Chapitre(5)

(Couple aléatoire réel et généralisation au vecteur aléatoire réel)

1. Vecteur aléatoire réel :

Définition1 : Soit $(\Omega, \Lambda, \mathrm{P})$ un espace de probabilité et soit un entier $d \geq 1$.
Soit X_1 , X_2 ,..., X_n des v.a.r. définies sur $(\Omega, \Lambda, \mathrm{P})$. On appelle **vecteur aléatoire** de dimension d ,
l'application $\vec{X}$ telle que : $\vec{X} : \Omega \quad \to IR^d$
et $\forall \omega \in \Omega, \quad \omega \mapsto \vec{X}(\omega) = (X_1(\omega), X_2(\omega),..., X_d(\omega))$.
Les v.a.r. X_i sont les coordonnées du vecteur aléatoire $\vec{X}$.

Définition2 : On dit qu'un vecteur aléatoire $\vec{X}$ est discret si chacune de ses coordonnées est discrète.

Définition3 : On dit qu'un vecteur aléatoire $\vec{X}$ est continu si chacune de ses coordonnées est continu.

Remarque1 : Dans la première partie de ce chapitre, on va s'intéresser au **couple aléatoire réel discret (C.A.D)**, du quel on va étendre les notions ainsi définies au **couple aléatoire réelle continu (C.A.C)** (on verra que les formules mathématiques du cas discret seront les mêmes dans le cas continu, en faisant remplacer le signe $\sum$ par le signe $\int$).
Dans la deuxième partie de ce chapitre, on va généraliser du couple aléatoire réel (C.A.R) au vecteur aléatoire réel (V.A.R) quelconque.

2. Couple aléatoire réel :

Remarque2 : Dans le cas discret, on s'intéresse au couple aléatoire discret (X, Y) où X et Y sont deux v.a.r. discrètes définies sur $(\Omega, \Lambda, \mathrm{P})$ avec $X(\Omega) = \{x_1, x_2,..., x_k\}$ et $Y(\Omega) = \{y_1, y_2,..., y_l\}$.
On peut étudier de façon analogue le cas où $X(\Omega)$ et $Y(\Omega)$ sont infinis dénombrables.
On rappelle $\quad X(\Omega) \times Y(\Omega) = \{(x, y) : x \in X(\Omega) \quad et \quad y \in Y(\Omega)\}$.
On notera dans la suite $\quad \{X = x, Y = y\} = \{X = x\} \cap \{Y = y\} = (X^{-1}\{x\}) \cap (Y^{-1}\{y\})$.

2.1.) Loi de probabilité du couple aléatoire :

Définition4 : Soit (X, Y) un couple aléatoire discret. On appelle la loi de probabilité ou loi conjointe de (X, Y) l'application $\mathrm{P}_{X,Y}$ définie par

$$\mathrm{P}_{X,Y} : \Omega \to [0,1]$$

$$\forall x \in (X(\Omega), Y(\Omega)), \qquad \mathrm{P}_{X,Y}(x, y) = \mathrm{P}(X = x, Y = y) = \mathrm{P}(\{X = x\} \cap \{Y = y\})$$
$$= \mathrm{P}\big((X^{-1}\{x\}) \cap (Y^{-1}\{y\})\big)$$

Propriétés1 :

1) $\mathrm{P}_{X,Y}(x, y) \in [0,1]$ (car c'est une probabilité)
2) Puisque $\{(x, y) : x \in X(\Omega) \quad et \quad y \in Y(\Omega)\}$ constitue un S.C. d'évènements de $(X, Y)(\Omega)$.
On a $\displaystyle \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} \mathrm{P}_{X,Y}(x, y) = 1$,
3) A un évènement par rapport à $(X, Y)(\Omega)$ (i.e. $A \subset (X, Y)(\Omega)$).
$\mathrm{P}(A)$ est donnée par la formule : $\displaystyle \mathrm{P}(A) = \sum_{(x,y) \in A} \mathrm{P}_{X,Y}(x, y)$

Exemple1 : Une urne contient 3 boules numérotées 1, 2 et 3. On tire successivement et sans remise deux boules de l'urne. Soit X et Y les numéros obtenus aux premier et second tirages. Donner en la représentant dans un tableau la loi du couple (X, Y) ?

2.2.) Loi de probabilité du couple aléatoire continu :

Définition5 : Soit (X,Y) un couple aléatoire continue. On appelle densité de probabilité du couple (X,Y), la fonction $f_{X,Y}$ définie pour tout pavé $I \times J$ de IR^2 par :

$$\mathrm{P}(X \in I, Y \in J) = \iint_{I \times J} f_{X,Y}(x,y)dxdy$$

Exemple(2) : Soit (X,Y) un couple aléatoire continu de densité

$$f_{X,Y}(x,y) = \begin{cases} 6xy^2 & si \quad (x,y) \in [0,1] \times [0,1] \\ \\ 0 & \sin on \end{cases}$$

Calculer $\mathrm{P}(X < Y)$? (rép : 3/5)

Propriétés2 : Soit (X,Y) un couple aléatoire continu de densité $f_{X,Y}$. Alors on a :

- $f_{X,Y}(x,y) \geq 0$
- $\iint_{IR^2} f_{X,y}(x,y)dxdy = 1$
- Pour tout domaine D de IR^2, $\mathrm{P}((X,Y) \in D) = \iint_{D} f_{X,Y}(x,y)dxdy$

Exemple(3) : Soit (X,Y) un couple aléatoire continu de densité

$$f_{X,Y}(x,y) = \begin{cases} axy^2 & \forall (x,y) \in IR^2 / 0 \leq x \leq y \leq 1 \\ \\ 0 & si \quad non \end{cases}$$

Calculer a ? (rép : $a = 10$)

2.3.) La fonction de répartition d'un couple aléatoire réel :

Définition6 : Soit (X,Y) un couple aléatoire réel. La fonction de répartition conjointe $F_{X,Y}$ du couple (X,Y) est définie par

$$\forall (x,y) \in IR^2, \quad F_{X,Y}(x,y) = \mathrm{P}(X \leq x, Y \leq y) = \begin{cases} \sum\limits_{x_i \leq x} \sum\limits_{y_j \leq y} \mathrm{P}_{X,Y}(x_i, y_j) & si \quad (X,Y) \, est\, discret \\ \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f_{X,Y}(u,v)dvdu & si \quad (X,Y) \, est\, continu \end{cases}$$

Dans le cas discret de la formule, on a noté : x_i les valeurs possibles de X et y_j celles de Y.

Remarque3 : On déduit de la formule précédente pour le cas (X,Y) continu que

$$f_{X,Y}(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$ pour tout couple (x,y) où la dérivée seconde mixte existe.

Propriétés3 : la fonction $F_{X,Y}$ possède des propriétés semblables à celles de la fonction F_X (voir Ch4).

i) $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0$ (évènement impossible),
ii) $F_{X,Y}(+\infty, +\infty) = 1$ (évènement certain),

iii) $F_{X,Y}(x_1,y_1) \le F_{X,Y}(x_2,y_2) \quad si \quad x_1 \le x_2 \quad et \quad y_1 \le y_2$
(car dans ce cas $\{X \le x_1, Y \le y_1\} \subset \{X \le x_2, Y \le y_2\}$)

iv) On montre que $\lim_{\varepsilon \to 0} F_{X,Y}(x+\varepsilon, y) = \lim_{\varepsilon \to 0} F_{X,Y}(x, y+\varepsilon) = F_{X,Y}(x,y)$.

v) On a : $F_{X,Y}(x,\infty) = \mathrm{P}(X \le x, Y < \infty) = \mathrm{P}(X \le x)$ et de même $F_{X,Y}(\infty, y) = P(Y \le y)$, on les notera dans la partie suivante respectivement F_X et F_Y et on les appellera les <u>fonctions de répartition</u> marginales de X et X .

2.4.) <u>Lois de probabilité marginales</u> :

<u>Définition7</u> : Soit (X,Y) un couple aléatoire réel. On appelle la **loi marginale (ou la densité marginale)** de X l'application P_X (ou la fonction f_X) de $X(\Omega)$ (ou IR)dans $[0,1]$ définie par $\forall x \in X(\Omega)$, $\quad \mathrm{P}_X(x) = \mathrm{P}(X = x) = \sum_{y \in Y(\Omega)} \mathrm{P}_{X,Y}(x,y) \quad si\ X\ est\ discrète.$

(ou $\forall x \in IR$, $\quad f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy \quad si\ X\ est\ continue$).

<u>Remarque4</u> : On définie de la même manière la loi P_Y (la fonction de densité f_Y) de Y en inter changeant entre X et Y dans la formule précédente.

<u>Remarque5</u> : on vérifie facilement que les applications de la définition7 sont bien des lois de probabilité.

<u>Exemple4</u>: (suite exemple1), calculer la loi la marginale de X et la loi marginale de Y et représenter les dans le même tableau de l'exemple1.

<u>Exemple5</u> : (suite exemple2), calculer la loi la marginale de X et la loi marginale de Y .

<u>Exercice</u>: (suite exemple 3), même question.

2.5.) <u>La covariance et le coefficient de corrélation linéaire</u> :

Avant de définir la covariance, on énonce un théorème (à admettre) semblable à celui énoncé dans le chapitre4.

<u>Théorème1</u> : Soit (X,Y) un C.A.R et g une fonction à valeurs réelles définie sur $X(\Omega) \times Y(\Omega)$. Alors, l'espérance de $g(X,Y)$ si elle existe est donnée par :

$$E[g(X,Y)] = \begin{cases} \displaystyle\sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} g(x,y)\mathrm{P}_{X,Y}(x,y) & si \quad (X,Y)\, est\ discret. \\ \displaystyle\int_{IR} \int_{IR} g(x,y) f_{X,Y}(x,y)dxdy & si \quad (X,Y)\, est\ continu \end{cases}$$

<u>Remarque6</u> : les mêmes remarques faites au théorème1 énoncé au chapitre4 seront étendues à ce théorème. Parmi ses applications, il permettra de calculer les <u>moments d'ordre(r, s)</u> :

<u>Moments non centrés d'ordre</u> (r,s) <u>du C.A.R</u> (X,Y) :

En appliquant le théorème1 avec $g(X,Y) = X^r Y^s, r,s \in IN^*$, ce moment noté $M_{r,s}(X,Y)$ est donné par la formule suivante :

$$M_{r,s}(X,Y) = E(X^r Y^s) = \begin{cases} \displaystyle\sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} x^r y^s \mathrm{P}_{X,Y}(x,y) & si \quad (X,Y)\, est\ discret. \\ \displaystyle\int_{IR} \int_{IR} x^r y^s f_{X,Y}(x,y)dxdy & si \quad (X,Y)\, est\ continu \end{cases}$$

Moments centrés d'ordre (r,s) **du C.A.R** (X,Y) **:**

De même, en posant $g(X,Y)=(X-E(X))^r(Y-E(Y))^s, r,s\in IN^*$,

ce moment noté $\mu_{r,s}(X,Y)$ est donné par la formule suivante :

$$\mu_{r,s}(X,Y)=\begin{cases}\displaystyle\sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)}(x-E(X))^r(y-E(Y))^s\mathrm{P}_{X,Y}(x,y) & si \quad (X,Y)\, est\ discret.\\ \displaystyle\int_{IR}\int_{IR}(x-E(X))^r(y-E(Y))^s f_{X,Y}(x,y)dxdy & si \quad (X,Y)\, est\ continu\end{cases}$$

Définition8 : Soit (X,Y) un couple aléatoire réel. On appelle **covariance** de (X,Y) notée $Cov(X,Y)$ le moment centré d'ordre (1,1) de (X,Y)

D'où

$$Cov(X,Y)=\mu_{1,1}(X,Y)=E[(X-E(X))(Y-E(Y))]$$

Dans le cas discret on aura : $Cov(X,Y)=\displaystyle\sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)}(x-E(X))(y-E(Y))\mathrm{P}_{X,Y}(x,y)$

Dans le cas continu on aura : $Cov(X,Y)=\displaystyle\iint_{IR^2}(x-E(X))(y-E(Y)).f_{X,Y}(x,y)dxdy$

Propriétés4 : Quelques résultats utiles et très faciles à démontrer concernant la covariance d'un couple aléatoire réel :

- $Cov(X,Y)=Cov(Y,X)$.
- $Cov(X,X)=V(X)$
- Soient X_1 , X_2 , Y trois variables aléatoires réelles sur Ω et α , β deux réels quelconques. On a :

  $$Cov(\alpha X_1+\beta X_2,Y)=\alpha Cov(X_1,Y)+\beta Cov(X_2,Y) .$$

  (On dit que la covariance est une forme bilinéaire symétrique)

- Soient X et Y deux v.a.r. sur Ω ayant chacune un moment d'ordre2.
  On a :

  $$|Cov(X,Y)|\le\sigma(X).\sigma(Y) \qquad \text{(Inégalité de Cauchy-Schwartz)}$$

- $Cov(X,Y)=E(XY)-E(X)E(Y)$ très souvent utiliser en pratique et qui s'écrit aussi

  $$\mu_{1,1}(X,Y)=M_{1,1}(X,Y)-M_1(X).M_1(Y) .$$

  Dans le cas discret, on aura :

  $$Cov(X,Y)=\Big\{\sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)}xy\mathrm{P}_{X,Y}(x,y)\Big\}-\Big\{\Big[\sum_{x\in X(\Omega)}x\mathrm{P}_X(\Omega)\Big]\times\Big[\sum_{y\in Y(\Omega)}y\mathrm{P}_Y(y)\Big]\Big\}$$

  Dans le cas continu, on aura :

  $$Cov(X,Y)=\Big\{\int_{IR}\int_{IR}x.y.f_{X,Y}(x,y)dxdy\Big\}-\Big\{\Big[\int_{IR}xf_X(x)dx\Big]\times\Big[\int_{IR}yf_Y(y)dy\Big]\Big\}$$

Exemple6 : Calculer $Cov(X,Y)$ pour les deux exemples 1 et 2.

Corollaire1 : (du théorème1)
Soit (X,Y) un couple aléatoire continu et $a,b \in IR$. Alors on montre que :

$$E[aX+bY] = a.E(X)+b.E(Y)$$

$$V[aX+bY] = a^2.V(X)+b^2.V(Y)+2.a.b.\mathrm{cov}(X,Y)$$

Généralisation : (à retenir)
Soient $X_1, X_2, ..., X_n$ n variables aléatoires réelles sur Ω ayant toutes les moments d'ordre2. On a :

$$E\left[\sum_1^n a_i X_i\right] = \sum_{i=1}^n a_i E(X_i)$$

$$V\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i^2 V(X_i) + 2\sum_{1\le i<j\le n} a_i a_j Cov(X_i, X_j)$$

Définition9 : On appelle **coefficient de corrélation linéaire** de X et Y le nombre réel, noté $\rho(X,Y)$ défini par :
$$\rho(X,Y)=\frac{Cov(X,Y)}{\sigma(X).\sigma(Y)}$$

Propriétés5 :

- $|\rho(X,Y)| \le 1$ (à partir de la quatrième des propriétés4)
- $|\rho(X,Y)| = 1 \Leftrightarrow \exists (a,b) \in IR^* \times IR$ tel que: $Y = aX + b$ avec probabilité égale à 1.
(Évènement quasi certain).

2.6.) Indépendance :

Soit (X,Y) un couple aléatoire réel.

Définition10 : Les v.a.r. X et Y sont **indépendantes** si et seulement si

$$\forall (x,y) \in IR^2,\quad F_{X,Y}(x,y) = F_X(x)F_Y(y)$$

Ce qui est équivalent à

Dans le cas discret : $\forall (x,y) \in X(\Omega)\times Y(\Omega),\quad P_{X,Y}(x,y) = \mathrm{P}_X(x)\times \mathrm{P}_Y(y)$

Dans le cas continu : $\forall (x,y) \in IR^2,\quad f_{X,Y}(x,y) = f_X(x)\times f_Y(y)$

Remarque7 : De manière générale, X et Y sont **indépendantes** si

$$\mathrm{P}(X \in A, Y \in B) = \mathrm{P}(X \in A).\mathrm{P}(Y \in B)$$

Où A (respectivement B) est un évènement quelconque qui n'implique que X (respectivement Y).

Théorème2 : (Simple à démontrer en considérant la remarque précédente)

Si X et Y sont deux v.a.r. indépendantes, alors $h(X)$ et $g(Y)$ sont aussi des v.a. indépendantes, où h et g deux fonctions définies respectivement sur $X(\Omega)$ et $Y(\Omega)$.

Remarque8 : Pour les variables aléatoires indépendantes, la loi du couple (la loi conjointe) correspond au produit de ses lois marginales.

Théorème3 : Soit (X,Y) un couple aléatoire réel. Soit h et g deux fonctions définies respectivement sur $X(\Omega)$ et $Y(\Omega)$, à valeurs réelles. Si X et Y sont indépendantes, alors on a :

$$E\big(h(X).g(Y)\big) = E\big(h(X)\big).E\big(g(Y)\big)$$

Corollaire2 : (du théorème2)

Soit (X,Y) un couple aléatoire continu. Si X et Y sont indépendants, alors $\mathrm{cov}(X,Y)=0$.

Remarque9 : La réciproque du théorème2 est fausse en général (voir exemple en TD), néanmoins on verra plus tard que si les variables X et Y sont normalement distribuées, on a une équivalence.

Théorème5 : (Propriété très intéressante de la fonction génératrice)

Si X et Y sont deux v.a.r. indépendantes alors pour $t \in IR, \quad G_{X+Y}(t) = G_X(t) \times G_Y(t)$

Généralisation :

- Si $X_1, X_2, ..., X_n$ n v.a. indépendante et si on note $X = \sum_{i=1}^{n} X_i$ alors on a:

$$t \in IR, \quad G_X(t) = \prod_{i=1}^{n} G_{X_i}(t)$$

- Si $X_1, X_2, ..., X_n$ n v.a. indépendantes et identiquement distribuées (**i.i.d**) et si $X = \sum_{i=1}^{n} X_i$ alors on a:

$$t \in IR, \quad G_X(t) = (G_{X_i}(t))^n$$

2.7.) Lois de probabilité conditionnelles :

Définition11 : Soit (X,Y) un couple aléatoire réel. On appelle la **loi conditionnelle de** X **sachant** Y (ou la **densité conditionnelle de** X **sachant** Y) l'application $\mathrm{P}_{X/Y}(x/y)$) de $X(\Omega)$ dans $[0,1]$ (ou la fonction f_X de IR dans $[0,1]$) définie par :

$$pour\ y \in Y(\Omega) \quad tel\,que \quad \mathrm{P}_Y(y) \neq 0,$$

$$\mathrm{P}_{X/Y}(x/y) = \mathrm{P}(X = x/Y = y) = \frac{\mathrm{P}(X = x, Y = y)}{\mathrm{P}(Y = y)} = \frac{\mathrm{P}_{X,Y}(x,y)}{\mathrm{P}_Y(y)} \qquad si\ X\ est\ discrète.$$

(ou $\forall (x,y) \in IR^2, \quad f_Y(y) \neq 0, \qquad f_{X/Y}(x) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)} \qquad si\ X\ est\ continue.$)

Remarque10 : On définie de la même manière la loi $\mathrm{P}_{Y/X}$ (la fonction de densité $f_{X/X}$) en inter--changeant entre X et Y dans la formule précédente.

Remarque11 : on vérifie facilement que les applications de la définition11 sont bien des lois de probabilité.

Exemple7 : (suite exemple1)

Sachant que le chiffre 1 a été tirée au premier tirage, donner la loi conditionnelle de Y .

Exemple8 : (suite exemple2)

Calculer la loi la marginale de X et la loi marginale de Y .

Théorème5 : Soit (X,Y) un couple aléatoire réel. Alors $\forall(x,y)\in X(\Omega)\times Y(\Omega)$, on a les formules des probabilités composées :

- (X,Y) est discret : $\forall(x,y)\in X(\Omega)\times Y(\Omega)$,

$$\mathrm{P}_{X,Y}(x,y)=\begin{cases}\mathrm{P}_X(x).\mathrm{P}_{Y/X}(y/x) & si \quad \mathrm{P}_X(x)\neq 0\\ \mathrm{P}_Y(y).\mathrm{P}_{X/Y}(x/y) & si \quad \mathrm{P}_Y(y)\neq 0\\ 0 & si \quad non\end{cases}$$

- (X,Y) est continu : $\forall(x,y)\in IR^2$

$$f_{X,Y}(x,y)=\begin{cases}f_X(x).f_{Y/X}(y/x) & si \quad f_X(x)>0\\ f_Y(y).f_{X/Y}(x/y) & si \quad f_Y(y)>0\\ 0 & si \quad non\end{cases}$$

Corollaire4: (du théorème5)

Soit (X,Y) un couple aléatoire réel. Si X et Y sont indépendantes, on a :

- (X,Y) est discret : $\forall(x,y)\in X(\Omega)\times Y(\Omega)$,

$$\begin{cases}\mathrm{P}_{X/Y}(x/y)=\mathrm{P}_X(x)\\ \\ \mathrm{P}_{Y/X}(y/x)=\mathrm{P}_Y(y)\end{cases}$$

- (X,Y) est continu : $\forall(x,y)\in IR^2$

$$\begin{cases}f_{X/Y}(x/y)= & f_X(x)\\ f_{Y/X}(y/x)= & f_Y(y)\end{cases}$$

2.8.) Fonction de répartition conditionnelle :

Soit X une variable aléatoire réelle et A un évènement aléatoire.

- **La fonction de répartition conditionnelle** de X sachant A qu'on note $F_X(x/A)$ est définie par :

$$F_X(x/A)=\frac{\mathrm{P}(\{X\leq x\}\cap A)}{\mathrm{P}(A)} \qquad si \quad \mathrm{P}(A)>0$$

- **La fonction de probabilité conditionnelle** (**la fonction de densité conditionnelle** pour le cas continu) de X sachant A est définie par :

Cas discret : $P_X(x/A)=\dfrac{\mathrm{P}(\{X=x\}\cap A)}{\mathrm{P}(A)} \qquad si \quad \mathrm{P}(A)>0$

Cas continu : $f_X(x/A)=\dfrac{d}{dx}F_X(x/A) \qquad si \quad \mathrm{P}(A)>0$,

- Particulièrement, soit (X,Y) est un couple aléatoire réel.
  - Si Y est une v.a.discrète, alors la fonction de répartition de X sachant $Y = y_k$ sera :

$$F_{X/Y}(x/y_k) = \frac{P(X \le x, y = y_k)}{P(Y = y_k)} = \frac{\sum_{x_i \le x} P(x_i, y_k)}{P_Y(y_k)} \qquad si \quad P_Y(y_k) > 0$$

Le résultat se déduit de la définition de $F_X(x/A)$ ci-dessus en posant $A = (Y = y_k)$

Remarque12 : On a utilisé la notation $F_X(x/Y = y_k)$ plutôt que $F_{X/Y}(x/y_k)$, on doit s'y habituer (Les notations diffèrent mais le sens est le même).

  - Si Y est une v.a.continu et X est aussi continu, alors la fonction de répartition de X sachant $Y = y$ sera :

$$F_{X/Y}(x/y) = \frac{\int_{-\infty}^{x} f_{X,Y}(u,y)du}{f_Y(y)} \qquad si \quad f_Y(y) > 0$$

- On termine cette section par un résultat simple à déduire et très souvent utilisé pour vérifier l'indépendance de deux variables aléatoires.
  Soit (X,Y) est un couple aléatoire réel.
  X et Y sont indépendantes si et seulement si $F_{X/Y}(x/y) = F_X(x) \qquad \forall (x,y) \in IR^2$
  (ou $F_{Y/X}(y/x) = F_Y(y) \qquad \forall (x,y) \in IR^2$)

3. Loi de probabilité de la somme :

Soit (X,Y) un couple aléatoire réel et soit la v.a.r $Z = X + Y$.
On cherche la loi P_Z (ou la densité f_Z) si (X,Y) est discret (ou (X,Y) est continu).

A) (X,Y) est discret :

Théorème6 :

$$\forall z \in Z(\Omega) = (X+Y)(\Omega),$$
$$P_Z(z) = P(Z = z) = \sum_{x+y=z} P_{X,Y}(x,y) = \sum_{x \in X(\Omega)} P_{X,Y}(x, z-x) = \sum_{y \in Y(\Omega)} P_{X,Y}(z-y, y)$$

Corollaire5 : (du théorème6)

Si de plus X et Y sont indépendantes, alors on aura :

$$\forall z \in Z(\Omega) = (X+Y)(\Omega),$$
$$P_Z(z) = P(Z = z) = \sum_{x+y=z} P_X(x)P_Y(y) = \sum_{x \in X(\Omega)} P_X(x)P_Y(z-x) = \sum_{y \in Y(\Omega)} P_X(z-y)P_Y(y)$$

Cas particulier :

- Si X et Y sont à valeurs dans IN, on aura :

$$\forall k \in Z(\Omega) = IN, \qquad P_Z(z) = P(Z = k) = \sum_{i+j=k} P_{X,Y}(i,j)$$

- Et si de plus X et Y sont indépendantes, on aura :

$$\forall k \in IN, \qquad P_Z(z) = P(Z = k) = \sum_{i=0}^{k} P_X(i)P_Y(k-i) = \sum_{j=0}^{k} P_X(k-j)P_Y(j)$$

B) (X,Y) est continu :

Théorème7 : Soit (X,Y) un couple aléatoire continu.

$Z = X + Y$ et f_Z sa densité, on a :

$$\forall z \in IR, \qquad f_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(x, z-x)dx = \int_{-\infty}^{+\infty} f_{X,Y}(z-y, y)dy$$

Corollaire6: (du théorème6)

Soit (X,Y) un couple aléatoire continu et $Z = X + Y.$

- Si X et Y sont indépendantes on aura :

$$\forall z \in IR \qquad f_Z(z) = \int_{-\infty}^{+\infty} f_X(x).f_Y(z-x)dx = \int_{-\infty}^{+\infty} f_X(z-y).f_Y(y)dy$$

- Et si de plus X et Y sont positives, on aura :

$$\forall z \in IR, \qquad f_Z(z) = \int_{0}^{z} f_X(x).f_Y(z-x)dx = \int_{0}^{z} f_X(z-y).f_Y(y)dy$$

Remarque13: La loi de $Z = X + Y$ correspond dans le cas indépendant au produit de convolution de la loi de X et la loi de Y .

4. Espérance conditionnelle et variance conditionnelle :

4.1.) l'espérance conditionnelle :

La notion d'espérance mathématique peut être généralisée en considérant l'espérance conditionnelle.
Pour un couple aléatoire réelle (X,Y), on avait défini la loi de probabilité de $X/Y = y$ par :

$$\begin{cases} Cas\,discret: & P_{X/Y}(x/y) = \dfrac{P_{X,Y}(x,y)}{P_Y(y)} & telle\,que & P_Y(y) > 0 \\ Cas\,continu: & f_{X/Y}(x/y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)} & telle\,que & f_Y(y) > 0 \end{cases}$$

Il arrive souvent de s'intéresser à l'espérance conditionnelle de X sous la condition $Y = y$ qui s'écrira donc

$$E(X/Y=y) = \begin{cases} \sum_x x.P_{X/Y}(x/y) & y\,telle\,que\,P_Y(y) > 0 & (Cas\,discret) \\ \int_{-\infty}^{+\infty} x.f_{X/Y}(x/y)dx & y\,telle\,que\,f_Y(y) > 0 & (Cas\,continu) \end{cases}$$

Remarque14 : Les espérances conditionnelles ont toutes les propriétés des espérances ordinaires et ceci est dû au fait que les probabilités conditionnelles satisfassent toutes les propriétés des probabilités ordinaires.

En particulier : $$E(h(X)/Y=y) = \begin{cases} \sum_x h(x).P_{X/Y}(x/y) & y\,telle\,que\,P_Y(y) > 0 & (Cas\,discret) \\ \int_{-\infty}^{+\infty} h(x).f_{X/Y}(x/y)dx & y\,telle\,que\,f_Y(y) > 0 & (Cas\,continu) \end{cases}$$

(où h est une fonction définie sur $X(\Omega)$ à valeurs réelles).

Et $$E\left(\sum_{i=1}^{n} X_i/Y = y\right) = \sum_{i=1}^{n} E(X_i/Y = y) .$$

- **Calcul de l'espérance par conditionnement :**

$E(X/Y=y)$ est une fonction faisant correspondre à un nombre y l'espérance conditionnelle de X sachant $Y=y$. La fonction obtenue en composant les fonctions Y et $E(X/Y=y)$ qu'on note $E(X/Y)$ est elle-même une variable aléatoire. On énonce dans le théorème suivant une propriété très intéressante de l'espérance conditionnelle.

Théorrème8 : $E(X)=E\big(E(X/Y)\big)$

Remarque15 : Ce théorème signifie que :

$$E(X)=\begin{cases}\sum_y E(X/Y=y).\mathrm{P}_Y(y) & si \quad Y\, est\, discrète.\\ \int_{IR} E(X/Y=y).\,f_Y(y)dy & si \quad Y\, est\, continue.\end{cases}$$

Démonstration : On démontre ce théorème dans le cas où les deux variables aléatoires sont discrètes, la même démarche pour le démontrer dans le cas où les variables seront continues.

$$\begin{aligned}\sum_y E(X/Y=y)\mathrm{P}_Y(y) &= \sum_y\sum_x x\mathrm{P}\big(X=x/Y=y\big)\mathrm{P}_Y(y)\\ &=\sum_y\sum_x x\frac{\mathrm{P}_{X,Y}(x,y)}{\mathrm{P}_Y(y)}\mathrm{P}_Y(y)\\ &=\sum_y\sum_x x\mathrm{P}_{X,Y}(x,y) \qquad ,\\ &=\sum_x x\sum_y \mathrm{P}_{X,Y}(x,y)\\ &=\sum_x x\mathrm{P}_X(x)=E(X)\end{aligned}$$

Remarque16 : Le théorème8 est aussi vraie lorsqu'on remplace X par une fonction $h(X)$:

$E(h(X))=E\big(E\big(h(X)/Y\big)\big)$, (même démarche dans la démonstration)

Remarque17 : Le résultat du théorème8 est très utilisé, il permet de calculer facilement une espérance après avoir conditionné la variable par une autre variable appropriée.

4.2.) La variance conditionnelle :

On définit la variance conditionnelle de X sachant X en suivant la démarche dans la définition de l'espérance conditionnelle, on obtient la variable aléatoire :

$$V(X/Y)=E\Big(\big(X-E(X/Y)\big)^2/Y\Big)=E(X^2/Y)-\big(E(X/Y)\big)^2$$

Théorème9 :(Formule de la variance conditionnelle)

$$V(X)=E\big(V(X/Y)\big)+V\big(E(X/Y)\big)$$

Démonstration : d'après la formule développée de la variance, on a :

$$V(X/Y)=E(X^2/Y)-\big(E(X/Y)\big)^2$$

Donc

$$\begin{aligned}E\big(V(X/Y)\big)&=E\big(E(X^2/Y)\big)-E\Big(\big(E(X/Y)\big)^2\Big)\\ &=E(X^2)-E\Big(\big(E(X/Y)\big)^2\Big) \qquad (*)\end{aligned}$$

Or $V\big(E(X/Y)\big)=E\Big(\big(E(X/Y)\big)^2\Big)-\underbrace{\big(E\big(E(X/Y)\big)}_{=E(X)}\big)^2$

D'où $E\left((E(X/Y))^2\right) = V(E(X/Y)) + (E(X))^2$

En remplaçant dans $(*)$, on aura :

$$E(V(X/Y)) = E(X^2) - V(E(X/Y) - (E(X))^2 = V(X) - V(E(X/Y))$$

D'où la formule citée dans le théprème.

4.3.) <u>Prédiction</u> : (ou **Estimation d'une variable aléatoire** ou **Regression)**

On souhaite prédire la valeur d'une v.a. Y en se basant sur ce que l'on sait de la v.a. X .
Soit $g(X)$ le prédicteur de Y (où g est une fonction), c'est-à-dire si X prend la valeur x alors $g(x)$ sera la valeur prédite de Y (on estime la v.a. Y par la fonction $g(X)$ de la v.a. X).

<u>4.3.1)</u> <u>Meilleur prédicteur</u> :

g sera de façon que la v.a. $g(X)$ soit aussi proche de Y selon le critère « minimiser **l'erreur quadratique moyenne** » qu'on définit par : $E\left((Y - g(X))^2\right)$.

<u>Remarque18</u> : Le théorème suivant énonce que le meilleur prédicteur de Y basé sur X selon le critère cité précédemment est $g(X) = E(Y/X)$.

<u>Théorème(10)</u> : Pour toute fonction g , $E\left((Y - g(X))^2\right) \geq E\left((Y - E(Y/X))^2\right)$.

<u>Démonstration</u> :

$$\begin{aligned} E\left((Y - g(X))^2/X\right) &= E\left[\left((Y - E(Y/X)) + (E(Y/X) - g(X))\right)^2/X\right] \\ &= E\left[(Y - E(Y/X))^2/X\right] + E\left[(E(Y/X) - g(X))^2/X\right] + \\ &\quad 2E\left[((Y - E(Y/X))/X)((E(Y/X) - g(X))/X)\right] \end{aligned}$$

$E(Y/X) - g(X)$ ne dépend que de X , donc peut être considérée comme constante, d'où

$$\begin{aligned} E\left[((Y - E(Y/X))/X)((E(Y/X) - g(X))/X)\right] &= (E(Y/X) - g(X))E((Y/X) - E(Y/X)/X) \\ &= (E(Y/X) - g(X))(E(Y/X) - E(Y/X)) = 0 \end{aligned}$$

On déduit $E\left((Y - g(X))^2/X\right) \geq E\left[(Y - E(Y/X))^2/X\right]$ et en prenant l'espérance des deux membres de l'inégalité, on trouve le résultat cité dans le théorème10.

<u>Remarque19</u> : (Cas particuliers)

- Si $g(X) = a$ (fonction constante), on démontre facilement que $a = E(Y)$ et que l'erreur minimale sera $V(Y)$.
- Si $Y = h(X)$ alors $g(X) = E(Y/X) = E(h(X)/X) = h(X)$ et l'erreur minimale sera nulle.
- Si X et Y sont indépendantes alors $g(X) = E(Y/X) = E(Y)$ et l'erreur minimale sera $V(Y)$.

<u>4.3.2)</u> <u>Meilleur prédicteur linéaire</u> :

Lorsque la loi conjointe de X et Y n'est pas connue ou que le calcul de $E(Y/X = x)$ n'est pas évident mais on connait leurs espérances, variances et leur corrélation, on peut trouver le **meilleur prédicteur <u>linéaire</u>** de Y basé sur X (i.e. $g(X) = aX + b$, a et b deux constantes réelles).
Cela revient à à déterminer a et b telles que $E\left((Y - (aX + b))^2\right)$ qu'note $e(a,b)$ soit minimum.

Or $e(a,b) = E\left(\left(Y-(aX+b)\right)^2\right) = E(Y^2) + a^2E(X^2) + 2abE(X) + b^2 - 2aE(XY) - 2bE(Y)$

et $\begin{cases} \dfrac{\partial}{\partial a} e(a,b) = 2aE(X^2) + 2bE(X) - 2E(XY) \\ \\ \dfrac{\partial}{\partial b} e(a,b) = 2aE(X) + 2b - 2E(Y) \end{cases}$

En identifiant les dérivées partielles à 0, la solution du système d'équations donne :

$$\begin{cases} a = \dfrac{E(XY)-E(X)E(Y)}{E(X^2)-(E(X))^2} = \dfrac{Cov(X,Y)}{\sigma_X^2} = \rho\dfrac{\sigma_Y}{\sigma_X} \qquad qu'on\,notera : \hat{a} \\ \\ b = E(Y) - aE(X) = E(Y) - \rho\dfrac{\sigma_Y}{\sigma_X}E(X) \qquad q'on\,notera : \hat{b} \end{cases}$$

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

On vient de démontrer le théorème suivant :

<u>Théorème11</u> : Le meilleur prédicteur linéaire de Y basé sur X au sens du carré moyen de l'erreur est : $\rho\dfrac{\sigma_Y}{\sigma_X}\left(X - E(X)\right) + E(Y)$.

<u>Remarque20</u> : le carré moyen de l'erreur de ce prédicteur linéaire est :

$$E\left[\left(Y - E(Y) - \rho\frac{\sigma_Y}{\sigma_X}(X - E(X)\right)^2\right] = \sigma_Y^2(1-\rho^2)$$

<u>Remarque21</u> :(Conclusion)

- $g(X) = E(Y/X)$ est en **<u>général</u>** le meilleur estimateur de Y en fonction de X .
- La fonction $\hat{a}X + \hat{b}$ est le meilleur estimateur **<u>linéaire</u>** de Y en fonction de X .
- Il arrive que le meilleur estimateur de Y en fonction de X soit linéaire, i.e.

  $$E(Y/X) = aX + b$$

  (le cas où X et Y sont des variables conjointement « normales »[Ch(6)])

  on dira dans ce cas que le prédicteur linéaire de Y basé sur X est le meilleur prédicteur.

5. <u>Généralités sur les variables aléatoires à plusieurs dimensions</u> :

<u>Introduction</u> : On s'intéressera à la généralisation du couple aléatoire vecteur $n-uplet, n \geq 2$, c'est à dire à un <u>vecteur</u> aléatoire prenant ses valeurs dans IR^n que nous noterons $X = (X_1, X_2, ..., X_n)$.

On remarque que dés le début de ce chapitre, on a considérer que toutes les coordonnées du vecteur sont discrètes ou bien toutes continues (vecteur discret ou continu)et ça sera le cas dans cette partie pour une raison : au deuxième semestre, notre intérêt va se porter sur le cas particulier d'un <u>échantillon</u> qui par définition sera une suite de v.a.r . indépendantes et identiquement distribuées (donc toutes discrètes ou toutes continues). Mais ça n'empêche pas que si par coïncidence on rencontre des vecteur <u>mixtes</u> (certaines de ses coordonnées sont discrètes et d'autres continues) de savoir s'y adapter.

On généralisera les notions de lois marginales et lois conditionnelles au vecteur aléatoire de la même façon.

On peut définir des lois marginales pour tout sous-ensemble de composantes du vecteur X .
On peut définir des lois conditionnelles d'un sous ensemble sachant la valeur d'un autre sous ensemble.
La notion de covariance se généralise en matrice des variances-covariances.
La notion d'indépendance se généralisera à la notion d'indépendance mutuelle (indépendance dans leur ensemble) des n composantes du vecteur.

- <u>Fonction de répartition</u> de $X=(X_1,X_2,...,X_n)$ où X_i , $i=1,...,n$ est une variable aléatoire, est la fonction :
  $\forall(x_1,x_2,...,x_n)\in IR^n$,

$$F(x_1,x_2,...,x_n)=\mathrm{P}\left(\bigcap_{i=1}^{n}(X_i\le x_i)\right)=\begin{cases}\displaystyle\sum_{u_1\le x_1}\sum_{u_2\le x_2}...\sum_{u_n\le x_n}\mathrm{P}_X(u_1,u_2,...u_n). & si\ X\ est\ discret.\\ \displaystyle\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}...\int_{-\infty}^{x_n}f_X(u_1,u_2,...,u_n)du_1du_2...du_n & si\ X\ est\ continu\end{cases}$$

Où on a noté P_X la loi de probabilité du vecteur X s'il est discret et f_X sa densité s'il est continu.
<u>**Remarque22**</u> **:** Dans le cas continu on a :

$$f_X(x_1,x_2,...,x_n)=\frac{\partial^n F(x_1,x_2,...,x_n)}{\partial x_1\partial x_2...\partial x_n}$$

- <u>Les fonctions densités de probabilité marginales</u> de $X_1,\ X_2,...,\ X_n$ sont définies par :
  pour $i=1,2,...,n$

$$\mathrm{P}_{X_i}(x_i)=\sum_{x_1\in X_1(\Omega)}\sum_{x_2\in X_2(\Omega)}...(\widehat{\sum_{x_i\in X_i(\Omega)}})...\sum_{x_n\in X\ (\Omega)}\mathrm{P}_X(x_1,x_2,...,x_i,...x_n)$$ [On somme (n-1) fois]

$$f_{X_i}(x_i)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}.....\int_{-\infty}^{+\infty}f(x_1,x_2,...,x_n)dx_1dx_2...\left(\widehat{dx_i}\right)\ ...dx_n$$ [On somme (intègre) (n-1) fois]

Avec la notation $\widehat{(\ )}$ signifiant que le terme le concernant n'est pas pris en considération.

- $X_1,\ X_2,...,\ X_n$ sont indépendantes si et seulement si

$$\begin{cases}F_X(x_1,x_2,..,x_n)=\prod_{i=1}^{n}F_{X_i}(x_i)\\ \quad ou\\ \begin{cases}\mathrm{P}_X(x_1,x_2,...,x_n)=\prod_{i=1}^{n}\mathrm{P}_{X_i}(x_i) & si\ X\ est\ discret\\ f_X(x_1,x_2,...,x_n)=\prod_{i=1}^{n}f_{X_i}(x_i) & si\ X\ est\ continu\end{cases}\end{cases}$$

- Les covariances des variables aléatoires $X_i,\ X_j$ notées c_{ij} sont définies par :

$$c_{ij}=Cov(X_i,X_j)=E\left[(X_i-E(X_i))(X_j-E(X_j))\right]$$

et on a la matrice <u>**variance-covariance**</u> :

$$\begin{bmatrix} c_{11} & c_{12} & . & . & . & c_{1n} \\ c_{21} & c_{22} & . & . & . & c_{2n} \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ c_{n1} & c_{n2} & . & . & . & c_{nn} \end{bmatrix}$$ tel que $c_{ii} = V(X_i) = E\left[\left(X_i - E(X_i)\right)^2\right], \quad i = 1,2,\ldots,n$

- **Notation matricielle des vecteurs aléatoires :**

On adopte la notation matricielle, on note $X = \begin{pmatrix} X_1 \\ X_2 \\ . \\ . \\ . \\ X_n \end{pmatrix}$ le vecteur aléatoire de dimension n .

On définit l'espérance mathématique de X par le vecteur des espérances (si elles existent) :

$$E(X) = \begin{pmatrix} E(X_1) \\ E(X_2) \\ . \\ . \\ . \\ E(X_n) \end{pmatrix}$$

On note $V(X) = \left[(\mathrm{cov}(X_i, X_j))_{1\le i, j\le n}\right]$ la matrice des variances-covariances de X (si elles existent). Cette matrice est symétrique et ses éléments diagonaux sont des variances.

<u>Propriétés6</u> : Soient A une matrice $(q \times n)$ et c un vecteur $(q \times 1)$. Alors :

1. la relation $Y = AX + c$ définit un vecteur aléatoire Y à valeurs dans IR^q .
2. $\begin{cases} E(Y) = AE(X) + c \\ V(Y) = AV(X)A^t \end{cases}$ ou A^t désigne la matrice transposée de A .

<u>Cas particulier</u> : Si A est le vecteur ligne $(1 \times n)$, $A = (1,1,\ldots,1)$, on retrouve le cas généralisé du corollaire1.

➢ <u>Changement de variables multidimensionnelles</u>:

Le changement de variables multidimensionnelles est un théorème qui devrait être étudié en cours d'analyse mathématique, ici on présentera les étapes de son application lorsque les conditions de son application sont bien justifiées.

A. <u>Cas bidimensionnel</u> :

$X_1,\ X_2$ deux v.a. conjointement <u>continues</u> de densité f_{X_1,X_2}

Supposons que $\begin{cases} Y_1 = & g_1(X_1, X_2) \\ Y_2 = & g_2(X_1, X_2) \end{cases}$

Les fonctions g_1 et g_2 satisfont les deux conditions suivantes :

a). On peut résoudre par rapport à x_1 et x_2 le système d'équations :

$$\begin{cases} y_1 = & g_1(x_1, x_2) \\ y_2 = & g_2(x_1, x_2) \end{cases}$$

Les solutions sont notées :

$$\begin{cases} x_1 = & h_1(y_1, y_2) \\ x_2 = & h_2(y_1, y_2) \end{cases}$$

b). Les fonctions g_1 et g_2 sont continûment différentiables et de Jacobien partout non nul ,

$$J(x_1, x_2) = \begin{vmatrix} \dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2} \\ \\ \dfrac{\partial g_2}{\partial x_1} & \dfrac{\partial g_2}{\partial x_2} \end{vmatrix} \neq 0, \qquad \forall (x_1, x_2)$$

Sous ces conditions, les v.a. Y_1 et Y_2 sont conjointement continues de densité :

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(x_1, x_2).\left|J(x_1, x_2)\right|^{-1}$$

Où il faut remplacer x_i par $h_i(y_1, y_2)$, $\quad (i = 1,2)$

<u>Exercice</u> : $(X_1,\ X_2)$ de densité f_{X_1,X_2}

Posons $Y_1 = X_1 + X_2$ et $Y_2 = X_1 - X_2$

Exprimer la densité conjointe de Y_1 et Y_2 en fontion de f_{X_1,X_2}

B. <u>Cas général</u> :

On a ici n variables $X_1, X_2, ..., X_n$ de densité conjointe continue, on s'intéresse à la densité conjointe de $Y_1, Y_2, ..., Y_n$ tel que :

$$\begin{cases} Y_1 = & g_1(X_1, ..., X_n) \\ Y_2 = & g_2(X_1, ..., X_n) \\ \quad . & \quad . \\ \quad . & \quad . \\ \quad . & \quad . \\ Y_n = & g_n(X_1, ..., X_n) \end{cases}$$

On admettra que ces fonctions g_i ont des dérivées partielles continues et que $J(x_1, ..., x_n) \neq 0$

Pour tout $(x_1, ..., x_n)$,

$$J(x_1,\dots,x_n)=\begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & . & . & . & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & . & . & . & \frac{\partial g_2}{\partial x_n} \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & & & & \frac{\partial g_n}{\partial x_n} \end{vmatrix}$$

On admettra encore que le système d'équations :

$$\begin{cases} y_1 = & g_1(x_1,\dots,x_n) \\ y_2 = & g_2(x_1,\dots,x_n) \\ . & . \\ . & . \\ . & . \\ y_n = & g_n(x_1,\dots,x_n) \end{cases}$$ a pour solution unique notée : $$\begin{cases} x_1 = & h_1(y_1,\dots,y_n) \\ x_2 = & h_2(y_1,\dots,y_n) \\ . & . \\ . & . \\ . & . \\ x_n = & h_n(y_1,\dots,y_n) \end{cases}$$

Sous ces conditions :

$$f_{Y_1,\dots,Y_n}(y_1,\dots,y_n) = f_{X_1,\dots,X_n}(x_1,\dots,x_n).\left|J(x_1,\dots,x_n)\right|^{-1}$$

Où il faut remplacer x_i par $h_i(y_1,\dots,y_n)$, $\quad (i=1,2,\dots,n)$

<u>Exercice</u> : X_1, X_2, X_3 indépendantes de densité f_{X_1,X_2,X_3} .

$$\begin{cases} Y_1 = & X_1+X_2+X_3 \\ Y_2 = & X_1-X_2 \\ Y_3 = & X_1-X_3 \end{cases}$$

- Déterminer é f_{Y_1,Y_2,Y_3} en fonction f_{X_1,X_2,X_3} .

6. <u>Statistiques d'ordre</u> : (Notions préliminaires)

Soient $X_1, X_2,\dots,X_n$ n v.a.r. $i.i.d$ de fonction de répartition commune F .
Pour $(\omega_1,\omega_2,\dots,\omega_n)\in\Omega$ «l'ensemble fondamental associé au vecteur $(X_1,X_2,\dots,X_n)$ »,
on définit les variables aléatoires suivantes :

$X_{(1)}(\omega_1,\omega_2,\dots,\omega_n)$ Vaudra le plus petit des $X_i(\omega_i)$.
$X_{(2)}(\omega_1,\omega_2,\dots,\omega_n)$ Vaudra le second plus petit des $X_i(\omega_i)$.
.
.
.
$X_{(j)}(\omega_1,\omega_2,\dots,\omega_n)$ Vaudra le $j^{ème}$ plus petit des $X_i(\omega_i)$.
.
.
.
$X_{(n)}(\omega_1,\omega_2,\dots,\omega_n)$ Vaudra le plus grand des $X_i(\omega_i)$.

Les fonctions ordonnées $X_{(1)}\le X_{(2)}\le\dots\le X_{(n)}$ sont appelées **<u>statistiques d'ordre</u>** associées à $X_1, X_2,\dots,X_n$.

Remarque23 : Ces variables aléatoires $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ font correspondre à un évènement conjoint les valeurs de $(X_1, X_2,..., X_n)$ classées par ordre croissant.

Pour simplifier, on donne la définition suivante :

Définition12 :

Soit h_k la fonction de IR^n dans IR qui à fait correspondre à $(x_1, x_2,..., x_n)$ la $k^{ème}$ valeur parmi $x_1, x_2,..., x_n$ lorsqu'on les range dans l'ordre croissant. On appelle alors **statistique d'ordre k**, la v.a. notée $X_{(k)}$ défini par : $X_{(k)} = h_k(X_1, X_2,..., X_n)$.

- **Fonction de répartition de $X_{(k)}$:**

On démontre que la fonction de répartition $F_{X_{(k)}}$ de $X_{(k)}$ est : $F_{X_{(k)}}(x) = \sum_{j=k}^{n} C_n^j (F(x))^j (1-F(x))^{n-j}$

Démonstration :

La démonstration est très simple en utilisant une propriété du chapitre6 (lois usuelles) :

«**La somme de v.a.** $i.i.d$ **Bernoulli** $\mathrm{B}(p)$ **est une binomiale** $\mathrm{B}(n, p)$ »

Car $(X_{(k)} \leq x)$ est équivalent qu'au moins k v.a. parmi $X_1, X_2,..., X_n$ sont inférieures à x et le nombre de v.a. parmi $X_1, X_2,..., X_n$ prenant une valeur inférieure ou égale à x sera une v.a. de loi $\mathrm{B}(n, p)$ avec $p = \mathrm{P}(X_i \leq x) = F(x)$ « probabilité de succès (Ch.6).

- **Note :** Pour cette partie de ce chapitre, on a donné la définition des statistiques d'ordre $X_{(k)}, k = 1,2,.., n$ ainsi que leurs fonctions de répartition, on va s'intéresser aux cas particuliers $k = 1$ (la v.a $\min$) et $k = n$ (la v.a $\max$) vu leurs utilisations fréquentes pour résoudre certaines situations rencontrées dans les exercices de probabilités et statistiques.
Les spécialistes en statistiques d'ordre s'intéressent aussi à d'autres notions par nécessité à résoudre des problèmes plus complexes. Parmi ces notions :
la loi conjointe de $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$, ses lois marginales $X_{(k)}$, la loi conjointe de deus statistiques d'ordre $(X_{(i)}, X_{(j)})$, la loi de l'étendu $X_{(j)} - X_{(i)}$, la loi de l'étendu de l'échantillon $X_{(n)} - X_{(1)}$ ...etc.
- **Cas particuliers :**
En se référant à la définition12, on aura pour $k = n$ et $k = 1$ respectivement,

i) Pour $k = n$, on définit donc une v.a. notée $X_{(n)}$ fonction de $X_1, X_2,..., X_n$ par :

$$X_{(n)} = \max\{X_1, X_2,..., X_n\}$$

Fonction de répartition de $X_{(n)}$:

$$(X_{(n)} \leq x) \Leftrightarrow (X_1 \leq x, X_2 \leq x,..., X_n \leq x)$$

D'où, $$\forall x \in IR,\quad F_{X_{(n)}}(x) = \mathrm{P}(X_1 \leq x, X_2 \leq x,..., X_n \leq x) = \underbrace{\prod_{i=1}^{n} \mathrm{P}(X_i \leq x)}_{\text{"indépendance"}} = \underbrace{(F(x))^n}_{\text{même loi}}$$

ii) Pour $k = 1$, on définit donc une v.a. notée $X_{(1)}$ fonction de $X_1, X_2,..., X_n$ par :

$$X_{(n)} = \min\{X_1, X_2,..., X_n\}$$

Fonction de répartition de $X_{(n)}$:

Puisque $\forall x \in IR,\ (X_{(1)} > x) \Leftrightarrow (X_1 > x, X_2 > x,..., X_n > x)$ et $\mathrm{P}(X_{(1)} \leq x) = 1 - \mathrm{P}(X_{(1)} > x)$

On aura : $$\forall x \in IR,\quad F_{X_{(1)}}(x) = 1 - \mathrm{P}(X_{(1)} > x) = 1 - \underbrace{\prod_{i=1}^{n} \mathrm{P}(X_i > x)}_{\text{"indépendance"}} = 1 - \underbrace{(1 - F(x))^n}_{\text{même loi}}$$