

ANNALES DE L'I. H. P., SECTION B

B. GAREL

Détection de plusieurs valeurs aberrantes dans un échantillon gaussien multidimensionnel

Annales de l'I. H. P., section B, tome 15, n° 1 (1979), p. 51-62

http://www.numdam.org/item?id=AIHPB_1979__15_1_51_0

© Gauthier-Villars, 1979, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P., section B » (<http://www.elsevier.com/locate/anihpb>) 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/>

Détection de plusieurs valeurs aberrantes dans un échantillon gaussien multidimensionnel

par

B. GAREL (*)

Laboratoire de probabilités et statistique n° 7, associé au C. N. R. S.,
B. P. 53, 38041 Grenoble Cedex

RÉSUMÉ. — Nous étudions ici le cas d'un échantillon gaussien multidimensionnel dont la matrice des variances-covariances Λ est inconnue. Nous recherchons des tests de détection de plusieurs valeurs aberrantes. Ces tests vérifient une propriété d'optimalité dans le cadre d'un modèle de glissements de la moyenne colinéaires. Le cas où l'on connaît la matrice Λ a fait l'objet d'une étude précédente dans [2].

SUMMARY. — Let a multinormal sample, the variance-covariance matrix of which is unknown. We are looking for tests of detection of outliers. A property of optimality of these tests is shown when the outlying observations have slipped in mean in parallel directions.

INTRODUCTION

Soit x_1, \dots, x_n un n -échantillon gaussien multidimensionnel d'une loi normale $\mathcal{N}_p(m, \Lambda)$ à valeurs dans \mathbb{R}^p de moyenne m et de matrice des variances-covariances Λ toutes deux inconnues. Nous adoptons pour notre étude le modèle \mathcal{A} , qui généralise à plusieurs dimensions l'un des deux

(*) Adresse professionnelle et pour toute correspondance : Faculté des Sciences et des Techniques, B. P. 143, 73011 Chambéry Cedex.

modèles proposés par Dixon et Grubbs en dimension 1. Ces modèles sont décrits dans [7]. Nous supposons donc qu'une valeur aberrante provient d'un glissement de la moyenne, soit d'une loi $\mathcal{N}_p(m + \Delta m, \Lambda)$, $\Delta m \in \mathbb{R}^p$, $\Delta m \neq 0$. Lorsque la matrice des variances-covariances Λ est connue, les tests correspondants ont été proposés dans [2].

Si $x_\mu = \begin{pmatrix} x_{\mu 1} \\ \vdots \\ x_{\mu p} \end{pmatrix}$ désigne un vecteur de \mathbb{R}^p et S une matrice $p \times p$, on notera

x'_μ le vecteur transposé de x_μ et S' la matrice transposée. On notera indifféremment une variable aléatoire et une réalisation de cette variable aléatoire.

De plus, on notera \bar{x} le vecteur moyenne et S la matrice des variances-covariances estimés sans biais à l'aide de cet échantillon, soit

$$\bar{x} = \frac{1}{n} \sum_{\mu=1}^n x_\mu \quad , \quad S = \frac{1}{n-1} \sum_{\mu=1}^n (x_\mu - \bar{x})(x_\mu - \bar{x})'$$

Enfin, on posera :

$$\underline{x} = (x_1, \dots, x_n) \quad \text{et} \quad \underline{dx} = dx_1 \dots dx_n$$

1. FORMULATION DU PROBLÈME ET THÉORÈME PRINCIPAL

Soient x_1, \dots, x_n n variables aléatoires gaussiennes à valeurs dans \mathbb{R}^p de moyenne m_1, \dots, m_n respectivement et de même matrice des variances-covariances Λ inconnues. Nous voulons tester l'hypothèse

$$H_0 = H(m_1 = \dots = m_n = m)$$

contre les hypothèses alternatives que k ($k \leq n$) moyennes des variables ont glissé suivant des vecteurs de \mathbb{R}^p colinéaires à un même vecteur $\Delta m \in \mathbb{R}^p$, k étant connu.

Nous noterons σ le groupe des permutations de l'ensemble $\{1, \dots, n\}$. Soit $\sigma \in \sigma$, l'hypothèse H_σ est que

$$(1-0) \quad \left\{ \begin{array}{l} E(X_1) = m + \alpha_{\sigma(1)} \Delta m \\ E(X_2) = m + \alpha_{\sigma(2)} \Delta m \\ \vdots \\ E(X_n) = m + \alpha_{\sigma(n)} \Delta m \end{array} \right.$$

où $\alpha_1, \dots, \alpha_n$ sont des scalaires réels connus. Nous avons donc $n!$ hypothèses alternatives. Les α_μ pour $\mu = 1, \dots, n$ ne sont pas nécessairement

distincts et certains peuvent être nuls, l'un au moins étant différent de zéro. On donnera, par exemple, un résultat dans la situation $k = 2$, $\alpha_1 = \alpha_2 = 1$, $\alpha_\mu = 0$ pour $\mu = 3, \dots, n$.

Nous utilisons la même méthode que dans [2] et formulons notre problème comme un problème de décision multiple dont voici les principaux éléments

1° Soit \mathcal{E} l'ensemble des états de la nature. Nous identifions \mathcal{E} à l'espace produit $\mathcal{H} \times \mathbb{R}^p$ où \mathcal{H} est l'ensemble discret des $r + 1$ hypothèses : $\mathcal{H} = \{H_0, H_1, \dots, H_r\}$ et où p est la dimension de l'espace d'un paramètre réel. \mathbb{R}^p est muni de sa tribu des boréliens, \mathcal{H} de son ensemble des parties $\mathcal{P}(\mathcal{H})$ et \mathcal{E} de la tribu produit.

2° Nous disposons sur \mathcal{E} d'une probabilité *a priori* que nous notons $b \otimes \eta$ où b est une probabilité discrète « *a priori* » sur \mathcal{H} et η une probabilité « *a priori* » absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^p : nous noterons $b[H_v] = b_v$, la probabilité « *a priori* » de l'hypothèse H_v , $v = 0, 1, \dots, r$. On a donc

$$\sum_{v=0}^r b_v = 1$$

3° Nous noterons \mathcal{D} l'ensemble discret des décisions possibles et de même cardinal que \mathcal{H} . $\mathcal{D} = \{D_0, D_1, \dots, D_r\}$ où D_v représente la décision d'accepter l'hypothèse H_v , pour $v = 0, 1, \dots, r$.

4° Nous noterons $f_v(x_1, \dots, x_n)$ la densité conjointe des x_μ pour $\mu = 1, \dots, n$ sous l'hypothèse H_v , pour $v = 0, 1, \dots, r$. De plus les f_v pour $v = 0, 1, \dots, r$ dépendent d'un paramètre vectoriel Δm à valeurs dans \mathbb{R}^p sur lequel on dispose justement de la probabilité « *a priori* » η .

5° Nous noterons \mathbb{D} l'ensemble des fonctions (ou règles) de décision aléatoires

$$\mathbb{D} = \{d/d : (\mathbb{R}^p)^n \rightarrow \mathcal{M}\}$$

où \mathcal{M} est un ensemble de probabilités sur \mathcal{D} . On peut identifier d (voir [8]) à une application $d : (\mathbb{R}^p)^n \rightarrow [0, 1]^{r+1}$

$$d = (d_0, d_1, \dots, d_r)$$

avec

$$0 \leq d_v(x_1, \dots, x_n) \leq 1 \quad \text{et} \quad \sum_{v=0}^r d_v(x_1, \dots, x_n) = 1.$$

$d_v(x_1, \dots, x_n)$ représente la probabilité de prendre la décision D_v sachant (x_1, \dots, x_n) . De plus, pour $v = 0, 1, \dots, r$, nous supposons que $d_v(\underline{x})$ est mesurable par rapport à la mesure de Lebesgue sur $\mathbb{R}^{n \cdot p}$.

6° Nous appellerons L la fonction de coût. L est une application de $\mathcal{H} \times \mathcal{D} \rightarrow \mathbb{R}^+$. $L(D_v | H_\xi)$ est le coût de la décision D_v lorsque l'état H_ξ est réalisé. Dans ce qui suit, nous prendrons la fonction L définie par

$$L(D_v | H_\xi) = \begin{cases} 0 & \text{si } v = \xi \\ 1 & \text{si } v \neq \xi \end{cases}$$

Il est alors possible de définir le risque de Bayes et les règles de Bayes relatifs à la probabilité *a priori* $b \otimes \eta$, et de caractériser ces dernières. Nous renvoyons pour une étude détaillée à [I]. Notons que dans le cas de glissements colinéaires décrit ci-dessus $r = n!$. On notera D_0 (resp. D_σ) la décision de considérer H_0 (resp. H_σ) comme réalisée, $d_0(\underline{x})$ (resp. $d_\sigma(\underline{x})$) la probabilité de prendre la décision D_0 (resp. D_σ) sachant \underline{x} .

Nous définissons sur \mathcal{H} la probabilité *a priori* b telle que

$$\begin{aligned} b[H_0] &= b_0 \\ b[H_\sigma] &= \frac{1 - b_0}{n!} \end{aligned}$$

Soit $c > 0$ une constante fixée telle que $\|\Delta m\|^2 = c$ où $\|\cdot\|$ désigne la norme euclidienne de \mathbb{R}^p .

Nous définissons la probabilité *a priori* $\eta_c(\Delta m)$ comme la mesure uniforme sur la sphère de \mathbb{R}^p d'équation $\|\Delta m\|^2 = c$.

On choisit de se limiter aux règles de décision vérifiant les propriétés suivantes.

1.1. La probabilité de choisir D_0 sous l'hypothèse H_0 est

$$P(D_0 | d, H_0) = 1 - a,$$

niveau de confiance du test. Donc a est fixé et ne dépend pas de d .

1.2. Les règles de décision sont invariantes lorsqu'on ajoute un vecteur constant $y \in \mathbb{R}^p$ à chaque observation.

$$\begin{aligned} \forall y \in \mathbb{R}^p \quad d_0(x_1 + y, \dots, x_n + y) &= d_0(x_1, \dots, x_n) \\ \text{et} \quad \forall \sigma \in \sigma \quad d_\sigma(x_1 + y, \dots, x_n + y) &= d_\sigma(x_1, \dots, x_n) \end{aligned}$$

Les règles de décision auxquelles on se limite sont donc invariantes par le sous-groupe noté G_1 , et défini ci-dessus, du groupe des translations de $\mathbb{R}^{n \cdot p}$.

1.3. Les règles de décision sont invariantes lorsqu'on multiplie les observations par une matrice $p \times p$, notée B , inversible.

$$\begin{aligned} \forall B \in Gl(p, \mathbb{R}), \quad d_0(Bx_1, \dots, Bx_n) &= d_0(x_1, \dots, x_n) \\ \text{et} \quad \forall \sigma \in \sigma, \quad d_\sigma(Bx_1, \dots, Bx_n) &= d_\sigma(x_1, \dots, x_n) \end{aligned}$$

Les règles de décision auxquelles on se limite sont donc invariantes par le sous-groupe noté G_2 et défini ci-dessus, du groupe linéaire de $\mathbb{R}^{n-p} Gl(n, p, \mathbb{R})$.

1.4. Enfin les règles de décision doivent vérifier : les probabilités de décision $P(D_\sigma | d, (H_\sigma, \Delta m))$ de choisir D_σ sous l'hypothèse H_σ sont indépendantes de $\sigma \in \sigma$ et ne dépendent que de $\| \Delta m \|$.

Il résulte des conditions 1.2 et 1.3 que les règles de décision auxquelles on se limite sont invariantes par le groupe noté G_3 engendré par G_1 et G_2 .

Nous utilisons par la suite la densité de l'invariant maximal relatif au groupe G_3 qui nous est donné par une généralisation à des observations de \mathbb{R}^p des résultats de [4], p. 213-218, généralisation que l'on peut trouver dans [1], p. 53-56. Si l'on note $z_\mu = x_\mu - x_1$ pour $\mu = 1, \dots, n$, B la matrice $p \times p$ formée des p vecteurs-colonnes (z_{n-p+1}, \dots, z_n) et $y_\mu = B^{-1}z_\mu$ pour $\mu = 1, \dots, n$, l'invariant maximal relatif au groupe G_3 nous est donné par (y_2, \dots, y_{n-p}) . Nous allons déduire la densité de l'invariant maximal relatif au groupe G_3 , sous les diverses hypothèses H_σ , du résultat ci-dessous. Soit \mathcal{I}_p la matrice identité de \mathbb{R}^p .

1.5. LEMME. — Soient x_1, \dots, x_n n variables aléatoires gaussiennes de lois respectives $\mathcal{N}_p(m_\mu, \mathcal{I}_p)$ pour $\mu = 1, \dots, n$. La densité de la loi conjointe des $n - p - 1$ variables $(y_2, \dots, y_{n-p}) = \underline{y}$ nous est donnée par

$$(1-6) \quad h(\underline{y}) = C_1 \exp \left\{ -\frac{1}{2} \sum_{\mu=1}^n \| m_\mu \|^2 + \frac{1}{2n} \sum_{i=1}^p \left(\sum_{\mu=1}^n m_{\mu i} \right)^2 \right\} \\ \int \exp \left\{ -\frac{1}{2} \sum_{\mu=1}^n (y_\mu - \bar{y})' B' B (y_\mu - \bar{y}) + \sum_{\mu=1}^n (y_\mu - \bar{y})' B' m_\mu \right\} | \det B |^{n-p-1} dB$$

où

$$C_1 = \frac{1}{n^{p/2} (2\pi)^{(n-1)p/2}} ; \quad \bar{y} = \frac{1}{n} \sum_{\mu=1}^n y_\mu$$

Preuve. — La densité $f(\underline{z})$ de la loi conjointe de $(z_2, \dots, z_n) = \underline{z}$ nous est fournie par le lemme 1.7 de [1] avec $\Lambda = \mathcal{I}_p$ et en remplaçant x_n par x_1 , soit :

$$f(\underline{z}) = C_1 \exp \left\{ -\frac{1}{2} \sum_{\mu=1}^n (z_\mu - \bar{z})' (z_\mu - \bar{z}) + \sum_{\mu=1}^n (z_\mu - \bar{z})' m_\mu \right. \\ \left. - \frac{1}{2} \sum_{\mu=1}^n m'_\mu m_\mu + \frac{1}{2n} \sum_{i=1}^p \left(\sum_{\mu=1}^n m_{\mu i} \right)^2 \right\}$$

Les vecteurs z_{n-p+1}, \dots, z_n étant stochastiquement indépendants, la matrice B définie plus haut est presque sûrement inversible.

Remarquons que pour $\mu = 1, \dots, p$, $B^{-1}(z_{n-p+\mu}) = e_\mu$ où e_μ est le $\mu^{\text{ième}}$ vecteur de base de \mathbb{R}^p .

Posons donc pour $\mu = 1, \dots, p$, $y_{n-p+\mu} = e_\mu$.

Opérons la transformation

$$y_\mu = B^{-1}z_\mu \quad \text{pour } \mu = 1, \dots, n-p$$

$$v_\mu = z_{n-p+\mu} \quad \text{pour } \mu = 1, \dots, p$$

soit

$$z_\mu = By_\mu \quad \text{pour } \mu = 1, \dots, n \quad \text{et} \quad \bar{z} = B\bar{y}$$

Le jacobien de cette transformation vaut $|\det B|^{n-p-1}$ et l'on obtient par intégration par rapport à $B = (v_1, \dots, v_p)$

$$h(\underline{y}) = C_1 \int_{\mathbb{R}^{p^2}} \exp \left\{ \frac{1}{2} \sum_{\mu=1}^n (y_\mu - \bar{y})' B' B (y_\mu - \bar{y}) \right. \\ \left. + \sum_{\mu=1}^n (y_\mu - \bar{y})' B' m_\mu - \frac{1}{2} \sum_{\mu=1}^n m'_\mu m_\mu + \frac{1}{2} \sum_{i=1}^p \left(\sum_{\mu=1}^n m_{\mu i} \right)^2 \right\} |\det B|^{n-p-1} dB$$

d'où le résultat.

Pour rendre utilisable la densité de l'invariant maximal sous les diverses hypothèses, nous introduisons les notations suivantes. Soit \mathcal{S} la matrice $p \times p$ symétrique inversible presque sûrement définie par

$$\mathcal{S} = \sum_{\mu=1}^n (y_\mu - \bar{y})(y_\mu - \bar{y})'$$

Soit $\mathcal{S}^{1/2}$ la matrice symétrique définie ≥ 0 vérifiant $\mathcal{S}^{1/2} \mathcal{S}^{1/2} = \mathcal{S}$.

Soit D la matrice telle que $D = B\mathcal{S}^{1/2}$. Posons :

$$(1-7) \quad \forall \sigma \in \sigma, \tilde{\mathcal{C}}_\sigma = \mathcal{S}^{-1/2} \left[\sum_{\mu=1}^n \alpha_{\sigma(\mu)} (y_\mu - \bar{y}) \right]$$

et

$$(1-8) \quad T_\sigma^2 = \left[\sum_{\mu=1}^n \alpha_{\sigma(\mu)} (x_\mu - \bar{x}) \right]' S^{-1} \left[\sum_{\mu=1}^n \alpha_{\sigma(\mu)} (x_\mu - \bar{x}) \right]$$

Alors on vérifie que

$$(1-9) \quad \sum_{\mu=1}^n (y_\mu - \bar{y})' B' B (y_\mu - \bar{y}) = \text{tr}(DD')$$

où $\text{tr}(\cdot)$ désigne la trace de la matrice, et que

$$(1-10) \quad \tilde{\sigma}'_o \tilde{\sigma}_o = \|\tilde{\sigma}_o\|^2 = T_o^2/n - 1$$

Posons enfin :

$$C_2 = \frac{1}{2n} \left[n \sum_{\mu=1}^n \alpha_{\sigma(\mu)}^2 - \left(\sum_{\mu=1}^n \alpha_{\sigma(\mu)} \right)^2 \right]$$

1.11. LEMME. — Sous l'hypothèse H_σ la densité de l'invariant maximal relatif au groupe G_3 nous est donnée par :

$$(1-12) \quad h_\sigma(\underline{y}, \Delta m) = C_1 |\det \mathcal{S}|^{-\frac{n-1}{2}} \exp \{ -C_2 \|\Delta m\|^2 \} \\ \times \int \exp \left\{ -\frac{1}{2} (\text{tr}(DD') + \tilde{\sigma}'_o D' \Delta m) \right\} |\det D|^{n-p-1} dD$$

Preuve. — Sous l'hypothèse H_σ , pour $\mu = 1, \dots, n$, $m_\mu = \alpha_{\sigma(\mu)} \Delta m$. En remplaçant dans l'expression (1.6) m_μ par sa valeur et en tenant compte de nos notations et de la propriété (1.9), on obtient le résultat indiqué.

1.13. LEMME. — Sous l'hypothèse H_0 , la densité de l'invariant maximal relatif au groupe G_3 nous est donnée par

$$(1-14) \quad h_0(\underline{y}) = C_1 |\det \mathcal{S}|^{-\frac{n-1}{2}} \int \exp \left\{ -\frac{1}{2} \text{tr}(DD') \right\} |\det D|^{n-p-1} dD$$

Preuve. — Sous l'hypothèse H_0 , $m_\mu = 0$. On obtient le résultat en remplaçant Δm par 0 dans l'expression (1.6).

Nous pouvons maintenant démontrer le résultat suivant.

1.15. THÉORÈME. — Dans le cadre du modèle décrit ci-dessus, la règle de décision vérifiant les propriétés 1.1, 1.2, 1.3, 1.4 qui maximise les probabilités de décision $P(D_\sigma | d, (H_\sigma, \Delta m))$ uniformément en $\Delta m \in \mathbb{R}^p$ et $\sigma \in \sigma$, nous est donnée par :

$$\left| \begin{array}{l} \text{Si } T_{(\sigma)}^2 < \lambda_a \text{ prendre la décision } D_0 \\ \text{Si } T_{(\sigma)}^2 > \lambda_a \text{ prendre la décision } D_{(\sigma)} \end{array} \right.$$

où (σ) est la permutation pour laquelle T_σ^2 est maximal, où T_σ^2 nous est donné par la formule (1.8), où λ_a est une constante qui dépend du niveau de signification, donc de a avec $P_{H_0}(T_{(\sigma)}^2 < \lambda_a) = 1 - a$.

1.16. REMARQUE. — La probabilité $P(D_\sigma | d, (H_\sigma, \Delta m))$ étant indépendante de $\sigma \in \sigma$ on l'appelle la probabilité de détection de k valeurs aberrantes

et l'on dit que la règle du théorème 1.15 maximise uniformément en Δm la probabilité de détection de k valeurs aberrantes éventuelles.

Preuve du théorème. — Les probabilités de décision d'une règle de décision vérifiant l'invariance par le groupe G_1 sont indépendantes de m , la moyenne inconnue. Nous supposons donc $m = 0$.

Les probabilités de décision d'une règle de décision vérifiant l'invariance par le groupe G_2 sont indépendantes de Λ , la matrice des variances-covariances inconnue. Nous supposons donc $\Lambda = \mathcal{I}_p$.

Alors (voir [1]) parmi les règles de décision possédant les propriétés 1.1, 1.2, 1.3, et 1.4, celles qui maximisent uniformément en $\sigma \in \sigma$ et $\Delta m \in \mathbb{R}^p$ les probabilités de décision $P(D_\sigma | d_\sigma, (H_\sigma, \Delta m))$ sont les règles de Bayes relatives à la probabilité *a priori* $b \otimes n_c$. D'autre part, les règles de Bayes relatives à la probabilité *a priori* $b \otimes \eta_c$ possédant les propriétés 1.1, 1.2 et 1.3, c'est-à-dire de niveau de signification $1 - \alpha$ et invariantes par le groupe de transformations G_3 sont données par

$$(1-17) \quad \left\{ \begin{array}{ll} d_0(\underline{x}) = 1 & \text{si} \quad Ah_0 > \int h_\sigma(\underline{y}, \Delta m) d\eta_c(\Delta m), \forall \sigma \in \sigma \\ d_\sigma(\underline{x}) = 1 & \text{si} \quad \int h_\sigma(\underline{y}, \Delta m) d\eta_c(\Delta m) > Ah_0 \\ & \text{et} \quad \int h_\sigma(\underline{y}, \Delta m) d\eta_c(\Delta m) > \int h_{\tilde{\sigma}}(\underline{y}, \Delta m) d\eta_c(\Delta m) \\ & \qquad \qquad \qquad \forall \tilde{\sigma} \in \sigma \quad \text{avec} \quad \tilde{\sigma} \neq \sigma \end{array} \right.$$

sauf peut-être sur un sous-ensemble E de $\mathbb{R}^{n \cdot p}$ de mesure nulle. h_0 (resp. h_σ) est la densité de l'invariant maximal relatif au groupe G_3 sous l'hypothèse H_0 (resp. H_σ).

A est une constante de \mathbb{R}^{+*} qui dépend de α .

Par un sous-ensemble $E \subset \mathbb{R}^{n \cdot p}$ de mesure nulle, nous voulons dire un ensemble E tel que

$$\iint_E f_\sigma(\underline{x}, \Delta m) d\underline{x} d(b \otimes \eta_c) = 0 ;$$

Nous allons donc calculer les intégrales de la formule (1.17), soit :

$$\begin{aligned} I_\sigma &= \int_{\|\Delta m\|^2=c} h_\sigma(\underline{y}, \Delta m) d\eta_c(\Delta m) \\ &= c_1 |\det \mathcal{S}|^{-\frac{n-1}{2}} e^{-c \times c_2} \\ &\quad \times \int_{\|\Delta m\|^2=c} \exp \left\{ -\frac{1}{2} \text{tr} (DD') + \tilde{c}'_\sigma D' \Delta m \right\} |\det D|^{n-p-1} dD d\eta_c(\Delta m) \end{aligned}$$

Posons

$$J_{\sigma} = \int \exp \left\{ -\frac{1}{2} \text{tr} (DD') + \tilde{c}'_{\sigma} D' \Delta m \right\} |\det D|^{n-p-1} dD$$

Effectuons une rotation R de telle sorte que

$$R \tilde{c}_{\sigma} = \|\tilde{c}_{\sigma}\| \cdot e_1 = \frac{1}{\sqrt{n-1}} T_{\sigma} \cdot e_1,$$

d'après la propriété (1.10).

Définissons la matrice $\Gamma = RD'$.

Alors on a :

$$J_{\sigma} = \int \exp \left\{ -\frac{1}{2} \text{tr} (\Gamma' \Gamma) + \frac{1}{\sqrt{n-1}} T_{\sigma} e'_1 \Gamma \Delta m \right\} |\det \Gamma|^{n-p-1} d\Gamma.$$

On va montrer que J_{σ} est fonction croissante de T_{σ}^2 . Écrivons l'exponentielle sous la forme

$$\exp \left\{ -\frac{1}{2} \text{tr} (\Gamma' \Gamma) \right\} \exp \left\{ \frac{1}{\sqrt{n-1}} T_{\sigma} \times (\Gamma \Delta m)_1 \right\}$$

où $(\Gamma \Delta m)_1$ est la 1^{re} composante du vecteur $\Gamma \Delta m$

$$= \exp \left\{ -\frac{1}{2} \text{tr} (\Gamma' \Gamma) \right\} \left(\sum_{l=0}^{+\infty} \frac{\left(\frac{1}{\sqrt{n-1}} T_{\sigma} \right)^l}{l!} (\Gamma \Delta m)_1^l \right)$$

ceci en utilisant le développement en série de l'exponentielle.

Intégrons, il vient

$$J_{\sigma} = \sum_{l=0}^{+\infty} C_l \left(\frac{1}{\sqrt{n-1}} T_{\sigma}^l \right)$$

où

$$C_l = \frac{1}{l!} \int \exp \left\{ -\frac{1}{2} \text{tr} (\Gamma' \Gamma) \right\} (\Gamma \Delta m)_1^l |\det \Gamma|^{n-p-1} d\Gamma$$

Effectuons le changement de variable de Γ en $-\Gamma$, alors

$$C_{2l+1} = \frac{1}{2l+1!} \int \exp \left\{ -\frac{1}{2} \text{tr} (\Gamma' \Gamma) \right\} (-\Gamma \Delta m)_1^{2l+1} |\det \Gamma|^{n-p-1} d\Gamma = -C_{2l+1}$$

d'où $C_{2l+1} = 0$ pour tout $l \in \mathbb{N}$. J_σ est donc la somme des coefficients pairs :

$$J_\sigma = \sum_{l=0}^{+\infty} C_{2l} \left(\frac{1}{n-1} T_\sigma^2 \right)^l$$

d'où il résulte que J_σ est une fonction croissante de T_σ^2 .

Il en est de même pour I_σ et l'on a :

$$Ah_0 > I_\sigma \Leftrightarrow T_\sigma^2 < \lambda_a \quad \text{pour tout } \sigma \in \sigma.$$

Choisir $\lambda_a > 0$ de telle sorte que la propriété 1.1 soit vérifiée revient à déterminer $A \in \mathbb{R}^{+*}$. D'autre part, pour σ et $\tilde{\sigma} \in \sigma$

$$I_\sigma > I_{\tilde{\sigma}} \Leftrightarrow J_\sigma > J_{\tilde{\sigma}} \Leftrightarrow T_\sigma^2 > T_{\tilde{\sigma}}^2$$

Relativement à la distribution *a priori* $\eta_c(\Delta m)$ les règles de Bayes vérifiant les propriétés 1.1, 1.2, 1.3 sont caractérisées par

$$(1-18) \quad \left| \begin{array}{ll} d_0(\underline{x}) = 1 & \text{si } T_{(\sigma)}^2 = \max_{\sigma \in \sigma} T_\sigma^2 < \lambda_a \\ d_{(\sigma)}(x) = 1 & \text{si } T_{(\sigma)}^2 > \lambda_a \end{array} \right.$$

Cette règle de décision est manifestement symétrique en x_1, \dots, x_n . Les densités des lois conjointes des variables x_1, \dots, x_n sous les diverses hypothèses sont également symétriques en x_1, \dots, x_n .

Enfin pour tout $\sigma \in \sigma$ on vérifie aisément que les probabilités de décision $P(D_\sigma | d, (H_0, \Delta m))$ de la règle ci-dessus ne dépendent que de $\|\Delta m\|$. La règle de décision (1.18) possède donc la propriété 1.4, ce qui termine la démonstration du théorème 1.15.

2. EXEMPLES D'APPLICATIONS

Nous donnons ici trois corollaires qui nous permettent de retrouver le résultat connu et, démontré dans [3] et de généraliser à la dimension $p > 1$ certains résultats en dimension 1.

2.1. COROLLAIRE. — Sous l'hypothèse qu'il n'y a au plus qu'une valeur aberrante, dans le cadre du modèle décrit au paragraphe 1, la règle de décision possédant les propriétés 1.1, 1.2, 1.3 et 1.4 qui maximise la proba-

bilité de détection de cette valeur uniformément en $\Delta m \in \mathbb{R}^p$ nous est donnée par :

$$\left| \begin{array}{ll} \text{prendre la décision } D_0 & \text{si } T_{(n)}^2 = \max_{\mu=1\dots n} (x_\mu - \bar{x})'S^{-1}(x_\mu - \bar{x}) < \lambda_a \\ \text{prendre la décision } D_{(n)} & \text{si } T_{(n)}^2 > \lambda_a \end{array} \right.$$

Preuve. — Dans ce cas, un seul des α_μ est non nul, égal à 1 et il y a n hypothèses alternatives.

Ce résultat est déjà connu et dû à [3].

2.2. COROLLAIRE. — Supposons que $1 < k < n$ des α_μ soient non nuls, tous égaux à 1. Alors la règle de décision satisfaisant aux hypothèses du théorème 1.15 et possédant les propriétés 1.1, 1.2, 1.3 et 1.4 nous est donnée par :

$$\left| \begin{array}{ll} \text{prendre la décision } D_0 & \text{si } T_{(\sigma)}^2 < \lambda_a \\ \text{prendre la décision } D_{(\sigma)} & \text{si } T_{(\sigma)}^2 > \lambda_a \end{array} \right.$$

où

$$T_{(\sigma)}^2 = \max_{\sigma \in \sigma} \left[\sum_{\mu=1}^k (x_{\sigma(\mu)} - \bar{x}) \right]' S^{-1} \left[\sum_{\mu=1}^k (x_{\sigma(\mu)} - \bar{x}) \right]$$

et λ_a est une constante dépendant de a .

(La preuve de ce corollaire est identique à celle du corollaire 2.2 de [2] en remplaçant Λ^{-1} par S^{-1}).

Ce résultat généralise à la dimension p le résultat connu de Murphy [5] en dimension 1.

En particulier lorsque $k = 2$, $\alpha_1 = \alpha_2 = 1$, $\alpha_\mu = 0$ pour $\mu = 3, \dots, n$, on trouve le résultat simple suivant : le test de détection optimal est associé à la statistique :

$$\max_{1 \leq i < j \leq n} (x_i + x_j - 2\bar{x})'S^{-1}(x_i + x_j - 2\bar{x}).$$

Supposons enfin que les n observations soient partagées en r classes de k éléments : $n = rk$.

Nous nous proposons de savoir si l'une des classes a glissé en moyenne.

2.3. COROLLAIRE. — La règle de décision correspondante, vérifiant les hypothèses du théorème 1.15 et possédant les propriétés 1.1, 1.2, 1.3 et 1.4, nous est donnée par :

$$\text{prendre la décision } \left| \begin{array}{ll} D_0 & \text{si } T_{(l)}^2 < \lambda'_a \\ D_{(l)} & \text{si } T_{(l)}^2 > \lambda'_a \end{array} \right.$$

où (l) désigne l'indice pour lequel

$$\begin{aligned} T_l^2 &= (\bar{x}(l) - \bar{x})' S^{-1} (\bar{x}(l) - \bar{x}) \quad \text{est maximal ;} \quad 1 \leq l \leq r \\ \lambda'_a &= \lambda_a / k^2 \end{aligned}$$

où $\bar{x}(l)$ désigne la moyenne de la $l^{\text{ième}}$ classe.

(La preuve de ce corollaire est identique à celle du corollaire 2.3 de [2] en remplaçant Λ^{-1} par S^{-1}).

Ce résultat généralise à la dimension p le résultat bien connu de Paulson [6] en dimension 1.

Bien que l'on ne connaisse pas Δm , on a supposé les α_μ connus. Nous pouvons donc dire qu'en l'absence de toute information sur la nature des glissements, il n'y a pas de règle optimale de détection de $k > 1$ valeurs aberrantes.

BIBLIOGRAPHIE

- [1] GAREL, Détection des valeurs aberrantes dans un échantillon gaussien multidimensionnel. *Thèse de 3^e cycle*, Grenoble, 1976, p. 39-62.
- [2] GAREL, Tests de détection de valeurs aberrantes multidimensionnelles. *Annales de l'I. H. P.*, section B, vol. XIV, n^o 3, 1978, p. 305-316.
- [3] KARLIN, TRUAX, Slippage problems. *A. M. S.*, t. 31, 1960, p. 296-324.
- [4] LEHMAN, *Testing statistical hypothesis*, Wiley and Sons, New York, 1959.
- [5] MURPHY, On tests for outlying observations. *Ph. D. Thesis*, Princeton University, 1951.
- [6] PAULSON, An optimum solution to the k -sample slippage problem for the normal distribution. *A. M. S.*, t. 23, 1952, p. 610-616.
- [7] QUITTARD-PINON, Quelques méthodes d'analyse statistique des données aberrantes. *Thèse de 3^e cycle*, Paris VI, 1973.
- [8] WALD, *Statistical decision functions*, Chelsea. Publ. Company, New York, 1971.

(Manuscrit reçu le 3 octobre 1978)