Chapitre 1

Rappels et compléments de probabilités

Ce chapitre contient des rappels d'élements de probabilités importants pour la statistique. Une bonne référence est [3]. On supposera connus quelques éléments de base du cours de théorie de la mesure comme les notions de mesure, d'application mesurable, d'intégration par rapport à une mesure etc..Il y sera cependant fait appel de manière minimale, essentiellement pour pouvoir traiter en même temps les cas discrêt et continu. On peut trouver dans [2, chap. 2 et chap. 3] et [1, chap. 6] un résumé des notions et résultats essentiels à connaitre en vue du calcul des probabilités.

1.1 Lois de probabilités

On rappelle qu'à toute expérience aléatoire on associe un triplet $(\Omega, \mathfrak{A}, P)$, dit espace de probabilité où :
Ω : est l'ensemble des résultats possibles de l'expérience
$\mathfrak{A}$: est l'ensemble des évènements (tribu de parties de Ω)
P : une mesure de probabilité.

1.1.1 Definitions

Définition 1.1. Soit $(\Omega, \mathfrak{A})$ et $(\mathfrak{X}, \mathfrak{B})$ deux espaces probabilisables. On appelle variable aléatoire une application mesurable

$$\begin{aligned} X \quad : \quad & \Omega \longrightarrow \mathfrak{X} \\ & \omega \longrightarrow X(\omega) \end{aligned}$$

i.e toute application telle que :

$$\forall\ B \in \mathfrak{B},\ X^{-1}(B) \in \mathfrak{A}.$$

Remarque 1.
1. Si $\mathfrak{X}$ est discret (fini ou infini dénombrable) X est dite variable aléatoire discrète.
2. Si $\mathfrak{X} = R$ on parle de variable aléatoire réelle (v.a.r). 0n prend alors $\mathfrak{B} = \mathfrak{B}_R$ (tribu des boréliens de R).

3. Si $\mathfrak{X} = R^n$ on parle de vecteur alétoire. Dans ce cas $\mathfrak{B} = \mathfrak{B}_{R^n}$ (tribu des boréliens de R^n).

Définition 1.2. Si l'espace $(\Omega, \mathfrak{A})$ est munie d'une mesure de probabilité P, on appelle loi de probabilité de la variable aléatoire X la mesure de probabilité image de P par X, i.e la mesure de probabilité notée P_X définie par

$$\begin{aligned} P_X \quad : \quad & \mathfrak{B} \longrightarrow [0,1] \\ & B \longrightarrow P_X(B) = P(X^{-1}(B)). \end{aligned}$$

Définition 1.3. On appelle fonction de répartition réelle une fonction $F : R \longrightarrow [0,1]$ possédant les propriétés

1. F croissante
2. F continue à gauche
3. $\lim_{x \longrightarrow -\infty} F(x) = 0$ et $\lim_{x \longrightarrow +\infty} F(x) = 1$

Théorème 1.1. *Soit P_X la loi de probabilité d'une v.a.r X alors la fonction*

$$\begin{aligned} F_X \quad : \quad & R \longrightarrow [0,1] \\ & x \longrightarrow F_X(x) = P(X < x) \end{aligned}$$

est une fonction de répartition dite fonction de répartition de X (ou de P_X).

Le théorème suivant est fondamental et montre que toute mesure de probabilité sur $(R, \mathfrak{B}_R)$ est caractérisée de manière unique par sa fonction de répartition.

Théorème 1.2. *Si X et Y sont deux v.a.r de fonctions de répartition respectives F_X et F_Y alors*

$$P_X = P_Y \iff F_X = F_Y$$

Pour une démonstration de ce théorème, on peut se reporter à [2].

Remarque 2. Dans certains ouvrages, la fonction de répartition est définie par $F_X(x) = P(X \leq x)$. c'est une définition équivalente à la précédente, à la différence que cette dernière est continue à droite et non continue à gauche.

En pratique, deux types de lois de probabilités sont utilisées : les lois discrêtes et les lois absolument continues.

Proposition 1.3. *La loi de probabilité d'une variable aléatoire discrête X à valeurs dans $\mathfrak{X} = \{x_0, x_1, ..., x_n, ...\}$ est entièrement caractérisée par la famille de nombres $p_i = P(X = x_i)$ $(i \in N)$. On a $\sum_{i=0}^{+\infty} P(X = x_i) = 1$ et*

$$\forall B \in \mathfrak{B} : \quad P(X \in B) = \sum_{i / x_i \in B} P(X = x_i)$$

Exemple 1.1. Soit X la variable aléatoire qui prend les valeurs $1, 2, 3, 4, 5, 6$ avec les probabilités respectives $\frac{3}{16}, \frac{1}{16}, \frac{2}{16}, \frac{5}{16}, \frac{3}{16}, \frac{2}{16}$. Sa fonction de répartition est donnée par :

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x \leq 1 \\ \frac{3}{16} & \text{si} \quad 1 \leq x \leq 2 \\ \frac{4}{16} & \text{si} \quad 2 \leq x \leq 3 \\ \frac{6}{16} & \text{si} \quad 3 \leq x \leq 4 \\ \frac{11}{16} & \text{si} \quad 4 \leq x \leq 5 \\ \frac{14}{16} & \text{si} \quad 5 \leq x \leq 6 \\ 1 & \text{si} \quad x > 6 \end{cases}$$

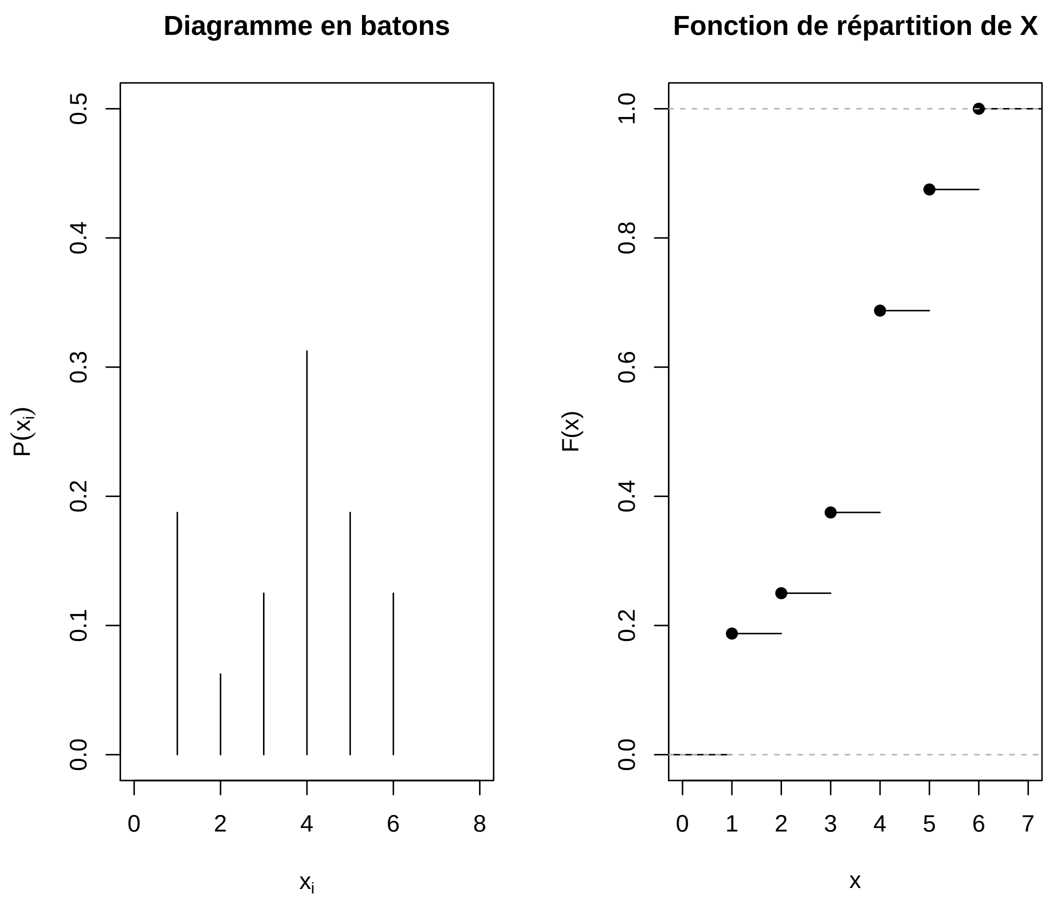

FIGURE 1.1 – Représentations graphiques d'une loi discrète

Définition 1.4. Une variable aléatoire réelle de fonction de répartition F_X est dite absolument continue (ou plus exactement que sa loi de probabilité P_X est absolument continue) s'il existe une fonction $f_X : R \longrightarrow R_+$ telle que

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

f_X est dite densité de probabilité de X.

Propriétés :

1. F_X est presque partout dérivable et on a $F_X^{'} = f_X$
2. f_X est presque partout positive
3. $\int_{-\infty}^{+\infty} f_X(t)dt = 1$
4. $P(X \in B) = \int_B f_X(t)dt \qquad \forall B \in \mathfrak{B}_R$

Définition 1.5. Toute fonction f vérifiant les propriétés 2 et 3 est dite fonction densité de probabilité .

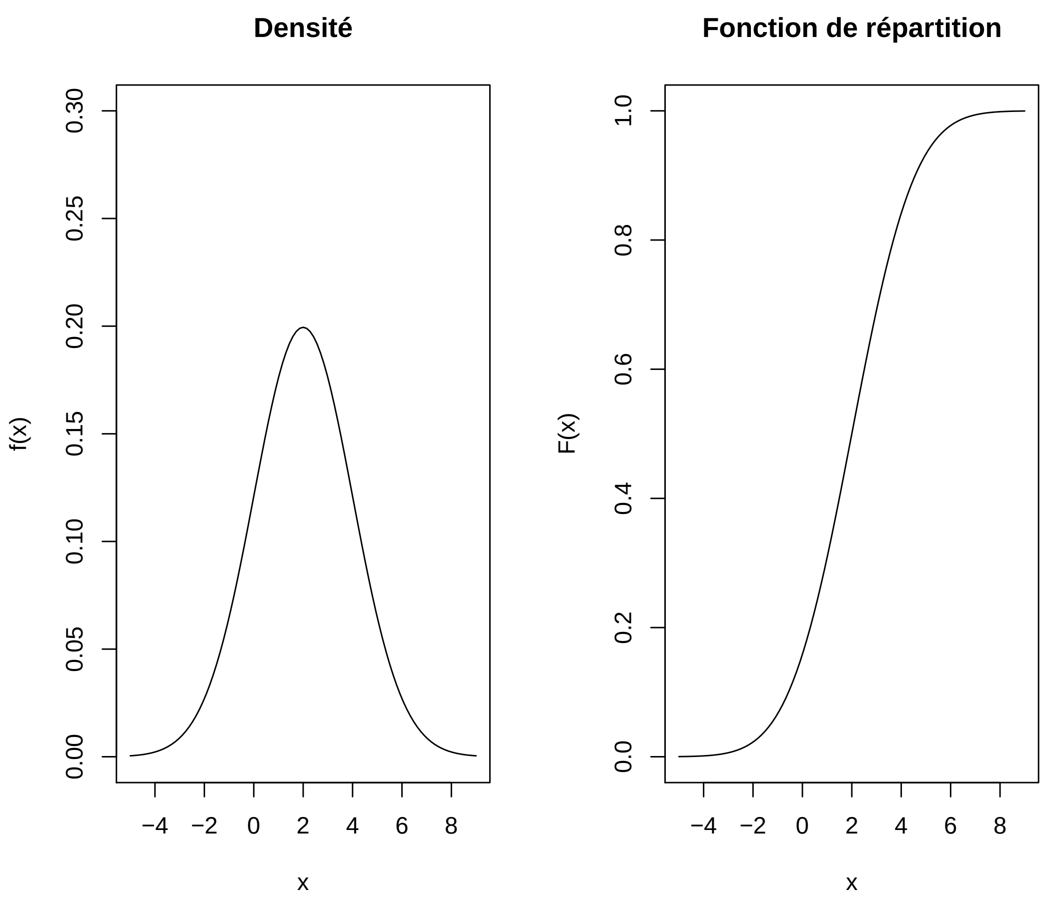

FIGURE 1.2 – Densité et fonction de répartition d'une loi continue

Définition 1.6. Soit P_X la loi de probabilité d'un vecteur aléatoire $X = (X_1, X_2, ..., X_n)$ alors la fonction

$$\begin{aligned} F_X \quad : \quad & R^n \longrightarrow [0,1] \\ & x \longrightarrow F_X(x) = P(X < x) = P(X_1 < x_1, X_2 < x_2, ..., X_n < x_n) \end{aligned}$$

est dite fonction de répartition de X (ou fonction de répartition conjointe de $X_1, X_2, ...$ et X_n).

Définition 1.7. Le vecteur aléatoire $X = (X_1, X_2, ..., X_n)$ est dit absolument continu s'il existe une fonction $f_X : R^n \longrightarrow R_+$ telle que :

$$F(x_1, x_2, ..., x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f_X(t_1, t_2, ..., t_n) dt_1 ... dt_n$$

f_X est dite densité de probabilité de X.

Propriétés :
On a :

1. $f_X(x_1, x_2, ..., x_n) = \dfrac{\partial}{\partial x_1} \dfrac{\partial}{\partial x_2} ... \dfrac{\partial}{\partial x_n} F(x_1, x_2, ..., x_n) \qquad \forall\ (x_1, x_2, ..., x_n) \in R^n$
2. $\int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} f_X(x_1, x_2, ..., x_n) dx_1 ... dx_n = 1$
3. $P(X \in B) = \int ... \int_B f_X(x_1, x_2, ..., x_n) dx_1 ... dx_n \qquad \forall B \in \mathfrak{B}_{R^n}$.

Proposition 1.4. *Si $X = (X_1, X_2, ..., X_n)$ est absolument continu alors les variables marginales X_i $(1 \leq i \leq n)$ sont absolument continues de densités*

$$f_{X_i}(x_i) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} f_X(x_1, x_2, ..., x_n) dx_1 .. dx_{i-1} dx_{i+1} ... dx_n$$

Théorème 1.5 (Changement de variables)**.** *Soit $X = (X_1, X_2, ..., X_n)$ un vecteur aléatoire absolûment continu de densité f_X et $\varphi : R^n \longrightarrow R^n$ définie par $\varphi(x_1, x_2, .., x_n) = (u_1, u_2, .., u_n) = u$ une transformation bijective et de classe C^1, alors le vecteur aléatoire $U = \varphi(X)$ est absolûment continu de densité*

$$f_U(u) = f_X(\varphi^{-1}(u)) \left| J_{\varphi^{-1}} \right|$$

où $\left| J_{\varphi^{-1}} \right|$ désigne la valeur absolue du jacobien (déterminant de la matrice jacobienne) de φ^{-1}

Exemple 1.2. Soit X et Y deux variables aléatoires indépendantes de densités respectives $f_X(x) = \dfrac{b^{a_1}}{\Gamma(a_1)} e^{-bx} x^{a_1 - 1} \mathbf{1}_{\mathbf{R}_+}^{(x)}$ et $f_Y(y) = \dfrac{b^{a_2}}{\Gamma(a_2)} e^{-by} y^{a_2 - 1} \mathbf{1}_{\mathbf{R}_+}^{(y)}$. Déterminer la loi de $(U, V) = \left(X + Y, \dfrac{X}{Y} \right)$.

On a $f_{X,Y}(x,y) = f_X(x)\ f_Y(y)$. et considèrons le changement de variables $\varphi : R^2 \longrightarrow R^2$ définie par $\varphi(x,y) = (x+y, \frac{x}{y}) = (u,v)$. on vérifie que cette application est bijective et de classe C^1 avec $\varphi^{-1}(u,v) = \left(\frac{uv}{v+1}, \frac{u}{v+1}\right)$ et de matrice jacobienne $J_{\varphi^{-1}} = \begin{pmatrix} \frac{v}{v+1} & \frac{1}{v+1} \\ \frac{u}{(v+1)^2} & \frac{-u}{(v+1)^2} \end{pmatrix}$. D'après le théorème du changement de variables le couple de v.a $(U,V) = \left(X+Y, \frac{X}{Y}\right)$ est absolument continu de densité $f_{U,V}(u,v) = f_{X,Y}(\varphi^{-1}(u,v)) \left|J_{\varphi^{-1}}\right|$. On obtient

$$f_{U,V}(u,v) = \frac{b^{a_1+a_2}}{\Gamma(a_1)\Gamma(a_2)} e^{-bu} u^{a_1+a_2-1} \frac{v^{a_1-1}}{(v+1)^{a_1+a_2}} 1_{R^+}(u) 1_{R^+}(v)$$

$$= \frac{b^{a_1+a_2}}{\Gamma(a_1+a_2)} e^{-bu} u^{a_1+a_2-1} 1_{R^+}(u). \frac{\Gamma(a_1+a_2)}{\Gamma(a_1)\Gamma(a_2)} \frac{v^{a_1-1}}{(v+1)^{a_1+a_2}} 1_{R^+}(v)$$

On a donc $f_{U,V}(u,v) = f_U(u) f_V(v)$ où

$$f_U(u) = \frac{b^{a_1+a_2}}{\Gamma(a_1+a_2)} e^{-bu} u^{a_1+a_2-1} 1_{R^+}(u)$$

et

$$f_V(v) = \frac{\Gamma(a_1+a_2)}{\Gamma(a_1)\Gamma(a_2)} \frac{v^{a_1-1}}{(v+1)^{a_1+a_2}} 1_{R^+}(v)$$

Remarque 3. Il sera vu à la section 1.2 que les lois de X, Y, $X+Y$, et $\frac{X}{Y}$ sont des lois connues et sont respectivement des lois $\gamma(a_1,b)$, $\gamma(a_2,b)$, $\gamma(a_1+a_2,b)$ et $\mathfrak{B}_{\mathbf{R}_+}(a,b)$.

Définition 1.8. Les n variables aléatoires $X_1, X_2, ..., X_n$ sont dites indépendantes si

$$P(X_1 \in B_1, X_2 \in B_2, ..., X_n \in B_n) = P(X_1 \in B_1) P(X_2 \in B_2)....P(X_n \in B_n)$$

$\forall B_1 \in \mathfrak{B}, \forall B_2 \in \mathfrak{B}...., \forall B_n \in \mathfrak{B}$

Proposition 1.6. *Les assertions suivantes sont équivalentes*

1. *Les n v.a.r $X_1, X_2, ..., X_n$ sont indépendantes*
2. $F_X(x_1,x_2,...,x_n) = F_{X_1}(x_1) F_{X_2}(x_2)...F_{X_n}(x_n), \qquad \forall\ (x_1,x_2,...,x_n) \in R^n$
3. *Pour toutes fonctions mesurables $h_1, h_2, ..., h_n$, les variables $h_1(X_1), h_2(X_2), ..., h_n(X_n)$ sont indépendantes.*
4. *Si de plus les variables sont absolument continues*
   $f_X(x_1,x_2,...,x_n) = f_{X_1}(x_1) f_{X_2}(x_2)...f_{X_n}(x_n), \qquad \forall\ (x_1,x_2,...,x_n) \in R^n$

Remarque 4. La proposition reste vraie si $X_1, X_2, ..., X_n$ sont des vecteurs aléatoires.

Application : Loi du maximum et du minimum de variables indépendantes.

Soit X_1 et X_2 deux v.a.r indépendantes et posons $U = \max(X_1, X_2)$ et $V = \min(X_1, X_2)$, les fonctions de répartition de U et V sont données respectivement par :

$$\begin{aligned} F_U(x) &= P(\max(X_1, X_2) < x) = P(X_1 < x, X_2 < x) \\ &= P(X_1 < x)P(X_2 < x) = F_{X1}(x)^2 \end{aligned}$$

et

$$\begin{aligned} F_V(x) &= P(\min(X_1, X_2) < x) = 1 - P(\min(X_1, X_2) \geq x) = 1 - P(X_1 \geq x, X_2 \geq x) \\ &= 1 - P(X_1 \geq x)P(X_2 \geq x) = 1 - (1 - F_{X1}(x))^2 \end{aligned}$$

Si X_1, X_2 sont absolument continues alors U et$V$ admettent des densités et on a :

$$f_U(x) = 2f_{X_1}(x)F_{X_1}(x)$$

et

$$f_V(x) = 2f_{X_1}(x)(1 - F_{X_1}(x))$$

Si on a n variables aléaroires $X_1, X_2, ..., X_n$ i.i.d, les fonctions de répartition de $U = \max\limits_{1 \leq i \leq n} (X_i)$ et $V = \min\limits_{1 \leq i \leq n} (X_i)$ sont données respectivement par :

$$F_U(x) = P(\max_{1 \leq i \leq n} (X_i) < x) = P(X_i < x)^n = F^n_{X_1}(x)$$

et

$$F_V(x) = P(\min_{1 \leq i \leq n} (X_i) < x) = 1 - P(\min_{1 \leq i \leq n} (X_i) \geq x) = 1 - (1 - F_{X_1}(x))^n$$

Dans le cas absolument continu, les densités respectives de U et V sont égales à :

$$f_U(x) = nF^{n-1}_{X_1}(x)f_{X_1}(x)$$

et

$$f_V(x) = n(1 - F_{X_1}(x))^{n-1}f_{X_1}(x)$$

1.1.2 Moments de variables aléatoires

Définition 1.9. Une variable aléatoire réelle X définie sur $(\Omega, \mathfrak{A}, P)$ admet une espérance mathématique si elle est P-intégrable i.e si

$$\int_\Omega |X(\omega)| \, dP(\omega) < +\infty$$

La quantité $\int_\Omega X(\omega)dP(\omega)$ est alors dite espérance mathématique de X et notée $E(X)$.

D'après le théorème du transfert (intégration par rapport à la mesure image) on a

$$E(X) = \int_\Omega X(\omega)dP(\omega) = \int_R x dP_X(x)$$

ceci permet d'avoir

$$E(X) = \begin{cases} \sum\limits_{i=0}^{+\infty} x_i P(X = x_i) & \text{si } X \text{ est discrête} \\ \int\limits_R x f_X(x) & \text{si } X \text{ est absolument continue} \end{cases}$$

Propriétés :

1. $E(aX + bY) = aE(X) + bE(Y) \quad \forall a \in R, \forall b \in R$ (linéarité)
2. $E(X) \geq 0$ si $X \geq 0$ (positivité)
3. Pour toute fonction mesurable g on a :
$$E(g(X)) = \int\limits_R g(x) dP_X(x) = \begin{cases} \sum\limits_{i=0}^{+\infty} g(x_i) P(X = x_i) & \text{si } X \text{ est discrête} \\ \int\limits_R g(x) f_X(x) & \text{si } X \text{ est absolument continue} \end{cases}$$
4. $E(XY) = E(X)E(Y)$ si X et Y sont indépendantes.

Définition 1.10.
- Une variable aléatoire réelle X admet un moment d'ordre k si X^k est P-intégrable, on note alors $m_k = E(X^k)$
- Une variable aléatoire réelle X , telle que $E(X) < +\infty$, admet un moment centré d'ordre k si $(X - E(X))^k$ est P-intégrable, on note alors $\mu_k = E((X - E(X))^k)$

Le moment centré d'ordre 2 est appelé variance de X et est noté $Var(X)$. C'est un indicateur de la dispersion d'une variable autour de sa moyenne.

Propriétés :

1. $Var(X) = E(X^2) - E(X)^2$.
2. $Var(X) = 0 \iff \exists\, c \in R \;/\; X = c \; P - p.s.$
3. $Var(aX + b) = a^2 Var(X)$
4. $Var(X + Y) = Var(X) + Var(Y)$

Théorème 1.7 (Inégalité de Bienaymé-Tchebychev)**.** *Si X est une v.a avec une espérance m et une variance σ^2 finies, on a :*

$$\forall \quad \epsilon > 0, \qquad P(|X - m| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

Définition 1.11. Soit un vecteur aléatoire $X = (X_1, X_2, ..., X_n)$.

1. On appelle espérance mathématique de X (lorsqu'elle existe) le vecteur
$$E(X) = (E(X_1), E(X_2), ..., E(X_n))^t.$$
2. On appelle matrice de variances-covariances du vecteur X, dont l'espérance existe, la matrice
$$C(X) = E((X - E(X))^t (X - E(X)))$$
On a $C(X) = (c_{i,j})_{1 \leq i \leq n, 1 \leq j \leq n}$ où $c_{i,j} = E(X_i - E(X_i)(X_j - E(X_j)) = cov(X_i, X_j)$

$$C(X) = \begin{pmatrix} Var(X_1) & cov(X_1,X_2) & . & . & cov(X_1,X_n) \\ cov(X_1,X_2) & Var(X_2) & . & . & cov(X_2,X_n) \\ . & . & . & . & . \\ . & . & . & . & . \\ cov(X_1,X_n) & cov(X_2,X_n) & . & . & Var(X_n) \end{pmatrix}$$

Propriétés : Pour toute matrice (p, n), A et tout vecteur b de dimension p on a :

1. $E(AX + b) = AE(X) + b$
2. $Cov(AX + b) = ACov(X)A^t$

Définition 1.12. Le coefficient de corrélation linéaire de deux variables X_1 et X_2 est défini par

$$\rho_{X_1,X_2} = \frac{cov(X_1, X_2)}{\sigma_{X_1}\sigma_{X_2}}$$

Proposition 1.8. *On a*

1. $-1 \leq \rho_{X_1,X_2} \leq 1$
2. $|\rho_{X_1,X_2}| = 1 \iff \exists\ a \in R, \exists\ b \in R\ /\ X_2 = aX_1 + b.$
3. *Si X_1 et X_2 sont indépendantes alors $\rho_{X_1,X_2} = 0$ (la réciproque est fausse).*

1.1.3 Fonctions génératrices des moments

Comme les fonctions de répartitions (voir proposition 1.2), les fonctions génératrices caractérisent entièrement les lois de probabilités. De plus, à l'instar des fonctions caractéristiques, elles constituent un outil puissant en probabilités. Elles permettent notamment de simplifier beaucoup de calculs. Elles sont à distinguer de la notion de fonction génératrice des moments factoriels définie uniquement pour les v.a entières (à valeurs dans N), quoique possédant des propriétés identiques.

Définition 1.13. On appele fonction génératrice des moments de la v.a réelle X la fonction définie sur une partie $D \subset R$ dans R_+ par

$$\Psi_X(t) = E(e^{tX})$$

Remarque 5.

1. Le domaine de définition D dépend de la v.a X
2. Si X est discrète $\Psi_X(t) = \sum\limits_{n=0}^{+\infty} e^{tx_i} P(X = x_i)$
3. Si X est absolument continue de densité f_X on a $\Psi_X(t) = \int_R e^{tx} f_X(x)dx$

Exemple 1.3.

1. $X \sim B(p) : \Psi_X(t) = p + qe^t$
2. $X \sim B(n,p); \Psi_X(t) = (p + qe^t)^n$
3. $X \sim P(\lambda) : \Psi_X(t) = e^{\lambda(e^t - 1)}$
4. $X \sim \gamma(a,b); \Psi_X(t) = \dfrac{1}{(1 - \frac{t}{b})^a}$
5. $X \sim N(m, \sigma^2) : \Psi_X(t) = e^{tm + \frac{1}{2}t^2\sigma^2}$ en particulier si $X \sim N(0,1) : \Psi_X(t) = e^{\frac{1}{2}t^2}$

Les trois théorèmes suivants sont très utiles et seront beaucoup utilisés par la suite.

Théorème 1.9. *Soit X et Y deux v.a.r*

$$P_X = P_Y \iff \Psi_X = \Psi_Y.$$

Théorème 1.10. *Soit X et Y deux v.a indépendantes. On a :*

$$\Psi_{X+Y}(t) = \Psi_X(t)\Psi_Y(t) \quad \forall t \in R$$

Démonstration. $\Psi_{X+Y}(t) = E(e^{t(X+Y)}) = E(e^{tX})E(e^{tY}) = \Psi_X(t)\Psi_Y(t)$ □

Théorème 1.11. *Une v.a.r X admet un moment d'ordre k si et seulement si Ψ_X est dérivable à l'ordre k au point 0 et on a :*

$$E(X^k) = \Psi_X^{(k)}(0)$$

Définition 1.14. On appelle fonction génératrice des moments du vecteur aléatoire $X = (X_1, X_2, ..., X_n)$ la fonction définie sur une partie $D \subset R^n$ dans R_+ par

$$\Psi_X(t) = E(e^{<t,X>}) = E(e^{\sum_{i=1}^{n} t_i X_i})$$

Exemple 1.4. $X \sim N_n(m, \Sigma) : \Psi_X(t) = e^{<t,m>+\frac{1}{2}t'\Sigma t}$. En particulier si $X \sim N_n(0, I_n)$ on a $\Psi_X(t) = e^{\frac{1}{2}t' t}$

Propriétés :

1. *Les fonctions génératrices marginales sont données par*

$$\Psi_{X_i}(t_i) = \Psi_X(0, .., 0, t_i, 0, ..0) \quad \text{pour tout } 1 \leq i \leq n$$

2. *$X_1, X_2, ..., X_n$ sont indépendantes si et seulement si*

$$\Psi_X(t) = \Psi_{X_1}(t_1)\Psi_{X_2}(t_2)....\Psi_{X_n}(t_n) forall\ t = (t_1, t_2, ..., t_n) \in R^n$$

1.2 Lois de probabilités usuelles univariées

1.2.1 Lois de probabilités discrêtes

Loi de Bernoulli

Définition 1.15. Une variable aléatoire X est dite suivre une loi de Bernoulli de paramètre p $(0 < p < 1)$ si elle prend ses valeurs dans $\{0, 1\}$ avec $P(X = 0) = 1 - p = q$ et $P(X = 1) = p$

Notation 1. $X \sim B(p)$

Propriétés : $E(X) = p$ et $Var(X) = pq$

Exemple 1.5. Tout résultat d'une expérience aléatoire à deux issues possibles (succés, échec ou encore 0,1) peut-être modélisée à l'aide d'une v.a de Bernoulli. Par exemple le résultat d'un lancer d'une pièce de monnaie, en posant $X = 1$ si pile sort et $X = 0$ si face sort. $p = P(X = 1)$ est la probabilité de pile et $q = 1 - p$ celle de face.

Loi binomiale

Définition 1.16. Une variable aléatoire X suit une loi binomiale de paramètres n et p ($n \in N$, $0 < p < 1$) si elle prend ses valeurs dans $\{0, 1, 2, .., n\}$ avec

$$P(X = k) = C_n^k p^k (1 - p)^{n-k}$$

Notation 2. $X \sim B(n, p)$

Propriétés :

1. $B(1, p) = B(p)$
2. $E(X) = np$ et $Var(X) = npq$
3. $X_i \sim B(p)$, $1 \leq i \leq n$ i.i.d $\Longrightarrow \sum_{i=1}^{n} X_i \sim B(n, p)$

Exemple 1.6. Considérons, une expérience de type succés-échec répétée n fois dans des conditions identiques et de manière indépendante telle que la probabilité d'avoir à chaque répétition un succès est p. Soit X la v.a qui compte le nombre total de succés obtenus, alors $X \sim B(n, p)$

Loi géométrique

Définition 1.17. Une variable aléatoire X suit une loi géométrique de paramètre p ($0 < p < 1$) si elle prend ses valeurs dans $N*$ avec

$$P(X = k) = p(1 - p)^{k-1} = pq^{k-1}$$

Notation 3. $X \sim G(p)$

Propriétés : $E(X) = 1/p$ et $Var(X) = \dfrac{q}{p^2}$

Exemple 1.7. Considérons, une expérience de type succés-échec répétée dans des conditions identiques et de manière indépendante jusqu'à obtenir pour la première fois un succès. On suppose que la probabilité d'avoir à chaque répétition un succès est p. Soit X la v.a qui compte le nombre de répétitions nécessaires, alors $X \sim G(p)$

Loi de Poisson

Définition 1.18. Une variable aléatoire X suit une loi de Poisson de paramètre λ ($\lambda \in R$) si elle prend ses valeurs dans N avec

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Notation 4. $X \sim P(\lambda)$

Propriétés :

1. $E(X) = \lambda$ et $Var(X) = \lambda$
2. $X_1 \sim P(\lambda_1)$, $X_2 \sim P(\lambda_2)$ indépendantes $\Longrightarrow X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$

La loi de Poisson est une loi limite comme le montre le théorème d'approximation de la loi binomiale par la loi de Poisson (voir section 1.7). Si n grand et p petit

$$B(n, p) \approx P(np)$$

1.2.2 Lois de probabilités continues

Loi Uniforme

Définition 1.19. Une variable aléatoire réelle X est dite de loi uniforme sur l'intervalle $[a,b]$ si elle admet la densité

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

Notation 5. $X \sim \mathfrak{U}_{[a,b]}$

Loi Gamma

Définition 1.20. On appelle fonction Gamma la fonction de R_+ dans R_+ définie par

$$\Gamma(a) = \int_0^{+\infty} x^{a-1}e^{-x}dx$$

Définition 1.21. On appelle fonction Béta la fonction définie sur $R_+ \times R_+$ par

$$\mathcal{B}(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$

Propriétés :

1. $\Gamma(a) = (a-1)\Gamma(a-1) \quad \forall a \in R^*$
2. $\Gamma(n) = (n-1)! \quad \forall n \in N^*$ avec $\Gamma(0) = 1$
3. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$
4. Pour n et m entiers naturels, et en supposant $n \geq m$, on a

$$\mathcal{B}(n,m) = \frac{(n-1)!(m-1)!}{(n+m-1)!} = \frac{1}{(n+m-1)}\frac{1}{C_{m-1}^{n-1}}.$$

Définition 1.22. Une variable aléatoire réelle X est dite suivre une loi gamma si elle admet la densité :

$$f(x) = \frac{b^a}{\Gamma(a)}x^{a-1}e^{-bx}\mathbf{1}_{\mathbf{R}_+}^{(x)}$$

a et b paramètres réels strictement positifs

Notation 6. $X \sim \gamma(a,b)$

Cas particuliers :

1. $\gamma(1,b) = Exp(b)$: loi exponentielle de paramètre b
2. $\gamma(\frac{n}{2}, \frac{1}{2}) = \chi_n^2$: loi du khi-deux à n degrés de liberté (loi fondamentale en statistique)

Propriétés :

1. $\int_0^{+\infty} x^{a-1}e^{-bx}dx = \frac{\Gamma(a)}{b^a}$ (intégrale utile à connaitre)

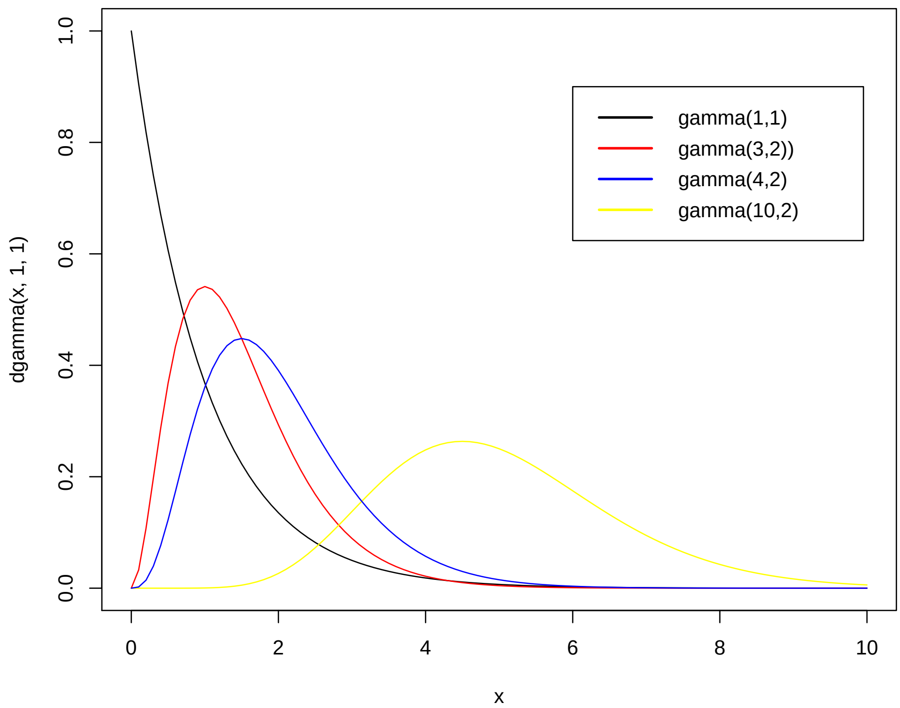

FIGURE 1.3 – Densités de lois gamma

2. $E(X) = \frac{a}{b}$, $Var(X) = \frac{a}{b^2}$ (démonstration directe ou en utilisant les fonctions génératrices)
3. $X_1 \sim \gamma(a_1, b), X_2 \sim \gamma(a_2, b) \implies X_1 + X_2 \sim \gamma(a_1 + a_2, b)$
   propriété de stabilité par convolution sur le premier paramètre
   (Démonstration directe ou par les fonctions génératrices voir section)

Lois béta

Définition 1.23. Une variable aléatoire réelle X est dite de loi Béta sur $[0, 1]$ (ou loi Béta de première espèce) si elle admet la densité

$$f(x) = \frac{1}{\mathcal{B}(a,b)} x^{a-1}(1-x)^{b-1}\mathbf{1}_{[\mathbf{0},\mathbf{1}]}^{(x)}$$

Définition 1.24. Une variable aléatoire réelle X est dite de loi Béta sur R_+ (ou loi Béta de deuxième espèce) si elle admet la densité

$$f(x) = \frac{1}{\mathcal{B}(a,b)} \frac{x^{a-1}}{(1+x)^{a+b}} \mathbf{1}_{\mathbf{R}_+}^{(x)}$$

Notation 7. $X \sim \mathfrak{B}_{[0,1]}(a, b)$ et $X \sim \mathfrak{B}_{\mathbf{R}_+}(a, b)$.

Propriétés :

1. La loi uniforme sur $[0, 1]$ correspond à la loi $\mathfrak{B}_{[0,1]}(1, 1)$.
2. Si $X \sim \mathfrak{B}_{[0,1]}(a, b)$ alors $\dfrac{X}{1-X} \sim \mathfrak{B}_{\mathbf{R}_+}(a, b)$.

Loi normale

Définition 1.25. Une variable aléatoire réelle X est dite suivre une loi normale si elle admet la densité :

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

μ, σ^2 paramètres réels respectivement quelconque et strictement positif

Notation 8. $X \sim N(\mu, \sigma^2)$
N(0,1) : loi normale centrée et réduite

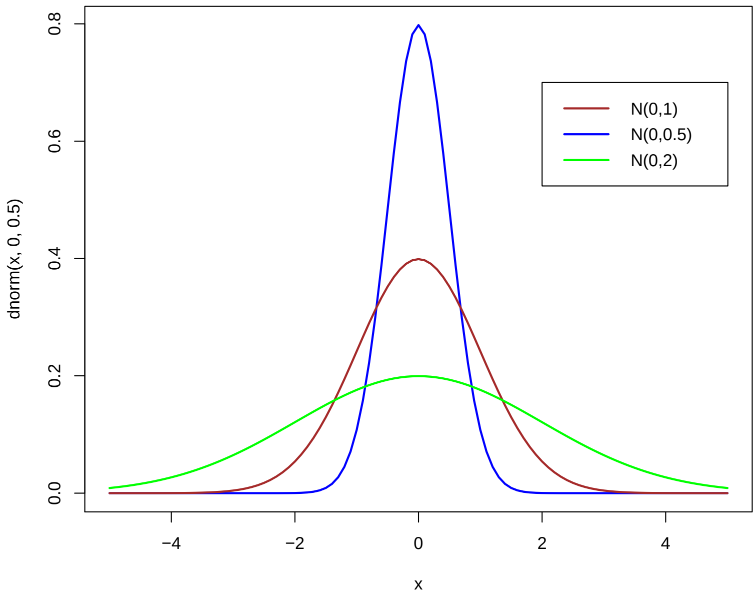

Figure 1.4 – Densités de lois normales

Propriétés :

1. $X \sim N(\mu, \sigma^2) \implies Y = \dfrac{X-\mu}{\sigma} \sim N(0, 1)$ (centrage et réduction)
2. $Y \sim N(0, 1) \implies X = \sigma Y + \mu \sim N(\mu, \sigma^2)$

3. La fonction de répartition $\Phi(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ d'une loi normale centrée réduite n'a pas d'expression analytique, elle est, en conséquence tabulée.
   La fonction de répartition d'une loi normale $N(m, \sigma^2)$ se déduit de Φ par la relation
   $$F_{N(m,\sigma^2)}(x) = \Phi(\frac{x-m}{\sigma})$$
   .
4. $E(X) = \mu, \quad Var(X) = \sigma^2$
5. $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$ indépendantes $\Longrightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
6. $X_i \sim N(\mu_i, \sigma_i^2)$, $1 \leq i \leq n$ indépendantes $\Longrightarrow \sum\limits_{i=1}^{n} X_i \sim N(\sum\limits_{i=1}^{n} \mu_i, \sum\limits_{i=1}^{n} \sigma_i^2)$
   propriété de stabilité par convolution sur les deux paramètres (Démonstration directe ou par les fonctions génératrices)
7. $X \sim N(0,1) \Longrightarrow X^2 \sim \chi_1^2$ (khi-deux à un ddl)
8. $X_i \sim N(0,1)$, $1 \leq i \leq n$ i.i.d $\Longrightarrow \sum\limits_{i=1}^{n} X_i^2 \sim \chi_n^2$ (khi-deux à n ddl)
9. $X_i \sim N(\mu_i, \sigma_i^2)$, $1 \leq i \leq n$ indépendantes $\Longrightarrow \sum\limits_{i=1}^{n} (\frac{X_i - \mu_i}{\sigma_i})^2 \sim \chi_n^2$

1.3 Lois multivariées

1.3.1 Loi multinomiale

Théorème 1.12 (Formule du multinôme)**.** *Soit $p_1, p_2, .., p_k$ k nombres réels et n un entier naturel. On a*

$$(p_1 + p_2, .. + p_k)^n = \sum_{(n_1, n_2, ..., n_k) / \sum\limits_{i=1}^{k} n_i = n} \frac{n!}{n_1! n_2! ... n_k!} p_1^{n_1} p_2^{n_2}, .., p_k^{n_k}$$

Remarque 6. Le coefficient $\frac{n!}{n_1! n_2! ... n_k!}$ est dit coefficient multinomial et est noté $\begin{pmatrix} n \\ n_1, n_2, ...n_k \end{pmatrix}$.
Pour $n = 2$ on a $\frac{n!}{n_1! n_2!} = C_n^{n_1}$ et la formule (1.12) se réduit à la formule du binôme.

Définition 1.26. Un vecteur aléatoire $X = (X_1, X_2, .., X_k)$ est dit de loi multinomiale $\mathfrak{M}(n, p_1, p_2, .., p_k)$ s'il prend ses valeurs dans $\mathcal{D} = \{(n_1, n_2, ..., n_k)/n_i \in N \ et \ \sum\limits_{i=1}^{k} n_i = n\}$ avec

$$P(X_1 = x_1, X_2 = x_2, .., X_k = x_k) = \frac{n!}{n_1! n_2! ... n_k!} p_1^{n_1} p_2^{n_2}, .., p_k^{n_k}$$

Notation 9. $X = (X_1, X_2, .., X_k) \sim \mathfrak{M}(n, p_1, p_2, .., p_k)$

Remarque 7. On a bien une loi de probabilité par la formule du multinôme.

Propriétés :

1. $X_i \sim B(n, p_i)$
2. $E(X_i) = np_i$ et $Var(X_i) = np_i(1 - p_i)$
3. Un calcul simple donne : $Cov(X_i, X_j) = -np_ip_j$

Exemple 1.8. Soit une urne contenant des boules de k couleurs différenres : $C_1, C_2, ...C_k$ en proportions respectives $p_1, p_2, .., p_k$. On fait n tirages avec remise, et soit $X_1, X_2, .., X_k$ le nombre de boules des différentes couleurs obtenues. Alors $X = (X_1, X_2, .., X_k)^t \sim \mathfrak{M}(n, p_1, p_2, .., p_k)$.

1.3.2 Loi normale multivariée

Définition 1.27 (Loi normale multivariée)**.** Un vecteur aléatoire $X = (X_1, X_2, .., X_n)$ à valeurs dans R^n est dit de loi normale s'il admet la densité

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\Sigma)}} e^{-\frac{1}{2}(x-m)^t \Sigma^{-1}(x-m)}$$

avec m $\in R^n$ et Σ matrice (n,n) symétrique définie positive.

Notation 10. $X \sim N_n(m, \Sigma)$

Cas particulier : $m = (0, 0, .., 0)$ et $\Sigma = I_n$
$X \sim N_n(0, I_n)$: loi normale centrée, réduite dans R^n

Exemple 1.9. Soit X un vecteur aléatoire de densité

$$f(x, y) = \frac{7}{18\pi} \exp(-\frac{4x^2 - 2xy + y^2)}{6})$$

On peut écrire :

$$-\frac{4x^2 - 2xy + y^2)}{6} = \frac{1}{2}(x, y) \begin{pmatrix} \frac{4}{3} & \frac{-1}{3} \\ \frac{-1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -\frac{1}{2}(x, y)\Sigma^{-1} \begin{pmatrix} x \\ y \end{pmatrix}$$

On vérifie que : $\frac{1}{2\pi \det(\Sigma)} = \frac{\det(\Sigma^{-1})}{2\pi} = \frac{7}{18\pi}$.
Le vecteur X est donc un vecteur gaussien avec pour moyenne et matrice de covariances respectivement $m = (0, 0)^t$ et Σ

Exemple 1.10. Loi normale bivariée $X \sim N_2(m, \Sigma)$ avec $m = (m_1, m_2)^t$ et $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\, \sigma_1\sigma_2 \\ \rho\, \sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$

On a alors :

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2}} \exp(-\frac{1}{2\sqrt{1 - \rho^2}}(\frac{(x - m_1)^2}{\sigma_1^2} - \frac{2\rho(x - m_1)(y - m_2)}{\sigma_1\sigma_2} + \frac{(y - m_2)^2}{\sigma_2^2}))$$

Théorème 1.13. *Soit $X \sim N_n(m, \Sigma)$ et A une matrice (p, n) alors $Y = AX \sim N_p(Am, A\Sigma A^t)$*

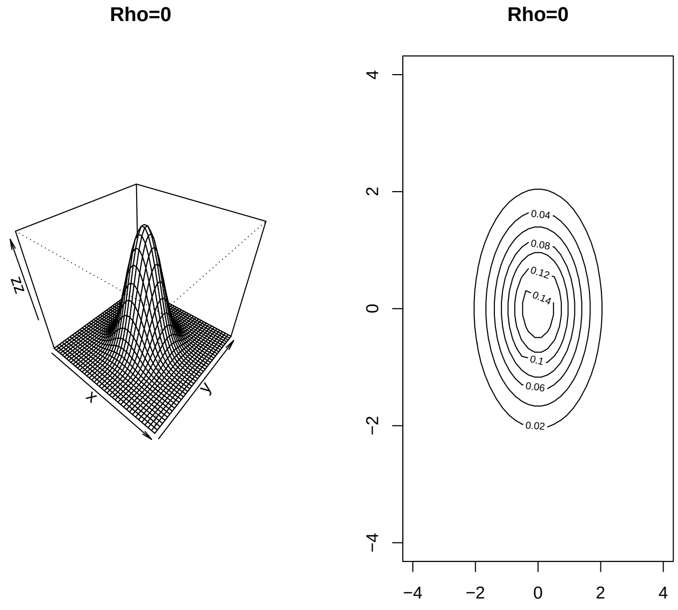

FIGURE 1.5 – Graphe et courbe de niveaux de lois normales bivariées

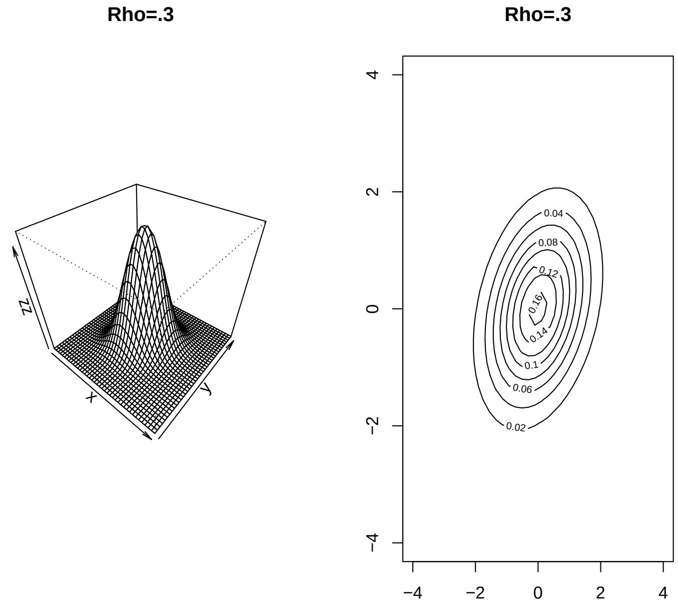

FIGURE 1.6 – Graphe et courbe de niveaux de lois normales bivariées

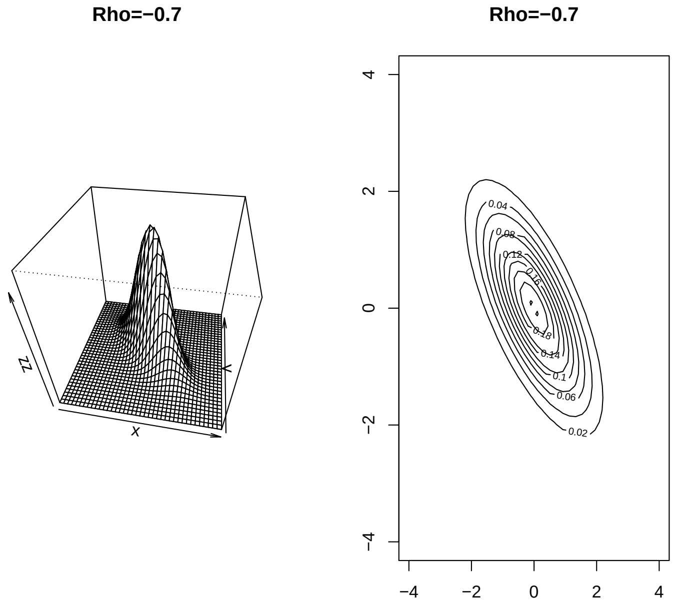

FIGURE 1.7 – Graphe et courbe de niveaux de lois normales bivariées

Démonstration. On utilise les fonctions génératrices (voir paragraphe 1.1.3) $\Psi_Y(u) = E(e^{<u,AX>}) = E(e^{<A^tu,X>}) = \Psi_X(A^tu) = e^{<A^tu,m>+\frac{1}{2}(A^tu)'\Sigma(A^tu)} = e^{<u,Am>+\frac{1}{2}u'(A\Sigma A^t)u}$ □

Corollaire 1.14. *Soit* $X = (X_1, X_2, .., X_n) \sim N_n(m, \Sigma)$

1. *Toute combinaison linéaire* $\sum_{i=1}^{n} a_iX_i$ *, pour tous réels* $a_1, ..., a_n$*, suit une loi normale*
2. *En particulier* $X_i \sim N(m_i, \sigma_i^2)$ *pour tout* i

Exemple 1.11. Soit $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \sim N_3(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix})$

Loi de $(-X+Y+Z, X-Y+Z, X+Y-Z)$? On a :

$$\begin{pmatrix} \text{U} \\ \text{V} \\ \text{W} \end{pmatrix} = \begin{pmatrix} \text{-1} & 1 & 1 \\ 1 & \text{-1} & 1 \\ 1 & 1 & \text{-1} \end{pmatrix} \begin{pmatrix} \text{X} \\ \text{Y} \\ \text{Z} \end{pmatrix} = A(X, Y, Z)^t$$

donc $(U, V, W)^t$ transformé linéaire de $(X, Y, Z)^t$ est un vecteur gaussien dans R^3 de moyenne $A\mu = (0, 0, 0)^t$ et

de matrice de variances-covariances $A\Sigma A^t = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.

Théorème 1.15. *Soit $X = (X_1, X_2, .., X_n) \sim N_n(m, \Sigma)$. $X_1, X_2, .., X_n$ sont indépendantes si et seulement si Σ est diagonale.*

Démonstration. ($\Longrightarrow$) évident

($\Longleftarrow$) soit $\Sigma = \begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ . & . & . & . \\ 0 & 0 & . & \sigma_n^2 \end{pmatrix}$

On a : $\det(\Sigma) = \sigma_1^2\sigma_2^2...\sigma_n^2$ et $(x-m)^t\Sigma^{-1}(x-m) = \sum_{i=1}^{n} \frac{(x_i - m_i)^2}{\sigma_i^2}$

$$\text{donc } f_X(x) = \frac{1}{(\sqrt{2\pi})^n \sigma_1..\sigma_n} e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-m_i)^2}{\sigma_i^2}} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}\frac{(x_i-m_i)^2}{\sigma_i^2}} = \prod_{i=1}^{n} f_{X_i}(x_i)$$

□

Théorème 1.16. *Soit $X_i \sim N(m, \sigma^2)$, $1 \le i \le n$ i.i.d, alors on a :*

1. *$\overline{X_n}$ et S_n^2 sont indépendantes*
2. *$\overline{X_n} \sim N(m, \frac{\sigma^2}{n})$*
3. *$\frac{nS_n^2}{\sigma^2} \sim \chi_{n-1}^2$*

1.3.3 Lois liées à la loi normale

Théorème 1.17. *Soit deux v.a X et Y indépendantes telles que $X \sim N(0,1)$, $Y \sim \chi_n^2$ alors $Z = \frac{X}{\sqrt{\frac{Y}{n}}}$ a pour densité*

$$f_Z(x) = \frac{1}{\sqrt{n}B(\frac{1}{2}, \frac{n}{2})} \frac{1}{(1 + \frac{x^2}{\sigma^2})^{\frac{n+1}{2}}}$$

dite loi de Student à n degrés de liberté.

Notation 11. St_n ou $t(n)$.

Exemple 1.12. Soit X_i, $1 \le i \le n$, un n-échantillon de $X \sim N(m_1, \sigma_1^2)$, alors :

1. $\sqrt{n}\dfrac{\sqrt{n}\dfrac{\overline{X_n} - m}{\sigma}}{\sqrt{\dfrac{\sum_{i=1}^{n}(X_i - m)^2}{\sigma^2}}} = \sqrt{n}\dfrac{\overline{X_n} - m}{\sqrt{\dfrac{\sum_{i=1}^{n}(X_i - m)^2}{n}}} \sim St_n$
2. $\sqrt{n-1}\dfrac{\sqrt{n}\dfrac{\overline{X_n} - m}{\sigma}}{\sqrt{\dfrac{nS^2}{\sigma^2}}} = \sqrt{n-1}\dfrac{\overline{X_n} - m}{S} = \sqrt{n}\dfrac{\overline{X_n} - m}{S'} \sim St_{n-1}$

Ces deux résultats découlent immédiatement du fait que

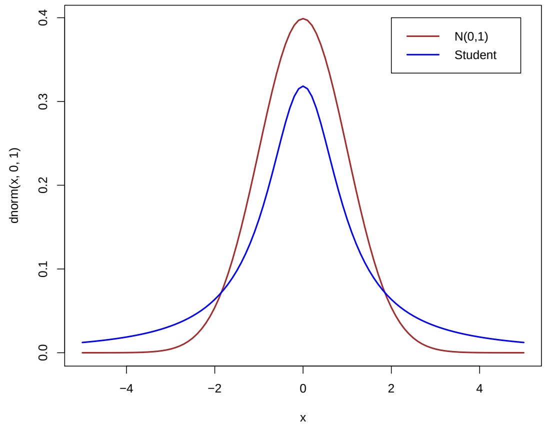

FIGURE 1.8 – Lois normale et de Student

1. $\dfrac{\overline{X_n} - m}{\dfrac{\sigma}{\sqrt{n}}} \sim N(0,1)$
2. $\dfrac{\sum\limits_{i=1}^{n}(X_i - m)^2}{\sigma^2} \sim \chi^2_n$
3. $\dfrac{nS_n^2}{\sigma^2} \sim \chi^2_{n-1}$

Théorème 1.18. *Soit deux v.a.r. U et V indépendantes telles que $U \sim \chi^2_n$, $V \sim \chi^2_m$ alors $F = \dfrac{\frac{U}{n}}{\frac{V}{m}} = \dfrac{m}{n}\dfrac{U}{V}$ a pour densité*

$$f_F(x) = \frac{(\frac{n}{m})^{\frac{n}{2}}(x)^{\frac{n}{2}-1}}{B(\frac{n}{2}, \frac{m}{2})} \frac{1}{(1 + \frac{n}{m}x)^{\frac{n+m}{2}}} 1_{R_+}(x)$$

dite loi de Fischer à n et m degrés de liberté notée $F_{n,m}$.

Exemple 1.13. Soit X_i, $1 \leq i \leq n$, un n-échantillon de $X \sim N(m_1, \sigma_1^2)$ et Y_i $1 \leq j \leq m$ un m-échantillon de $Y \sim N(m_2, \sigma_2^2)$, avec X et Y indépendantes alors

1. $\dfrac{\dfrac{\sum\limits_{i=1}^{n}(X_i - m_1)^2}{n\sigma_1^2}}{\dfrac{\sum\limits_{j=1}^{m}(Y_j - m_2)^2}{m\sigma_2^2}} \sim F_{n,m}.$

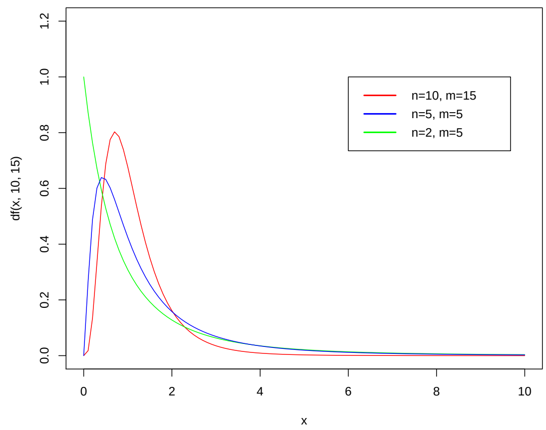

FIGURE 1.9 – Densités de lois de Fischer

2. $$\frac{\dfrac{\sum_{i=1}^{n}(X_i-\overline{X_n})^2}{(n-1)\sigma_1^2}}{\dfrac{\sum_{j=1}^{m}(Y_j-\overline{Y_m})^2}{(m-1)\sigma_2^2}} = \frac{(n/(n-1))\sigma_2^2}{(m/(m-1))\sigma_1^2}\frac{S_1^2}{S_2^2} = \frac{\sigma_2^2}{\sigma_1^2}\frac{S_1'^2}{S_2'^2} \sim F_{n-1,m-1}.$$

1.4 Famille des lois exponentielles

La famille des lois exponentielles est une classe de lois trés générale qui contient comme cas particuliers la plupart des lois usuelles. Elle joue un rôle important en statistique paramétrique.

Définition 1.28. *La famille de lois* $\mathcal{F} = \{f_\theta(.)/\theta \in \Theta\}$ *de* $\boldsymbol{R}^p$ est dite appartenir à la famille des lois exponentielles $\mathcal{E}_1$ d'ordre 1 s'il existe des fonctions h, a_1 de $\boldsymbol{R}^p$ dans $\boldsymbol{R}$ et c, α_1 de Θ dans $\boldsymbol{R}$ telles que

$$f_\theta(x) = c(\theta)h(x)\exp(a(x)\alpha(\theta))$$

Cela est équivalent à

$$f_\theta \in \mathcal{E}_1 \iff \log(f_\theta(x)) = C(\theta) + H(x) + a(x)\alpha(\theta)$$

en posant $C(\theta) = \log(c(\theta))$, $H(x) = \log(h(x))$

Exemple 1.14. La plupart des lois de probabilités usuelles appartiennent à cette famille

1. Loi de Poisson
   $f_\lambda(x) = e^{-\lambda}\dfrac{\lambda^x}{x!} = e^{-\lambda}\dfrac{1}{x!}e^{x\log\lambda} = c(\lambda)h(x)\exp(a(x)\alpha(\lambda))$

ou encore :
$\log(f_\lambda(x)) = -\lambda + \log(\frac{1}{x!}) + x\log\lambda$

2. Loi de Bernoulli
$f_p(x) = p^x(1-p)^{1-x} = (1-p)(\frac{p}{1-p})^x = (1-p)\exp(x\log(\frac{p}{1-p})) = c(p)\exp(a(x)\alpha(p))$

   ou bien
   $\log(f_p(x)) = \log(1-p) + x\log(\frac{p}{1-p}) = C(p) + a(x)\alpha(p)$

3. Loi $N(m,\sigma^2)$, σ^2 connu
$$f_m(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\frac{m^2}{\sigma^2}}e^{-\frac{1}{2}\frac{x^2}{\sigma^2}}e^{\frac{xm}{\sigma^2}} = c(m)h(x)\exp(a(x)\alpha(m))$$

4. Loi $N(m,\sigma^2)$, m connu
$$f_m(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}(x-m)^2.\frac{1}{\sigma^2}} = c(\sigma^2)\exp(a(x)\alpha(\sigma^2))$$

Définition 1.29. *La famille de lois $\mathcal{F} = \{f_\theta(.)/\theta \in \Theta\}$ de $\boldsymbol{R}^p$* est dite appartenir à la famille des lois exponentielles $\mathcal{E}_k$ d'ordre k s'il existe des fonctions $h, a_i(1 \le i \le k)$ de $\boldsymbol{R}^p$ dans $\boldsymbol{R}$ et $c, \alpha_i(1 \le i \le k)$ de Θ dans $\boldsymbol{R}$ telles que

$$f_\theta(x) = c(\theta)h(x)\exp(\sum_{i=1}^{k} a_i(x)\alpha_i(\theta))$$

Cela est équivalent à :

$$f_\theta \in \mathcal{E}_k \iff \log(f_\theta(x)) = C(\theta) + H(x) + \sum_{i=1}^{k} a_i(x)\alpha_i(\theta)$$

avec : $C(\theta) = \log(c(\theta))$, $H(x) = \log(h(x))$

Exemple 1.15. Loi $N(m,\sigma^2)$

$$\begin{aligned} f_{m,\sigma^2}(x) &= \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\frac{m^2}{\sigma^2}}e^{\frac{xm}{\sigma^2}-\frac{1}{2}\frac{x^2}{\sigma^2}} \\ &= c(m,\sigma^2)h(x)\exp(a_1(x)\alpha_1(m) + a_2(x)\alpha_2(m)) \end{aligned}$$

avec $a_1(x) = x$, $a_2(x) = x^2$

1.5 Lois conditionnelles

Définition 1.30. Soit X et Y deux v.a discrêtes à valeurs respectivement dans $\mathcal{X} = \{x_1, x_2, ..., x_n, ...\}$ et $\mathcal{Y} = \{y_1, y_2, ..., y_p, ...\}$.

On appelle loi conditionnelle de X sachant que $Y = y_p$ la loi définie par :

$$P(X = x_n/Y = y_p) = \frac{P(X = x_n, Y = y_p)}{P(Y = y_p)}$$

pour tout $n \in N$.

Exemple 1.16. $X_1 \sim B(p)$, $X_2 \sim B(p)$ indépendantes. Loi de $X_1/X_1 + X_2 = 1$?

Il faut déterminer $P(X_1 = 0/X_1 + X_2 = 1)$ et $P(X_1 = 1/X_1 + X_2 = 1)$

On a : $X_1 + X_2 \sim B(2, p)$

$$P(X_1 = 0/X_1 + X_2 = 1) = \frac{P(X_1 = 0, X_1 + X_2 = 1)}{P(X_1 + X_2 = 1)} = \frac{P(X_1 = 0, X_2 = 1)}{P(X_1 + X_2 = 1)}$$
$$= \frac{P(X_1 = 0)P(X_2 = 1)}{P(X_1 + X_2 = 1)} = \frac{pq}{2pq} = \frac{1}{2}$$

De même $P(X_1 = 1/X_1 + X_2 = 1) = \frac{1}{2}$. Donc

$$X_1/X_1 + X_2 = 1 \sim B(\frac{1}{2})$$

Définition 1.31. Soit (X, Y) un vecteur aléatoire absolument continu de densité $f_{(X,Y)}$.

On appelle loi conditionnelle de X sachant que $Y = y$ la loi définie par la densité :

$$f_X(x/Y = y) = \frac{f_{(X,Y)}(x, y)}{f_Y(y)}$$

pour tout $x \in R$.

Exemple 1.17. Soit un vecteur aléatoire (X, Y) de densité $f(x, y) = xe^{-xy}\ 1_{\{x>0\ ,\ y>1\}}$. Loi de $X/X + Y = y$?

On a après calculs : $f_Y(y) = \frac{1}{y^2}\ 1_{\{\ y>1\}} \implies$

$$f_X(x/Y = y) = \frac{xe^{-xy}1_{\{x>0\ ,\ y>1\}}}{\frac{1}{y^2}1_{\{\ y>1\}}} = xy^2e^{-xy}1_{\{x>0\ \}}$$

Par exemple : $f_X(x/Y = 1) = xe^{-x}1_{\{x>0\ \}}$ donc $X/Y = 1 \sim \gamma(2, 1)$

$f_X(x/Y = 2) = 4xe^{-2x}1_{\{x>0\ \}}$, donc $X/Y = 2 \sim \gamma(2, 2)$

D'une manière générale on remarque que :

$$X/Y = y \sim \gamma(2, y)$$

Définition 1.32. On appelle espérance conditionnelle de X sachant que $Y = y$ l'espérance de la loi conditionnelle de X sachant que $Y = y$, i.e

1. Dans le cas discret :$E(X/Y = y_p) = \sum_{i=1}^{+\infty} x_i P(X = x_i/Y = y_p)$
2. Dans le cas continu : :$E(X/Y = y) = \int_R x \ f_X(x/Y = y)dx$

Exemple 1.18. 1. $E(X_1/X_1 + X_2 = 1) = \frac{1}{2}$ (puisque $X_1/X_1 + X_2 = 1 \sim B(\frac{1}{2})$)

2. $E(X/Y = 1) = 2$, $E(X/Y = 2) = 1$. D'une manière générale

$$E(X/Y = y) = \frac{2}{y}$$

Exemple 1.19. $(X, Y) \sim N_2(m, \Sigma)$ avec $m = (m_1, m_2)^t$ et $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\ \sigma_1\sigma_2 \\ \rho\ \sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ on a alors :

$$f_X(x/Y = y) = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} \exp(-\frac{1}{2\sigma_1^2(1-\rho^2)}(x - m_1 - \frac{\rho\sigma_1(y - m_2)}{\sigma_2})^2)$$

donc

$$X/Y = y \sim N_1(m_1 - \frac{\rho\sigma_1(y - m_2)}{\sigma_2}, \sigma_1^2(1-\rho^2))$$

On déduit :

$$E(X/Y = y) = m_1 - \frac{\rho\sigma_1(y - m_2)}{\sigma_2}$$

et

$$Var(X/Y = y) = \sigma_1^2(1-\rho^2)$$

.

Exemple 1.20 (Généralisation)**.** Soit $\mathbf{X} = (\mathbf{Y}_1, \mathbf{Y}_2) = (X_1, X_2, .X_p, X_{p+1}., X_n)\ R^n$. où $\mathbf{Y}_1 = (X_1, X_2, .X_p)$ et $\mathbf{Y}_2 = (X_{p+1}..., X_n)$

Considérons les décompositions suivantes : $\mathbf{m} = (\mathbf{m}_1, \mathbf{m}_2)$, avec $\mathbf{m}_1 = E(\mathbf{Y}_1), \mathbf{m}_2 = E(\mathbf{Y}_2)$ et

$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}$

où : $\mathbf{\Sigma}_{11} = Var(\mathbf{Y}_1), \mathbf{\Sigma}_{22} = Var(\mathbf{Y}_2)$ et $\mathbf{\Sigma}_{12} = Cov(\mathbf{Y}_1, \mathbf{Y}_2)$.

Alors on peut montrer que

$$\mathbf{Y}_1 \mid \mathbf{Y}_2 = y_2 \sim N_n(\mathbf{m}_1 + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(y_2 - \mathbf{m}_2), \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21})$$

Les lois conditionnelles de tout sous-vecteur gaussien sachant l'autre sont donc gaussiennes.

0n a, en particulier :

$$E(\mathbf{Y}_1 \mid \mathbf{Y}_2 = y_2) = \mathbf{m}_1 + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(y_2 - \mathbf{m}_2)$$

et

$$Var(\mathbf{Y}_1 \mid \mathbf{Y}_2 = y_2) = \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21})$$

La courbe de $y \to E(X/Y = y)$ est dite courbe de régression de X sur Y. Dans le cas gaussien (bivarié), c'est donc une droite dite droite de régression.

On peut définir de même les autres moments conditionnels

1. $E(X^k/Y = y) = \int\limits_R x^k \ f_X(x/Y = y)dx$ moment d'ordre k de la loi conditionnelle
2. Notamment la variance conditionnelle

$$Var(X/Y = y) = E(X^2/Y = y) - (E(X/Y = y))^2$$

Définition 1.33. On appelle espérance conditionnelle de X sachant Y, notée $E(X/Y)$, la variable aléatoire qui prend la valeur $E(X/Y = y)$ lorsque Y prend la valeur y.

Définition 1.34. On appelle moment conditionnel d'ordre k de X sachant Y la variable aléatoire, notée $E(X^k/Y)$ qui prend la valeur $E(X^k/Y = y)$ lorsque Y prend la valeur y.

En particulier

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

Remarque 8. Les variables aléatoires $E(X/Y)$, $E(X^k/Y)$, $Var(X/Y)$ sont des fonctions de la v.a Y. Elles admettent à leur tour des moments, et on a par exemple :

$$E(E(X/Y)) = \int\limits_R E(X/Y = y) f_Y(y)dy = \int\limits_R (\int\limits_R x f_X(x/Y = y)dx) f_Y(y)dy$$

Exemple 1.21. Exemple 2 : $E(X/Y = y) = \dfrac{2}{y} \implies E(X/Y) = \dfrac{2}{Y}$

$Var(X/Y = y) = \dfrac{2}{y^2} \implies Var(X/Y) = \dfrac{2}{Y^2}$

Cas gaussien : $E(X/Y = y) = m_1 - \dfrac{\rho\sigma_1(y - m_2)}{\sigma_2} \implies E(X/Y) = m_1 - \dfrac{\rho\sigma_1(Y - m_2)}{\sigma_2}$

$Var(X/Y = y) = \sigma 1^2(1 - \rho^2) \implies Var(X/Y) = \sigma_1^2(1 - \rho^2)$

Théorème 1.19 (Théorème de la moyenne conditionnelle)**.** *On a*

$$E(E(X/Y)) = E(X).$$

Démonstration. (Dans le cas continu)
On a :

$$\begin{aligned} E(E(X/Y)) &= \int\limits_R (\int\limits_R x f_X(x/Y = y)dx) f_Y(y)dy = \int\limits_R \int\limits_R x f_X(x/Y = y) f_Y(y)dxdy \\ &= \int\limits_R \int\limits_R x \frac{f_{(X,Y)}(x, y)}{f_Y(y)} f_Y(y)dxdy = \int\limits_R \int\limits_R x f_{(X,Y)}(x, y)dxdy \\ &= E(X) \end{aligned}$$

□

Théorème 1.20 (Théorème de la variance conditionnelle)**.** *On a*

$$Var(X) = E(Var(X/Y)) + Var(E(X/Y)).$$

1.6 Convergences

Différents types de convergence existent pour les suites de fonctions (CV simple, CV uniforme), certaines étant plus fortes que d'autres. De la même manière on va définir différents types de convergence pour des suites de v.a.

En général , toute notion de convergence est associée à une notion de distance,. Des distances différentes définissent des notions de convergence différentes.

$X_n \underset{n\longrightarrow+\infty}{\longrightarrow} X$ dans un certain sens si pour une certaine distance d, $\lim\limits_{n\to+\infty} d(X_n, X) = 0$.

En probabilité, pour certains types de convergence, cette notion de distance n'apparait pas de manière explicite, bien qu'elle soit sous-jacente.

1.6.1 Convergence en loi

Définition 1.35. Soit $(X_n)_{n\in N}$ une suite de v.a. La suite $(X_n)_{n\in N}$ est dite converger en loi vers la v.a X si

$$F_{X_n}(x) \underset{n\longrightarrow+\infty}{\longrightarrow} F_X(x)$$

en tout x point de continuité de F_X.

Notation : $X_n \underset{n\longrightarrow+\infty}{\overset{Loi}{\longrightarrow}} X$

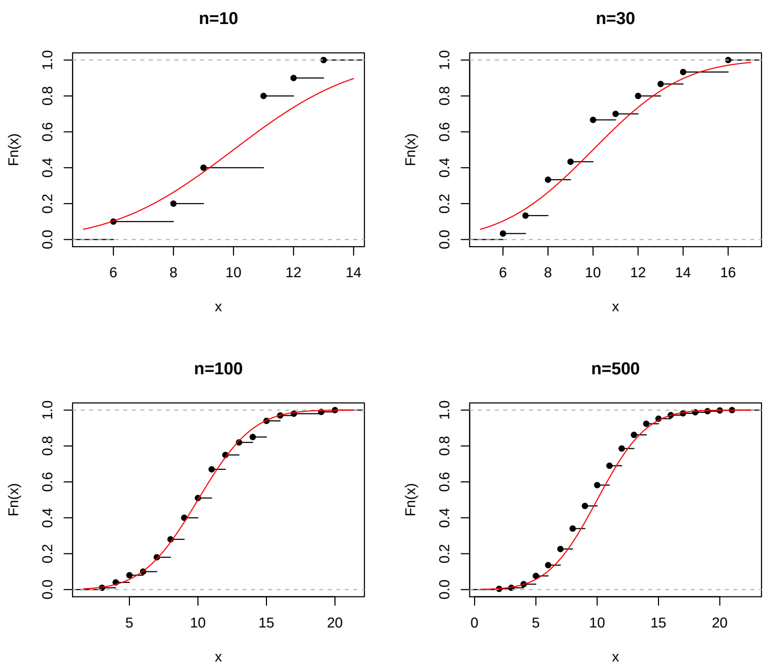

FIGURE 1.10 – Evolution des f.r dans la CV en loi

Théorème 1.21. *Soit Ψ_X et φ_X respectivement les fonctions génératrice et caractéristique de X. Les trois assertions suivantes sont équivalentes :*

1. $X_n \underset{n \longrightarrow +\infty}{\overset{Loi}{\longrightarrow}} X$.
2. $\Psi_{X_n}(u) \underset{n \longrightarrow +\infty}{\longrightarrow} \Psi_X(u)$ *en tout u point du domaine de définition de* Ψ_X.
3. $\varphi_{X_n}(x) \underset{n \longrightarrow +\infty}{\longrightarrow} \varphi_X(x)$ *en tout x de R.*

Si $X_n \underset{n \longrightarrow +\infty}{\overset{Loi}{\longrightarrow}} X$ on a $P(X_n < x) = F_{X_n}(x) \underset{n \longrightarrow +\infty}{\longrightarrow} P(X < x) = F_X(x)$ (en tout x point de continuité de F_X) et donc

$P(a < X_n < b) = F_{X_n}(b) - F_{X_n}(a) \underset{n \longrightarrow +\infty}{\longrightarrow} P(a < X < b) = F_X(b) - F_X(a)$

Conclusion : si n est grand

$$P(X_n < x) \approx P(X < x)$$

$$P(a < X_n < b) \approx P(a < X < b)$$

1.6.2 Convergence en probabilité

Définition 1.36. Soit $(X_n)_{n\in N}$ une suite de v.a. La suite $(X_n)_{n\in N}$ est dite converger en probabilité vers la v.a X si

$$\forall \epsilon > 0 \quad P(|X_n - X| > \epsilon) \underset{n \longrightarrow +\infty}{\longrightarrow} 0$$

Exemple 1.22. Soit $(X_n)_{n\in N}$ une suite de v.a telles que $X_n \sim B(\frac{1}{n})$. Montrons que $X_n \underset{n \longrightarrow +\infty}{\overset{\Pr oba}{\longrightarrow}} 0$

On a : $P(|X_n - 0| > \epsilon) = \begin{cases} 0 & \text{si} \quad \epsilon \geq 1 \\ \frac{1}{n} & \text{si} \quad \epsilon < 1 \end{cases}$

$\Longrightarrow \lim\limits_{n \to +\infty} P(|X_n - 0| > \epsilon) = 0 \quad \forall \epsilon > 0$

Exemple 1.23. Soit $(X_n)_{n\in N}$ une suite de v.a telles que $P(X_n = 0) = 1 - \frac{1}{n^2}$ et $P(X_n = n) = \frac{1}{n^2}$

On a $P(|X_n - 0| > \epsilon) = \frac{1}{n^2}$

$\lim\limits_{n \to +\infty} P(|X_n - 0| > \epsilon) = 0 \quad \forall \epsilon > 0 \Longrightarrow X_n \underset{n \longrightarrow +\infty}{\overset{\Pr oba}{\longrightarrow}} 0$

1.6.3 Convergence presque sûre

Définition 1.37. Soit $(X_n)_{n\in N}$ une suite de v.a. La suite $(X_n)_{n\in N}$ est dite converger presque sûrement vers la v.a X s'il existe un ensemble $A \subset \Omega$, avec P(A)=1, telle que :

$$\forall \omega \in A \quad \lim_{n \to +\infty} X_n(\omega) = X(\omega)$$

Proposition 1.22. *Soit* $(X_n)_{n\in N}$ *telle que :* $X_n \underset{n \longrightarrow +\infty}{\overset{P.S}{\longrightarrow}} X$ *alors* $X_n \underset{n \longrightarrow +\infty}{\overset{\Pr oba}{\longrightarrow}} X$

Théorème à admettre

1.6.4 Convergence en moyenne quadratique

Les convergences en probabilité et presque sûre ne sont pas généralement faciles à établir. On a donc besoin de critères suffisants. La notion de convergence en moyenne quadratique est plus simple, car c'est la convergence de moments d'ordre 2. Elle implique la C.V en probabilité et donc est une condition suffisante de cette dernière.

Définition 1.38. Soit $(X_n)_{n\in N}$ une suite de v.a amettant des moments d'ordre 2. La suite $(X_n)_{n\in N}$ est dite converger en moyenne quadratique vers la v.a X si

$$\lim_{n\to+\infty} E((X_n - X)^2) = 0$$

Exemple 1.24. Soit $(X_n)_{n\in N}$ une suite de v.a telles que $X_n \sim B(\frac{1}{n})$. Montrons que $X_n \underset{n\longrightarrow+\infty}{\overset{m.q}{\longrightarrow}} 0$

On a $E(|X_n - 0|^2) = E({X_n}^2) = \dfrac{1}{n} \underset{n\longrightarrow+\infty}{\longrightarrow} 0$

Exemple 1.25. Soit $(X_n)_{n\in N}$ une suite de v.a telles que $P(X_n = 0) = 1 - \dfrac{1}{n^2}$ et $P(X_n = n) = \dfrac{1}{n^2}$ $X_n \underset{n\longrightarrow+\infty}{\overset{m.q}{\longrightarrow}} 0$?

On a $E(|X_n - 0|^2) = E({X_n}^2) = \dfrac{n^2}{n^2} = 1 \underset{n\longrightarrow+\infty}{\not\rightarrow} 0$

Donc $X_n \underset{n\longrightarrow+\infty}{\overset{m.q}{\not\rightarrow}} 0$

Proposition 1.23. $X_n \underset{n\longrightarrow+\infty}{\overset{m.q}{\longrightarrow}} X \iff \lim_{n\to+\infty} E(X_n) = E(X)$ *et* $\lim_{n\to+\infty} Var(X_n - X) = 0$

Démonstration. $E((X_n - X)^2) = Var(X_n - X) + (E(X_n - X))^2$ □

Corollaire 1.24. *Soit a une constante réelle, on a :* $X_n \underset{n\longrightarrow+\infty}{\overset{m.q}{\longrightarrow}} a \iff \lim_{n\to+\infty} E(X_n) = a$ *et* $\lim_{n\to+\infty} Var(X_n) = 0$

Démonstration. conséquence directe du théorème précédent $E((X_n - X)^2) = Var(X_n - X) + (E(X_n - X))^2$ □

Proposition 1.25. *Soit* $(X_n)_{n\in N}$ *telle que :* $X_n \underset{n\longrightarrow+\infty}{\overset{m.q}{\longrightarrow}} X$ *alors* $X_n \underset{n\longrightarrow+\infty}{\overset{Proba}{\longrightarrow}} X$

Démonstration. On applique l'inégalité de Bienaymé-Tchebychev à $Y = X_n - X$ qui est une variable centrée (de moyenne nulle)

$P(|X_n - X| > \epsilon) \leq \dfrac{E((X_n - X)^2)}{\epsilon^2} \implies \lim_{n\to+\infty} P(|X_n - X| > \epsilon) \leq \lim_{n\to+\infty} \dfrac{E((X_n - X)^2)}{\epsilon^2} = 0$ □

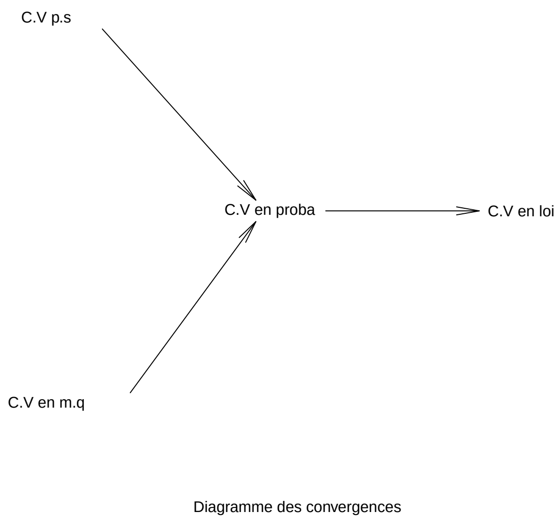

FIGURE 1.11 – Diagramme des CV

1.7 Théorèmes limites

Ce qu'on appelle théorèmes limites, est constitué de deux familles de théorèmes (lois des grands nombres et théorèmes central limite ou de la limite centrale) absolument fondamentaux en probabilités et statistique. Ce sont des résultats complémentaires dans le sens ou les lois des grands nombres établissent que, sous certaines hypothèses, la moyenne d'une suite de v.a de même loi, converge vers une limite fixe (constante) et que le théorème central-limite donne la loi asymptotique de cette moyenne et donc permet de "mesurer" la variation (ou l'erreur) autour de cette limite.

Théorème 1.26 (Loi faible des grands nombres)**.** *Soit* $(X_n)_{n\in N}$ *une suite de v.a i.i.d admettant une espérance finie m. Alors*

$$\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n} \underset{n\longrightarrow+\infty}{\overset{\Pr oba}{\longrightarrow}} m$$

Démonstration. On donne la démonstration, sous l'hypothèse supplémentaire que les v.a (X_n) admettent une variance σ^2 finie. On applique l'inégalité de Bienaymé-Tchebychev à $\overline{X_n}$

$$\forall \quad \epsilon > 0, \quad P(\left|\overline{X_n} - m\right| > \epsilon) \leq \frac{Var(\overline{X_n})}{\epsilon^2} = \frac{Var(X)}{n\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \underset{n\longrightarrow+\infty}{\longrightarrow} 0$$

□

Théorème 1.27 (Loi forte des grands nombres)**.** *Soit* $(X_n)_{n\in N}$ *une suite de v.a i.i.d* *admettant une espérance m et une variance σ^2 finies. Alors*

$$\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow[n \longrightarrow +\infty]{P.S} m$$

Démonstration. La démonstration dans le cas général est complexe. Une démonstration simple existe, cependant, si on suppose l'existence de moments jusqu'à l'ordre 4 pour X_n. $\square$

Exemple 1.26.

1. Soit $(X_n)_{n \in N}$ une suite de v.a i.i.d de loi de Bernoulli $B(p)$. On a
$$\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\textit{nombre de 1 obtenus}}{n} = F_n \xrightarrow[n \longrightarrow +\infty]{P.S} p$$
2. $X_i \sim N(m, \sigma^2)$, $1 \leq i \leq n$ i.i.d $\Longrightarrow \overline{X_n} \xrightarrow[n \longrightarrow +\infty]{P.S} m$
3. $X_i \sim P(\lambda)$, $1 \leq i \leq n$ i.i.d $\Longrightarrow \overline{X_n} \xrightarrow[n \longrightarrow +\infty]{P.S} \lambda$

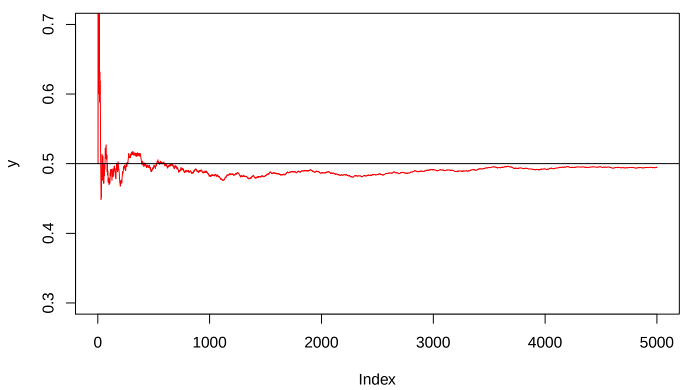

FIGURE 1.12 – Evolution de la moyenne : n=5000

Théorème 1.28 (Théorème central-limite)**.** *Soit $(X_n)_{n \in N}$ une suite de v.a i.i.d admettant une espérance m et une variance σ^2 finies. Alors*

$$\frac{\sum_{i=1}^{n} X_i - nm}{\sqrt{n}\sigma} \xrightarrow[n \longrightarrow +\infty]{Loi} X$$

avec $X \sim N(0, 1)$

Corollaire 1.29. *Sous les mêmes hypothèses, si on pose*

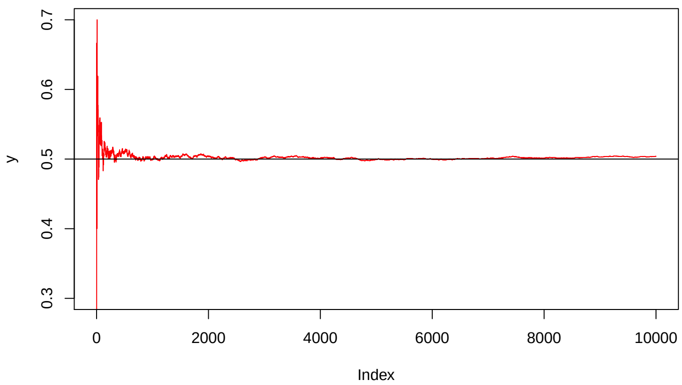

FIGURE 1.13 – Evolution de la moyenne : n=10000

$$\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n} \quad \textit{on a}$$

$$\frac{\overline{X_n} - m}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n \longrightarrow +\infty]{Loi} X$$

avec $X \sim N(0, 1)$

Démonstration. Les démonstrations du théorème et du corollaire sont équivalentes. Un dévelloppement limité à l'ordre 2 de la fonction génératrice de $\frac{\overline{X_n} - m}{\sigma/\sqrt{n}}$ permet d'avoir le résultat.Posons :

$$\frac{\left(\sum_{i=1}^{n} X_i - nm\right)}{\sqrt{n}\sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - m}{\sigma}\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i$$

On a :

$$\Psi_{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i}(t) = \left(\Psi_{Y_i}\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

$$\begin{aligned}
\log\left(\Psi_{Y_i}\left(\frac{t}{\sqrt{n}}\right)\right)^n &= n \log \Psi_{Y_i}\left(\frac{t}{\sqrt{n}}\right) \\
&= n \log\left(\Psi_{Y_i}(0) + \frac{t}{2\sqrt{n}}\Psi'_{Y_i}(0) + \frac{t^2}{2n}\Psi'_{Y_i}(0) + o\left(\frac{t^2}{2n}\right)\right) \\
&= n\left(\frac{t^2}{2n} + o\left(\frac{t^2}{2n}\right)\right)
\end{aligned}$$

$$\lim_{n \to +\infty} \log \Psi_{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i}(t) = \frac{t^2}{2},$$

d'où

$$\lim_{n\to+\infty} \Psi_{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i}(t) = e^{\dfrac{t^2}{2}}$$

et donc

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i \underset{n\longrightarrow+\infty}{\overset{Loi}{\longrightarrow}} N(0,1)$$

$\square$

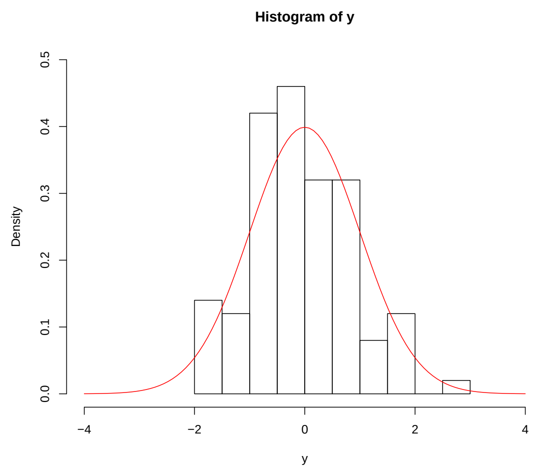

FIGURE 1.14 – Distribution de la moyenne empirique : n=10

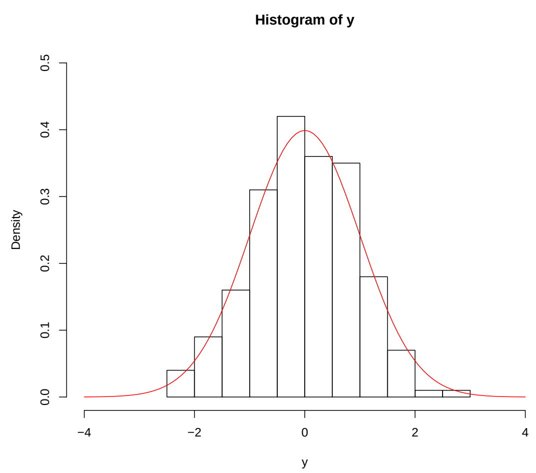

FIGURE 1.15 – Distribution de la moyenne empirique : n=50

Exemple 1.27. Approximation normale de la loi binomiale

Soit $(X_n)_{n\in N}$ une suite de v.a i.i.d de loi de Bernoulli $B(p)$. On a $\sum\limits_{i=1}^{n} X_i \sim B(n,p)$ et si n grand

$$\sum_{i=1}^{n} X_i \approx N(np, npq) \implies B(n,p) \approx N(np, npq)$$

Exemple 1.28. Approximation normale de la loi de Poisson.

Soit $(X_n)_{n\in N}$ une suite de v.a i.i.d de loi de Bernoulli $P(\lambda)$. On a $\sum\limits_{i=1}^{n} X_i \sim P(n\lambda)$ et si n grand

$$\sum_{i=1}^{n} X_i \approx N(\lambda, \lambda) \implies P(n\lambda) \approx N(\lambda, \lambda)$$

Il existe une version multidimensionnelle du TCL qui est donnée dans le théorème suivant :

Théorème 1.30 (Théorème central-limite multidimensionnel)**.** *Soit* $(\mathbf{X}_n)$ *une suite de vecteurs aléatoires, à valeurs dans* R^p, *i.i.d admettant une moyenne* $\mathbf{m}$ *et une matrice de variances-covariances* $\boldsymbol{\Sigma}$ *alors*

$$\frac{\sum\limits_{i=1}^{n} \mathbf{X}_i - n\mathbf{m}}{\sqrt{n}} \underset{n\longrightarrow+\infty}{\overset{Loi}{\longrightarrow}} \mathbf{X}$$

avec $\mathbf{X} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$ *ou de manière équivalente*

$$\sqrt{n}(\overline{\mathbf{X}} - \mathbf{m}) \underset{n\longrightarrow+\infty}{\overset{Loi}{\longrightarrow}} \mathbf{X}$$

Bibliographie

[1] P. Brémaud. *Initiation aux Probabilités et aux chaînes de Markov.* Springer, 2009.

[2] Jean Jacod, Philip Protter, et al. *L'essentiel en théorie des probabilités.* Cassini, 2003.

[3] Sheldon M Ross. *Initiation aux probabilités.* PPUR presses polytechniques, 2007.