Statistique Mathématique

Salim Lardjane

Université Bretagne Sud

Partie I
Rappels et Compléments de Probabilités

Probabilités

L'objectif de cette première partie est de regrouper les résultats probabilistes utilisés en Statistique Mathématique ; on décrira les lois de probabilités usuelles et on fournira les théorèmes de convergence les plus utiles.

La plupart des résultats ne seront qu'énoncés.

Les résultats avancés seront étoilés.

Pour plus de détails sur les aspect mathématiques, on pourra consulter :

Méthodes Statistiques, 3ème édition, Philippe Tassi, Economica 2004.

Cours de Probabilités, 3ème édition, Alain Monfort, Economica 1996.

Cours de Statistique Mathématique, 3ème édition, Alain Monfort, Economica 1997.

NOTIONS THEORIQUES

Espace probabilisé

On note $(\Omega, \mathcal{A}, \mathbb{P})$ l'**espace probabilisé fondamental**.

Ω est appelé **référentiel**.

$\mathcal{A}$ est une **tribu** sur Ω. C'est un ensemble de sous-ensembles de Ω, appelés **événements**

$\mathbb{P}$ est une **mesure de probabilité** définie sur l'espace probabilisable $(\Omega, \mathcal{A})$.

Concrètement, Ω est l'ensemble des résultats possibles d'une expérience aléatoire et un événement est un ensemble de résultats dont on peut calculer la probabilité à l'aide de $\mathbb{P}$.

Exemple classiqe : Lancer d'un dé

$\Omega = \{1, 2, 3, 4, 5, 6\}$

$\mathcal{A} = \mathcal{P}(\Omega)$, ensemble des parties de Ω.

$\mathbb{P}(A) = \frac{\#A}{6}$ où A est un événement et $\#A$ son cardinal.

Ainsi, si $A = \{2, 4, 6\}$, A correspond au fait d'obtenir un nombre pair à l'issue du lancer et $\mathbb{P}(A) = 3/6 = 1/2$.

Variables aléatoires

On appelle **variable aléatoire** une application X définie sur $(\Omega, \mathcal{A})$ à valeurs dans un espace probabilisable $(\mathcal{H}, \mathcal{B})$, où $\mathcal{B}$ est une tribu d'événements de $\mathcal{H}$, et qui est **mesurable**, c'est-à-dire que :

pour tout événement B de $\mathcal{B}$, l'ensemble $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$ est un événement de $\mathcal{A}$, c'est-à-dire qu'on peut calculer sa probabilité à l'aide de $\mathbb{P}$.

Variables aléatoires

Exemple : Lancer d'un dé.

Soit X l'application sur $(\Omega, \mathcal{A})$ définie par $X(\omega) = 2\omega$ à valeurs dans $\mathcal{H} = \{2, 4, 6, 8, 10, 12\}$.

Alors, pour toute partie B de $\mathcal{H}$, on a

$$X^{-1}(B) = \{\omega \in \Omega : 2\omega \in B\}$$

C'est un sous-ensemble de Ω et donc un événement de $\mathcal{A} = \mathcal{P}(\Omega)$.

Donc X est mesurable.

C'est donc une variable aléatoire.

Loi de probabilité

La **loi de probabilité** de X est la **mesure de probabilité** $\mathbb{P}^X$ sur $(\mathcal{H}, \mathcal{B})$ telle que pour tout événement B de $\mathcal{B}$, on a

$$\mathbb{P}^X(B) = \mathbb{P}(X \in B) = \mathbb{P}(X^{-1}(B)).$$

Si $\mathcal{H} = \mathbb{R}$, X est appelée **variable aléatoire réelle** (v.a.r.).

Si $\mathcal{H} = \mathbb{R}^p$, on dit que X est un **vecteur aléatoire** de dimension p.

Si $\mathcal{H}$ est fini ou infini dénombrable, c'est-à-dire si ses éléments peuvent être indicés par tout ou partie de $\mathbb{Z}$, X est une **variable aléatoire discrète**.

Notations

Si X est une variable aléatoire réelle donnée, on note respectivement F_X, f_X, g_X et ϕ_X la **fonction de répartition**, la **fonction de densité**, la **fonction génératrice** et la **fonction caractéristique** de X et on adopte les mêmes notations si X est un vecteur aléatoire.

La fonction de densité est parfois appelée **fonction de masse** lorsque X une v.a.r. discrète.

Toutes ces fonctions **caractérisent** la loi de X.

On en verra des exemples sur des cas particuliers.

Paramètres

De façon générale, la loi de probabilité d'une variable X, et donc les fonctions la caractérisant, dépendent d'un certain nombre de paramètres qui influent sur leurs forme.

Le symbole θ représente les paramètres de la loi $\mathbb{P}^X$, qui est notée $\mathbb{P}^X(\theta)$ ou, plus souvent, P_θ, avec $\theta \in \Theta$.

Θ désigne l'**espace des paramètres**.

Il peut être contenu dans $\mathbb{R}$ – cas d'un seul paramètre – ou dans $\mathbb{R}^p$ – cas de plusieurs paramètres.

Dans ce dernier cas, on dit que $\theta = (\theta_1, \theta_1, \ldots, \theta_p)$ est un **paramètre vectoriel**.

Paramètres fonctionnels

Considérons une v.a. réelle ou vectorielle X de loi P.

On appelle **paramètre fonctionnel** (ou intrinsèque) un paramètre θ tel que $\theta = T(P)$ où T est une **fonctionnelle** de $\mathcal{P}$ dans Θ, où $\mathcal{P}$ désigne l'ensemble de toutes les lois de probabilité.

Exemple : l'espérance mathématique

$$\begin{aligned}
\mathbb{E}(X) &= \int X d\mathbb{P} = \int x dP(x) \\
&= \int_{-\infty}^{\infty} x f_X(x) dx \quad (X \text{ v.a.r. continue}) \\
&= \sum_{i \in \mathbb{Z}} x_i f_X(x_i) \quad (X \text{ v.a.r discrète})
\end{aligned}$$

On en verra des exemples sur des cas particuliers.

Indépendance

Dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, la probabilité $\mathbb{P}(B)$ d'un événement B mesure la "chance" que B a de se réaliser.

Si on suppose que l'on dispose de l'information supplémentaire "l'événement A va se réaliser", A ayant une probabilité $\mathbb{P}(A) > 0$, la "chance" que B a de réaliser est modifiée.

En particulier, si $A \cap B = A$, c'est-à-dire si A implique B, cette "chance" devient une certitude.

Si $A \cap B = \emptyset$, cette "chance" devient une impossibilité.

Indépendance

Ainsi, dans l'hypothèse où l'on sait que A se réalise, il est naturel de mesurer la "chance" qu'un événement quelconque B a de se réaliser par un nombre proportionnel à $\mathbb{P}(A \cap B)$, c'est-à-dire de la forme $k\mathbb{P}(A \cap B)$ avec $k > 0$.

Pour que la fonction ainsi définie soit une probabilité, on impose que

$$k\mathbb{P}(A \cap \Omega) = k\mathbb{P}(A) = 1$$

d'où $k = 1/\mathbb{P}(A)$.

Ceci nous amène aux définitions suivante.

Indépendance

Etant donné un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et un événement $A \in \mathcal{A}$ tel que $\mathbb{P}(A) > 0$, on appelle **probabilité conditionnelle** d'un événement B **sachant** A, le nombre

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

Indépendance

Deux événement A et B sont dits **indépendants** si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Plus généralement $A_1, \ldots, A_n$ sont dits **indépendants dans leur ensemble** si

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \cdots \mathbb{P}(A_{i_k})$$

pour toute suite $(i_1, \ldots, i_k)$ d'entiers distincts appartenant à $\{1, \ldots, n\}$.

Indépendance

Exemple (exercice) : On jette deux fois une pièce de monnaie : $\Omega = \{P, F\}^2$ et on suppose $(\Omega, \mathcal{P}(\Omega)$ muni de la **probabilité uniforme** : $\mathbb{P}(A) = \#A/\#\Omega$.

Montrer que les événements

$$\begin{aligned} A &= \{PP, PF\}, \\ B &= \{PP, FP\}, \\ C &= \{PP, FF\} \end{aligned}$$

sont indépendants deux à deux mais pas indépendants dans leur ensemble.

Indépendance

On dit que n variables ou vecteurs aléatoires $X_1, \ldots, X_n$ de $(\Omega, \mathcal{A}, \mathbb{P})$ dans $(\mathbb{R}^{n_i}, \mathcal{B}^{n_i})$ respectivement, sont **indépendants** si

$$\mathbb{P}(X_1 \in A_1, \ldots, X_n \in A_n) = \mathbb{P}(X_1 \in A_1) \cdots \mathbb{P}(X_n \in A_n)$$

pour tous $A_i \in \mathcal{B}^{n_i}$, $i = 1, \ldots, n$.

Convergence presque sûre

Soit $X_1, X_2, \ldots, X_n, \ldots$ une suite de variables aléatoires réelles. On dit que la suite converge **presque surement** vers la variable aléatoire X et on note

$$X_n \rightarrow_{p.s.} X$$

si $X_n(\omega) \rightarrow X(\omega)$ pour tout $\omega \in C$ avec $\mathbb{P}(C) = 1$, ce qu'on note également

$$\mathbb{P}(X_n \rightarrow X) = 1.$$

Convergence presque sûre

On peut montrer que la définition précédente équivaut à avoir pour tout $\varepsilon > 0$,

$$\mathbb{P}(\sup_{m \geq n} |X_m - X| > \varepsilon) \to 0$$

lorsque $n \to \infty$.

Convergence en probabilité

Soit $X_1, X_2, \ldots, X_n, \ldots$ une suite de variables aléatoires réelles. On dit que la suite converge **en probabilité** vers la variable aléatoire X et on note

$$X_n \to_{\mathbb{P}} X$$

si pour tout $\varepsilon > 0$

$$\mathbb{P}(|X_n - X| > \varepsilon) \to 0$$

lorsque $n \to \infty$.

Convergence en loi

Soit Π l'ensemble des probabilités dans $(\mathbb{R}^m, \mathcal{B}^m)$.

On souhaite définir une notion de convergence sur Π.

On dit qu'une suite de mesures de probabilité (P_n) converge **étroitement** vers P, et on note $P_n \rightarrow_{et.} P$ si

$$\int f dP_n \rightarrow \int f dP$$

pour toute fonction f de $\mathbb{R}^m$ dans $\mathbb{R}$ continue et bornée, où les intégrales sont des **intégrales de Lebesgue** (cf. A. Monfort - Cours de Probabilité).

Convergence en loi

Un critère pratique de convergence étroite est le suivant :

Si chaque probabilité P_n a une densité g_n et si $g_n(x) \to g(x)$ pour tout x, g étant une densité de probabilité associée à une mesure de probabilité P, alors $P_n \to_{et.} P$.

En fait il suffit que g_n converge vers g **presque partout** (cf. A. Monfort - Cours de Probabilités).

Convergence en loi

Soit (X_n) une suite de variables ou de vecteurs aléatoires et X une variable ou un vecteur aléatoire, respectivement, à valeurs dans $(\mathbb{R}^m, \mathcal{B}^m)$.

On dit que X_n converge **en loi** vers X, lorsque $n \to \infty$ et on note $X_n \to_L X$ si

$$\mathbb{P}^{X_n} \to_{et.} \mathbb{P}^X,$$

c'est-à-dire si

$$\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$$

pour toute fonction f continue bornée.

Théorème porte-manteau*

Les cinq assertions suivantes sont équivalentes :

1. X_n converge en loi vers X ;
2. pour toute fonction φ bornée et uniformément continue,

$$\mathbb{E}\left[\varphi(X_n)\right] \;\to\; \mathbb{E}\left[\varphi(X)\right];$$

3. pour tout fermé F,

$$\limsup_n \; \mathbb{P}\left(X_n \in F\right) \;\leq\; \mathbb{P}\left(X \in F\right);$$

4. pour tout ouvert O,

$$\liminf_n \mathbb{P}\left(X_n \in O\right) \;\geq\; \mathbb{P}\left(X \in O\right);$$

5. pour tout borélien A dont la frontière ∂A vérifie $\mathbb{P}\left(X \in \partial A\right) = 0$,

$$\mathbb{P}\left(X_n \in A\right) \to \; \mathbb{P}\left(X \in A\right).$$

Loi forte des Grands Nombres

Soit (X_i) une suite de variables aléatoires réelles indépendantes et identiquement distribuées (i.e. de même loi), en abrégé i.i.d.. On note

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

leur moyenne arithmétique et $m = \mathbb{E}(X_i)$ pour tout i.

Alors,

$$\bar{X}_n \to_{p.s.} m.$$

Lois faibles des Grands Nombres

Soit (X_i) une suite de v.a.r. de même loi, **non corrélées**.

On note $\mathbb{E}(X_i) = m$ l'espérance des X_i.

On note $V(X_i) = \sigma^2$ leur variance :

$$V(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right).$$

On note $cov(X_i, X_j)$ leurs covariances :

$$cov(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right).$$

Ces dernières sont par hypothèses (variables non corrélées) égales à 0.

Alors,

$$\bar{X}_n \to_{\mathbb{P}} m.$$

Lois faibles des Grands Nombres

Soit (X_i) une suite de v.a.r. i.i.d. et soit $\mathbb{E}(X_i) = m < \infty$ pour tout i.

Alors

$$\bar{X}_n \rightarrow_{\mathbb{P}} m.$$

Convergence faible vers une constante

Une condition suffisante pour qu'une suite de v.a.r. (X_n) converge en probabilité vers une **constante** réelle a est :

$$\mathbb{E}(X_n) \to a$$

et

$$V(X_n) \to 0.$$

Théorème Central Limite

Soit (X_i) une suite de v.a.r. i.i.d. avec $\mathbb{E}(X_i) = m$, $V(X_i) = \sigma^2$ pour tout i. On a alors le résultat suivant :

$$\frac{\sqrt{n}(\bar{X}_n - m)}{\sigma} \rightarrow_L \mathcal{N}(0, 1).$$

Théorème Central Limite Multidimensionnel

Soit X un vecteur aléatoire de $\mathbb{R}^p$, (X_i) une suite de vecteurs indépendants et de même loi que X.

On note $m = \mathbb{E}(X)$, et $V(X) = \Sigma$ désigne la matrice de variances-covariances de X.

Alors,

$$\sqrt{n}(\bar{X}_n - m) \rightarrow_L \mathcal{N}(0, \Sigma).$$

Théorème de Paul Levy

Soit (X_i) une suite de v.a.r. et on note $\phi_i(u)$ la fonction caractéristique de X_i. Alors :

a) Si $X_i \to_L X$ et si X admet pour fonction caractéristique $\phi(u)$, on a

$$\lim_{i \to \infty} \phi_i(u) = \phi(u) \quad \forall u \in \mathbb{R}$$

b) Si $\phi_i(u)$ converge simplement vers une certaine fonction $\phi(u)$ lorsque $i \to \infty$, et si $\phi(u)$ est continue en zéro, alors $\phi(u)$ est la fonction caractéristique d'une v.a.r. X telle que

$$X_i \to_L X.$$

Convergence en probabilité et en loi

Si $X_n \to_{\mathbb{P}} X$, alors $X \to_L X$.

Si $X_n \to_L a$ avec $a \in \mathbb{R}$, alors $X_n \to_{\mathbb{P}} a$.

Si $X_n \to_L X$ et $Y_n \to_{\mathbb{P}} a$, alors :

a) $X_n + Y_n \to_L X + a$

b) $X_n Y_n \to_L aX$

c) $\frac{X_n}{Y_n} \to_L \frac{X}{a} \quad (a \neq 0)$

Théorèmes de Slutsky

Si $X_n \to_{\mathbb{P}} X$ et $g : \mathbb{R} \to \mathbb{R}$ est une application continue, alors $g(X_n) \to_{\mathbb{P}} g(X)$.

Soient X_n et X des vecteurs aléatoires de $\mathbb{R}^p$. Si $X_n \to_{\mathbb{P}} X$ et $g : \mathbb{R}^p \to \mathbb{R}^q$ est une application continue, alors $g(X_n) \to_{\mathbb{P}} g(X)$.

Soient X_n et X des vecteurs aléatoires de $\mathbb{R}^p$. Si $X_n \to_L X$ et $g : \mathbb{R}^p \to \mathbb{R}^q$ est une application continue, alors $g(X_n) \to_L g(X)$.

Applications et convergence en loi 1

Soit (b_n) une suite de constante réelles positives telle que $b_n \to \infty$ lorsque $n \to \infty$, soit a une constante réelle et soit (X_n) une suite de v.a.r. telle que

$$b_n(X_n - a) \to_L \mathcal{N}(0, \sigma^2)$$

Soit $g : \mathbb{R} \to \mathbb{R}$ une application dérivable. Alors,

$$b_n(g(X_n) - g(a)) \to_L \mathcal{N}(0, \left(g'(a)\right)^2 \sigma^2)$$

Applications et convergence en loi 2

Soit (b_n) une suite de constante réelles positives telle que $b_n \to \infty$ lorsque $n \to \infty$, soit a un vecteur constant de $\mathbb{R}^p$ et soit (X_n) une suite de vecteurs aléatoires de $\mathbb{R}^p$ tels que

$$b_n(X_n - a) \to_L \mathcal{N}(0, \Sigma)$$

Soit $g : \mathbb{R}^p \to \mathbb{R}^q$ une application différentiable et G sa matrice $q \times p$ des dérivées premières $\frac{\partial g_i(u)}{\partial u_j}$. Alors,

$$b_n(g(X_n) - g(a)) \to_L \mathcal{N}(0, G(a)\Sigma G'(a))$$

où $'$ désigne la transposée.

Transformations de variables aléatoires

Supposons que $k \geq 1$ variables aléatoires $Y_1, \ldots, Y_k$ sont obtenus comme transformation de k variables aléatoires $X_1, \ldots, X_k$ dont on connait la densité conjointe $f(x_1, \ldots, x_k)$.

Quelle est la densité conjointe de $Y_1, \ldots, Y_k$?

Notons g la transformation étudiée.

On suppose que g est bijective, donc inversible et on note

$$x_i = g_i^{-1}(y_1, \ldots, y_k)$$

pour $i = 1, \ldots, k$.

Transformations de variables aléatoires

Le *Jacobien* $J(g^{-1})$ de la transformation est le déterminant de la matrice des dérivée partielles

$$\frac{\partial g_i^{-1}(y_1, \ldots, y_k)}{\partial y_j}, \quad i, j = 1, \ldots, k$$

La fonction de densité conjointe de $Y_1, \ldots, Y_k$ est alors donnée par :

$$p(y_1, \ldots, y_k) = f(g^{-1}(y_1, \ldots, y_k)) \cdot |J(g^{-1})|$$

Des exemples seront vus en TD.

LOIS DE PROBABILITES USUELLES

Loi de Bernoulli

On appelle *variable indicatrice* X la variable à à valeurs dans $\{0, 1\}$ telle que

$$\mathbb{P}(X = 1) = \mathbb{P}^X(\{1\}) = p$$

$$\mathbb{P}(X = 0) = \mathbb{P}^X(\{0\}) = q \quad (q = 1 - p)$$

avec $p \in [0, 1]$.

On note $\mathcal{B}(p)$ la loi de la variable X et on l'appelle *loi de Bernoulli* de paramètre p.

On peut démontrer (le faire à titre d'exercice) que

$$\mathbb{E}(X) = p$$

$$V(X) = pq = p(1 - p)$$

La fonction caractéristique correspondante est

$$\phi_X(t) = \mathbb{E}(e^{iXt}) = q + pe^{it}.$$

Loi de Poisson

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

$$\mathbb{P}(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad \lambda > 0,\ x \in \mathbb{N}$$

Les *moments* de la loi $\mathcal{P}(\lambda)$ sont donnés par

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

$$V(x) = \lambda$$

La loi de Poisson est la seule loi discrète, définie sur $\mathbb{N}$, dont l'espérance est égale à la variance.

Sa fonction caractéristique est

$$\phi_X(t) = \mathbb{E}(e^{iXt}) = e^{\lambda(e^{it}-1)}.$$

Loi de Poisson

Soient X et Y deux variables aléatoires indépendantes suivant respectivement des lois de Poisson $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$. Alors, la variable $X+Y$ suit une loi de Poisson de paramètre $\lambda+\mu$.

Le démontrer à titre d'exercice.

Comportement asymptotique. Soit X une variable de loi $\mathcal{P}(\lambda)$. Alors, la variable

$$\frac{X-\lambda}{\sqrt{\lambda}}$$

converge en loi vers la loi normale centrée réduite $\mathcal{N}(0,1)$ lorsque λ tend vers l'infini.

Illustrer ce résultat par une simulation sous R.

Loi binomiale

On considère n tirage équiprobables indépendants dans une population composée de deux types d'éléments, le premier en proportion p, le deuxième en proportion $q = 1 - p$.

Soit X le nombre d'éléments du premier type présents dans 'échantillon de taille n ainsi obtenu.

Alors, X est une variable aléatoire à valeurs dans $\{0, 1, \ldots, n\}$.

La loi de X est appelée *loi binomiale* de paramètres n et p et elle est notée $\mathcal{B}(n, p)$.

Plus généralement, on dit qu'une v.a. X suit une loi binomiale de paramètres n et p si

$$\mathbb{P}^X(\{x\}) = \mathbb{P}(X = x) = C_n^x p^x (1-p)^{n-x}$$

pour $x \in \{0, 1, \ldots, n\}$.

Loi binomiale

Les moments de la loi $\mathcal{B}(n,p)$ sont donnés par

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

$$V(X) = npq.$$

Le démontrer à titre d'exercice.

Sa fonction caractéristique est donnée par

$$\phi_X(t) = \mathbb{E}(e^{iXt}) = (q + pe^{it})^n$$

Loi binomiale

Soient $X_1, \ldots, X_n$ des variables aléatoires indépendants suivant la même loi de Bernoulli $\mathcal{B}(p)$. Alors, la variable aléatoire $\sum_{i=1}^{n} X_i$ suit la loi binomiale $\mathcal{B}(n, p)$.

Le démontrer à titre d'exercice.

Soient X et Y deux variables indépendantes suivant respectivement les lois $\mathcal{B}(n, p)$ et $\mathcal{B}(m, p)$. Alors, la variable $X + Y$ suit la loi binomiale $\mathcal{B}(n + m, p)$.

Le démontrer à titre d'exercice. On pourra utiliser le résultat suivant : Soient X et Y deux v.a. indépendants de fonction caractéristiques $\phi_X(t)$ et $\phi_Y(t)$ respectivement. Alors, la fonction caractéristique de $X + Y$ est donnée par $\phi_X(t)\phi_Y(t)$.

Loi binomiale

Comportement asymptotique. On a les deux résultats suivants :

a) Sous les hypothèses : i) p est fonction de n, ii) $\lim_{n\to\infty} np(n) = \lambda > 0$, on a

$$\lim_{n\to\infty} \mathbb{P}^X(\{x\}) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad \forall x \in \mathbb{N}$$

Par conséquent, sous les hypothèses i) et ii), il y a a convergence en loi de la loi binomiale vers la loi de Poisson de paramètre λ.

b) Sous l'hypothèse : p fixé, on a

$$\frac{X - np}{\sqrt{npq}} \to_L \mathcal{N}(0,1)$$

Ceci résulte du TCL.

Loi multinomiale

Il s'agit d'une généralisation de la loi binomiale.

Considérons une population composée d'éléments de m types, chaque type j en proportion p_j, avec $p_j > 0$, $\sum_{j=1}^{m} p_j = 1$.

On tire dans cette population, de façon équiprobable et indépendante, un échantillon de taille n et on défini les variables aléatoires N_j ($1 \leq j \leq m$), à valeurs dans $\{0, 1, \ldots, n\}$ représentant le nombre d'éléments de type j figurant dans l'échantillon.

On dit que le vecteur $(N_1, \ldots, N_m)$ suit une loi multinomiale de paramètres $n, p_1, \ldots, p_m$, notée $\mathcal{M}(n, p_1, \ldots, p_m)$.

Loi multinomiale

Plus généralement, on dit que le vecteur aléatoire $N = (N_1, \ldots, N_m)$ suit la loi $\mathcal{M}(n, p_1, \ldots, p_m)$ si

$$\mathbb{P}^N(\{(n_1, \ldots, n_m)\}) = \frac{n!}{n_1! \ldots n_m!} p_1^{n_1} \cdots p_m^{n_m}$$

avec $p_j > 0$ pour $j = 1, \ldots, m$,

$$\sum_{j=1}^{m} p_j = 1$$

et

$$\sum_{j=1}^{m} N_j = n, \quad \sum_{j=1}^{m} n_j = n.$$

Démontrer à titre d'exercice que la loi de N_j est $\mathcal{B}(n, p_j)$ pour $j = 1, \ldots, m$.

Loi multinomiale

Les moments de la loi $\mathcal{M}(n, p_1, \ldots, p_m)$ sont donnés par

$$\mathbb{E}(N_j) = np_j$$

$$V(N_j) = np_jq_j$$

$$\text{cov}(N_j, N_l) = -np_jp_l, \quad j \neq l$$

La fonction caractéristique de la loi $\mathcal{M}(n, p_1, \ldots, p_m)$ est donné par :

$$\phi_N(t) = \mathbb{E}(e^{it'N}) = \left(\sum_{j=1}^{m} p_j e^{it_j}\right)^n$$

Loi hypergéométrique

On considère un tirage équiprobable *sans remise* de n éléments dans une population de taille N ($n \leq N$).

On s'intéresse à un type donné d'éléments de la population, que l'on supposera être en proportion p (Np est donc entier).

Soit X le nombre d'éléments du type étudié présents dans l'échantillon de taille n obtenu.

La loi de X est appelée loi hypergéométrique de paramètres N, n, p et elle est notée $\mathcal{H}(N, n, p)$.

Loi hypergéométrique

Plus généralement, X suit la loi $\mathcal{H}(N, n, p)$ si :

$$\mathbb{P}(X = x) = \mathbb{P}^X(\{x\}) = \frac{C_{Np}^{x} C_{nq}^{n-x}}{C_N^n}$$

pour $\max(0, n - Nq) \leq x \leq \min(n, Np)$ et $q = 1 - p$.

Ses moments sont donnés par :

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

$$V(X) = \frac{N-n}{N-1} npq$$

Loi hypergéométrique

Comportement asymptotique. Sous l'hypothèse n et p fixés, on a :

$$\lim_{N \to \infty} \mathbb{P}(X = x) = C_n^x p^x q^{n-x}$$

Ainsi, X converge en loi vers la loi binomiale $\mathcal{B}(n, p)$.

Loi binomiale négative

On considère une population dont une proportion p est composée d'éléments d'un type donné.

On désire obtenir n éléments de ce type en procédant à une suite de tirages équiprobables et indépendants.

Soit Y la variable aléatoire désignant le nombre de tirages nécessaires pour obtenir les n éléments voulus.

La loi de $X = Y - n$ est appelée loi binomiale négative de paramètres n et p et elle est notée $\mathcal{BN}(n, p)$.

Loi binomiale négative

Plus généralement, X suit une loi binomiale négative de paramètres n et p si :

$$\mathbb{P}(X = x) = C_{n+x-1}^{n-1} p^n q^x$$

pour $x \in \mathbb{N}$.

Ses moments sont donnés par :

$$\mathbb{E}(X) = n\frac{q}{p}$$

$$V(X) = n\frac{q}{p^2}$$

et sa fonction caractéristique par :

$$\phi_X(t) = \left(\frac{p}{1 - qe^{it}}\right)^n$$

Loi géométrique

Dans le cas $n = 1$, la loi de la variable Y désignant le nombre de tirages effectués jusqu'à l'obtention du premier éléments du type désiré porte le nom de loi géométrique ou loi de Pascal de paramètre p.

Sa fonction de masse est donnée par :

$$\mathbb{P}(Y = y) = pq^{y-1}$$

pour $y \in \mathbb{N} - \{0\}$.

On a :

$$\mathbb{E}(Y) = \frac{1}{p}, \quad V(Y) = \frac{q}{p^2}$$

Sa fonction caractéristique est :

$$\phi_Y(t) = \frac{pe^{it}}{1 - qe^{it}}$$

Loi uniforme continue

X suit une loi uniforme sur le segment $[a, b]$ $(a < b)$ si sa densité est donnée par :

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

et on note $X \sim \mathcal{U}_{[a,b]}$.

Si A est un ensemble quelconque, $\mathbb{I}_A(x)$ désigne la *fonction indicatrice* de A.

Celle-ci est telle que $\mathbb{I}_A(x) = 1$ si $x \in A$ et 0 sinon.

Loi uniforme continue

Si $X \sim \mathcal{U}_{[a,b]}$, ses moments sont donnés par :

$$\mathbb{E}(X) = \frac{a+b}{2}$$

$$V(X) = \frac{(b-a)^2}{12}$$

Sa fonction caractéristique est donnée par :

$$\phi_X(t) = \frac{e^{itb} - e^{ita}}{it(b-a)}$$

Si $a = 0$ et $b = 1$, on a :

$$\phi_X(t) = \frac{e^{it} - 1}{it}$$

Loi gamma

X suit une loi gamma de paramètres $p > 0$ et $\theta > 0$, notée $\gamma(p, \theta)$ si sa densité est donné par :

$$f(x) = \frac{\theta^p}{\Gamma(p)} e^{-\theta x} x^{p-1} \mathbb{I}_{\mathbb{R}+}(x)$$

où

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

Si le paramètre d'échelle θ est égal à 1 on note la loi $\gamma(p)$; si $\theta \neq 1$, la v.a. $Y = \theta X$ est de loi $\gamma(p)$.

La densité de la loi $\gamma(p)$ est

$$f(x) = \frac{1}{\Gamma(p)} e^{-x} x^{p-1} \mathbb{I}_{\mathbb{R}+}(x)$$

Si $p = 1$, la loi $\gamma(1, \theta)$ porte le nom de *loi exponentielle* de paramètre θ et elle est notée $\mathcal{E}(\theta)$.

Propriétés de la fonction gamma

$$
\begin{aligned}
\Gamma(p) &= (p-1)\Gamma(p-1), \quad p > 0 \\
\Gamma(p) &= (p-1)!, \quad p \in \mathbb{N} - \{0\} \\
\Gamma(1/2) &= \sqrt{\pi} \\
\Gamma(p+1) &\sim \left(\frac{p}{e}\right)^p \sqrt{2\pi p}\left(1 + \frac{1}{12p}\right), \quad p \to \infty
\end{aligned}
$$

Loi gamma

Si $X \sim \gamma(p, \theta)$, on a pour tout $r \geq 0$:

$$\mathbb{E}(X^r) = \frac{\Gamma(p+r)}{\theta^r \Gamma(p)}$$

On en déduit :

$$\mathbb{E}(X) = \frac{p}{\theta}$$

$$V(X) = \frac{p}{\theta^2}$$

La fonction caractéristique de la loi $\gamma(p, \theta)$ est donnée par :

$$\phi_X(t) = \frac{1}{(1 - it/\theta)^p}$$

Loi gamma

Soient X et Y deux v.a.r. indépendantes suivant respectivement les lois $\gamma(p, \theta)$ et $\gamma(q, \theta)$.

Alors, la v.a. $X + Y$ suit la loi $\gamma(p + q, \theta)$.

Comportement asymptotique. Si $X \sim \gamma(p)$, on peut montrer que :

$$\frac{X - p}{\sqrt{p}} \rightarrow_L \mathcal{N}(0, 1)$$

lorsque $p \rightarrow \infty$.

Loi logistique

On dit qu'une v.a.r. X suit la loi logistique si sa densité est donné par :

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

La loi logistique est fréquemment définie par sa fonction de répartition :

$$F(x) = \frac{1}{1+e^{-x}}, \quad x \in \mathbb{R}$$

On en déduit la relation simple entre la densité et la fonction de répartition de la loi logistique :

$$f(x) = \frac{F(x)}{1-F(x)}$$

Loi logistique

Les moments de la loi logistique sont donnés par :

$$\mathbb{E}(X) = 0$$

$$V(X) = \frac{\pi^2}{3}$$

La forme la plus générale de la loi logistique se déduit de la précédente par la transformation $Y = \alpha + \beta X$, où $\alpha \in \mathbb{R}$ est le paramètre de position et $\beta > 0$ le paramètre d'échelle.

Loi logistique

La densité et la fonction de répartition de Y sont :

$$f_Y(y) = \frac{\exp(-\frac{y-\alpha}{\beta})}{\beta(1 + \exp(-\frac{y-\alpha}{\beta}))^2}$$

$$F_Y(y) = \frac{1}{1 + \exp(-\frac{y-\alpha}{\beta})}$$

On en déduit également les moments de Y :

$$\mathbb{E}(Y) = \alpha$$

$$V(Y) = \frac{\beta^2 \pi^2}{3}.$$

Retrouver les résultats précédents à titre d'exercice.

Loi de Laplace

Une v.a.r. X suit une loi de Laplace si sa densité est donnée par :

$$f_X(x) = \frac{1}{2} e^{-|x|}, \quad x \in \mathbb{R}$$

Ses moments sont donnés par :

$$\mathbb{E}(X) = 0 \quad \text{(symétrie)}$$

$$\mathbb{E}(X^r) = 0 \quad (r \text{ impair})$$

$$\mathbb{E}(X^r) = r! \quad (r \text{ pair})$$

Démontrer ce dernier résultat à titre d'exercice, en utilisant les propriétés de la fonction gamma.

Sa fonction caractéristique est donnée par :

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

Loi beta de seconde espèce

Soient X et Y deux v.a.r. indépendantes suivant respectivement les lois $\gamma(p)$ et $\gamma(q)$, avec $p, q > 0$.

Alors, la variable $Z = \frac{X}{Y}$ suit une loi beta de seconde espèce de paramètres p et q. Celle-ci est notée $\beta(p, q)$.

Sa densité est donnée par :

$$f(z) = \frac{1}{B(p,q)} \frac{z^{p-1}}{(1+z)^{p+q}} \mathbb{I}_{\mathbb{R}^+}(z)$$

où

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

Le démontrer à titre d'exercice.

Pour ce faire, on pourra considérer la transformation $g : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+ \times \mathbb{R}^+$ définie par

$$g(X,Y) = (X, \frac{X}{Y}).$$

Propriétés de la fonction beta

$$\begin{aligned} B(p,q) &= \int_0^{+\infty} \frac{x^{p-1}}{(1+x)^{p+q}} dx \\ &= \int_0^1 x^{p-1}(1-x)^{q-1} dx \end{aligned}$$

$$B(p,q) = B(q,p)$$

$$B(1/2, 1/2) = \pi$$

Loi beta de première espèce

La variable aléatoire $T = \frac{Z}{1+Z}$ suit la loi beta de première espèce $\beta(p, q)$ de paramètres p et q, où $p, q > 0$.

T est une variable à valeurs dans $[0, 1]$.

La densité de T est donnée par :

$$f(t) = \frac{1}{B(p,q)} t^{p-1}(1-t)^{q-1}\mathbb{I}_{[0,1]}(t)$$

Le démontrer à titre d'exercice.

$\beta(1, 1) = \mathcal{U}_{[0,1]}$.

Pour tout $r \geq 0$, on a :

$$\mathbb{E}(T^r) = \frac{B(p+r, q)}{B(p,q)}$$

$$\mathbb{E}(Z^r) = \frac{B(p+r, q-r)}{B(p,q)} \quad (r < q)$$

Moments des lois beta

On en déduit :

$$\mathbb{E}(T) = \frac{p}{p+q}$$

$$V(T) = \frac{pq}{(p+q)^2(p+q+1)}$$

$$\mathbb{E}(Z) = \frac{p}{q-1} \quad (q > 1)$$

$$V(Z) = \frac{p(p+q-1)}{(q-1)^2(q-2)} \quad (q > 2)$$

La variable Z n'a donc des moments d'ordre 1 et 2 que pour $q > 2$.

Loi normale unidimensionnelle

La v.a.r. X suit une *loi normale* d'espérance m et de variance σ^2 si sa densité est donnée par

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

On note $x \sim \mathcal{N}(m, \sigma^2)$.

Loi normale centrée réduite

La v.a.r. $U = \frac{X-m}{\sigma}$ suit une loi normale d'espérance nulle et de variance 1, notée $\mathcal{N}(0,1)$, dite *loi normale centrée réduite*, de densité :

$$f_U(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}, \quad u \in \mathbb{R}$$

Moments de la loi normale

Si $X \sim \mathcal{N}(m, \sigma^2)$, alors :

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

et

$$V(X) = \sigma^2.$$

Loi normale et loi gamma

Théorème : La v.a.r. $\frac{1}{2}U^2$ suit la loi $\gamma(\frac{1}{2}, 1)$.

Le démontrer à titre d'exercice.

Moments centrés de la loi normale

Une conséquence du théorème précédent est que les *moments centrés* de la loi normale $\mathcal{N}(m, \sigma^2)$ sont donnés par :

$$\mathbb{E}(X-m)^r = \sigma^r 2^{r/2} \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{1}{2})}, \quad (r \text{ pair})$$

$$\mathbb{E}(X-m)^r = 0 \quad (r \text{ impair})$$

Le démontrer à titre d'exercice.

Fonction caractéristique

Si $X \sim \mathcal{N}(m, \sigma^2)$ et $U \sim \mathcal{N}(0, 1)$, alors :

$$\phi_U(t) = e^{-t^2/2}$$

$$\phi_X(t) = e^{itm} e^{-\sigma^2 t^2/2}.$$

Somme de variables normales indépendantes

Théorème : Soient X et Y deux v.a.r. indépendantes suivant respectivement les lois $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_1, \sigma_2^2)$. Alors, la v.a.r. $X_1 + X2$ suit la loi normale $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

Le démontrer à titre d'exercice en utilisant les fonctions caractéristiques.

Loi normale multidimensionnelle

Soit X un vecteur aléatoire de dimension p, noté

$$X = \begin{pmatrix} X_1 \\ \vdots \\ \vdots \\ X_p \end{pmatrix}$$

On dit que X est *normal* (ou gaussien) si pour tout vecteur $u = \begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix}$ de $\mathbb{R}^p$, $u'X$ est une v.a.r. normale unidimensionnelle.

Composantes

En particulier, en posant $u_i = 1$ et $u_j = 0$, $j \neq i$, on voit que les composantes X_i sont des variables normales unidimensionnelles.

Transformations linéaires et densité

On déduit de la définition que si A est une application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^d$, et si X est un vecteur aléatoire normal de $\mathbb{R}^p$, alors AX est un vecteur aléatoire normal dans $\mathbb{R}^d$.

Il est également possible de définir la loi normale multidimensionnelle par sa densité.

En effet, soit $m \in \mathbb{R}^p$ et Σ une matrice $p \times p$ symétrique définie positive.

X est un vecteur aléatoire normal de dimension p, de loi $\mathcal{N}(m, \Sigma)$, si sa densité est donnée par :

$$f_X(x) = \frac{1}{(2\pi)^{p/2}\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(x-m)'\Sigma^{-1}(x-m)\right).$$

Loi log-normale

La v.a.r. Y suivant une loi normale $\mathcal{N}(m, \sigma^2)$, la v.a.r. définie par $X = \exp(Y)$ suit une loi *log-normale*.

La densité de X est donnée par :

$$f_X(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left(-\frac{(\log(x) - m)^2}{2\sigma^2}\right) \cdot \mathbb{I}_{\mathbb{R}^+}(x)$$

Ses moments sont donnés par :

$$\mathbb{E}(X) = e^{m+\sigma^2/2}$$

$$V(X) = e^{2m+\sigma^2} \cdot (e^{\sigma^2} - 1)$$

A titre d'exercice, représenter graphiquement sous R la densité de la loi log-normale pour différentes valeurs des paramètres m et σ^2.

Loi de Laplace

La v.a.r. X suit la *loi de Laplace* si sa densité est donnée par :

$$f_X(x) = \frac{1}{2} e^{-|x|}, \quad x \in \mathbb{R}$$

Ses moments sont donnés par :

$$\mathbb{E}(X) = 0 \quad \text{(symétrie)}$$

$$\mathbb{E}(X^r) = 0 \quad (r \text{ impair})$$

$$\mathbb{E}(X^r) = r! \quad (r \text{ pair})$$

Le démontrer à titre d'exercice.

Sa fonction caractéristique est donnée par

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

Loi du Khi-deux

On considère n v.a.r. $X_1, \ldots, X_n$ *indépendantes* suivant toutes la loi normale $\mathcal{N}(0, 1)$.

Alors, la v.a.r. $U_n = \sum_{i=1}^n X_i^2$ suit la *loi du Khi-deux* à n degrés de liberté, notée $\chi^2(n)$.

Sa densité est donnée par :

$$f_{U_n}(x) = \frac{1}{2^{n/2}\Gamma(n/2)}\, e^{-x/2} \cdot x^{\frac{n}{2}-1} \cdot \mathbb{I}_{\mathbb{R}^+}(x)$$

Le démontrer à titre d'exercice. Notons que $U_n/2$ est de loi $\gamma(n/2)$ d'après les résultats vus précédemment.

Loi du Khi-deux

Le théorème suivant est très utilisé en Statistique.

Théorème : Soit X un vecteur aléatoire de dimension p, de loi $\mathcal{N}(m, \Sigma)$. Alors, la v.a.r.

$$Y = (X - m)'\Sigma^{-1}(X - m)$$

suit la loi $\chi^2(p)$.

Preuve : Admis (voir le manuel du cours pour la preuve, p. 32).

Loi du Khi-deux : moments et f.c.

De la relation existant entre les lois $\chi^2(n)$ et $\gamma(n/2)$, on déduit :

$$\mathbb{E}(U_n^r) = 2^r \, \frac{\Gamma(\frac{n}{2} + r)}{\Gamma(\frac{n}{2})}$$

$$\mathbb{E}(U_n) = n$$

$$V(U_n) = 2n$$

La fonction caractéristique de la loi $\chi^2(n)$ est donnée par :

$$\phi_{U_n}(t) = \frac{1}{(1 - 2it)^{n/2}}$$

Loi du Khi-deux : additivité

Théorème : Soient X et Y deux v.a.r. indépendantes suivant respectivement des lois du Khi-deux à n et m degrés de liberté. Alors, la v.a.r. $X+Y$ suit la loi $\chi^2(n+m)$.

Le démontrer à titre d'exercice, en utilisant la méthode des fonctions caractéristiques.

Loi du Khi-deux : asymptotique

Soit $U_n \sim \chi^2(n)$.

Alors, d'après le Théorème de Lévy,

$$\frac{U_n - n}{\sqrt{2n}} \rightarrow_L \mathcal{N}(0,1) \quad (n \rightarrow \infty)$$

Un autre résultat est celui de Fisher :

$$\sqrt{2U_n} - \sqrt{2n-1} \rightarrow_L \mathcal{N}(0,1) \quad (n \rightarrow \infty)$$

La meilleure approximation asymptotique est toutefois celle déduite du résultat de Wilson-Hilferty :

$$\frac{\sqrt[3]{\frac{U_n}{n}} - \left(1 - \frac{2}{9n}\right)}{\sqrt{\frac{2}{9n}}} \rightarrow_L \mathcal{N}(0,1) \quad (n \rightarrow \infty)$$

Loi du Khi-deux décentrée

On considère n v.a.r. indépendantes $X_1, \ldots, X_n$, de lois respectives $\mathcal{N}(m_i, 1)$.

La v.a.r. $U_{n,\theta} = \sum_{i=1}^n X_i^2$ suit alors une *loi du Khi-deux décentrée* à n degrés de liberté et de *paramètre d'excentricité* (ou de décentrage)

$$\theta = \sum_{i=1}^n m_i^2.$$

On note cette loi $\chi^2(n, \theta)$.

Un résultat important est que **la loi de** $U_{n,\theta}$ **ne dépend de** $m_1, \ldots, m_n$ **qu'à travers** θ.

Une démonstration de ce résultat est fournie dans le manuel du cours, p. 34.

Loi du Khi-deux décentrée

Les moments de la loi $\chi^2(n,\theta)$ sont donnés. par :

$$\mathbb{E}(U_{n,\theta}) = n + \theta$$

$$V(U_{,n,\theta}) = 2(n + 2\theta)$$

Théorème : Soit le couple de v.a.r. (X, Y) tel que la loi marginale de X est une loi de Poisson $\mathcal{P}(\theta/2)$ et la loi conditionnelle de Y sachant $X = x$ est la loi $\chi^2(2x + n)$. Alors, la loi marginale de Y est la loi $\chi^2(n,\theta)$.

Preuve : Admis.

Loi du Khi-deux décentrée

Asymptotique.

On a :

$$\frac{U_{n,\theta} - (n+\theta)}{\sqrt{2(n+2\theta)}} \to_L \mathcal{N}(0,1)$$

Cette propriété est valable pour :

a) θ fixe, $n \to \infty$

b) n fixe, $\theta \to \infty$

Loi de Student

Soient X et Y deux v.a.r. indépendantes suivant respectivement les lois $\mathcal{N}(0,1)$ et $\chi^2(n)$.

On appelle *loi de Student* à n degrés de liberté la loi suivie par le rapport :

$$T_n = \frac{X}{\sqrt{\frac{Y}{n}}}$$

Cette loi est notée St_n ou T_n.

Loi de Student

La densité de la loi St_n est donnée par :

$$f_{T_n}(x) = \frac{1}{\sqrt{n}B(1/2, n/2)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

Pour le démontrer, il suffit de remarquer que $T_n^2/n = X^2/Y$ est aussi le rapport de deux variables $X^2/2$ et $Y/2$ indépendantes et de lois respectives $\gamma(1/2)$ et $\gamma(n/2)$. Par conséquent, T_n^2/n suit une loi $\beta(1/2, n/2)$ de seconde espèce. On en déduit la densité de T_n (le faire à titre d'exercice).

Loi de Student

Les moments de la loi St_n sont donnés par :

$$\mathbb{E}(T_n) = 0 \quad (\text{symétrie}, n > 1)$$

$$V(T_n) = \frac{n}{n-2} \quad (n > 2)$$

Les moments s'obtiennent à partir des moments de la loi bêta de seconde espèce, d'où les conditions d'existence sur les degrés de liberté n.

Loi de Student

Asymptotique. T_n converge en loi vers $\mathcal{N}(0,1)$ lorsque n tend vers l'infini :

$$T_n \to_L \mathcal{N}(0,1).$$

Ceci peut être démontré en remarquant que la v.a.r. Y/n converge en probabilité vers la variable certaine 1. En effet,

$$\mathbb{E}(Y/n) = 1 \quad (\forall n)$$

$$V(Y/n) = 2/n \to 0 \quad (n \to \infty).$$

Il suffit alors d'appliquer les résultats sur la convergence en loi vus précédemment.

Cas particulier : Loi de Cauchy

Pour $n = 1$, $T_1 = \frac{X}{\sqrt{Y}}$ où X et Y suivent indépendamment, respectivement les lois $\mathcal{N}(0,1)$ et $\chi^2(1)$.

La loi suivie par T_1 porte le nom de *loi de Cauchy*.

C'est aussi la loi du rapport de deux v.a.r. normales centrées réduites indépendantes.

Sa densité est :

$$f_{T_1}(x) = \frac{1}{\pi}\frac{1}{1+x^2}$$

Sa fonction caractéristique est donnée par :

$$\phi_{T_1}(x) = e^{-|t|}$$

Cas particulier : Loi de Cauchy

T_1^2 suit la loi $\beta(1/2, 1/2)$ de deuxième espèce et les conditions d'existence des moments de cette dernière font que T_1 ne possède aucun moment et donc, a fortiori, **ni espérance, ni variance**.

Loi de Student non centrée

Soient X et Y deux variables aléatoires indépendantes de lois respectives $\mathcal{N}(\delta, 1)$ et $\chi^2(n)$.

On appelle *loi de Student supérieurement non centrée*, de paramètre de décentrage δ, la loi du rapport :

$$\frac{X}{\sqrt{\frac{Y}{n}}}$$

On la note $St_n(\delta)$ ou $T_n(\delta)$.

Loi de Student non centrée

Ses moments sont donnés par :

$$\mathbb{E}(T_n(\delta)) = \sqrt{\frac{n}{2}}\,\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})}\,\delta$$

$$V(T_n(\delta)) = \frac{n}{n-2}(1+\delta^2) - \frac{\delta^2 n}{2}\,\frac{\Gamma^2(\frac{n-1}{2})}{\Gamma^2(\frac{n}{2})}$$

Loi de Student doublement non centrée

Soient X et Y deux v.a.r. indépendantes suivant respectivement les lois $\mathcal{N}(\delta, 1)$ et $\chi^2(n, \theta)$.

On appelle *loi de Student doublement non centrée*, de paramètre de décentrage (δ, θ), la loi du rapport :

$$\frac{X}{\sqrt{\frac{Y}{n}}}$$

On la note $St_n(\delta, \theta)$ ou $T_n(\delta, \theta)$.

Loi de Student doublement non centrée

Les formules donnant les moments de la loi de Student doublement non centrée sont très complexes.

Cependant, on peut établir, asymptotiquement, au premier ordre en $\frac{1}{n}$:

$$\mathbb{E}(T_n(\delta,\theta)) = \delta + \frac{\delta}{n}\left(\frac{3}{4} - \frac{\theta}{2}\right)$$

$$\mathbb{E}(T_n^2(\delta,\theta)) = (\delta^2+1)\left(1 + \frac{2-\theta}{n}\right)$$

Notons que si $\delta = 0$, la loi de $T_n(0,\theta)$ est appelée *loi de Student inférieurement décentrée*.

Loi de Fisher-Snedecor

Soient X et Y deux v.a.r. indépendantes de lois respectives $\chi^2(n)$ et $\chi^2(m)$.

Alors, la variable

$$U = \frac{X/n}{Y/m}$$

suit une *loi de Fisher-Snedecor* à n et m degrés de liberté, notée $F_{n,m}$.

Sa densité est donnée par :

$$f_F(x) = \frac{1}{B(n/2, m/2)}\, n^{n/2}\, m^{m/2} \frac{x^{n/2-1}}{(m+nx)^{(m+n)/2}}\, \mathbb{I}_{\mathbb{R}^+}(x)$$

Celle-ci peut être obtenue en remarquant que la variable $\frac{n}{m}U$ suit une loi $\beta(n/2, m/2)$ de deuxième espèce.

Loi de Fisher-Snedecor

Remarques.

a) T_n^2, carré d'une variable de Student à n degrés de liberté, suit une loi de Fisher-Snedecor à 1 et n degrés de liberté.

b) La v.a.r. $Y = \log F_{2,2}$ suit une loi logistique.

Moments.

$$\mathbb{E}(F_{n,m}) = \frac{m}{m-2} \quad (m > 2)$$

$$V(F_{n,m}) = \frac{2m^2(n+m-2)}{n(m-4)(m-2)^2} \quad (m > 4)$$

Asymptotique.

Plutôt que $F_{n,m}$, c'est souvent le comportement asymptotique de $Z_{n,m} = \frac{1}{2} \log F_{n,m}$ que l'on étudie :

$$\frac{Z_{n,m} - \frac{1}{2}(\frac{1}{m} - \frac{1}{n})}{\sqrt{\frac{1}{2}(\frac{1}{n} + \frac{1}{m})}} \to_L \mathcal{N}(0,1)$$

pour $n, m \to \infty$.

Remarques.

a) La transformation $F \mapsto \frac{1}{2} \log F$ s'appelle transformation Z de Fisher.

b) Notons $f_\alpha(n,m)$ le fractile d'ordre α de la loi $F_{n,m}$. On a alors :

$$f_{1-\alpha}(m,n) = \frac{1}{f_\alpha(n,m)}$$

Le démontrer à titre d'exercice.

Loi de Fisher non centrée

Soient X et Y deux v.a.r. indépendantes suivant respectivement des lois du Khi-deux décentrée $\chi^2(n, \theta)$ et $\chi^2(m, \lambda)$.

On appelle *loi de Fisher doublement décentrée* de paramètre de décentrage (θ, λ) la loi du rapport :

$$\frac{X/n}{Y/m}$$

On la note $F_{n,m}(\theta, \lambda)$.

Loi de Fisher non centrée

On se contentera d'examiner le cas où $\lambda = 0, \theta \neq 0$, dit *loi de Fisher simplement décentrée* $F_{n,m}(\theta, 0)$, notée $F_{n,m}(\theta)$.

On remarque que, par construction :

$$F_{n,m}(0, \lambda) = F_{m,n}(\lambda, 0).$$

On peut donc passer de loi de Fisher simplement décentrée au numérateur à la loi décentrée au dénominateur par cette relation.

Loi de Fisher non centrée

Moments.

$$\mathbb{E}(F_{n,m}(\theta)) = \frac{m(n+\theta)}{n(m-2)} \quad (m > 2)$$

$$V(F_{n,m}(\theta)) = 2\,\frac{m^2}{n^2}\,\frac{(n+\theta)^2 + (n+2\theta)(m-2)}{(m-2)^2(m-4)} \quad (m > 4)$$

La famille exponentielle

Il s'agit d'une famille de lois de probabilité (parfois appelée famille de Darmois ou de Koopman) englobant la plupart des lois précédentes et qui est très importante en Statistique. En effet, des résultats généraux existent, pour cette famille, en théorie de l'estimation et des tests.

Le nom de famille exponentielle *n'a aucun rapport* avec la loi exponentielle.

La famille exponentielle

Définition. Une loi de probabilité P_θ, $\theta \in \Theta$, de densité $f(x, \theta)$ est dite appartenir à la *famille exponentielle* s'il existe des fonctions

$$\alpha_j(\theta), T_j(x), c(\theta), h(x)$$

avec $h(x) > 0$, $\forall x$, telles que

$$f(x, \theta) = c(\theta) h(x) \exp\left[\sum_{j=1}^{r} \alpha_j(\theta) T_j(x)\right]$$

autrement dit, telles que

$$\log f(x, \theta) = \beta(\theta) + b(x) + \sum_{j=1}^{r} \alpha_j(\theta) T_j(x).$$

La famille exponentielle : exemples

On pourra vérifier, à titre d'exercice, que les lois suivantes appartiennent à la famille exponentielle :

– Loi binomiale

– Loi de Poisson

– Loi normale

– Loi gamma

– Loi beta

La loi uniforme, la loi de Student et la loi de Cauchy *ne se mettent pas* sous forme exponentielle.

Symétrie et aplatissement

X étant une v.a.r., nous notons respectivement par m_k et μ_k les moments non centrés et centrés de X :

$$m_k = \mathbb{E}(X^k)$$

$$\mu_k = \mathbb{E}[(X - m_1)^k].$$

Les *coefficients de Pearson* de symétrie et d'aplatissement sont respectivement définis par :

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

Symétrie et aplatissement

Lorsqu'une distribution est symétrique, les moments centrés d'ordre impair sont nul ; β_1 est donc nul.

Le coefficient β_1, dit *coefficient de symétrie* (ang. *skewness*), est un indicateur de symétrie de la loi de X.

Le coefficient β_2 quantifie le degré d'aplatissement (ang. *kurtosis*) de la loi de X, en un sens précisé dans la suite.

Symétrie et aplatissement

En pratique, les coefficients de Pearson sont moins utilisés que les *coefficients de Fisher*, γ_1 et γ_2, définis par :

$$\gamma_1 = \sqrt{\beta_1}$$

$$\gamma_2 = \beta_2 - 3$$

L'interprétation de γ_1 est identique à celle de β_1 (skewness).

La définition du coefficient de kurtosis γ_2 fait référence à la loi normale.

A titre d'exercice, en utilisant les résultats vus précédemment sur les moments de la loi normale, montrer que $\beta_2 = 3$.

Pour une loi normale $\mathcal{N}(0, 1)$, on a donc :

$$\gamma_1 = \gamma_2 = 0.$$