Chapitre III :

Loi de probabilité d’une variable aléatoire à deux dimensions

Section 1 : Loi de probabilité d'une variable à deux dimensions

1.1 Définition :

Soit Ω un espace fondamental et Δ une σ - algèbre muni d'une probabilité P. P est l'application qui associe a chaque élément A de Δ, son image dans l'intervalle [0,1], qui est en fait sa probabilité :

$$A \subset \Omega,\ A \in \Delta$$

$$P : (\Omega,\Delta) \longmapsto [0,1]$$

$$A \longmapsto P\{A\} \in [0,1]$$

Soit une application qui associe a chaque $e \in \Omega$ un vecteur

$$z(e) = (X_1(e), X_2(e)) \in IR^2$$

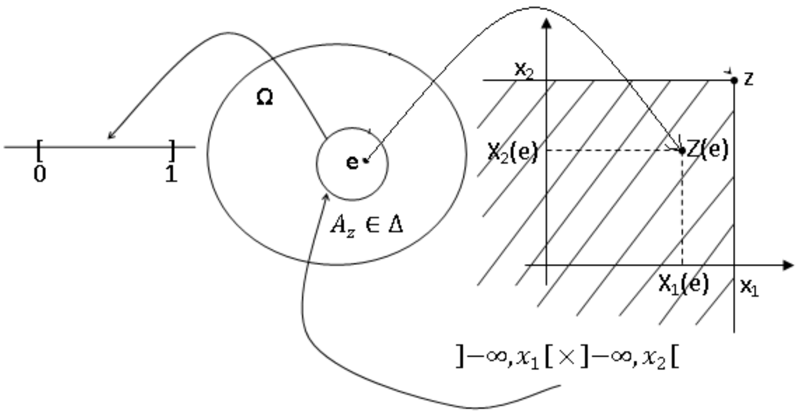
Soit A_z L'image inverse du rectangle ouvert

$$\{\ ]-\infty, x_1[\ \times\ ]-\infty, x_2[\ \} \quad \in A_z \Leftrightarrow [X_1(e) < x_1 \text{ et } X_2(e) < x_2]$$

Ou encore $\quad e \in A_z \Leftrightarrow (Z(e) < z)\ \ avec\ z = (x_1, x_2)$

On dit que l'application Z est une variable aléatoire à deux démenions si :

$$\forall\, z \in IR^2 ;\ A_z = Z^{-1}(\ ]-\infty, x_1[\ \times\ ]-\infty, x_2[\ ) \in \Delta.$$



1.2 Fonction de répartition :

La définition d'une variable aléatoire a deux dimensions Z affirme l'existence de A_Z dans Δ (c'est-à-dire $A_Z \in \Delta$). Maintenant d'après L'application P, déjà définit, on peut faire correspondre à z, la probabilité de A_Z.

$$z \longmapsto P(A_z) = P(e \in A_z) = P(z(e) < z) = F(z)$$

La fonction F est appelée fonction de répartition de la variable aléatoire Z.

L'ensemble des valeurs possibles de Z s'appelle ensemble de définition de Z, Il est noté Z (Ω) ou encore $D_{(x_1, x_2)}$ ou encore D_z

La projection de $D_{(x_1, x_2)}$ sur l'axe $\overline{ox_1}$ donne l'ensemble de définition de la distribution marginale de la variable X_1. la projection de $D_{(x_1, x_2)}$ sur l'axe

$\overrightarrow{ox_2}$ donne l'ensemble de définition de la distribution marginale de la variable X_2.

Propriété de la fonction de répartition F :

a) F(z) appartient à l'intervalle $[0, 1]$: $\forall\, z \in IR^2 \;; 0 \leq F(z) \leq 1$

b) $\text{si } z \to +\infty \text{ alors } F(z) \to 1$

$z \to +\infty \Leftrightarrow x_1 \to +\infty \text{ et } x_2 \to +\infty$

c) Si au moins l'une des coordonnées de z tend vers - ∞ $\text{alors } F(z) \to 0$:

Si $\exists\, x_i \text{ au moins tel que } x_i \to -\infty \text{ alors } F(z) \to o$

Remarque: au moins ici veut dire l'un des cas suivants :

Seulement $x_1 \to -\infty$ ou seulement $x_2 \to -\infty$

Ou encore $(x_1 \to -\infty \text{ et } x_2 \to -\infty)$

d) F est totalement non décroissante :

$\forall\, (x_1, x_2) \text{ et } \forall\, a_1 \geq 0 \;\forall\, a_2 \geq 0$:

$F(x_1 + a_1, x_2) - F(x_1, x_2) \geq 0$

$F(x_1, x_2 + a_2) - F(x_1, x_2) \geq 0$

$F(x_1 + a_1, x_2 + a_2) - F(x_1 + a_1, x_2) - F(x_1, x_2 + a_2) + F(x_1, x_2) \geq 0$

e) F est continue à gauche :

$$\text{Si } \varepsilon \to o \text{ par valeurs positives alors } F(z - \varepsilon) \to F(z)$$

Remarque :
ici $\varepsilon \to 0^+$veut dire que chaque composante $\varepsilon_i \to 0^+$cad que ε est un vecteur

dont les composantes tendent vers 0^+

Section 2 : variable aléatoire, à deux démenions discrète :

2.1 Tableau de contingence :
Soit Z une variable aléatoire discrète a deux dimensions. Appelons les composantes de Z par X et Y : Z=(X, Y).

L'ensemble des valeurs possibles du couple (X, Y)e, c'est –à-dire $D_{(X,Y)}$ est un ensemble fini on infini dénombrable de points du plan .En supposant qu'on peut ordonner à la fois les valeurs possibles x_i de X et y_j de Y, la loi de probabilité de Z est définie par :

$D_{(X,Y)} = \{z_i = (x_i, y_i) \ / \ F(z) = P(Z < z) = P(e \in Az\}$.

et $P_{ij} = P(X_i = x_i \text{ et } Y_j = y_j) \ \forall (x_i, y_j) \in D^2_{(x,y)}$

$$\sum\sum_{P_{ij}} = \sum_i \sum_j P(X = x_i \text{ et } Y = y_j) = 1$$

Le tableau de contingence est un tableau à double entrée qui montre les probabilités P_{ij} des couples de $D_{(x,\, y)}$, puis les deux distributions marginales de X et de Y.

Tableau de contingence :

y \ x	y_1	…	Y_j	…	Loi marginale de X
x_1	P_{11}	…	P_{1j}	…	$P_{1.}$
. . .	. . .				. . .
x_i	P_{i1}	…	P_{ij}	…	$P_{i.}$
. . .	. . .		. . .		. . .
Loi marginale de Y	$P_{.1}$	…	$P_{.j}$	…	1

2.2 Distribution marginales :

Distribution marginale de X :

Distribution marginale de X est donnée par :

$$\forall\ x_i \in D_x\ ;\ P(\ X = x_i) = \sum_j P\big(X = x_i \text{ et } Y = \ y_j\big) = \sum_j P_{ij} = P_{i.}$$

Distribution marginale de Y :

Distribution marginale de Y est donnée par :

$$\forall\ y_j \in D_y\ ; P\big(\ Y = y_j\big) = \sum_i P\big(X = x_i \text{ et } Y = y_j\big) = \sum_i P_{ij} = P_{.j}$$

Les ensembles D_x et D_y sont déterminés par projection sur l'axe des abscisses et des ordonnées respectivement de $D_{(X,\ Y)}$.

Les $P_{i.}$, $P_{.j}$ et les P_{ij} doivent vérifier les conditions générales d'une loi de probabilité à savoir :

$$P(D_X) = \sum_i P_{i.} = 1$$

$$P(D_Y) = \sum_j p_{.j} = 1$$

$$\text{et } P\left[D_{(X,Y)}\right] = \sum_i \sum_j P_{ij} = 1$$

2.3 Distribution conditionnelles :

Définition 1 : variable conditionnelle Y liée à $X=x_i$

On appelle variable conditionnelle de Y liée par $X=x_i$ la variable discrète à une dimension dont l'ensemble de définition D_{Y/x_i} est donné par les points d'intersection de $D_{(x,y)}$ avec la droite passant par l'abscisse x_i , et, dont les probabilités correspondantes sont données par :

$$P_j^i = P(Y = y_j/X = x_i) = \frac{P(X = x_i \text{ et } Y = y_i)}{P(X = x_i)} = \frac{P_{ij}}{P_{i.}} \;; \forall\, y_j \in D_{Y/x_i}$$

Définition2 : Variable conditionnelle de X liée à $Y=y_j$

On définit d'une manière similaire la loi de probabilité de $X / Y=y_j$ par : son domaine D_{X/y_j} formé par les points d'intersection de $D_{(X,Y)}$ et de l'axe passant par $y= y_j$

$$P_i^j = P(X = x_i/Y = y_j) = \frac{P(X = x_i \text{ et } Y = Y_j)}{P(Y = y_j)} = \frac{P_{ij}}{p_{.j}} \;;\; \forall\, x_i \in D_{X/y_j}$$

Loi conditionnelle de $X/Y = y_j$

Valeurs de $(X/Y = y_j)$	$P(X/y = y_j)$
x_1	$\frac{P_{1j}}{P_{.j}}$
. . .	. . .
x_i	$\frac{P_{ij}}{P_{.j}}$
. . .	. . .
$\sum$	1

Loi conditionnelle de $Y/X = x_i$

Valeurs de $(Y/X = x_i)$	$P(Y/X = x_i)$
y_1	$\frac{P_{i1}}{P_{i.}}$
. . .	. . .
y_j	$\frac{P_{ij}}{P_{i.}}$
. . .	. .
$\sum$	1

Propriété :

$$\forall\, P_{i.} \neq 0\ \forall\, P_{.j} \neq 0 \qquad P_{ij} = P_{i.}\, P_j^i = P_{.j}\, P_i^j$$

2.4 Fonction de répartition du couple (X, Y) :

$$F(x,y) = P(\,X < x \text{ et } Y < y) = \sum_{x_i < x} \sum_{y_j < y} P(X = x_i \text{ et } Y = y_j)$$

2.5 Indépendance et corrélation linéaire :

2.5.1 Indépendance

Définition : On dit que deux variables X et Y, composant un couple, sont indépendantes,

Si on a :

$$\forall\, x_i \in D_x\ \ \forall\, y_j \in D_y \quad P\big(X = x_i \text{ et } Y = y_j\big) = P(X = x_i).P(\,Y = y_j)$$

Ou encore : $\forall i\ \forall j \quad P_{ij} = P_{i.}\, P_{.j}$

Propriété : Soit un couple de variables aléatoires (X,Y). Si X et Y sont indépendantes alors :

a) Les distributions marginales et les distributions conditionnelles sont confondues

$$P_j^i = \frac{P_{ij}}{P_{i.}} = P_{.j} \quad \forall\, y_j \in D_{Y/x_i} \quad \text{et} \quad D_{Y/x_i} = D_Y$$

$$P_i^j = \frac{P_{ij}}{P_{.j}} = P_{i.} \quad \forall\, x_i \in D_{X/y_j} \quad \text{et} \quad D_{X/y_j} = D_X$$

b) La fonction de répartition du couple (X, Y) est le produit des fonctions de répartitions marginales :

Preuve :

$$F(x,y) = \sum_{x_i < x} \sum_{y_j < y} P_{ij} = \sum_{x_i < x} \sum_{y_j < y} P_{i.}\, P_{.j}$$

$$F(x,y) = \sum_{x_i<x} P_{i.} \sum_{y_j<y} P_{.j} = F(x.)\,F(.y)$$

Ou encore : $P(X < x\ et\ Y < y) = P(X < x).P(Y < y)$

$$= F(x).F(y)$$

Remarque importante :

Si X et Y sont indépendantes alors :

- ✓ L'égalité $P_{ij}=P_{i.}P_{.j}$ est une condition nécessaire et suffisante d'indépendance;
- ✓ L'égalité F(x,y)= F(x.) f(.y) est une condition nécessaire et suffisante d'indépendance;
- ✓ L'égalité $D_{(X,Y)} = D_X.D_Y$ est une condition nécessaire mais non suffisante d'indépendance.

2.5.2. Corrélation linéaire :

En pratique, il peut arriver de rencontrer des phénomènes, faisant penser qu'il existe une certaine relation entre deux variables. Par exemple, on peut supposer l'existence d'une liaison entre le budget de la publicité, et, le volume des ventes d'un produit ; ou encore entre le revenu d'un ménage et ces dépenses alimentaires ; ou encore entre le taux de criminalité et la densité de la population dans une région…etc.

Définition :

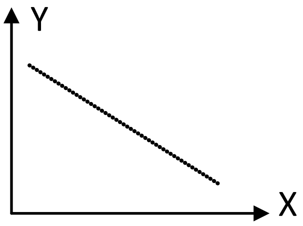
Deux variables X et Y sont dit corrélés, si les variations de ces deux variables se réalisent dans le même sens, ou bien dans le sens contraire. Dans le premier cas, on dit qu'il y a corrélation positive, dans le second, on dit qu'il existe une corrélation négative. La confirmation de l'existence de cette corrélation peut se faire progressivement par :

- ✓ L'observation visuel du nuage des points (x_i, y_j) appartenant à $D_{(X,\ Y)}$;
- ✓ Le calcul du coefficient de corrélation ;

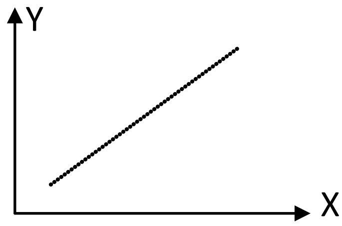
- ✓ Le calcul des tests sur le coefficient de corrélation. Cette dernière méthode fait partie du cours sur la théorie des tests, et, ne fait pas partie du cours sur les probabilités de ce semestre.

Observation du nuage de points :

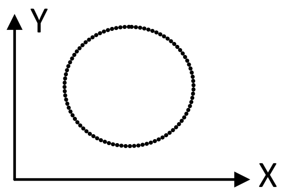
Si nous représentons les couples (x_i, y_j) de $D_{(X, Y)}$, nous obtenons un nuage de points, pouvant prendre ou pas l'une des allures suivantes :



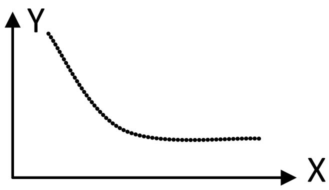
Corrélation négative



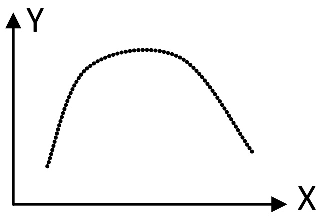
Corrélation positive



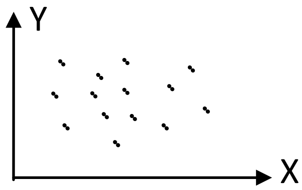
Corrélation circulaire



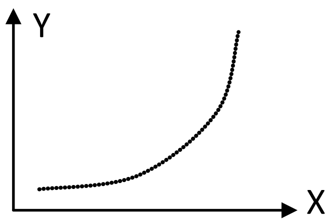
Corrélation hyperbolique



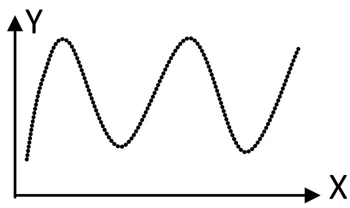
Corrélation parabolique



Absence de corrélation



Corrélation exponentielle



Corrélation sinusoïdale

Mesure de la corrélation linéaire :

Pour mesurer le degré de corrélation linéaire entre deux variable X et Y, on utilise le coefficient de corrélation linéaire ρ définit par :

$$\rho(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sigma(X)\sigma(Y)} = \frac{E(X.Y) - E(X).E(Y)}{\sigma(X)\sigma(Y)}$$

Avec :

$\sigma(X)$ est l'écart type de la variable X

$\sigma(Y)$ est l'écart type de la variable Y

Cov(X,Y) est la covariance entre X et Y

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

Propriété :

a. $-1 \leq \rho \leq 1$

Preuve :

$$\rho^2 - 1 = \frac{[\mathrm{Cov}(X, Y)]^2}{\sigma^2(X)\sigma^2(Y)} - 1$$

D'après l'inégalité de Schwartz on a :

$$\mathrm{Cov}\ (X, Y)^2 \leq \sigma^2(X)\sigma^2(Y)$$

Donc : $\rho^2 - 1 \leq 0 \Rightarrow \rho^2 \leq 1$

D'où $-1 \leq \rho \leq 1$

b-Si X et Y sont indépendantes alors Cov(X, Y)=0 mais l'inverse n'est pas vraie

$$E(XY) = \sum_{x_i} \sum_{y_j} x_i y_j\ P(X = x_i \text{ et } Y = y_j)$$

$$= \sum_{x_i} \sum_{y_j} x_i y_j\ P(X = x_i)\ P\big(Y = y_j\big)$$

(Car X et Y sont indépendantes)

$$E(XY) = \sum_{x_i} x_i\ P(X = x_i) \sum_{y_j} y_j\ P\big(Y = y_j\big) = E(X)\ E(Y)$$

Donc : Cov(X, Y)=0

Section 3 : variable aléatoire absolument continue à deux dimensions :

3.1 Définition générale :

Une variable Z définit une variable aléatoire absolument continue à deux dimensions si la fonction de répartition F du couple aléatoire Z est continue et dérivable à l'ordre deux par rapport à chacune des deux variables X et Y,

c'est-a-dire que :

$$F''_{xy}(x, y) = \frac{\partial^2 F}{\partial x\, \partial y} \text{ existe}$$

Si cette dérivée existe, elle va déterminer la densité de probabilité du couple (X, Y) : $F''_{xy}(x, y) = f(x, y)$ (derivéé seconde croisée)

L'ensemble de définition $D_{(X, Y)}$ est alors un domaine (ou réunion de domaines) du plan IR².

3.2 Probabilité attachée a une droite parallèle à un axe et à un point:

Puisque la fonction de répartition est dérivable donc continue par rapport à x et y. De ce fait la probabilité attachée à toute droite parallèle à un axe de coordonnées est nulle.

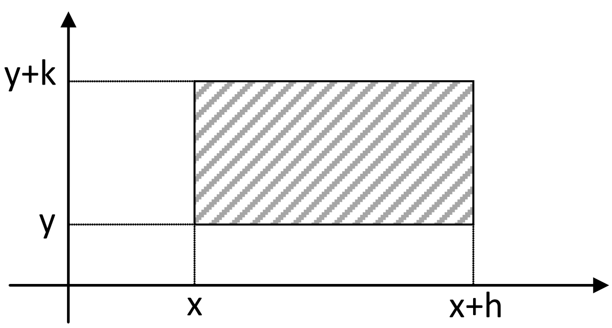
De même, la probabilité attachée à un point est nulle.

3.3 Probabilité attachée à un rectangle :

La probabilité attachée à un rectangle (en faisant abstraction des frontières, puisque toute partie d'une droite porte une probabilité nulle) est obtenue comme différence entre les probabilités attachées à des bandes comme suit :

$$P[x \leq X < x + h \; et \; y \leq Y < y + k]$$

$$= F(x + h, y + k) - F(x + h, y) - F(x, y + k) + F(x, y)$$



3.4 Densité moyenne et densité en un point :

On appelle densité moyenne sur le rectangle tel qu'il a été définit précédemment :

$$\Delta(x, x+h, y+k) = \frac{F(x+h, y+k) - F(x+h, y) - [F(x, y+k) - F(x, y)]}{h.k}$$

Pour trouve la densité moyenne en un point on pose h= dx et k = dy et on fait tendre dx→0 et dy →0.

Autrement dit, la densité moyenne en un point est la valeur limite de la densité moyenne de probabilité à l'intérieur du rectangle de sommet (x, y), lorsqu'on fait tendre les côtes du rectangle dx et dy indépendamment vers zéro.

La densité moyenne à l'intérieur du rectangle de sommet (x, y) s'écrit :

$$\Delta(x, x+dx, y+dy) = \frac{F(x+dx, y+dy) - F(x, y+dy) - F(x+dx, y) + F(x, y)}{dx\ dy}$$

$$\Delta(x, x+dx, y+dy) = \frac{\dfrac{F(x+dx, y+dy) - F(x, y+dy) - F(x+dx, y) + F(x, y)}{dy}}{dx}$$

Comme F est supposée avoir une dérivée seconde par rapport à x et y, alors la valeur limite de cette quantité si on fait tendre : dx et dy vers 0 indépendamment n'est que l'expression de la dérivé seconde F''_{xy}:

$$\lim_{\substack{dx \to 0 \\ dy \to 0}} \Delta(x, x+dx, y+dy) = F''_{xy} = f(x, y)$$

3.5 Probabilité élémentaire :

Il s'ensuit de ce qui a précédé que la probabilité élémentaire attachée à un rectangle infinitésimal d'aire dx dy, entourant le point (x, y) est égal à :

$$\underbrace{f(x, y)}_{\text{Densité moyenne de probabilité au point } (x, y)}\ \underbrace{dx\ dy}_{\text{Aire du rectangle entourant } (x, y)}$$

3.6 Probabilité d'un domaine D :

De ce qui précède, on déduit que la probabilité attachée au domaine D est égal à la somme des probabilités élémentaires entourant le point (x , y), lorsque

(x, y) décrit tout le domaine D :

$$P[(x, y) \in D] = \iint_D f(x, y) dx\, dy$$

Propriétés :

1. Si l'aire de D=0 ⇒ P[(x, y) ∈ D] = 0

Cela résulté du fait que la probabilité attachée à une droite, à un segment ou a une courbe du plan IR^2 est nulle, donc pour les points (x, y) appartenant a ce type de domaines, la densité moyenne et sa limite (de la densité moyenne) sont nulles, donc : $f(x, y) = 0 \Longrightarrow P[(x, y) \in D] = 0$

2. $f(x, y) \geq 0$

   Car f(x, y) est la dérivée seconde croisée d'une fonction totalement non décroissante F.

3. $$\iint_D f(x, y) dx\, dy = 1$$

Preuve :

$$\iint_{IR^2} f(x,y)dx\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dx\,dy$$

$$= F(+\infty,+\infty) - F(-\infty,-\infty) = 1 - 0 = 1$$

Or

$$\iint_{IR^2} f(x,y)dx\,dy = \iint_{D} f(x,y)dx\,dy + \iint_{\overline{D}} f(x,y)dx\,dy = 1$$

$$= \iint_{D} f(x,y)dx\,dy + \iint_{\overline{D}} 0\ dx\,dy = 1$$

$$= \iint_{D} f(x,y)dx\,dy = 1$$

$$\overline{D} = \{(x,y)/(x,y) \notin D\}$$

$$\overline{D} \text{ est le complémentaire de D dans } IR^2$$

Remarque :

D'une manière analogue à la densité de probabilité d'une seule variable, f(x, y) peut être discontinue et même être infinie en certains points du domaine du couple (X, Y).

3.7 Les lois marginales :

On peut définir les distributions marginales à partir de la fonction de répartition.

3.7.1 La distribution marginale de X :

$$P(X < x) = F(x,+\infty) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(x,y)dx\,dy$$

3.7.2 La densité de probabilité marginale de X :

$$f(x.) = F'(x.) = F'(x,+\infty)$$

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

3.7.3 La distribution marginale de Y :

$$P(Y < y) = F(+\infty, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{y} f(x, y)dx\,dy$$

3.7.4 La densité de probabilité marginale de Y :

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

3.8 Les lois conditionnelles :

3.8.1 La loi de X conditionnée par y :

$$f_y(X/Y = y) = \frac{f(x, y)}{f_y(y)} = \frac{f(x, y)}{f(.\,y)}$$

3.8.2 La loi de Y conditionnée par x :

$$f_x(Y/X = x) = \frac{f(x, y)}{f_x(x)} = \frac{f(x, y)}{f(x.\,)}$$

3.9 Indépendance et corrélation linéaire :

Définition : On dit que X et Y sont indépendantes si on a :

$$f(x, y) = f(x.\,)\, f(.\,y)$$

Ou encore :

$$f(X/Y = y) = \frac{f(x, y)}{f_y(y)} = \frac{f(x.\,).\,(f.\,y)}{f(.\,y)} = f(x.\,)$$

$$f(Y/X = x) = \frac{f(x, y)}{f_x(x)} = \frac{f(x.\,)f(.\,y)}{f(x.\,)} = f(.\,y)$$

3.10 Coefficient de corrélation :

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

Si X et Y sont indépendantes alors Cov (x, y) = 0

$$Cov(x, y) = E(X.\,Y) - E(X).\,E(Y)$$

$$E(X.Y) = \int\int_{D_{(X,Y)}} x\, y\, f(x, y) dx\, dy$$

$$E(X.Y) = \int_{D_X} x\, f(x.) dx \int_{D_Y} y\, f(.y) dy = E(X).E(Y)$$

Donc : Cov(X, Y) = 0

Exercices :

On veut former un comité de 3 étudiants. Les candidats à ce comité sont :

4 garçons, 2 filles et 4 étudiants étrangers.

Soit X : « variable aléatoire : nombre de garçons dans le comité »

Y : « variable aléatoire : nombre de filles dans le comité »

a. Calculer la loi du couple (X, Y) ?
b. Calculer les distributions marginales de X et Y ?
c. Calculer les distributions conditionnelles suivantes : $X/y = 2$; $Y/x = 1$?
d. Calculer Cov(x, y) et ρ(x, y) ? Interpréter ?
e. X et Y sont-elles indépendantes ?

Solution :

a. La loi du couple(X, Y) est donnée par le domaine $D_{(x,y)}$ et les probabilités des couples du domaine. Ce que nous pouvons résumer dans le tableau de contingence :

X \ Y	0	1	2	∑
0	1/30	3/30	1/30	5/30
1	6/30	8/30	1/30	15/30
2	6/30	3/30	-	9/30
3	1/30	-	-	1/30
∑	14/30	14/30	2/30	1

b. Distribution marginale de X :

x_i	$P_{i.}$
0	5/30
1	15/30
2	9/30
3	1/30
∑	1

Distribution marginale de Y :

y_j	$P_{.j}$
0	14/30
1	14/30
2	2/30
∑	1

c. Distribution de X/Y=2 :

x_i	$P(X=x_i / Y=2)$
0	1/2
1	1/2
2	0
3	0
∑	1

Distribution de Y/X=1 :

y_j	$P(Y=y_j / X=1)$
0	6/15
1	8/15
2	1/15
∑	1

$$\mathrm{P}(\mathrm{X}=0 \big/ \mathrm{Y}=2)=\frac{\mathrm{P}(\mathrm{X}=0, \mathrm{Y}=2)}{\mathrm{p}(\mathrm{Y}=2)}=\frac{\frac{1}{30}}{\frac{2}{30}}=\frac{1}{2}$$

$$\mathrm{P}(\mathrm{X}=1 \big/ \mathrm{Y}=2)=\frac{\mathrm{P}(\mathrm{X}=1, \mathrm{Y}=2)}{\mathrm{p}(\mathrm{Y}=2)}=\frac{\frac{1}{30}}{\frac{2}{30}}=\frac{1}{2}$$

$$\mathrm{P}(\mathrm{X}=2/\mathrm{Y}=2)=0$$

$$\mathrm{P}(\mathrm{X}=3/\mathrm{Y}=2)=0$$

$$P(Y=0 \big/ X=1)=\frac{\frac{6}{30}}{\frac{15}{30}}=\frac{6}{15}$$

$$P(Y=1/X=1)=\frac{8}{15}$$

$$P(Y = 2/X = 1) = \frac{1}{15}$$

d. Cov(X, Y) = E(X.Y) - E(X) E(Y)

$$\mathrm{E(X,Y)} = \sum\sum \mathrm{x_i y_j P_{ij}} = 0 + \frac{8}{30} + \frac{2 \times 1}{30} + \frac{2 \times 3}{30} = \frac{16}{30}$$

$$\mathrm{E(X)} = \sum \mathrm{x_i\ P_{i.}} = \left(0 \times \frac{5}{30}\right) + \left(1 \times \frac{15}{30}\right) + \left(\frac{2 \times 9}{30}\right) + \left(\frac{3}{30}\right) = \frac{36}{30}$$

$$\mathrm{E(Y)} = \sum \mathrm{y_j\ P_{.j}} = \left(0 \times \frac{14}{30}\right) + \left(1 \times \frac{14}{30}\right) + \left(2 \times \frac{2}{30}\right) = \frac{14 + 4}{30} = \frac{18}{30}$$

$$\mathrm{Cov(X,Y)} = \frac{16}{30} - \frac{36}{30} \times \frac{18}{30} = \frac{(30 \times 16) - 36 \times 18}{900} = \frac{480 - 648}{900} = -\frac{168}{900}$$

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

$$\mathrm{V(X)} = \mathrm{E(X^2)} - (\mathrm{E(X)})^2 = \frac{60}{30} - (\frac{36}{30})^2 = \frac{(60 \times 30) - (36)^2}{30 \times 30} = \frac{1800 - 1296}{900}$$

$$V(X) = \frac{1800 - 1296}{900} = \frac{504}{900} = 0.56 = (0.749)^2$$

$$E(X^2) = \sum x_i^2\ P_{i.} = \left(0 \times \frac{5}{30}\right) + \left(1 \times \frac{15}{30}\right) + \left(4 \times \frac{9}{30}\right) + \left(9 \times \frac{1}{30}\right) = \frac{60}{30}$$

$$\mathrm{V(Y)} = \mathrm{E(Y^2)} - \left(\mathrm{E(Y)}\right)^2 = \frac{22}{30} - (\frac{18}{30})^2 = \frac{(22 \times 30) - 18^2}{900} = \frac{336}{900}$$

$$\mathrm{E(Y^2)} = \sum \mathrm{y_j^2\ P_{.j}} = \left(0 \times \frac{14}{30}\right) + (1 \times \frac{14}{30}) + \left(4 \times \frac{2}{30}\right) = \frac{22}{30}$$

$$V(Y) = \frac{336}{900} = 0.373$$

$$\rho(X,Y) = \frac{-\frac{168}{900}}{\sqrt{\frac{504}{900}}\sqrt{\frac{336}{900}}} = \frac{-\frac{168}{900}}{\frac{\sqrt{504 \times 336}}{900}} = \frac{-168}{\sqrt{504 \times 336}} = -\frac{168}{\sqrt{169344}}$$

$$= -\frac{168}{411.5} \approx -0.41$$

Cov(X,Y)≈ $-0.19 \neq 0$. Donc X et Y ne sont pas indépendantes.

ρ(X, Y)= - 0.41 cela veut dire qu' il existe une corrélation linéaire négative entre X et Y c'est-à-dire, encore, que X et Y varient dans le sens contraire.

Exercice 2 :

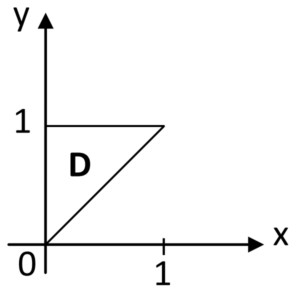
Soit (X, Y) un couple de variables aléatoires absolument continues. Le domaine D de définition du couple est le triangle défini par l'origine le point (x=0, y=1) et le point (x=1, y=1). La densité de probabilité f(x, y) est défini par :

$$\begin{cases} f(x,y) = \dfrac{k}{\sqrt{xy}} & \text{si } (x,y) \in D \\ \quad\quad = 0 & \text{si } (x,y) \notin D \end{cases}$$

a. Déterminer k ?
b. Déterminer les lois marginales et conditionnelles ?

Solution :

a. Détermination de k :



Pour que f(x, y) constitue une densité de probabilité, il faut que :

$$I = \int_0^1 \int_x^1 \frac{k}{\sqrt{xy}} dx\, dy = k \int_0^1 \int_x^1 \frac{1}{\sqrt{xy}}\ dxdy = 1$$

$$= k \int_0^1 \frac{1}{\sqrt{x}} dx[\int_x^1 \frac{1}{\sqrt{y}} dy] = k \int_0^1 \frac{1}{\sqrt{x}} (2 - 2\sqrt{x}) dx = 1$$

$$\text{car: } \int_x^1 \frac{1}{\sqrt{y}} dy = \left[2\sqrt{y}\,\right]_x^1 = 2 - 2\sqrt{x}$$

$$I = 2k\ [\left[2\sqrt{x}\right]_0^1 - [x]_0^1\ ] = 1$$

$$I = 2k\ [2 - 1] = 1\ \Leftrightarrow k = \frac{1}{2}$$

b. Densités marginales :

Densité marginal de x :

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

$$= \frac{1}{2\sqrt{x}}\left[2\sqrt{y}\right]_{x}^{1} = \frac{2-2\sqrt{x}}{2\sqrt{x}} = \frac{1-\sqrt{x}}{\sqrt{x}}$$

$$f(x.) = \begin{cases} \dfrac{1-\sqrt{x}}{\sqrt{x}} & \text{pour } 0 < x < 1 \\ 0 & \text{ailleurs} \end{cases}$$

Densité marginale de Y :

$$f(.y) = \int_{-\infty}^{+\infty} \frac{1}{2\sqrt{xy}}dx = \frac{1}{2\sqrt{y}}\int_{-\infty}^{+\infty} \frac{1}{\sqrt{x}}dx = \frac{1}{2\sqrt{y}}\int_{0}^{y} \frac{1}{\sqrt{x}}dx$$

$$= \frac{1}{2\sqrt{y}}\left[2\sqrt{x}\right]_{0}^{y} = \frac{2\sqrt{y}}{2\sqrt{y}} = 1$$

$$f(.y) = \begin{cases} 1 & \text{pour } 0 < y < 1 \\ 0 & \text{ailleurs} \end{cases}$$

Loi de probabilité conditionnelle :

Loi de X conditionnée par Y = y :

$$f_y(X/Y = y) = \frac{f(x,y)}{f(.y)} = \frac{\frac{1}{2\sqrt{xy}}}{1} \quad \text{pour } 0 < x < y$$

$$= \begin{cases} \dfrac{1}{2\sqrt{xy}} & \text{pour } 0 < x < y \\ 0 & \text{ailleurs} \end{cases}$$

Loi de Y condition par X=x :

$$\mathrm{f_x(Y/X = x)} = \frac{\mathrm{f(x,y)}}{\mathrm{f(x.)}} = \frac{\frac{1}{2\sqrt{\mathrm{xy}}}}{\frac{1-\sqrt{\mathrm{x}}}{\sqrt{\mathrm{x}}}} \quad \text{pour} \quad \mathrm{x} < y < 1$$

$$\mathrm{fx(Y} \Big/ \mathrm{X = x)} = \begin{cases} \frac{1}{2\sqrt{\mathrm{y}}} \frac{1}{\left(1-\sqrt{\mathrm{x}}\right)} & \text{pour} \quad \mathrm{x} < y < 1 \\ 0 & \text{ailleurs} \end{cases}$$

Exercice 1 :

On distribue quatre étudiants inscrits en retard dans trois groupes de TD (A, B, C). On note X le nombre d'étudiants affectés à A et Y le nombre d'étudiants affectes à B.

1) Donner la loi du couple (X,Y).
2) Calculer P(X=1/Y=2).
3) X et Y sont elles indépendantes ?
4) Calculer Cov(X,Y) ; ρ(X, Y).

Exercice 2 :

X et Y représentent respectivement le nombre d'enfants mis au monde et le nombre d'enfants décédés d'une mère tirée au hasard.

X \ Y	0	1	2	3	4	Total
0	0,02	0	0	0	0	0,02
1	0,50	0,10	0	0	0	0,60
2	0.05	0,15	0,05	0	0	0,25
3	0,02	0,03	0,01	0,01	0	0.07
4	0,01	0,01	0,02	0,01	0,01	0,06
Total	0,60	0,29	0,08	0,02	0,01	1,00

1) X et Y sont elles indépendantes.
2) Déterminer Cov(X,Y) et $\rho_{(X,Y)}$
3) Loi de probabilité de (X/Y=0) , (X/Y=1), (Y/X=0), (Y/X=1), E(X/Y=0) et E(Y/X=1).

(NB : $\rho_{(X,Y)}$ désigne le coefficient de corrélation)

SOLUTIONS

Exercice 1 :

1.

X : « nombre d'étudiants affectés à A » ;

Y : « nombre d'étudiants affectés à B ».

$D_X = \{0, 1, 2, 3, 4\}; \quad D_Y = \{0, 1, 2, 3, 4\}$

$$P(X = i, Y = j) = C_4^i \left(\frac{1}{3}\right)^i C_{4-i}^j \left(\frac{1}{3}\right)^j \left(\frac{1}{3}\right)^{4-i-j}$$

X \ Y	0	1	2	3	4	∑
0	1/81	4/81	6/81	4/81	1/81	16/81
1	4/81	12/81	12/81	4/81	0	32/81
2	6/81	12/81	6/81	0	0	24/81
3	4/81	4/81	0	0	0	8/81
4	1/81	0	0	0	0	1/81
∑	16/81	32/81	24/81	8/81	1/81	1

2.

$$P(X = 1/Y = 2) = \frac{P(X = 1, Y = 2)}{P(Y = 2)} = \frac{12}{24} = 0.5$$

3. **1ère façon :**

On remarque que : $\mathrm{P(X = 1/Y = 2) \neq P(X = 1)}$

Donc X et Y ne sont pas indépendants.

2ème façon :

On vérifie qu'il existe des couples $(i, j) \in D_{(X,Y)}$ / $P_{ij} \neq P_{i.} P_{.j}$

Il suffit de donner un contre exemple infirmant l'indépendance

4.

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

$$\mathrm{Cov(X, Y)} = \frac{108}{81} - \left(\frac{108}{81}\right)^2 = -0.44$$

$\rho(X, Y) \cong -0.49$ Donc il existe une corrélation linéaire négative entre les variables X et Y, c'est-à-dire qu'elles varient dans le sens contraire.

Exercice 2 :

Il suffit d'appliquer les formules du cours.

1. Non
2. Cov(X, Y)= E(X.Y) - E(X).E(Y)