Chapitre 4 : Variables aléatoires continues

I) Fonction de répartition d'une variable aléatoire

Dans ce chapitre on ne suppose plus que l'ensemble $X(\Omega)$ des valeurs prises par une v.a. X, est discret. $(\Omega, \mathcal{F}, \mathbb{P})$ est toujours un espace probabilisé sur lequel seront définies les v.a. que nous considérerons.

Il convient de préciser ce qu'on entend par variable aléatoire dans ce contexte général :

Définition : 1) Une application $X : \Omega \to \mathbb{R}$ est appelée variable aléatoire (v.a.) si <u>pour tout intervalle</u> $I (\subset \mathbb{R})$,

$$[X \in I] := \{\omega \in \Omega \mid X(\omega) \in I\} \in \mathcal{F}$$

(condition de mesurabilité)

2) Une application $X : \Omega \to \mathbb{R}^d$ est un vecteur aléatoire si <u>pour tout pavé</u>, $\mathbf{I} = \prod_{i=1}^{d} I_i$ (I_i intervalle de $\mathbb{R}$, $\forall i$),

$$[X \in \mathbf{I}] = \{\omega \in \Omega \mid X(\omega) \in \mathbf{I}\} \in \mathcal{F}$$

Remarque : Soit X un vecteur aléatoire à valeurs dans $\mathbb{R}^d$. Pour tout $\omega \in \Omega$ écrivons

$$X(\omega) = (X_1(\omega), \dots, X_d(\omega))$$

(coordonnées de $X(\omega)$ dans la base canonique de $\mathbb{R}^d$. Alors les coordonnées X_i de X sont des v.a. réelles et on assimile le vecteur aléatoire X, au n-uplet

$$(X_1, X_2, \dots, X_n)$$

de ses coordonnées dans la base canonique de $\mathbb{R}^d$. Parfois on aura à changer de base ; dans ce cas les coordonnées de X devront être calculées dans la nouvelle base.

Définition 2 : On appelle <u>fonction de répartition</u> d'une v.a. X, la fonction $F_X : \mathbb{R} \to [0,1]$ définie par :

$$F_X(x) = \mathbb{P}(X \leq x) \quad (x \in \mathbb{R}).$$

Proposition : La fonction de répartition F_X possède les propriétés suivantes :

1) F_X est croissante (au sens large)
2) F_X est continue à droite en tout point $x \in \mathbb{R}$
3) $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$.

dém^n :

1) si $x < x'$, $[X \leq x] \subset [X \leq x'] \Rightarrow F_X(x) \leq F_X(x')$

2) Soit $x \in \mathbb{R}$ et $t_n \downarrow x$ (une suite de points décroissante et tendant vers x). Alors $[X \leq t_n]$ est une suite décroissante d'événements et $[X \leq x] = \bigcap_{n=1}^{+\infty} [X \leq t_n]$

Donc $F_X(x) = \lim_{n \to +\infty} F_X(t_n)$ (continuité de $\mathbb{P}$ par limite décroissante) d'où la continuité à droite

3) exercice.

__Remarques__ : 1) La notion de fonction de répartition est la généralisation de la notion de probabilités cumulées dans le cas d'une v.a. discrète. En effet soient $a < b \in \mathbb{R}$, on a

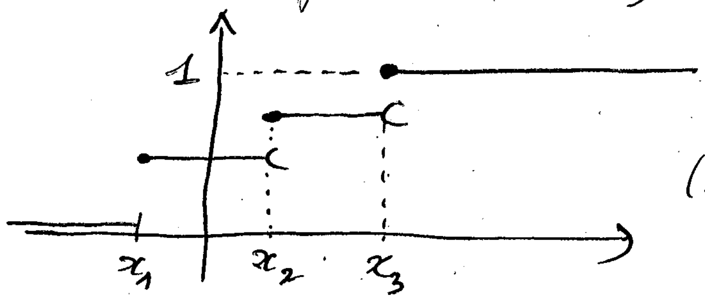
$$\mathbb{P}(X \in]a,b]) = F_X(b) - F_X(a)$$

puisque

$$[X \in]-\infty, b]] = [X \in]-\infty, a]] \cup [X \in]a,b]]$$

(réunion disjointe).

2) Toute v.a. admet une fonction de répartition. En particulier les v.a. discrètes. Par exemple si X prend un nombre fini de valeurs $x_1 < x_2 < \dots < x_N$, F_X est de la forme (en escalier)



(Ici $N = 3$)

le saut $F_X(x_i) - F_X(x_i^-) = \mathbb{P}(X = x_i)$

(où on note $F_X(x_i^-) = \lim_{\substack{t<x \\ t\to x}} F_X(t)$ la limite à gauche de F_X en x_i)

3) Pour les vecteurs aléatoires on a aussi une notion de fonction de répartition mais nous ne l'utiliserons pas.

4) Etant donnée une fonction F possédant les 3 propriétés de la Proposition (croissante, continue à droite en tout point et $\lim_{t\to-\infty} F(t) = 0$ et $\lim_{t\to+\infty} F(t) = 1$), __nous admettrons__ qu'il existe une v.a. X telle que $F_X = F$.

5) Si F_X est __continue en un point__ $x_0 \in \mathbb{R}$, alors

$$F_X(x_0) - F_X(x_0^-) = \mathbb{P}(X = x_0) = 0$$

6) Si F_X est __continue sur $\mathbb{R}$__, on dit que X est une __v.a. continue.__

__Exemple__ : soit $\lambda > 0$ et F la fonction définie par

$$F(t) = \begin{cases} 0 & \text{si } t < 0 \\ 1 - e^{-\lambda t} & \text{si } t \geq 0. \end{cases}$$

F est la fonction de répartition d'__une v.a. appelée exponentielle de paramètre λ__. Cette v.a. est continue et joue un rôle important dans les applications

(__attention__ : __Ne confondez pas__ v.a. de Poisson de paramètre $\lambda > 0$ et v.a. exponentielle de paramètre $\lambda > 0$!!!!)

__Exercice__ : Soit X une v.a. de loi exponentielle de paramètre $\lambda > 0$ et $t_0 > 0$ un réel fixé. On pose $Y = \min(X, t_0)$. Montrer que Y est une v.a. de fonction de répartition

$$F_Y(t) = \begin{cases} 0 & \text{si } t < 0 \\ 1 - e^{-\lambda t} & \text{si } 0 \leq t < t_0 \\ 1 & \text{si } t \geq t_0 \end{cases}$$

II Variables aléatoires ayant une densité

DÉFINITION

Soit $X : \Omega \to \mathbb{R}^d$ un vecteur aléatoire. On dit que X a une densité de probabilité s'il existe une fonction $f : \mathbb{R}^d \to \mathbb{R}$, $f \geq 0$ (sur $\mathbb{R}^d$ tout entier) f intégrable sur $\mathbb{R}^d$ (au sens de Lebesgue) t.q. $\forall I$ pavé de $\mathbb{R}^d$,

$$\mathbb{P}(X \in I) = \int_I f(x_1, \dots, x_d)\, dx_1 \dots dx_d \quad (*)$$

(noter qu'en particulier avec $I = \mathbb{R}^d$, la fonction f est telle que $\int_{\mathbb{R}^d} f(x_1, \dots, x_d)\, dx_1 \dots dx_d = 1$).

Si $d = 1$, la condition $(*)$ équivaut à dire que

$$\forall x \in \mathbb{R}, \quad F_X(x) = \int_{-\infty}^{x} f(t)\, dt \quad (**)$$

Remarque 1 Un vecteur aléatoire (resp. une v.a.) n'a pas forcément de densité de probabilité. Par exemple si $d = 1$, $(**)$ montre que si X a une densité, alors X est une v.a. continue. Mais attention la réciproque est fausse ! Il existe des v.a. continues qui n'ont pas de densité de probabilité.

Remarque 2 : Si la v.a. X a une densité de probabilité f continue sur $\mathbb{R}$, alors la fonction de répartition F_X est dérivable en tout point $x \in \mathbb{R}$ et

$$\frac{dF_X}{dx}(x) = f(x)$$

(d'après le théorème fondamental du calcul intégral élémentaire).

Remarque 3 : Si une v.a. X a une densité f, alors pour tout intervalle $I = [a,b]$ ou $[a,b[$ ou $]a,b]$ ou $]a,b[$, on a :

$$\mathbb{P}(X \in I) = \int_a^b f(x)\, dx$$

Remarque 4 (non mathématique) : En général on n'a pas à démontrer qu'une v.a. X admet une densité. Le fait que X admet une densité, est une hypothèse avec laquelle on travaille : c'est le cas pour le programme du Master MME donc pour la 1ère moitié du cours.

Exemples :

1) Une v.a. exponentielle de paramètre $\lambda > 0$ admet une densité de probabilité donnée par :

$$f(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ \lambda e^{-\lambda t} & \text{si } t > 0. \end{cases}$$

2) Soient $m \in \mathbb{R}$ et $\sigma > 0$ des paramètres fixés, la fonction

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right) \quad x \in \mathbb{R}$$

est une densité de probabilité appelée densité normale (ou de Laplace-Gauss) de paramètres m et σ^2 (on verra plus tard la signification de ces paramètres) (exercice : voir cours oral)

3) la fonction

$$(x,y) \mapsto f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\frac{x^2-2\rho xy+y^2}{1-\rho}\right)$$

(où $\rho \in]-1,1[$ est un paramètre) est la densité de probabilité d'un vecteur aléatoire de $\mathbb{R}^2$ (exercice : voir cours oral)

Remarque 5 : Etant donnée une densité de probabilité f sur $\mathbb{R}^d$ (i.e. une fonction $f \geq 0$, intégrable et d'intégrale totale égale à 1 :

$$\int_{\mathbb{R}^d} f(x_1,\dots,x_d)\,dx_1\dots dx_d = 1,$$

il existe des vecteurs aléatoires X de $\mathbb{R}^d$ ayant f pour densité de probabilité (admis).

III) Densités marginales d'un vecteur aléatoire et variables aléatoires indépendantes.

Soit X un vecteur aléatoire de $\mathbb{R}^d$ ayant une densité de probabilité f.

Soit $(X_1, X_2, \dots, X_d)$ le d-uplet des coordonnées de X dans la base canonique de $\mathbb{R}^d$.

$$\boxed{\text{Théorème : Chaque v.a. } X_k \ (k=1,\dots,d) \text{ a une densité de probabilité donnée par la formule : } f_k(t) = \int_{\mathbb{R}^{d-1}} f(x_1,\dots,x_{k-1},t,x_{k+1},\dots,x_d)\,dx_1\dots dx_{k-1}dx_{k+1}\dots dx_d}$$

(autrement dit on obtient la valeur $f_k(t)$ en fixant la $k^{\text{ième}}$ variable égale à t et en intégrant sur $\mathbb{R}^{d-1}$ la fonction $f(x_1,\dots,x_{k-1},t,x_{k+1},\dots,x_d)$).

Terminologie : la fonction f_k est appelée densité marginale d'indice k de la densité f.

dém^n : (dans le cas $n=2$ pour simplifier l'écriture). Soit (X,Y) un couple de v.a. de densité f sur $\mathbb{R}^2$. Cherchons la fonction de répartition de X :

$\forall t \in \mathbb{R}$,

$$F_X(t) = \mathbb{P}(X \leq t) = \mathbb{P}((X,Y) \in]-\infty,t] \times \mathbb{R})$$

$$= \int_{]-\infty,t]\times\mathbb{R}} f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{t}\left(\int_{\mathbb{R}} f(x,y)\,dy\right)dx \quad \text{(Th. de Fubini)}$$

$$= \int_{-\infty}^{t} f_1(x)\,dx \quad \text{où } f_1(x) = \int_{\mathbb{R}} f(x,y)\,dy$$

donc X admet (par définition) une densité égale à f_1. Même chose pour Y cqfd.

Application (v.a. indépendantes)

Définition : Soit $(X_1, X_2, \dots, X_n)$ un n-uplet de variables aléatoires. On dit que les v.a. X_i $(i=1,\dots,n)$ sont indépendantes si pour tous les intervalles I_i $(\subset \mathbb{R})$ $(i=1,\dots,n)$, on a

$$\mathbb{P}(X_1 \in I_1, X_2 \in I_2, \dots, X_n \in I_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in I_i)$$

Exercice : Montrer que si les v.a. X_i sont discrètes la notion d'indépendance donnée dans le chapitre 3 est équivalente à celle donnée ici.

Théorème (condition d'indépendance pour des v.a. ayant une densité) : On suppose que X_i a une densité de probabilité f_i (pour tout $i=1,\dots,n$). Alors les v.a. X_i $(i=1,\dots,n)$ sont indépendantes si et seulement si la fonction $f : \mathbb{R}^n \to \mathbb{R}$ t.q. :

$$f(x_1, \dots, x_n) = f_1(x_1) f_2(x_2) \cdots f_n(x_n),$$

est une densité de probabilité du vecteur aléatoire $X = (X_1, X_2, \dots, X_n)$.

dém^n : a) ($\Rightarrow$) supposons que les v.a. X_i sont indépendantes. Pour tout pavé $I = I_1 \times I_2 \times \dots \times I_n$ de $\mathbb{R}^n$ on a donc

$$\begin{aligned}
\mathbb{P}(X \in I) &= \mathbb{P}(X_1 \in I_1, \dots, X_n \in I_n) \\
&= \prod_{j=1}^{n} \mathbb{P}(X_j \in I_j) \quad \text{(indépendance)} \\
&= \prod_{j=1}^{n} \int_{I_j} f_j(x_j)\,dx_j \quad (\text{car } f_j \text{ densité de } X_j) \\
&= \int_{I_1 \times \dots \times I_n} f_1(x_1) f_2(x_2) \cdots f_n(x_n)\, dx_1\, dx_2 \cdots dx_n \quad \text{(Fubini)}
\end{aligned}$$

Donc $f(x_1, \dots, x_n) = f_1(x_1) f_2(x_2) \cdots f_n(x_n)$ est une densité du vecteur X.

b) ($\Leftarrow$) exercice facile.

Cas particulier : Pour un couple (X, Y) de densité $f(x,y)$ on note

$$f(x, \cdot) = \int_{-\infty}^{\infty} f(x,y)\,dy \quad \text{et} \quad f(\cdot, y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

les densités marginales. La condition d'indépendance de X et Y est :

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

(densité du couple = produit des densités marginales).

Application : Densité d'une somme de deux variables aléatoires indépendantes :

Théorème : Soient X et Y des v.a. indépendantes ayant des densités respectivement égales à f_1 et f_2. Alors la v.a. $X+Y$ a une densité donnée par la formule

$$f(t) = \int_{-\infty}^{\infty} f_1(x) f_2(t-x)\,dx = \int_{-\infty}^{\infty} f_1(t-y) f_2(y)\,dy$$

(on dit que $f = f_1 * f_2$ est le produit de convolution des fonctions f_1 et f_2).

dém^n : pour simplifier la démonstration on va supposer que f_1 et f_2 sont des fonctions continues.

La fonction de répartition F de la v.a. $X+Y$ est égale à

$$F(t) = \mathbb{P}(X+Y \leq t) = \mathbb{P}((X,Y) \in D_t)$$

où $D_t = \{(x,y) \in \mathbb{R}^2 \mid x+y \leq t\}$. D'où

$$F(t) = \int_{D_t} f_1(x) f_2(y)\,dx\,dy \quad (\text{car } X \text{ et } Y \text{ indép.})$$

$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{t-y} f_1(x)\,dx \right) f_2(y)\,dy$$

Mais grâce au théorème de dérivation de Lebesgue sous le signe intégrale, on voit que F est continuement dérivable et que

$$F'(t) = \int_{-\infty}^{\infty} f_1(t-y) f_2(y)\,dy$$

Cette fonction $t \mapsto F'(t)$ est donc une densité de probabilité de la v.a. $X+Y$.

L'autre formule s'obtient en commençant par intégrer par rapport à l'autre variable. CQFD.

IV Loi de probabilité d'une v.a. (resp. d'un vecteur aléatoire).

Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow \mathbb{R}^d$ un vecteur aléatoire.

Soit $\mathcal{B}_d$ la tribu de Borel de $\mathbb{R}^d$ (i.e. la plus petite tribu de parties de $\mathbb{R}^d$ contenant les pavés $I = \prod_{i=1}^{d} I_i$ (I_i intervalle de $\mathbb{R}$)).

On peut justifier que pour tout $B \in \mathcal{B}_d$,

$$[X \in B] = \{\omega \in \Omega \,;\, X(\omega) \in B\} \in \mathcal{F}$$

(admis pour l'instant pour le Master MME).

Proposition : L'application $\mu_X : \mathcal{B}_d \to [0,1]$ définie par :

$$\forall B \in \mathcal{B}_d, \ \mu_X(B) = \mathbb{P}(X \in B),$$

est une mesure de probabilité sur $(\mathbb{R}^d, \mathcal{B}_d)$.

Définition : La mesure μ_X est appelée *loi de probabilité* du vecteur aléatoire X.

dém^{on} de la proposition :

si $B \in \mathcal{B}_d$, $[X \in B] \in \mathcal{F}$ donc le nombre $\mu_X(B) = \mathbb{P}(X \in B)$ est bien défini. De plus:

a) $\mu_X(\mathbb{R}^d) = \mathbb{P}(X \in \mathbb{R}^d) = \mathbb{P}(\Omega) = 1$.

b) Pour toute suite $(B_n)(\subset \mathcal{B}_d)$ de boréliens de $\mathbb{R}^d$ 2 à 2 disjoints, on a

$$\mu_X(\bigcup_n B_n) = \mathbb{P}(X \in \bigcup_n B_n) = \mathbb{P}(\bigcup_n [X \in B_n])$$

$$= \sum_n \mathbb{P}(X \in B_n) \text{ (car les } [X \in B_n]$$

sont 2 à 2 incompatibles)

$$= \sum_n \mu_X(B_n)$$

a) et b) $\Longrightarrow$ μ_X est une mesure de probabilité sur $(\mathbb{R}^d, \mathcal{B}_d)$ cqfd.

Notation de l'intégrale relative à la mesure μ_X

Etant donné une fonction $g : \mathbb{R}^d \longrightarrow \mathbb{R}$ mesurable et intégrable par rapport à la mesure μ_X, on notera

$$\int_{\mathbb{R}^d} g(x)\, d\mu_X(x) \quad \text{ou} \quad \int_{\mathbb{R}^d} g(x)\, \mu_X(dx)$$

son intégrale par rapport à μ_X.

Remarque: les fonctions g boréliennes bornées sur $\mathbb{R}^d$ sont μ_X-intégrables.

Cas des variables aléatoires : lien entre loi de probabilité et fonction de répartition:

Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow \mathbb{R}$ une variable aléatoire de loi de probabilité μ_X et de fonction de répartition F_X.

En prenant comme borélien B l'intervalle $]-\infty, t]$, on voit que

$$\mathbb{P}(X \leq t) = \mathbb{P}(X \in]-\infty, t]) = \mu_X(]-\infty, t])$$

Mais on a aussi: $\mathbb{P}(X \leq t) = F_X(t)$ par définition.

Donc :

$$\forall t \in \mathbb{R}, \quad F_X(t) = \mu_X(]-\infty, t])$$

De plus le lien entre mesure de probabilité sur $(\mathbb{R}, \mathcal{B}_1)$ et fonction de répartition est plus précis:

Convenons d'appeler fonction de répartition (sans faire référence à une v.a.) toute fonction $F : \mathbb{R} \to [0,1]$ t.q.: a) F est croissante (au sens large), b) F est continue à droite en tout point c) $\lim_{t \to -\infty} F(t) = 0$ et $\lim_{t \to +\infty} F(t) = 1$. Alors:

$$\boxed{\begin{array}{l}\text{Théorème (admis): } F : \mathbb{R} \to [0,1] \text{ est une fonction de répartition si et seulement s'il} \\ \text{existe une (unique) mesure de probabilité } \mu \text{ sur } (\mathbb{R}, \mathcal{B}_1) \text{ telle que :} \\ \forall t \in \mathbb{R}, \quad F(t) = \mu(]-\infty, t]).\end{array}}$$

Cas particulier des v.a. ($d=1$) :

Pour $f : \mathbb{R} \to \mathbb{R}$ borélienne μ_X-intégrable, on note aussi

$$\int_{\mathbb{R}} g(x)\,d\mu_X(x) = \int_{-\infty}^{\infty} g(x)\,dF_X(x)$$

l'intégrale de Lebesgue-Stieltjes par rapport à la fonction de répartition F_X de la v.a. X.

Remarque : Lorsque X a une densité de probabilité $f(x)$, alors

$$\mu_X = f(x)dx$$

i.e. μ_X est une mesure de densité f par rapport à la mesure de Lebesgue dx, et on a simplement

$$\int_{\mathbb{R}} g(x)\,d\mu_X(x) = \int_{\mathbb{R}} g(x) f(x)dx.$$

V Espérance, variance et moments d'une v.a. (resp d'un vecteur aléatoire)

Définition : On dit que la v.a. X a un moment d'ordre 1 si

$$\int_{\mathbb{R}} |x|\,d\mu_X(x) < +\infty$$

(i.e. si la fonction $x \mapsto |x|$ est μ_X-intégrable).

alors l'intégrale $\int_{\mathbb{R}} x\,d\mu_X(x)$ existe et on l'appelle l'espérance mathématique de X. On note

$$\mathbb{E}(X) = \int_{\mathbb{R}} x\,d\mu_X(x)$$

Dans le cas où X a une densité f, on a :

$$\mathbb{E}(X) = \int_{\mathbb{R}} x f(x)dx$$

Fonctions déterministes d'une v.a. (resp vecteur aléatoire)

Soit $g : \mathbb{R}^d \to \mathbb{R}$ une fonction borélienne et X un vecteur aléatoire de $\mathbb{R}^d$. Alors l'application composée $g\circ X : \Omega \to \mathbb{R}$ est une variable aléatoire :

$$\omega \longmapsto g\circ X(\omega) = g(X(\omega))$$

qu'on note aussi $g(X)$.

Théorème : La v.a. $g(X)$ a un moment d'ordre 1 si et seulement si la fonction g est μ_X-intégrable (i.e. g borélienne et $\int_{\mathbb{R}^d} |g(x)|\,d\mu_X(x) < +\infty$).

Dans ce cas, on a

$$\mathbb{E}(g(X)) = \int_{\mathbb{R}^d} g(x)\,d\mu_X(x).$$

démonstration : résultat admis pour le moment. On notera l'analogie avec le cas des vecteurs aléatoires discrets vu au chapitre 3.

Moments d'une v.a.

Définition : On dit que la v.a. X a un **moment d'ordre $r \in \mathbb{N}$** si la v.a. X^r a un moment d'ordre 1. Dans ce cas,

$$\mathbb{E}(X^r) = \int_{\mathbb{R}} x^r \, d\mu_X(x)$$

Si X a un moment d'ordre 2, on appelle variance de X le nombre :

$$\mathrm{Var}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$$
$$= \int_{\mathbb{R}} (x - \mathbb{E}(X))^2 \, d\mu_X(x)$$

On notera qu'on a aussi

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Formules sur l'espérance et la variance :

Comme pour les v.a. discrètes on a les résultats suivants :

Théorème : Soient $X_1, \dots, X_n$ des v.a. réelles, $\alpha_1, \dots, \alpha_n$ des constantes. Alors :

1) si les X_i ont un moment d'ordre 1, $\alpha_1 X_1 + \dots + \alpha_n X_n$ aussi et

$$\mathbb{E}\left(\sum_{i=1}^{n} \alpha_i X_i\right) = \sum_{i=1}^{n} \alpha_i \mathbb{E}(X_i) \quad \text{(linéarité de l'espérance)}$$

2) Si les X_i sont indépendantes et ont un moment d'ordre 1, alors la v.a. $Y = X_1 \cdots X_n$ a aussi un moment d'ordre 1 et

$$\mathbb{E}(Y) = \prod_{i=1}^{n} \mathbb{E}(X_i)$$

3) Si les X_i sont indépendantes et ont un moment d'ordre 2, alors la v.a. $S = X_1 + \dots + X_n$ a aussi un moment d'ordre 2 et

$$\mathrm{Var}(S) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

Remarque : l'inégalité de Bienaymé-Tchebychev est aussi valable pour les v.a. ayant un moment d'ordre 2 (comme dans le cas discret) et la loi faible des grands nombres est aussi valable.

Proposition (inégalité de Markov) : Soit X une v.a. ayant un moment d'ordre 1. Alors pour tout réel $a > 0$, on a

$$\mathbb{P}(|X| \geq a) \leq \frac{1}{a} \mathbb{E}(|X|)$$

démon : $\mathbb{P}(|X| \geq a) = \mu_X(\{x \in \mathbb{R} \, ; \, |x| \geq a\})$ (par définition de μ_X) et

$$\mathbb{E}(|X|) = \int_{\mathbb{R}} |x| \, d\mu_X(x) \geq \int_{\{|x| \geq a\}} |x| \, d\mu_X(x) \geq$$

$$\geq a \int_{\{|x|\geq a\}} d\mu_X(x) = \mu_X(\{|x|\geq a\}) = \mathbb{P}(|X|\geq a)$$

d'où le résultat.

<u>Exercice</u> : Déduire de l'inégalité de Markov, l'inégalité de Bienaymé-Tchebychev i.e. si X a un moment d'ordre 2, pour tout $a>0$,

$$\mathbb{P}\left(\left|\frac{X-m}{\sigma}\right|\geq a\right) \leq \frac{1}{a^2}\mathrm{Var}(X).$$

$\left(m = \mathbb{E}(X) \text{ et } \sigma = \sqrt{\mathrm{Var}(X)}\right)$.

<u>Application à la détermination pratique de la loi de probabilité d'une v.a. (resp. d'un vecteur aléatoire)</u>

Une mesure de probabilité μ sur $(\mathbb{R}^d, \mathcal{B}_d)$ est parfaitement déterminée par les valeurs des intégrales

$$\int_{\mathbb{R}^d} g(x)\, d\mu(x)$$

pour g décrivant l'ensemble des fonctions boréliennes bornées sur $\mathbb{R}^d$. En particulier si μ et ν sont des mesures de probabilité telles que

$$\forall g \text{ borélienne bornée}, \int_{\mathbb{R}^d} g(x)\, d\mu(x) = \int_{\mathbb{R}^d} g(x)\, d\nu(x),$$

alors, $\mu = \nu$ (résultat de théorie de la mesure).

On tire de ceci une réciproque très utile dans la pratique du théorème (voir p. 15) permettant le calcul de $\mathbb{E}(g(X))$ lorsqu'on connait la loi de X.

Notons que si $g : \mathbb{R}^d \to \mathbb{R}$ est borélienne et <u>bornée</u> sur $\mathbb{R}^d$, alors $g(X)$ a <u>toujours</u> un moment d'ordre 1 et $\mathbb{E}(g(X))$ <u>existe</u>.

<u>Théorème</u> : Soit X un vecteur aléatoire de $\mathbb{R}^d$. Supposons qu'il existe une mesure de probabilité μ sur $(\mathbb{R}^d, \mathcal{B}_d)$ telle que :

$\forall g : \mathbb{R}^d \to \mathbb{R}$ borélienne et <u>bornée</u> sur $\mathbb{R}^d$, on ait

$$\mathbb{E}(g(X)) = \int_{\mathbb{R}^d} g(x)\, d\mu(x),$$

<u>Alors</u> μ est la loi de probabilité du vecteur aléatoire X.

<u>Exemple</u> : Soit (X,Y) un couple de variables aléatoires réelles (i.e. un vecteur aléatoire de $\mathbb{R}^2$) de densité de probabilité

$$f(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} \quad ((x,y)\in\mathbb{R}^2).$$

Trouver la loi de la v.a. $Z = X^2+Y^2$.

<u>Solution</u> : Soit $g : \mathbb{R} \to \mathbb{R}$ une fonction borélienne bornée sur $\mathbb{R}$.

On va calculer $\mathbb{E}(g(Z))$:

$$g(Z) = g(X^2+Y^2) = (g\circ h)(X,Y)$$

où $h : (x,y) \longmapsto x^2+y^2$.

D'où

$$\mathbb{E}(g(Z)) = \mathbb{E}\big((g\circ h)(X,Y)\big)$$

$$= \iint_{\mathbb{R}^2} g\circ h(x,y)\, f(x,y)\,dx\,dy$$

$$= \iint_{\mathbb{R}^2} g(x^2+y^2)\frac{1}{2\pi}\, e^{-\frac{1}{2}(x^2+y^2)}\,dx\,dy \quad (*)$$

Posons $x = \rho\cos\theta$, $y = \rho\sin\theta$ $\quad (\theta\in[0,2\pi],\ \rho\in\mathbb{R}_+)$

Alors

$$(*) = \int_{\rho=0}^{+\infty}\int_{\theta=0}^{2\pi} g(\rho^2)\frac{1}{2\pi}e^{-\frac{1}{2}\rho^2}\rho\, d\rho\, d\theta$$

$$= \int_0^{+\infty} g(\rho^2)e^{-\frac{1}{2}\rho^2}\rho\, d\rho.$$

Faisons le chg^t de variable $z = \rho^2$ $(\Rightarrow dz = 2\rho\, d\rho)$

Alors

$$(*) = \int_0^{+\infty} g(z)\underbrace{\tfrac{1}{2}e^{-\frac{1}{2}z}}\, dz$$

Remarque : c'est la densité de la loi exponentielle de paramètre $\lambda = \frac{1}{2}$

Conclusion :

la fonction $f(z) = \frac{1}{2}e^{-\frac{1}{2}z}\mathbb{1}_{[0,+\infty[}(z)$ est la densité de la v.a. Z