Université Cheikh Anta Diop de Dakar

Ecole Supérieure Polytechnique

Probabilité – Statistique

Chapitre 3 : Variables aléatoires et fonction de variables aléatoires

Dr. Dahirou WANE

Année universitaire : 2021 – 2022

Plan

1 Quelques Définitions

2 Fonction de répartition

3 Variable aléatoire discrète

4 Variable aléatoire continue

5 Couple de variables aléatoires

6 Variables aléatoires indépendantes

Plan

1. Quelques définitions

Quelques définitions

Dans une épreuve, une variable aléatoire (v.a) X est une quantité dont la valeur, a priori aléatoire, est déterminée à l'issue d'un tirage.

Elle est donc représentée comme une application X définie sur l'ensemble Ω des résultats possibles de l'épreuve.

Si X a un nombre fini ou infini dénombrable d'éléments, X est une v.a discrète

Si X a un nombre fini non dénombrable d'éléments, X est une v.a continue

Fonction de répartition

La fonction de répartition d'une v.a. X à valeurs dans un ensemble $\mathrm{E} \subset \mathbb{R}$ est la fonction, notée F_x, qui à tout réel x associe la probabilité que X prenne une valeur inférieure ou égale à x :

$$F(X) : \begin{array}{l} \mathbb{R} \rightarrow [0,1] \\ x \rightarrow P(X \leq x) \end{array}$$

$$\mathrm{F(X) = P(X \leq x)}$$

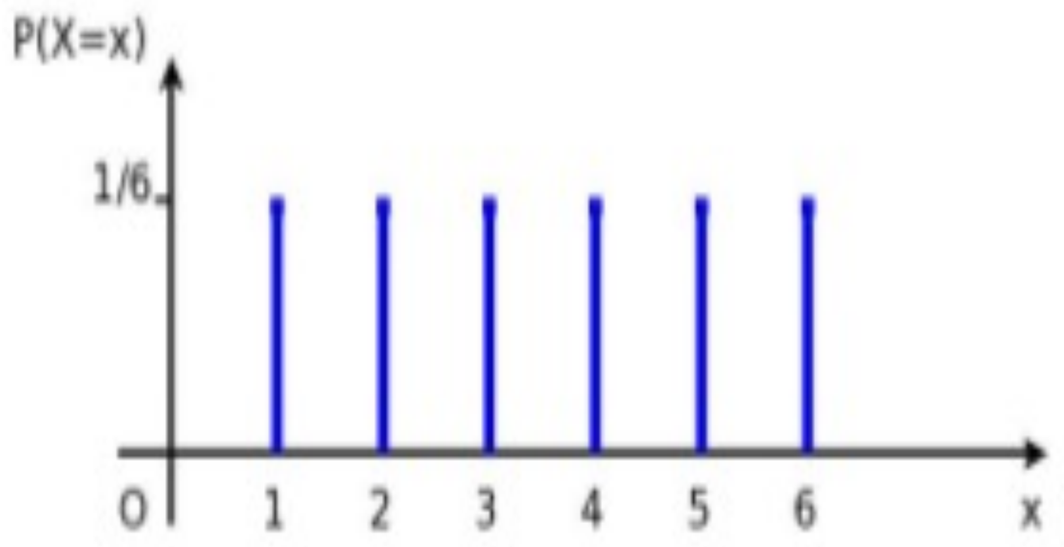
Propriétés

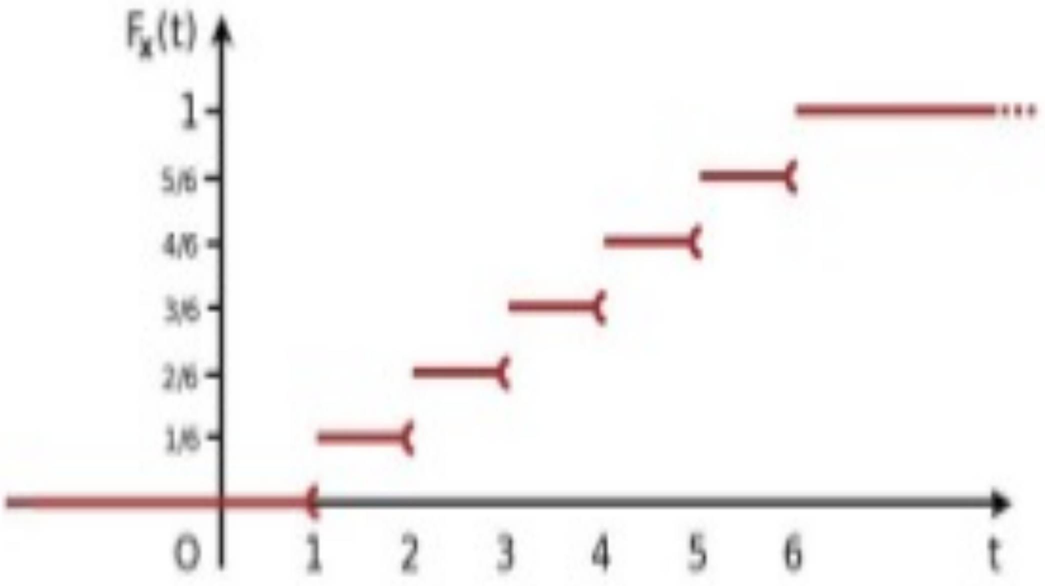
- F(x) est positive et non décroissante : $x_2 \geq x_1 \rightarrow F(x_2) \geq F(x_1)$
- F(x) $\longrightarrow$ 1, lorsque x $\longrightarrow$ + ∞
- F(x) $\longrightarrow$ 0, lorsque x $\longrightarrow$ - ∞
- $P(x_1 \leq x \leq x_2) = F(x_2) - F(x_1)$
- F(x) = P(X ≤ x)

Fonction de répartition

Exemple 1 : X est le résultat du lancer d'un dé à 6 faces.

x	1	2	3	4	5	6
$\mathbb{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$





Variable aléatoire discrète

v.a. dont les différentes valeurs sont finies ou infinies dénombrables.

Exemple 2 : jet d'une pièce

Soit X le nombre de jets successifs nécessaires pour obtenir la face 'Pile' au

$i^{ème}$ lancer.

$E = (1,2,3, \ldots x)$ Face 'F' : x-1 fois et Face 'P' : au x^e lancer

$P(X=x) = (1/2)^{x-1}.(1/2) = (1/2)^x$

Loi de probabilité

X	1	2	3	…	…	X	Somme
P(X=x)	1/2	1/4	1/8	…	…	$(1/2)^x$	1

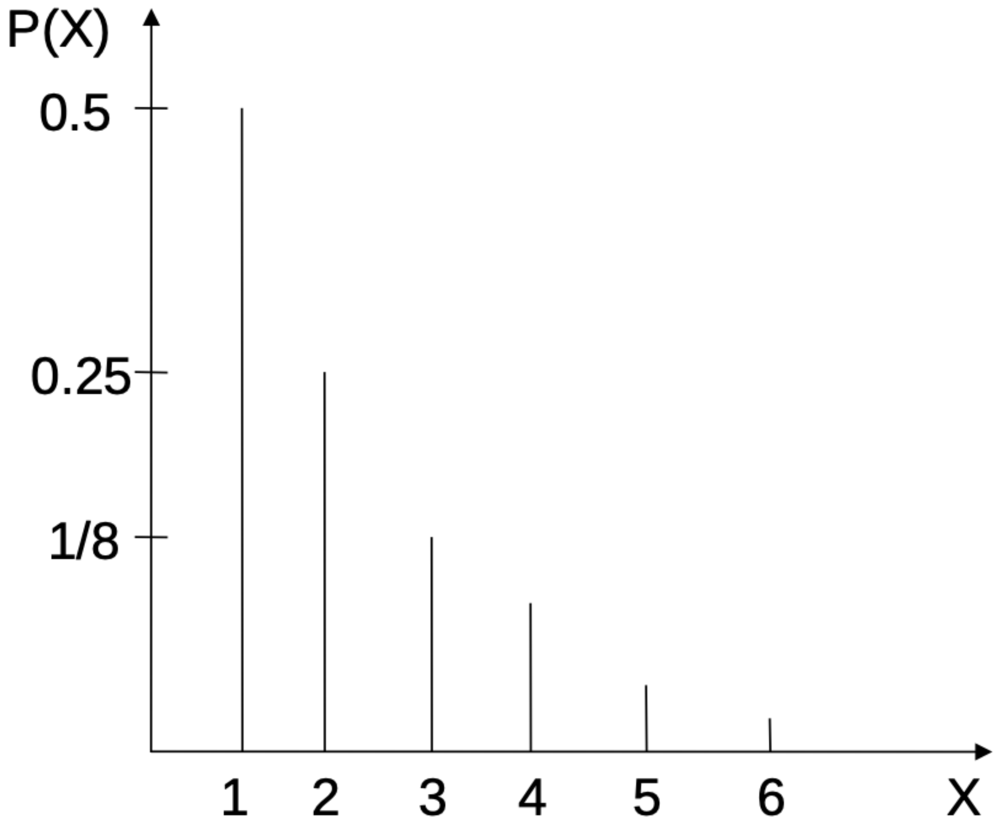
$$\sum_{x\to\infty} P(X=x) = \sum_{x=1}^{\infty}\left(\frac{1}{2}\right)^x$$

$$= \frac{a}{1-q} = \frac{\frac{1}{2}}{1-\frac{1}{2}} = 1$$

Suite géométrique de raison q =1/2 et premier terme a = 1/2.

Variable aléatoire discrète

X	1	2	3	…	…	X	Somme
P(X=x)	1/2	1/4	1/8	…	…	$(1/2)^x$	1
F(x)	1/2	3/4	7/8	…	…	1	



Variable aléatoire continue

Soit X une v.a qui peut prendre n'importe quelle valeur de l'ensemble de définition.

$$\exists\ f(x)\ /\ \begin{cases} 1)\ f(x) \geq 0 \\ 2)\ \int_{-\infty}^{+\infty} f(x).dx = 1 \end{cases}$$

$$P(a < x < b) = \int_a^b f(x).dx$$

f(x) est appelée fonction de distribution de probabilité ou fonction de densité de probabilité.

Exemple 3 : Soit la fonction f(x) définie par :

$$f(x) = \begin{cases} c.x^2; \quad 0 < x < 3 \\ 0\ ;\ ailleurs. \end{cases}$$

a) Calculer la constante c en utilisant la définition de la fonction de densité de probabilité

b) Calculer P(1 < x < 2)

c) Déterminer la fonction de répartition F(x).

a) Détermination de la valeur de c : $\int_{-\infty}^{+\infty} c.x^2.dx = 1 = \left[\frac{c.x^3}{3}\right]_0^3 = 9c \Rightarrow c = \frac{1}{9}$

b) Calcul de P(1 < x < 2) : $P(1 < x < 2) = \int_1^2 \frac{1}{9}x^2dx = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$

C) Détermination de F(x) : $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x).dx$

1) x < 0 : $F(x) = \int_{-\infty}^{x} f(x).dx = 0$

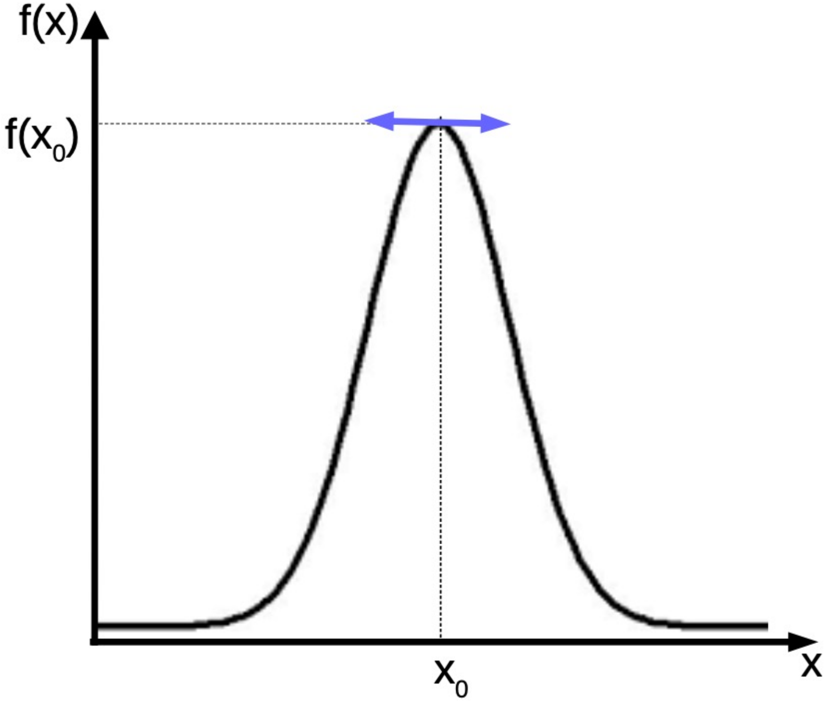
2) 0 ≤ x < 3 : $F(x) = \int_{-\infty}^{0} f(x).dx + \int_0^x f(x).dx = \int_0^x \frac{1}{9}x^2.dx = \frac{x^3}{27}$

3) x ≥ 3 : $F(x) = \int_{-\infty}^{0} f(x).dx + \int_0^3 f(x).dx + \int_3^x f(x).dx = 1$

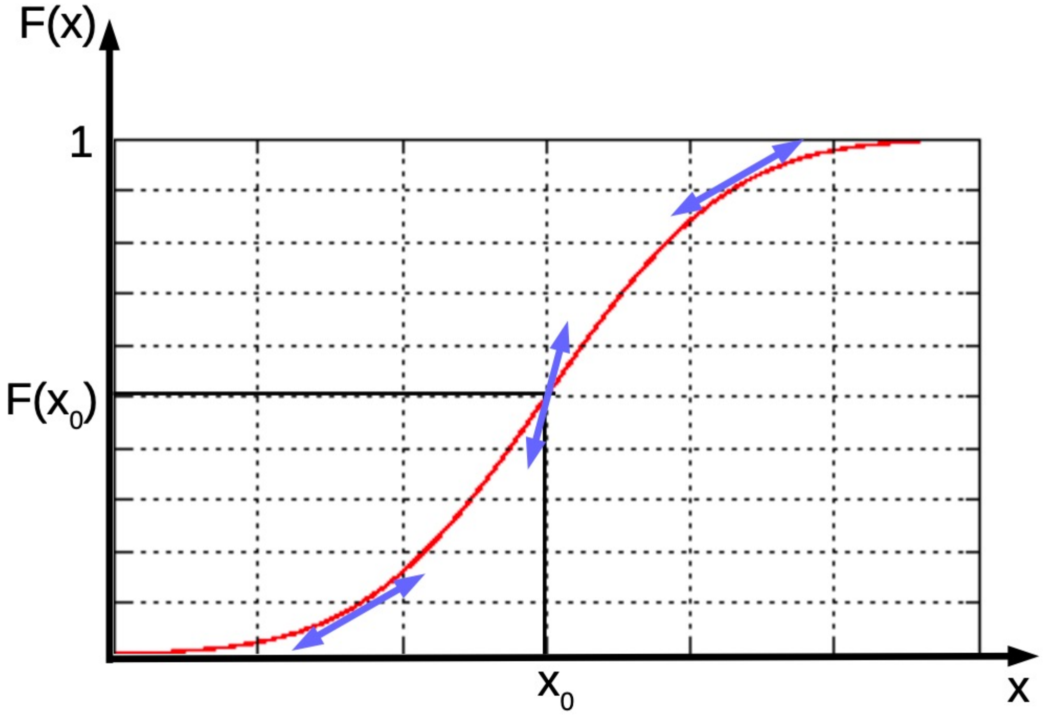
Variable aléatoire continue

Relation F(x) et la fonction densité de probabilité en x_0

$$F(x_0) = \int_{-\infty}^{x_0} f(x).dx$$



$f'(x_0) = 0\,;\; f''(x_0) < 0$



$F''(x_0) = 0\,;\; F'''(x_0) < 0$

M_0 est appelé point d'inflexion

Variable aléatoire à 2 dimensions

Variable aléatoire discrète

Si X et Y sont 2 v.a. discrètes, on définit la fonction de probabilité conjointe de X et Y par

$$P(X = x, Y = y) = f(x, y) \qquad Avec \begin{cases} f(x,y) \geq 0. \\ \sum_x \sum_y f(x,y) = 1 \end{cases}$$

Si X = $(x_1, x_2, \ldots, x_n)$ et Y = $(y_1, y_2, \ldots, y_n)$

$P(X = xi, Y = yj) = f(xi, yj)$: fonction de probabilité conjointe

Couple de variables aléatoires

Variable aléatoire discrète

$$X = (x_i)_{i \in I} \; et \; Y = (y_j)_{j \in J} \qquad P_{ij} = P(X = xi, Y = yj)$$

Lois marginales

$$P(X = xi) = \sum_{j \in J} P(X = xi, Y = yj) = \sum_{j \in J} P_{ij} = P_{i.}$$

$$P(Y = yj) = \sum_{i \in I} P(X = xi, Y = yj) = \sum_{i \in I} P_{ij} = P_{.j}$$

Loi conditionnelle

$$P(X = x_i / Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{P_{ij}}{P_{.j}} = P_i^j$$

Indépendance

$$P(X = x_i / Y = y_j) = P(X = xi) . P(Y = yj)$$

Couple de variables aléatoires

Variable aléatoire discrète

Considérons deux familles complètes d'événements (A_1, A_2, …, A_n) et (B_1, B_2, …, B_n)

$$P(A_i \cap B_j) = P_{ij}$$

Tableau de contingence

$$P(A_i) = P_{i.} \ et \ P(B_j) = P_{.j}$$

Image du tableau

$$P(A_i) = P\left(\bigcup_{j=1}^{n}(A_i \cap B_j)\right) = \sum_{j=1}^{n} P(A_i \cap B_j)$$

Par conséquent : $P_{i.} = \sum_{j=1}^{n} P_{ij}$ et de même $P_{.j} = \sum_{i=1}^{n} P_{ij}$

Définition des probabilités conditionnelles

$$P(A_i/B_j) = \frac{P(A_i \cap B_j)}{P(B_j)} = \frac{P_{ij}}{P_{.j}} \qquad P(B_j/A_i) = \frac{P(A_i \cap B_j)}{P(A_i)} = \frac{P_{ij}}{P_{i.}}$$

Couple de variables aléatoires

Variable aléatoire discrète

Loi des probabilités totales $$P(A_i) = \sum_j P(B_j)P(A_i/B_j)$$

A_i et B_j indépendants $$P(A_i \cap B_j) = P(A_i).P(B_j)$$

$$P_{ij} = P_{i.}P_{.j}$$

B / A	B_1	B_2	
A_1	P_{11}	P_{12}	$P_{1.}$
A_2	P_{21}	P_{22}	$P_{2.}$
A_3	P_{31}	P_{32}	$P_{3.}$
	$P_{.1}$	$P_{.2}$	1

Couple de variables aléatoires

Tableau de probabilité conjointe

X \ Y	y_1	y_2	…	y_j	…	y_n	Total
x_1	f_{11}	f_{12}		f_{1j}		f_{1n}	$f_1(x_1)$
x_2	f_{21}	f_{22}		f_{2j}		f_{2n}	$f_1(x_2)$
…	…	…		…		…	…
x_i	f_{i1}	f_{i2}		f_{ij}		f_{in}	$f_1(x_i)$
…	…	…		…		…	…
x_n	f_{n1}	f_{n2}		f_{nj}		f_{nn}	$f_1(x_n)$
Total	$f_2(y_1)$	$f_2(y_2)$	…	$f_2(y_j)$	…	$f_2(y_n)$	1

Couple de variables aléatoires

Tableau de probabilité conjointe

Les probabilités

$$P(X = x_i) = f_1(x_i) = \sum_{j}^{n} f(x_i, y_j) = P_{i.}$$

$$P(Y = y_j) = f_2(y_j) = \sum_{i}^{n} f(x_i, y_j) = P_{.j}$$

Sont les fonctions de probabilité marginales de X et Y, respectivement

On note

$$\sum_{i=1}^{n} f_1(x_i) = 1 \quad et \quad \sum_{j=1}^{n} f_2(y_j) = 1$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} f(x_i, y_j) = 1$$

Fonction de répartition conjointe de X et Y

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{u \leq x}\sum_{v \leq y} f(u, v)$$

Couple de variables aléatoires

Tableau de probabilité conjointe

Exemple 4 : jet de deux dés équilibrés. On définit la v.a. X comme le nombre de points amenés par le premier dé et la v.a. Y la somme des points donnés par les deux dés.

Loi de probabilité

X \ Y	2	3	4	5	6	7	8	9	10	11	12	Loi marginale de X
1	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	0	1/6
2	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	1/6
3	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	1/6
4	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	1/6
5	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	1/6
6	0	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	1/6
Loi marginale de Y	1/36	1/18	1/12	1/9	5/26	1/6	5/26	1/9	1/12	1/18	1/36	1

Couple de variables aléatoires

Tableau de probabilité conjointe

Formule des probabilités conditionnelles

Si $P_{i.} \neq 0 \;\; P(Y = y_j / X = x_i) = P_{j/i} = \dfrac{P_{ij}}{P_{i.}}$

De même, si $P_{j.} \neq 0 \;\; P(X = x_i / Y = y_j) = P_{i/j} = \dfrac{P_{ij}}{P_{.j}}$

Formule des probabilités composées

$$P_{ij} \neq 0 \,;\; P_{i.}P_{j/i} = P_{.j}P_{i/j}$$

Couple de variables aléatoires

Variable aléatoire continue

$$F(x,y) = P(X < x, Y < y)$$

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x.\partial y}$$

Lois marginales

$$F_X(x) = P(X < x) = F(x, +\infty)$$

$$F_Y(y) = P(Y < y) = F(+\infty, y)$$

Densités marginales

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y).dy$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y).dx$$

Lois Conditionnelles

$$f_X(x/Y = y) = \frac{f(x,y)}{f_Y(y)}$$

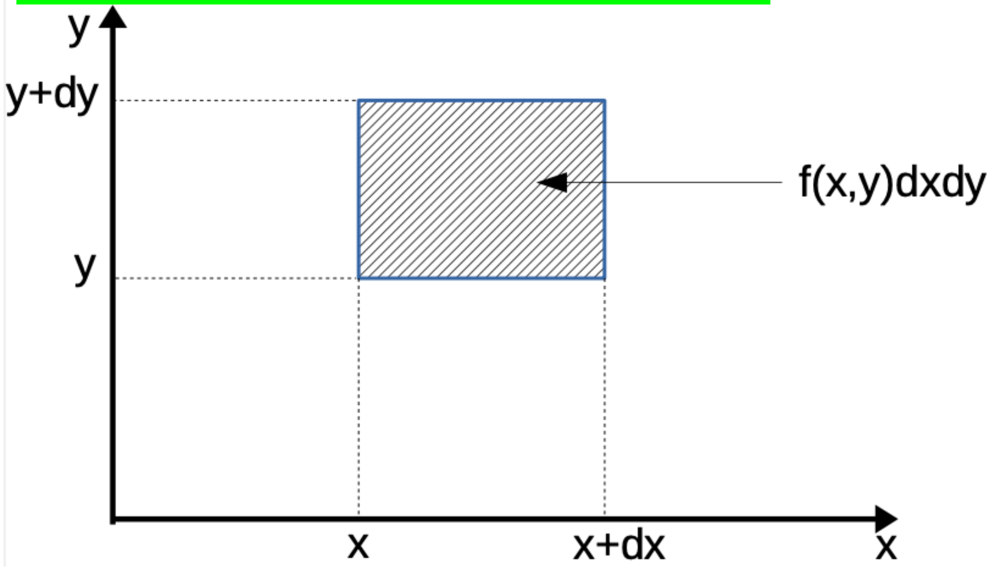
$$f_Y(y/X = x) = \frac{f(x,y)}{f_X(x)}$$

Indépendance

$$f(x,y) = f_X(x) f_Y(y)$$

Couple de variables aléatoires

Variable aléatoire continue



Par analogie avec le cas discret, en remplaçant les sommes par des intégralités, la fonction densité de probabilité conjointe X et Y est définie

$$1°)\ f(x,y) \geq 0$$

$$2°)\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y).dx.dy = 1$$

$$P(x \leq X \leq x+dx, y \leq Y \leq y+dy) = f(x,y)dxdy$$

Densités marginales

$$\int_{-\infty}^{+\infty} f(x,y).dx = f(y)$$

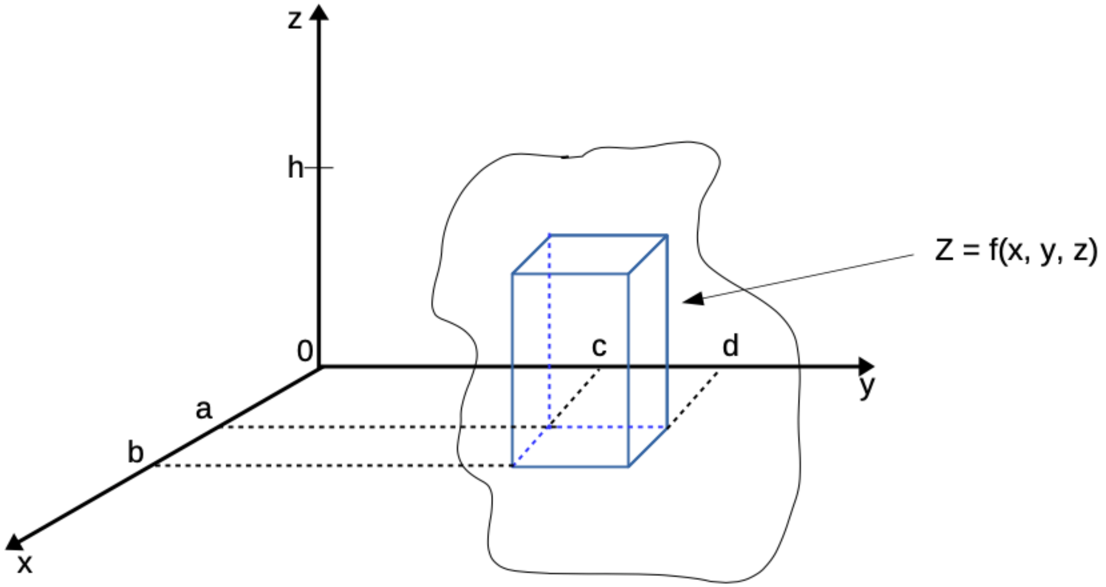
$$\int_{-\infty}^{+\infty} f(x,y).dy = f(x)$$

$$\boldsymbol{Z = f(x,y)}$$

Probabilité de surface

Couple de variables aléatoires

Variable aléatoire continue



Le volume total limité par la surface formée par le plan (x,y) est égale à 1

$$P(a \leq x \leq b, c \leq y \leq d, 0 \leq z \leq h) = \int_{x=a}^{b} \int_{y=c}^{d} \int_{z=0}^{h} f(x,y,z) dx.dy.dz$$

$$\boldsymbol{Z = f(x,y,z)}$$

Probabilité de volume

Couple de variables aléatoires

Variable aléatoire continue

Fonction de distribution conjointe de X et Y est :

$$F(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)du.dv$$

Donc $$\frac{\partial^2 F}{\partial x \partial y} = f(x,y)$$

Fonction de distributions marginales

$$P(X \leq x) = F_1(x) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{+\infty} f(u,v)du.dv$$

$$P(Y \leq y) = F_2(y) = \int_{u=-\infty}^{+\infty} \int_{v=-\infty}^{y} f(u,v)du.dv$$

variables aléatoires indépendantes

Soient deux v.a. X et Y. Si les événements X = x et Y = y sont indépendantes pour tout x et y, on dit que X et Y sont v.a. indépendantes

Dans ce cas:

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$ Cas discret

$$f(x, y) = f_1(x)f_2(y)$$ Cas continu

Soient X et Y deux v.a. Si les événements $X \leq x$ et $Y \leq y$ sont indépendantes on dit que les deux v.a. X et Y sont indépendantes.

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

$$F(x, y) = F_1(x)F_2(y)$$

Où $F_1(x)$ et $F_2(y)$ sont les fonctions de distributions marginales de X et Y respectivement.

Inversement X et Y sont indépendants si pour tout x et y leur fonction de distribution conjointe F(X,Y) peut être exprimée comme un point $F(x,y) = F_1(x).F_2(y)$

Distributions conditionnelles

La densité de probabilité conditionnelle de y liée par X = x est :

$$Z(y/x) = \frac{f(x,y)}{f(x)} = \frac{f(x,y)}{\int_{-\infty}^{+\infty} f(x,y)dy}$$

De même la densité de probabilité conditionnelle de X liée par Y = y est :

$$Z(x/y) = \frac{f(x,y)}{f(y)} = \frac{f(x,y)}{\int_{-\infty}^{+\infty} f(x,y)dx}$$

X et Y sont indépendantes si pour tout couple (x,y) : f(x, y) = f(x).f(y)

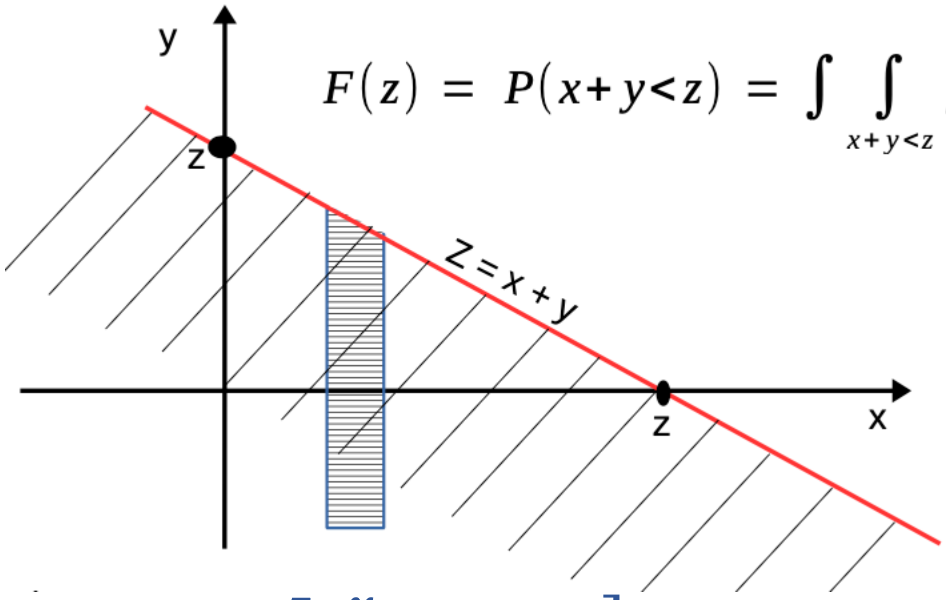
Loi de densité : variable aléatoire à densité

(x,y) est dite à densité si $\exists\ (x,y),\quad F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)du.dv$

$1°)\ f(x,y) \geq 0\ \forall\ (x,y) \in \mathbb{R}^2$ $\qquad$ $2°)\ \iint_{\mathbb{R}^2} f_{X,Y}(x,y)dx.dy = 1$

Somme de deux v.a. indépendantes

$Z = x + y \quad F(Z) = P(Z < z)$



$$F(z) = P(x+y<z) = \int\int_{x+y<z} f(x,y)dxdy$$

Si X et Y sont indépendantes :

$$F(z) = \iint_{x+y<z} f(x).g(y)dx.dy = \int_{-\infty}^{+\infty}\left[f(x).dx \int_{-\infty}^{z-x} g(y).dy\right]$$

$$F(z) = \int_{-\infty}^{+\infty} G(z-x)f(x).dx$$ De même $$F(z) = \int_{-\infty}^{+\infty} F(z-y)g(y).dy$$

$$f(z) = \frac{d}{dz}\int_{-\infty}^{+\infty} F(z-y)g(y).dy$$

$$f(z) = \int_{-\infty}^{+\infty} f(z-y)g(y).dy = \int_{-\infty}^{+\infty} g(z-x)f(x).dx$$

Produit de convolution

Notion de moment et fonction

Théorie de la moyenne

Soit une fonction g(X), X = v.a. discrète ou continue. La moyenne de g(X) est définie par:

Cas discret : $E[g(X)] = \sum_i P_i g(x_i)$ Cas continu : $E[g(X)] = \int_D f(x)g(x)dx$

Si g(X) est l'espérance mathématique, alors $E[g(X)] = E(X)$

Moment d'ordre 1

Cas discret : $m_1 = E(X) = \sum_i P_i x_i$ Cas continu : $m_1 = E(X) = \int_D xf(x)dx$

La variable aléatoire est centrée si : $\bar{X} = E(X) = 0$

Moment d'ordre k

Cas discret : $m_k = E(X^k) = \sum_i P_i x_i^k$ Cas continu : $m_k = E(X^k) = \int_D x^k f(x)dx$

Variance et Écart-type

Variance $V(X) = \sigma_X^2 = E[(X - \bar{X})^2] = E[(X - E(X))^2]$

Cas discret : $\sigma_X^2 = \sum_i P_i(x_i - m_1)^2$ Cas continu : $\sigma_X^2 = \int_{-\infty}^{+\infty} (x - m_1)^2 f(x).dx$

Théorique : $V(X) = \sigma_X^2 = E[(X^2 - 2XE(X) + E^2(X))] = E(X^2) - m_1^2$

Changement de variable :

Translation $\sigma_{X+a}^2 = E\left[\left(X + a - \overline{(X+a)}\right)\right]^2 = E(X - \bar{X})^2 = \sigma_X^2$

Conclusion : il y a une invariance lorsque $X \rightarrow X + a$

Homothétie : $\sigma_{aX}^2 = E\left[\left(aX - \overline{(aX)}\right)\right]^2 = a^2[E(X^2) - (E(X))^2] = a^2(m_2 - m_1^2) = a^2\sigma_X^2$

Conclusion : $X \rightarrow X + a$; $\sigma_{aX}^2 = a^2.\sigma_X^2$

Écart-type $\sigma_X = \sqrt{V(X)}$

Variance et Écart-type

Exemple 5 : Soit X une variable aléatoire discrète

X	1	2	3	…	n
X^2	1	4	9	…	n^2
P(X)	P	qp	q^2p	…	q^np

Avec p + q = 1

a) Vérifier que la somme des probabilités est égale à 1

b) Déterminer l'espérance mathématique et l'écart-type de cette v.a.

Solution

a) $0 \leq p \leq 1$ et $0 \leq q \leq 1$ $\quad \sum_i P_i = p(1 + q + q^2 + \ldots) = p\dfrac{1 - q^n}{1 - q} = 1$

b) $E(X) = p + 2qp + 3q^2p + \ldots = p(1 + 2q + 3q^2 + \ldots)$

$$= p\frac{d}{dq}(q + q^2 + q^3 + \ldots) = p\frac{d}{dq}\left(\frac{q}{1-q}\right) = p\left(\frac{1-q+q}{(1-q)^2}\right) = p\left(\frac{1}{(1-q)^2}\right) \rightarrow E(X) = \frac{1}{p} = m_1$$

Variance et Écart-type

Exemple 5 : Soit X une variable aléatoire discrète

Solution (suite)

$$E(X^2) = p(1 + q2^2 + q^2 3^2 + \ \dots) = p(1 + 4q + 9q^2 + \ \dots)$$

$$= \mathrm{p}\frac{d}{dq}(\mathrm{q} + 2q^2 + 3q^3 + \ \dots) = p\frac{d}{dq}\left(q\left(\frac{d}{dq}\left(\frac{1}{1 - q}\right)\right)\right)$$

$$E(X^2) = \mathrm{p}\frac{d}{dq}\left(\frac{q}{(1 - q)^2}\right) = p(\frac{1+q}{(1-q)^3} = \frac{1+q}{p^2}$$

$$\mathrm{Var}(X) = \sigma_X^2 = m_2 - m_1^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}$$

$$\mathrm{Var}(X) = \sigma_X^2 = \frac{q}{p^2} \rightarrow \sigma_X = \frac{\sqrt{q}}{p}$$

Covariance et Coefficient de corrélation linéaire

X et Y deux variables aléatoires $\mathrm{Var}(X+Y) = V(\mathrm{X}) + \mathrm{V}(\mathrm{Y}) + 2E[(X-E(X))(Y-E(Y))]$

On pose par définition $\mathrm{Cov}(X,Y) = E[(X-E(X))(Y-E(Y))]$

Propriétés :

$$\mathrm{Cov}(X,Y) = \mathrm{Cov}(\mathrm{Y},\mathrm{X})$$

$$\mathrm{Cov}(X,Y+Z) = \mathrm{Cov}(\mathrm{X},\mathrm{Z}) + \mathrm{Cov}(\mathrm{X},\mathrm{Y})$$

$$\mathrm{Cov}(X,\lambda Y) = \lambda\mathrm{Cov}(\mathrm{Y},\mathrm{X})$$

$$\mathrm{Cov}(X,Y) = E(XY) - \mathrm{E}(\mathrm{X})\mathrm{E}(\mathrm{Y})$$

Cas particulier : X et Y deux v.a. indépendantes $Cov(X,Y) = 0$, La réciproque est fausse.

Mesure de la dépendance entre deux v.a. :

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \qquad -1 \le \rho \le 1$$

***ρ* : coefficient de corrélation linéaire**

Application

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

$\rho = \mp 1 \; si \; |Cov(X,Y)| = \sigma_X \sigma_Y$

Donc $(X - E(X))$ et $(Y - E(Y))$ $sont \; proportionnelles$

Soit $\mathrm{X} - \mathrm{E(X)} = a[Y - E(Y)])$

$\rho = \pm 1$ si $\exists$ $une \; relation \; linéaire \; entre \; X \; et$ Y)

$\rho = 0$: $pas \; de \; relation \; linéaire \; mais \; possibilité \; de \; relation \; fonctionnelle \; (E_X, X \; et \; X^2)$
X et Y non corrélées.

$\rho > 0$: $X \; et$ Y positivement corrélées

$\rho < 0$: $X \; et$ Y négativement corrélées