

B. LECOUTRE

H. ROUANET

Deux structures statistiques fondamentales en analyse de la variance univariée et multivariée

Mathématiques et sciences humaines, tome 75 (1981), p. 71-82

http://www.numdam.org/item?id=MSH_1981__75__71_0

© Centre d'analyse et de mathématiques sociales de l'EHESS, 1981, tous droits réservés.

L'accès aux archives de la revue « Mathématiques et sciences humaines » (<http://msh.revues.org/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

DEUX STRUCTURES STATISTIQUES FONDAMENTALES EN ANALYSE DE LA VARIANCE UNIVARIEE ET MULTIVARIEE*

B. LECOUTRE**, H. ROUANET***

1. INTRODUCTION

Cet article sera consacré à deux structures statistiques que l'on rencontre constamment en analyse de la variance, univariée et multivariée : la structure "multinormale - khi-deux" et la structure "multinormale - Wishart".

Nous avons précédemment distingué (Lecoutre [4]) le problème de l'étude de l'effet associé à un contraste et le problème de l'étude de l'effet associé à une comparaison à un nombre quelconque de degrés de liberté. En analyse univariée, le premier problème peut être ramené à la structure "normale - khi-deux". Les deux structures développées ici constituent des généralisations de cette structure "normale - khi-deux" :

- . la structure "multinormale - khi-deux" intervient dans les problèmes de comparaison à plusieurs degrés de liberté en analyse de la variance univariée (voir Lecoutre [3] et [4]) ;
- . la structure "multinormale - Wishart" intervient dans les problèmes d'inférence sur un contraste en analyse de la variance multivariée (ce dernier point constitue un prolongement des méthodes développées en [3] et [4]).

Pour chacune de ces deux structures, nous envisagerons successivement les tests de signification et les solutions bayésiennes.

* Cette recherche a bénéficié de l'aide financière de l'ATP 3447, "Analyse des comparaisons pour protocoles multidimensionnels".

**Laboratoire de Psychologie (ERA 235), Université de PARIS-VIII et Groupe Mathématiques et Psychologie, Université René Descartes

***Groupe Mathématiques et Psychologie, Université René Descartes

2. STRUCTURE STATISTIQUE "MULTINORMALE - KHI-DEUX"

Nous appellerons structure statistique "multinormale - khi-deux", ou encore "m-normale - khi-deux", la donnée de deux statistiques, d et s^2 , qui, sous un certain modèle d'échantillonnage, sont distribuées indépendamment, respectivement

$$N_m(\delta, b\sigma^2 I_m) \quad \text{et} \quad \sigma^2 \frac{\chi_q^2}{q}$$

ce que nous noterons :

$$\begin{aligned} d \| s^2 / \delta, \sigma^2 \\ d/\delta, \sigma^2 \sim N_m(\delta, b\sigma^2 I_m) \quad \text{bCR}_+ \quad \sigma > 0 \\ s^2/\delta, \sigma^2 \sim \sigma^2 \frac{\chi_q^2}{q} \end{aligned}$$

La structure "m-normale - khi-deux" fait donc intervenir le paramètre multinumérique δ et le paramètre numérique σ^2 . Pour l'inférence, nous ferons jouer au paramètre $\lambda^2 = \delta'\delta$ le rôle de *paramètre numérique principal* si $m > 2$ (si $m = 1$ le paramètre numérique principal est bien entendu δ), σ^2 jouant le rôle de paramètre numérique secondaire. En particulier, l'hypothèse nulle $\delta = 0$ est équivalente à $\lambda^2 = 0$. Nous définirons en conséquence la statistique numérique $\ell^2 = d'd$.

Le test de signification sera calqué sur les *tests F* classiques en analyse de la variance (ou encore sur les tests dits du *t de Student* dans le cas particulier $m = 1$). Les procédures bayésiennes, en revanche, mettront en évidence de nouvelles distributions, qui, à notre connaissance, n'ont guère été étudiées à l'heure actuelle.

2.1. Test de signification de l'hypothèse nulle $\lambda^2 = 0$

On envisagera la statistique de test $F = \frac{\ell^2}{mbs^2}$, dont la distribution

d'échantillonnage est celle d'un F de Fisher-Snedecor à m et q degrés, non-centré en général (d'excentricité $\frac{\lambda^2}{b\sigma^2}$), centré lorsque $\lambda^2 = 0$; ce que nous écrirons :

$$\text{sous } H_0 : \lambda^2 = 0, \quad F \sim F_{m,q}$$

(F centré usuel à m et q degrés de liberté)

Dans le cas particulier $m = 1$ (structure "normale - khi-deux"), on envisagera également la statistique de test $t = \frac{d}{\sqrt{bs}}$, dont la distribution d'échantillonnage est celle d'un t de Student à q degrés de liberté, non-centré en général (d'excentricité $\frac{\delta}{\sqrt{b\sigma}}$), centré lorsque $\delta = 0$; ce que nous écrirons :

$$\text{si } m = 1, \text{ sous } H_0 : \delta = 0, \quad t \sim t_q \quad (t \text{ usuel à } q \text{ degrés de liberté})$$

2.2. Procédures bayésiennes

On se donnera pour (δ, σ^2) une distribution initiale, et on en déduira, à partir de la distribution d'échantillonnage conjointe $(d, s^2)/\delta, \sigma^2$, la distribution finale donnée par le théorème de Bayes.

2.2.1. Cas de distributions initiales marginales indépendantes

Dans le cas particulier où les distributions initiales (marginales) relatives à chacun des paramètres δ et σ^2 sont indépendantes, on pourra, en utilisant l'indépendance des distributions d'échantillonnage, décomposer l'inférence :

. d'une part en une inférence relative à la structure "multinormale"

$$d/\delta, [\sigma^2, s^2] \sim N_m(\delta, b\sigma^2 \mathbf{I}_m)$$

. d'autre part en une inférence relative à la structure "khi-deux"

$$s^2/\sigma^2, [d] \sim \sigma^2 \frac{\chi_q^2}{q}$$

On en déduira les distributions finales relatives, d'une part à δ étant donné σ^2 , et d'autre part à σ^2 , ce qui déterminera la distribution finale conjointe.

2.2.2. Distributions fiducio-bayésiennes

Comme distribution initiale on peut prendre pour δ la distribution (conjuguée) $N_m(d_o, b_o \mathbf{U}_o)$ ($|\mathbf{U}_o| > 0$) et pour σ^{-2} la distribution (conjuguée)

$s_o^{-2} \frac{\chi_{q_o}^2}{q_o}$ avec $\delta \perp \sigma^{-2}$; on obtient dans ce cas la distribution finale,

caractérisée par les propriétés suivantes :

$$\delta/\sigma^2, d, s^2 \sim N_m(d_1, \frac{b_o b}{b_o + b} \mathbf{U}_1)$$

$$\sigma^{-2}/d, s^2 \sim s_1^{-2} \frac{\chi_{q_o + q}^2}{q_o + q}$$

où $d_1 = (b_o \Sigma^{-1} + b \mathbf{U}_o^{-1})^{-1} (b_o \Sigma^{-1} d_o + b \mathbf{U}_o^{-1} d_o)$

$$\mathbf{U}_1 = \left(\frac{b_o \Sigma^{-1} + b \mathbf{U}_o^{-1}}{b_o + b} \right)^{-1}$$

avec ici $\Sigma = \sigma^2 \mathbf{I}_m$

$$s_1^2 = \frac{q_o s_o^2 + q s^2}{q_o + q}$$

Quand $b_o \rightarrow +\infty$ et $q_o \rightarrow 0$, ce qui revient à exprimer un état initial d'ignorance, on obtient, en passant à la limite, la distribution finale, caractérisée par les propriétés suivantes :

$$\delta^* / \sigma^2, d, s^2 \sim N_m(d, b\sigma^2 \mathbf{I}_m)$$

$$\sigma^{-2} / \mathbf{d}, s^2 \sim s^{-2} \frac{\chi_q^2}{q}$$

Cette distribution, que nous appelons distribution *fiducio-bayésienne* pour la structure "m-normale - khi-deux" (et que nous marquons d'un astérisque), est généralement justifiée dans le cadre bayésien par le recours à la distribution initiale *non-informative* définie en prenant la densité $f(\delta, \sigma^2)$ proportionnelle à σ^{-2} (distribution impropre) : cf. Lindley [5], Box et Tiao [2].

La distribution fiducio-bayésienne (marginale) relative au paramètre δ est une distribution qu'il est d'usage en statistique bayésienne d'appeler distribution du *t de Student m-dimensionnel généralisé*, de centre $\mathbf{d}$ et de matrice d'échelle $bs^2 \mathbf{I}_m$; nous noterons :

$$\delta^* / \mathbf{d}, s^2 \sim \mathbf{t}_{m,q} (\mathbf{d}, bs^2 \mathbf{I}_m)$$

Lorsque $m = 1$, il s'agit simplement de la distribution usuelle du *t de Student* à q degrés de liberté, à un facteur de centrage et d'échelle près :

$$\text{pour } m = 1, \quad \delta^* / d, s^2 \sim d + \sqrt{bs} t_q$$

La distribution fiducio-bayésienne relative au couple (λ^2, σ^2) est obtenue à partir de la distribution :

$$\lambda^{2*} / \sigma^2, \mathbf{d}, s^2 \sim bs^2 \chi_m^2 \left(\frac{1}{bs^2} \right)$$

La distribution fiducio-bayésienne (marginale) relative au paramètre numérique λ^2 est une distribution, que nous appelons distribution du *psi-deux*, à m et q degrés de liberté, qui généralise la distribution du khi-deux non-centré, et dont le rôle apparaît central dans les extensions bayésiennes de l'analyse de la variance (voir Lecoutre [3] et [4]) ; nous écrirons :

$$\lambda^{2*} / \mathbf{d}, s^2 \sim bs^2 \psi_{m,q}^2 \left(\frac{1}{bs^2} \right)$$

Il pourra encore être intéressant de considérer les distributions fiducio-bayésiennes relatives, d'une part à $\frac{\delta}{\sigma}$ et d'autre part à $\frac{\lambda^2}{\sigma^2}$.

La première est une distribution, que nous appelons *distribution $\mathbf{L}'$ m-dimensionnelle*, à q degrés de liberté (voir sa définition en annexe) ; nous écrirons :

$$\frac{\delta^*}{\sigma^*} / \mathbf{d}, s^2 \sim \mathbf{L}'_{m,q} \left(\frac{\mathbf{d}}{s}, b \mathbf{I}_m \right)$$

La seconde est une distribution, que nous appelons *distribution $\mathbf{L}^2$* , à m et q degrés de liberté (voir sa définition en annexe) ; nous écrirons :

$$\frac{\lambda^{2*}}{\sigma^{2*}} / \mathbf{d}, s^2 \sim b \mathbf{L}_{m,q}^2 \left(\frac{1}{bs^2} \right)$$

Par ailleurs, si nous désignons par F_{obs} la valeur prise par la statistique de test F définie en 2.1., nous pouvons écrire les distributions fiducio-bayésiennes relatives à λ^2 et au rapport $\frac{\lambda^2}{\sigma^2}$ sous la forme :

$$\begin{aligned} \lambda^{2*} / d, s^2 &\sim \frac{1^2}{mF_{\text{obs}}} \psi_{m,q}^2 (mF_{\text{obs}}) \\ &\quad \text{si } 1^2 \neq 0 \\ \frac{\lambda^{2*}}{\sigma^{2*}} / d, s^2 &\sim \frac{\frac{1^2}{s^2}}{mF_{\text{obs}}} L_{m,q}^2 (mF_{\text{obs}}) \end{aligned}$$

Et, dans le cas particulier $m = 1$, si nous désignons par t_{obs} la valeur prise par la statistique de test t définie en 2.1., nous avons encore :

$$\begin{aligned} \delta^* / d, s^2 &\sim t_q \left(d, \left(\frac{d}{t_{\text{obs}}} \right)^2 \right) \\ &\quad \text{si } d \neq 0 \\ \frac{\delta^*}{\sigma^{2*}} / d, s^2 &\sim L'_q \left(\frac{d}{s}, \left(\frac{\frac{d}{s}}{t_{\text{obs}}} \right)^2 \right) \quad (\text{en notant } L'_q \text{ pour } \mathbf{L}'_{1,q}) \end{aligned}$$

2.2.3. Cas d'une distribution initiale de la même famille que la distribution fiducio-bayésienne

Il peut être intéressant de considérer le cas d'une distribution initiale de la même famille que la distribution fiducio-bayésienne, soit une distribution initiale caractérisée par :

$$\begin{aligned} \delta / \sigma^2 &\sim N_m(d_o, b_o \sigma^2 \mathbf{I}_m) \\ \sigma^{-2} &\sim s_o^{-2} \frac{\chi^2_{q_o}}{q_o} \end{aligned}$$

La distribution finale est encore une distribution de la même famille, donnée par :

$$\begin{aligned} \delta / \sigma^2, d, s^2 &\sim N_m \left(d_1, \frac{b_o b}{b_o + b} \sigma^2 \mathbf{I}_m \right) \\ \text{où } d_1 &= \frac{b_o d + b d_o}{b_o + b} \\ \sigma^{-2} / d, s^2 &\sim s_1^{-2} \frac{\chi^2_{q_o + q + m}}{q_o + q + m} \\ \text{où } s_1^2 &= \frac{q_o s_o^2 + q s^2 + \frac{(d_o - d)'(d_o - d)}{b_o + b}}{q_o + q + m} \end{aligned}$$

On en déduit :

$$\delta / d, s^2 \sim t_{m, q_o + q + m} \left(d_1, \frac{b_o b}{b_o + b} s_1^2 \mathbf{I}_m \right)$$

$$\lambda^2/\sigma^2, \mathbf{d}, s^2 \sim \frac{b_o b}{b_o + b} \sigma^2 \chi_m^2 \left(\frac{\mathbf{d}_1' \mathbf{d}_1}{\frac{b_o b}{b_o + b} \sigma^2} \right)$$

et encore :

$$\frac{\delta}{\sigma} / \mathbf{d}, s^2 \sim \mathbf{L}'_{m, q_o + q + m} \left(\frac{\mathbf{d}_1}{s_1}, \frac{b_o b}{b_o + b} \mathbf{I}_m \right)$$

$$\frac{\lambda^2}{\sigma^2} / \mathbf{d}, s^2 \sim \frac{b_o b}{b_o + b} L_{m, q_o + q + m}^2 \left(\frac{\mathbf{d}_1' \mathbf{d}_1}{\frac{b_o b}{b_o + b} s_1^2} \right)$$

On remarquera que, lorsque $b_o \rightarrow +\infty$ et $q_o \rightarrow 0$, la distribution précédente diffère de la distribution fiducio-bayésienne "par m degrés de liberté", ce qui est dû à l'hypothèse de non-indépendance des distributions initiales de δ et σ^{-2} .

3. STRUCTURE STATISTIQUE "MULTINORMALE - WISHART"

Nous appellerons structure statistique "multinormale - Wishart", ou encore "r-normale - Wishart", la donnée de deux statistiques, $\mathbf{d}$ et $\mathbf{S}$, qui, sous un certain modèle d'échantillonnage, sont distribuées indépendamment, respectivement $N_r(\delta, b\Sigma)$ et $\frac{1}{q} \omega_{r,q}(\Sigma)$ (distribution de *Wishart* r-dimensionnelle à q degrés de liberté, de matrice de covariance Σ , cf. Anderson [1] et Rao [7]) ; ce que nous noterons :

$$\mathbf{d} \perp \mathbf{S} / \delta, \Sigma$$

$$\mathbf{d} / \delta, \Sigma \sim N_r(\delta, b\Sigma) \quad b \in \mathbb{R}_+ \quad |\Sigma| > 0$$

$$\mathbf{S} / \delta, \Sigma \sim \frac{1}{q} \omega_{r,q}(\Sigma)$$

La structure "r-normale - Wishart" fait donc intervenir les paramètres multinumériques δ et Σ . Pour l'inférence, nous ferons jouer le rôle de *paramètre numérique principal* au paramètre $\Delta^2 = \delta' \Sigma^{-1} \delta$, où $\Delta = (\delta' \Sigma^{-1} \delta)^{1/2}$ généralise la *distance de Mahalanobis* classique, si $r \geq 2$ (si $r=1$, on retrouve la structure "normale - khi-deux" déjà étudiée en 2.). En particulier, l'hypothèse nulle $\delta = \mathbf{0}$ est équivalente à $\Delta^2 = 0$. Nous définirons en conséquence la statistique numérique $D^2 = \mathbf{d}' \mathbf{S}^{-1} \mathbf{d}$.

Le test de signification sera calqué sur le test classique de *Hotelling*. Les procédures bayésiennes mettront en évidence de nouvelles distributions, comme dans le cas précédent.

3.1. Test de signification de l'hypothèse nulle $\Delta^2=0$

On envisagera la statistique de test $F = \frac{q-r+1}{q} \frac{D^2}{rb} \quad (1)$, dont la distribution d'échantillonnage est celle d'un F de Fisher-Snedecor à r et q-r+1 degrés de liberté, non-centré en général (d'excentricité $\frac{\Delta^2}{b}$), centré lorsque $\Delta^2=0$; ce que nous écrirons :

$$\text{sous } H_0 : \Delta^2 = 0, F \sim F_{r, q-r+1}$$

(F centré usuel à r et q-r+1 degrés de liberté)

3.2. Procédures bayésiennes

On se donnera pour (δ, Σ) une distribution initiale, et on en déduira, à partir de la distribution d'échantillonnage $(d, S)/\delta, \Sigma$, la distribution finale donnée par le théorème de Bayes.

3.2.1. Cas de distributions initiales marginales indépendantes

Dans le cas particulier où les distributions initiales (marginales) relatives à chacun des paramètres δ et Σ sont indépendantes, on pourra, en utilisant l'indépendance des distributions d'échantillonnage, décomposer l'inférence :

. d'une part en une inférence relative à la structure "multinormale"

$$d/\delta, [\Sigma, S] \sim N_r(d, b\Sigma)$$

. d'autre part en une inférence relative à la structure "Wishart"

$$S/\Sigma, [d] \sim \frac{1}{q} \omega_{r, q}(\Sigma)$$

On en déduira les distributions finales relatives, d'une part à δ étant donné Σ , et d'autre part à Σ , ce qui déterminera la distribution finale conjointe.

3.2.2. Distributions fiducio-bayésiennes

Comme distribution initiale on peut prendre pour δ la distribution (conjuguée) $N_r(d_0, b_0 U_0)$ ($|U_0| > 0$) et pour Σ^{-1} la distribution (conjuguée) $\frac{1}{q_0} \omega_{r, q_0}(S_0^{-1})$

avec $\delta \propto \Sigma^{-1}$; la distribution finale relative à δ étant donné Σ est de la même forme que celle donnée en 2.2.2. pour δ étant donné σ^2 (mais avec Σ quelconque) ; celle relative à Σ^{-1} est :

$$\Sigma^{-1}/d, S \sim \frac{1}{q_0+q} \omega_{r, q_0+q} \left(\left(\frac{q_0 S_0 + q S}{q_0+q} \right)^{-1} \right)$$

Quand $b_0 \rightarrow +\infty$ et $q_0 \rightarrow 0$, ce qui revient à exprimer un état initial d'ignorance, on obtient, en passant à la limite, la distribution finale, caractérisée par les propriétés suivantes :

(1) On a encore $F = \frac{q-r+1}{q} \frac{T^2}{r}$, en posant $T^2 = \frac{D^2}{b}$ où T^2 correspond à la statistique de test de Hotelling.

$$\begin{aligned}\delta^* / \Sigma, d, S &\sim N_r(d, b\Sigma) \\ \Sigma^{-1*} / d, S &\sim \frac{1}{q} W_{r,q}(S^{-1})\end{aligned}$$

Cette distribution, que nous appelons distribution *fiducio-bayésienne* pour la structure "r-normale - Wishart" (et que nous marquons d'un astérisque), est généralement justifiée dans le cadre bayésien par le recours à la distribution initiale *non-informative* définie en prenant la densité $f(\delta, \Sigma)$ proportionnelle à $|\Sigma|^{-\frac{r+1}{2}}$ (distribution impropre) : cf. Box et Tiao [2], Press [6].

La distribution fiducio-bayésienne (marginale) relative au paramètre δ est une distribution du t de Student r-dimensionnel généralisé, à q-r+1 degrés de liberté :

$$\delta^* / d, S \sim t_{r, q-r+1} \left(d, b \frac{q}{q-r+1} S \right)$$

On en déduit en particulier :

$$\delta^* ' S^{-1} \delta^* / d, S \sim b \frac{q}{q-r+1} \psi_{r, q-r+1}^2 \left(\frac{D^2}{b \frac{q}{q-r+1}} \right)$$

La distribution fiducio-bayésienne relative au couple (Δ^2, Σ) est obtenue à partir de la distribution :

$$\Delta^{2*} / \Sigma, d, S \sim b \chi_r^2 \left(\frac{d' \Sigma^{-1} d}{b} \right)$$

La distribution fiducio-bayésienne (marginale) relative au paramètre Δ^2 est une distribution L^2 , à r et q degrés de liberté :

$$\Delta^{2*} / d, S \sim b L_{r,q}^2 \left(\frac{D^2}{b} \right)$$

Par ailleurs, si nous désignons par F_{obs} la valeur prise par la statistique de test F définie en 3.1., nous pouvons écrire les distributions fiducio-bayésiennes relatives à $\delta^* ' S^{-1} \delta^*$ et à Δ^2 sous la forme :

$$\begin{aligned}\delta^* ' S^{-1} \delta^* / d, S &\sim \frac{D^2}{r F_{\text{obs}}} \psi_{r, q-r+1}^2 (r F_{\text{obs}}) \\ \Delta^{2*} / d, S &\sim \frac{q-r+1}{q} \frac{D^2}{r F_{\text{obs}}} L_{r,q}^2 \left(\frac{q}{q-r+1} r F_{\text{obs}} \right)\end{aligned} \quad \text{si } D^2 \neq 0$$

3.2.3. Cas d'une distribution initiale de la même famille que la distribution fiducio-bayésienne

Il peut être intéressant de considérer le cas d'une distribution initiale de la même famille que la distribution fiducio-bayésienne, soit une distribution initiale caractérisée par :

$$\begin{aligned}\delta/\Sigma &\sim N_r(d_o, b_o \Sigma) \\ \Sigma^{-1} &\sim \frac{1}{q_o} \omega_{r, q_o} (S_o^{-1})\end{aligned}$$

La distribution finale est encore une distribution de la même famille, donnée par :

$$\begin{aligned}\delta/\Sigma, d, S &\sim N_r(d_1, \frac{b_o b}{b_o + b} \Sigma) \\ \text{où} \quad d_1 &= \frac{b_o d + b d_o}{b_o + b} \\ \Sigma^{-1} / d, S &\sim \frac{1}{q_o + q + 1} \omega_{r, q_o + q + 1} (S_1^{-1}) \\ S_1 &= \frac{q_o S_o + q S + \frac{(d_o - d)(d_o - d)'}{b_o + b}}{q_o + q + 1}\end{aligned}$$

On en déduit :

$$\begin{aligned}\delta/d, S &\sim t_{r, q_o + q - r + 2} \left(d_1, \frac{b_o b}{b_o + b} \frac{q_o + q + 1}{q_o + q - r + 2} S_1 \right) \\ \delta' S_1^{-1} \delta / d, S &\sim \frac{b_o b}{b_o + b} \frac{q_o + q + 1}{q_o + q - r + 2} \psi_{r, q_o + q - r + 2}^2 \left(\frac{d_1' S_1^{-1} d_1}{\frac{b_o b}{b_o + b} \frac{q_o + q - 1}{q_o + q - r + 2}} \right) \\ \Delta^2 / d, S &\sim \frac{b_o b}{b_o + b} L_{r, q_o + q + 1}^2 \left(\frac{d_1' S_1^{-1} d_1}{\frac{b_o b}{b_o + b}} \right)\end{aligned}$$

On remarquera que, lorsque $b_o \rightarrow +\infty$ et $q_o \rightarrow 0$, la distribution précédente diffère de la distribution fiducio-bayésienne "par un degré de liberté", ce qui est dû à l'hypothèse de non-indépendance des distributions initiales de δ et Σ^{-1} .

4. ANNEXES : DISTRIBUTIONS INTRODUITES

4.1. Distributions $\mathbf{L}'_{p,q}(\mathbf{a})$

Si $\mathbf{x}/y \sim \mathbf{N}_p(\sqrt{y} \mathbf{a}, b\mathbf{I}_p)$ et $y \sim \frac{\chi^2_q}{q}$, alors nous dirons que $\mathbf{x}$ a la distribution $\mathbf{L}'$ p-dimensionnelle, à q degrés de liberté, de position $\mathbf{a}$ et d'échelle $b\mathbf{I}_p$; nous noterons :

$$\mathbf{x} \sim \mathbf{L}'_{p,q}(\mathbf{a}, b\mathbf{I}_p)$$

On a les propriétés suivantes (écrites symboliquement) :

$$\mathbf{L}'_{p,q}(\mathbf{0}, b\mathbf{I}_p) = \mathbf{N}_p(\mathbf{0}, b\mathbf{I}_p)$$

$$\mathbf{L}'_{p,\infty}(\mathbf{a}, b\mathbf{I}_p) = \mathbf{N}_p(\mathbf{a}, b\mathbf{I}_p)$$

Si $\mathbf{x} \sim \mathbf{L}'_{p,q}(\mathbf{a})$, on montre que $\mathbf{x}$ a pour densité :

$$f(\mathbf{x}) = (2\pi b)^{-\frac{p}{2}} \frac{1}{\Gamma(\frac{q}{2})} \left(\frac{qb}{qb + \mathbf{a}'\mathbf{a}} \right)^{\frac{q}{2}} \exp \left(-\frac{\mathbf{x}'\mathbf{x}}{2b} \right) \sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma\left(\frac{q+j}{2}\right) \left(\frac{2}{b(qb + \mathbf{a}'\mathbf{a})} \right)^{\frac{1}{2}} \mathbf{a}'\mathbf{x})^j$$

pour moyenne :

$$\frac{\sqrt{2}}{\sqrt{q}} \frac{\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q}{2})} \mathbf{a}$$

et pour matrice de variances et covariances :

$$b\mathbf{I}_p + \left(1 - \frac{2}{q} \left[\frac{\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q}{2})} \right]^2 \right) \mathbf{a}\mathbf{a}'$$

Enfin, si $\mathbf{x} \sim \mathbf{L}'_{p,q}(\mathbf{a}, b\mathbf{I}_p)$ et si $\mathbf{M}$ est une matrice à p lignes et k ($k \leq p$) colonnes, telle que $\mathbf{M}'\mathbf{M} = c\mathbf{I}_k$, alors :

$$\mathbf{M}'\mathbf{x} \sim \mathbf{L}'_{k,q}(\mathbf{M}'\mathbf{a}, bc\mathbf{I}_k)$$

En particulier, en prenant pour $\mathbf{M}$ un vecteur colonne comportant un 1 à la ligne i et des 0 partout ailleurs, on obtient la distribution marginale du ième composant de $\mathbf{x}$, soit :

$$x^i \sim \mathbf{L}'_{1,q}(a^i, b)$$

qui, suivant la définition précédente, est la distribution $\mathbf{L}'$ unidimensionnelle à q degrés de liberté, d'indice de position a^i et d'indice d'échelle b ; nous noterons encore :

$$x^i \sim \mathbf{L}'_q(a^i, b)$$

4.2. Distribution $L_{p,q}^2(a)$

Si $x/y \sim \chi_p^2(ay)$ et $y \sim \frac{\chi_q^2}{q}$, alors nous dirons que x a la distribution L^2 , à p et q degrés de liberté, d'excentricité a ; nous noterons :

$$x \sim L_{p,q}^2(a)$$

On peut également donner la caractérisation équivalente suivante :

si $x \sim L_{p,q}'(a, l_p)$ alors $x'x \sim L_{p,q}^2(a'a)$.

On a les propriétés suivantes (écrites symboliquement) :

$$L_{p,q}^2(0) = \chi_p^2$$

$$L_{p,\infty}^2(a) = \chi_p^2(a)$$

Si $x \sim L_{p,q}^2(a)$, on montre que x a pour densité :

$$f(x) = 2^{-\frac{p}{2}} \frac{1}{\Gamma(\frac{q}{2})} \left(\frac{q}{q+a}\right)^{\frac{q}{2}} x^{\frac{p}{2}-1} \exp\left(-\frac{x}{2}\right) \sum_{j=0}^{+\infty} \frac{1}{j!} \frac{\Gamma(\frac{q}{2}+j)}{\Gamma(\frac{p}{2}+j)} \left(\frac{ax}{2(q+a)}\right)^j$$

pour moyenne : $p+a$

et pour variance : $2(p+2a + \frac{a^2}{q})$

BIBLIOGRAPHIE

- [1] ANDERSON T.W., *An introduction to multivariate statistical analysis*, New York, John Wiley and Sons, 1958.
- [2] BOX G.E.P., TIAO G.C., *Bayesian inference in statistical analysis*, Wesley Publishing Company, 1973.
- [3] LECOUTRE B., *Extensions bayésiennes de l'analyse de la variance*, Paris, Université René Descartes, thèse de doctorat de troisième Cycle, 1980.
- [4] LECOUTRE B., "Extensions de l'analyse de la variance : l'analyse bayésienne des comparaisons", *Math. Sci. hum.*, 75, 1981, p.49-69.
- [5] LINDLEY D.V., *Introduction to probability and statistics from a bayesian viewpoint, Part 2 - Inference*, Cambridge University Press, 1965.
- [6] PRESS S.J., *Applied multivariate analysis*, New York, Holt, Rinehart and Winston, Inc., 1972.
- [7] RAO C.R., *Linear statistical inference and its applications*, New York, John Wiley and Sons, 1965.