Université de Batna 2
L2 SAD S4

Variables aléatoires à plusieurs dimensions

1- Généralités dans le cas unidimensionnel

Soit (Ω, P) un espace probabilisé. Une variable aléatoire (v.a) est toute application $X : \Omega \to \mathbb{R}$ et $\Omega_1 = X(\Omega)$ est appelé l'univers image. Donc si Ω_1 est dénombrable la v.a. X est dite discrète sinon elle sera continue.

Exemple Dans un groupe d'élèves il y a 12 filles et 8 garçons. L'enseignant veut présenter à la fête de fin d'année une pièce de théâtre avec 6 acteurs. Soit X la v.a. égale au nombre de filles dans la pièce.

Dans ce cas $X \in \{0,1,2,3,4,5,6\} = \Omega_1$ et $P(X = k) = p_i = \frac{C_{12}^k \times C_8^{6-k}}{C_{20}^6}, i = 1, \dots, 7$, est la loi de probabilité où bien-sûr $\sum_{i=1}^{7} p_i = 1$.

Les paramètres de la v.a. discrète X sont l'espérance mathématique $E(X) = \sum_{i=1}^{l} x_i . p_i$ (la valeur moyenne de la v.a. X) et la variance $var(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2 \geq 0$, où $E(X^2) = \sum_{i=1}^{l} x_i^2 . p_i$. En plus de l'écart-type $\sigma_X = \sqrt{var(X)}$.

Exercice Revenir à l'exemple précédent et calculer $E(X)$ et $var(X)$.

Propriétés

$$E(X + a) = E(X) + a, a \in \mathbb{R},$$
$$E(\lambda X) = \lambda E(X), \lambda \in \mathbb{R},$$
$$E(X \pm Y) = E(X) \pm E(Y),$$
$$var(X + a) = var(X), a \in \mathbb{R},$$
$$var(\lambda X) = \lambda^2 var(X), \lambda \in \mathbb{R}.$$

Pour le cas où la v.a. X est continue on disposera soit d'une fonction densité de probabilité f ou d'une fonction de répartition F_X. La fonction f est intégrable (continue), positive et en plus $\int_{-\infty}^{+\infty} f(x)dx = 1$. Et la fonction de répartition $F_X(x) = \int_{-\infty}^{x} f(t)dt, \forall x \in \mathbb{R}$.

Exemple Soit la fonction $f(x) = \begin{cases} axe^{-\lambda x}, x \geq 0 \\ 0 \ sinon \end{cases}, a, \lambda > 0$. Vérifier que f est une densité de probabilité. Et donner les valeurs de a, λ pour lesquels $E(X) = 1000$.

Comme propriétés on a $P(a \leq X \leq b) = \int_a^b f(t)dt = F_X(b) - F_X(a)$.

Les paramètres de la v.a. continue X sont alors $E(X) = \int_{-\infty}^{+\infty} t . f(t)dt$ et bien sûr $var(X) = E(X^2) - (E(X))^2$ où $E(X^2) = \int_{-\infty}^{+\infty} t^2 . f(t)dt$. Et donc $\sigma_X = \sqrt{var(X)}$.

2- Couple de v.a.

2.1- Le cas discret

Soient X et Y deux v.a. sur le même espace de probabilité (Ω, P) où $X(\Omega) = \{x_1, x_2, \dots\}$ et $Y(\Omega) = \{y_1, y_2, \dots\}$. La loi du couple (X, Y) est définie par

$$0 \leq p_{ij} = P\left(X = x_i, Y = y_j\right) \leq 1, \qquad \forall i, j \geq 1, \text{où} \sum_{i \geq 1} \sum_{j \geq 1} p_{ij} = 1.$$

Exemple Les centres de transfusion sanguine disposent de la répartition suivante des principaux groupes sanguins

Groupe sanguin	O	A	B	AB
Rhésus +	37%	38,1%	6,2%	2,8%
Rhésus -	7%	7,2%	1,2%	0,5%

1- Quelle est la probabilité pour qu'une personne ait un sang de facteur Rhésus - ? Il s'agit ici de toute la ligne Rhésus – et donc c'est (7+7,2+1,2+0,5)%=15,9%

2- 10 personnes prises au hasard donnent leur sang. Soit X la v.a. qui prend pour valeurs le nombre de personnes appartenant au groupe A. Calculer P(X=4). C'est le cas $X \sim B(n, p)$ où n=10 et p=(38,1+7,2)%=45,3% et donc pour $k = 0, \dots, 10$ $P(X = k) = C_{10}^{k} p^k (1 - p)^{10-k}$.

3- Pour une intervention chirurgicale, on doit avoir au moins trois donneurs de groupe O et facteur Rhésus +. Dix personnes vont faire ce don ignorant leur groupe sanguin. Calculer la probabilité d'avoir au moins les donneurs nécessaires parmi les dix volontaires. Soit Y la v.a. nombre de personnes de groupe O avec Rhésus +, alors $Y \sim B(n, p)$ avec p=37%,n=10 et on veut P(Y≥3)=1-P(Y<3)=1-[P(Y=0)+P(Y=1)+P(Y=2)].

Lois marginales

$$P(X = x_i) = p_{i.} = \sum_{j \geq 1} p_{ij} = p_{i1} + p_{i2} + \cdots, \forall i \geq 1$$

$$P\left(Y = y_j\right) = p_{.j} = \sum_{i \geq 1} p_{ij} = p_{1j} + p_{2j} + \cdots, \forall j \geq 1.$$

Fonction de répartition du couple (X,Y)

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y) = \sum_{x_i \leq x} (\sum_{y_j \leq y} p_{ij}) = \sum_{y_j \leq y} \left(\sum_{x_i \leq x} p_{ij} \right).$$

Indépendance de deux v.a.

Les deux v.a. X et Y sont dites indépendantes si et seulement si les deux événements $(X = x_i)$ et $(Y = y_j)$ sont indépendants, c-à-d

$$P\left(X = x_i et\ Y = y_j\right) = p_{ij} = p_{i.} \times p_{.j}, \forall i, j.$$

Probabilités conditionnelles

$$\forall j \geq 1\ et\ \ i\ fixé\ \ p_{j/i} = P\left(Y = y_j / x = x_i\right) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{p_{i.}},$$

$$\forall i \geq 1\ et\ \ j\ \ fixé\ \ p_{i/j} = P\left(X = x_i / Y = y_j\right) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{.j}}.$$

Caractéristiques du couple (X,Y)

Il faut noter en premier que pour X et Y deux v.a définies sur un même espace probabilisé (Ω, P) on a

La somme X+Y est la v.a. définie sur Ω par $\forall\omega \in \Omega, (\mathrm{X}+\mathrm{Y})(\omega) = \mathrm{X}(\omega) + \mathrm{Y}(\omega)$, l'univers image Z=X+Y est constitué par les réels $z_k = x_i + y_j$ et on a $P(Z = z_k) = \sum p_{ij}$.

Le produit X.Y est la v.a. définie sur Ω par $\forall\omega \in \Omega, (\mathrm{X}.\mathrm{Y})(\omega) = \mathrm{X}(\omega).\mathrm{Y}(\omega)$, l'univers image T=X.Y est constitué par les réels $t_k = x_i.y_j$ et on a $P(T = t_k) = \sum p_{ij}$.

Exemple

X\Y	0	1	3	5	10	Loi marg X
0	1/25	1/25	1/25	1/25	1/25	5/25
1	0	2/25	1/25	1/25	1/25	5/25
3	0	0	3/25	1/25	1/25	5/25
5	0	0	0	4/25	1/25	5/25
10	0	0	0	0	5/25	5/25
Loi marg Y	1/25	3/25	5/25	7/25	9/25	1

Pour Z=X+Y et T=X.Y donner les lois de probabilités de Z et de T.

Z=X+Y et donc Z=0,1,2,3,4,5,6,8,10,11,13,15,20, par exemple si

X=0 et Y=0 alors Z=0 et donc P(Z)=P(X=0,Y=0)=1/25,

X=0 et Y=1 ou X=1 et Y=0 alors Z=1 et donc P(Z=1)=P((X=0,Y=1)∪(X=1,Y=0))=1/25+0=1/25,

X=0 et Y=3 ou X=3 et Y=0 alos Z=3 et P(Z=3)=P((X=0,Y=3)∪(X=3,Y=0)=1/25+0=1/25,

.
.
.

Pour T=X.Y alors T=0,1,3,5,9,10,15,25,30,50,100, par exemple si

X=0 et Y quelconque on aura T=0 et de même X quelconque et Y=0 on aura T=0 et donc P(T=0)=5/25

.
.
.

Remarque

Soit (X,Y) un couple de v.a. indépendantes alors E(X.Y)=E(X).E(Y).

En plus

$$E(X) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xP(X = x, Y = y), \quad E(Y) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} yP(X = x, Y = y).$$

$$E(XY) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X = x, Y = y).$$

Pour $Z = \phi(X,Y)$ où $\phi: D \longrightarrow \mathbb{R}$ pour D un ouvert de $\mathbb{R}^2$ et $X(\Omega) \times Y(\Omega) \subset D$, alors

$$E(Z) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} \varphi(x,y).P(X = x, Y = y).$$

Et $E(X,Y) = (E(X), E(Y))^t$.

La **covariance** et le **coefficient de corrélation** du couple de v.a. (X,Y) sont

$$cov(X,Y) = E(X.Y) - E(X).E(Y).$$

$$\rho_{XY} = \frac{cov(X,Y)}{\sigma_X.\sigma_Y}.$$

Et on a les propriétés suivantes pour X,Y,X' et Y' des v.a.et $\lambda \in \mathbb{R}$

$$cov(X,Y) = cov(Y,X)$$
$$cov(X+X',Y) = cov(X,Y) + cov(X',Y)$$
$$cov(X,Y+Y') = cov(X,Y) + cov(X,Y')$$
$$cov(X,\lambda Y) = \lambda cov(X,Y)$$
$$cov(X,X) = var(X)$$
$$var(X+Y) = var(X) + var(Y) + 2cov(X,Y)$$
$$cov(X,Y) = E\left[\left(X - E(X)\right).\left(Y - E(Y)\right)\right]$$
$$|cov(X,Y)| \leq \sigma_X.\sigma_Y$$
$$-1 \leq \rho_{XY} \leq 1$$

Et si X et Y sont indépendants alors $E(XY) = E(X).E(Y)$, $var(X+Y) = var(X) + var(Y)$, $cov(X,Y) = 0$ et $\rho_{XY} = 0$.

Exemple

Soit la loi de probabilité

k	-2	-1	0	1	2
P(X=k)	1/6	1/4	1/6	1/4	1/6

Pour $Y = X^2$ donner la loi de probabilité du couple (X,Y) est-ce que y a indépendance ? Et calculer $cov(X,Y)$.

k	0	1	4
P(Y=k)	1/6	1/4+1/4=1/2	1/6+1/6=1/3

X\Y	0	1	4
-2	0	0	1/6
-1	0	1/4	0
0	1/6	0	0
1	0	1/4	0
2	0	0	1/6

X et Y ne sont pas indépendants car P(X=-2,Y=0)=0 et P(X=-2).P(Y=0)=1/6.1/6.

On sait que

$$cov(X,Y) = E(X.Y) - E(X).E(Y)$$

Avec la loi de probabilité de X.Y

X.Y	-8	-4	-2	-1	0	1	2	4	8	∑
P(X.Y)	1/6	0	0	1/4	1/6	1/4	0	0	1/6	1
xi.yj.pij	-8/6	0	0	-1/4	0	1/4	0	0	8/6	0

E(X.Y)=0, E(X)=0,E(Y)=11/6 et donc cov(X,Y)=0.

2.2 Le cas continu

On dit que l'application $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$ est une densité de probabilité sur $\mathbb{R}^2$ si elle est positive et vérifie

$$\int_{\mathbb{R}^2} f(x,y)dxdy = 1.$$

En plus la fonction de répartition sera

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(\sigma,\tau)d\sigma d\tau,$$

Et l'on dit que le couple (X,Y) et un couple aléatoire continue.

Exemple

$$f(x,y) = \begin{cases} 2(x+y-2xy) \; si \; (x,y) \in [0,1]\times[0,1] \\ 0 \qquad\qquad\qquad\qquad sinon \end{cases}$$

Prouver que f est une densité de probabilité et donner F_{XY}.

Pour ceci il faut en premier prouver que $\forall x,y \; f(x,y) \geq 0$.

Si $(x,y) \notin [0,1]\times[0,1]$ $f(x,y) = 0$, sinon

On a

$$(x-y)^2 = x^2+y^2-2xy \geq 0 \Rightarrow -2xy \geq -x^2-y^2 \Rightarrow x+y-2xy \geq x-x^2+y-x^2$$

Or sur [0,1] on sait que $x-x^2 = x(1-x) \geq 0$ et pareil pour $y-y^2 = y(1-y) \geq 0$.

Ensuite il faut prouver que $\int_{\mathbb{R}^2} f(x,y)dxdy = 1$.

On a

$$\int_{\mathbb{R}^2} f(x,y)dxdy = \int_{\mathbb{R}}\left[\int_{\mathbb{R}} f(x,y)dx\right]dy = 2\int_0^1\left[\int_0^1(x+y-2xy)dx\right]dy =$$

$$= 2\int_0^1\left[\frac{1}{2}x^2+yx-yx^2\Big|_0^1 dy = 2\int_0^1\left[\frac{1}{2}dy\right] = 1\right.$$

Ainsi f est bien une densité de probabilité.

Les **lois marginales** (densités marginales) sont alors

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

Et donc les fonctions de répartition marginales seront

$$F_X(x) = F_{XY}(x,+\infty), F_Y(y) = F_{XY}(+\infty,y),$$

avec

$$F_X(x) = P(X \leq x) \;\; et \;\; F_Y(y) = P(Y \leq y).$$

On reprendra l'exemple précédent.

Donc les moyennes marginales sont

$$E(X) = \int_{-\infty}^{+\infty} xf_X(x)dx = \int_{-\infty}^{+\infty}[\int_{-\infty}^{+\infty} xf(x,y)dx]dy,$$

$$E(Y) = \int_{-\infty}^{+\infty} yf_Y(y)dy = \int_{-\infty}^{+\infty}[\int_{-\infty}^{+\infty} yf(x,y)dy]d\,x.$$

Avec

$$E(XY) = \int_{-\infty}^{+\infty}[\int_{-\infty}^{+\infty} xyf(x,y)dy]d\,x.$$

$$cov(X,Y) = E(XY) - E(X)E(Y) = E([X-E(X)][Y-E(Y)]).$$

On notera que

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

Si $f(x,y) = f_X(x).f_Y(y)$ on dira alors que X et Y sont indépendantes. De même avec la relation $F_{XY}(x,y) = F_X(x).F_Y(y)$.

Lois conditionnelles

La densité de probabilité conditionnelle de Y sachant $X = x$ est

$$f(y\,/x) = f_{Y\,/\,X}(\,{}^{y}/_{x}) = \begin{cases} \dfrac{f(x,y)}{f(x)} & si\ f(x) > 0 \\ 0 & si\ f(x) = 0 \end{cases},$$

De même la densité de probabilité conditionnelle de X sachant $Y = y$ est

$$f(x\,/y) = f_{X\,/\,Y}(\,x/\,y) = \begin{cases} \dfrac{f(x,y)}{f(y)} & si\ f(y) > 0 \\ 0 & si\ f(x) = 0 \end{cases},$$

Remarque

Si X et Y sont indépendantes on a $f_{X\,/\,Y} = f_X$ et $f_{Y/\,X} = f_Y$.

Changement de variables (X,Y)

Soit le couple de v.a. (X,Y) et considérons le nouveau couple (u,v) où $X = \Phi(U,V)$ et $Y = \Psi(U,V)$ trouvons alors la loi du couple (U,V) :

Si $(X,Y) \in D \subset \mathbb{R}^2$ et $(U,V) \in \Delta \subset \mathbb{R}^2$, alors pour $J = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}$ la matrice jacobienne de la transformation $\Phi: (X,Y) \to (U,V)$ on aura :

$f(x,y)$ se transforme en $f(U,V)|J|$,

$$\iint_D f(x,y)dxdy = \iint_\Delta f(u,v)|J|dudv.$$

Exemple

Considérons $f(x,y) = e^{-x-y}, x > 0\ et\ y > 0.$

La **matrice de covariance** est une matrice carrée symétrique donnée par

$$\Sigma_{XY} = \begin{pmatrix} var(X) & cov(X,Y) \\ cov(X,Y) & var(Y) \end{pmatrix}.$$

3- Vecteurs Gaussiens

Soit X une v.a. continue qui suit la loi normale $N(\mu, \sigma^2)$, alors elle admet la densité suivante

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right),$$

Dans ce cas on note $X \sim N(\mu, \sigma^2)$ où $E(X) = \mu$ et $var(X) = \sigma^2$.

En particulier $Y = \frac{X-\mu}{\sigma} \sim N(0,1)$ avec $f(y) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right)$.

On a si $X_1 \sim N(\mu_1, \sigma_1^2)$ et $X_2 \sim N(\mu_2, \sigma_2^2)$ et si $X_1 \perp\!\!\!\perp X_2$ (indépendants), alors

$$X_1 + X_2 \sim N(\mu_1, \sigma_1^2 + \sigma_2^2).$$

Le vecteur aléatoire $X = (x_1, x_2, \dots, x_n)^T$ est appelé ***vecteur Gaussien*** si et seulement si pour tout $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$ la v.a. $\sum_{i=1}^n \alpha_i X_i$ est gaussienne.

Exemple Prenons $X \sim N(0,4)$ et $Y \sim N(1,9)$ où $X \perp\!\!\!\perp Y$, en plus soit $U = 2X - Y$ et $V = X + \frac{1}{2}Y$.

Quelles sont les lois de U et V ?

Par linéarité on a $E(U) = -1$ et $E(V) = \frac{1}{2}$.

Par indépendance de X et Y on a

$$var(U) = var(2X - Y) = var(2X) + var(Y) = 4var(X) + var(Y) = 25$$

$$var(V) = var\left(X + \frac{1}{2}Y\right) = var(X) + var\left(\frac{1}{2}Y\right) = var(X) + \frac{1}{4}var(Y) = \frac{25}{4}$$

Ainsi $U \sim N(-1,25)$ et $V \sim N(\frac{1}{2}, \frac{25}{4})$.

Il faut démontrer en plus que le vecteur (X,Y) est Gaussien.

Quelle est la loi d'un vecteur Gaussien ?

Pour $X = (X_1, X_2, \dots, X_n)^T$ v.a. gaussien on a $\mu_X = E(X) = (E(X_1), E(X_2), \dots, E(X_n))^T$ et $var(X) = \Sigma_X$ la matrice de covariance et on notera $X \sim V(\mu_X, \Sigma_X)$. En plus pour tout *i* les v.a. sont toutes gaussiennes.

Remarques

1- Si pour tout $i = 1, \dots, n$ les composantes X_i des v.a. gaussiennes sont mutuellement indépendantes, alors le vecteur X est Gaussien.

2- Si pour tout $i = 1, \dots, n$ les composantes X_i des v.a. gaussiennes, mais pas indépendantes alors il se pourra que X ne soit pas Gaussien.

Exemple

Soit $X = (X_1, X_2, X_3)^T$ un vecteur Gaussien de vecteur moyenne $\mu_X = (3, -5, 10)^T$ et matrice de covariance $\Sigma_X = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et posons $Y = 3X_1 + 2X_2 - X_3$ donner la loi de Y.

Proposition

Soit $\mu_X \in \mathbb{R}^n$ et Σ_X une matrice carrée symétrique et définie positive, alors il existe un vecteur Gaussien X de moyenne μ_X et de matrice de covariance Σ_X.

Exemple

Soit $X = (X_1, X_2, X_3)^T$ un vecteur de vecteur moyenne $\mu_X = (1, 0, -1)^T$ et matrice de covariance $\Sigma_X = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. Montrer que X est un vecteur Gaussien.

Remarque
Si $X = (X_1, \dots, X_n)^T$ où $X_i \sim N(0,1)$, $\forall i = 1, \dots, n$, on aura $\mu_X = (0, \dots, 0)^T$ et $\Sigma_X = I_n$.

Proposition
Si $X \sim N(\mu_X, \Sigma_X)$, alors pour tout $\lambda \in \mathbb{R}^n$ la v.a. $\lambda^T X \sim N(\lambda^T \mu_X, \lambda^T \Sigma_X \lambda)$.

Théorème
Une v.a. $X \sim N(\mu_X, \Sigma_X)$ si et seulement si Σ_X est inversible. En ce cas

$$f_X(x_1, \dots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{det \Sigma_X}} . exp\left(-\frac{1}{2}(X - \mu_X)^T \Sigma_X^{-1} (X - \mu_X)\right)$$

avec $X = (X_1, \dots, X_n)^T$ et $x = (x_1, \dots, x)^T$.

Exemple

$X = (X_1, X_2, X_3)^T$, $X \sim N(\mu_X, \Sigma_X)$ où $\mu_X = (3, -5, 10)^T$ et $\Sigma_X = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Considérons la v.a. $Y = 3X_1 + 2X_2 - X_3$ et trouvons la loi de Y.