UNE MÉTHODE DES MOMENTS STABLE POUR LES MÉLANGES DE DEUX DISTRIBUTIONS NORMALES.

Rev. Roumaine Math. Pures Appl., 49 (2004), I, 45-62

Emmanuel MONFRINI,
Courriel : monfrini@univ-lyon1.fr

Résumé. A travers l'étude de l'unicité des solutions du système classique des moments du mélange de deux distributions normales, nous mettons en évidence la nécessité d'élargir la méthode de Pearson jusqu'ici réputée trop instable. Nous considérons, alors, un second système que nous montrons être complémentaire du premier, et que nous inversons. Cette étude nous permet de rendre stable la méthode des moments pour ce type de mélanges.

A stabilized method of moments for mixtures of two normal distributions.

Abstract. *By studying uniqueness, we demonstrate that Pearson's unstable method of moments for mixtures of two normal distributions must be completed. We then invert a second system of moment equations, which is constructed to complete the classical one. We are thus able to stabilize the method of moments for this particular case of mixture, by adequately exploiting the first six moment equations.*

Mots clés : Method of moments (78M05); Moment problem (44A60); Mixture distribution (statistics 62-XX); Nonlinear equations and systems, general (34a34); General algebraic systems (08-XX).

1. Introduction.

Nous nous intéressons ici à l'estimation des paramètres d'un mélange de deux distributions gaussiennes inconnues par la méthode des moments. La méthode des moments se résume principalement à l'inversion d'un système d'équations algébriques avec contraintes (cf [15] ou [2]). Ainsi, pour un échantillon de taille n, en notant $\overrightarrow{M}$ et $\overrightarrow{M^{(n)}}$, les vecteurs (éventuellement de dimension infinie) dont la $k^{ième}$ composante est le moment d'ordre k, respectivement le moment empirique d'ordre k, d'une distribution H, et en notant θ le vecteur de ses paramètres, il existe une fonctionnelle notée $\mathfrak{H}$ telle que

$\overrightarrow{M} = \mathfrak{H}(\theta)$. Nous définissons, alors, un estimateur des moments $\widehat{\theta}$ de θ en résolvant l'équation $\overrightarrow{M^{(n)}} = \mathfrak{H}(\widehat{\theta})$. Il s'agit donc d'inverser $\mathfrak{H}$ en considérant la restriction de cette fonctionnelle à un système constitué d'un nombre fini d'équations, dont la détermination mérite quelques précautions. En effet, l'estimation d'un nombre fini, noté q, de paramètres nécessite, en général, le recours à un même nombre q d'équations de moment, qu'il faut choisir en tenant compte de la perte de précision de l'estimation lors de l'augmentation de l'ordre des moments. Nous cherchons, ici, à préciser cette problématique pour un mélange de deux distributions gaussiennes, mélange ayant révélé des cas d'instabilité jusqu'ici inexpliqués (voir sur ce point les deux articles de référence [7] et [8]).

Fixons les notations. Soient deux distributions F_1 et F_2 sur $\mathbb{R}$, mélangées en proportions $p_1 = p$ et $p_2 = 1-p$ respectivement, et conduisant à une distribution de mélange notée $H = p_1F_1 + p_2F_2$. Nous notons M_1 la moyenne et, pour tout $k \geqslant 1$, M_k^c le $k^{ième}$ moment centré de la distribution H. Alors, pour tout $k \geqslant 1$, la $k^{ième}$ équation de moment centré s'écrit :

$$M_k^c = p\int_{\mathbb{R}} (x-M_1)^k\, dF_1(x) + (1-p)\int_{\mathbb{R}} (x-M_1)^k\, dF_2(x)\,, \tag{1.1}$$

Pour des raisons de symétrie, nous posons : $x_1 = \mu_1 - M_1$ et $x_2 = \mu_2 - M_1$, où $\mu_1 = \int_{\mathbb{R}} x dF_1(x)$ et $\mu_2 = \int_{\mathbb{R}} x dF_2(x)$. Ainsi (1.1) devient pour tout $k \geqslant 1$:

$$M_k^c = p\int_{\mathbb{R}} (x_1 + (x-\mu_1))^k\, dF_1(x) + (1-p)\int_{\mathbb{R}} (x_2 + (x-\mu_2))^k\, dF_2(x)\,,$$

et :

$$M_k^c = \sum_{i=0}^{k} C_i^k \left[p x_1^i \int_{\mathbb{R}} (x-\mu_1)^{k-i} dF_1(x) + (1-p) x_2^i \int_{\mathbb{R}} (x-\mu_1)^{k-i} dF_2(x) \right]. \tag{1.2}$$

Dans le cas qui nous intéresse ici, du mélange en proportions p_1 et p_2 de distributions gaussiennes $\mathcal{N}(\mu_1, \sigma_1^2)$ et $\mathcal{N}(\mu_2, \sigma_2^2)$, les six premières équations de moment sont, en appliquant (1.2) au cas des moments de la distribution normale (cf [9]) :

$$\begin{array}{ll}
(E_1): & 0 = px_1 + (1-p)\,x_2, \\
(E_2): & M_2^c = px_1^2 + (1-p)\,x_2^2 + \left(p\sigma_1^2 + (1-p)\,\sigma_2^2\right), \\
(E_3): & M_3^c = px_1^3 + (1-p)\,x_2^3 + 3\left(p\sigma_1^2 x_1 + (1-p)\,\sigma_2^2 x_2\right), \\
(E_4): & M_4^c = px_1^4 + (1-p)\,x_2^4 + 6\left(p\sigma_1^2 x_1^2 + (1-p)\,\sigma_2^2 x_2^2\right) + 3\left(p\sigma_1^4 + (1-p)\,\sigma_2^4\right), \\
(E_5): & M_5^c = px_1^5 + (1-p)\,x_2^5 + 10\left(p\sigma_1^2 x_1^3 + (1-p)\,\sigma_2^2 x_2^3\right) + 15\left(p\sigma_1^4 x_1 + (1-p)\,\sigma_2^4 x_2\right), \\
(E_6): & M_6^c = p\,x_1^6 + (1-p)\,x_2^6 + 15\left(p\,\sigma_1^2 x_1^4 + (1-p)\,\sigma_2^2 x_2^4\right) + 15\left(p\sigma_1^6 + (1-p)\,\sigma_2^6\right) \\
 & \qquad + 45\left(p\,\sigma_1^4 x_1^2 + (1-p)\,\sigma_2^4 x_2^2\right).
\end{array}$$

Soient, alors, les deux systèmes S_1 et S_2, formés des quatre premières équations ($E_1 - E_4$), auxquelles on adjoint respectivement la cinquième (E_5) et la sixième (E_6) équation.

Le fait de travailler sur un mélange fini de distributions normales nous assure de l'identifiabilité de cette classe de mélanges finis (cf [17] ou [18]). Soulignons, pour la suite, que la notion d'identifiabilité d'une classe de mélanges finis est définie à une permutation des indices près (ici, $\mu_1 \leftrightarrow \mu_2$, $\sigma_1^2 \leftrightarrow \sigma_2^2$ et $p_1 \leftrightarrow p_2$, cf [13]) et que le même problème survient, bien sûr, dans l'étude de l'unicité des solutions des systèmes S_1 et S_2. De plus, il convient de signaler que l'unicité de la solution du problème des moments $\overrightarrow{M} = \mathfrak{H}(\theta)$, donnée par l'identifiabilité, ne suffit pas à assurer l'unicité des solutions des systèmes S_1 et S_2 qui ne sont composés chacun que de cinq équations de moments.

Nos résultats, sont alors les suivants. Nous montrons l'existence de deux variétés, notées V_1 et V_2, sur lesquelles l'inversion locale du système S_1 n'est pas possible. Ces deux variétés sont directement à l'origine des problèmes d'instabilité constatés dans l'utilisation de la méthode des moments de Pearson, comme nous le signalons plus loin. La mise en évidence de l'existence de V_1 et V_2 nous conduit naturellement à étudier le système S_2 susceptible de suppléer à S_1. Ce système, lui, ne peut être inversé localement sur une troisième variété, notée V_3. Une étape importante de notre travail consiste à établir que $V_3 \cap [V_1 \cup V_2] = \emptyset$. Ce résultat qui a pour conséquence de montrer l'injectivité locale du système surdéterminé $S_1 \cup S_2$ constitué des six premières équations de moment, permet surtout d'envisager la construction d'une méthode des moments stable exploitant la complémentarité de S_1 et S_2. Nous effectuons, ensuite, l'inversion algébrique de S_1 et S_2. Nous inversons ces deux systèmes en suivant la démarche de Pearson [15], par une méthode de substitution qui les ramène à la résolution d'équations polynomiales. Pour S_1, nous retrouvons le polynôme de Pearson $P_{Pearson}$. Si k_4 et k_5 sont les quatrième et cinquième cumulants du mélange, nous avons :

$$\begin{aligned}P_{Pearson}(x) = {} & 24x^9 + 84k_4x^7 + 36M_3^{c2}x^6 + (90k_4^2 + 72k_5M_3^c)x^5 + (444k_4M_3^{c2} - 18k_5^2)x^4 \\ & + (288M_3^{c4} - 108M_3^ck_4k_5 + 27k_4^3)x^3 - (63M_3^{c2}k_4^2 + 72M_3^{c3}k_5)x^2 - 96M_3^{c4}k_4x - 24M_3^{c6}.\end{aligned}$$

Grâce à la méthode originale, que nous développons en détail ici, pour obtenir l'inversion de S_1, l'inversion de S_2 se réduit alors à une substitution d'une partie des résultats obtenus pour S_1 dans la sixième équation de moment (E_6), au lieu de la cinquième (E_5). Pour inverser S_2, nous nous ramenons à la recherche des racines positives d'un polynôme de degré douze, noté ici P_{12}.

La mise en évidence de V_1 et V_2 nous montre que, même au niveau local, le système S_1 ne peut apporter qu'une réponse partielle au problème des moments pour un mélange de deux distributions gaussiennes. Pour préciser les cas de non unicité locale dans l'inversion de S_1 rappelons que, dans son article fondateur [15], Pearson signale les insuffisances de S_1 sur la variété V_1, qui correspond au cas d'une distribution symétrique avec $\mu_1 = \mu_2$. En effet, (E_3) et (E_5) sont toujours vérifiées et le système des cinq premières équations de moment est sous-déterminé. Dans ce cas, Pearson suggère dans [15] de recourir à la sixième équation de moment, à travers un système formé des équations (E_1), (E_2), (E_4) et (E_6). En ce qui concerne le probléme lié à V_2, il n'est mentionné ni par Pearson, ni après lui. Nous montrons ici que dans le cas de V_2 la racine de $P_{Pearson}$ permettant l'inversion de S_1 est une racine double. Observons aussi que le système S_2 qui est, en fait, une généralisation du système envisagé par Pearson pour résoudre le cas où les ensembles de paramètres appartiennent à V_1, permet de compléter la réponse que nous pouvons apporter au problème des moments. Soulignons, cependant, que la connaissance de l'unicité locale, ne nous permet pas de règler la question de l'unicité globale. Dans un premier temps, il est nécessaire d' éliminer le problème lié à la symétrie de l'ensemble des paramètres ($\mu_1 \leftrightarrow \mu_2$, $\sigma_1^2 \leftrightarrow \sigma_2^2$ et $p_1 \leftrightarrow p_2$). Pour cela, conformément à ce que suggèrent Mc Lachlan et Peel [13], nous introduisons un ordre total sur l'espace des paramètres du mélange. Cette opération s'effectue naturellement lors de l'inversion algébrique des systèmes S_1 et S_2 que nous proposons. Le second point de discussion, qui concerne l'unicité globale des solutions de notre problème des moments, est lié à l'exploitation, par nos deux systèmes S_1 et S_2, d'un nombre fini d'équations de moments. Pearson avait déjà constaté dans [15] l'existence de plusieurs solutions à S_1 et suggérait de recourir à la sixième équation de moment, à travers un test statistique, pour éliminer les éventuelles solutions parasites à S_1. Nous renvoyons, à ce propos, à la discussion qui suit la remarque 5 ci-dessous.

Signalons, pour terminer cette introduction, les points suivants qui permettent de situer notre travail par rapport aux autres méthodes d'estimation utilisées pour les mélanges de distributions.

Plusieurs méthodes d'estimation des paramètres de mélange ont été proposées et l'on pourra consulter en particulier [13] très complet sur le sujet, ou [8] qui propose une comparaison entre la méthode des moments, une méthode particulière d'estimation par le maximum de vraisemblance et une méthode de minimum d'une distance du χ^2. Mentionnons aussi la méthode EM, la plus couramment utilisée dans le cadre des mélanges de lois (cf [13]). Cette dernière méthode permet de maximiser la vraisemblance et peut être adaptée au cas du mélange de deux distributions normales où la vraisemblance est infinie (cf [4], [13] et [16]). Cette technique itérative d'estimation est souvent considérée comme meilleure que la méthode des moments, historiquement la première (cf [15]), mais qui a révélé des cas d'instabilité inexpliqués jusqu'ici (voir la section 2.3. ci-dessous pour une discussion plus complète). Notons, pourtant, que la méthode des moments est fréquemment utilisée pour initialiser la méthode EM, et que cette phase est déterminante pour pouvoir éviter les maxima locaux de la vraisemblance (cf [6]) et améliorer la vitesse de convergence de l'algorithme (cf [13]). Ne serait-ce qu'à ce titre, la question que nous étudions ici revêt une certaine importance pratique en plus de son intérêt théorique évident (cf [10], [11] et [12]).

La suite de ce document est organisée de la façon suivante. Nous présentons l'ensemble de nos résultats d'unicité et d'inversion de S_1 et S_2 dans la section 2. Nous concluons la section 2. en considérant les systèmes de moments empiriques, $\overrightarrow{M^{(n)}} = \mathfrak{H}(\widehat{\theta})$, et en énonçant une méthode originale d'estimation des paramètres d'un mélange de deux distributions gaussiennes basée sur la méthode des moments. La section 3, quant à elle, est consacrée au développement des preuves des résulats présentés dans la section 2.

2. Présentation des résultats et commentaires.

2.1. Unicité locale des solutions.

Soient les trois sous-variétés V_1, V_2 et V_3 de $\mathbb{R}^2 \times \mathbb{R}^{+2} \times \left]0,1\right[$ définies par :

$$\begin{aligned} V_1 &= \left\{ \left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p\right), \mu_1 = \mu_2 \right\}, \\ V_2 &= \left\{ \left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p\right), \left|\mu_1 - \mu_2\right| = \left(6\sqrt{6} - 9\right)^{1/4} \sqrt{\left|\sigma_1^2 - \sigma_2^2\right|} \right\}, \\ V_3 &= \left\{ \begin{array}{c} \left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p\right),\ p = \left(\frac{1}{2} + \frac{3\left(\sigma_1^2-\sigma_2^2\right)\left(15\left(\sigma_1^2-\sigma_2^2\right)^4 - 10(\mu_1-\mu_2)^4\left(\sigma_1^2-\sigma_2^2\right)^2 - (\mu_1-\mu_2)^8\right)}{2(\mu_1-\mu_2)^2\left(135\left(\sigma_1^2-\sigma_2^2\right)^4 - 18(\mu_1-\mu_2)^4\left(\sigma_1^2-\sigma_2^2\right)^2 - (\mu_1-\mu_2)^8\right)} \right) \\ \text{et} \quad (\mu_1-\mu_2)^2 \left(135\left(\sigma_1^2-\sigma_2^2\right)^4 - 18(\mu_1-\mu_2)^4\left(\sigma_1^2-\sigma_2^2\right)^2 - (\mu_1-\mu_2)^8\right) \neq 0 \end{array} \right\}. \end{aligned}$$

Remarquons que $V_3 \cap [V_1 \cup V_2] = \emptyset$ par construction. Nous sommes alors en mesure d'énoncer deux résultats d'unicité.

PROPOSITION 2.1. – *Il y a unicité locale des solutions du système S_1 si et seulement si $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p)$ n'appartient pas à $V_1 \cup V_2$.*

PROPOSITION 2.2. – *Il y a unicité locale des solutions du système S_2 si et seulement si $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p)$ n'appartient pas à V_3.*

Remarque 1. – Le cas des mélanges correspondants à des distributions symétriques coïncide avec V_1 lorsque $\mu_1 = \mu_2$, et avec V_3 lorsque $\sigma_1^2 = \sigma_2^2$ et $p = 0.5$.

Sur le plan local, la résolution de la question de l'existence et de l'unicité des solutions du problème théorique des moments $\overrightarrow{M} = \mathfrak{H}\left(\theta\right)$ est donnée par le corollaire suivant :

COROLLAIRE 2.3. – *Les systèmes S_1 et S_2 sont complémentaires pour la résolution du problème des moments pour le mélange de deux distributions normales.*

COROLLAIRE 2.4. – *Il y a unicité locale des solutions du système $S_1 \cup S_2$ constitué des six premières équations de moments (injectivité locale).*

2.2. Inversion algébrique des systèmes S_1 et S_2.

Nous n'envisageons ici que le cas où les moyennes μ_1 et μ_2 sont différentes, le cas où les moyennes sont identiques est développé dans [3], et les résultats sont rappelés dans la section 3.2.2. Pour des raisons de commodités techniques nous effectuons, dans les équations ($E_2 - E_6$), les changements de variables :

$$\sigma = p\sigma_1^2 + (1-p)\,\sigma_2^2 \quad \text{et} \quad \sigma' = p\sigma_1^2 x_1 + (1-p)\,\sigma_2^2 x_2. \tag{2.3}$$

Nous obtenons ainsi les six équations transformées :

$$
\begin{array}{ll}
(E_2') : & px_1^2 + (1-p)\,x_2^2 = M_2^c - \sigma, \\
(E_3') : & px_1^3 + (1-p)\,x_2^3 = M_3^c - 3\sigma', \\
(E_4') : & px_1^4 + (1-p)\,x_2^4 = M_4^c - 6\left(\sigma'\left(x_1+x_2\right) - \sigma x_1 x_2\right) - \frac{3\left(\left(\sigma'\right)^2 + \sigma^2\left(px_1^2+(1-p)x_2^2\right)\right)}{p(1-p)(x_1-x_2)^2}, \\
(E_5') : & px_1^5 + (1-p)\,x_2^5 = M_5^c - 10\left(\sigma'\left(\left(x_1+x_2\right)^2 - x_1x_2\right) - \sigma x_2\, x_1\,(\,x_1+x_2)\right) \\
 & \qquad + \frac{15}{x_1x_2}\left(\sigma'^2\left(x_1+x_2\right) - 2\sigma\sigma'\, x_1x_2\right), \\
(E_6') : & \left(px_1^6 + (1-p)\,x_2^6\right)\left(x_1x_2\right)^2 = -15\ \left(x_1x_2\right)^2\left[\sigma'\left(x_1+x_2\right)\left(\left(x_1+x_2\right)^2 - 2x_2x_1\right)\right] \\
 & \qquad -15\ \left(x_1x_2\right)^2\left[\sigma x_1x_2\ \left(\left(x_1+x_2\right)^2 - x_1x_2\right)\right] + \left(x_1x_2\right)^2 M_6^c \\
 & \qquad -15\left(x_1-x_2\right)\left[\left(1-2p\right)\sigma'^3\ + \left(3\sigma'^2\sigma - \sigma^3x_1x_2\right)\left[\frac{-(1-2p)x_1x_2}{(x_1+x_2)}\right]\right] \\
 & \qquad +45x_1x_2\left[\sigma'^2\left(\left(x_1+x_2\right)^2 - x_2x_1\right) - 2\sigma'\sigma x_1x_2(x_1+x_2) + \left(\sigma x_1x_2\right)^2\right],
\end{array}
$$

d'où nous pouvons tirer les nouvelles expressions de S_1 et S_2 après changement de variable.

Nous exposons maintenant nos résultats, en commençant par énoncer deux lemmes. Nous désignons par $sgn\,(x)$ la fonction qui à x associe $\frac{x}{|x|}$ (avec la convention $sgn\,(0) = 1$) et nous posons $X = M_2^c - \sigma$, ainsi que $Y = M_3^c - 3\sigma'$. On notera que nous avons toujours $X > 0$ lorsque $\mu_1 \neq \mu_2$.

LEMME 2.5. – *Lorsque $\mu_1 \neq \mu_2$, si on cherche $p \in [0.5, 1[$ avec $\mu_1 < \mu_2$ lorsque $p = 0.5$, les paramètres μ_1, μ_2, σ_1^2, σ_2^2 et p solutions de (E_1), (E_2') et (E_3') vérifient alors les équations :*

$$
p = \frac{\sqrt{4X^3+Y^2} + |Y|}{2\sqrt{4X^3+Y^2}}, \tag{2.4}
$$

$$
\mu_1 = M_1 - sgn\,(Y)\sqrt{\frac{\sqrt{4X^3+Y^2} - |Y|}{\sqrt{4X^3+Y^2} + |Y|}X}, \tag{2.5}
$$

$$
\mu_2 = M_1 + sgn\,(Y)\sqrt{\frac{\sqrt{4X^3+Y^2} + |Y|}{\sqrt{4X^3+Y^2} - |Y|}X}, \tag{2.6}
$$

ainsi que

$$
\sigma_1^2 = \frac{\sigma' - \sigma\left(\mu_2 - M_1\right)}{p\left(\mu_1 - \mu_2\right)} \quad \textit{et} \quad \sigma_2^2 = \frac{\sigma' - \sigma\left(\mu_1 - M_1\right)}{\left(1-p\right)\left(\mu_1 - \mu_2\right)}. \tag{2.7}
$$

LEMME 2.6. – *Sous les conditions du lemme 2.5, nous avons pour le système S_1,*

$$
\sigma' = \frac{2M_3^cX^3 + k_5X^2 - 3k_4XM_3^c + 2M_3^{c3}}{-2X^3 - 3k_4X + 4M_3^{c2}}, \tag{2.8}
$$

et, lorsque $M_3^c \neq 0$, nous avons pour S_2,

$$
\sigma' = \frac{4X^6 + 8X^4k_4 + \left(k_6 + 2M_3^{c2}\right)X^3 - k_4^2X^2 - 4k_4M_3^{c2}X + 4M_3^{c4}}{M_3^c\left(-2X^3 - 7Xk_4 + 10M_3^{c2}\right)}. \tag{2.9}
$$

Nous en déduisons, alors, le résultat initial de Pearson pour le système S_1 et nous présentons un résultat équivalent pour le système S_2.

THÉORÈME 2.7. – *Lorsque M_c^3 et M_c^5 ne sont pas simultanément nuls, la résolution du système S_1 se ramène à la recherche des racines positives du polynôme de degré 9 :*

$$P_\sigma(X) = \frac{-P_{Pearson}(-X)}{3}.$$

THÉORÈME 2.8. – *La résolution du système S_2, dans le cas où $M_3^c \neq 0$, se ramène à la recherche des racines positives du polynôme de degré 12 :*

$$\begin{gathered} P_{12}(X) = 96X^{12} + 384k_4X^{10} + 8(6\ k_6 + 13\ M_3^{c2})X^9\ + 336k_4^2X^8 \\ +12k_4(8k_6 + 5M_3^{c2})X^7 + 6\left(-16k_4^3 + k_6^2 + 4\ k_6\ M_3^{c2} + 22\ M_3^{c4}\right)X^6 \\ -6k_4^2\left(2k_6 + 47M_3^{c2}\right)X^5 + 6k_4\left(k_4^3 - 2\left(4k_6M_3^{c2} + 5M_3^{c4}\right)\right)X^4 \\ +\left(97k_4^3M_3^{c2} + 48M_3^{c4}\left(k_6 + 7M_3^{c2}\right)\right)X^3 - 141k_4^2M_3^{c4}X^2 + 48\ \ k_4\ M_3^{c6}X - 4\ M_3^{c8}. \end{gathered}$$

Nous pouvons aussi expliciter le comportement du système S_1 sur la variété V_2 grâce au résultat suivant.

THÉORÈME 2.9. – *Si l'on se place sur la variété V_2, l'ensemble des solutions de S_1 est fini. Le polynôme P_σ admet une racine double positive, qui est celle permettant d'estimer les paramètres.*

Les remarques suivantes permettent de préciser certaines situations.

Remarque 2. – La restriction sur l'ensemble des paramètres dans le lemme 2.5 correspond à un ordre total et à l'existence de solutions symétriques ($\mu_1 \leftrightarrow \mu_2$, $\sigma_1^2 \leftrightarrow \sigma_2^2$ et $p_1 \leftrightarrow p_2$) à S_1 et S_2.

Remarque 3. – Lorsque $M_3^c \neq 0$, comme $P_\sigma(0) = 24M_3^{c6}$ et que le terme en X^9 de $P_\sigma(X)$ est positif, si le polynôme P_σ admet une racine positive, il en admettra au moins une deuxième, éventuellement une racine double. Dans ce dernier cas S_1 peut admettre au moins deux solutions.

Remarque 4. – Dans le cas où $M_3^c = 0$, $P_\sigma(X) = 8X^9 + 28k_4X^7 + 30k_4^2X^5 + 6k_5^2X^4 + 9k_4^3X^3$, et P_σ n'a pas de racine strictement positive si $k_4 \geqslant 0$. Observons que dans ce cas les moyennes sont égales.

Remarque 5. – P_{12} a toujours au moins une racine positive.

Comme le laissent présager les remarques 3 et 5, et comme nous le montrons dans [14], chacun des systèmes S_1 et S_2 pris séparément peut admettre pour solutions plusieurs ensembles de paramètres statistiquement acceptables (i.e. tels que $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2,\ p) \in \mathbb{R}^2 \times \mathbb{R}^{+2} \times]0,1[$) et ceci même en éliminant la possibilité de solutions symétriques. Ce cas de figure semble pouvoir être contourné grâce à l'équation de moment inexploitée. En effet, dans [14], nous n'avons jamais constaté l'existence de solutions multiples pour le système $S_1 \cup S_2$ formé des six premières équations ($E_1 - E_6$). Dans le cas, où il y aurait plusieurs ensembles de solutions, à l'un des systèmes S_1 ou S_2, ayant une existence statistique, nous choisirons parmi les solutions possibles de S_1, celle qui minimisera la quantité :

$$T_6 = \left|\frac{p\left(15\sigma_1^6 + 45\sigma_1^4 x_1^2 + 15\sigma_1^2 x_1^4 + x_1^6\right) + (1-p)\left(x_2^6 + 15\sigma_2^2 x_2^4 + 45\sigma_2^4 x_2^2 + 15\sigma_2^6\right) - M_6^c}{M_6^c}\right|,$$

ou parmi les solutions possibles de S_2, celle qui minimisera la quantité :

$$T_5 = \left|\frac{p\left(x_1^5 + 10\sigma_1^2 x_1^3 + 15\sigma_1^4 x_1\right) + (1-p)\left(x_2^5 + 10\sigma_2^2 x_2^3 + 15\sigma_2^4 x_2\right) - M_5^c}{M_5^c}\right|.$$

Cette méthode nous a toujours permis de ne garder qu'une solution au problème des moments dans l'ensemble des exemples traités dans [14].

Rappelons, pour terminer et pour être complet, le résultat de l'inversion du système dans le cas des variances identiques (cf [15]).

THÉORÈME 2.10. – *L'inversion des systèmes S_1 et S_2 dans le cas où les variances sont égales, se ramène à la recherche des racines positives du polynôme :*

$$P_{var\acute{e}gales}(X) = 2X^3 + k_4 X - M_3^{c2}.$$

Remarque 6. – $P_{var\acute{e}gales}$ n'a qu'une seule racine positive (ou nulle).

2.3. Systèmes des moments empiriques (Estimation).

Les problèmes d'unicité étudiés pour les systèmes de moments théoriques, S_1 et S_2, se compliquent lors du passage au système de moments empiriques $\overrightarrow{M^{(n)}} = \mathfrak{H}(\widehat{\theta})$. En effet, il n'y a pas continuité, en fonction des moments du mélange, des solutions des systèmes théoriques au voisinage de $V_1 \cup V_2$ ou V_3 selon les cas. La méthode d'estimation que nous proposons dans cette section permet de choisir le système pour lequel l'estimation est la plus fiable, et apporte ainsi des réponses claires aux cas d'instabilité présentés dans les articles de référence [7] et [8] dont nous commençons par rappeler le contenu.

Dans leur article [7] consacré à l'étude de la stabilité de la méthode des moments pour le mélange de deux distributions normales inconnues Fryer et Robertson, envisagent neuf paramétrages différents pour le mélange de deux distributions normales. Ils mettent ainsi en évidence trois cas pour lesquels la méthode présente un biais très important, qu'ils ne parviennent pas à expliquer. Ces trois exemples sont situés dans un voisinage immédiat de V_2. Grâce au théorème 2.9, nous pouvons maintenant préciser les travers de la méthode des moments basée sur l'équivalent empirique de S_1 au voisinage de V_2. En effet, le passage des moments théoriques aux moments empiriques, qui entraîne un "décalage" du

polynôme P_σ, peut soit conduire à l'existence de deux racines très proches et générer des jeux de paramètres proches, soit conduire à la disparition de la racine positive qui devrait permettre l'inversion convenable du système empirique associé à S_1. Nous sommes par conséquent en mesure de justifier le biais important constaté par Fryer et Robertson par le fait que S_1 n'est pas approprié à l'estimation des paramètres dans les trois exemples défavorables présentés. Il est à noter que la méthode de mesure du biais proposée par Fryer et Robertson ne fait pas apparaître de biais important dans le seul cas de mélange choisi au voisinage de V_1. Les résultats présentés dans [14] montrent pourtant une épaisseur non négligeable du voisinage de V_1 sur lequel l'utilisation de la méthode des moments basée sur S_1 n'est pas satisfaisante.
Dans leur article [8] de comparaison de différentes méthodes d'estimation des paramètres du mélange de deux distributions normales inconnues, Fryer et Robertson envisagent neuf nouveaux exemples de paramétrages. Ils constatent dans quatre cas, l'infériorité de la méthode des moments par rapport à une méthode de maximisation de la vraisemblance et à une méthode de minimisation d'une distance du χ^2, les cinq autres cas n'étant pas défavorables à la méthode des moments. Ces quatre cas, à l'origine de l'abandon progressif de la méthode des moments pour l'estimation des paramètres de distributions de mélanges, sont, d'après notre étude, à traiter avec le système S_2. En effet, deux cas sont issus d'un proche voisinage de V_1 et deux autres d'un proche voisinage de V_2.

Ainsi l'intérêt de l'étude du second système des moments apparaît clairement lors du passage aux systèmes de moments empiriques. De plus, pour exploiter convenablement la complémentarité entre S_1 et S_2, il est nécessaire de privilégier les résultats issus du système pertinent faisant intervenir les moments empiriques donnant les meilleures estimations des moments réels. Pour cela, nous favorisons systématiquement les cinq premières équations de moments, et nous n'aurons recours au second système que dans un voisinage de V_1 ou V_2. Nous proposons, en tirant les enseignements des quarante six exemples traités dans [14] et des dix-huit cas étudiés dans les articles de Fryer et Robertson ([7] et [8]), de procéder à l'estimation des paramètres du mélange de deux distributions gaussiennes inconnues par la méthode que nous présentons maintenant. Notons $\mathfrak{M}_3^c$ et κ_4 respectivement le troisième moment centré empirique et le quatrième cumulant empirique du mélange et notons $\mathcal{T}_6$ l'équivalent du test T_6 pour les moments empiriques :

► **1**. <u>Si $|\mathfrak{M}_3^c| < 0.5$ et $\mathfrak{M}_3^c \neq 0$:</u>

▷ Si $\kappa_4 > 0$, l'estimation vient de S_2.

▷ Si $\kappa_4 \leqslant 0$, solution issue de P_σ ou $P_{varégales}$ minimisant le test $\mathcal{T}_6$.

► **2**. <u>Si $\mathfrak{M}_3^c = 0$:</u>

▷ Si $\kappa_4 > 0$, cas particulier de S_2 où $\mu_1 = \mu_2$.

▷ Si $\kappa_4 \leqslant 0$, choix de la solution issue de $P_{varégales}$.

► **3**. <u>On choisira la solution de S_2 si :</u>

$\triangleright$ La solution issue de S_2 vérifie :

$$T_{V_2} = 10\left(|x_1 - x_2| - \left(6\sqrt{6} - 9\right)^{1/4}\sqrt{|\sigma_1^2 - \sigma_2^2|}\right)^2 < 2,$$

$\triangleright$ Si S_1 ne donne pas de solution acceptable.

$\blacktriangleright$ **4**. Sinon la solution vient de S_1.

Pour la mise en oeuvre pratique sur des exemples de cette méthodologie, nous renvoyons à [14].

3. Démonstrations des résultats.

3.1. Unicité des solutions.

Preuve de la proposition 2.1. Nous calculons le jacobien J_1 de S_1. Il s'écrit :

$$J_1 = p^2\left(1-p\right)^2 \times \begin{vmatrix} 1 & 1 & x_1 - x_2 & 0 & 0 \\ 2x_1 & 2x_2 & \begin{pmatrix} x_1^2 + \sigma_1^2 \\ -x_2^2 - \sigma_2^2 \end{pmatrix} & 1 & 1 \\ 3(x_1^2+\sigma_1^2) & 3(x_2^2+\sigma_2^2) & \begin{pmatrix} x_1^3 + 3\sigma_1^2 x_1 \\ -x_2^3 - 3\sigma_2^2 x_2 \end{pmatrix} & 3x_1 & 3x_2 \\ 4(x_1^3 + 3\sigma_1^2 x_1) & 4(x_2^3 + 3\sigma_2^2 x_2) & \begin{pmatrix} x_1^4 + 6\sigma_1^2 x_1^2 + 3\sigma_1^4 \\ -x_2^4 - 6\sigma_2^2 x_2^2 - 3\sigma_2^4 \end{pmatrix} & 6(x_1^2+\sigma_1^2) & 6(x_2^2+\sigma_2^2) \\ 5\begin{pmatrix} x_1^4 + 6\sigma_1^2 x_1^2 \\ +3\sigma_1^4 \end{pmatrix} & 5\begin{pmatrix} x_2^4 + 6\sigma_2^2 x_2^2 \\ +3\sigma_2^4 \end{pmatrix} & \begin{pmatrix} x_1^5 + 10\sigma_1^2 x_1^3 + 15\sigma_1^4 x_1 \\ -x_2^5 - 10\sigma_2^2 x_2^3 - 15\sigma_2^4 x_2 \end{pmatrix} & \begin{matrix} 10x_1 \times \\ (x_1^2 + 3\sigma_1^2) \end{matrix} & \begin{matrix} 10x_2 \times \\ (x_2^2 + 3\sigma_2^2) \end{matrix} \end{vmatrix}.$$

J_1 se simplifie alors en :

$$J_1 = -p^2\left(1-p\right)^2\left(x_1 - x_2\right)\left(\left(x_1 - x_2\right)^8 + 18\left(x_1 - x_2\right)^4\left(\sigma_1^2 - \sigma_2^2\right)^2 - 135\left(\sigma_1^2 - \sigma_2^2\right)^4\right),$$

et se factorise sous la forme :

$$\begin{aligned} J_1 = p^2\left(1-p\right)^2\left(x_2 - x_1\right) & \\ & \times\left(\left(x_1 - x_2\right)^4 + \left(9 + 6\sqrt{6}\right)\left(\sigma_1^2 - \sigma_2^2\right)^2\right)\left(\left(x_1 - x_2\right)^2 + \sqrt{6\sqrt{6} - 9}\left|\sigma_1^2 - \sigma_2^2\right|\right) \\ & \times\left(|x_1 - x_2| + \left(6\sqrt{6} - 9\right)^{1/4}\sqrt{|\sigma_1^2 - \sigma_2^2|}\right)\left(|x_1 - x_2| - \left(6\sqrt{6} - 9\right)^{1/4}\sqrt{|\sigma_1^2 - \sigma_2^2|}\right), \end{aligned}$$

ce qui permet de conclure grâce au théorème des fonctons implicites.⧫

Remarque 7. – Le développement du jacobien peut s'effectuer (et a été vérifié) en utilisant un logiciel de calcul formel. La factorisation demande toutefois une intervention "manuelle" de l'utilisateur.

Preuve de la proposition 2.2. Calculons le Jacobien J_2 de S_2. Il s'écrit :

$$J_2 = p^2\left(1-p\right)^2 \times \begin{vmatrix} 1 & 1 & x_1 - x_2 & 0 & 0 \\ 2x_1 & 2x_2 & x_1^2 + \sigma_1^2 - x_2^2 - \sigma_2^2 & 1 & 1 \\ 3(x_1^2+\sigma_1^2) & 3(x_2^2+\sigma_2^2) & x_1^3 + 3\sigma_1^2 x_1 - x_2^3 - 3\sigma_2^2 x_2 & 3x_1 & 3x_2 \\ 4(x_1^3 + 3\sigma_1^2 x_1) & 4(x_2^3 + 3\sigma_2^2 x_2) & \begin{matrix} x_1^4 + 6\sigma_1^2 x_1^2 + 3\sigma_1^4 \\ -x_2^4 - 6\sigma_2^2 x_2^2 - 3\sigma_2^4 \end{matrix} & 6(x_1^2+\sigma_1^2) & 6(x_2^2+\sigma_2^2) \\ \begin{matrix} 6(x_1^5 + 10\sigma_1^2 x_1^3 \\ +15\sigma_1^4 x_1) \end{matrix} & \begin{matrix} 6(x_2^5 + 10\sigma_2^2 x_2^3 \\ +15\sigma_2^4 x_2) \end{matrix} & \begin{matrix} x_1^6 + 15\sigma_1^2 x_1^4 + 45\sigma_1^4 x_1^2 + \sigma_1^6 \\ -x_2^6 - 15\sigma_2^2 x_2^4 - 45\sigma_2^4 x_2^2 - \sigma_2^6 \end{matrix} & \begin{matrix} 15(x_1^4 + \\ 6\sigma_1^2 x_1^2 + 3\sigma_1^4) \end{matrix} & \begin{matrix} 15(x_2^4 + \\ 6\sigma_2^2 x_2^2 + 3\sigma_2^4) \end{matrix} \end{vmatrix},$$

ou encore, après développement :

$$\begin{array}{rcl} J_2 & = & -3(1-p)^2p^2 \times \\ & & \left[\begin{array}{c} (x_1-x_2)^9(x_1+x_2)+3(x_1-x_2)^8(\sigma_1^2-\sigma_2^2)+18(x_1-x_2)^5(x_1+x_2)(\sigma_1^2-\sigma_2^2)^2 \\ +30(x_1-x_2)^4(\sigma_1^2-\sigma_2^2)^3-\ 135(x_1+x_2)(\,x_1-x_2)(\sigma_1^2-\sigma_2^2)^4-45(\sigma_1^2-\sigma_2^2)^5) \end{array} \right]. \end{array}$$

Et si l'on pose : $Z=\sigma_1^2-\sigma_2^2$, et :

$$F_1=3Z\left[(x_1-x_2)^8+10(x_1-x_2)^4Z^2-15Z^4\right],$$

$$F_2=(x_1+x_2)(x_1-x_2)\left[(x_1-x_2)^8+18(x_1-x_2)^4Z^2-135Z^4\right],$$

on obtient alors :

$$J_2=3(1-p)^2p^2\left[F_1+F_2\right]. \tag{3.10}$$

Remarquons que F_2 s'annule pour $|x_1|=|x_2|$ et pour $|x_1-x_2|=\left(6\sqrt{6}-9\right)^{1/4}\sqrt{|Z|}$. Le terme F_2 s'annule donc, en particulier, sur V_1 et V_2. Nous montrons maintenant que le terme F_1 ne s'annule pas sur $V_1\cup V_2$. En effet, si $x_1=x_2$, alors $x_1=x_2=0$ et $J_2=-45Z^5$ n'est pas nul pour un “vrai” mélange. Si l'on se place sur V_2, on a alors $|x_1-x_2|=\left(6\sqrt{6}-9\right)^{1/4}\sqrt{|Z|}$, et

$$F_1=3Z^5\left[\left(6\sqrt{6}-9\right)^2+10\left(6\sqrt{6}-9\right)-15\right],$$

qui n'est pas non plus nul. Ainsi les deux jacobiens J_1 et J_2 ne sont pas simultanément nuls.

Pour terminer la preuve de la proposition *2.2* et comme J_1 et J_2 ne sont pas simultanément nuls, on peut désormais se placer hors de V_1 et V_2, ce qui assure que $(x_1-x_2)^{10}+18(x_1-x_2)^6Z^2-135(x_1-x_2)^2Z^4\neq 0$. Comme $x_1+x_2=(1-2p)\,(x_1-x_2)$, et en posant $D=(x_1-x_2)$, le jacobien devient :

$$\begin{array}{rcl} J_2 & = & 3(1-p)^2p^2\left[45Z^5-3D^8Z-30D^4Z^3+\,(1-2p)\,D\left(135DZ^4-18D^5Z^2-D^9\right)\right] \\ & = & -6D^2\left(135Z^4-18D^4Z^2-D^8\right)(1-p)^2p^2\left[p-\left(\dfrac{1}{2}+\dfrac{3Z\left(15Z^4-10D^4Z^2-D^8\right)}{2D^2\left(135Z^4-18D^4Z^2-D^8\right)}\right)\right].\blacklozenge \end{array}$$

3.2. Inversion algébrique de S_1 et S_2.

3.2.1 Cas général où $\mu_1\neq\mu_2$.

Preuve du lemme 2.5. Des équations (E_1) et (E_2'), ainsi que de l'expression $X=M_2^c-\sigma$, on tire, compte tenu de la symétrie du problème $(x_1\mapsto x_2,\ p_1\mapsto p_2)$:

$$x_1=\pm\sqrt{\frac{(1-p)}{p}X}\text{ et }x_2=-sgn\,(x_1)\sqrt{\frac{p}{(1-p)}X}.$$

La quantité p est alors solution de l'équation :

$$p^2 - p + \frac{X^3}{4X^3 + Y^2} = 0,$$

dont les deux solutions, qui sont positives, sont les deux coefficients de mélanges p et $1-p$. On tire alors les expressions (2.4), (2.5) et (2.6) des trois premières équations (E_1), (E_2') et (E_3'), communes à S_1 et S_2, des expressions $X = M_2^c - \sigma$ et $Y = M_3^c - 3\sigma'$, en privilégiant le paramétrage pour lequel $p \geqslant 1-p$ et $\mu_1 < \mu_2$ lorsque $p = 1-p$. (2.7) se déduit alors des changements de variables (2.3) en se rappelant que $x_1 = \mu_1 - M_1$, $x_2 = \mu_2 - M_1$.♦

Preuve du lemme 2.6 : Prenant en compte les résultats du lemme *2.5*, après simplification de (E_4') on obtient :

$$\sigma'^2 = \frac{M_3^{c2} - (2X^2 + k_4)\,X}{6}. \tag{3.11}$$

Par ailleurs, la substitution de (2.4) dans (E_5') conduit, après simplification, à :

$$18\sigma'^2 M_3^c - \sigma'\left(4X^3 + M_3^{c2} + 18\sigma'^2\right) + (k_5 + 8M_3^c X)\,X^2 - M_3^{c3} = 0, \tag{3.12}$$

et comme $4X^3 + M_3^{c2} + 18\sigma'^2 > 0$, on obtient, en substituant (3.11) dans (3.12) :

$$\sigma' = \frac{2M_3^c X^3 + k_5 X^2 - 3k_4 X M_3^c + 2M_3^{c3}}{-2X^3 - 3k_4 X + 4M_3^{c2}}. \tag{3.13}$$

(E_6'), quant à elle, devient :

$$\begin{aligned}&\sigma'^2(72X^3 - 90X^2 M_2^c - 36M_3^{c2} + 36\sigma'^2) + \sigma'(12X^3 M_3^c + 3M_3^{ixc3} + 42\sigma'^2 M_3^c)\\ &+ 16X^6 - 30X^5 M_2^c + \left(-15k_4 M_2^c - 22M_3^{c2} - k_6\right) X^3 + 15X^2 M_2^c M_3^{c2} + M_3^{c4} = 0.\end{aligned}$$

Il est, alors, nécessaire de considérer deux cas :

$$M_3^c = 0 \text{ ou } M_3^c \neq 0,$$

en prenant soin de préciser que l'on a toujours :

$$12X^3 + 3M_3^{c2} + 42\sigma'^2 \geqslant 0.$$

Si $M_3^c = 0$, deux cas sont envisageables. Le premier correspond à $M_3^c = 0$ et $k_4 > 0$. Les moyennes sont alors égales et la résolution du système doit être menée différemment (voir la section 3.2.2). Le second correspond à $M_3^c = 0$ et $k_4 \leqslant 0$. On est alors dans le cas où les variances sont égales avec un coefficient de mélange $p = 0.5$. Nous développons le cas général du système formé des quatre premières équations, adapté au mélange de deux distributions de même variance, dans la section 3.2.3.

Si $M_3^c \neq 0$, on obtient, en substituant (3.11) dans (E_6') :

$$\sigma' = \frac{4X^6 + 8X^4 k_4 + \left(k_6 + 2M_3^{c2}\right) X^3 - k_4^2 X^2 - 4k_4 M_3^{c2} X + 4M_3^{c4}}{M_3^c \left(-2X^3 - 7Xk_4 + 10M_3^{c2}\right)}.\blacklozenge \tag{3.14}$$

Preuve des théorèmes 2.7 et 2.8. Il suffit désormais de substituer les expressions de σ' établies dans (3.13) et (3.14) respectivement.♦

3.2.2 Si $\mu_1 = \mu_2$.

Nous nous ramenons au système S_2' étudié par Pearson dans le cadre du mélange de deux distributions de moyennes égales. Il s'agit du cas où $M_3^c = 0$ et $k_4 > 0$. Si $\mu_1 = \mu_2 = M_1$, on a $x_1 = x_2 = 0$. Il reste trois paramètres à déterminer. Le système à inverser est donc :

$$\left\{ \begin{array}{l} \left(p\sigma_1^2 + (1-p)\,\sigma_2^2\right) = M_2^c \\ 3\left(p\sigma_1^4 + (1-p)\,\sigma_2^4\right) = M_4^c \\ 15\left(p\sigma_1^6 + (1-p)\,\sigma_2^6\right) = M_6^c \end{array} \right. .$$

D'après [3] l'inversion de ce système donne :

$$\begin{aligned} \sigma_1^2 &= M_2^c + \frac{1}{2}\left[\frac{k_6}{5k_4} - \sqrt{\left(\frac{k_6}{5k_4}\right)^2 + \frac{4k_4}{3}}\right], \sigma_2^2 = M_2^c + \frac{1}{2}\left[\frac{k_6}{5k_4} + \sqrt{\left(\frac{k_6}{5k_4}\right)^2 + \frac{4k_4}{3}}\right], \\ p &= \frac{\sigma_2^2 - M_2^c}{\sigma_2^2 - \sigma_1^2}, \mu_1 = \mu_2 = M_1, \end{aligned}$$

où k_6 est le sixième cumulant du mélange.

3.2.3 Si $\sigma_2^2 = \sigma_1^2$.

Preuve du théorème 2.10 : L'utilisation de ce sous-système s'avère très utile dans la pratique. Dans le cas où les variances sont égales, les systèmes S_1 et S_2 se ramènent à :

$$\left\{ \begin{array}{l} px_1 + (1-p)\,x_2 = 0 \\ px_1^2 + (1-p)\,x_2^2 = M_2^c - \sigma \\ px_1^3 + (1-p)\,x_2^3 = M_3^c \\ px_1^4 + (1-p)\,x_2^4 = M_4^c - 6\sigma\left(px_1^2 + (1-p)\,x_2^2\right) - 3\sigma^2 \end{array} \right. .$$

Des trois premières équations on tire, comme précédemment, p, μ_1, μ_2, ((cf (2.4), (2.5) et (2.6)). Dans le cas où $M_3^c = 0$, on a $p = \frac{1}{2}$ et on choisira, par exemple, $\mu_1 < \mu_2$.

La quatrième équation devient en posant $X = M_2^c - \sigma$:

$$M_3^{c2} = 2X^3 + k_4 X.\blacklozenge$$

3.3. Caractérisation de V_2.

Preuve du théorème 2.9. P_σ s'exprime en fonction de x_1+x_2, x_1-x_2 et de la différence $\sigma_1^2-\sigma_2^2$. Posons :

$$S = x_1 + x_2,\ D = x_1 - x_2 \text{ et } Z = \sigma_1^2 - \sigma_2^2.$$

Si l'on tire p de la première équation de moments, il est possible d'exprimer M_3^c, k_4 et k_5 en fonction de S, D et Z. On tire en effet :

$$\begin{aligned}
M_3^c &= \frac{(D-S)(D+S)(DS+3Z)}{4D}, \\
k_4 &= -\frac{(D-S)(D+S)(D^4-3(D^2S^2+4DSZ+2Z^2))}{8D^2}, \\
k_5 &= -\frac{(D-S)(D+S)(2D^4S-3D^2S^3+5D^3Z-15DS^2Z-15SZ^2)}{4D^2},
\end{aligned}$$

et nous avons :

$$P_\sigma(X) = \frac{D^2-S^2-4X}{1024D^6}\times$$
$$\left[\begin{array}{c}
-2048\, D^6X^8 \\
-512\, D^6\, \left(D^2-S^2\right) X^7 \\
+256\, D^4\, \left(D^2-S^2\right)\left(3\, D^4-10\, D^2\, S^2-42\, D\, S\, Z-21\, Z^2\right) X^6 \\
+64\, D^4 \left(D^2-S^2\right)^2 \left(3\, D^4-7\, D^2\, S^2-24\, D\, S\, Z+6\, Z^2\right) X^5 \\
-8D^2\left(D^2-S^2\right)^2 \left[\begin{array}{c} D^4\left(9D^4-166D^2S^2+265S^4\right)-168D^3\left(5D^2S-13S^3\right)\, Z \\ -24D^2\left(38D^2-233S^2\right)\, Z^2+4320DSZ^3+540Z^4 \end{array}\right] X^4 \\
-2D^2\left(D^2-S^2\right)^2 \left[\begin{array}{c} D^4\left(9D^6+165D^4S^2-737D^2S^4+611S^6\right) \\ +48D^3S\left(21D^4-138D^2S^2+137S^4\right) Z \\ +12D^2\left(135D^4-1704D^2S^2+2089S^4\right) Z^2 \\ -576DS\left(42D^2-67S^2\right) Z^3+108\left(-69D^2+169S^2\right) Z^4 \end{array}\right] X^3 \\
+(D^2-S^2)^3(D\,S+3\,Z)^2 \left[\begin{array}{c} D^4(21D^4-254D^2S^2+317S^4) \\ +24D^3S(-49D^2+95\,S^2)\,Z \\ +12D^2(-93\,D^2+379S^2)Z^2+1584DSZ^3-108Z^4 \end{array}\right] X^2 \\
+8\left(D^2-S^2\right)^4 (DS+3Z)^4 \left(2D^4-5D^2S^2-18DSZ-3Z^2\right) X \\
+2(D^2-S^2)^5(DS+3Z)^6
\end{array}\right].$$

On remarque alors que $-x_1x_2 = M_2^c - \sigma$ est la racine du polynôme qui permet l'inversion du système. Comme $-x_1x_2 = \frac{D^2-S^2}{4}$, $-x_1x_2$ est une racine double si le reste, noté r, de la division polynomiale de $P_\sigma(X)$ par $(X+x_1x_2)^2 = \frac{1}{16}\left(4X-D^2+S^2\right)^2$ est nul. Si l'on note

$$F_3 = D^4 + 6\, D\, S\, Z + 27\, Z^2,$$

et

$$F_4 = D^8 + 18\, D^4\, Z^2 - 135\, Z^4,$$

ce reste s'écrit :

$$r = -\frac{(D-S)^5\ (D+S)^5(D^2-S^2-4\ X)}{4096D^6}\ F_3 \times F_4.$$

Le premier facteur, F_3, de ce produit ne s'annule que pour $x_1 = x_2$ et $\sigma_1^2 = \sigma_2^2$. En effet, le discriminant du polynôme en Z est $\Delta = 36D^2\left(|S| + \sqrt{3}\,|D|\right)\left(|S| - \sqrt{3}\,|D|\right)$. Comme $|D| = |x_1| + |x_2|$ et $|S| = ||x_1| - |x_2||$, on aura $|D| \geqslant |S|$, et donc $\Delta \leqslant 0$, avec un cas de nullité si et seulement si $x_1 = x_2 = 0$. Le facteur F_3 ne s'annule, par conséquent, que dans le cas où les deux distributions sont identiques, qui n'entre pas dans le cadre des mélanges. Le second facteur F_4 s'annule si et seulement si l'on se place sur la variété V_2. La racine de P_σ permettant d'estimer les paramètres est donc effectivement une racine double lorsque l'ensemble des paramètres appartient à V_2. C'est, de plus, le seul endroit où cela peut se produire.⧫

Remerciements. J'adresse mes sincères remerciements à Monsieur Pierre-Loti-Viaud.

RÉFÉRENCES BIBLIOGRAPHIQUES

[1] M. Aitkin, D. B. Rubin, *Estimation and hypothesis testing in finite mixture models.* J. Roy. Stat. Soc., **B. 47** (1985), 67-75.

[2] A.A. Borovkov, *Statistique Mathématique*. Mir, Moscou, 1987.

[3] A.C. Cohen, *Estimation in mixtures of two normal distributions.* Technometrics **9** No 1. (1967), 15-28.

[4] A.P. Dempster, N.M. Laird, D.B. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*. J. Roy. Stat. Soc., **B. 39** (1977), 1-38.

[5] B.S. Everitt, D.J. Hand, *Finite Mixture Distributions.* Chapman and Hall, London, 1981.

[6] S. Fiorin, *Inconsistency for roots of likelihood equations which are relative maxima of the likelihood function.* Rapport technique ISUP-LSTA, 2001-**6**.

[7] J. G. Fryer, C. A. Robertson, *The bias and accuracy of moment estimators.* Biometrika, **57** (1970), 57-65.

[8] J. G. Fryer, C. A. Robertson, *A comparison of some methods for estimating mixed normal distributions.* Biometrika, **59** (1972), 639-648.

[9] M.G. Kendall, A. Stuart, J.K. Ord, *Kendall's Advanced Theory of Statistics, fifth edition of Vol1 - Distribution theory*. Charles Griffin & Company, Limited, London, 1987.

[10] B.G. Lindsay, *On the determinant of moment matrices*. Ann.Statist., **17**, (1989), 711-721.

[11] B.G. Lindsay, *Moment matrices : applications in mixtures*. Ann. statist., **17**, (1989), 722-740.

[12] B.G. Lindsay, *Mixture Models : Theory, geometry and Applications*. NSFCBMS Regional Conference Series in Probability and Statistics, **vol.5**, IMS, Hayward, CA, 1995.

[13] G. McLachlan, D. Pell, *Finite Mixture Models.* John Wiley & Sons, New-York, 2000.

[14] E. Monfrini, *Identifiabilité et méthode des moments pour les mélanges de distributions du système de Pearson.* Thèse de doctorat, Université Paris 6, 2002.

[15] K. Pearson, *Contribution to the mathematical theory of evolution*. Phil.Trans.Roy. Soc., **A. 185** (1895), 71-110.

[16] R. Redner, H.F. Walker, *Mixture densities, maximum likelihood and the E.M. algorithm*. SIAM Review, **26** (1984), 195-239.

[17] H. Teicher, *Identifiability of finite mixtures*. Ann. Math. Statist., **34** (l963), 1265-1269.

[18] D.M. Titterington, A.F.M. Smith, U.E. Makov, *Statistical Analysis of Finite Mixture Distributions*. Wiley, New York, 1985.