Cours de probabilité - 3IF

Victor Lezaud

23 avril 2018

Sommaire

1 Notions de dénombrement

Dans cette partie on traite la combinatoire de p éléments pris parmi n éléments d'un ensemble E (sauf pour les permutations).

1.1 Permutations

On appelle **permutation** des n éléments l'ensemble E toute disposition ordonnée de ces n éléments. Il existe $n!$ permutations.

1.2 Arrangements sans répétition

On appelle **arrangement sans répétition** toute disposition **ordonnée** de p éléments pris parmi les n éléments de E. Il existe A_n^p arrangements sans répétition.

$$A_n^p = \frac{n!}{(n-p)!}$$

1.3 Combinaisons sans répétition

On appelle **combinaisons sans répétition** toute disposition **non ordonnée** de p éléments pris parmi les n éléments de E. Il existe C_n^p combinaisons sans répétition.

$$C_n^p = \frac{n!}{p!(n-p)!}$$

1.4 Arrangements avec répétition

On appelle **arrangements avec répétition** toute disposition **ordonnée** de p éléments, **non nécessairement distincts**, pris parmi les n éléments de E. Il existe n^p arrangements avec répétition.

1.5 Combinaisons avec répétition

On appelle **combinaisons avec répétition** toute disposition **non ordonnée** de p éléments, **non nécessairement distincts**, pris parmi les n éléments de E. Il existe C_{n+p-1}^p combinaisons sans répétition.

$$C_{n+p-1}^p = \frac{(n+p-1)!}{p!(n-1)!}$$

1.6 Permutations avec répétitions

Supposons que les éléments n éléments de l'ensemble E, se répartissent en l catégories avec n_i le nombre d'éléments du type i et $n = \sum n_i$. On appelle permutation avec répétition de n élémens de l'ensemble E toute disposition ordonnée de n éléments où figure, pour tout i, n_i éléments du type i. Il en existe :

$$\frac{n!}{\prod n_i!}$$

2 Fondements de la théorie des probabilités

2.1 Espace d'événements

2.1.1 Univers

On appelle **ensemble des possibles**, ou **ensemble des éventualités**, ou encore **univers**, l'ensemble Ω des résultats possibles d'une expérience aléatoire. Les éléments de ω de Ω sont appelés **issues**, ou **événements élémentaires**, ou encore **épreuves** de l'expérience aléatoire.

2.1.2 Evénement

On appelle événement l'**ensemble des issues** de l'expérience qui vérifie une **propriété** donnée. C'est donc une partie A de Ω ($A \in P(\Omega)$).

2.1.3 Tribu

On associe à l'univers Ω un ensemble d'événements T appelé **tribu**. Cet ensemble doit vérifier les propriétés suivantes :
- $\Omega \in T$
- $A \in T \rightarrow A^c \in T$
- $\forall i \in \mathbb{N}, A_i \in T \rightarrow \cup_{i \in \mathbb{N}} A_i \in T$

Dans ce cas, alors le couple (Ω, T) est appelé **espace d'événements**. Lorsque Ω est fini et dénombrable on prend pour tribu l'ensemble $P(\Omega)$.

2.2 Espace probabilisé

2.2.1 Définition

Soit (Ω, T) un espace d'événements. On appelle **probabilité** sur (Ω, T) une application P définie sur T vérifiant :

1. $\forall A \in T, P(A) \in [0, 1]$
2. $P(\Omega) = 1$
3. Quelle que soit la suite $(A_i)_{i \in \mathbb{N}}$ d'éléments de T deux à deux disjoints, on a
$$P\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i=1}^{+\infty} P(A_i)$$

Le triplet (Ω, T, P) est appelé **espace probabilisé**.

2.2.2 Propriétés

Soit A et B deux événements d'un espace probabilisé (Ω, T, P). La probabilité P est une application qui vérifie les propriétés suivantes :
- $P(A^c) = 1 - P(A)$
- $A \subset B \Rightarrow P(A) < P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

2.3 Probabilités conditionnelles

2.3.1 Définition

Soit (Ω, T, P) un espace probabilisé, $A \in T, B \in T)$ tel que $P(B) \neq 0$. La **probabilité de A sachant B** est définie par

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

L'application

$$P_B : A \in T \longmapsto P(A|B)$$

définit une probabilité sur l'espace des événements (Ω, T). Cette probabilité est appelée **probabilité conditionnelle sachant B**.

2.3.2 Propriétés

Soient $A, B \in T$. On a les relations suivantes :

— $P(A^c|B) = 1 - P(A|B)$

— $P(A \cap B) = P(A|B) \times P(B) = P(B|A) \times P(A)$

2.3.3 Formule de Bayes

Une partition d'un univers est un **ensemble fini d'événement disjoint** deux à deux et dont l'**union forme l'univers**.

Théorème : Soit (Ω, T, P) un espace probabilisé et soient $A_1, \ldots, A_n$ une **partition** de Ω. On a les relations suivantes :

$$\forall B \in T \quad P(B) = \sum_{i=1}^{n} P(B|A_i) \times P(A_i)$$
$$\forall B \in T \quad P(A_k|B) = \frac{P(B|A_k) \times P(A_k)}{\sum_{i=1}^{n} P(B|A_i) \times P(A_i}$$

2.3.4 Indépendance

Définition : Soient (Ω, T, P) un espace probabilisé et A et B deux éléments de T. Les événements A et B sont dits **indépendants** si et seulement si

$$P(A \cap B) = P(A) \times P(B)$$
$$P(A|B) = P(A)$$

Complémentaires : Si A et B ont **indépendants**, alors il en est de même pour A et B^c, A^c et B et A^c et B^c.

Indépendance mutuelle : Les événements $A_1, \ldots, A_n \in T$ sont dits **mutuellement indépendants** si $\forall k \in \{1, \ldots, n\}$ et $\forall (i_1, \ldots, ik) \in \mathbb{N}^k$ tel que $1 \leqslant i_1 < \ldots \leqslant n$, on a :

$$P(\cap_{j=1}^{k} A_{i_j}) = \prod_{j=1}^{k} P(A_{i_j})$$

3 Variables aléatoires réelles

3.1 Définition

3.1.1 Variable aléatoire

Variable aléatoire : Soient (Ω, T) et (Ω', T') deux espaces d'événements. On appelle **variable aléatoire** de l'espace d'événements (Ω, T) vers (Ω', T') **une application** $X : \Omega \mapsto \Omega'$ telle que :

$$\forall A' \in T', \ X^{-1}(A') \in T$$

où $X^{-1}(A') = \{\omega \in \Omega, X(\omega) \in A'\}$

Variable aléatoire réelle : Une application $X : \Omega \mapsto \mathbb{R}$ est une **variable aléatoire réelle** sur l'espace d'événements (Ω, T) si :

$$\forall x \in \mathbb{R}, \ X^{-1}\left(] - \infty, x]\right) \in T$$

3.1.2 Fonction de répartition

Définition : Soit X une v.a.r. sur l'espace probabilisé (Ω, T, P). L'application

$$\begin{aligned} F_x : &\mathbb{R} \longmapsto [0,1] \\ &x \longmapsto P\left(X^{-1}\left(]-\infty, x]\right)\right) \end{aligned}$$

est appelé **fonction de répartition**.

Proposition : Soit X une v.a.r. sur (Ω, T, P). La **fonction de répartition** associée à X vérifie les propriétés suivantes :
- $\lim\limits_{x \to -\infty} F_X(x) = 0$
- $\lim\limits_{x \to +\infty} F_X(x) = 1$
- F_X est croissante
- F_X est continue à droite

3.1.3 Variables aléatoires discrètes et variables aléatoires continues

V.a.r discrète : Une v.a.r. X est dite **discrète** si elle ne prend qu'un **nombre fini ou dénombrable de valeurs** dans $\mathbb{R}$, c'est-à-dire si $X(\Omega) = \{x_j \in \mathbb{R}, j \in J\}$ où $J \subset \mathbb{N}$. Dans ce cas la fonction

$$\begin{aligned} p : &J \longmapsto [0,1] \\ &j \longmapsto p_j \end{aligned}$$

où $p_j = P(X = x_j)$ est appelée **fonction de masse** de la v.a.r. C.

Proposition : Si X est une **v.a.r. discrète** prenant les valeurs $\{x_i \in \mathbb{R}, i \in I\}$ avec $I \subset \mathbb{N}$, alors sa **fonction de répartition** F_X est **constante par morceaux** et a pour points de discontinuités $\{x_i \in \mathbb{R}, i \in I\}$.

V.a.r. absolument continue : Une v.a.r. X est dite **absolument continue** s'il existe une fonction de $\mathbb{R}$ dans $\mathbb{R}$, telle que la **fonction de répartition** F_X de la v.a.r. X admette la représentation **intégrale** suivante :

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

Ceci est équivalent à dire que F_X est **dérivable** sur $\mathbb{R}$, de dérivée f_X. La fonction f_X est appelée **densité** X. On parle aussi de **v.a.r. à densité** pour désigner une **v.a.r. absolument continue**.

3.1.4 Loi d'une variable aléatoire

Soient (Ω, T) et (Ω', T') deux espaces d'événements. Soit X une variable aléatoire de (Ω, T) vers (Ω', T') et P une probabilité sur (Ω, T). L'**application**

$$\begin{aligned} P_X :& T' \longmapsto [0,1] \\ & A' \longmapsto P(X^{-1}(A')) \end{aligned}$$

définit une **probabilité** sur l'espace d'événements $(\Omega' T')$. Cette probabilité est appelée **loi de la variable aléatoire** X.

3.2 Exemples de lois

3.2.1 Variables aléatoires discrètes

Loi de Bernoulli notée $B(1,p)$, $p \in]0,1[$, permet de décrire une expérience possédant uniquement **deux issues** possibles (exemple : pile ou face, succès ou réussite). La loi es définie par :

$$P(X = 1) = p \quad et \quad P(X = 0) = 1 - p$$

Loi binomiale notée $B(1,p)$, $n \in \mathbb{N}^*$, $p \in]0,1[$ permet de décrire une expérience composée d'une **suite d'épreuve de Bernoulli** où l'on s'intéresse au **nombre de réussite** (exemple : nombre de face en n lancers). La loi est définie par :

$$\forall i \in \{0, \ldots, n\}, \quad P(X = i) = C_n^i p^i (1-p)^{n-i}$$

Loi géométrique notée $G(p)$, $p \in]0,1[$ décrit une expérience composée d'une **suite d'épreuve de Bernoulli** où l'on s'intéresse à la probabilité d'une **série d'échecs suivie d'une réussite** (Exemple : nombre de piles avant une face). La loi est définie par :

$$\forall i \in \mathbb{N}^*, \quad P(X = i) = p(1-p)^{i-1}$$

Loi uniforme sur $\{1, \ldots, \mathbb{N}\}$ notée $U_{\mathbb{N}}, \mathbb{N} \in \mathbb{N}^*$ décrit une expérience possédant $\mathbb{N}$ **issues possibles, toutes équiprobables** (Exemple : lancer de dé). La loi est définie par :

$$\forall k \in \{1, \ldots, \mathbb{N}\}, \quad P(X = k) = \frac{1}{\mathbb{N}}$$

Loi de Poisson notée $P(\lambda)$, $\lambda > 0$ permet de décrire les **files d'attentes**. La loi est définie par :

$$\forall k \in \mathbb{N}, P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

3.2.2 Variables aléatoires continues

Loi uniforme sur $[a, b]$ notée $U([a, b])$, $a < b$ a pour densité la fonction

$$f_X(t) = \frac{1}{b-a} I_{[a,b]}(t)$$

et pour fonction de répartition

$$F_X = \begin{cases} 0 & \text{si } x \leqslant a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x \geqslant b \end{cases}$$

Loi normale notée $\mathbb{N}(m, \sigma^2)$, $m \in \mathbb{R}$, $\sigma > 0$ a pour densité

$$f_X(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right)$$

Loi exponentielle notée $\varepsilon(\lambda)$, $\lambda > 0$ a pour densité :

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

Loi de Cauchy a pour densité

$$f_X(t) = \frac{1}{\pi(1+t^2)}$$

Loi Gamma notée $\Gamma(a, b)$, $a > 0$, $b > 0$ a pour densité sur $\mathbb{R}^+$ la fonction

$$f_X(t) = \frac{1}{\Gamma(a)} b^a t^{a-1} e^{-bt}$$

où

$$\Gamma(a) = \int_0^{+\infty} u^{a-1} e^{-u} du, \quad \forall a > 0$$

3.3 Moments d'une variable aléatoire

3.3.1 Espérance

Définition Soit X une v.a.r. sur l'espace probabilisé (Ω, T, P).

— Si X est une v.a.r. **discrète** avec $X(\Omega) = \{x_j \in \mathbb{R}, j \in J\}$ où $J \subset \mathbb{N}$, on appelle **espérance** de X et on note $E(X)$, la **moyenne des valeurs** prises par X **pondérés par leurs probabilités** de réalisation, autrement dit, lorsque cette quantité existe

$$E(X) = \sum_{j \in J} x_j \times P(X = x_j)$$

— Si X est une v.a.r. **continue** de densité f_X, on appelle **espérance** de X et on note $E(X)$ la quantité

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

Proposition

— Soit X une **v.a.r. discrète** sur l'espace probabilisé (Ω, T, P) avec $X(\Omega) = \{x_j \in \mathbb{R}, j \in J\}$ où $J \subset \mathbb{N}$, et $g : X(\Omega) \mapsto \mathbb{R}$ une application. Lorsque cette quantité existe, **l'espérance de la v.a.r.** $g(X)$ est définie par

$$E(g(X)) = \sum_{j \in J} g(x_j) \times P(X = x_j)$$

— Soit X une **v.a.r. absolument continue** sur l'espace probabilisé (Ω, T, P) de densité f_X et $g : \mathbb{R} \mapsto \mathbb{R}$ une **application continue par morceaux**. Lorsque cette quantité existe, **l'espérance de la v.a.r.** $g(X)$ est définie par

$$E(g(X)) = \int_{\mathbb{R}} g(t) \times f_X(t)dt$$

Propriété de l'espérance :

Linéarité Soient X une **v.a.r. continue**, g_1 et g_2 deux **fonctions** de $\mathbb{R}$ dans $\mathbb{R}$ **continues par morceaux** et λ un réel. On a

$$E(g_1(X) + \lambda g_2(X)) = E(g_1(X)) + \lambda E(g_2(X))$$

Positivité Soient X une **v.a.r. continue** et g une fonction de $\mathbb{R}$ dans $\mathbb{R}$ **positive et continue par morceaux**. On a

$$E(g(X)) \geqslant 0$$

Croissance Soient X une **v.a.r. continue** g_1 et g_2 deux fonctions de $\mathbb{R}$ dans $\mathbb{R}$ **continues par morceaux** vérifiant $g_1 \leqslant g_2$. On a

$$E(g_1(X)) \leqslant E(g_2(X))$$

3.3.2 Variance

La variance de la v.a.r. X est définie, lorsque cette quantité existe, par

$$\begin{aligned} Var(X) &= E\left((X - E(X))^2\right) \\ &= E(X^2) - E(X)^2 \end{aligned}$$

Propriétés de la variance Soit X une v.a.r. et λ un réel

— $Var(X + \lambda) = Var(X)$

— $Var(\lambda X) = \lambda^2 Var(X)$

— $Var(X) \geqslant 0$

Variable centrée réduite Si X est une v.a.r., pour laquelle $E(X)$ et $Var(X)$ existent, alors la v.a.r. Y définie par

$$Y = \frac{X - E(X)}{\sqrt{Var(X)}}$$

appelée **variable centrée réduite associée à la v.a.r. X**, vérifie :

$$E(Y) = 0$$
$$Var(Y) = 1$$

L'écart-type d'une v.a.r. X est défini par

$$\sigma_X = \sqrt{Var(X)}$$

3.3.3 Espérance et variance pour les v.a.r. de lois usuelles

Nom de loi	Notation math	Espérance	Variance
Uniforme (discrète)	$U_N,\ N \in \mathbb{N}^*$	$\frac{N+1}{2}$	$\frac{N^2-1}{12}$
Bernoulli	$B(1,p),\ p \in]0,1[$	p	$p(1-p)$
Géométrique	$G(p), p \in]0,1[$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Uniforme (continue)	$U(a,b),\ a < b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normale	$N(m,\sigma^2),\ m,\sigma \in \mathbb{R}$	m	σ^2

3.3.4 Inégalités remarquables

Inégalité de Markov Soit X une v.a.r.. Pour tout réel a **strictement positif**, on a

$$P(|X| \geqslant a) \leqslant E(|X|)$$

Inégalité de Bienaymé-Tchébychev Soit X une v.a.r. et a un réel **strictement positif**, on a

$$P(|X - E(X)| \geqslant a) \leqslant \frac{Var(X)}{a^2}$$

3.4 Caractérisation de la loi d'une variable aléatoire

La loi d'une v.a.r. est caractérisée par sa fonction de densité (ou masse pour les v.a.r. discrètes). Il existe d'autre manière de caractériser la loi d'une v.a.r..

3.4.1 Caractérisation par la fonction de répartition

La **fonction de répartition est suffisante pour caractériser la loi** de la v.a.r. associée.

3.4.2 Autre caractérisation

Soit X une v.a.r. admettant une densité f_X. S'il existe, **pour toute fonction Φ bornée et continue par morceaux** sur $\mathbb{R}$, une fonction f **ne dépendant pas** de Φ telle que l'espérance de $\Phi(X)$ admette la formule de représentation intégrale

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

alors $f = f_X$, autrement dit f **est la densité de** X.

3.4.3 Fonction caractéristique

Définition : On appelle **fonction caractéristique** de la v.a.r. X la fonction de $\mathbb{R}$ dans $\mathbb{C}$ définie par

$$\Phi_X(t) = E(e^{itX}) \quad \forall t \in \mathbb{R}$$

Théorème : Deux v.a.r. définies sur un même espace probabilisé qui ont la **même fonction caractéristique** ont la **même loi**.

Propriétés de la fonction caractéristique

— Φ_X est une **application continue** à valeurs dans $\{z \in \mathbb{C}, |z| \leqslant 1\}$ et on a $\Phi_X(0) = 1$

— Si Φ_X est **deux fois dérivable** en 0, alors $E(X)$ et $E(X^2)$ existent et

$$E(X) = -i\Phi'_X(0)$$
$$E(X^2) = -\Phi''_X(0)$$

— Soient a et b deux réels et Y la v.a.r. définie par $Y = aX + b$. La fonction caractéristique Φ_Y de la v.a.r. Y vérifie

$$\Phi_Y(t) = e^{itb}\Phi_X(at) \quad \forall t \in \mathbb{R}$$

3.4.4 Fonction génératrice

Définition : Soit X une v.a.r.**discrète à valeurs entières**. On appelle **fonction génératrice** de X la fonction définie pour tout $s \in [-1, 1]$ par

$$G_X(s) = E(s^X) = \sum_{j \in \mathbb{N}} s^j \times P(X = j)$$

Théorème : Deux v.a.r. entières définies sur un même espace probabilisé qui ont la **même fonction génératrice** ont la **même loi**.

Propriétés de la fonction génératrice

— La série $\sum_{j\in\mathbb{N}} s^j \times P(X = j)$ **converge** pour tout $s \in [-1, 1]$

— G_X est **continue** sur $[-1, 1]$ et **infiniment dérivable** sur $]-1, 1[$

— $G_X(1) = 1$, $G_X(0) = P(X = 0)$

— $G'_X(1) = E(X)$

— $\forall k \in \mathbb{N}^*, \quad G_X^{(k)} = E(X \times (X-1) \times \cdots \times (X-k+1)$

Nom de loi	Notation math	fct caractéristique	fct génératrice
Binomiale	$B(n,p)$	$(pe^{it} + (1-p))^n$	$(ps + (1-p))^n$
Poisson	$P(\lambda)$	$e^{\lambda(e^{it}-1)}$	$e^{\lambda(s-1)}$
Normale	$N(m,\sigma^2)$	$e^{itm}e^{-\frac{t^2\sigma^2}{2}}$	

4 Vecteurs aléatoires

4.1 Couple de variables aléatoires réelles

4.1.1 Couple aléatoire

Soient (Ω, T, P) un espace probabilisé et X l'application

$$\begin{aligned} X : &\Omega \longmapsto \mathbb{R}^2 \\ &\omega \longmapsto (X_1(\omega), X_2(\omega)) \end{aligned}$$

On dit que X est un **couple aléatoire** si pour tout borélien B de $\mathbb{R}^2$, $X^{-1}(B) \in T$.

4.1.2 Lois d'un couple de variables aléatoires réelles

Fonction de répartition : On appelle **fonction de répartition** du couple de v.a.r. $X = (X_1, X_2)$ l'application $F_X : \mathbb{R}^2 \mapsto \mathbb{R}$ définie par

$$F_X(x_1, x_2) = P(X_1 \leqslant x_1, X_2 \leqslant x_2) \quad \forall (x_1, x_2) \in \mathbb{R}^2$$

avec

$$\begin{aligned} P\left(X_1 \leqslant x_1, X_2 \leqslant x_2\right) &= P\left([X_1 \leqslant x_1] \cap [X_2 \leqslant x_2]\right) \\ P\left(X_1 \leqslant x_1, X_2 \leqslant x_2\right) &= P\left(\{\omega \in \Omega, X_1(\omega) \leqslant x_1, X_2(\omega) \leqslant x_2\}\right) \end{aligned}$$

Limite de la fonction de répartition Si $X = (X_1, X_2)$ est un couple de v.a.r. de fonction de répartition F_X alors on a

$$\lim_{x \to -\infty} F_X(x,x) = 0 \qquad et \qquad \lim_{x \to +\infty} F_X(x,x) = 1$$

Couple de v.a.r. discret : Le couple de v.a.r. $X = (X_1, X_2)$ défini sur l'espace probabilisé (Ω, T, P) est dit **discret** s'il est à valeurs dans un sous-ensemble $D = X(\Omega)$ **fini ou dénombrable** de $\mathbb{R}^2$. La fonction

$$\begin{aligned} D &\longrightarrow [0,1] \\ (k_1, k_2) &\longmapsto P(X_1 = k_1, X_2 = k_2) \end{aligned}$$

est alors appelée **fonction de masse** du couple de v.a.r. $X = (X_1, X_2)$.

Couple de v.a.r. absolument continu Le couple de v.a.r. $X = (X_1, X_2)$, est dit **absolument continu** de densité f_X si sa **fonction de répartition** F_X admet la **représentation intégrale** suivante :

$$F_X(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_X(t_1, t_2) dt_2 dt_1$$

Loi marginale : On appelle k^{ieme} loi marginale du couple de v.a.r. $X = (X_1, X_2)$ la loi de la variable aléatoire réelle X_k.

4.1.3 Moments d'un couple de variables aléatoires réelles

Espérance : On appelle espérance du couple de v.a.r. (X_1, X_2) l'élément de $\mathbb{R}^2$ défini par

$$E(X_1, X_2) = (E(X_1), E(X_2))$$

Covariance : Soit (X_1, X_2) un couple de v.a.r.. Si $E(X_1)$ et $E(X_2)$ existent, on appelle **covariance** du couple de v.a.r. (X_1, X_2) le réel

$$Cov(X_1, X_2) = E\left((X_1 - E\left(X_1\right)) \times (X_2 - E\left(X_2\right))()\right.$$

Propriétés de la covariance :

- $Cov(X_1, X_1) = Var(X_1)$
- $Cov(X_1, X_2) = E(X_1, X_2) - E(X_1) \times E(X_2)$
- L'application $(X_1, X_2) \mapsto Cov(X_1, X_2) \in \mathbb{R}$ est une **forme bilinéaire symétrique**. Ainsi pour toutes v.a.r. X_1, X_2, Y_1 et pour tout réel λ on a :

$$Cov(X_1, X_2) = Cov(X_2, X_1)$$
$$Cov(X_1 + \lambda Y_1, X_2 = Cov(X_1, X_2) + \lambda Cov(Y_1, X2)$$

4.1.4 Fonction caractéristique et fonction génératrice

Fonction caractéristique : On appelle **fonction caractéristique** du couple de v.a.r. $X = (X_1, X_2)$ la fonction Φ_X définie de $\mathbb{R}^2$ dans $\mathbb{C}^{\not\vDash}$ par

$$\Phi_X(t_1, t_2) = E\left(e^{i(t_1 X_1 + t_2 X_2)}\right)$$

Fonction génératrice : On appelle **fonction génératrice** du couple de v.a.r. $X = (X_1, X_2)$ la fonction G_X définie de $\mathbb{R}^2$ dans $\mathbb{R}$ par

$$G_X(s_1, s_2) = E\left(s_1^{X_1} s_2^{X_2}\right)$$

4.1.5 Variables aléatoires conditionnelles

Cas discret : Soient (X, Y) un couple de v.a.r. **discrètes** à valeurs dans $\mathbb{N}^2$ et j un entier tel que $P(Y = j) \neq 0$. On appelle **loi conditionnelle de** X **sachant** $[Y = j]$ la loi définie par la **fonction de masse**

$$p : i \in \mathbb{N} \longmapsto P(X = i | Y = j)$$

On appelle **fonction de répartition conditionnelle de X sachant** $[Y = j]$ l'application $F_X^{[Y=j]}$ de $\mathbb{R}$ dans $[0, 1]$ définie pour tout $x \in \mathbb{R}$ par

$$\begin{aligned} F_X^{[Y=j]}(x) &= P(X \leqslant x | Y = j) \\ &= \frac{P(X \leqslant x, Y = j)}{P(Y = j)} \end{aligned}$$

Cas continu : Soient (X, Y) un couple de v.a.r. de densité $f_{X,Y}$ et f_Y la densité de Y. Pour $y \in \mathbb{R}$ fixé tel que $f_Y(y) \neq 0$, la fonction $f_X^{[Y=y]}$ définie par

$$f_X^{[Y=y]}(s) = \frac{f_{X,Y}(s, y)}{f_Y(y)}$$

est une densité de probabilité appelée **densité conditionnelle de X sachant** $[Y = y]$

4.2 Indépendance de 2 variables aléatoires réelles

4.2.1 Définition

Soient X et Y deux v.a.r., X et Y sont dites **indépendantes** si et seulement si :

— $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \quad P(X \leqslant x, Y \leqslant y) = P(X \leqslant x) \times P(Y \leqslant y)$

— Avec les **fonctions caractéristiques** :

$$\forall (s, t) \in \mathbb{R}^2, \quad \Phi_{X,Y}(s, t) = \Phi_X(s) \times \Phi_Y(t)$$

— Avec les **fonctions génératrices** :

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

— Pour des v.a.r. **discrètes** :

$$\forall i \in \mathbb{N}, \forall j \in \mathbb{N}, \quad P(X = i, Y = j) = P(X = i) \times P(X = j)$$

— Pour des v.a.r. **absolument continues** de densités marginales f_X et f_Y :

$$\forall s \in \mathbb{R}, \forall t \in \mathbb{R}, \quad f_{X,Y}(s, t) = f_X(s) \times f_Y(t)$$

4.2.2 Indépendance et moments

Soient X et Y deux v.a.r. **indépendantes** et g et h deux fonctions de $\mathbb{R}$ dans $\mathbb{R}$. On a :

— $E(XY) = E(X) \times E(Y)$

— $Var(X + Y) = Var(X) + Var(Y)$

— $E(g(X)h(Y)) = E(g(X)) \times E(h(Y))$

4.2.3 Coefficient de corrélation

Définition : On appelle **coefficient de corrélation** de X et Y le réel défini par :

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

Propriétés du coefficient de corrélation :

— Pour toutes v.a.r.X et Y on a $|\rho(X,Y)| \leqslant 1$

— si X et Y sont indépendantes, alors $\rho(X,Y) = 0$

— $\exists (a,b) \in \mathbb{R}^2,\ Y = aX + b \Leftrightarrow |\rho(X,Y)| = 1$

4.3 Somme de deux variables aléatoires indépendantes

4.3.1 Loi de la somme de deux variables aléatoires

Somme de v.a.r : Soient X et Y deux v.a.r. **indépendantes**. La loi de la v.a.r. $X+Y$ est obtenue en effectuant le **produit de convolution** des lois de X et de Y, autrement dit,

— si X et Y dont des v.a.r. **discrètes**, on a $\forall k \in \mathbb{N}$,

$$P(X+Y=k) = \sum_{i\in\mathbb{N}} P(X=k-i) \times P(Y=i)$$

— si X et Y sont des v.a.r. **continues** de densités respectives f_X et f_Y, on a $\forall x \in \mathbb{R}$,

$$f_{X+Y}(x) = \int_{\mathbb{R}} f_X(x-t) \times f_Y(t)dt$$

Fonction de la somme Soient X, Y deux v.a.r. **continues** de densités respectives f_X et f_Y et h une fonction **continue par morceaux**. On a

$$E(h(X+Y)) = \int_{\mathbb{R}} h(t) \times (f_X \otimes f_Y)(t)dt$$

4.3.2 Fonction caractéristique et génératrice de la somme de deux variables aléatoires indépendantes

Fonction caractéristique : Si X et Y sont deux v.a.r. **indépendantes** de fonctions caractéristiques Φ_X et Φ_Y alors la fonction caractéristique Φ_{X+Y} de la v.a.r. $X+Y$ st donnée par

$$\Phi_{X+Y}(t) = \Phi_X(t) \times \Phi_Y(t)$$

Fonction générique : Si X et Y sont deux v.a.r. **discrètes indépendantes** de fonctions génératrices G_X et G_Y alors la fonction génératrice G_{X+Y} de $X+Y$ est donnée par

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

5 Théorèmes limites

5.1 Convergences stochastiques

5.1.1 Convergence en loi

Définition : Soit X une v.a.r. de **fonction de répartition** F_X et $(X_n)_{n\in\mathbb{N}}$ une **suite de v.a.r.** de **fonctions de répartitions** respectives $(F_{X_n})_{n\in\mathbb{N}}$. On dit que la suite $(X_n)_{n\in\mathbb{N}}$ **converge en loi** vers X si pour tout réel x tel que F_X est continue en x, on a

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x)$$

On note alors

$$X_n \xrightarrow[n\to+\infty]{l} X$$

Proposition : La suite $(X_n)_{n\in\mathbb{N}}$ **converge en loi** vers X si et seulement si on a la relation suivante entre les **fonctions caractéristiques** Φ_X et Φ_{X_n} des v.a.r. X et X_n,

$$\forall t\in\mathbb{R},\quad \lim_{n\to+\infty} \Phi_{X_n}(t) = \Phi_X(t)$$

De manière équivalente la suite $(X_n)_{n\in\mathbb{N}}$ **converge en loi** vers X si et seulement si pour **toute fonction** h de $\mathbb{R}$ dans $\mathbb{R}$, **bornée et continue par morceaux**

$$\lim_{n\to+\infty} E(h(X_n)) = E(h(X))$$

5.1.2 Convergence en probabilité

Définition : Soient X une v.a.r. et $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. sur un même espace probabilisé (Ω, T, P). On dit que la suite $(X_N)_{n\in\mathbb{N}}$ **converge en probabilité** vers X si

$$\forall t>0,\quad \lim_{n\to+\infty} P(|X_n - X| \geqslant t) = 0$$

On note

$$X_n \xrightarrow[n\to+\infty]{P} X$$

Proposition : **Si** la suite $(X_N)_{n\in\mathbb{N}}$ converge vers X **en probabilité**, **alors** la suite $(X_N)_{n\in\mathbb{N}}$ converge vers X **en loi**.

5.1.3 Convergence presque sûre

Définition Soient X une v.a.r. et $(X_N)_{n\in\mathbb{N}}$ une **suite de v.a.r.** définies sur un même espace probabilisé (Ω, T, P). Soit $A\in T$ **l'ensemble des éventualités** $\omega\in\Omega$ telles que la suite numérique $(X_N(\omega)_{n\in\mathbb{N}}$ converge vers $X(\omega)$.

$$A = \{\omega\in\Omega,\ \lim_{n\to+\infty} X_n(\omega) = X(\omega)\} \quad \text{avec } A\in T$$

On dit que la suite $(X_N)_{n\in\mathbb{N}}$ **converge presque sûrement** vers X si $\mathbf{P(A) = 1}$.
On note

$$X_n \xrightarrow[n\to+\infty]{\text{p.s.}} X$$

5.2 Loi faible des grands nombres

Soient $X_1, \ldots, X_n$ n v.a.r. de **même loi et non corrélées**. On suppose que ces v.a.r. admettent une **espérance** m et une **variance** σ^2 **finies**.
La v.a.r. $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$, **converge en probabilité** vers m, noté

$$\overline{X_n} \xrightarrow[n\to+\infty]{P} m$$

5.3 Loi forte des grands nombres

Soient $X_1, \ldots, X_n$ n v.a.r. de **même loi et indépendantes**. On suppose que ces v.a.r.admettent une **espérance finie**.
La v.a.r. $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$, **converge presque sûrement** vers m

5.4 Théorème de la limite centrale

Théorème : Soient $X_1, \ldots, X_n$ n v.a.r. **indépendantes de même loi**, de moyenne m et de variance σ^2. La v.a.r. $S_n = \sum_{i=1}^{n} X_i$ vérifie,

$$\frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow[n\to+\infty]{L} Y$$

où $Y \sim N(0,1)$

Valeurs particulières :

Bernoulli : Soient $X_1, \ldots, X_n$ n v.a.r. **indépendantes de loi de Bernoulli** de paramètre p. La v.a.r. $S_n = \sum_{i=1}^{n} X_i$ suit une **loi binomiale** $B(n,p)$ et la suite $(S_n)_{n\in\mathbb{N}}$ **converge en loi** vers une **loi normale** de paramètres $(np, np(1-p))$.

Poisson : Soient $X_1, \ldots, X_n$ n v.a.r. **indépendantes de loi de Poisson** de paramètre λ. La v.a.r. $S_n = \sum_{i=1}^{n} X_i$ suit une **loi de Poisson** de paramètre $n\lambda$ et **converge en loi** vers une **loi normale** de paramètres $(n\lambda, n\lambda)$.

6 Chaînes de Markov discrètes

6.1 Chaîne de Markov homogène

6.1.1 Chaîne de Markov

Une **chaîne de Markov** sur (Ω, T, P) à valeurs dans (E, ε) est une **suite de v.a.r.** $(X_k)_{k\in\mathbb{N}}$ définies sur (Ω, T, P) et à valeurs dans (E, ε) vérifiant $\forall n \in \mathbb{N}$, $\forall (e_0, \ldots, e_{n+1}) \in E^{n+2}$,

$$P(X_{n+1} = e_{n+1}|X_n = e_n, \ldots, X_0 = 0) = P(X_{n+1} = e_{n+1}|X_n = e_n)$$

$$\Leftrightarrow$$

$$P(X_{n+1} = e_{n+1}, X_{n-1} = e_{n-1}, ..., X_0 = 0|X_n = e_n) = P(X_{n+1} = e_{n+1}|X_n = e_n) \times P(X_{n-1} = e_{n-1}, \ldots, X_0 = 0|X_n = e_n)$$

6.1.2 Chaîne de Markov homogène

Définition On dit qu'une chaîne $X = (X_k)_{k\in\mathbb{N}}$ est une **chaîne de Markov homogène** si pour tout l les **probabilités de transitions** en l étapes **sont les mêmes** indépendamment de l'instant n de la première transition. Autrement-dit si $\forall (n,l) \in \mathbb{N}^2, l < n, \forall (i,j) \in E^2$

$$P(X_n = j|X_{n-l} = i) = P(X_l = j|X_0 = i)$$

Matrice de transition : On appelle **matrice de transition** d'une **chaîne de Markov homogène** X la matrice $G = (G_{i,j})_{i,j \in E} \in M_N(\mathbb{R})$ d'éléments

$$G_{i,j} = P(X_1 = j | X_0 = i)$$

Propriétés de la matrice de transition : La **matrice de transition** est une matrice **stochastique**, elle vérifie donc :
- $G_{i,j} \geqslant 0 \quad \forall (i,j) \in E^2$
- $\forall i \in E, \sum_{j \in E} G_{ij} = 1$

Transition en n étapes : Soit $X = (X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène de matrice de transition G. La **probabilité de transition en n étapes** de l'état i à l'état j est donnée par,

$$P(X_n = j | X_0 = i) = (G^n)_{ij}$$

Caractérisation d'une chaîne de Markov

Equation de Chapman-Kolmogorov : Soit $X = (X_n)_{n \in \mathbb{N}}$ une **chaîne de Markov homogène** d'espace d'états E. $\forall i, j \in E, \forall n, m \in \mathbb{N}$, on a

$$P(X_{m+n} = j | X_0 = i) = \sum_{k \in E} P(X_m = j | X_0 = k) \times P(X_n = k | X_0 = i)$$

Fonction de masse Soit $X = (X_n)_{n \in \mathbb{N}}$ une chaîne de Markov **homogène** de matrice de transition G. Pour tout entier n, la **fonction de masse de la v.a.r. discrète $\mathbf{X_n}$** est l'application $\Pi^{(n)}$ définie par :

$$\begin{aligned} \Pi^{(n)} : & E \longrightarrow [0,1] \\ & k \longmapsto \pi_k^{(n)} \end{aligned}$$

où $\pi_k^{(n)}$ est la k^e composante du vecteur ligne $\pi^{(n)}$ obtenu en effectuant le produit du vecteur ligne π par la matrice G^n. Autrement-dit $\pi^{(n)} = \pi G^n$, et $\pi^{(0)} = \pi$

Proposition : Soit $X = (X_n)_{n \in \mathbb{N}}$ une chaîne de Markov **homogène** sur un espace d'états E, de matrice de transition G et de condition initiale X_0 ayant pour fonction de masse Π de E dans $[0,1]$. Pour tout entier n et pour tout $(e_0, \ldots, e_n) \in E^{n+1}$, on a

$$P(X_0 = e_0, \ldots, X_n = e_n) = \pi_{e_0} G_{e_0,e_1} \ldots G_{e_{n-1},e_n}$$

6.2 Comportement asymptotique d'une chaîne de Markov

Loi de probabilité invariante : On appelle **loi de probabilité invariante** de la chaîne de Markov **homogène** $X = (X_n)_{n \in \mathbb{N}}$ une **fonction de masse** $\Xi : k \in E \mapsto \xi_k \in [0,1]$ où le vecteur $\xi = (\xi_0, \ldots, \xi_N)$ est solution du système linéaire

$$\xi = \xi G$$

Théorème : Soit $X = (X_n)_{n\in\mathbb{N}}$ une chaîne de Markov **homogène, irréductible et apériodique**, sur un espace d'états E, de matrice de transition G.

— Il existe une **unique loi de probabilité** Ξ invariante et $\forall i \in E, \xi_i > 0$.

— Pour tout $(i, j) \in E^2$, on a

$$\lim_{n\to+\infty} (G^n)_{ij} = \xi_j$$

— Quelle que soit la loi de X_0, la suite de v.a.r. $(X_n)_{n\in\mathbb{N}}$ **converge en loi** vers la **loi de probabilité invariante** Ξ.