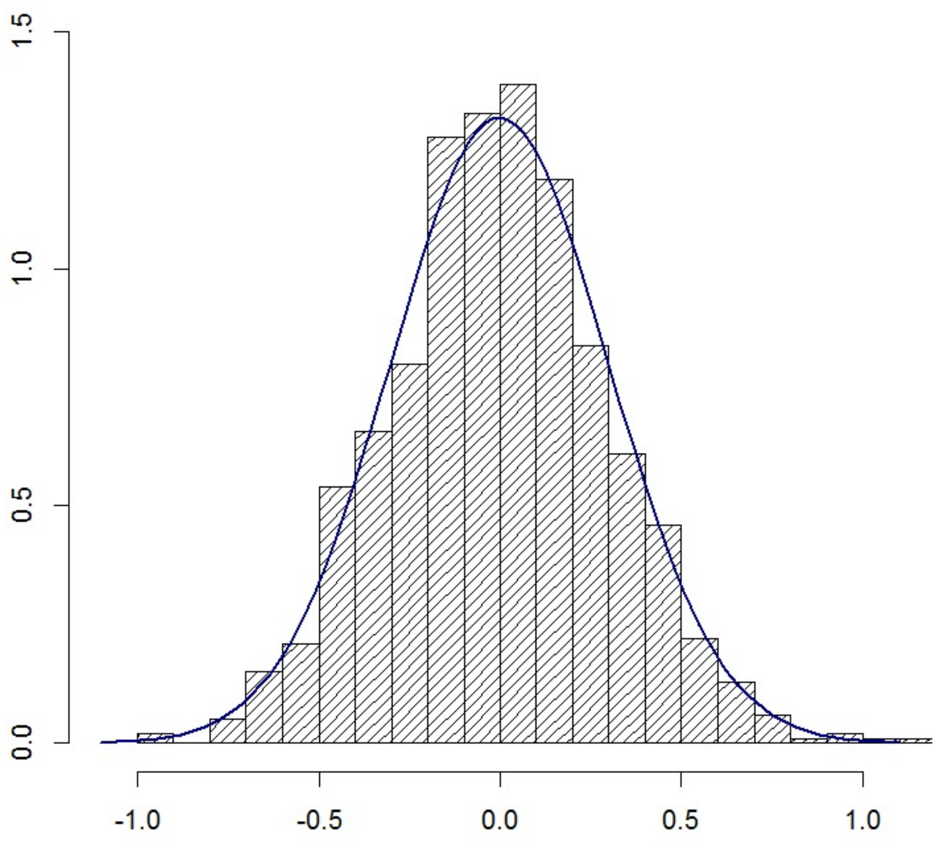
Université de Caen

Socle théorique de la Statistique

Christophe Chesneau

http://www.math.unicaen.fr/~chesneau/



Caen, le 07 Février 2017

Table des matières

∼ Note ∼

L'objectif de ce document est de présenter quelques bases mathématiques sur lesquelles reposent la Statistique paramétrique. Notamment, les principales lois de probabilités utilisées en Statistique seront étudiées (lois normale, de Student, du Chi-deux et de Fisher. . .).

Contact : `christophe.chesneau@gmail.com`

Bonne lecture !

1 Variables aléatoires réelles à densité : l'essentiel

Point de départ : On suppose construit un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Lorsqu'une quantité est introduite (dérivée, intégrale, espérance, moments, variance ...), il est supposé que celle-ci existe.

Variable aléatoire réelle (*var*) : On appelle *var* toute application $X : \Omega \to \mathbb{R}$.

Support : Le support d'une *var* X, noté $X(\Omega)$, est l'ensemble des valeurs atteintes par X. C'est la première chose à préciser lorsqu'on considère une *var*.

Densité : On appelle densité toute fonction $f : \mathbb{R} \to \mathbb{R}$ vérifiant :

- pour tout $x \in \mathbb{R}$, $f(x) \geq 0$,
- $\int_{-\infty}^{\infty} f(x)dx = 1$.

Var à densité : On dit qu'une *var* X est à densité s'il existe une densité f telle que, pour tous $a \leq b$, on a

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx.$$

On dit alors que X est de densité f ou X possède la densité f ou f est une densité de X. La loi de X est caractérisée par f.

Calculs usuels :

$\mathbb{P}(X = a)$	$\mathbb{P}(X \leq b)$	$\mathbb{P}(X \geq a)$	$\mathbb{P}(a \leq X \leq b)$
0	$\int_{-\infty}^{b} f(x)dx$	$\int_{a}^{\infty} f(x)dx$	$\int_a^b f(x)dx = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx$

On peut donc remplacer les inégalités larges par des inégalités strictes :

$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a < X < b) \ldots$

Fonction de répartition :

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt, \qquad x \in \mathbb{R}.$$

Retour sur le support : Soit F la fonction de répartition d'une *var* X de densité f. Alors $X(\Omega)$ est l'adhérence de $\{x \in \mathbb{R};\ F(x) \in]0,1[\}$.

Caractérisation d'une fonction de répartition :

- F est continue sur $\mathbb{R}$,
- F est croissante,
- $\lim_{x \to -\infty} F(x) = 0$ et $\lim_{x \to \infty} F(x) = 1$.

Densité et fonction de répartition : Soit X une *var* à densité de fonction de répartition F. Alors une densité f de X est donnée par

$$f(x) = F'(x)$$

pour tout $x \in X(\Omega)$ sauf les points où F n'est pas dérivable, et par ce qu'on veut ailleurs (du moment que f reste une densité).

Égalité en loi : X et Y suivent la même loi $\Leftrightarrow$ $F_X(x) = F_Y(x)$ pour tout $x \in \mathbb{R} \Leftrightarrow$ $f_X(x) = f_Y(x)$ pour tout $x \in \mathbb{R}$ sauf, éventuellement, en un nombre fini de points.

Var symétrique : X est symétrique $\Leftrightarrow$ X et $-X$ suivent la même loi.

X possède une densité f paire $\Rightarrow$ X est symétrique.

Propriétés d'une *var* symétrique :

- pour tout $x \in \mathbb{R}$, $\mathbb{P}(X \leq -x) = 1 - \mathbb{P}(X \leq x)$,
- pour tout $x \geq 0$, $\mathbb{P}(|X| \leq x) = 2\mathbb{P}(X \leq x) - 1$.

Espérance :

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

L'espérance de X est la valeur moyenne de X.

Formule du transfert :

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

Propriétés élémentaires :

- $\mathbb{E}(a) = a$,
- $X(\Omega) \subseteq [0, \infty[$ $\Rightarrow$ $\mathbb{E}(X) \geq 0$,
- $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$,
- $\mathbb{E}(aX^2 + bX + c) = a\mathbb{E}(X^2) + b\mathbb{E}(X) + c$.

Moment d'ordre r :

$$\mathbb{E}(X^r) = \int_{-\infty}^{\infty} x^r f(x)dx.$$

Espérance et *var* symétrique : X est symétrique $\Rightarrow$ pour toute fonction $g : \mathbb{R} \to \mathbb{R}$ impaire (si existence) $\mathbb{E}(g(X)) = 0$.

Variance :

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

La variance de X mesure la dispersion des valeurs de X autour de son espérance.

Propriétés élémentaires :

- $\mathbb{V}(X) \geq 0$,
- $\mathbb{V}(X) = 0 \Leftrightarrow X$ est une *var* constante,
- $\mathbb{V}(aX + b) = a^2\mathbb{V}(X)$.

Formule de König-Huyghens :

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

Écart-type :

$$\sigma(X) = \sqrt{\mathbb{V}(X)}.$$

Tout comme la variance, l'écart-type de X mesure la dispersion des valeurs de X autour de son espérance. Il a l'avantage d'être de la même unité que X contrairement à la variance qui élève l'unité au carré.

Transformée de Laplace :

$$L(t) = \mathbb{E}(e^{tX}),$$

pour tout t telle que $\mathbb{E}(e^{tX})$ existe.

Moments et transformée de Laplace :

$$\mathbb{E}(X^n) = L^{(n)}(0).$$

C'est pourquoi on appelle aussi L fonction génératrice des moments de X.

Égalité en loi et transformée de Laplace : X et Y suivent la même loi $\Leftrightarrow$ $L_X(t) = L_Y(t)$ pour tout $t \in \mathbb{R}$.

Vecteur de *var* : Soient $X_1, \ldots, X_n$ n *var*. On appelle vecteur de *var* l'application $(X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ vérifiant $(X_1, \ldots, X_n)(\omega) = (X_1(\omega), \ldots, X_n(\omega))$, $\omega \in \Omega$.

Support : Le support d'un vecteur de *var* $(X_1, \ldots, X_n)$, noté $(X_1, \ldots, X_n)(\Omega)$, est l'ensemble de ses valeurs possibles.

Densité (n-dimensionnelle) : On appelle densité (n-dimensionnelle) toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ vérifiant :

- pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n$, $f(x_1, \ldots, x_n) \geq 0$,
- $\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1 \ldots dx_n = 1$.

Vecteur de *var* à densité : On dit qu'un vecteur de *var* $(X_1, \ldots, X_n)$ est à densité s'il existe une densité f, telle que, pour tout $\mathcal{D} \subseteq \mathbb{R}^n$, on a

$$\mathbb{P}\left((X_1, \ldots, X_n) \in \mathcal{D}\right) = \int \ldots \int_{\mathcal{D}} f(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

On dit alors que $(X_1, \ldots, X_n)$ est de densité f ou $(X_1, \ldots, X_n$ possède la densité f ou f est une densité de $(X_1, \ldots, X_n)$. La loi de $(X_1, \ldots, X_n)$ est caractérisée par f.

Fonction de répartition :

$$F(x_1, \ldots, x_n) = \mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i \leq x_i\}\right), \qquad (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Retour sur le support : Soit F la fonction de répartition d'un vecteur de *var* $(X_1, \ldots, X_n)$ à densité. On définit $(X_1, \ldots, X_n)(\Omega)$ par l'adhérence de $\{(x_1, \ldots, x_n) \in \mathbb{R}^n;\ F(x_1, \ldots, x_n) \in]0, 1[\}$.

Densités marginales : Une densité de X_1 est donnée par

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_2 \ldots dx_n, \qquad x_1 \in \mathbb{R}.$$

On définit de la même manière des densités pour d'autres *var* $X_2, \ldots, X_n$.

Densités de vecteurs de *var* composantes de $(X_1, \ldots, X_n)$: Une densité de (X_1, X_2) est donnée par

$$f_{(X_1,X_2)}(x_1, x_2) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_3 \ldots dx_n, \qquad (x_1, x_2) \in \mathbb{R}^2.$$

On définit de la même manière des densités d'autres vecteurs de *var*.

Indépendance mutuelle de n *var* à densité : $X_1, \ldots, X_n$ sont (mutuellement) indépendantes $\Leftrightarrow$

$(X_1, \ldots, X_n)$ possède la densité produit : $\prod_{i=1}^{n} f_{X_i}(x_i)$, *i.e.*

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n$ sauf, éventuellement, sur un ensemble de volume nul

(donc $(X_1, \ldots, X_n)(\Omega) = X_1(\Omega) \times \ldots \times X_n(\Omega)$) $\Leftrightarrow$ pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n$, on a

$$F(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$$

$\Leftrightarrow$ pour tous $a_i \leq b_i$, $i \in \{1, \ldots, n\}$,

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \{a_i \leq X_i \leq b_i\}\right) = \prod_{i=1}^{n} \mathbb{P}(a_i \leq X_i \leq b_i).$$

Sur l'indépendance :

- Si $(X_1, \ldots, X_n)(\Omega)$ n'est pas un ensemble produit, alors $X_1, \ldots, X_n$ ne sont pas (mutuellement) indépendantes.
- $X_1, \ldots, X_n$ sont indépendantes $\Rightarrow$ $g_1(X_1), \ldots, g_n(X_n)$ sont indépendantes.
- $X_1, \ldots, X_n$ sont indépendantes $\Rightarrow$ $g(X_1, \ldots, X_q)$ et $h(X_{q+1}, \ldots, X_n)$ sont indépendantes.

Formule du transfert :

$$\mathbb{E}\left(g(X_1, \ldots, X_n)\right) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) f(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

Espérance et indépendance : $X_1, \ldots, X_n$ sont indépendantes $\Rightarrow$

$$\mathbb{E}\left(\prod_{i=1}^{n} g_i(X_i)\right) = \prod_{i=1}^{n} \mathbb{E}(g_i(X_i)).$$

Covariance :

$$\mathbb{C}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Propriétés de la covariance :

- $\mathbb{C}(X, Y) = \mathbb{C}(Y, X)$,
- $\mathbb{C}(X, a) = 0$,
- $\mathbb{C}(aX + b, cY + d) = ac\mathbb{C}(X, Y)$,
- $|\mathbb{C}(X, Y)| \leq \sigma(X)\sigma(Y)$,
- $\mathbb{C}(aX + bY, cU + dV) = ac\mathbb{C}(X, U) + ad\mathbb{C}(X, V) + bc\mathbb{C}(Y, U) + bd\mathbb{C}(Y, V)$.

Matrice de covariance :

$$\Sigma = \begin{pmatrix} \mathbb{C}(X_1, X_1) & \mathbb{C}(X_1, X_2) & \dots & \dots & \mathbb{C}(X_1, X_n) \\ \mathbb{C}(X_2, X_1) & \mathbb{C}(X_2, X_2) & \dots & \dots & \mathbb{C}(X_2, X_n) \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \mathbb{C}(X_{n-1}, X_n) \\ \mathbb{C}(X_n, X_1) & \dots & \dots & \dots & \mathbb{C}(X_n, X_n) \end{pmatrix}.$$

Espérance et variance d'une somme de n *var* :

$$\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i), \quad \mathbb{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{V}(X_i) + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} \mathbb{C}(X_i, X_j).$$

$X_1, \dots, X_n$ sont indépendantes $\Rightarrow \mathbb{C}(X_i, X_j) = 0$ pour tout $i \neq j$ avec $(i, j) \in \{1, \dots, n\}^2 \Rightarrow$

$$\mathbb{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{V}(X_i).$$

Bilinéarité de la covariance :

$$\mathbb{C}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{q} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{q} a_i b_j \mathbb{C}(X_i, Y_j).$$

Théorème du changement de variable (pour un couple de *var*, *i.e.* $n = 2$) : Soient (X, Y) un couple de *var* de densité $f_{(X,Y)}$, et $\phi : (X, Y)(\Omega) \to \mathbb{R}$ et $\psi : (X, Y)(\Omega) \to \mathbb{R}$ deux fonctions.

On considère le couple de *var* :

$$(U,V)=(\phi(X,Y),\psi(X,Y)).$$

On suppose que $g(x,y)=(\phi(x,y),\psi(x,y))$ est injective et qu'il existe deux fonctions $h:\mathbb{R}^2\to X(\Omega)$ et $k:\mathbb{R}^2\to Y(\Omega)$ différentiables telles que

$$\begin{cases} u=\phi(x,y), \\ v=\psi(x,y), \end{cases} \quad \Leftrightarrow \quad \begin{cases} x=h(u,v), \\ y=k(u,v). \end{cases}$$

Soit $J(u,v)$ le jacobien associé à $(x,y)=(h(u,v),k(u,v))$ défini par le déterminant :

$$J(u,v)=\begin{vmatrix} \frac{\partial h}{\partial u}(u,v) & \frac{\partial h}{\partial v}(u,v) \\ \\ \frac{\partial k}{\partial u}(u,v) & \frac{\partial k}{\partial v}(u,v) \end{vmatrix}=\frac{\partial h}{\partial u}(u,v)\frac{\partial k}{\partial v}(u,v)-\frac{\partial k}{\partial u}(u,v)\frac{\partial h}{\partial v}(u,v).$$

Alors une densité de (U,V) est donnée par

$$f_{(U,V)}(u,v)=f_{(X,Y)}(h(u,v),k(u,v))|J(u,v)|, \qquad (u,v)\in\mathbb{R}^2.$$

2 Loi normale

2.1 Généralités

Densité

On dit qu'une *var* X suit la loi normale $\mathcal{N}(\mu, \sigma^2)$, avec $\mu \in \mathbb{R}$ et $\sigma > 0$, si elle possède la densité :

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

(On adopte désormais la notation : $X \sim \ldots \;\Leftrightarrow\; X$ suit la loi $\ldots$).

Le résultat suivant est alors central :

La fonction f définie précédemment est bien une densité.

Preuve :

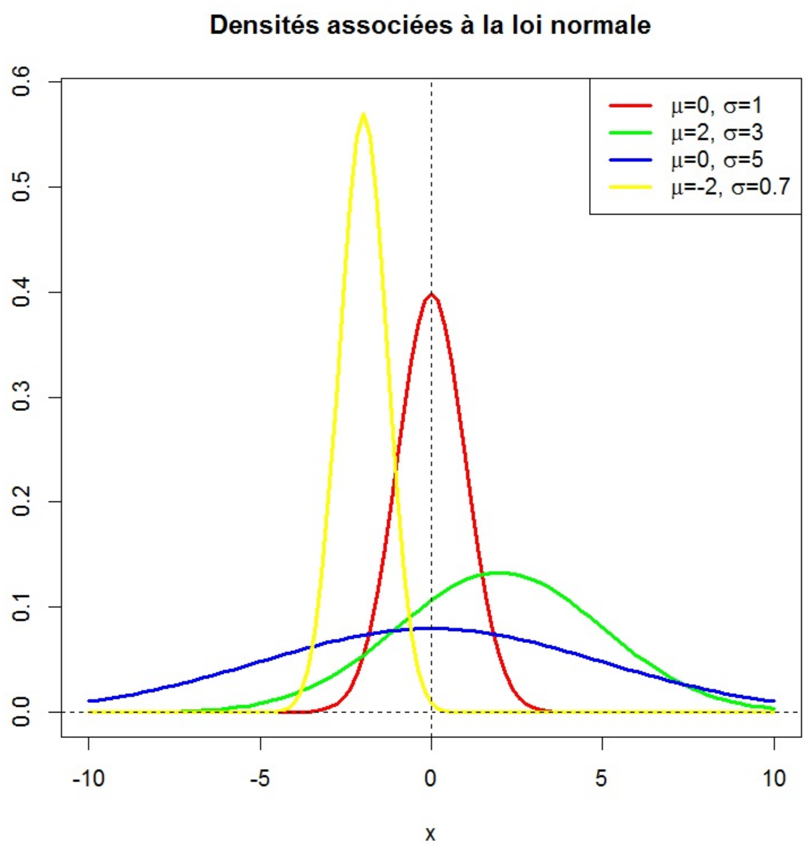
- Pour tout $x \in \mathbb{R}$, on a $f(x) \geq 0$.
- En faisant le changement de variable : $x = \sigma y + \mu$, en écrivant l'intégrale simple obtenue comme une intégrale double, puis en faisant un changement de variable en coordonnées polaires : $(y, z) = (r\cos(\theta), r\sin(\theta))$, $r \geq 0$, $\theta \in [0, 2\pi]$ (de jacobien r), il vient

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)dx &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(\sigma y+\mu-\mu)^2}{2\sigma^2}} \sigma dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \\
&= \frac{1}{\sqrt{2\pi}} \sqrt{\left(\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy\right)^2} = \frac{1}{\sqrt{2\pi}} \sqrt{\left(\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy\right)\left(\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz\right)} \\
&= \frac{1}{\sqrt{2\pi}} \sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{y^2+z^2}{2}} dydz} = \frac{1}{\sqrt{2\pi}} \sqrt{\int_{0}^{2\pi}\int_{0}^{\infty} e^{-\frac{r^2\cos^2(\theta)+r^2\sin^2(\theta)}{2}} rdrd\theta} \\
&= \frac{1}{\sqrt{2\pi}} \sqrt{\left(\int_{0}^{2\pi} d\theta\right)\left(\int_{0}^{\infty} re^{-\frac{r^2}{2}} dr\right)} = \frac{1}{\sqrt{2\pi}} \sqrt{[\theta]_0^{2\pi} \left[-e^{-\frac{r^2}{2}}\right]_0^{\infty}} \\
&= \frac{1}{\sqrt{2\pi}} \sqrt{2\pi \times 1} = 1.
\end{aligned}$$

□

Représentation graphique

Des densités associées à la loi normale $\mathcal{N}(\mu, \sigma^2)$, avec plusieurs valeurs de μ et σ, sont représentées ci-dessous :



De manière générale, le graphe de la densité f est symétrique par rapport à μ. Celui-ci est en forme de "cloche" plus ou moins arrondie selon les valeurs de σ.

En outre, ce graphe a

- un point maximum de coordonnées : $(x_*, y_*) = (\mu, f(\mu)) = \left(\mu, \frac{1}{\sqrt{2\pi\sigma^2}}\right)$,
- deux points d'inflexion de coordonnées

$$(x_0, y_0) = \left(\mu - \sigma, \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}}\right), \qquad (x_1, y_1) = \left(\mu + \sigma, \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}}\right).$$

Paramètres

Soit $X \sim \mathcal{N}(\mu, \sigma^2)$. Alors on a

$$\mathbb{E}(X) = \mu, \qquad \mathbb{V}(X) = \sigma^2, \qquad \sigma(X) = \sigma.$$

Preuve :

- On a $\mathbb{E}(X) = \mathbb{E}(X - \mu) + \mu$, avec

$$\begin{aligned}
\mathbb{E}(X-\mu) &= \int_{-\infty}^{\infty} (x-\mu) f(x) dx = \int_{-\infty}^{\infty} (x-\mu) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (x-\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \left[-\sigma^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right]_{-\infty}^{\infty} = 0.
\end{aligned}$$

  D'où

$$\mathbb{E}(X) = 0 + \mu = \mu.$$

- En utilisant l'égalité : $\mathbb{E}(X) = \mu$, en faisant une intégration par parties et en utilisant l'égalité : $\int_{-\infty}^{\infty} f(x)dx = 1$, il vient

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}((X-\mu)^2) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx \\
&= \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \left(\left[(x-\mu) \times \left(-\sigma^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \left(-\sigma^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) dx \right) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \left(0 + \sigma^2 \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right) = \sigma^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \sigma^2 \int_{-\infty}^{\infty} f(x) dx = \sigma^2 \times 1 = \sigma^2.
\end{aligned}$$

- On a

$$\sigma(X) = \sqrt{\mathbb{V}(X)} = \sqrt{\sigma^2} = \sigma.$$

$\square$

Transformée de Laplace

Soit $X \sim \mathcal{N}(\mu, \sigma^2)$. Alors la transformée de Laplace de X est

$$L(t) = e^{t\mu + \frac{t^2\sigma^2}{2}}, \qquad t \in \mathbb{R}.$$

Preuve : En faisant le changement de variable : $x = \sigma y + \mu$, et en utilisant l'égalité : $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-\alpha)^2}{2}} dy = 1$, $\alpha \in \mathbb{R}$ (car c'est l'intégrale d'une densité associée à la loi normale $\mathcal{N}(\alpha, 1)$), pour tout $t \in \mathbb{R}$, il vient

$$\begin{aligned}
L(t) &= \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{t(\sigma y+\mu)} e^{-\frac{(\sigma y+\mu-\mu)^2}{2\sigma^2}} \sigma dy = \frac{1}{\sqrt{2\pi}} e^{t\mu} \int_{-\infty}^{\infty} e^{t\sigma y} e^{-\frac{y^2}{2}} dy \\
&= \frac{1}{\sqrt{2\pi}} e^{t\mu} e^{\frac{t^2\sigma^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{(y-t\sigma)^2}{2}} dy = e^{t\mu + \frac{t^2\sigma^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-t\sigma)^2}{2}} dy = e^{t\mu + \frac{t^2\sigma^2}{2}} \times 1 \\
&= e^{t\mu + \frac{t^2\sigma^2}{2}}.
\end{aligned}$$

□

Remarque : On a $L'(t) = (\mu + t\sigma^2) e^{t\mu + \frac{t^2\sigma^2}{2}}$ et $L''(t) = (\sigma^2 + (\mu + t\sigma^2)^2) e^{t\mu + \frac{t^2\sigma^2}{2}}$.

On retrouve alors l'espérance et la variance de X :

$$\mathbb{E}(X) = L'(0) = \mu, \qquad \mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = L''(0) - (L'(0))^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2.$$

Combinaison linéaire

Soient $n \in \mathbb{N}^*$, $a_0, a_1, \ldots, a_n$ $n+1$ réels et $X_1, \ldots, X_n$ n *var* indépendantes avec $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ pour tout $i \in \{1, \ldots, n\}$. On pose

$$Y_n = a_0 + \sum_{i=1}^{n} a_i X_i.$$

Alors on a

$$Y_n \sim \mathcal{N}\left(a_0 + \sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right).$$

Preuve : Étudions la transformée de Laplace de Y_n. Par l'indépendance de $X_1, \ldots, X_n$ et l'égalité : $L_{X_i}(t) = \mathbb{E}(e^{tX_i}) = e^{t\mu_i + \frac{t^2\sigma_i^2}{2}}$, $i \in \{1, \ldots, n\}$, $t \in \mathbb{R}$, pour tout $t \in \mathbb{R}$, il vient

$$\begin{aligned} L_{Y_n}(t) &= \mathbb{E}(e^{tY_n}) = \mathbb{E}\left(e^{t\left(a_0 + \sum\limits_{i=1}^n a_i X_i\right)}\right) = e^{ta_0}\mathbb{E}\left(\prod_{i=1}^n e^{ta_iX_i}\right) = e^{ta_0}\prod_{i=1}^n \mathbb{E}\left(e^{ta_iX_i}\right) \\ &= e^{ta_0}\prod_{i=1}^n L_{X_i}(ta_i) = e^{ta_0}\prod_{i=1}^n e^{ta_i\mu_i + \frac{t^2a_i^2\sigma_i^2}{2}} = e^{t\left(a_0 + \sum\limits_{i=1}^n a_i\mu_i\right) + \frac{t^2\sum\limits_{i=1}^n a_i^2\sigma_i^2}{2}}. \end{aligned}$$

On reconnaît la transformée de Laplace associée à la loi $\mathcal{N}\left(a_0 + \sum\limits_{i=1}^n a_i\mu_i, \sum\limits_{i=1}^n a_i^2\sigma_i^2\right)$. Comme la transformée de Laplace d'une *var* caractérise sa loi, on a

$$Y_n \sim \mathcal{N}\left(a_0 + \sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n a_i^2\sigma_i^2\right).$$

□

Applications :

- Soit $X \sim \mathcal{N}(\mu, \sigma^2)$. Pour tout $a \in \mathbb{R}$, on a $X + a \sim \mathcal{N}(\mu + a, \sigma^2)$. C'est une application du résultat précédent avec $n = 1$, $a_0 = a$, $a_1 = 0$ et $X_1 = X$.
- Soit $X \sim \mathcal{N}(\mu, \sigma^2)$. Pour tout $b \in \mathbb{R}^*$, on a $bX \sim \mathcal{N}(b\mu, b^2\sigma^2)$. C'est une application du résultat précédent avec $n = 1$, $a_0 = 0$, $a_1 = b$ et $X_1 = X$.
- Soit $X \sim \mathcal{N}(\mu, \sigma^2)$. Alors on a
$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$
C'est une application du résultat précédent avec $n = 1$, $a_0 = -\frac{\mu}{\sigma}$, $a_1 = \frac{1}{\sigma}$ et $X_1 = X$. Ainsi, la transformation opérée par Z a centrée et réduite la *var* X.
- Soient $X_1, \ldots, X_n$ n *var iid* suivant chacune la loi normale $\mathcal{N}(\mu, \sigma^2)$. Alors on a
$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$
C'est une application du résultat précédent avec $a_0 = 0$, et $a_i = \frac{1}{n}$ pour tout $i \in \{1, \ldots, n\}$.
On a également
$$\sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right) \sim \mathcal{N}(0, 1).$$

2.2 Sur loi normale centrée réduite $\mathcal{N}(0,1)$

Densité

On dit qu'une *var* X suit la loi normale centrée réduite $\mathcal{N}(0,1)$ si elle possède la densité :

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

On a $\mathbb{E}(X) = 0$, d'où le "centrée", et $\mathbb{V}(X) = 1$, d'où le "réduite".

Caractérisation

Si $X \sim \mathcal{N}(\mu, \sigma^2)$, alors on a $Z = \dfrac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$. On a alors centré X et réduit X.

Quelques propriétés

Comme f est paire, X est symétrique. Cela implique :

- pour tout $x \in \mathbb{R}$, on a $\mathbb{P}(-X \leq x) = \mathbb{P}(X \leq x)$,
- pour tout $x \in \mathbb{R}$, on a $\mathbb{P}(X \leq -x) = 1 - \mathbb{P}(X \leq x)$,
- pour tout $x \geq 0$, on a $\mathbb{P}(|X| \leq x) = 2\mathbb{P}(X \leq x) - 1$,
- pour toute fonction impaire g, si existence, on a $\mathbb{E}(g(X)) = 0$.

Règle des 4σ

Si $X \sim \mathcal{N}(\mu, \sigma^2)$, alors on a $\mathbb{P}(\mu - 4\sigma \leq X \leq \mu + 4\sigma) \simeq 1$. Ainsi, la plupart des réalisations d'une *var* $X \sim \mathcal{N}(\mu, \sigma^2)$ sont contenues dans l'intervalle $[\mu - 4\sigma, \mu + 4\sigma]$.

Moments

Soit $X \sim \mathcal{N}(0,1)$. Pour tout $k \in \mathbb{N}$, on a $\mathbb{E}(X^{2k}) = \prod_{i=1}^{k}(2i-1)$ et $\mathbb{E}(X^{2k+1}) = 0$. On peut dresser le tableau des moments suivant :

r	1	2		3		4		5		6		7		8		9
$\mathbb{E}(X^r)$	0	1	$\nearrow^{\times}$	0	$\searrow_{=}$	3	$\nearrow^{\times}$	0	$\searrow_{=}$	15	$\nearrow^{\times}$	0	$\searrow_{=}$	105	$\nearrow^{\times}$	945

Les flèches précisent que les moments non nuls peuvent se calculer simplement avec un produit croisé.

Table de valeurs et loi normale centrée réduite

Soit $X \sim \mathcal{N}(0,1)$. La fonction de la répartition de X est

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt, \qquad x \in \mathbb{R}.$$

On a $F(x) = F(y) \Leftrightarrow x = y$, et $f(x) = F'(x)$, $x \in \mathbb{R}$. La fonction F ne peut pas s'écrire sous une forme analytique. On utilise alors une table de valeurs donnant $F(x)$ avec $x \in [0, 3.99]$ et $x = x_1 + x_2$:

x_1 \ x_2	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

3 Loi du Chi-deux

Densité

On dit qu'une *var* X suit la loi du Chi-deux $\chi^2(\nu)$, avec $\nu > 0$, si elle possède la densité :

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} & \text{si } x > 0, \\ 0 & \text{sinon,} \end{cases}$$

où Γ désigne la fonction gamma d'Euler définie par : $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$, $x > 0$

(cette fonction vérifie, entre autre, pour tout $m \in \mathbb{N}^*$ *:* $\Gamma(x+m) = \Gamma(x) \prod_{i=0}^{m-1} (x+i)$*).*

Le réel ν est appelé degré de liberté.

On peut remarquer que la loi du Chi-deux $\chi^2(2)$ coïncide avec la loi exponentielle $\mathcal{E}(\frac{1}{2})$;

elles possèdent la densité :

$$f(x) = \begin{cases} \dfrac{1}{2} e^{-\frac{x}{2}} & \text{si } x \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

Le résultat suivant est alors central :

La fonction f définie précédemment est bien une densité.

Preuve :

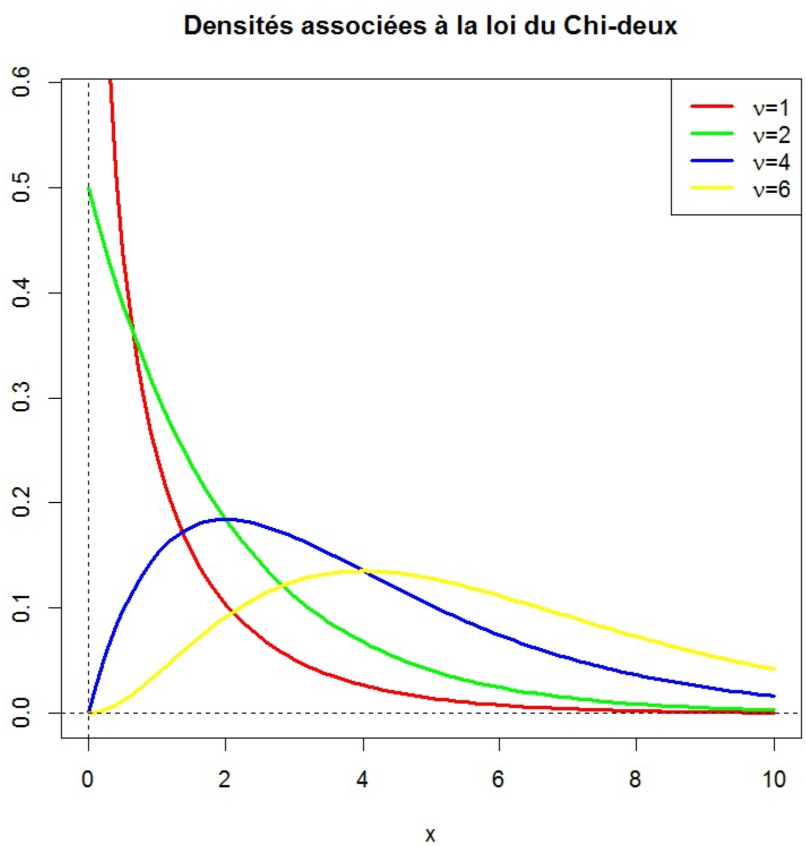
- Pour tout $x \in \mathbb{R}$, on a $f(x) \geq 0$.
- En faisant le changement de variable : $x = 2y$, et en utilisant la définition de $\Gamma(\frac{\nu}{2})$, il vient

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= \int_0^\infty \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} dx = \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty (2y)^{\frac{\nu}{2}-1} e^{-\frac{2y}{2}} 2dy \\ &= \frac{1}{\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty y^{\frac{\nu}{2}-1} e^{-y} dy = \frac{1}{\Gamma\left(\frac{\nu}{2}\right)} \Gamma\left(\frac{\nu}{2}\right) = 1. \end{aligned}$$

$\square$

Représentation graphique

Des densités associées à la loi du Chi-deux $\chi^2(\nu)$, avec plusieurs valeurs de ν, sont représentées ci-dessous :



Paramètres

Soit $X \sim \chi^2(\mu)$. Alors on a

$$\mathbb{E}(X) = \nu, \qquad \mathbb{V}(X) = 2\nu, \qquad \sigma(X) = \sqrt{2\nu}.$$

Preuve :

- En faisant le changement de variable : $x = 2y$, et en utilisant l'égalité : $\Gamma\left(\frac{\nu}{2}+1\right) = \frac{\nu}{2}\Gamma(\frac{\nu}{2})$, il vient

$$\begin{aligned}\mathbb{E}(X) &= \int_{-\infty}^{\infty} xf(x)dx = \int_0^{\infty} x\frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)}x^{\frac{\nu}{2}-1}e^{-\frac{x}{2}}dx = \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)}\int_0^{\infty}(2y)^{\frac{\nu}{2}}e^{-\frac{2y}{2}}2dy \\ &= 2\frac{1}{\Gamma\left(\frac{\nu}{2}\right)}\int_0^{\infty} y^{\frac{\nu}{2}}e^{-y}dy = 2\frac{1}{\Gamma\left(\frac{\nu}{2}\right)}\Gamma\left(\frac{\nu}{2}+1\right) = 2\times\frac{\nu}{2} = \nu.\end{aligned}$$

- On a $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$. Comme $\mathbb{E}(X) = \nu$, il reste à calculer $\mathbb{E}(X^2)$. En faisant le changement de variable : $x = 2y$, et en utilisant l'égalité : $\Gamma\left(\frac{\nu}{2}+2\right) = \left(\frac{\nu}{2}+1\right)\frac{\nu}{2}\Gamma(\frac{\nu}{2})$, il vient

$$\begin{aligned}
\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^{\infty} x^2 \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} dx = \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^{\infty} (2y)^{\frac{\nu}{2}+1} e^{-\frac{2y}{2}} 2dy \\
&= 4\frac{1}{\Gamma\left(\frac{\nu}{2}\right)} \int_0^{\infty} y^{\frac{\nu}{2}+1} e^{-y} dy = 4\frac{1}{\Gamma\left(\frac{\nu}{2}\right)} \Gamma\left(\frac{\nu}{2}+2\right) = 4\left(\frac{\nu}{2}+1\right)\frac{\nu}{2} = (\nu+2)\nu.
\end{aligned}$$

  Il s'ensuit

$$\mathbb{V}(X) = (\nu+2)\nu - (\nu)^2 = 2\nu.$$

- On a

$$\sigma(X) = \sqrt{\mathbb{V}(X)} = \sqrt{2\nu}.$$

□

Transformée de Laplace

Soit $X \sim \chi^2(\nu)$. Alors la transformée de Laplace de X est

$$L(t) = \frac{1}{(1-2t)^{\frac{\nu}{2}}}, \qquad t \in \left]-\infty, \frac{1}{2}\right[.$$

Preuve : En faisant le changement de variable : $y = (1-2t)x$, et en utilisant l'égalité : $\int_{-\infty}^{\infty} f(y)dy = 1$, pour tout $t \in \left]-\infty, \frac{1}{2}\right[$, on obtient

$$\begin{aligned}
L(t) &= \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_0^{\infty} e^{tx} \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} dx = \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^{\infty} x^{\frac{\nu}{2}-1} e^{-\frac{(1-2t)x}{2}} dx \\
&= \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^{\infty} \left(\frac{y}{1-2t}\right)^{\frac{\nu}{2}-1} e^{-\frac{y}{2}} \frac{1}{1-2t} dy = \frac{1}{(1-2t)^{\frac{\nu}{2}}} \int_0^{\infty} \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} y^{\frac{\nu}{2}-1} e^{-\frac{y}{2}} dy \\
&= \frac{1}{(1-2t)^{\frac{\nu}{2}}} \int_{-\infty}^{\infty} f(y)dy = \frac{1}{(1-2t)^{\frac{\nu}{2}}} \times 1 = \frac{1}{(1-2t)^{\frac{\nu}{2}}}.
\end{aligned}$$

□

Remarque : On a

$$L'(t) = \frac{\nu}{(1-2t)^{\frac{\nu}{2}+1}} \qquad L''(t) = \frac{\nu(\nu+2)}{(1-2t)^{\frac{\nu}{2}+2}}.$$

On retrouve alors l'espérance et la variance de X :

$$\mathbb{E}(X) = L'(0) = \nu, \qquad \mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = L''(0) - (L'(0))^2 = \nu(\nu + 2) - \nu^2 = 2\nu.$$

Liaison acec la loi normale

Soit $X \sim \mathcal{N}(0, 1)$. Alors on a $X^2 \sim \chi^2(1)$.

Preuve : Étudions la transformée de Laplace de X^2. En faisant le changement de variable : $y = x\sqrt{1-2t}$, et en utilisant l'égalité : $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 1$ (car c'est l'intégrale d'une densité associée à la loi normale $\mathcal{N}(0, 1)$), pour tout $t \in \left]-\infty, \frac{1}{2}\right[$, on obtient

$$\begin{aligned} L_{X^2}(t) &= \mathbb{E}(e^{tX^2}) = \int_{-\infty}^{\infty} e^{tx^2} f(x)dx = \int_{-\infty}^{\infty} e^{tx^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(1-2t)x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \frac{1}{\sqrt{1-2t}} dy = \frac{1}{\sqrt{1-2t}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\ &= \frac{1}{\sqrt{1-2t}} \times 1 = \frac{1}{\sqrt{1-2t}}. \end{aligned}$$

On reconnaît la transformée de Laplace associée à la loi $\chi^2(1)$. Comme la transformée de Laplace d'une *var* caractérise sa loi, on a $X^2 \sim \chi^2(1)$.

□

Somme

Soit X et Y deux *var* indépendantes avec $X \sim \chi^2(\alpha)$ et $Y \sim \chi^2(\beta)$. Alors on a

$$X + Y \sim \chi^2(\alpha + \beta).$$

Preuve : Étudions la transformée de Laplace de $X + Y$. Par l'indépendance de X et Y, et les égalités : $L_X(t) = \frac{1}{(1-2t)^{\frac{\alpha}{2}}}$ et $L_Y(t) = \frac{1}{(1-2t)^{\frac{\beta}{2}}}$, $t \in \left]-\infty, \frac{1}{2}\right[$, pour tout $t \in \left]-\infty, \frac{1}{2}\right[$, il vient

$$\begin{aligned} L_{X+Y}(t) &= \mathbb{E}(e^{t(X+Y)}) = \mathbb{E}(e^{tX} e^{tY}) = \mathbb{E}(e^{tX})\mathbb{E}(e^{tY}) = L_X(t)L_Y(t) \\ &= \frac{1}{(1-2t)^{\frac{\alpha}{2}}} \times \frac{1}{(1-2t)^{\frac{\beta}{2}}} = \frac{1}{(1-2t)^{\frac{\alpha+\beta}{2}}}. \end{aligned}$$

On reconnaît la transformée de Laplace associée à la loi $\chi^2(\alpha+\beta)$. Comme la transformée de Laplace d'une *var* caractérise sa loi, on a $X+Y \sim \chi^2(\alpha+\beta)$.

□

Applications :

- Soient X et Y deux *var iid* suivant chacune la loi normale $\mathcal{N}(0,1)$.

$$X^2+Y^2 \sim \chi^2(2).$$

  C'est une application du résultat précédent en remarquant que X^2 et Y^2 sont indépendantes avec $X^2 \sim \chi^2(1)$ et $Y^2 \sim \chi^2(1)$.

- Soient $X_1,\ldots,X_n$ n *var iid* suivant chacune la loi normale $\mathcal{N}(\mu,\sigma^2)$. Alors on a

$$K_n = \sum_{i=1}^{n} \left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n).$$

  C'est une application du résultat précédent utilisant un raisonnement par récurrence et le fait que, comme $\frac{X_1-\mu}{\sigma},\ldots,\frac{X_n-\mu}{\sigma}$ sont *iid* suivant chacune la loi normale $\mathcal{N}(0,1)$, $\left(\frac{X_1-\mu}{\sigma}\right)^2,\ldots,\left(\frac{X_n-\mu}{\sigma}\right)^2$ sont également *iid* suivant chacune la loi du Chi-deux $\chi^2(1)$.

4 Loi de Student

Densité

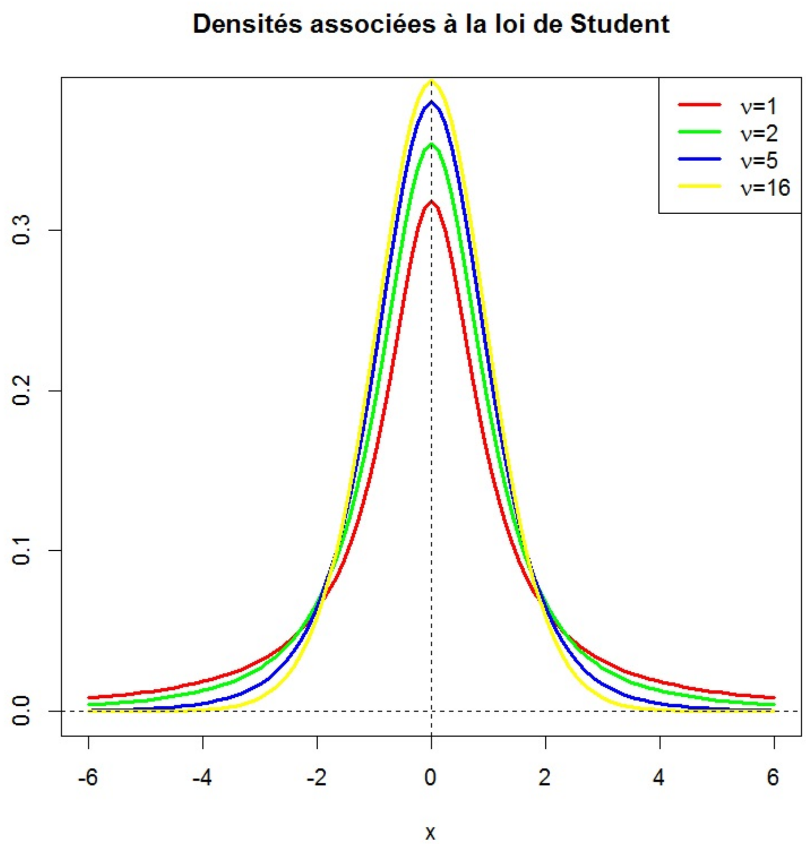
On dit qu'une *var* X suit la loi de Student $\mathcal{T}(\nu)$, avec $\nu > 0$, si elle possède la densité :

$$f(x) = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \qquad x \in \mathbb{R}.$$

Le réel ν est appelé degré de liberté.

Représentation graphique

Des densités associées à la loi de Student $\mathcal{T}(\nu)$, avec plusieurs valeurs de ν, sont représentées ci-dessous :



Caractérisation

Soient X et Y deux *var* indépendantes avec $X \sim \mathcal{N}(0,1)$ et $Y \sim \chi^2(\nu)$. Alors on a

$$T = \frac{X}{\sqrt{\frac{Y}{\nu}}} \sim \mathcal{T}(\nu).$$

Preuve :

- Dans un premier temps, déterminons une densité du couple (T, Y). On a $(T, Y) = \left(\frac{X}{\sqrt{\frac{Y}{\nu}}}, Y\right)$ et $(u, v) = \left(\frac{x}{\sqrt{\frac{y}{\nu}}}, y\right) \Leftrightarrow (x, y) = \left(u\sqrt{\frac{v}{\nu}}, v\right)$. Le théorème du changement de variable nous assure qu'une densité de (T, Y) est donnée par

$$f_{(T,Y)}(u, v) = f_{(X,Y)}\left(u\sqrt{\frac{v}{\nu}}, v\right)|J(u, v)|, \qquad (u, v) \in \mathbb{R}^2,$$

où $f_{(X,Y)}$ est une densité de (X, Y) et $J(u, v)$ est le jacobien associé à $(x, y) = \left(u\sqrt{\frac{v}{\nu}}, v\right)$.

Comme $(X, Y)(\Omega) = \mathbb{R} \times [0, \infty[$, on a $(T, Y)(\Omega) = \mathbb{R} \times [0, \infty[$. Par conséquent, pour tout $(u, v) \notin \mathbb{R} \times [0, \infty[$, on a $f_{(T,Y)}(u, v) = 0$. Pour tout $(u, v) \in \mathbb{R} \times [0, \infty[$, en utilisant l'indépendance de X et Y, *i.e.* $f_{(X,Y)}(x, y) = f_X(x)f_Y(y)$, et

$$J(u, v) = \begin{vmatrix} \frac{\partial}{\partial u}\left(u\sqrt{\frac{v}{\nu}}\right) & \frac{\partial}{\partial v}\left(u\sqrt{\frac{v}{\nu}}\right) \\ \frac{\partial}{\partial u}v & \frac{\partial}{\partial v}v \end{vmatrix} = \begin{vmatrix} \sqrt{\frac{v}{\nu}} & \frac{u}{2\sqrt{\nu v}} \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{v}{\nu}},$$

il vient

$$\begin{aligned} f_{(T,Y)}(u, v) &= f_X\left(u\sqrt{\frac{v}{\nu}}\right) f_Y(v)\sqrt{\frac{v}{\nu}} = \frac{1}{\sqrt{2\pi}}e^{-\frac{\left(u\sqrt{\frac{v}{\nu}}\right)^2}{2}} \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)}v^{\frac{\nu}{2}-1}e^{-\frac{v}{2}}\sqrt{\frac{v}{\nu}} \\ &= \frac{1}{\sqrt{2\pi\nu}2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)}v^{\frac{(\nu+1)}{2}-1}e^{-\frac{v}{2}\left(1+\frac{u^2}{\nu}\right)}. \end{aligned}$$

Au final, une densité de (T, Y) est

$$f_{(T,Y)}(u, v) = \begin{cases} \frac{1}{\sqrt{2\pi\nu}2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)}v^{\frac{(\nu+1)}{2}-1}e^{-\frac{v}{2}\left(1+\frac{u^2}{\nu}\right)} & \text{si } (u, v) \in \mathbb{R} \times [0, \infty[, \\ 0 & \text{sinon.} \end{cases}$$

- Une densité de T est donnée par : $f_T(u) = \int_{-\infty}^{\infty} f_{(T,Y)}(u,v)dv$, $u \in \mathbb{R}$. On a $T(\Omega) = \mathbb{R}$. En faisant le changement de variable : $y = \frac{v}{2}\left(1+\frac{u^2}{\nu}\right)$, pour tout $u \in \mathbb{R}$, on a

$$\begin{aligned}
f_T(u) &= \int_{-\infty}^{\infty} f_{(T,Y)}(u,v)dv = \int_0^{\infty} \frac{1}{\sqrt{2\pi\nu}2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} v^{\frac{(\nu+1)}{2}-1} e^{-\frac{v}{2}\left(1+\frac{u^2}{\nu}\right)} dv \\
&= \frac{1}{\sqrt{2\pi\nu}2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^{\infty} v^{\frac{(\nu+1)}{2}-1} e^{-\frac{v}{2}\left(1+\frac{u^2}{\nu}\right)} dv \\
&= \frac{1}{\sqrt{2\pi\nu}2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^{\infty} \left(\frac{2y}{1+\frac{u^2}{\nu}}\right)^{\frac{(\nu+1)}{2}-1} e^{-y} \frac{2}{1+\frac{u^2}{\nu}} dy \\
&= \frac{1}{\sqrt{\nu\pi}} \frac{1}{\Gamma\left(\frac{\nu}{2}\right)} \left(1+\frac{u^2}{\nu}\right)^{-\frac{\nu+1}{2}} \int_0^{\infty} y^{\frac{(\nu+1)}{2}-1} e^{-y} dy = \frac{1}{\sqrt{\nu\pi}} \frac{1}{\Gamma\left(\frac{\nu}{2}\right)} \left(1+\frac{u^2}{\nu}\right)^{-\frac{\nu+1}{2}} \Gamma\left(\frac{\nu+1}{2}\right).
\end{aligned}$$

  Ainsi, une densité de T est

$$f_T(u) = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1+\frac{u^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \qquad u \in \mathbb{R}.$$

$\square$

Remarque : Le résultat précédent prouve que f est bien une densité.

Paramètres

Soit $X \sim \mathcal{T}(\nu)$.

- Si $\nu \leq 1$, l'espérance de X n'existe pas, sinon on a

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

- Si $\nu \leq 2$, la variance de X n'existe pas, sinon on a

$$\mathbb{V}(X) = \frac{\nu}{\nu-2}, \qquad \sigma(X) = \sqrt{\frac{\nu}{\nu-2}}.$$

Preuve :

- Comme X possède une densité f paire, X est symétrique ; quand elle existe, sont espérance est nulle : $\mathbb{E}(X) = 0$.
- On a $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$. Comme $\mathbb{E}(X) = 0$, il reste à calculer $\mathbb{E}(X^2)$.

Le plus rapide est d'utiliser la caractérisation de la loi de Student démontrée précédemment : X suit la même loi que $\frac{A}{\sqrt{\frac{B}{\nu}}}$, avec A et B deux *var* indépendantes, $A \sim \mathcal{N}(0,1)$ et $B \sim \chi^2(\nu)$. Comme A et B sont indépendantes, il en est de même pour A^2 et $\frac{1}{B}$, ce qui entraîne

$$\mathbb{E}(X^2) = \mathbb{E}\left(\left(\frac{A}{\sqrt{\frac{B}{\nu}}}\right)^2\right) = \nu\mathbb{E}\left(\frac{A^2}{B}\right) = \nu\mathbb{E}(A^2)\mathbb{E}\left(\frac{1}{B}\right).$$

Comme $A \sim \mathcal{N}(0,1)$, on a $\mathbb{E}(A^2) = \mathbb{V}(A) + (\mathbb{E}(A))^2 = 1 + 0^2 = 1$. D'autre part, comme $B \sim \chi^2(\nu)$, en faisant le changement de variable : $x = 2y$ et en utilisant l'égalité : $\Gamma\left(\frac{\nu}{2}\right) = \left(\frac{\nu}{2} - 1\right)\Gamma\left(\frac{\nu}{2} - 1\right)$, il vient

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{B}\right) &= \int_{-\infty}^{\infty} \frac{1}{x} f_B(x)dx = \int_0^{\infty} \frac{1}{x} \times \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} dx \\
&= \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^{\infty} x^{\frac{\nu}{2}-2} e^{-\frac{x}{2}} dx = \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^{\infty} (2y)^{\frac{\nu}{2}-2} e^{-y} 2dy \\
&= \frac{1}{2\Gamma\left(\frac{\nu}{2}\right)} \int_0^{\infty} y^{\frac{\nu}{2}-2} e^{-y} dy = \frac{1}{2\Gamma\left(\frac{\nu}{2}\right)} \Gamma\left(\frac{\nu}{2} - 1\right) = \frac{1}{2} \times \frac{1}{\frac{\nu}{2} - 1} = \frac{1}{\nu - 2}.
\end{aligned}$$

On en déduit que

$$\mathbb{E}(X^2) = \nu \times 1 \times \frac{1}{\nu - 2} = \frac{\nu}{\nu - 2}.$$

Il s'ensuit

$$\mathbb{V}(X) = \frac{\nu}{\nu - 2} - 0^2 = \frac{\nu}{\nu - 2}.$$

- On a

$$\sigma(X) = \sqrt{\mathbb{V}(X)} = \sqrt{\frac{\nu}{\nu - 2}}.$$

□

Approximation

Lorsque $\nu \geq 31$, on peut approcher la loi de Student $\mathcal{T}(\nu)$ par la loi normale $\mathcal{N}(0,1)$.

Éléments de preuve : Cette approximation, tournée vers pratique, repose sur le lemme de Scheffé et la convergence :

$$\lim_{\nu \to \infty} \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad x \in \mathbb{R}.$$

□

5 Loi de Fisher

Densité

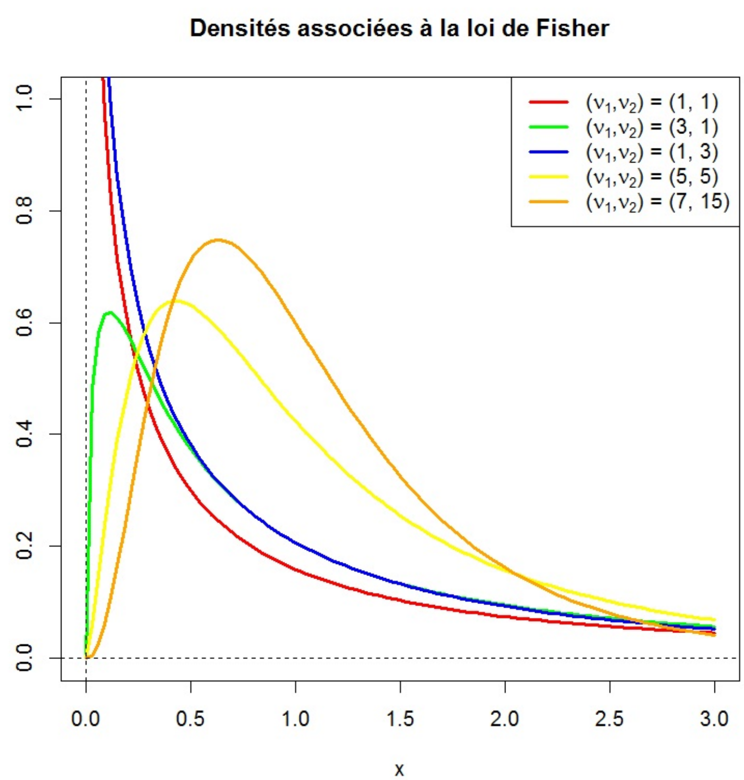
On dit qu'une *var* X suit la loi de Fisher $\mathcal{F}(\nu_1, \nu_2)$, avec $\nu_1 > 0$ et $\nu_2 > 0$, si elle possède la densité :

$$f(x) = \begin{cases} \dfrac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)x}\left(\dfrac{\nu_1 x}{\nu_1 x+\nu_2}\right)^{\frac{\nu_1}{2}}\left(1-\dfrac{\nu_1 x}{\nu_1 x+\nu_2}\right)^{\frac{\nu_2}{2}} & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}$$

Les réels ν_1 et ν_2 sont appelés degrés de liberté.

Représentation graphique

Des densités associées à la loi de Fisher $\mathcal{F}(\nu_1, \nu_2)$, avec plusieurs valeurs de ν_1 et ν_2, sont représentées ci-dessous :



Caractérisation

Soient X et Y deux *var* indépendantes avec $X \sim \chi^2(\nu_1)$ et $Y \sim \chi^2(\nu_2)$. Alors on a

$$F = \frac{\frac{X}{\nu_1}}{\frac{Y}{\nu_2}} \sim \mathcal{F}(\nu_1, \nu_2).$$

Preuve :

- Dans un premier temps, déterminons une densité du couple (F, Y). On a $(F, Y) = \left(\frac{\frac{X}{\nu_1}}{\frac{Y}{\nu_2}}, Y\right)$ et $(u, v) = \left(\frac{\frac{x}{\nu_1}}{\frac{y}{\nu_2}}, y\right) \Leftrightarrow (x, y) = \left(\frac{\nu_1}{\nu_2} uv, v\right)$. Le théorème du changement de variable nous assure qu'une densité de (F, Y) est donnée par

$$f_{(F,Y)}(u, v) = f_{(X,Y)}\left(\frac{\nu_1}{\nu_2} uv, v\right) |J(u, v)|, \qquad (u, v) \in \mathbb{R}^2,$$

où $f_{(X,Y)}$ est une densité de (X, Y) et $J(u, v)$ est le jacobien associé à $(x, y) = \left(\frac{\nu_1}{\nu_2} uv, v\right)$.
Comme $(X, Y)(\Omega) = [0, \infty[^2$, on a $(F, Y)(\Omega) = [0, \infty[^2$. Par conséquent, pour tout $(u, v) \notin [0, \infty[^2$, on a $f_{(F,Y)}(u, v) = 0$. Pour tout $(u, v) \in \mathbb{R} \times [0, \infty[$, en utilisant l'indépendance de X et Y, *i.e.* $f_{(X,Y)}(x, y) = f_X(x) f_Y(y)$, et

$$J(u, v) = \begin{vmatrix} \frac{\partial}{\partial u}\left(\frac{\nu_1}{\nu_2} uv\right) & \frac{\partial}{\partial v}\left(\frac{\nu_1}{\nu_2} uv\right) \\ \frac{\partial}{\partial u} v & \frac{\partial}{\partial v} v \end{vmatrix} = \begin{vmatrix} \frac{\nu_1}{\nu_2} v & \frac{\nu_1}{\nu_2} u \\ 0 & 1 \end{vmatrix} = \frac{\nu_1}{\nu_2} v,$$

il vient

$$\begin{aligned} f_{(F,Y)}(u, v) &= f_X\left(\frac{\nu_1}{\nu_2} uv\right) f_Y(v) \left|\frac{\nu_1}{\nu_2} v\right| \\ &= \frac{1}{2^{\frac{\nu_1}{2}} \Gamma\left(\frac{\nu_1}{2}\right)} \left(\frac{\nu_1}{\nu_2} uv\right)^{\frac{\nu_1}{2} - 1} e^{-\frac{\frac{\nu_1}{\nu_2} uv}{2}} \frac{1}{2^{\frac{\nu_2}{2}} \Gamma\left(\frac{\nu_2}{2}\right)} v^{\frac{\nu_2}{2} - 1} e^{-\frac{v}{2}} \frac{\nu_1}{\nu_2} v \\ &= \frac{1}{2^{\frac{\nu_1+\nu_2}{2}} \Gamma\left(\frac{\nu_1}{2}\right) \Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} u^{\frac{\nu_1}{2} - 1} v^{\frac{\nu_1+\nu_2}{2} - 1} e^{-\frac{v}{2}\left(1 + \frac{\nu_1}{\nu_2} u\right)}. \end{aligned}$$

Au final, une densité de (F, Y) est

$$f_{(F,Y)}(u, v) = \begin{cases} \frac{1}{2^{\frac{\nu_1+\nu_2}{2}} \Gamma\left(\frac{\nu_1}{2}\right) \Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} u^{\frac{\nu_1}{2} - 1} v^{\frac{\nu_1+\nu_2}{2} - 1} e^{-\frac{v}{2}\left(1 + \frac{\nu_1}{\nu_2} u\right)} & \text{si } (u, v) \in [0, \infty[^2, \\ 0 & \text{sinon.} \end{cases}$$

- Une densité de F est donnée par : $f_F(u) = \int_{-\infty}^{\infty} f_{(F,Y)}(u,v)dv$, $u \in \mathbb{R}$. On a $F(\Omega) = [0,\infty[$. En faisant le changement de variable : $y = \frac{v}{2}\left(1 + \frac{\nu_1}{\nu_2}u\right)$, pour tout $u \in [0,\infty[$, on a

$$\begin{aligned}
f_F(u) &= \int_{-\infty}^{\infty} f_{(F,Y)}(u,v)dv = \int_0^{\infty} \frac{1}{2^{\frac{\nu_1+\nu_2}{2}}\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} u^{\frac{\nu_1}{2}-1} v^{\frac{\nu_1+\nu_2}{2}-1} e^{-\frac{v}{2}\left(1+\frac{\nu_1}{\nu_2}u\right)} dv \\
&= \frac{1}{2^{\frac{\nu_1+\nu_2}{2}}\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} u^{\frac{\nu_1}{2}-1} \int_0^{\infty} v^{\frac{\nu_1+\nu_2}{2}-1} e^{-\frac{v}{2}\left(1+\frac{\nu_1}{\nu_2}u\right)} dv \\
&= \frac{1}{2^{\frac{\nu_1+\nu_2}{2}}\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} u^{\frac{\nu_1}{2}-1} \int_0^{\infty} \left(\frac{2y}{1+\frac{\nu_1}{\nu_2}u}\right)^{\frac{\nu_1+\nu_2}{2}-1} e^{-y} \frac{2}{1+\frac{\nu_1}{\nu_2}u} dy \\
&= \frac{1}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} u^{\frac{\nu_1}{2}-1} \left(1+\frac{\nu_1}{\nu_2}u\right)^{-\frac{\nu_1+\nu_2}{2}} \int_0^{\infty} y^{\frac{\nu_1+\nu_2}{2}-1} e^{-y} dy \\
&= \frac{1}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} u^{\frac{\nu_1}{2}-1} \left(1+\frac{\nu_1}{\nu_2}u\right)^{-\frac{\nu_1+\nu_2}{2}} \Gamma\left(\frac{\nu_1+\nu_2}{2}\right) \\
&= \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)u}\left(\frac{\nu_1 u}{\nu_1 u+\nu_2}\right)^{\frac{\nu_1}{2}}\left(1-\frac{\nu_1 u}{\nu_1 u+\nu_2}\right)^{\frac{\nu_2}{2}}.
\end{aligned}$$

Ainsi, une densité de F est

$$f_F(u) = \begin{cases} \dfrac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)u}\left(\dfrac{\nu_1 u}{\nu_1 u+\nu_2}\right)^{\frac{\nu_1}{2}}\left(1-\dfrac{\nu_1 u}{\nu_1 u+\nu_2}\right)^{\frac{\nu_2}{2}} & \text{si } u \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

$\square$

Remarque : Le résultat précédent prouve que f est bien une densité.

Applications :

- Si $X \sim \mathcal{F}(\nu_1,\nu_2)$, alors on a $\frac{1}{X} \sim \mathcal{F}(\nu_2,\nu_1)$.
  C'est une application du résultat précédent : si $X \sim \mathcal{F}(\nu_1,\nu_2)$, alors X suit la même loi que $\frac{\frac{A}{\nu_1}}{\frac{B}{\nu_2}}$, avec A et B deux *var* indépendantes, $A \sim \chi^2(\nu_1)$ et $B \sim \chi^2(\nu_2)$. Donc $\frac{1}{X} = X^{-1}$ suit la même loi que $\left(\frac{\frac{A}{\nu_1}}{\frac{B}{\nu_2}}\right)^{-1} = \frac{\frac{B}{\nu_2}}{\frac{A}{\nu_1}}$, laquelle suit la loi de Fisher $\mathcal{F}(\nu_2,\nu_1)$.
- Si $X \sim \mathcal{T}(\nu)$, alors on a $X^2 \sim \mathcal{F}(1,\nu)$. C'est une application du résultat précédent : si $X \sim \mathcal{T}(\nu)$, alors X suit la même loi que $\frac{A}{\sqrt{\frac{B}{\nu}}}$, avec A et B deux *var* indépendantes, $A \sim \mathcal{N}(0,1)$ et $B \sim \chi^2(\nu)$. Donc X^2 suit la même loi que $\frac{A^2}{\frac{B}{\nu}} = \frac{\frac{A^2}{1}}{\frac{B}{\nu}}$, laquelle suit la loi de Fisher $\mathcal{F}(1,\nu)$ car A^2 et B sont indépendantes, $A^2 \sim \chi^2(1)$ et $B \sim \chi^2(\nu)$.

Paramètres

Soit $X \sim \mathcal{F}(\nu_1, \nu_2)$.

- Si $\nu_2 < 3$, l'espérance de X n'existe pas, sinon on a
$$\mathbb{E}(X) = \frac{\nu_2}{\nu_2 - 2}.$$
- Si $\nu_2 < 4$, la variance de X n'existe pas, sinon on a
$$\mathbb{V}(X) = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}, \qquad \sigma(X) = \sqrt{\frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}}.$$

Preuve :

- Le plus rapide est d'utiliser la caractérisation de la loi de Fisher démontrée précédemment : X suit la même loi que $\frac{\frac{A}{\nu_1}}{\frac{B}{\nu_2}}$, avec A et B deux *var* indépendantes, $A \sim \chi^2(\nu_1)$ et $B \sim \chi^2(\nu_2)$. Comme A et B sont indépendantes, il en est de même pour A et $\frac{1}{B}$, ce qui entraîne
$$\mathbb{E}(X) = \mathbb{E}\left(\frac{\frac{A}{\nu_1}}{\frac{B}{\nu_2}}\right) = \frac{\nu_2}{\nu_1}\mathbb{E}\left(\frac{A}{B}\right) = \frac{\nu_2}{\nu_1}\mathbb{E}(A)\mathbb{E}\left(\frac{1}{B}\right).$$
Comme $A \sim \chi^2(\nu_1)$, on a $\mathbb{E}(A) = \nu_1$. D'autre part, comme $B \sim \chi^2(\nu_2)$, en faisant le changement de variable : $x = 2y$ et en utilisant l'égalité : $\Gamma\left(\frac{\nu_2}{2}\right) = \left(\frac{\nu_2}{2} - 1\right)\Gamma\left(\frac{\nu_2}{2} - 1\right)$, il vient
$$\begin{aligned}
\mathbb{E}\left(\frac{1}{B}\right) &= \int_{-\infty}^{\infty} \frac{1}{x} f_B(x)dx = \int_0^{\infty} \frac{1}{x} \times \frac{1}{2^{\frac{\nu_2}{2}}\Gamma\left(\frac{\nu_2}{2}\right)} x^{\frac{\nu_2}{2}-1} e^{-\frac{x}{2}} dx \\
&= \frac{1}{2^{\frac{\nu_2}{2}}\Gamma\left(\frac{\nu_2}{2}\right)} \int_0^{\infty} x^{\frac{\nu_2}{2}-2} e^{-\frac{x}{2}} dx = \frac{1}{2^{\frac{\nu_2}{2}}\Gamma\left(\frac{\nu_2}{2}\right)} \int_0^{\infty} (2y)^{\frac{\nu_2}{2}-2} e^{-y} 2dy \\
&= \frac{1}{2\Gamma\left(\frac{\nu_2}{2}\right)} \int_0^{\infty} y^{\frac{\nu_2}{2}-2} e^{-y} dy = \frac{1}{2\Gamma\left(\frac{\nu_2}{2}\right)} \Gamma\left(\frac{\nu_2}{2} - 1\right) = \frac{1}{2} \times \frac{1}{\frac{\nu_2}{2} - 1} = \frac{1}{\nu_2 - 2}.
\end{aligned}$$
On en déduit que
$$\mathbb{E}(X) = \frac{\nu_2}{\nu_1} \times \nu_1 \times \frac{1}{\nu_2 - 2} = \frac{\nu_2}{\nu_2 - 2}.$$
- On a $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$. Comme $\mathbb{E}(X) = \frac{\nu_2}{\nu_2 - 2}$, il reste à calculer $\mathbb{E}(X^2)$. Nous allons utiliser de nouveau la caractérisation de la loi de Student : X suit la même loi que $\frac{\frac{A}{\nu_1}}{\frac{B}{\nu_2}}$, avec A et B deux *var* indépendantes, $A \sim \chi^2(\nu_1)$ et $B \sim \chi^2(\nu_2)$.

Comme A et B sont indépendantes, il en est de même pour A^2 et $\frac{1}{B^2}$, ce qui entraîne

$$\mathbb{E}(X^2) = \mathbb{E}\left(\left(\frac{\frac{A}{\nu_1}}{\frac{B}{\nu_2}}\right)^2\right) = \frac{\nu_2^2}{\nu_1^2}\mathbb{E}\left(\frac{A^2}{B^2}\right) = \frac{\nu_2^2}{\nu_1^2}\mathbb{E}(A^2)\mathbb{E}\left(\frac{1}{B^2}\right).$$

Comme $A \sim \chi^2(\nu_1)$, on a

$$\mathbb{E}(A^2) = \mathbb{V}(A) + (\mathbb{E}(A))^2 = 2\nu_1 + (\nu_1)^2 = \nu_1(\nu_1 + 2).$$

D'autre part, comme $B \sim \chi^2(\nu_2)$, en faisant le changement de variable : $x = 2y$ et en utilisant l'égalité : $\Gamma\left(\frac{\nu_2}{2}\right) = \left(\frac{\nu_2}{2} - 1\right)\left(\frac{\nu_2}{2} - 2\right)\Gamma\left(\frac{\nu_2}{2} - 2\right)$, il vient

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{B^2}\right) &= \int_{-\infty}^{\infty} \frac{1}{x^2} f_B(x)dx = \int_0^{\infty} \frac{1}{x^2} \times \frac{1}{2^{\frac{\nu_2}{2}}\Gamma\left(\frac{\nu_2}{2}\right)} x^{\frac{\nu_2}{2}-1} e^{-\frac{x}{2}} dx \\
&= \frac{1}{2^{\frac{\nu_2}{2}}\Gamma\left(\frac{\nu_2}{2}\right)} \int_0^{\infty} x^{\frac{\nu_2}{2}-3} e^{-\frac{x}{2}} dx = \frac{1}{2^{\frac{\nu_2}{2}}\Gamma\left(\frac{\nu_2}{2}\right)} \int_0^{\infty} (2y)^{\frac{\nu_2}{2}-3} e^{-y} 2dy \\
&= \frac{1}{2^2\Gamma\left(\frac{\nu_2}{2}\right)} \int_0^{\infty} y^{\frac{\nu_2}{2}-3} e^{-y} dy = \frac{1}{2^2\Gamma\left(\frac{\nu_2}{2}\right)}\Gamma\left(\frac{\nu_2}{2} - 2\right) = \frac{1}{2^2} \times \frac{1}{\left(\frac{\nu_2}{2} - 1\right)\left(\frac{\nu_2}{2} - 2\right)} \\
&= \frac{1}{(\nu_2 - 2)(\nu_2 - 4)}.
\end{aligned}$$

On en déduit que

$$\mathbb{E}(X^2) = \frac{\nu_2^2}{\nu_1^2} \times \nu_1(\nu_1 + 2) \times \frac{1}{(\nu_2 - 2)(\nu_2 - 4)} = \frac{\nu_2^2(\nu_1 + 2)}{\nu_1(\nu_2 - 2)(\nu_2 - 4)}.$$

Au final, on a

$$\begin{aligned}
\mathbb{V}(X) &= \frac{\nu_2^2(\nu_1 + 2)}{\nu_1(\nu_2 - 2)(\nu_2 - 4)} - \left(\frac{\nu_2}{\nu_2 - 2}\right)^2 = \frac{\nu_2^2(\nu_1 + 2)(\nu_1 - 2) - \nu_2^2\nu_1(\nu_2 - 4)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)} \\
&= \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}.
\end{aligned}$$

∘ Il vient

$$\sigma(X) = \sqrt{\mathbb{V}(X)} = \sqrt{\frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}}.$$

□

6 Loi du Chi-deux non centrée

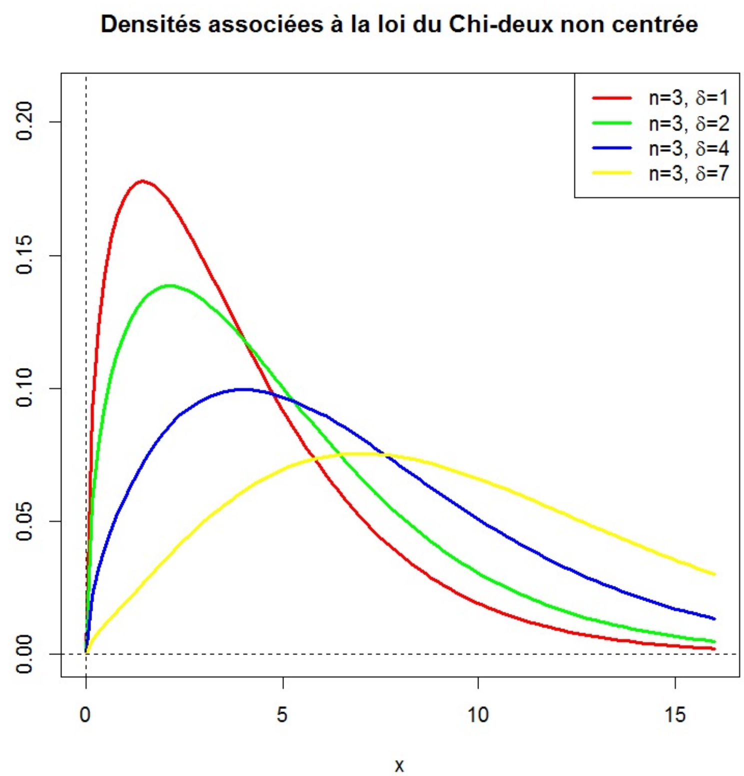
Densité

On dit qu'une *var* X suit la loi du Chi-deux non centrée $\chi^2(n,\delta)$, avec $n \in \mathbb{N}^*$ et $\delta > 0$, si elle possède la densité :

$$f(x) = \begin{cases} e^{-\frac{\delta}{2}} \displaystyle\sum_{r=0}^{\infty} \frac{1}{r!} \left(\frac{\delta}{2}\right)^r \frac{1}{2^{\frac{n}{2}+r}\Gamma\left(\frac{n}{2}+r\right)} x^{\frac{n}{2}+r-1} e^{-\frac{x}{2}} & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}$$

Représentation graphique

Des densités associées à la loi du Chi-deux non centrée $\chi^2(n,\delta)$, avec $n = 3$ et plusieurs valeurs de δ, sont représentées ci-dessous :



Caractérisation

Soient $X_1, \ldots, X_n$ n *var* indépendantes avec $X_i \sim \mathcal{N}(\mu_i, 1)$ pour tout $i \in \{1, \ldots, n\}$. On pose $\delta = \sum_{i=1}^{n} \mu_i^2$. Alors on a

$$K_n = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n, \delta).$$

Paramètres

Soit $X \sim \chi^2(n, \delta)$. Alors on a

$$\mathbb{E}(X) = n + \delta, \qquad \mathbb{V}(X) = 2(n + 2\delta), \qquad \sigma(X) = \sqrt{2(n + 2\delta)}.$$

Preuve :

- Le plus rapide est d'utiliser la caractérisation de loi du Chi-deux non centrée présentée précédemment : X suit la même loi que $\sum_{i=1}^{n} A_i^2$, avec $A_1, \ldots, A_n$ indépendantes et $A_i \sim \mathcal{N}(\mu_i, 1)$ pour tout $i \in \{1, \ldots, n\}$. En utilisant l'égalité : $\mathbb{E}(A_i^2) = \mathbb{V}(A_i) + (\mathbb{E}(A_i))^2 = 1 + \mu_i^2$, il vient

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{n} A_i^2\right) = \sum_{i=1}^{n} \mathbb{E}(A_i^2) = \sum_{i=1}^{n} (1 + \mu_i^2) = n + \sum_{i=1}^{n} \mu_i^2 = n + \delta.$$

- Nous allons utiliser de nouveau la caractérisation de la loi du Chi-deux non centrée : X suit la même loi que $\sum_{i=1}^{n} A_i^2$, avec $A_1, \ldots, A_n$ indépendantes et $A_i \sim \mathcal{N}(\mu_i, 1)$ pour tout $i \in \{1, \ldots, n\}$. Ainsi, l'indépendance des *var* $A_1, \ldots, A_n$ entraîne

$$\mathbb{V}(X) = \mathbb{V}\left(\sum_{i=1}^{n} A_i^2\right) = \sum_{i=1}^{n} \mathbb{V}(A_i^2).$$

D'autre part, pour tout $i \in \{1, \ldots, n\}$, comme $A_i - \mu_i \sim \mathcal{N}(0, 1)$, on a $\mathbb{E}((A_i - \mu_i)^4) = 3$, $\mathbb{E}((A_i - \mu_i)^3) = 0$, $\mathbb{E}((A_i - \mu_i)^2) = 1$ et $\mathbb{E}(A_i - \mu_i) = 0$. Il s'ensuit :

$$\begin{aligned}
\mathbb{V}(A_i^2) &= \mathbb{E}((A_i^2 - \mathbb{E}(A_i^2))^2) = \mathbb{E}((A_i^2 - (1 + \mu_i^2))^2) = \mathbb{E}(((A_i - \mu_i + \mu_i)^2 - 1 - \mu_i^2)^2) \\
&= \mathbb{E}(((A_i - \mu_i)^2 + 2\mu_i(A_i - \mu_i) - 1)^2) = \mathbb{E}((A_i - \mu_i)^4) + 2\mu_i \mathbb{E}((A_i - \mu_i)^3) - \mathbb{E}((A_i - \mu_i)^2) \\
&+ 2\mu_i \mathbb{E}((A_i - \mu_i)^3) + 4\mu_i^2 \mathbb{E}((A_i - \mu_i)^2) - 2\mu_i \mathbb{E}(A_i - \mu_i) - \mathbb{E}((A_i - \mu_i)^2) - 2\mu_i \mathbb{E}(A_i - \mu_i) + 1 \\
&= 3 - 1 + 4\mu_i^2 - 1 + 1 = 2 + 4\mu_i^2 = 2(1 + 2\mu_i^2).
\end{aligned}$$

On en déduit que

$$\mathbb{V}(X) = \mathbb{V}\left(\sum_{i=1}^{n} A_i^2\right) = \sum_{i=1}^{n} 2(1+2\mu_i^2) = 2\left(n + 2\sum_{i=1}^{n} \mu_i^2\right) = 2(n+2\delta).$$

∘ Il vient

$$\sigma(X) = \sqrt{\mathbb{V}(X)} = \sqrt{2(n+2\delta)}.$$

□

Transformée de Laplace

Soit $X \sim \chi^2(n, \delta)$. La transformée de Laplace de X est

$$L(t) = \frac{1}{(1-2t)^{\frac{n}{2}}} e^{\frac{t\delta}{1-2t}}, \qquad t \in \left]-\infty, \frac{1}{2}\right[.$$

Preuve : Nous allons utiliser la caractérisation de loi du Chi-deux non centrée présentée précédemment : X suit la même loi que $\sum_{i=1}^{n} A_i^2$, avec $A_1, \ldots, A_n$ indépendantes et $A_i \sim \mathcal{N}(\mu_i, 1)$ pour tout $i \in \{1, \ldots, n\}$. On a

$$L(t) = \mathbb{E}(e^{tX}) = \mathbb{E}\left(e^{t\sum_{i=1}^{n} A_i^2}\right) = \mathbb{E}\left(\prod_{i=1}^{n} e^{tA_i^2}\right) = \prod_{i=1}^{n} \mathbb{E}(e^{tA_i^2}).$$

Pour tout $i \in \{1, \ldots, n\}$, en notant f_{A_i} une densité de A_i, et en utilisant l'égalité : $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\beta^2}} e^{-\frac{(x-\alpha)^2}{2\beta^2}} dx = 1$, $\alpha \in \mathbb{R}$, $\beta > 0$ (car c'est l'intégrale d'une densité associée à la loi normale $\mathcal{N}(\alpha, \beta^2)$), on obtient

$$\begin{aligned}
\mathbb{E}(e^{tA_i^2}) &= \int_{-\infty}^{\infty} e^{tx^2} f_{A_i}(x)dx = \int_{-\infty}^{\infty} e^{tx^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu_i)^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(1-2t)x^2 - 2\mu_i x + \mu_i^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu_i^2}{2}} \int_{-\infty}^{\infty} e^{-(1-2t)\frac{1}{2}\left(x^2 - 2\frac{\mu_i}{1-2t}x\right)} dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu_i^2}{2}} e^{\frac{1}{2}\frac{\mu_i^2}{1-2t}} \int_{-\infty}^{\infty} e^{-(1-2t)\frac{1}{2}\left(x - \frac{\mu_i}{1-2t}\right)^2} dx \\
&= \frac{1}{\sqrt{1-2t}} e^{\frac{t\mu_i^2}{1-2t}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\left(\frac{1}{1-2t}\right)}} e^{-\frac{\left(x - \frac{\mu_i}{1-2t}\right)^2}{2\left(\frac{1}{1-2t}\right)}} dx = \frac{1}{\sqrt{1-2t}} e^{\frac{t\mu_i^2}{1-2t}} \times 1 = \frac{1}{\sqrt{1-2t}} e^{\frac{t\mu_i^2}{1-2t}}.
\end{aligned}$$

Il s'ensuit :

$$L(t) = \prod_{i=1}^{n} \frac{1}{\sqrt{1-2t}} e^{\frac{t\mu_i^2}{1-2t}} = \frac{1}{(1-2t)^{\frac{n}{2}}} e^{\sum_{i=1}^{n} \frac{t\mu_i}{1-2t}} = \frac{1}{(1-2t)^{\frac{n}{2}}} e^{\frac{t\delta}{1-2t}}.$$

□

Remarque : En fait, c'est grâce à la caractérisation de la loi du Chi-deux non centrée et l'expression de sa transformée de Laplace que l'on peut expliciter une densité associée.

Complément : loi de Fisher non centrée

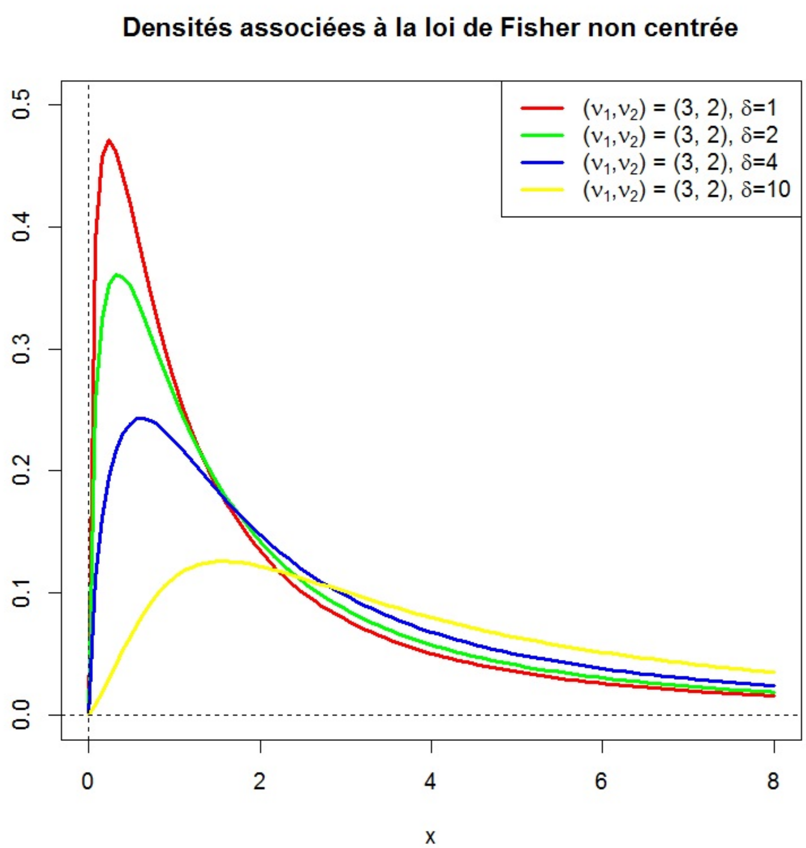
La loi de Fisher non centrée, notée $\mathcal{F}(n_1, n_2, \delta)$, $n_1 \in \mathbb{N}^*$ et $n_2 \in \mathbb{N}^*$ et $\delta > 0$, est caractérisée par le résultat suivant.

Soient X et Y deux *var* indépendantes avec $X \sim \chi^2(n_1, \delta)$ et $Y \sim \chi^2(n_2)$. Alors on a

$$F = \frac{\frac{X}{n_1}}{\frac{Y}{n_2}} \sim \mathcal{F}(n_1, n_2, \delta).$$

On est alors en mesure d'expliciter une densité associée et les paramètres de cette loi.

Des densités associées à la loi de Fisher non centrée $\mathcal{F}(n_1, n_2, \delta)$, avec $(\nu_1, \nu_2) = (3, 2)$ et plusieurs valeurs de δ, sont représentées ci-dessous :



7 Forme quadratique aléatoire

Définition

Soient $X_1, \ldots, X_n$ n *var* indépendantes et $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$. On appelle forme quadratique en X toute *var* Q telle qu'il existe une matrice carré symétrique $A = (a_{i,j})_{(i,j)\in\{1,\ldots,n\}^2}$ vérifiant

$$Q = X^t A X = \sum_{i=1}^{n} a_{i,i} X_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} X_i X_j = \sum_{i=1}^{n} a_{i,i} X_i^2 + 2 \sum_{i=2}^{n} \sum_{j=1}^{i-1} a_{i,j} X_i X_j.$$

Résultats en loi

Soient $X_1, \ldots, X_n$ n *var* indépendantes avec $X_i \sim \mathcal{N}(\mu_i, \sigma^2)$ pour tout $i \in \{1, \ldots, n\}$ et $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$. Soit A une matrice carré symétrique A de dimension $n \times n$ telle que $A^2 = A$.

On pose $Q = X^t A X$, $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$, $\delta = \frac{1}{\sigma^2} \mu^t A \mu$ et $d = \mathrm{rang}(A)$. Alors on a

$$\frac{Q}{\sigma^2} \sim \chi^2(d, \delta).$$

Commentaire : La preuve utilise des propriétés d'indépendance de la loi normale multidimensionnelle.

Soient $X_1, \ldots, X_n$ n *var* indépendantes avec $X_i \sim \mathcal{N}(\mu_i, \sigma^2)$ pour tout $i \in \{1, \ldots, n\}$. On pose :

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2.$$

Alors on a

$$\frac{n-1}{\sigma^2} S_n^2 \sim \chi^2(n-1).$$

Preuve : Nous cherchons à appliquer le résultat précédent. En posant $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$, après décomposition, on peut écrire :

$$(n-1)S_n^2 = \sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = \left(1 - \frac{1}{n}\right)\sum_{i=1}^{n} X_i^2 - \frac{2}{n}\sum_{i=2}^{n}\sum_{j=1}^{i-1} X_i X_j = X^t A X,$$

où A est la matrice symétrique de dimension $n \times n$ définie par

$$A = \mathbb{I}_n - \frac{1}{n} 1_{n\times n} = \begin{pmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \ldots & \ldots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \ldots & \ldots & -\frac{1}{n} \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & -\frac{1}{n} \\ -\frac{1}{n} & \ldots & \ldots & \ldots & 1-\frac{1}{n} \end{pmatrix}.$$

Comme $1_{n\times n}1_{n\times n} = n 1_{n\times n}$, on a

$$A^2 = \left(\mathbb{I}_n - \frac{1}{n}1_{n\times n}\right)\left(\mathbb{I}_n - \frac{1}{n}1_{n\times n}\right) = \mathbb{I}_n - \frac{2}{n}1_{n\times n} + \frac{1}{n^2}1_{n\times n}1_{n\times n} = \mathbb{I}_n - \frac{1}{n}1_{n\times n} = A.$$

En posant $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$, la somme des valeurs des éléments de chaque colonne de A étant nulle, on obtient $\delta = \frac{1}{\sigma^2}\mu^t A \mu = 0$. De plus, comme $A^2 = A$, on a $d = \mathrm{rang}(A) = \mathrm{trace}(A) = n\left(1 - \frac{1}{n}\right) = n - 1$. Le résultat précédent implique :

$$\frac{n-1}{\sigma^2} S_n^2 \sim \chi^2(n-1, 0) = \chi^2(n-1).$$

□

Résultats sur l'indépendance

Soient $X_1, \ldots, X_n$ n *var* indépendantes avec $X_i \sim \mathcal{N}(\mu_i, \sigma^2)$ pour tout $i \in \{1, \ldots, n\}$ et $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$. Soient A et B deux matrices carrés symétriques A et B de dimension $n \times n$ telles que $A^2 = A$, $B^2 = B$ et $AB = BA = 0_{n,n}$. Alors les *var* $Q_1 = X^t AX$ et $Q_2 = X^t BX$ sont indépendantes.

Commentaire : La preuve utilise des propriétés d'indépendance de la loi normale multidimensionnelle.

Soient $X_1, \ldots, X_n$ n *var* indépendantes avec $X_i \sim \mathcal{N}(\mu_i, \sigma^2)$ pour tout $i \in \{1, \ldots, n\}$. On pose :

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2.$$

Alors les *var* $\overline{X}_n$ et S_n sont indépendantes.

Preuve : Nous cherchons à appliquer le résultat précédent pour montrer que $n\overline{X}_n^2$ et $(n-1)S_n^2$ sont indépendances. On pose $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$. On peut alors écrire :

$$n\overline{X}_n^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 + \frac{2}{n} \sum_{i=2}^{n} \sum_{j=1}^{i-1} X_i X_j,$$

donc $n\overline{X}_n^2 = X^t AX$ avec $A = \frac{1}{n} 1_{n \times n}$.

De même, on peut écrire :

$$(n-1)S_n^2 = \sum_{i=1}^{n} (X_i - \overline{X}_n)^2 = \left(1 - \frac{1}{n}\right) \sum_{i=1}^{n} X_i^2 - \frac{2}{n} \sum_{i=2}^{n} \sum_{j=1}^{i-1} X_i X_j,$$

donc $(n-1)S_n^2 = X^t BX$ avec $B = \mathbb{I}_n - \frac{1}{n} 1_{n \times n}$.

Comme $1_{n \times n} 1_{n \times n} = n 1_{n \times n}$, on vérifie que $A^2 = A$, $B^2 = B$ et $AB = BA = 0_{n,n}$. Le résultat précédent implique l'indépendance de $n\overline{X}_n^2$ et $(n-1)S_n^2$, entraînant ainsi l'indépendance de $\overline{X}_n$ et S_n.

□

Var et lois usuelles en Statistique

S'il fallait ne retenir qu'une seule chose de ce document, cela serait le résultat suivant :

Soient $X_1, \ldots, X_n$ n *var iid* suivant chacune la loi normale $\mathcal{N}(\mu, \sigma^2)$. On pose

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2.$$

Le tableau ci-dessous résume les *var* usuelles et les lois associées couramment utilisées en Statistique :

var	$\sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right)$	$\frac{n-1}{\sigma^2} S_n^2$	$\sqrt{n}\left(\frac{\overline{X}_n - \mu}{S_n}\right)$
Loi	$\mathcal{N}(0,1)$	$\chi^2(n-1)$	$\mathcal{T}(n-1)$

Preuve : Les deux premiers résultats ont déjà été prouvé, il reste à montrer le dernier. On peut écrire :

$$T_n = \sqrt{n}\left(\frac{\overline{X}_n - \mu}{S_n}\right) = \frac{Z_n}{\sqrt{\frac{K_n}{n-1}}},$$

avec $Z_n = \sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right) \sim \mathcal{N}(0,1)$ et $K_n = (n-1)\frac{S_n^2}{\sigma^2} \sim \chi^2(n-1)$. Comme $\overline{X}_n$ et S_n sont indépendantes, il en de même pour Z_n et K_n. Cela caractérise la loi de Student :

$$T_n \sim \mathcal{T}(n-1).$$

$\square$

8 Complément : loi normale multidimensionelle

Vecteur gaussien

Soient $n \in \mathbb{N}^*$ et $X_1, \ldots, X_n$ n *var*. On dit que $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ est un vecteur gaussien si et seulement si toute combinaison linéaire de $X_1, \ldots, X_n$ suit une loi normale : pour tout $(a_1, \ldots, a_n) \in \mathbb{R}^n$,

$$a_1 X_1 + \ldots + a_n X_n \sim \mathcal{N}.$$

Un vecteur gaussien est caractérisé par son espérance μ et sa matrice de covariance Σ.

La loi de X est la loi normale multidimensionnelle notée $\mathcal{N}_n(\mu, \Sigma)$.

Critère d'indépendance

Soient $n \in \mathbb{N}^*$, $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur gaussien et $(i, j) \in \{1, \ldots, n\}^2$ avec $i \neq j$. Alors X_i et X_j sont indépendantes si et seulement si $\mathbb{C}(X_i, X_j) = 0$.

Loi normale multidimensionelle

Soient, $n \in \mathbb{N}^*$, $\mu \in \mathbb{R}^n$ et Σ une matrice de dimension $n \times n$ symétrique définie positive vérifiant $\det(\Sigma) > 0$. Alors $X \sim \mathcal{N}_n(\mu, \Sigma)$ si et seulement si X possède la densité :

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(x-\mu)^t \Sigma^{-1} (x - \mu) \right), \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

Notons que, si $X \sim \mathcal{N}_n(0_n, \mathbb{I}_n)$, donc $\mu = 0_n$ et $\Sigma = \mathbb{I}_n$, alors on a

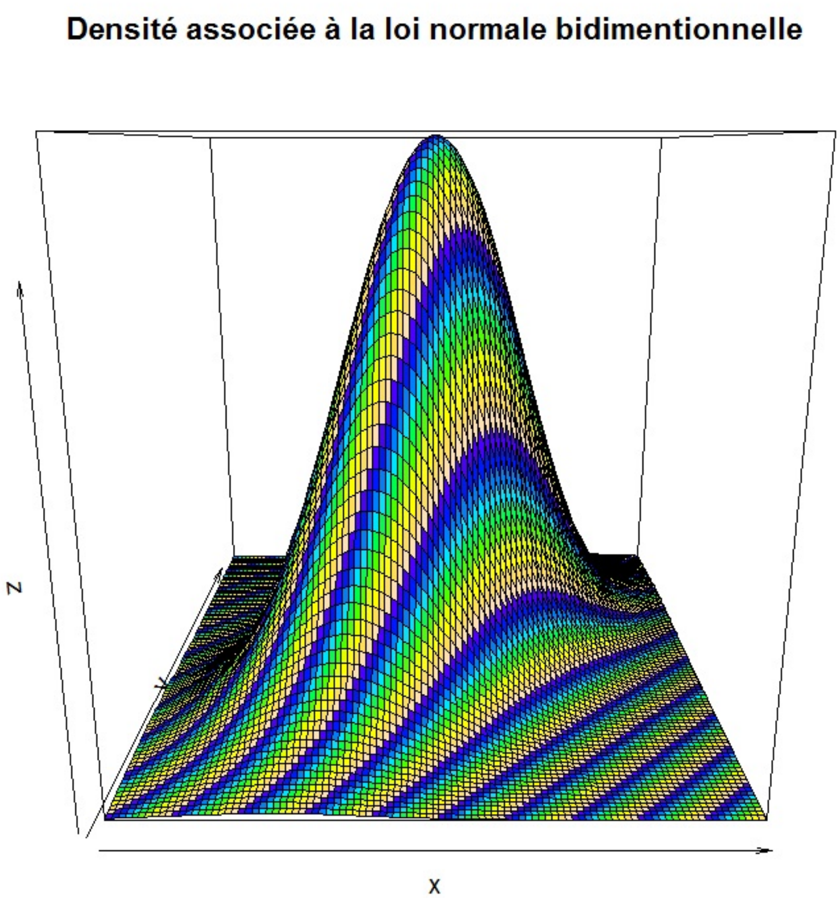
$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}x^t x} = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum\limits_{i=1}^{n} x_i^2} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}.$$

Cette forme multiplicative entraîne que :

- $X_1, \ldots, X_n$ sont indépendantes,
- $X_1, \ldots, X_n$ suivent la même loi avec $X_i \sim \mathcal{N}(0,1)$ pour tout $i \in \{1, \ldots, n\}$.

Représentation graphique

Une densité associée à la loi $\mathcal{N}_2(0_2, \Sigma)$, avec $\Sigma = \begin{pmatrix} 0.5 & 1 \\ 1 & 0.5 \end{pmatrix}$ est présentée ci-dessous :



Notations

Soient $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$, $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$ et $Z = \begin{pmatrix} Z_{1,1} & Z_{2,1} & \dots & Z_{n,1} \\ Z_{1,2} & Z_{2,2} & \dots & Z_{n,2} \\ \vdots & \vdots & & \vdots \\ Z_{1,n} & Z_{2,n} & \dots & Z_{n,n} \end{pmatrix}$ des matrices de *var.*

On adopte les notations :

- Espérance de X :

$$\mathbb{E}_n(X) = \begin{pmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_n) \end{pmatrix}.$$

- Espérance de Z :

$$\mathbb{E}_{n,n}(Z) = \begin{pmatrix} \mathbb{E}(Z_{1,1}) & \mathbb{E}(Z_{2,1}) & \dots & \mathbb{E}(Z_{n,1}) \\ \mathbb{E}(Z_{1,2}) & \mathbb{E}(Z_{2,2}) & \dots & \mathbb{E}(Z_{n,2}) \\ \vdots & \vdots & & \vdots \\ \mathbb{E}(Z_{1,n}) & \mathbb{E}(Z_{2,n}) & \dots & \mathbb{E}(Z_{n,n}) \end{pmatrix}.$$

- Covariance de X et Y :

$$\mathbb{C}_n(X,Y) = \mathbb{E}_{n,n}\left((X - \mathbb{E}_n(X))(Y - \mathbb{E}_n(Y))^t\right) = \begin{pmatrix} \mathbb{C}(X_1,Y_1) & \mathbb{C}(X_1,Y_2) & \dots & \mathbb{C}(X_1,Y_n) \\ \mathbb{C}(X_2,Y_1) & \mathbb{C}(X_2,Y_2) & \dots & \mathbb{C}(X_2,Y_n) \\ \vdots & \vdots & & \vdots \\ \mathbb{C}(X_n,Y_1) & \mathbb{C}(X_n,Y_2) & \dots & \mathbb{C}(X_n,Y_n) \end{pmatrix}.$$

- Matrice de covariance de X :

$$\mathbb{V}_n(X) = \mathbb{C}_n(X,X) = \mathbb{E}_{n,n}\left((X - \mathbb{E}_n(X))(X - \mathbb{E}_n(X))^t\right).$$

Paramètres

Si $X \sim \mathcal{N}_n(\mu, \Sigma)$, alors X est un vecteur gaussien avec

$$\mathbb{E}_n(X) = \mu, \qquad \mathbb{V}_n(X) = \Sigma.$$

Forme linéaire

Soit $X \sim \mathcal{N}_n(\mu, \Sigma)$.

- Soient A une matrice de dimension $p \times n$ avec $p \leq n$ et $a \in \mathbb{R}^n$ un vecteur colonne. Alors on a
$$AX + a \sim \mathcal{N}_n(A\mu + a, A\Sigma A^t)$$
(donc $\mathbb{E}_n(AX + a) = A\mu + a$ et $\mathbb{V}_n(AX + a) = A\Sigma A^t$).
- Soient A une matrice de dimension $p \times n$ avec $p \leq n$ et B une matrice de dimension $q \times n$ avec $q \leq n$. Alors on a
$$\mathbb{C}_n(AX, BX) = A\Sigma B^t.$$

Vecteurs gaussiens et indépendance

Soit $X \sim \mathcal{N}_n(\mu, \Sigma)$. Soient A une matrice de dimension $p \times n$ avec $p \leq n$ et B une matrice de dimension $q \times n$ avec $q \leq n$. Alors AX et BX sont indépendantes si et seulement si $A\Sigma B^t = 0_{p,q}$.

Ainsi, les composantes de tout sous vecteur de X sont indépendantes si et seulement si leurs covariances sont nulles.

Retour sur deux résultats

En utilisant certaines propriétés des vecteurs gaussiens, nous allons donner des éléments de preuves à deux résultats présentés dans la section précédente.

Soient $X_1, \ldots, X_n$ n *var* indépendantes avec $X_i \sim \mathcal{N}(\mu_i, \sigma^2)$ pour tout $i \in \{1, \ldots, n\}$ et $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$. Soit A une matrice carré symétrique A de dimension $n \times n$ telle que $A^2 = A$.

On pose $Q = X^t AX$, $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$, $\delta = \frac{1}{\sigma^2}\mu^t A\mu$ et $d = \text{rang}(A)$. Alors on a
$$\frac{Q}{\sigma^2} \sim \chi^2(d, \delta).$$

Preuve : Comme $A^2 = A$, par diagonalisation de A, il existe une matrice P de dimension $n \times d$ vérifiant : $A = PP^t$ et $P^tP = \mathbb{I}_d$. De plus, par hypothèse, on a $X \sim \mathcal{N}_n(\mu, \sigma^2\mathbb{I}_n)$. On en déduit que

$$\frac{P^tX}{\sigma} \sim \mathcal{N}_d\left(\frac{P^t\mu}{\sigma}, \Sigma\right),$$

avec $\Sigma = \frac{1}{\sigma^2}P^t\sigma^2\mathbb{I}_nP = P^tP = \mathbb{I}_d$. Par conséquent, en posant $Y_i = \frac{[P^tX]_i}{\sigma}$, la i-ème composante du vecteur $\frac{P^tX}{\sigma}$ pour tout $i \in \{1, \ldots, d\}$, alors $Y_1, \ldots, Y_n$ sont indépendantes avec $Y_i \sim \mathcal{N}\left(\frac{[P^t\mu]_i}{\sigma}, 1\right)$ et

$$\frac{Q}{\sigma^2} = \frac{X^tAX}{\sigma^2} = \frac{X^tPP^tX}{\sigma^2} = \left(\frac{P^tX}{\sigma}\right)^t\left(\frac{P^tX}{\sigma}\right) = \sum_{i=1}^{d} Y_i^2.$$

La caractérisation de la loi du Chi-deux non centrée nous assure que

$$\frac{Q}{\sigma^2} \sim \chi^2(d, \delta),$$

avec

$$\delta = \sum_{i=1}^{n}\left(\frac{[P^t\mu]_i}{\sigma}\right)^2 = \frac{1}{\sigma^2}(P^t\mu)^t(P^t\mu) = \frac{1}{\sigma^2}\mu^tPP^t\mu = \frac{1}{\sigma^2}\mu^tA\mu.$$

□

Soient $X_1, \ldots, X_n$ n *var* indépendantes avec $X_i \sim \mathcal{N}(\mu_i, \sigma^2)$ pour tout $i \in \{1, \ldots, n\}$ et $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$. Soient A et B deux matrices carrés symétriques A et B de dimension $n \times n$ telles que $A^2 = A$, $B^2 = B$ et $AB = BA = 0_{n,n}$. Alors les *var* $Q_1 = X^tAX$ et $Q_2 = X^tBX$ sont indépendantes.

Preuve : On pose $p = \mathrm{rang}(A)$ et $q = \mathrm{rang}(B)$. Comme $A^2 = A$, par diagonalisation de A, il existe une matrice P de dimension $n \times p$ vérifiant $A = PP^t$ et $P^tP = \mathbb{I}_p$. De même, comme $B^2 = B$, il existe une matrice Q de dimension $n \times q$ vérifiant $B = QQ^t$ et $Q^tQ = \mathbb{I}_q$. Remarquons que

$$Q_1 = X^tAX = X^tPP^tX = (P^tX)^tP^tX, \qquad Q_2 = X^tBX = X^tQQ^tX = (Q^tX)^tQ^tX.$$

Ainsi, Q_1 est fonction de P^tX et Q_2 est fonction de Q^tX. Pour conclure à l'indépendance de Q_1 et Q_2, il suffit de montrer que P^tX et Q^tX sont indépendantes.

Comme $X \sim \mathcal{N}_n(\mu, \sigma^2 \mathbb{I}_n)$, avec $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$, on a $P^t X$ et $Q^t X$ indépendantes si et seulement si

$$P^t \sigma^2 \mathbb{I}_n Q = 0_{p,q} \Leftrightarrow P^t Q = 0_{p,q}.$$

La condition $AB = 0_{n,n}$ entraîne :

$$PP^t QQ^t = 0_{n,n} \Rightarrow P^t PP^t QQ^t Q = P^t 0_{n,n} Q = 0_{p,q} \Rightarrow \mathbb{I}_p P^t Q \mathbb{I}_q = 0_{p,q} \Rightarrow P^t Q = 0_{p,q}.$$

On en déduit que $P^t X$ et $Q^t X$ sont indépendantes, entraînant ainsi l'indépendance de Q_1 et Q_2.

$\square$

9 Annexe : reproduction des graphiques

Les graphiques de ce document ont été obtenus avec le logiciel R :

`https://cran.r-project.org/`

Cette section propose les commandes R permettant leur reproduction.

Loi normale

```
curve(dnorm(x, 0, 1), xlim = c(-10, 10), ylim = c(0, 0.58), col = "red", lwd = 3,
ylab = "", main = "Densités associées à la loi normale")
curve(dnorm(x, 2, 3), col = "green", lwd = 3, add = TRUE)
curve(dnorm(x, 0, 5), col = "blue", lwd = 3, add = TRUE)
curve(dnorm(x, -2, 0.7), col = "yellow", lwd = 3, add = TRUE)
abline(h = 0, lty = 2)
abline(v = 0, lty = 2)
legend("topright",
legend = c(
expression(paste(mu, "=0, ", sigma, "=1")),
expression(paste(mu, "=2, ", sigma, "=3")),
expression(paste(mu, "=0, ", sigma, "=5")),
expression(paste(mu, "=-2, ", sigma, "=0.7"))),
 lwd = 3, col = c("red", "green", "blue", "yellow"))
```

Loi du Chi-deux

```
curve(dchisq(x, 1), xlim = c(0, 10), ylim = c(0, 0.58), col = "red", lwd = 3,
ylab = "", main = "Densités associées à la loi du Chi-deux")
curve(dchisq(x, 2), col = "green", lwd = 3, add = TRUE)
curve(dchisq(x, 4), col = "blue", lwd = 3, add = TRUE)
curve(dchisq(x, 6), col = "yellow", lwd = 3, add = TRUE)
abline(h = 0, lty = 2)
abline(v = 0, lty = 2)
legend("topright",
legend = c(
expression(paste(nu, "=1")),
expression(paste(nu, "=2")),
expression(paste(nu, "=4")),
expression(paste(nu, "=6"))),
 lwd = 3, col = c("red", "green", "blue", "yellow"))
```

Loi de Student

```
curve(dt(x, 1), xlim = c(-6, 6), ylim = c(0, 0.38), col = "red", lwd = 3,
ylab = "", main = "Densités associées à la loi de Student")
curve(dt(x, 2), col = "green", lwd = 3, add = TRUE)
curve(dt(x, 5), col = "blue", lwd = 3, add = TRUE)
curve(dt(x, 16), col = "yellow", lwd = 3, add = TRUE)
abline(h = 0, lty = 2)
abline(v = 0, lty = 2)
legend("topright",
legend = c(
expression(paste(nu, "=1")),
expression(paste(nu, "=2")),
expression(paste(nu, "=5")),
expression(paste(nu, "=16"))),
 lwd = 3, col = c("red", "green", "blue", "yellow"))
```

Loi de Fisher

```
curve(df(x, 1, 1), xlim = c(0, 3), ylim = c(0, 1), col = "red", lwd = 3,
ylab = "", main = "Densités associées à la loi de Fisher")
curve(df(x, 3, 1), col = "green", lwd = 3, add = TRUE)
curve(df(x, 1, 3), col = "blue", lwd = 3, add = TRUE)
curve(df(x, 5, 5), col = "yellow", lwd = 3, add = TRUE)
curve(df(x, 7, 15), col = "orange", lwd = 3, add = TRUE)
abline(h = 0, lty = 2)
abline(v = 0, lty = 2)
legend("topright",
legend = c(
expression(paste( "(", nu[1] , ",", nu[2], ")", " = (1, 1)")),
expression(paste( "(", nu[1] , ",", nu[2], ")", " = (3, 1)")),
expression(paste( "(", nu[1] , ",", nu[2], ")", " = (1, 3)")),
expression(paste( "(", nu[1] , ",", nu[2], ")", " = (5, 5)")),
expression(paste( "(", nu[1] , ",", nu[2], ")", " = (7, 15)"))),
 lwd = 3, col = c("red", "green", "blue", "yellow", "orange"))
```

Loi du Chi-deux non centrée

```
curve(dchisq(x, 3, 1), xlim = c(0, 16), ylim = c(0, 0.21), col = "red", lwd = 3,
ylab = "", main = "Densités associées à la loi du Chi-deux non centrée")
curve(dchisq(x, 3, 2), col = "green", lwd = 3, add = TRUE)
curve(dchisq(x, 3, 4), col = "blue", lwd = 3, add = TRUE)
curve(dchisq(x, 3, 7), col = "yellow", lwd = 3, add = TRUE)
abline(h = 0, lty = 2)
abline(v = 0, lty = 2)
legend("topright",
legend = c(
expression(paste(n, "=3, ", delta, "=1")),
expression(paste(n, "=3, ", delta, "=2")),
expression(paste(n, "=3, ", delta, "=4")),
expression(paste(n, "=3, ", delta, "=7"))),
 lwd = 3, col = c("red", "green", "blue", "yellow"))
```

Loi de Fisher non centrée

```
curve(df(x, 3, 2, 1), xlim = c(0, 8), ylim = c(0, 0.5), col = "red", lwd = 3,
ylab = "", main = "Densités associées à la loi de Fisher non centrée")
curve(df(x, 3, 2, 2), col = "green", lwd = 3, add = TRUE)
curve(df(x, 3, 2, 4), col = "blue", lwd = 3, add = TRUE)
curve(df(x, 3, 2, 10), col = "yellow", lwd = 3, add = TRUE)
abline(h = 0, lty = 2)
abline(v = 0, lty = 2)
legend("topright",
legend = c(
expression(paste( "(", nu[1] , ",", nu[2], ")", " = (3, 2), " , delta, "=1")),
expression(paste( "(", nu[1] , ",", nu[2], ")", " = (3, 2), ", delta, "=2")),
expression(paste( "(", nu[1] , ",", nu[2], ")", " = (3, 2), ", delta, "=4")),
expression(paste( "(", nu[1] , ",", nu[2], ")", " = (3, 2), ", delta, "=10"))),
 lwd = 3, col = c("red", "green", "blue", "yellow"))
```

Loi normale multidimensionnelle

```
library(mvtnorm)
N = 100
Mean = c(0, 0)
Sigma = matrix(c(1, 0.5, 0.5, 1), 2, 2)
x = seq(-3, 3, length = N)
y = seq(-3, 3, length = N)
z = matrix(0, N, N)
for (i in 1:N) for (j in 1:N) {
z[i, j] = dmvnorm(c(x[i], y[j]), Mean, Sigma)
}
persp(x, y, z, main = "Densité associée à la loi normale bidimentionnelle",
col= topo.colors(10))
```