PC Variables aléatoires réelles

I. Variables aléatoires réelles (var)

I.1. Définition

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé

définition : Une var est une application $X : \Omega \longrightarrow \mathbb{R}$ vérifiant

Pour tout $x \in X(\Omega)$ $X^{-1}(\{x\}) = \{\omega \in \Omega; X(\omega) = x\} \in \mathcal{A}$

$X^{-1}(\{x\})$ est noté $(X = x)$ ou $\{X = x\}$

alors si I intervalle de $\mathbb{R}$, $X^{-1}(I) = \{\omega \in \Omega; X(\omega) \in I\} \in \mathcal{A}$

$X^{-1}(I)$ est noté $(X \in I)$ ou $\{X \in I\}$

Par exemple $I = [a, b[$ $X^{-1}(I)$ se note $(a \leq X < b)$

définition : X est une var discrète (vard) si $X(\Omega)$ est un ensemble au plus dénombrable.

Remarque : Si $\mathcal{A}=\mathcal{P}(\Omega)$, alors toute application de Ω dans $\mathbb{R}$ est une var.

Exemples : 1. var certaine : $\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ x &\longmapsto \alpha \end{aligned} \qquad \alpha \in \mathbb{R}$

2. var indicatrice : $A \in \mathcal{A}$

$$\begin{aligned} \mathbb{1}_A : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longmapsto \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{ si } \omega \notin A \end{cases} \end{aligned}$$

3. X vard et $u : \mathcal{D}_u \longrightarrow \mathbb{R}$ avec $X(\Omega) \subset \mathcal{D}_u$

alors $\begin{aligned} uoX : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longmapsto u(X) \end{aligned}$ est une vard

I.2. Loi de probabilité d'une vard

définition : X vard définie sur $(\Omega, \mathcal{A}, \mathbb{P})$, alors l'application

$\begin{aligned} \mathbb{P}_X : X(\Omega) &\longrightarrow \mathbb{R} \\ x &\longmapsto \mathbb{P}(X = x) \end{aligned}$ est appelée loi de probabilité de X ou distribution de probabilité de X

Remarques : 1. $\mathbb{P}_X(x) = \mathbb{P}(X = x) \in [0,1]$

2. Comme $\Omega = \bigcup_{i \in I} (X = x_i)$ I fini ou dénombrable et les événements $(X = x_i)$ sont deux à deux incompatibles $\mathbb{P}(\Omega) = 1 = \sum_{i \in I} \mathbb{P}(X = x_i)$

3. $J \subset I$ $\mathbb{P}(X \in J) = \sum_{j \in J} \mathbb{P}(X = x_j)$

théorème : (*admis*) : $X : \Omega \longrightarrow \mathbb{R}$, X prenant ses valeurs dans $\{x_n, n \in I\}$, I fini ou dénombrable, $(p_n)_{n\in I}$ suite de réels positifs avec $\sum_{i\in I} p_n = 1$, alors il existe une probabilité $\mathbb{P}$ définie sur $(\Omega, \mathcal{A})$ telle que

$\forall\ n \in I \quad \mathbb{P}(X = x_n) = p_n$

Remarque : Si X et Y sont deux vard suivant la même loi (donc $X(\Omega) = Y(\Omega)$) , on écrit $X \sim Y$ et si $u : \mathcal{D}_u \longrightarrow \mathbb{R} \quad$ avec $X(\Omega) \subset \mathcal{D}_u$ (donc $Y(\Omega) \subset \mathcal{D}_u$), alors $uoX \sim uoY$

II. Moments d'une vard

II.1. Espérance

définition : X vard définie sur $(\Omega, \mathcal{A},\mathbb{P})$, alors

X admet une espérance (ou est d'espérance finie) si $(x\mathbb{P}(X = x))_{x\in X(\Omega)}$ est une famille sommable.

(ie si $X(\Omega) = \{x_n, n \in I\}$, la série $\sum x_n\mathbb{P}(X = x_n)$ cv absolument)

On note alors dans ce cas $\mathbb{E}(X) = \sum_{x\in X(\Omega)} x\mathbb{P}(X = x)(= \sum_{n\in I} x_n\mathbb{P}(X = x_n))$,

appelée espérance de X (ou moment d'ordre 1 de X)

Remarque : Si $X(\Omega)$ est fini , $(x\mathbb{P}(X = x))_{x\in X(\Omega)}$ est une famille sommable et son espérance existe.

théorème de transfert : (*admis*) X vard, $u : \mathcal{D}_u \longrightarrow \mathbb{R}$ avec $X(\Omega) \subset \mathcal{D}_u$, alors $u(X)$ est une vard et $u(X)$ admet une espérance si et seulement si $(u(x)\mathbb{P}(X = x))_{x\in X(\Omega)}$ est une famille sommable.

Dans ce cas $\mathbb{E}(u(X)) = \sum_{x\in X(\Omega)} u(x)\mathbb{P}(X = x)$

Remarque : donc $\mathbb{E}(u(X))$ se calcule à partir de la loi de X sans déterminer la loi de $u(X)$.

définition : X est centrée si $\mathbb{E}(X) = 0$

Propriété : X, Y vard avec $|X| \leq Y$ et Y admet une espérance, alors X admet une espérance.

Propriétés : X, Y vard admettant des espérances

1. **Linéarité** $\lambda \in \mathbb{R}$, alors $\lambda X + Y$ admet une espérance et $\mathbb{E}(\lambda X + Y) = \lambda\mathbb{E}(X)+\mathbb{E}(Y)$
2. $(a, b) \in \mathbb{R}^2$, alors $aX + b$ admet une espérance et $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$
3. $X-\mathbb{E}(X)$ est centrée.
4. **Positivité** si $X \geq 0 \quad$ alors $\mathbb{E}(X) \geq 0$
et $\mathbb{E}(X = 0)$ alors $\mathbb{P}(X = 0) = 1$ (ie X nulle presque sûrement)
5. **Croissance** si $X \geq Y$ alors $\mathbb{E}(X) \geq \mathbb{E}(Y)$

Propriété : Si X vard positive et d'espérance nulle, alors $(X = 0)$ est un événement presque sûr.

théorème : Si $X(\Omega) \subset \mathbb{N} \quad X$ admet une espérance $\Longleftrightarrow \sum\mathbb{P}(X \geq n)$ cv

Dans ce cas $\mathbb{E}(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geq n)$ $(= \sum_{n=0}^{+\infty} \mathbb{P}(X > n))$

II.2. Variance

théorème de Huyghens-Koenig : X vard, on suppose que X^2 admet une espérance (ie X admet un moment d'ordre 2), alors X est d'espérance finie.

HUYGENS Christiaan 1629 La Haye - 1695 La Haye : Astronome et physicien Huygens fait des études de droit et de mathématiques à Leyde et à Breda. Présenté à Mersenne et à Descartes par son père, il se tourne vers les mathématiques et la recherche scientifique. À l'invitation de Colbert, Huygens s'établit et sera membre de l'Académie des Sciences. Il découvre le premier la nature ondulatoire de la lumière expliquant ainsi les effets de réfraction et de diffraction. Au moyen d'une lunette astronomique de sa fabrication, il découvre en 1656 les anneaux de Saturne pressentis par Galilée et son premier satellite. Huygens définit les notions de force centrifuge et de moment d'inertie, invente le ressort spirale. Il contribue avec Pascal au développement du calcul des probabilités.

KOENIG Johann Samuel 1712 Büdingen - 1757 Zuilenstein : Physicien, philosophe, juriste, ami de Voltaire, il a été élève de Jean Bernoulli et de Leibniz. Il enseigne les mathématiques, la philosophie et le droit à La Haye. Ses recherches portent en mécanique, et en calcul des probabilités. En physique, un repère de Koenig est synonyme de repère barycentrique. En mathématiques, son nom est jumelé à Huygens, pour le calcul de la variance d'une série statistique

définition : X vard, on suppose que X^2 admet une espérance, on définit la variance de X par $V(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$ et l'écart-type de X par $\sigma(X) = \sqrt{V(X)}$

Remarque : $V(X) \geq 0$

Propriété : Si X vard admet un moment d'ordre 2, alors $V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$

théorème : Soit X vard admettant un moment d'ordre 2, alors $\forall\ (a, b) \in \mathbb{R}^2$, $aX + b$ admet une variance et $V(aX + b) = a^2 V(X) \qquad \sigma(aX + b) = |a|\sigma(X)$

définition : X vard est centrée réduite si $\mathbb{E}(X) = 0$ et $V(X) = 1$

Exemple : X vard avec $V(X) \neq 0$, alors $X^* = \dfrac{X - \mathrm{E}(X)}{\sigma(X)}$ est centrée réduite, appelée va centrée réduite associée à X

III. Cas des var discrètes à valeurs dans $\mathbb{N}$

X vard avec $X(\Omega) \subset \mathbb{N}$

théorème : Le rayon de convergence de la série entière $\sum \mathbb{P}(X = n) t^n$ est supérieur ou égal à 1, noté R_X

définition : La série génératrice d'une var à valeurs dans $\mathbb{N}$ est la série $G_X(t) = \mathbb{E}(t^X) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n) t^n$

Remarque : G_X converge normalement sur [-1,1] au moins.

Propriété : Si X et Y sont deux var à valeurs dans $\mathbb{N}$, si $G_X = G_Y$, par unicité du DSE, on a $\forall\, n \in \mathbb{N} \quad \mathbb{P}(X = n) = \mathbb{P}(Y = n)$, ie X et Y ont la même loi de probabilité, donc la loi de probabilité d'une var à valeurs dans $\mathbb{N}$ est caractérisée par sa série génératrice.

théorème : X vard à valeurs dans $\mathbb{N}$

X admet une espérance $\Longleftrightarrow G_X$ est dérivable en 1

Dans ce cas, $\mathbb{E}(X) = G'_X(1)$

théorème : X vard à valeurs dans $\mathbb{N}$

X admet une variance $\Longleftrightarrow G_X$ est deux fois dérivable en 1

Dans ce cas, $V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$

IV. Lois usuelles

IV.1. Va certaine

X va certaine si $X : \Omega \longrightarrow \mathbb{R}$

$\qquad\qquad\qquad\quad \omega \longmapsto a$

$X(\Omega) = \{a\} \qquad \mathbb{P}(X = a) = 1$

$\mathbb{E}(X) = a \qquad V(X) = 0$

IV.2. Loi uniforme

X suit une loi uniforme sur $\{1, \cdots, n\}$, notée $X \hookrightarrow \mathcal{U}(\{1, \cdots, n\}$

si $X(\Omega) = \{1, \cdots, n\} \qquad$ et $\forall\, k \in \{1, \cdots, n\} \quad \mathbb{P}(X = k) = \dfrac{1}{n}$

modèle probabiliste : On choisit au hasard (ie équiprobabilité) un objet parmi n objets numérotés et on appelle X le numéro de l'objet choisi.

$\mathbb{E}(X) = \dfrac{n+1}{2} \qquad V(X) = \dfrac{n^2 - 1}{12}$

La série génératrice de X est définie sur $\mathbb{R}$ $\quad G_X(t) = \begin{cases} \frac{t(1-t^n)}{n(1-t)} \text{ si } t \neq 1 \\ 1 \text{ si } t = 1 \end{cases}$

IV.2. Loi de Bernoulli

BERNOULLI Jakob (Jacques) 1654 Bâle - 1705 Bâle : Les Bernoulli sont issus d'une famille protestante d'Anvers qui se réfugia à Francfort pour échapper aux persécutions catholiques, puis la famille émigre à Bâle. Les Bernoulli constituent une véritable dynastie de mathématiciens dont le premier est Jakob, surnommé Jacques 1er. Son frère Johann (Jean), mathématicien et physicien sera Johann 1er et les fils de ce dernier seront Daniel 1er, Johann II, Nicolas III... Les travaux de Jakob portent sur l'analyse fonctionnelle, le calcul différentiel, le calcul intégral. Il est avec son frère Johann un des grands artisans du développement en série des fonctions entamé par Mercator, Gregory et Leibniz. Précurseur de l'analyse fonctionnelle, il est *le premier à parler de fonction et à utiliser la notation fx qu'utilisera Euler, proche du f(x) d'aujourd'hui qui est dû à Lagrange.*

BERNOULLI Johann (Jean) 1667 Bâle - 1748 Bâle : Frère cadet de Jacques Bernoulli (Jakob), Johann, dit Jean 1er, fait ses études à Bâle. Médecin, physicien, il fut, entre autres, le professeur de Leonhard Euler et un ami de Gabriel Cramer. Admirateur de Leibniz, on doit à Johann Bernoulli d'importants travaux en calcul différentiel et intégral qu'il introduisit en France par l'intermédiaire de Guillaume de L'hospital. Johann obtint une chaire de mathématiques à l'université de Groningue. En sciences physiques, on lui doit la notation g pour désigner l'accélération de la pesanteur. Johann établit la méthode de décomposition des fractions rationnelles en éléments simples. Johann Bernoulli prouva la divergence de la série harmonique.

BERNOULLI Daniel 1700 Groningen - 1782 Bâle : Fils de Jean Bernoulli , Daniel étudie la philosophie et s'initie aux mathématiques par son père et son frère Nicolas II. Après des études médicales à Heidelberg, sa rencontre avec Euler, ami de la famille, sera décisive pour l'orientation de sa carrière : il s'adonne alors aux mathématiques appliquées à la physique. Ses travaux portent sur l'élasticité, l'hydrodynamique, les cordes vibrantes, les tuyaux sonores... Dans les années 1730, Daniel Bernoulli sera aussi avec Buffon, le premier à introduire le calcul intégral dans la branche toute neuve des mathématiques qu'est le calcul des probabilités appliqué à des phénomènes aléatoires continus.

X suit une loi de Bernoulli de paramètre $p \in]0,1[$, notée $X \hookrightarrow \mathcal{B}(p)$

si $X(\Omega) = \{0,1\}$ et $\mathbb{P}(X=1) = p$ $\mathbb{P}(X=0) = 1-p = q$

modèle probabiliste : On a une épreuve à deux possibilités: succès avec probabilité p, échec avec probabilité $1-p$

$\mathbb{E}(X) = p$ $V(X) = p - p^2 = pq$

La série génératrice de X est définie sur $\mathbb{R}$ $G_X(t) = 1 - p + pt$

IV.4 Loi binômiale

X suit une loi binômiale de paramètres (n,p) $p \in]0,1[$, $n \in \mathbb{N}^*$, notée $X \hookrightarrow \mathcal{B}(n,p)$

si $X(\Omega) = \{0, \cdots, n\}$ et $\forall\ k \in \{0, \cdots, n\}$ $\mathbb{P}(X=k) = \binom{n}{k} p^k q^{n-k}$

modèle probabiliste : On répète n fois de manière identique et indépendante la même épreuve de Bernoulli de paramètre p et X désigne le nombre de succès obtenus.

$\mathbb{E}(X) = np$ $V(X) = npq$

La série génératrice de X est définie sur $\mathbb{R}$ $G_X(t) = (1-p+pt)^n$

IV.5. Loi géométrique

X suit une loi géométrique de paramètre p, $p \in]0,1[$, notée $X \hookrightarrow \mathcal{G}(p)$

si $X(\Omega) = \mathbb{N}^*$ et $\forall\ k \in \mathbb{N}^*$ $\mathbb{P}(X=k) = p(1-p)^{k-1} = pq^{k-1}$

On a bien $\displaystyle\sum_{k=1}^{+\infty} \mathbb{P}(X=k) = 1$

modèle probabiliste : On considère une suite d'épreuves identiques et indépendantes de Bernoulli de paramètre p et X désigne le rang du premier succès (ou temps d'attente).

$k \in \mathbb{N}^* \qquad \mathbb{P}(X = k) = \mathbb{P}(\text{échec}, \cdots, \text{échec}, \text{succès}) = q^{k-1}p$

$\mathbb{E}(X) = \dfrac{1}{p} \qquad V(X) = \dfrac{q}{p^2}$

La série génératrice de X vérifie : $R_X = \dfrac{1}{q} \qquad$ et $\forall\ t \in \left]-\dfrac{1}{q}, \dfrac{1}{q}\right[\quad G_X(t) = \dfrac{pt}{1-qt}$

théorème : $X \hookrightarrow \mathcal{G}(p)$, alors $\forall\ n \in \mathbb{N}^* \quad \mathbb{P}(X > n) = (1-p)^n$

théorème : La loi géométrique est une loi sans mémoire, ie

$\forall\ (n,k) \in \mathbb{N}^2 \quad \mathbb{P}_{X>n}(X > n+k) = \mathbb{P}(X > k)$

IV.Loi de Poisson

POISSON Siméon Denis 1781 Pithiviers - 1840 Sceaux : Polytechnicien, élève de Fourier et de Laplace, astronome et physicien, il a été professeur à l'École Polytechnique, professeur de mécanique à la faculté des sciences de Paris, directeur de l'enseignement mathématique des collèges de France. Élevé à la dignité de pair de France par Louis-Philippe, Poisson fut nommé doyen de la faculté des sciences. Ses travaux portent en électricité, magnétisme, mécanique céleste, mouvements vibratoires. À la demande de Poisson et de Fourier une chaire de Calcul des probabilités et de physique mathématique sera créée à la faculté des sciences de Paris. Membre éminent de l'Académie des Sciences, section de physique générale, Poisson rejettera cependant et curieusement les travaux du jeune Galois en 1831. Ce n'est que tardivement que Poisson se tourne vers le calcul des probabilités ("Recherches sur la probabilité des jugements en matière criminelle et en matière civile"). Cette implication des mathématiques et du hasard dans la société fut à l'époque l'objet de houleux débats philosophiques.

X suit une loi de Poisson de paramètre λ, $\lambda > 0$, notée $X \hookrightarrow \mathcal{P}(\lambda)$

si $X(\Omega) = \mathbb{N} \qquad$ et $\forall\ k \in \mathbb{N} \quad \mathbb{P}(X = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$

On a bien $\displaystyle\sum_{k=0}^{+\infty} \mathbb{P}(X = k) = 1$

modèle probabiliste : X désigne le nombre de véhicules franchissant un péage pendant une période T, le nombre de clients pendant une période $T, \cdots$.

$\mathbb{E}(X) = \lambda \qquad V(X) = \lambda$

La série génératrice de X vérifie : $R_X = +\infty \qquad$ et $\forall\ t \in \mathbb{R} \quad G_X(t) = e^{\lambda(t-1)}$

théorème : $X \hookrightarrow \mathcal{P}(\lambda)$, $Y \hookrightarrow \mathcal{P}(\mu)$, X et Y var indépendantes, alors $X + Y \hookrightarrow \mathcal{P}(\lambda + \mu)$

V. Couples de var

V.1. Définition

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé

définition : X, Y deux var définies sur Ω , on pose $Z : \Omega \longrightarrow \mathbb{R}^2$

$$\omega \longmapsto (X(\omega), Y(\omega))$$

$Z = (X, Y)$ est appelé couple de var et Z est une var

On a $Z(\Omega) \subset X(\Omega) \times Y(\Omega)$

Propriété : X, Y deux vard définies sur Ω et $Z = (X, Y)$

Soit u une fonction de deux variables définie sur $X(\Omega) \times Y(\Omega)$, alors $u(Z) = u(X, Y)$ est une vard

V.2. Loi conjointe d'un couple de vard

Définition : $Z = (X, Y)$ couple de vard

L'application $X(\Omega) \times Y(\Omega) \longrightarrow \mathbb{R}$ est appelée loi conjointe du couple (X, Y)

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

Si $X(\Omega) = \{x_i, i \in I\}$ $I \subset \mathbb{N}$, I fini ou dénombrable

$Y(\Omega) = \{y_j, j \in J\}$ $J \subset \mathbb{N}$, J fini ou dénombrable

La loi conjointe est déterminée par la donnée de $\mathbb{P}(X = x_i, Y = y_j) = p_{ij}$,

$p_{ij} \in [0,1]$ et $\displaystyle\sum_{i \in I} \sum_{j \in J} p_{ij} = 1$

Réciproque : *(admise)* Toute famille $(p_{ij})_{i \in I, j \in J}$ de réels positifs de somme 1 définit la loi conjointe d'un couple (X, Y) de vard alors $\mathbb{P}(Z \in A \times B) = \displaystyle\sum_{i, x_i \in A} \sum_{j, j \in B} p_{ij}$

Généralisation : $X_1, \cdots, X_n$ n var définies sur $(\Omega, \mathcal{A}, \mathbb{P})$,

$Z = (X_1, \cdots, X_n)$ est une var à valeurs dans $\mathbb{R}^n$

La loi conjointe de $(X_1, \cdots, X_n)$ est la loi de Z définie par la donnée des nombres $\mathbb{P}(X_1 = x_1, \cdots, X_n = x_n)$

V.3. Lois marginales

définition : $Z = (X, Y)$ couple de vard définies sur $(\Omega, \mathcal{A}, \mathbb{P})$

$X(\Omega) = \{x_i, i \in I\}$ $I \subset \mathbb{N}$, I fini ou dénombrable

$Y(\Omega) = \{y_j, j \in J\}$ $J \subset \mathbb{N}$, J fini ou dénombrable

On appelle lois marginales du couple (X, Y) les lois de X (première loi marginale) et de Y (deuxième loi marginale)

théorème : $\forall\, i \in I$ $\quad \mathbb{P}(X = x_i) = \displaystyle\sum_{j \in J} \mathbb{P}(X = x_i, Y = y_j)$

$$p_i = \sum_{j \in J} p_{ij} \quad \text{et} \quad \sum_{i \in I} p_i = 1$$

$$\forall\ j \in J \qquad \mathbb{P}(Y=y_j)=\sum_{i\in I}\mathbb{P}(X=x_i,Y=y_j)$$

$$q_j=\sum_{i\in I}p_{ij} \quad \text{et} \sum_{j\in J}q_j=1$$

ie on obtient les lois marginales à partir de la loi conjointe

Remarque : Si $X(\Omega)$ et $Y(\Omega)$ sont finis, on peut représenter les lois dans un tableau

$Y\backslash X$	$\cdots$	x_i	$\cdots\cdots$	
$\cdots$				
y_j		p_{ij}		$=\mathrm{P}(Y=y_j)$
$\cdots$				
		$=\mathrm{P}(X=x_i)$		1

Généralisation : $X_1,\cdots,X_n$ n vard définies sur $(\Omega,\mathcal{A},\mathbb{P})$ et $Z=(X_1,\cdots,X_n)$

Les lois marginales de $X_1,\cdots,X_n$ dont données par si $x_i\in X_i(\Omega)$

$\mathbb{P}(X_i=x_i)=\displaystyle\sum_{x_1\in X_1(\Omega),\cdots,x_{i-1}\in X_{i-1}(\Omega),x_{i+1}\in X_{i+1}(\Omega),\cdots,x_n\in X_n(\Omega)}\mathbb{P}(X_1=x_1,\cdots,X_{i-1}=x_{i-1},X_i=x_i,X_{i+1}=x_{i+1},\cdots,X_n=x_n)$

V.4. Loi conditionnelle

$(\Omega,\mathcal{A},\mathbb{P})$ espace probabilisé, $B\in\mathcal{P}$ avec $\mathbb{P}(X\in B)\neq 0$

X vard définie sur $(\Omega,\mathcal{A},\mathbb{P})$

définition : La loi conditionnelle de X sachant B (ou loi conditionnée par B) est l'application

$$X(\Omega)\longrightarrow\mathbb{R}$$
$$x\longmapsto\mathbb{P}_B(X=x)=\frac{\mathbb{P}(X=x\cap B)}{\mathbb{P}(B)}$$

définition : (X,Y) couple de vard définies sur $(\Omega,\mathcal{A},\mathbb{P})$

$X(\Omega=\{x_i,i\in I\}\quad Y(\Omega)=\{y_j,j\in J\}$

Soit $m\in J$ tel que $\mathbb{P}(Y=y_m)\neq 0$, la loi conditionnelle de X sachant $(Y=y_m)$ est la donnée de

$$\mathbb{P}_{Y=y_m}(X=x_n)=\frac{\mathbb{P}(X=x_n,Y=y_m)}{\mathbb{P}(Y=y_m)}$$

De même si $\mathbb{P}(X=x_n)\neq 0$, $\mathbb{P}_{X=X_n}(Y=y_m)=\dfrac{\mathbb{P}(X=x_n,Y=y_m)}{\mathbb{P}(X=x_n)}$

Remarque : $\displaystyle\sum_{n\in I}\mathbb{P}_{y=y_m}(X=x_n)=\frac{1}{\mathbb{P}(Y=y_m)}\sum_{n\in I}\mathbb{P}(X=x_n,Y=y_m)=\frac{\mathbb{P}(Y=y_m)}{\mathbb{P}(Y=y_m)}=1$

Propriété : Si $\forall\ m \in J \quad \mathbb{P}(Y = y_m) \neq 0$,

$\forall\ (n,m) \in I \times J \qquad \mathbb{P}(X = x_n, Y = y_m)$=$\mathbb{P}_{Y=y_m}(X = x_n)\mathbb{P}(Y = y_m)$

et $\forall\ n \in I \quad \mathbb{P}(X = x_n) = \sum_{m\in J} \mathbb{P}_{Y=y_m}(X = x_n)\mathbb{P}(Y = y_m)$

formule des probabilités totales avec le SCE $(Y = y_m)_{m\in J}$

ie on obtient la loi de X à partir de la loi de Y et des lois conditionnelles de X sachant $Y = y_m$

De même avec X

V.5. Moments

X, Y vard, si X et Y admettent des espérances, alors $X+Y$ admet une espérance et $\mathbb{E}(X+Y)$ =$\mathbb{E}(X)$+$\mathbb{E}(Y)$

Propriété : Si X^2 et Y^2 admettent une espérance alors XY admet une espérance

et $\mathbb{E}(XY)^2 \leq$$\mathbb{E}(X^2)\mathbb{E}(Y^2)$ avec égalité si et seulement si $\exists\ (\alpha, \beta) \in \mathbb{R}^2 \setminus \{(0,0)\} \quad \alpha X + \beta Y$ est nulle presque sûrement.

et $X + Y$ admet une variance avec $V(X + Y) = V(X) + V(Y) + 2(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y))$

définition : Si X^2 et Y^2 admettent une espérance , on appelle covariance de (X, Y),

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

et coefficient de corrélation de (X, Y) (si de variance non nulle) $\quad \rho(X, Y) = \dfrac{Cov(X, Y)}{\sigma(X)\sigma(Y)}$

On a donc $V(X + Y)V(X) + V(Y) + 2Cov(X, Y) = V(X) + V(Y) + Cov(X, Y) + Cov(Y, X)$

théorème : Si X et Y admettent des moments d'ordre 2, alors $\rho(X, Y) \in$ [-1,1]

Généralisation : Si $X_1, \cdots, X_n$ sont n vard admettant des moments d'ordre 2,

alors $\mathbb{E}((X_1 + \cdots + X_n)^2)$ existe

et $V(X_1 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1\leq i<j\leq n} Cov(X_i, X_j) = \sum_{i=1}^{n} V(X_i) + \sum_{i\neq j} Cov(X_i, X_j)$

VI. Var indépendantes

VI.1. Indépendance de deux var

définition : X , Y vard définies sur $(\Omega, \mathcal{A}, \mathbb{P})$

X et Y sont indépendantes si $\forall\ A \subset X(\Omega) \quad \forall\ B \subset Y(\Omega) \qquad \mathbb{P}(X \in A, Y \in B)$ =$\mathbb{P}(X \in A)\mathbb{P}(Y \in B)$

On note alors $X \perp\!\!\!\perp Y$

théorème :

X et Y sont indépendantes $\Longleftrightarrow \forall\ (x, y) \in X(\Omega) \times Y(\omega) \qquad \mathbb{P}(X = x, Y = y)$ =$\mathbb{P}(X = x)\mathbb{P}(Y = y)$

théorème : X, Y var indépendantes, f et g deux fonctions définies respectivement sur $X(\Omega)$ et $Y(\Omega)$, alors $f(X) \perp\!\!\!\perp g(Y)$

théorème : $(\Omega, \mathcal{A}, \mathbb{P})$ espace probabilisé $(A, B) \in \mathcal{A}^2$

A et B sont indépendants $\Longleftrightarrow \mathbb{1}_A \perp\!\!\!\perp \mathbb{1}_B$

théorème : X , Y vard indépendantes, $m \in J$ tel que $\mathbb{P}(Y = y_m) \neq 0$,

alors la loi conditionnelle de X sachant $Y = y_m$ est égale à la loi de X

théorème : X , Y vard indépendantes admettant une espérance, alors XY admet une espérance et

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

Corollaire : X , Y vard indépendantes, alors

$V(X+Y) = V(X) + V(Y)$ et $Cov(X, Y) = 0$

théorème : X , Y vard à valeurs dans $\mathbb{N}$ et indépendantes, alors

$\forall\ t \in$ [-1,1] $G_X(t)G_Y(t) = G_{X+Y}(t)$

VI.2. Indépendance de n vard

définition : $X_1, \cdots, X_n$ n vard définies sur $(\Omega, \mathcal{A}, \mathbb{P})$,

$(X_1, \cdots, X_n)$ sont mutuellement indépendantes si

$$\forall\ i \in [\![1, \cdots, n]\!] \quad \forall\ A_i \subset X_i(\Omega) \quad \mathbb{P}(X_1 \in A_1, \cdots, X_n \in A_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in A_i)$$

Remarques : 1. $(X_1, \cdots, X_n)$ n mutuellement indépendantes, alors

$\forall\ k \leq n \quad X_{i_1}, \cdots, X_{i_k}$ sont mutuellement indépendantes

2. $(X_1, \cdots, X_n)$ n mutuellement indépendantes, alors les X_i sont deux à deux indépendantes,

la réciproque est fausse

théorème : $(X_1, \cdots, X_n)$ sont mutuellement indépendantes $\Longleftrightarrow$

$$\forall\ (x_1, \cdots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega) \quad \mathbb{P}(X_1 = x_1, \cdots, X_n = x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i)$$

Généralisation : $(X_n)_{n \in \mathbb{N}}$ suite de vard définies sur $(\Omega, \mathcal{A}, \mathbb{P})$,

$(X_n)_{n \in \mathbb{N}}$ sont mutuellement indépendantes si

$\forall\ n \in \mathbb{N} \quad X_0, \cdots, X_n$ sont mutuellement indépendantes

définition : La suite $(X_n)_{n \in \mathbb{N}}$ de vard définies sur $(\Omega, \mathcal{A}, \mathbb{P})$est i.i.d. si les variables suivent toutes la même loi(identiquement distribuées) et mutuellement indépendantes.

théorème : Lemme des coalitions :

$(X_1, \cdots, X_n)$ sont mutuellement indépendantes, alors si $1 \leq m \leq n-1$, alors

$f(X_1, \cdots, X_m)$ et $g(X_{m+1}, \cdots, X_n)$ sont indépendantes.

Généralisation : Si $(X_1, \cdots, X_n)$ sont mutuellement indépendantes, alors
$f_1(X_1, \cdots, X_{m_1})$, $f_2(X_{m_1+1}, \cdots, X_{m_1+m_2})$, $\cdots$, $f_p(X_{m_1+\cdots+m_{p-1}+1}, \cdots, X_{m_1+\cdots+m_p})$
sont mutuellement indépendantes.

VII. Inégalités probabilistes

VII.1. Inégalité de Markov

MARKOV Andreï Andreïevitch, 1856 Razian - 1922 Saint-Pétersbourg : Élève de Tchebychev à l'université de Saint-Pétersbourg où il fut nommé en 1886, ses travaux portent sur le calcul des probabilités et la théorie du potentiel. Markov crée l'analyse "markovienne", qui a permis de grands progrès dans le cryptage à vocation militaire mais aussi dans l'analyse de documents anciens dégradés dans le but de retrouver des textes partiellement effacés.

théorème : Soit X vard admettant un moment d'ordre 2, alors

$$\forall\ t > 0 \quad \mathbb{P}(|X| \geq t) \leq \frac{\mathrm{E}(|X|)}{t}$$

$$\forall\ t > 0 \quad \mathbb{P}(|X| \geq t) \leq \frac{\mathrm{E}(|X|^2)}{t^2}$$

VII.2. Inégalité de Bienaymé-Tchebychev

BIENAYMÉ Irénée Jules 1796 Paris - 1878 Paris : Polytechnicien, inspecteur général des Finances, Bienaymé se spécialise en statistique en appliquant la théorie des probabilités aux calculs financiers et aux problèmes démographiques. La révolution de 1848 le prive de son poste d'inspecteur des finances. Il obtient alors une chaire de probabilités à la Sorbonne. Dans le domaine de la statistique et des probabilités, on lui doit en particulier des compléments sur la méthode des moindres carrés et sur le thèorème central limite énoncé auparavant par De Moivre puis par Laplace. Bienaymé entre à l'Acadmie des sciences en 1852 et est un des membres fondateurs de la société mathématique de France.

TCHEBYCHEV Pafnouty Lvovitch 1821 Raïon de Zhukovsky - 1894 Saint-Pétersbourg : Tchebychev étudie à l'université de Moscou avant de s'installer à Saint-Pétersbourg. Après un doctorat portant sur l'intégration des fonctions elliptiques, Tchebychev crée sa propre école de mathématiques. Ses travaux portent sur l'intégration et l'approximation des fonctions tant algébriques qu'irrationnelles, la théorie des nombres et le calcul des probabilités où il définit le concept de quantité (variable) aléatoire.

théorème : Soit X vard admettant un moment d'ordre 2, alors

$$\forall\ \varepsilon > 0 \quad \mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

Corollaire : Soit X vard admettant un moment d'ordre 2, alors

$$\forall\ \varepsilon > 0 \quad \mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon) > 1 - \frac{V(X))}{\varepsilon^2}$$

VII.3. Loi faible des grands nombres:

théorème: $(X_n)_{n\geq 1}$ une suite de vard i.i.d. avec $\mathbb{E}(X^2)$ existe, on pose $S_n = \displaystyle\sum_{k=1}^{n} X_k$

$m = \mathbb{E}(X) \quad \sigma = \sigma(X)$

alors $\forall\ \varepsilon > 0 \qquad \lim\limits_{n \mapsto +\infty} \mathbb{P}\left(\left|\dfrac{S_n}{n} - m\right| \geq \varepsilon\right) = 0$