

# Distance/Divergence entre distributions $t$ multivariées

Pierre Santagostini 1\*

Nizar Bouhlel 2†

## Résumé

Différentes mesures peuvent être utilisées pour le calcul de la distance entre deux distributions de probabilité. Parmi celles-ci, on trouve la divergence de Rényi, de Bhattacharyya, de Hellinger ou de Kullback Leibler. Les applications sont nombreuses en image, en signal et dans d'autres domaines connexes en informatique telles que la reconnaissance de forme, l'apprentissage automatique, la détection de changements, la sélection de modèles, la recherche d'image par le contenu, etc. Le package **mstudentd** fournit des fonctions pour le calcul de ces divergences entre lois  $t$  multivariées (lois de Student multivariées) en utilisant des expressions analytiques récemment introduites (Bouhlel and Rousseau 2023).

**Mots-clefs** : Package - Divergence de Rényi - Divergence de Kullback-Leibler - Distribution  $t$  multivariée

## Package **mstudentd** : divergence entre deux distributions $t$ multivariées centrées

Le package **mstudentd** (Santagostini and Bouhlel 2024) fournit des fonctions pour le calcul de la divergence entre une distribution  $t$  multivariée centrée à **nu1** degrés de liberté, de matrice de corrélation **Sigma1** et une distribution à **nu2** degrés de liberté, de matrice de corrélation **Sigma2** :

- Divergence de Rényi d'ordre  $\beta = 0.25$  :

```
diststudent(nu1, Sigma1, nu2, Sigma2, dist = "renyi", bet = 0.25)
```

- Divergence de Bhattacharyya :

```
diststudent(nu1, Sigma1, nu2, Sigma2, dist = "bhattacharyya")
```

- Divergence de Hellinger :

```
diststudent(nu1, Sigma1, nu2, Sigma2, dist = "hellinger")
```

- Divergence de Kullback-Leibler :

```
kldstudent(nu1, Sigma1, nu2, Sigma2)
```

Des expressions analytiques des divergences de Rényi et Kullback-Leibler, exposées ci-après, ont récemment été introduites par Bouhlel and Rousseau (2023). Les calculs implémentés dans ces fonctions **diststudent** et **kldstudent** sont ainsi basés sur des expressions exactes des divergences. Cela permet un calcul plus précis, comparé à un calcul approché par la méthode de Monte Carlo, ou à un calcul de la divergence entre des distributions de probabilité discrétisées (différents packages R fournissent des fonctions pour la divergence entre des distributions discrètes, par exemple **rd** du package **divo** ou **kld** du package **dlookr**).

## Calcul de la divergence entre distributions $t$ multivariées

Soient  $\mathbf{X}_1$  et  $\mathbf{X}_2$  deux vecteurs aléatoires distribués selon des lois  $t$  multivariées centrées (à  $p$  variables) à  $\nu_i$  degrés de liberté, de vecteurs moyennes  $\boldsymbol{\mu}_i = \mathbf{0}$  et de matrices de corrélation  $\Sigma_i$ , de densités de probabilité :

$$f(\mathbf{x}|\nu_i, \boldsymbol{\mu}_i = \mathbf{0}, \Sigma_i) = f(\mathbf{x}|\nu_i, \Sigma_i) = \frac{\Gamma\left(\frac{\nu_i+p}{2}\right) |\Sigma_i|^{-1/2}}{\Gamma\left(\frac{\nu_i}{2}\right) (\nu_i\pi)^{p/2}} \left(1 + \frac{1}{\nu_i} \mathbf{x}^T \Sigma_i^{-1} \mathbf{x}\right)^{-\frac{\nu_i+p}{2}}, \quad i = 1, 2 \quad (1)$$

\*Institut Agro, Univ Angers, INRAE, IRHS, SFR QUASAV, F-49000 Angers, France, pierre.santagostini@agrocampus-ouest.fr

†Institut Agro, Univ Angers, INRAE, IRHS, SFR QUASAV, F-49000 Angers, France, nizar.bouhlel@agrocampus-ouest.fr

## Divergence de Rényi

Une expression de la divergence de Rényi d'ordre  $\beta$  ( $\beta > 0$ ,  $\beta \neq 1$ ) entre ces deux distributions est donnée par Bouhlel and Rousseau (2023) :

$$D_R^\beta(\mathbf{X}_1 \parallel \mathbf{X}_2) = \frac{1}{\beta - 1} \left[ \beta \ln \left( \frac{\Gamma\left(\frac{\nu_1 + p}{2}\right) \Gamma\left(\frac{\nu_2}{2}\right) \nu_2^{\frac{p}{2}}}{\Gamma\left(\frac{\nu_2 + p}{2}\right) \Gamma\left(\frac{\nu_1}{2}\right) \nu_1^{\frac{p}{2}}} \right) + \ln \left( \frac{\Gamma\left(\frac{\nu_2 + p}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)} \right) + \ln \left( \frac{\Gamma(\delta_1 + \delta_2 - \frac{p}{2})}{\Gamma(\delta_1 + \delta_2)} \right) \right. \\ \left. - \frac{\beta}{2} \sum_{i=1}^p \ln \lambda_i + \ln F_D^{(p)} \left( \delta_1, \underbrace{\frac{1}{2}, \dots, \frac{1}{2}}_p; \delta_1 + \delta_2; 1 - \frac{\nu_2}{\nu_1} \frac{1}{\lambda_1}, \dots, 1 - \frac{\nu_2}{\nu_1} \frac{1}{\lambda_p} \right) \right] \quad (2)$$

où  $\delta_1 = \frac{\nu_1 + p}{2} \beta$ ,  $\delta_2 = \frac{\nu_2 + p}{2} (1 - \beta)$ ,  $\lambda_1 < \dots < \lambda_{p-1} < \lambda_p$  sont les valeurs propres de la matrice carrée  $\Sigma_1 \Sigma_2^{-1}$  et  $F_D^{(p)}$  est la fonction de Lauricella  $D$ -hypergéométrique :

$$F_D^{(p)}(a; b_1, \dots, b_p; g; x_1, \dots, x_p) = \sum_{m_1 \geq 0} \dots \sum_{m_p \geq 0} \frac{(a)_{m_1 + \dots + m_p} (b_1)_{m_1} \dots (b_p)_{m_p}}{(g)_{m_1 + \dots + m_p}} \frac{x_1^{m_1}}{m_1!} \dots \frac{x_p^{m_p}}{m_p!}$$

$(q)_i$  étant le symbole de Pochhammer :  $(q)_i = \frac{\Gamma(q+i)}{\Gamma(q)}$  ( $i = 1, 2, \dots$ ).

La condition  $\left| 1 - \frac{\nu_2}{\nu_1} \frac{1}{\lambda_i} \right| < 1$ ,  $i = 1, \dots, p$  assure la convergence de la série. Quand cette condition n'est pas vérifiée, des transformations de la fonction de Lauricella permettent d'obtenir la convergence.

## Divergence de Bhattacharyya, divergence de Hellinger

Les divergences de Bhattacharyya et de Hellinger sont calculées à partir de la divergence de Rényi :

- Divergence de Bhattacharyya entre  $\mathbf{X}_1$  et  $\mathbf{X}_2$  :  $D_B(\mathbf{X}_1 \parallel \mathbf{X}_2) = \frac{1}{2} D_R^{1/2}(\mathbf{X}_1 \parallel \mathbf{X}_2)$
- Divergence de Hellinger :  $D_H(\mathbf{X}_1 \parallel \mathbf{X}_2) = 1 - \exp\left(-\frac{1}{2} D_R^{1/2}(\mathbf{X}_1 \parallel \mathbf{X}_2)\right)$

## Divergence de Kullback-Leibler

La divergence de Kullback-Leibler entre  $\mathbf{X}_1$  et  $\mathbf{X}_2$  est donnée par :  $D_{KL}(\mathbf{X}_1 \parallel \mathbf{X}_2) = \lim_{\beta \rightarrow 1} D_R^\beta(\mathbf{X}_1 \parallel \mathbf{X}_2)$ . Soit  $\psi$  la fonction digamma :  $\psi(x) = \frac{d}{dx} \ln \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ . D'après (Bouhlel and Rousseau 2023) :

$$D_{KL}(\mathbf{X}_1 \parallel \mathbf{X}_2) = \ln \left( \frac{\Gamma\left(\frac{\nu_1 + p}{2}\right) \Gamma\left(\frac{\nu_2}{2}\right) \nu_2^{\frac{p}{2}}}{\Gamma\left(\frac{\nu_2 + p}{2}\right) \Gamma\left(\frac{\nu_1}{2}\right) \nu_1^{\frac{p}{2}}} \right) + \frac{\nu_2 - \nu_1}{2} \left[ \psi\left(\frac{\nu_1 + p}{2}\right) - \psi\left(\frac{\nu_1}{2}\right) \right] - \frac{1}{2} \sum_{i=1}^p \ln \lambda_i \\ - \frac{\nu_2 + p}{2} \times \frac{\partial}{\partial a} \left\{ F_D^{(p)} \left( a, \underbrace{\frac{1}{2}, \dots, \frac{1}{2}}_p; a + \frac{\nu_1 + p}{2}; 1 - \frac{\nu_1}{\nu_2} \lambda_1, \dots, 1 - \frac{\nu_1}{\nu_2} \lambda_p \right) \right\} \Big|_{a=0} \quad (3)$$

La convergence est assurée quand  $|1 - \frac{\nu_1}{\nu_2} \lambda_i| < 1$ ,  $i = 1, \dots, p$ .

## Références

- Bouhlel, N., and D. Rousseau. 2023. "Exact Rényi and Kullback-Leibler Divergences Between Multivariate  $t$ -Distributions." *IEEE Signal Processing Letter*. <https://doi.org/10.1109/LSP.2023.3324594>.
- Santagostini, P., and N. Bouhlel. 2024. *mstudentd: Multivariate t Distribution*. <https://CRAN.R-project.org/package=mstudentd>.