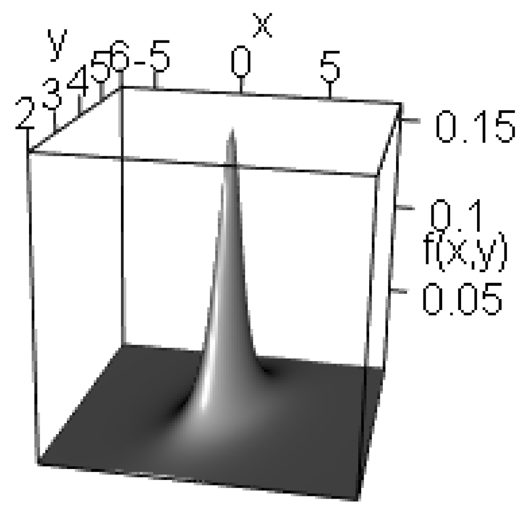
# Distance/divergence entre distributions $t$ multivariées

**Pierre Santagostini, Nizar Bouhlel**

**Institut Agro, Univ Angers, INRAE, IRHS, SFR QUASAV, F-49000 Angers**




## Package `mstudentd`

Un package pour la distribution $t$ multivariée.

- Outils pour la manipulation de distributions $t$ multivariées : fonction densité de probabilité, simulation d'échantillon, représentation graphique (2 variables)
- Mesure de divergence entre distribution $t$ multivariée : Rényi, Bhattacharyya, Hellinger, Kullback-Leibler

## Distribution $t$ multivariée

**Densité de probabilité** d'un vecteur aléatoire, de loi $t$ multivariée (dimension $p$) à $\nu$ degrés de liberté, de vecteur moyenne $\boldsymbol{\mu}$ et de matrice de corrélation $\Sigma$ :

$$f(\boldsymbol{x}|\nu,\boldsymbol{\mu},\Sigma) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)|\Sigma|^{-1/2}}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{p/2}}\left(1+\frac{1}{\nu}\,(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)^{-\frac{\nu+p}{2}}$$

Le cas particulier $\nu = 1$ correspond à la loi de Cauchy multivariée.

Applications nombreuses en analyse d'image, signal, autres domaines : reconnaissance de forme, apprentissage automatique, détection de changements, sélection de modèles, recherche d'image par le contenu…

## Divergence entre distributions $t$ multivariées

### Divergence entre deux vecteurs aléatoires de dimension $p$, de lois $t$ multivariées centrées de paramètres $(\nu_1, \boldsymbol{0}, \Sigma_1)$ et $(\nu_2, \boldsymbol{0}, \Sigma_2)$ :

Soient $\delta_1 = \frac{\nu_1+p}{2}\beta$, $\delta_2 = \frac{\nu_2+p}{2}(1-\beta)$ et $\lambda_1 < \cdots < \lambda_p$ les valeurs propres de la matrice carrée $\Sigma_1\,\Sigma_2^{-1}$,

- Divergence de Rényi (Bouhlel & Rousseau) :

$$D_R^\beta(\boldsymbol{X}_1\|\boldsymbol{X}_2) = \frac{1}{\beta-1}\left[\beta\ln\left(\frac{\Gamma\left(\frac{\nu_1+p}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)\nu_2^{p/2}}{\Gamma\left(\frac{\nu_2+p}{2}\right)\Gamma\left(\frac{\nu_1}{2}\right)\nu_1^{p/2}}\right)+\ln\left(\frac{\Gamma\left(\frac{\nu_2+p}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)+\ln\left(\frac{\Gamma\left(\delta_1+\delta_2-\frac{p}{2}\right)}{\Gamma(\delta_1+\delta_2)}\right)-\frac{\beta}{2}\sum_{i=1}^{p}\ln\,\lambda_i+\ln F_D\right]$$

où $F_D$ est donné par :

- $F_D = F_D^{(p)}\left(\delta_1,\underbrace{\frac{1}{2},\ldots,\frac{1}{2}}_{p};\delta_1+\delta_2;1-\frac{\nu_2}{\nu_1}\frac{1}{\lambda_1},\ldots,1-\frac{\nu_2}{\nu_1}\frac{1}{\lambda_p}\right)$, si $\frac{\nu_1}{\nu_2}\lambda_i > 1,\ \forall i=1,\ldots,p$
- $F_D = \prod_{i=1}^{p}\left(\frac{\nu_1}{\nu_2}\lambda_i\right)^{1/2}F_D^{(p)}\left(\delta_2,\underbrace{\frac{1}{2},\ldots,\frac{1}{2}}_{p};\delta_1+\delta_2;1-\frac{\nu_1}{\nu_2}\lambda_1,\ldots,1-\frac{\nu_1}{\nu_2}\lambda_p\right)$, si $\frac{\nu_1}{\nu_2}\lambda_i < 1, \forall i=1,\ldots,p$
- $F_D = \left(\frac{\nu_2}{\nu_1}\frac{1}{\lambda_p}\right)^{\delta_2}\prod_{i=1}^{p}\left(\frac{\nu_1}{\nu_2}\lambda_i\right)^{1/2}F_D^{(p)}\left(\delta_2,\underbrace{\frac{1}{2},\ldots,\frac{1}{2},\delta_1+\delta_2-\frac{p}{2}}_{p};\delta_1+\delta_2;1-\frac{\lambda_1}{\lambda_p},\ldots,1-\frac{\lambda_{p-1}}{\lambda_p},1-\frac{\nu_2}{\nu_1}\frac{1}{\lambda_p}\right)$ sinon

$F_D^{(p)}$ étant la fonction $D$-hypergéométrique de Lauricella :

$$F_D^{(n)}(a,b_1,\ldots,b_n;c;x_1,\ldots x_n) = \sum_{m_1\geq 0}\ldots\sum_{m_n\geq 0}\frac{(a)_{m_1+\cdots+m\_n}(b_1)_{m_1}\ldots(b_n)_{m_n}}{(c)_{m_1+\cdots+m_n}}\frac{x_1^{m_1}}{m_1!}\ldots\frac{x_n^{m_n}}{m_n!}$$

et $(x)_m = \frac{\Gamma(x+m)}{\Gamma(x)} = x(x+1)\ldots(x+m-1)$est le symbole de Pochhammer.

- Divergence de Bhattacharyya :

$$D_B(\boldsymbol{X}_1\|\boldsymbol{X}_2) = \frac{1}{2}D_R^{1/2}(\boldsymbol{X}_1\|\boldsymbol{X}_2)$$

- Divergence de Hellinger :

$$D_H(\boldsymbol{X}_1\|\boldsymbol{X}_2) = 1-\exp\left(-\frac{1}{2}\,D_R^{1/2}(\boldsymbol{X}_1\|\boldsymbol{X}_2)\right)$$

- Divergence de Kullback-Leibler (Bouhlel & Rousseau) :

$$D_{KL}(\boldsymbol{X}_1\|\boldsymbol{X}_2) = \ln\left(\frac{\Gamma\left(\frac{\nu_1+p}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)\nu_2^{p/2}}{\Gamma\left(\frac{\nu_2+p}{2}\right)\Gamma\left(\frac{\nu_1}{2}\right)\nu_1^{p/2}}\right)+\frac{\nu_2-\nu_1}{2}\left(\psi\left(\frac{\nu_1+p}{2}\right)-\psi\left(\frac{\nu_1}{2}\right)\right)-\frac{1}{2}\sum_{i=1}^{p}\ln\lambda_i-\frac{\nu_2+p}{2}\,D$$

où $D$ est donné par :

- $D = \frac{\partial}{\partial a}\left\{F_D^{(p)}\left(a,\frac{1}{2},\ldots,\frac{1}{2};a+\frac{\nu_1+p}{2};1-\frac{\nu_1}{\nu_2}\lambda_1,\ldots,1-\frac{\nu_1}{\nu_2}\lambda_p\right)\right\}\Big|_{a=0}$, si $\frac{\nu_1}{\nu_2}\lambda_i < 1,\ \forall i=1,\ldots,p$
- $D = \prod_{i=1}^{p}\left(\frac{\nu_2}{\nu_1}\frac{1}{\lambda_i}\right)^{1/2}\frac{\partial}{\partial a}\left\{F_D^{(p)}\left(\frac{\nu_1+p}{2},\frac{1}{2},\ldots,\frac{1}{2};a+\frac{\nu_1+p}{2};1-\frac{\nu_2}{\nu_1}\frac{1}{\lambda_1},\ldots,1-\frac{\nu_2}{\nu_1}\frac{1}{\lambda_p}\right)\right\}\Big|_{a=0}$, si $\frac{\nu_1}{\nu_2}\lambda_i > 1, \forall i$
- $D = -\ln\left(\frac{\nu_1}{\nu_2}\lambda_p\right)+\frac{\partial}{\partial a}\left\{F_D^{(p)}\left(a,\frac{1}{2},\ldots,\frac{1}{2},a+\frac{\nu_1}{2};a+\frac{\nu_1+p}{2};1-\frac{\lambda_1}{\lambda_p},\ldots,1-\frac{\lambda_{p-1}}{\lambda_p},1-\frac{\nu_2}{\nu_1}\frac{1}{\lambda_p}\right)\right\}\Big|_{a=0}$ sinon

et $\psi(.)$ est la fonction digamma : $\psi(x) = \frac{d}{dx}\Gamma(x)$

## Exemple

```
nu1 <- 2
Sigma1 <- rbind(c(2, 1.2, 0.4),
                c(1.2, 2, 0.6),
                c(0.4, 0.6, 2)
)
Sigma1
     [,1] [,2] [,3]
[1,]  2.0  1.2  0.4
[2,]  1.2  2.0  0.6
[3,]  0.4  0.6  2.0
nu2 <- 4
Sigma2 <- rbind(c(1, 0.3, 0.1),
                c(0.3, 1, 0.4),
                c(0.1, 0.4, 1)
)
Sigma2
     [,1] [,2] [,3]
[1,]  1.0  0.3  0.1
[2,]  0.3  1.0  0.4
[3,]  0.1  0.4  1.0
```

### Divergence de Rényi : $D_R^\beta(\boldsymbol{X}_1\|\boldsymbol{X}_2)$

```
diststudent(nu1, Sigma1, nu2, Sigma2, dist = "renyi", bet = 0.5, eps = 1e-6)
[1] 0.1541504
attr(,"epsilon")
[1] 1e-06
attr(,"k")
[1] 20
```

### Divergence de Bhattacharyya : $D_B(\boldsymbol{X}_1\|\boldsymbol{X}_2)$

```
diststudent(nu1, Sigma1, nu2, Sigma2, dist = "bhattacharyya", eps = 1e-6)
[1] 0.07707521
attr(,"epsilon")
[1] 1e-06
attr(,"k")
[1] 20
```

### Divergence de Hellinger : $D_H(\boldsymbol{X}_1\|\boldsymbol{X}_2)$

```
diststudent(nu1, Sigma1, nu2, Sigma2, dist = "hellinger", eps = 1e-6)
[1] 0.07417978
attr(,"epsilon")
[1] 1e-06
attr(,"k")
[1] 20
```

### Divergence de Kullback-Leibler : $D_{KL}(\boldsymbol{X}_1\|\boldsymbol{X}_2)$

```
kldstudent(nu1, Sigma1, nu2, Sigma2, eps = 1e-6)
[1] 0.5796759
attr(,"epsilon")
[1] 1e-06
attr(,"k")
[1] 15
```

Arguments :

- `nu1`, `Sigma1` : paramètres de la première distribution.
- `nu2`, `Sigma2` : paramètres de la deuxième distribution.
- `dist` : la mesure de divergence qu'on veut calculer.
- `eps` : précision souhaitée pour le calcul de la fonction de Lauricella (Rényi, Bhattacharyya, Hellinger) ou de sa dérivée partielle (Kullback-Leibler).

Résultat :

La valeur de la divergence,

avec les attributs :

- `attr(, "epsilon")` : précision atteinte pour la fonction de Lauricella (ou sa dérivée partielle).
- `attr(, "k")` : le nombre d'éléments de la somme qui ont été calculés pour atteindre cette précision.

## Bibliographie

[1] Bouhlel, N., Rousseau, D., 2023. Exact Rényi and Kullback-Leibler Divergences Between Multivariate $t$-Distributions. IEEE Signal Processing Letter. https://doi.org/10.1109/LSP.2023.3324594

[2] Santagostini, P., Bouhlel, N., 2024. mstudentd: Multivariate t Distribution. https://CRAN.Rproject.org/package=mstudentd

## Perspectives

- Ajout au package mstudentd de l'estimation des paramètres d'une loi $t$ multivariée.
- Création d'un package pour des distributions multivariées, qui regroupera les outils actuellement proposés par les packages **mggd** (loi normale généralisée multivariée), **mcauchyd** (loi de Cauchy multivariée) et **mstudentd**, avec extension au cas multivarié complexe, et aux lois elliptiques en général.