 0$, $\mathbb{P}(|X_n - X| < \epsilon) \to 1$ quand $n \to \infty$. On note : $X_n \overset{\mathbb{P}}{\to} X$ Remarque : la convergence en probabilité implique la convergence en loi.
- presque sûrement si $X_n(\omega) \to X(\omega)$ $\forall \omega \in \Omega$ quand $n \to \infty$ sauf sur un ensemble de probabilité nulle. On note : $X_n \overset{p.s.}{\to} X$

Inégalité de Markov

Soit X une variable aléatoire à valeurs positives ; soit un réel $a > 0$.
Alors $\mathbb{P}(X \geq a) \leq \dfrac{1}{a}\,\mathbb{E}(X)$.

Inégalité de Bienaymé-Tchebycheff

Soit X une variable aléatoire de moyenne $\mathbb{E}(X)$ et d'écart-type $\sigma > 0$. Pour tout $t > 0$, $\mathbb{P}\left(|X - \mathbb{E}(X)| \geq t\right) \leq \dfrac{\sigma^2}{t^2}$.

Loi faible des grands nombres

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, de même loi et intégrables. Alors, lorsque n tend vers l'infini, on a :

$$\frac{X_1 + \ldots + X_n}{n} \to \mathbb{E}(X_1) \text{ en probabilité}$$

Théorème Central Limite

Soit une suite de n variables aléatoires $X_1, X_2, ..., X_n$ indépendantes et de même loi (i.i.d.), d'espérance m et de variance σ^2. Soit la variable définie comme la moyenne $\overline{X_n} = \frac{1}{n} \sum_{i=1}^{n} X_i$. Alors lorsque $n \to \infty$, avec Z variable de loi $\mathcal{N}(0, 1)$:

$$\sqrt{n}\,\frac{\overline{X_n} - m}{\sigma} \to_{n\to\infty} Z \text{ en loi}$$

- Etabli par De Moivre, puis Laplace, version finale dûe à Lindenberg en 1922.
- Une v.a.r. résultant de la somme de plusieurs v.a.r. de même loi et mêmes paramètres, est distribuée selon une loi normale centrée réduite lorsque le nombre d'épreuves n tend vers l'infini.
- S'applique quelle que soit la loi des variables X_n !
- Première conséquence : grande importance de la loi normale ! de plus, les phénomènes dont la variation est engendrée par un nombre important de causes indépendantes sont susceptibles d'être modélisés par une loi normale.

- Seconde conséquence : approximation de certaines lois de probabilité par d'autres.
  Exemple : $(X_n)_{n\geq 1}$ suite i.i.d. de loi $\mathcal{B}e(p)$; rappels :
  $\mathbb{E}(X_n) = p$; $\text{Var}(X_n) = p(1-p)$.
  Alors $\sqrt{n}\,\dfrac{\overline{X_n} - p}{\sqrt{p(1-p)}} = \dfrac{X_1 + \ldots + X_n - n\,p}{\sqrt{n\,p(1-p)}} = Z_n \to_{n\to\infty} Z$ en loi

  On remarque que $X_1 + \ldots + X_n$ suit une loi binomiale $\mathcal{B}(n,p)$, donc on approchera une loi binomiale par une loi normale centrée réduite quand n est grand et p éloigné de 0 et 1.
  En pratique dès que $n\,p > 15$, on approche la loi binomiale $\mathcal{B}(n,p)$ par une loi normale de paramètres $n.p$ et $\sqrt{n.p.(1-p)}$.

Lois binomiale et normale