Probabilité & Statistique

Chapitre 3: Variables aléatoires simultanées

Mohamed Essaied Hamrita

Définition (Fonction de répartition conjointe)

On définit la fonction F de répartition ***simultanée****, ou* ***conjointe****, pour toute paire de variables aléatoires X et Y :*

$$F(x, y) = \mathbb{P}(X \leq x, Y \leq y), \qquad -\infty < x, y < +\infty$$

Définition (Fonction de répartition conjointe)

*On définit la fonction F de répartition **simultanée**, ou **conjointe**, pour toute paire de variables aléatoires X et Y:*

$$F(x,y) = \mathbb{P}(X \leq x, Y \leq y), \qquad -\infty < x, y < +\infty$$

Définition (Fonction de répartition marginale)

Les fonction de répartition de X et de Y peuvent être déduites de la fonction de répartition conjointe de X et Y comme suit:

$$F_X(x) = \mathbb{P}(X \leq x) = \lim_{y \longrightarrow \infty} F(x,y); \;\; F_Y(y) = \mathbb{P}(X \leq x) = \lim_{x \longrightarrow \infty} F(x,y)$$

*Les fonctions $F_X(x)$ et $F_Y(y)$ sont appelées fonctions de répartition **marginales** en X et en Y respectivement.*

$$
\begin{aligned}
F_X(x) &= \mathbb{P}(X \leq x) = \mathbb{P}(X \leq x, Y \leq \infty) \\
&= \mathbb{P}\left(\lim_{y \longrightarrow \infty} (X \leq x, Y \leq y)\right) \\
&= \lim_{y \longrightarrow \infty} \mathbb{P}(X \leq x, Y \leq y) \\
&= \lim_{y \longrightarrow \infty} F(x, y) = F(x, \infty).
\end{aligned}
$$

$$
\begin{aligned}
F_X(x) = \mathbb{P}(X \leq x) &= \mathbb{P}(X \leq x, Y \leq \infty) \\
&= \mathbb{P}\left(\lim_{y \longrightarrow \infty} (X \leq x, Y \leq y)\right) \\
&= \lim_{y \longrightarrow \infty} \mathbb{P}(X \leq x, Y \leq y) \\
&= \lim_{y \longrightarrow \infty} F(x, y) = F(x, \infty).
\end{aligned}
$$

Soit F une fonction de répartition conjointe du couple aléatoire X et Y. Déterminer les probabilités suivantes:

- $\mathbb{P}(X > a, Y > b)$, a, b deux réels
- $\mathbb{P}(a_1 \leq X \leq a_2, b_1 \leq Y \leq b_2)$, $a_1 < a_2$ et $b_1 < b_2$.

$$
\begin{aligned}
\mathbb{P}(X > a, Y > b) &= 1 - \mathbb{P}(\overline{X > a, Y > b}) = 1 - \mathbb{P}(\{\overline{X > a}\} \cup \{\overline{Y > b}\}) \\
&= 1 - \mathbb{P}(\{X \leq a\} \cup \{Y \leq b\}) \\
&= 1 - [\mathbb{P}(X \leq a) + \mathbb{P}(Y \leq b) - \mathbb{P}(X \leq x, Y \leq y)] \\
&= 1 - F_X(a) - F_Y(b) + F(a, b)
\end{aligned}
$$

$$\begin{aligned}\mathbb{P}(X > a, Y > b) &= 1 - \mathbb{P}(\overline{X > a, Y > b}) = 1 - \mathbb{P}(\{\overline{X > a}\} \cup \{\overline{Y > b}\}) \\ &= 1 - \mathbb{P}(\{X \leq a\} \cup \{Y \leq b\}) \\ &= 1 - [\mathbb{P}(X \leq a) + \mathbb{P}(Y \leq b) - \mathbb{P}(X \leq x, Y \leq y)] \\ &= 1 - F_X(a) - F_Y(b) + F(a, b)\end{aligned}$$

$$\begin{aligned}&\mathbb{P}\left(a_1 \leq X \leq a_2, b_1 \leq Y \leq b_2\right) \\ &\qquad = F(a_2, b_2) + F(a_1, b1) - F(a_1, b_2) - F(a_2, b_1)\end{aligned}$$

Définition (Densité conjointe discrète)

La densité de probabilité jointe d'un couple aléatoire discrète est définie par:

$$p(x, y) = \mathbb{P}\left(X = x, Y = y\right)$$

Définition (Densité conjointe discrète)

La densité de probabilité jointe d'un couple aléatoire discrète est définie par:

$$p(x, y) = \mathbb{P}(X = x, Y = y)$$

Définition (Densités marginales discrètes)

Les densités marginales d'un couple aléatoire discrète sont données par:

$$p_X(x) = \sum_y p(x, y), \qquad p_Y(y) = \sum_x p(x, y)$$

Soit la densité jointe d'un couple aléatoire discrète définie par:

X/Y	0	1	2
-1	0.05	0.1	0.3
1	0.15	0.3	0.1

1) Déterminer les densités marginales en X et en Y.

2) En déduire $\mathbb{E}(X)$ et $\mathbb{E}(Y)$.

Définition

Les variables X et Y sont dites ***conjointement continues*** *s'il existe une fonction f de deux variables réels ayant pour tout sous-ensemble E du plan la propriété suivante:*

$$\mathbb{P}((X,Y)\in E)=\int\int_{X,Y\in E} f(x,y)dxdy$$

La fonction f est appelée densité ***conjointe*** *ou* ***simultanée*** *de X et Y.*

La fonction de répartition est:

$$F(x,y)=\int_{-\infty}^{x}\int_{\infty}^{y} f(x,y)dxdy$$

il suffit de dériver pour obtenir

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

1) La densité conjointe de X et Y est donnée par

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & \text{si } x > 0,\ y > 0 \\ 0 & \text{sinon} \end{cases}$$

Calculer $\mathbb{P}(X > 1, Y < 1)$ et $\mathbb{P}(X < Y)$.
2) Soit g la densité conjointe définie par

$$g(x,y) = \begin{cases} e^{-(x+y)} & \text{si } x > 0,\ y > 0 \\ 0 & \text{sinon} \end{cases}$$

Déterminer la densité de la variable aléatoire $Z = X/Y$.

Définition (Densités marginales)

*Les densités **marginales** d'un couple aléatoire X et Y sont données par:*

$$f_X(x) = \int_{\mathbb{R}} f(x,y)dy, \qquad f_Y(y) = \int_{\mathbb{R}} f(x,y)dx$$

Définition (Densités marginales)

Les densités ***marginales*** *d'un couple aléatoire X et Y sont données par:*

$$f_X(x) = \int_{\mathbb{R}} f(x,y)dy, \qquad f_Y(y) = \int_{\mathbb{R}} f(x,y)dx$$

Déterminer les densités marginales des variables aléatoires X et Y définies par la densité conjointe g de l'exercice précédent.

On peut définir les éléments de la distribution conjointe de n variables en suivant la même démarche que celle utilisée dans le cas $n = 2$. Par exemple, la fonction de répartition conjointe F de n variables $X_1, X_2, \ldots, X_n$ a pour définition

$$F(x_1, x_2, \ldots, x_n) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n)$$

Définition

*Les variables X et Y sont dites **indépendantes** si $f(x,y) = f_X(x)f_Y(y)$, pour tout x, y.*

*$X_1, X_2, \ldots, X_n$ sont dites **mutuellement indépendantes** si*

$$f(x_1, x_2, \ldots, x_n) = f_{x_1}(x_1)f_{x_2}(x_2)\ldots f_{x_n}(x_n)$$

Définition

Les variables X et Y sont dites ***indépendantes*** *si $f(x, y) = f_X(x)f_Y(y)$, pour tout x, y.*

$X_1, X_2, \ldots, X_n$ sont dites ***mutuellement indépendantes*** *si*
$f(x_1, x_2, \ldots, x_n) = f_{x_1}(x_1)f_{x_2}(x_2)\ldots f_{x_n}(x_n)$

Le couple aléatoire X et Y défini par la densité g de l'exercice 3 sont-elles indépendantes?

Définition

La loi conditionnelle de X sachant Y, notée $X|Y$ est définie par:

$$f_{X|Y}(x|y) = \frac{f(x,y}{f_Y(y)}$$

Définition

La loi conditionnelle de X sachant Y, notée $X|Y$ est définie par:

$$f_{X|Y}(x|y) = \frac{f(x, y}{f_Y(y)}$$

Si X et Y sont **indépendantes**, alors $f_{X|Y}(x|y) = f_X(x)$ et $f_{Y|X}(y|x) = f_Y(y)$.

Définition

La loi conditionnelle de X sachant Y, notée $X|Y$ est définie par:

$$f_{X|Y}(x|y) = \frac{f(x,y}{f_Y(y)}$$

Si X et Y sont **indépendantes**, alors $f_{X|Y}(x|y) = f_X(x)$ et $f_{Y|X}(y|x) = f_Y(y)$.

1) Déterminer la loi de $Y|X = -1$ de la densité définie dans l'exercice 2.

2) Déterminer la loi de $X|Y$ de la densité f définie dans l'exercice 3.

Proposition

Si $X_1, X_2, \ldots, X_n$ sont mutuellement indépendantes alors

- $\mathbb{E}(X_1 X_2 \ldots X_n) = \mathbb{E}(X_1)\mathbb{E}(X_2) \ldots \mathbb{E}(X_n)$

Proposition

Si $X_1, X_2, \ldots, X_n$ *sont mutuellement indépendantes alors*

- $\mathbb{E}(X_1X_2\ldots X_n) = \mathbb{E}(X_1)\mathbb{E}(X_2)\ldots\mathbb{E}(X_n)$
- $\mathbb{E}(\phi_1(X_1)\phi_2(X_2)\ldots\phi_n(X_n)) = \mathbb{E}(\phi_1(X_1))\mathbb{E}(\phi_2(X_2))\ldots\mathbb{E}(\phi_n(X_n))$

Proposition

Si $X_1, X_2, \ldots, X_n$ *sont mutuellement indépendantes alors*

- $\mathbb{E}(X_1 X_2 \ldots X_n) = \mathbb{E}(X_1)\mathbb{E}(X_2)\ldots\mathbb{E}(X_n)$
- $\mathbb{E}(\phi_1(X_1)\phi_2(X_2)\ldots\phi_n(X_n)) = \mathbb{E}(\phi_1(X_1))\mathbb{E}(\phi_2(X_2))\ldots\mathbb{E}(\phi_n(X_n))$

Proposition

Si $X_1, X_2, \ldots, X_n$ *sont mutuellement indépendantes alors*

- $\mathbb{E}(X_1 X_2 \ldots X_n) = \mathbb{E}(X_1)\mathbb{E}(X_2)\ldots\mathbb{E}(X_n)$
- $\mathbb{E}(\phi_1(X_1)\phi_2(X_2)\ldots\phi_n(X_n)) = \mathbb{E}(\phi_1(X_1))\mathbb{E}(\phi_2(X_2))\ldots\mathbb{E}(\phi_n(X_n))$

Définition (Covariance)

La covariance entre X *et* Y *est définie par*

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

Si X *et* Y *sont indépendantes, alors* $Cov(X, Y) = 0$.

Proposition

La covariance vérifie les propriétés suivantes:

- $Cov(X, X) = (X)$

Proposition

La covariance vérifie les propriétés suivantes:

- $Cov(X, X) = (X)$
- $Cov(X, Y) = Cov(Y, X)$

Proposition

La covariance vérifie les propriétés suivantes:

- $Cov(X, X) = (X)$
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(aX + b, Y) = aCov(X, Y)$, $a, b \in \mathbb{R}$.

Proposition

La covariance vérifie les propriétés suivantes:

- $Cov(X, X) = (X)$
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(aX + b, Y) = aCov(X, Y), \ a, b \in \mathbb{R}.$
- $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z).$

Proposition

La covariance vérifie les propriétés suivantes:

- $Cov(X, X) = (X)$
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(aX + b, Y) = aCov(X, Y)$, $a, b \in \mathbb{R}$.
- $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$.
- $(X + Y) = (X) + (Y) + 2Cov(X, Y)$.

Proposition

La covariance vérifie les propriétés suivantes:

- $Cov(X, X) = (X)$
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(aX + b, Y) = aCov(X, Y)$, $a, b \in \mathbb{R}$.
- $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$.
- $(X + Y) = (X) + (Y) + 2Cov(X, Y)$.

Proposition

La covariance vérifie les propriétés suivantes:

- $Cov(X, X) = (X)$
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(aX + b, Y) = aCov(X, Y)$, $a, b \in \mathbb{R}$.
- $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$.
- $(X + Y) = (X) + (Y) + 2Cov(X, Y)$.

Définition (Coefficient de corrélation)

Le coefficient de corrélation entre X et Y est: $\rho = \dfrac{Cov(X, Y)}{\sigma_X \sigma_Y}$.

$\rho \in [-1, 1]$.

Définition (Espérance conditionnelle)

*L'**espérance conditionnelle** de $X|Y=y$ est définie par:*

$$\mathbb{E}\left(X|Y=y\right)=\begin{cases}\displaystyle\sum_{x} x\mathbb{P}\left(X=x|Y=y\right) & \textit{cas discret}\\ \displaystyle\int_{\mathbb{R}} x f_{X|Y}(x|y)dx & \textit{cas continu}\end{cases}$$

Soit la densité jointe d'un couple aléatoire définie par:

$$f\left(x,y\right)=\frac{2}{xy},\quad \text{pour} \quad 1<y<x<e$$

Déterminer $\mathbb{E}\left(Y|X=x\right)$

Par définition $\mathbb{E}\left(Y|X=x\right)=\displaystyle\int_{\mathbb{R}} y f_{Y|X}(x,y)dy$.

Or $f_{Y|X}(x,y)=\dfrac{f\left(x,y\right)}{f_X\left(x\right)}$ avec

$$\begin{aligned} f_X(x) &= \int_1^x \frac{2}{xy}dy = \frac{2}{x}\Big[\ln(y)\Big]_1^x \\ &= \frac{2\ln(x)}{x}, \quad \text{pour } \; 1 < x < e. \end{aligned}$$

D'où $f_{Y|X}(x,y)=\dfrac{2}{xy}\dfrac{x}{\ln(x)}=\dfrac{1}{y\ln(x)}$ pour $1<y<x$, et par conséquent,

$$\mathbb{E}\left(Y|X=x\right)=\int_1^x \frac{y}{y\ln(x)}dy=\frac{1}{\ln(x)}\Big[y\Big]_1^x=\frac{x-1}{\ln(x)}$$

Proposition

- ***Linéarité:*** *Pour toutes constantes a, b et X, Y et Z des variables aléatoires,*

$$\mathbb{E}\left(aY + bZ|X = x\right) = a\mathbb{E}\left(Y|X = x\right) + b\mathbb{E}\left(Z|X = x\right)$$

Proposition

- ***Linéarité:*** *Pour toutes constantes a, b et X, Y et Z des variables aléatoires,*

$$\mathbb{E}\left(aY + bZ|X = x\right) = a\mathbb{E}\left(Y|X = x\right) + b\mathbb{E}\left(Z|X = x\right)$$

- ***Indépendance:*** *Si X et Y sont deux variables aléatoires indépendantes, alors $\mathbb{E}\left(Y|X = x\right) = \mathbb{E}\left(Y\right)$*

Proposition

- ***Linéarité:*** *Pour toutes constantes* a, b *et* X, Y *et* Z *des variables aléatoires,*

$$\mathbb{E}\left(aY + bZ|X = x\right) = a\mathbb{E}\left(Y|X = x\right) + b\mathbb{E}\left(Z|X = x\right)$$

- ***Indépendance:*** *Si* X *et* Y *sont deux variables aléatoires indépendantes, alors* $\mathbb{E}\left(Y|X = x\right) = \mathbb{E}\left(Y\right)$
- $\mathbb{E}\left(g(X)|X = x\right) = \mathbb{E}\left(X\right)$ *où* g *est une transformation.*

Proposition

- ***Linéarité:*** *Pour toutes constantes* a, b *et* X, Y *et* Z *des variables aléatoires,*

$$\mathbb{E}\left(aY + bZ|X = x\right) = a\mathbb{E}\left(Y|X = x\right) + b\mathbb{E}\left(Z|X = x\right)$$

- ***Indépendance:*** *Si* X *et* Y *sont deux variables aléatoires indépendantes, alors* $\mathbb{E}\left(Y|X = x\right) = \mathbb{E}\left(Y\right)$
- $\mathbb{E}\left(g(X)|X = x\right) = \mathbb{E}\left(X\right)$ *où* g *est une transformation.*
- ***Espérance totale:*** $\mathbb{E}\left(Y\right) = \mathbb{E}\Big(\mathbb{E}\left(Y|X = x\right)\Big)$*: La moyenne totale est la moyenne des moyennes.*

Similairement à l'espérance conditionnelle, la variance conditionnelle est une variance prise par rapport à une distribution conditionnelle.

Définition

La variance conditionnelle de Y sachant $X = x$ est

$$\mathbb{V}\left(Y|X=x\right) = \begin{cases} \sum_{y}(y-\mathbb{E}(Y|X=x))\mathbb{P}\left(Y=y|X=x\right) & \textit{cas discret} \\ \int_{\mathbb{R}}(y-\mathbb{E}(Y|X=x))^2 f_{Y|X}(x,y)dy & \textit{cas continu} \end{cases}$$

La variance conditionnelle possède les propriétés suivantes:

Proposition

- $\mathbb{V}\left(Y|X=x\right)=\mathbb{E}\left(Y^2|X=x\right)-\mathbb{E}\left(Y|X=x\right)^2.$

La variance conditionnelle possède les propriétés suivantes:

Proposition

- $\mathbb{V}\left(Y|X=x\right)=\mathbb{E}\left(Y^2|X=x\right)-\mathbb{E}\left(Y|X=x\right)^2.$
- $\mathbb{V}(aY+b|X=x)=a^2\mathbb{V}(Y|X=x).$

La variance conditionnelle possède les propriétés suivantes:

Proposition

- $\mathbb{V}\left(Y|X=x\right) = \mathbb{E}\left(Y^2|X=x\right) - \mathbb{E}\left(Y|X=x\right)^2$.
- $\mathbb{V}(aY+b|X=x) = a^2\mathbb{V}(Y|X=x)$.
- ***Variance totale:*** $\mathbb{V}(Y) = \mathbb{E}\big(\mathbb{V}(Y|X)\big) + \mathbb{V}\big(\mathbb{E}(Y|X)\big)$.

Considérons deux variables aléatoires X et Y conjointement continues de densité $f(X,Y)$. Il arrive qu'on s'intéresse à la densité conjointe de U et V, deux fonctions de X et Y. Supposons que $U = g_1(X,Y)$ et
$V = g_2(X,Y)$.
les variables U et V sont conjointement continues et de densité

$$g(u,v) = |J(x,y)|^{-1} f(g_1^{-1}(u,v), g_2^{-1}(u,v))$$

où $J(x,y)$ est le jacobien définit par:

$$J(x,y) = \begin{vmatrix} \dfrac{\partial g_1}{\partial x} & \dfrac{\partial g_1}{\partial y} \\ \dfrac{\partial g_2}{\partial x} & \dfrac{\partial g_2}{\partial y} \end{vmatrix}$$

Exemple

Soit $f(x,y) = 2\exp(-(x+y)), \quad$ si $0 < x < y$. Déterminer la densité jointe du couple (U, V) tel que $U = Y - X$ et $V = Y$.

Exemple

Soit $f(x,y) = 2\exp(-(x+y)), \quad$ si $0 < x < y$. Déterminer la densité jointe du couple (U,V) tel que $U = Y - X$ et $V = Y$.

On a $g_1(x,y) = x - y$ et $g_2(x,y) = y$, d'où $x = v - u$, $y = v$ et

$$J(x,y) = \begin{vmatrix} -1 & 1 \\ 0 & 1 \end{vmatrix} = -1 \Longrightarrow |J(x,y)| = 1$$

On obtient, donc

$$g(u,v) = f(v-u,v) = 2\exp(-(v-u+v)) = 2\exp(-(2v-u))$$

$$0 < x < y \Longrightarrow \begin{cases} v - u > 0 \\ v - u < v \end{cases} \Longrightarrow \begin{cases} v > u \\ u > 0 \end{cases} \Longrightarrow 0 < u < v$$

Ainsi, $g(u,v) = 2\exp(-(2v-u))$ si $0 < u < v$.