CHAPITRE 6

VARIABLES ALÉATOIRES GÉNÉRALES LOIS de PROBABILITÉ, INDÉPENDANCE.

Introduction : Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $X : \Omega \to \mathbb{R}$ une application. On va considérer le cas général où l'ensemble $X(\Omega)$ des valeurs prises par X n'est pas forcément discret.

Pour que X soit une variable aléatoire, il est raisonnable de supposer que pour tout intervalle $I \subset \mathbb{R}$, on puisse dire si $X \in I$, autrement dit il convient que l'ensemble

$$\{\omega \in \Omega \, ; \, X(\omega) \in I\}$$

soit un événement, c'est à dire appartienne à $\mathcal{F}$.

On notera cet événement $[X \in I]$. C'est en fait l'image inverse $X^{-1}(I)$ de I par X.

Cette condition

$$\forall a < b, \, [X \in I] \in \mathcal{F}.$$

s'appelle condition de mesurabilité.

La raison de cette dénomination est la suivante :

<u>Rappel</u> : Si $(\Omega, \mathcal{F})$ et $(\Omega', \mathcal{F}')$ sont des espaces mesurables et $X : \Omega \to \Omega'$ une application, on dit que X est <u>mesurable</u> si

$$\forall B \in \mathcal{F}', \quad X^{-1}(B) := [X \in B] \in \mathcal{F}$$

(où $[X \in B] = \{\omega \in \Omega ;\ X(\omega) \in B\}$)

<u>Définition</u> : si $(\Omega', \mathcal{F}') = (\mathbb{R}^d, \mathcal{B}_d)$ où $\mathcal{B}_d$ est <u>la tribu des boréliens</u> de $\mathbb{R}^d$, toute application mesurable $X : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}^d, \mathcal{B}_d)$ est appelée <u>vecteur aléatoire</u> à valeurs dans $\mathbb{R}^d$. Si $d = 1$, X s'appelle <u>variable aléatoire</u> (v.a. en abrégé dans la suite).

(I) <u>Conditions de mesurabilité et Opérations sur les variables aléatoires</u>

(A) Voici des conditions plus simples pour vérifier qu'une application $X : \Omega \to \mathbb{R}^d$ est un vecteur aléatoire

Soit $\mathcal{C}$ une partie génératrice de la tribu $\mathcal{B}_d$ (i.e. la tribu $\mathcal{T}(\mathcal{C})$ engendrée par $\mathcal{C}$ est égale à $\mathcal{B}_d$.

Proposition : Soit $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}_d)$ telle que
$\forall C \in \mathcal{C}$, $[X \in C] \in \mathcal{F}$.
Alors X est un vecteur aléatoire

démonstration : Soit $\mathcal{G} = \{A \subset \mathbb{R}^d ; [X \in A] \in \mathcal{F}\}$.
Par définition $\mathcal{G} \supset \mathcal{C}$. Montrons que $\mathcal{G}$ est une tribu :

i) $\phi \in \mathcal{G}$ (clair)

ii) si $A \in \mathcal{G}$, $[X \in A] \in \mathcal{F}$ donc $[X \in A]^c \in \mathcal{F}$.
Mais $[X \in A]^c = [X \in A^c]$. D'où $A^c \in \mathcal{G}$.

iii) si $(A_n) \in \mathcal{G}$, $[X \in A_n] \in \mathcal{F}$ pour tout entier n,
donc $\bigcup_n [X \in A_n] \in \mathcal{F}$ (car $\mathcal{F}$ est une tribu).
Mais $\bigcup_n [X \in A_n] = [X \in \bigcup_n A_n]$. Donc $\bigcup_n A_n \in \mathcal{F}$.

D'où $\mathcal{G}$ est une tribu telle que $\mathcal{G} \supset \mathcal{C}$. Donc $\mathcal{G} \supset \mathcal{T}(\mathcal{C})$ (car $\mathcal{T}(\mathcal{C})$ est la plus petite tribu contenant $\mathcal{C}$) i.e. $\mathcal{G} \supset \mathcal{B}_d$ CQFD.

Corollaire 1 $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_1)$ est une v.a. si et seulement si pour tout I intervalle de $\mathbb{R}$, $[X \in I] \in \mathcal{F}$.

Corollaire 2 : $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_1)$ est une r.a.
si et seulement si

$\forall x \in \mathbb{R}$, $[X \leq x] \in \mathcal{F}$

(resp. $\forall x \in \mathbb{R}$, $[X < x] \in \mathcal{F}$)

démⁿ : Soit $\mathcal{C} = \{]-\infty, x] \, ; \, x \in \mathbb{R} \}$.
La tribu engendrée par $\mathcal{C}$ contient les ensembles

$$]a, b] =]-\infty, b] \cap]-\infty, a]^c$$

donc $\mathcal{C}$ contient aussi les ouverts

$$]a, b[= \bigcup_{n=1}^{\infty}]a, b - \tfrac{1}{n}]$$

Donc $\mathcal{T}(\mathcal{C}) \supset \mathcal{B}_1$. D'où $\mathcal{T}(\mathcal{C}) = \mathcal{B}_1$.

CQFD.

(B) Opérations sur les v.a. (resp les vecteurs aléatoires) :

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

Théorème : Soit $f : (\mathbb{R}^d, \mathcal{B}_d) \to (\mathbb{R}^{d'}, \mathcal{B}_{d'})$ une fonction mesurable. Alors la fonction composée

$$f(X) = f \circ X : (\Omega, \mathcal{F}) \to (\mathbb{R}^{d'}, \mathcal{B}_{d'})$$

est un vecteur aléatoire.

<u>dém^n</u> : Soit $B \in \mathcal{B}_{d'}$. On a

$$[f \circ X \in B] = [X \in f^{-1}(B)].$$

Comme f est mesurable, $f^{-1}(B) \in \mathcal{B}_d$ et comme X est un v.a. $[X \in f^{-1}(B)] \in \mathcal{F}$. <u>CQFD</u>.

<u>Corollaire 1</u> : Si $f : (\mathbb{R}^d, \mathcal{B}_d) \to (\mathbb{R}^{d'}, \mathcal{B}_{d'})$ est continue, $f(X)$ est un vecteur aléatoire.

<u>dém^n</u> : Pour tout ouvert $\mathcal{O}$ de $\mathbb{R}^{d'}$, $f^{-1}(\mathcal{O})$ est un ouvert de $\mathbb{R}^d$ donc $f^{-1}(\mathcal{O}) \in \mathcal{B}_d$ et alors

$$[X \in f^{-1}(\mathcal{O})] \in \mathcal{F}.$$

Donc $[f(X) \in \mathcal{O}] \in \mathcal{F}$.

Comme les ouverts engendrent $\mathcal{B}_{d'}$ la proposition p.3 montre aussitôt que $f(X)$ est un vecteur aléatoire

<u>Corollaire 2</u> : Si $X, Y : (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}_d)$ sont des vecteurs aléatoires alors :

1) $\forall \alpha, \beta \in \mathbb{R}$, $\alpha X + \beta Y : \omega \mapsto \alpha X(\omega) + \beta Y(\omega)$ est un vecteur aléatoire

2) si $d = 1$, $XY : \omega \mapsto X(\omega) \cdot Y(\omega)$ est une v.a.

3) si $d = 1$ et si $\forall \omega \in \Omega, Y(\omega) \neq 0$,

$$\frac{X}{Y} : \omega \mapsto \frac{X(\omega)}{Y(\omega)} \text{ est une v.a.}$$

<u>dém^n</u> :

1) $\alpha X + \beta Y$ est l'application composée des deux applications mesurables

$$\omega \mapsto (X(\omega), Y(\omega)) \quad \text{de } \Omega \text{ dans } \Omega \times \Omega$$

et

$$(x,y) \mapsto \alpha x + \beta y \quad \text{de } \mathbb{R}^d \times \mathbb{R}^d \text{ dans } \mathbb{R}^d$$

2) $X \cdot Y$ est l'application composée de $\omega \mapsto (X(\omega), Y(\omega))$ et $(x,y) \mapsto xy$

3) analogue à 2) avec $(x,y) \mapsto \frac{x}{y}$.

<u>Théorème</u> :

1) Si $X_1, \dots, X_N$ est une suite finie de v.a. de $(\Omega, \mathcal{F})$ dans $(\mathbb{R}, \mathcal{B}_1)$, les applications

$$M_N = \underset{1 \leq k \leq N}{\mathrm{Max}}\, X_k \quad \text{et} \quad I_N = \underset{1 \leq k \leq N}{\mathrm{Min}}\, X_k$$

sont des v.a.

2) Si (X_n) est une suite infinie dénombrable de v.a. et si les applications

$$S = \sup_{n \in \mathbb{N}} X_n \quad \text{et} \quad I = \inf_{n \in \mathbb{N}} X_n$$

<u>sont bien définies</u> de Ω dans $\mathbb{R}$ (i.e. $\forall \omega \in \Omega$, $S(\omega) = \sup_n X_n(\omega) < +\infty$ et $I(\omega) = \inf_n X_n(\omega) > -\infty$) ce sont des v.a.

dém^n : Voyons l'assertion 2 : $\forall x \in \mathbb{R}$, on a

$$[S \leq x] = [\operatorname{Sup}_n X_n \leq x] = \bigcap_{n=1}^{\infty} [X_n \leq x] \in \mathcal{F}$$

car tous les $[X_n \leq x] \in \mathcal{F}$ (qui est une tribu).

Donc S est une v.a. d'après le corollaire 2 p. 4.

De même

$$[I \leq x]^c = [I > x] = \bigcap_{n=1}^{\infty} [X_n > x] \in \mathcal{F}$$

donc $[I \leq x] \in \mathcal{F}$.

Corollaire : Soit $X_n : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_1)$ une suite de v.a. qui converge simplement vers une limite finie X

(i.e. $\forall \omega \in \Omega$, $\lim_{n \to +\infty} X_n(\omega) = X(\omega) \in \mathbb{R}$ existe)

Alors X est une v.a.

dém^n : $X = \lim_n X_n = \liminf X_n =$

$$\operatorname{Sup}_n \left(\operatorname{Inf}_{k \geq n} X_k \right) = \operatorname{Sup}_n Y_n \quad \text{où } Y_n = \operatorname{Inf}_{k \geq n} X_k$$

est une v.a. Donc $\operatorname{Sup}_n Y_n$ est aussi une v.a. par le théorème p. 6.

(C) Variables aléatoires simples et théorème de structure des v.a.

Définition : 1) Soit $A \in \mathcal{F}$ un événement.
L'application

$$\mathbb{1}_A : \omega \mapsto \mathbb{1}_A(\omega) = \begin{cases} 0 \text{ si } \omega \notin A \\ 1 \text{ si } \omega \in A \end{cases}$$

s'appelle variable indicatrice de A. C'est clairement une v.a.

2) On appelle v.a. simple toute combinaison linéaire de v.a. indicatrices :

$$X = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i} \quad (\alpha_i \in \mathbb{R}, A_i \in \mathcal{F})$$

Remarque (et exercice) Toute v.a. simple peut s'écrire $X = \sum_{j=1}^{m} \beta_j \mathbb{1}_{B_j}$ où les $B_j \in \mathcal{F}$ sont deux à deux disjoints et les $\beta_j \in \mathbb{R}$. Autrement dit une v.a. simple est une v.a. qui prend un nombre fini de valeurs (i.e. $\operatorname{Card} X(\Omega) < +\infty$)

Théorème (de structure) : Toute v.a. $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_1)$ est limite simple d'une suite de v.a. simples.

démo : Pour chaque $n > 0$, considérons les événements

$$A_k^{(n)} = \left[X \in \left[\frac{k}{n}, \frac{k+1}{n} \right[\right], \; k \in \{-n^2, \dots, n^2\}$$

et posons

$$X_n = \sum_{k=-n^2}^{n^2} \frac{k}{n} \mathbb{1}_{A_k^{(n)}} .$$

Montrons que $X = \lim_n X_n$

Soit $\omega \in \Omega$. Il existe $n_0 \in \mathbb{N}$ tel que $X(\omega) \in [-n_0, n_0]$ et pour tout $n > n_0$ il existe un entier k tel que

$$X(\omega) \in \left[\frac{k}{n}, \frac{k+1}{n}\right[$$

donc $\omega \in A_k^{(n)}$. On a alors $X_n(\omega) = \frac{k}{n}$ et par conséquent

$$|X(\omega) - X_n(\omega)| < \frac{1}{n}$$

ce qui prouve que $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ <u>CQFD</u>.

<u>Exercice</u> Montrer que si $X \geq 0$ est une v.a. alors X est limite d'une <u>suite croissante</u> de v.a. simples

II Loi de probabilité et fonction de répartition des v.a.

(A) Loi de probabilité d'un vecteur aléatoire

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

<u>Proposition</u>: L'application $\mu_X : \mathcal{B}_d \longrightarrow [0,1]$ définie par :

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

est une mesure de probabilité sur $(\mathbb{R}^d, \mathcal{B}_d)$. On l'appelle <u>loi de probabilité</u> du vecteur X.

<u>dém^n</u>: 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 :

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

ii) Pour toute suite (B_n) de boréliens de $\mathbb{R}^d$, 2 à 2 disjoints, on a

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

car les $[X \in B_n]$ sont des événements incompatibles

Exemples

1) Les <u>v.a. discrètes</u> sont celles telles que $X(\Omega)$ est fini ou dénombrable. Si $X(\Omega) = \{x_k \,;\, k \in \mathbb{N}\}$ ou $\{x_k \,;\, 1 \leq k \leq N\}$, la loi de probabilité μ_X est la mesure de probabilité

$$\mu_X = \sum_k p_k \, \delta_{x_k} \qquad p_k = \mathbb{P}(X = x_k)$$

où δ_{x_k} est la mesure de Dirac au point x_k (i.e. $\forall B \in \mathcal{B}_d$, $\delta_{x_k}(B) = \begin{cases} 1 & \text{si } x_k \in B \\ 0 & \text{si } x_k \notin B \end{cases}$).

Une telle mesure est appelée mesure discrète de "masses" p_k aux points x_k.

2) Les v.a. <u>continues</u> sont celles telles que $\forall x \in \mathbb{R}^d$, $\mathbb{P}(X = x) = 0$ (i.e. elles n'ont de la masse en aucun point). Les cas particuliers les plus importants de v.a. continues sont celles qui ont une <u>densité de probabilité</u> on les appelle aussi <u>v.a. absolument continues</u> car la loi μ_X de X est absolument continue par rapport à la mesure de Lebesgue de $\mathbb{R}^d$:

<u>Définition</u> : on dit que le vecteur aléatoire $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^d, \mathcal{B}_d)$ a une <u>densité de probabilité</u> $f : \mathbb{R}^d \to \mathbb{R}_+$ si sa loi de probabilité μ_X est de la forme : $\forall B \in \mathcal{B}_d$

$$\mu_X(B) = \int_B f(x_1, \dots, x_d)\, dx_1 \dots dx_d \quad (*)$$

La fonction densité $f : \mathbb{R}^d \to \mathbb{R}_+$ est une fonction borélienne positive et $\int_{\mathbb{R}^d} f(x_1, \dots, x_d)\, dx_1 \dots dx_d = 1$

(Noter que $(*)$ est une intégrale de Lebesgue. La densité n'est pas unique au sens strict mais si g est une autre densité alors on montre facilement (exercice de L3) que $f = g$ presque partout pour la mesure de Lebesgue).

<u>Remarque</u> : Il existe des lois de probabilité qui ne sont ni discrètes ni absolument continues par exemple un <u>mélange</u> de ces deux types ($\mu_X = p\mu_X^{(1)} + (1-p)\mu_X^{(2)}$ où $p > 0$, $\mu_X^{(1)}$ probabilité discrète et $\mu_X^{(2)}$ probabilité absolument continue) mais il existe des lois de probabilité beaucoup plus compliquées (par exemple non discrète mais

quand même singulière par rapport à la mesure de Lebesgue). Dans les applications on se restreindra aux deux cas : discret ou à densité de probabilité.

(B) Cas des variables aléatoires - Fonction de répartition.

Dans le cas $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}_1)$ d'une variable aléatoire, la loi de probabilité de X qui est une <u>mesure</u>, peut être caractérisée par une fonction (c'est un objet <u>plus simple</u> qu'une mesure !)

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

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

<u>Proposition</u>: F_X satisfait les propriétés suivantes :

1) $0 \leq F_X(t) \leq 1$
2) F_X est croissante (au sens large) et <u>continue à droite</u> en tout point
3) $\lim_{t \to -\infty} F_X(t) = 0$ et $\lim_{t \to +\infty} F_X(t) = 1$

<u>dém^n</u> : 1) clair

2) si $t_1 < t_2$ alors $]-\infty, t_1] \subset]-\infty, t_2]$ donc

$\mu_X(]-\infty, t_1]) \leq \mu_X(]-\infty, t_2])$ i.e. $F_X(t_1) \leq F_X(t_2)$.

Soit $\varepsilon_n \downarrow 0$. Pour tout $t \in \mathbb{R}$, on a

$$\mu_X(]t, t+\varepsilon_n]) = \mu_X(]-\infty, t+\varepsilon_n]) - \mu_X(]-\infty, t])$$
$$= F_X(t+\varepsilon_n) - F_X(t)$$

Or $]t, t+\varepsilon_n] \downarrow \emptyset$ donc $\mu_X(]t, t+\varepsilon_n]) \to 0$ (résultat du chapitre 2) donc F continue à droite

3) facile (exercice).

<u>Théorème (admis)</u> : A toute fonction de répartition correspond une et une seule mesure de probabilité μ sur $(\mathbb{R}, \mathcal{B}_1)$ telle que

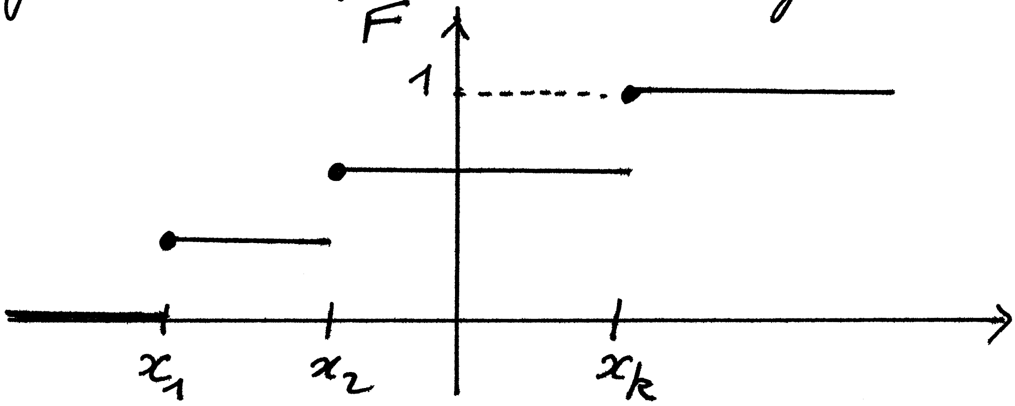
$$\forall t \in \mathbb{R}, \quad F(t) = \mu(]-\infty, t])$$

<u>Conséquence</u> : La fonction de répartition d'une v.a. X caractérise parfaitement la loi μ_X de X.

<u>Cas particuliers</u> 1) Si X a une densité f, la fonction de répartition est de la forme

$$F(t) = \int_{-\infty}^{t} f(x)\,dx .$$

2) Si X est discrète et si l'ensemble $X(\Omega)$ des valeurs prises par X est constitué de points isolés, la fonction de répartition est étagée :



et le saut de F en x_k est égal à

$$F(x_k^+) - F(x_k^-) = \mathbb{P}(X = x_k).$$

Exercice : Montrer que si la fonction de répartition F de la v.a. X est de classe $\mathcal{C}^1$ alors X a une densité

$$f(x) = \frac{dF}{dx}(x) \quad (x \in \mathbb{R}).$$

III) Densités de probabilité

A) Exemples des densités de probabilité des v.a. les plus classiques

1) La densité uniforme sur un intervalle compact $[a, b]$:

$$f(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$$

Une v.a. X ayant une telle densité est appelée v.a. uniforme sur $[a,b]$. On dit aussi que X est un nombre au hasard entre a et b.

Par définition quand on dit qu'on tire un nombre au hasard entre a et b, c'est qu'on considère une v.a. X uniforme sur $[a,b]$.

2) la densité de Gauss ou densité normale centrée réduite est la fonction

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

La densité normale de paramètres (m, σ^2) est la fonction

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

La loi d'une v.a. X ayant une telle densité est notée $N(m, \sigma^2)$.

3) La densité exponentielle de paramètre $\lambda > 0$ est la fonction

$$f(x) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x).$$

Ⓑ Densités de probabilité des vecteurs aléatoires densités marginales

Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^d, \mathcal{B}_d)$ un vecteur aléatoire. Dans la base canonique de $\mathbb{R}^d$, écrivons

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

la i^e coordonnée X_i de X est une v.a. et on écrira $X = (X_1, X_2, \dots, X_d)$.

Proposition : Si X a une densité de probabilité $f(x_1, \dots, x_d)$, alors la i^e coordonnée de X admet aussi une densité (appelée densité marginale d'ordre i) qui est la fonction définie sur $\mathbb{R}$:

$$f_i(x) = \int_{\mathbb{R}^{d-1}} f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_d)\, dx_1 \dots dx_{i-1}\, dx_{i+1} \dots dx_d$$

dém^n : supposons $d=2$ pour alléger les notations et soit $X = (X_1, X_2)$ de densité $f(x_1, x_2)$.

Supposons $i=1$ (le même raisonnement marche si $i=2$) et soit $A \in \mathcal{B}_1$. On a

$$\mathbb{P}(X_1 \in A) = \mathbb{P}(X_1 \in A, X_2 \in \mathbb{R}) = \mathbb{P}(X \in \tilde{A})$$

où $\tilde{A} = A \times \mathbb{R}$. Par définition de la densité f, on a donc

$$\mathbb{P}(X_1 \in A) = \iint_{\tilde{A}} f(x_1, x_2)\, dx_1\, dx_2 =$$

$$\int_A \left(\int_{\mathbb{R}} f(x_1, x_2)\, dx_2 \right) dx_1 = \int_A f_1(x_1)\, dx_1 ,$$

ce qui montre que X_1 a une densité f_1 donné par la formule annoncée CQFD.

<u>Remarque</u> : Attention la réciproque de la proposition est fausse ! On peut avoir un vecteur aléatoire $X = (X_1, X_2)$ tel que X_1 et X_2 ont des densités de probabilité mais X n'a pas de densité !. Par exemple soit X_1 une v.a. ayant une densité f alors le vecteur $X = (X_1, 2X_1)$ n'a pas de densité (exercice).

<u>Exercice</u> : Soit $V = (X, Y)$ un couple de v.a. (i.e. un vecteur aléatoire de $\mathbb{R}^2$) de loi uniforme sur le disque centré à l'origine et de rayon 1 (i.e. V a une densité de probabilité

$$f(x,y) = \begin{cases} \frac{1}{\pi} & \text{si } x^2+y^2 \leq 1 \\ 0 & \text{sinon} \end{cases}$$

Montrer que les lois marginales ont des densités de la forme :

$$f_X(x) = \begin{cases} 0 & \text{si } x \notin [-1,1] \\ \frac{2}{\pi}\sqrt{1-x^2} & \text{si } x \in [-1,1] \end{cases}$$

$$f_Y(y) = \begin{cases} 0 & \text{si } y \notin [-1,1] \\ \frac{2}{\pi}\sqrt{1-y^2} & \text{si } y \in [-1,1] \end{cases}$$

IV) Variables aléatoires indépendantes

Soient $X_i : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}_1)$ $(i = 1, \dots, n)$ des v.a.

<u>Définition</u> : On dit que les v.a. sont <u>indépendantes</u> si quel que soient les boréliens $A_1, \dots, A_n$, on a

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n)$$
$$= \prod_{i=1}^{n} \mathbb{P}(X_i \in A_i)$$

(la virgule signifiant l'intersection des événements)

<u>Exemple</u> : Comme dans le cas discret, des variables aléatoires relatives à des expériences aléatoires indépendantes, sont indépendantes. La justification donnée au chapitre 2 s'applique dans le cas général aussi.

<u>Remarque</u> : On va caractériser la propriété d'indépendance à l'aide de la loi des v.a. X_i et de celle du vecteur $X = (X_1, \dots, X_n)$ dont les coordonnées dans la base canonique de $\mathbb{R}^n$ sont les X_i :

Désignons par μ_i la loi de probabilité de X_i (c'est une mesure de probabilité sur $\mathbb{R}$) et μ_X la loi de probabilité du vecteur X (appelée <u>loi conjointe</u> des v.a. X_i) qui est une proba sur $(\mathbb{R}^n, \mathcal{B}_n)$.

<u>Théorème</u> : les v.a. $X_1, \dots, X_n$ sont indépendantes si et seulement si

$$\mu_X = \mu_1 \otimes \mu_2 \otimes \cdots \otimes \mu_n$$

(<u>mesure produit</u> des mesures $\mu_1, \dots, \mu_n$)

<u>En particulier</u> <u>si</u> la loi conjointe μ_X a une densité $f(x_1, \dots, x_n)$ et si f_i désigne la densité marginale d'ordre i (i.e. de X_i), les v.a. X_i sont indépendantes si et seulement si

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

<u>dém</u>^{n} : (on suppose $n=2$ pour simplifier l'écriture)

Soit $R = A_1 \times A_2$ avec $A_i \in \mathcal{B}_1$ $(i=1,2)$, un rectangle mesurable de $\mathbb{R}^2$.

($\Rightarrow$) on a $\mu_X(R) = \mathbb{P}(X \in R) = \mathbb{P}(X_1 \in A_1, X_2 \in A_2)$

$$= \mu_1(A_1)\mu_2(A_2) \text{ (par hypothèse)}$$

μ_X coïncide avec $\mu_1 \otimes \mu_2$ sur les rectangles donc $\mu_X = \mu_1 \otimes \mu_2$

($\Leftarrow$) supposons $\mu_X = \mu_1 \otimes \mu_2$. Alors pour $A_1, A_2 \in \mathcal{B}_1$

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2) = \mu_X(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) =$$

$= \mathbb{P}(X_1 \in A_1)\mathbb{P}(X_2 \in A_2)$ donc X_1 et X_2 sont indépendantes

CQFD.

L'assertion concernant le cas où μ_X a une densité résulte de la 1^{re} partie du théorème et du fait que si $\mu_i = f_i(x_i)dx_i$ alors

$$\mu_1 \otimes \cdots \otimes \mu_n = f_1(x_1)\cdots f_n(x_n)\,dx_1 \cdots dx_n .$$

<u>Corollaire</u> Soient X et Y des v.a. discrètes prenant les valeurs x_i (resp. y_j) avec la probabilité $p_{i\bullet}$ (resp $p_{\bullet j}$) et soit $p_{ij} = \mathbb{P}(X=x_i, Y=y_j)$. Alors X et Y sont indépendantes si et seulement si

$$\forall i,j \quad p_{ij} = p_{i\bullet}\, p_{\bullet j}$$

<u>dém^n</u>: résulte du théorème et du fait que si $\mu_1 = \sum_i p_{i\bullet}\,\delta_{x_i}$ et $\mu_2 = \sum_j p_{\bullet j}\,\delta_{y_j}$, on a

$$\mu_1 \otimes \mu_2 = \sum_{(i,j)} p_{i\bullet}\, p_{\bullet j}\, \delta_{(x_i, y_j)}$$

(on a utilisé la distributivité du produit tensoriel et le fait que $\delta_{x_i} \otimes \delta_{y_j} = \delta_{(x_i, y_j)}$) <u>CQFD</u>.

<u>Exercice</u> (transformation déterministe de v.a. indépendantes) Soient X_i ($i=1,\dots,d$) des v.a. indépendantes et $g_i : \mathbb{R} \to \mathbb{R}$ des fonctions boréliennes. Montrer que les v.a. $Y_i = g(X_i)$ sont indépendantes.