Zeitschrift: L'Enseignement Mathématique
Herausgeber: Commission Internationale de l'Enseignement Mathématique
Band: 6 (1960)
Heft: 1: L'ENSEIGNEMENT MATHÉMATIQUE

Artikel: LES MODÈLES LINÉAIRES EN ANALYSE STATISTIQUE
Autor: Breny, H.
DOI: https://doi.org/10.5169/seals-36341

Nutzungsbedingungen
Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation
L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use
The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 15.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

LES MODÈLES LINÉAIRES EN ANALYSE STATISTIQUE

par H. Breny

(suite)

3. Cas ou rg $\mathfrak{A} = \dim \mathbf{B}$.

3, 1. *Remarques théoriques.*

3, 11. Lorsque $r = p$, $\mathfrak{G} \equiv \mathfrak{A}^T \mathfrak{A}$ est une matrice symétrique $p \times p$, régulière, dont on peut former l'inverse $\mathfrak{G}^{-1}$. Alors, de

$$\mathfrak{G}\, \hat{\mathfrak{b}}_H = \mathfrak{A}^T \mathfrak{x}$$

on tire

$$\hat{\mathfrak{b}}_H = \mathfrak{G}^{-1} \mathfrak{A}^T \mathfrak{x} \,,$$

et on en déduit

$$\begin{aligned} \mathbf{C}\, \hat{\mathfrak{b}}_H &= \mathbf{E} \left(\hat{\mathfrak{b}}_H - \mathfrak{b}_H\right) \left(\hat{\mathfrak{b}}_H - \mathfrak{b}_H\right)^T \\ &= \mathbf{E}\, \mathfrak{G}^{-1} \mathfrak{A}^T (\mathfrak{x} - \mathbf{E}\, \mathfrak{x}) (\mathfrak{x} - \mathbf{E}\, \mathfrak{x})^T \mathfrak{A}\, \mathfrak{G}^{-1} \,. \end{aligned}$$

Mais

$$\mathbf{E}\, (\mathfrak{x} - \mathbf{E}\, \mathfrak{x}) (\mathfrak{x} - \mathbf{E}\, \mathfrak{x})^T = \mathbf{C}\, \mathfrak{x} = \mathfrak{I}_n\, \sigma^2 \,,$$

donc

$$\mathbf{C}\, \hat{\mathfrak{b}}_H = \mathfrak{G}^{-1} \mathfrak{A}^T (\mathfrak{I}_n\, \sigma^2)\, \mathfrak{A}\, \mathfrak{G}^{-1} = (\mathfrak{G}^{-1} \mathfrak{A}^T \mathfrak{A}\, \mathfrak{G}^{-1})\, \sigma^2 = \mathfrak{G}^{-1} \sigma^2 \,. \quad (13)$$

Le calcul de $\mathfrak{G}^{-1}$ équivaut donc à celui de $\mathbf{C}\, \hat{\mathfrak{b}}_H$.

3, 12. Pour le calcul de **SCN**, il faut former les équations (7); ici, les $\mathfrak{u}_i^\star$ sont représentés par les lignes de $\mathfrak{A}^T$, $\mathfrak{c}_i^T$, de sorte que, aux inconnues près, le système (7) n'est autre que le système normal (donc, $\lambda_i = \hat{\mathfrak{b}}_{H,i}$); il résulte alors de (8) que

$$\begin{aligned} \mathbf{SCN} &= \sum_1^p \hat{\mathfrak{b}}_{H,i}\, \mathfrak{c}_i^T \mathfrak{x} = \hat{\mathfrak{b}}_H^T \mathfrak{A}^T \mathfrak{x} \\ &= (\mathfrak{x}^\star \mathfrak{A})\, \mathfrak{G}^{-1} (\mathfrak{A}^T \mathfrak{x}) \,. \quad (14) \end{aligned}$$

3, 13. **SCE** acquiert ici une signification très simple. En effet, si on pose

$$\mathfrak{x} = \mathfrak{A}\,\hat{\mathfrak{b}}_H + \mathfrak{r}$$

(de sorte que les composantes $r_{H,i}$ de r_H sont les « résidus »), on a, d'une part,

$$\mathfrak{A}^T\,\mathfrak{x} = \mathfrak{A}^T\,\mathfrak{A}\,\hat{\mathfrak{b}}_H\,,$$

et, d'autre part,

$$\mathfrak{A}^T\,\mathfrak{x} = \mathfrak{A}^T\,\mathfrak{A}\,\hat{\mathfrak{b}}_H + \mathfrak{A}^T\,\mathfrak{r}\;;$$

on a donc $\mathfrak{A}^T\,\mathfrak{r} = 0$; dès lors

$$\begin{aligned}\mathfrak{x}^\star\,\mathfrak{x} &= \left(\hat{\mathfrak{b}}_H^T\,\mathfrak{A}^T + \mathfrak{r}^T\right)\left(\mathfrak{A}\,\hat{\mathfrak{b}}_H + \mathfrak{r}\right)\\ &= \hat{\mathfrak{b}}_H^T\,\mathfrak{A}^T\,\mathfrak{A}\,\hat{\mathfrak{b}}_H + \mathfrak{r}^T\,\mathfrak{r} + \hat{\mathfrak{b}}_H^T\,\mathfrak{A}^T\,\mathfrak{r} + \mathfrak{r}^T\,\mathfrak{A}\,\hat{\mathfrak{b}}_H\\ &= \hat{\mathfrak{b}}_H^T\,\mathfrak{A}^T\,\mathfrak{x} + \mathfrak{r}^T\,\mathfrak{r} = \mathbf{SC}\mathrm{N} + \mathfrak{r}^T\,\mathfrak{r}\,,\end{aligned}$$

d'où, puisque $\mathfrak{x}^\star\,\mathfrak{x} = \mathbf{SCT}$,

$$\mathbf{SCE} = \mathfrak{r}^T\,\mathfrak{r} = \sum_1^n \mathfrak{r}_{H,i}^2\,.$$

3, 14. Si $\mathfrak{G}$ est diagonale (ce qui arrive si et seulement si les lignes de $\mathfrak{A}^T$ correspondent à des vecteurs deux à deux orthogonaux de V_+), les $\hat{\mathfrak{b}}_{H,i}$ sont deux à deux orthogonaux, et

$$\mathbf{SC}\mathrm{N} = \sum_1^p \mathrm{SC}\left\{\,\hat{\mathfrak{b}}_{H,i}\,\right\}.$$

3, 2. *Exécution des calculs.*

3, 21. La résolution des équations normales peut évidemment se faire par un procédé quelconque. On sait toutefois, depuis Benoît et Banachiewicz, que les procédés « compacts » habituels constituent tous des variations plus ou moins heureuses de la méthode de « factorisation triangulaire », particulièrement simple à appliquer dans le cas d'une matrice symétrique, comme l'est $\mathfrak{G}$ (cfr. [VI, VII]).

3, 22. Il s'agit, en principe, de trouver une matrice $\mathfrak{S}$, triangulaire supérieure, telle que

$$\mathfrak{G} = \mathfrak{S}^T\,\mathfrak{S} = \mathfrak{S} \times^0 \mathfrak{S}\,.$$

Le calcul de $\mathfrak{S}$ se fait ligne par ligne, suivant le schéma [13])

$$\begin{cases} (\mathfrak{S}\,\varepsilon_i) \times^0 (\mathfrak{S}\,\varepsilon_j) = \varepsilon_i^T\,\mathfrak{G}\,\varepsilon_j\ , \\ i = 1,\ j = 1, 2, \ldots;\ i = 2,\ j = 2, 3, \ldots;\ \text{etc.} \end{cases}$$

Si on désigne par $\mathfrak{g}_\Sigma$ la colonne des sommes des éléments des lignes de $\mathfrak{G}$, et par $\mathfrak{s}_\Sigma$ la colonne des sommes des éléments des lignes de $\mathfrak{S}$, on a

$$(\mathfrak{S}\,\varepsilon_i) \times^0 \sum_{j=1}^{p} (\mathfrak{S}\,\varepsilon_j) = \sum_{j=1}^{p} \varepsilon_i^T\,\mathfrak{G}\,\varepsilon_j$$

ou

$$(\mathfrak{S}\,\varepsilon_i) \times^0 \mathfrak{s}_\Sigma = \varepsilon_i^T\,\mathfrak{g}_\Sigma\ ;$$

chaque ligne de $\mathfrak{S}$ est ainsi vérifiée dès qu'elle est calculée.

3, 23. $\mathfrak{S}$ étant calculée, on a le choix, pour poursuivre, entre deux méthodes; la première fournit SCN et $\mathfrak{G}^{-1}$ (donc, à un facteur près, $\mathbf{C}\,\hat{\mathfrak{b}}_H$); la seconde fournit une décomposition de SCN en termes de la forme red $[\mathsf{U}_1^\star \mid \mathsf{U}_i^\star, \ldots, \mathsf{U}_{i-1}^\star]$, les $\mathsf{U}_i^\star$ étant orthogonaux et unidimensionnels.

3, 231. Dans la première méthode, on calcule $\mathfrak{G}^{-1}$ sans calculer $\mathfrak{S}^{-1}$, à partir de la relation $\mathfrak{G}^{-1} = \mathfrak{S}^{-1}\,(\mathfrak{S}^T)^{-1}$, qui donne

$$\mathfrak{S} \times_0 (\mathfrak{G}^{-1})^T = (\mathfrak{S}^T)^{-1}\ ;$$

mais, d'une part, $(\mathfrak{G}^{-1})^T = \mathfrak{G}^{-1}$ et, d'autre part, $(\mathfrak{S}^T)^{-1}$ est une matrice triangulaire inférieure, de sorte que

$$\varepsilon_i^T\,(\mathfrak{S}^T)^{-1}\,\varepsilon_j = 0 \quad (i < j)\ , \quad \varepsilon_i^T\,(\mathfrak{S}^T)^{-1}\,\varepsilon_i = \varepsilon_i^T\,\mathfrak{S}^{-1}\,\varepsilon_i = \left[\varepsilon_i^T\,\mathfrak{S}\,\varepsilon_i\right]^{-1}.$$

On part donc de

$$\left(\varepsilon_i^T\,\mathfrak{S}\right) \times_0 \left(\varepsilon_j^T\,\mathfrak{G}^{-1}\right) = \varepsilon_i^T\,\left(\mathfrak{S}^T\right)^{-1}\,\varepsilon_j$$

et on prend successivement

$$\begin{array}{l} j = p\ ,\quad i = p\ ,\quad p-1\ ,\quad p-2\ ,\quad \ldots; \\ j = p-1\ ,\quad i = p-1\ ,\quad p-2\ ,\quad \ldots\ ; \\ \text{etc.} \end{array}$$

On calcule ensuite aisément

$$\hat{\mathfrak{b}}_H^T = \left(\mathfrak{G}^{-1}\,\mathfrak{A}^T\,\mathfrak{x}\right)^T = \mathfrak{G}^{-1} \times_0 \left(\mathfrak{A}^T\,\mathfrak{x}\right)^T$$

puis

$$SCN = \hat{\mathfrak{b}}_H^T \mathfrak{A}^T \mathfrak{x} = \hat{\mathfrak{b}}_H^T \times_0 (\mathfrak{A}^T \mathfrak{x})^T .$$

Une vérification des calculs est possible: comme $\mathfrak{G}^{-1} \mathfrak{G} = \mathfrak{J}_p$, on a

$$\mathfrak{G}^{-1} \left(\sum_1^p \mathfrak{G} \, \varepsilon_j \right) = \sum_1^p \mathfrak{J}_p \, \varepsilon_j = \| 1, \ldots, 1 \| ,$$

donc

$$\mathfrak{G}^{-1} \times^0 \mathfrak{g}_\Sigma = \| 1, \ldots, 1 \|^T ,$$

ce qui vérifie chaque colonne de $\mathfrak{G}^{-1}$ dès qu'elle est calculée.

3, 2321. Dans la seconde méthode, on pose (cfr. [VIII]), rapportant **B** à la base $\mathfrak{K}$ au lieu de $\mathfrak{H}$:

$$\mathfrak{b}_K = \mathfrak{S} \, \mathfrak{b}_H , \quad \hat{\mathfrak{b}}_K = \mathfrak{S} \, \hat{\mathfrak{b}}_H ,$$

d'où

$$\mathfrak{S}^T \hat{\mathfrak{b}}_K = \mathfrak{S}^T \mathfrak{S} \, \hat{\mathfrak{b}}_H = \mathfrak{G} \, \hat{\mathfrak{b}}_H = \mathfrak{A}^T \mathfrak{x}$$

ou encore

$$\mathfrak{S} \times^0 \hat{\mathfrak{b}}_K = \mathfrak{A}^T \mathfrak{x} ,$$

formule qui se prête immédiatement au calcul numérique.

$\hat{\mathfrak{b}}_K$ est évidemment l'estimateur privilégié de $\mathfrak{b}_K$. Il jouit de la propriété suivante:

$$\mathbf{C} \, \mathfrak{b}_K = \mathbf{E} \left[\mathfrak{S} \left(\hat{\mathfrak{b}}_H - \mathfrak{b}_H \right) \right] \left[\mathfrak{S} \left(\hat{\mathfrak{b}}_H - \mathfrak{b}_H \right) \right]^T = \mathfrak{S} \left[\mathbf{C} \, \hat{\mathfrak{b}}_H \right] \mathfrak{S}^T$$
$$= \mathfrak{S} \left(\mathfrak{G}^{-1} \sigma^2 \right) \mathfrak{S}^T = \mathfrak{J}_p \, \sigma^2 ;$$

$\hat{\mathfrak{b}}_K$ est donc orthonormal, et l'on a

$$\operatorname{var} \left(\varepsilon_i^T \hat{\mathfrak{b}}_K \right) = \sigma^2 ,$$

$$\operatorname{cov} \left(\varepsilon_i^T \hat{\mathfrak{b}}_K , \varepsilon_j^T \hat{\mathfrak{b}}_K \right) = 0 \quad (i \neq j) ,$$

$$\mathbf{SCN} = \sum_1^p \mathbf{SC} \left\{ \varepsilon_i^T \hat{\mathfrak{b}}_K \right\} .$$

Mais, si l'on pose $\varepsilon_i^T \hat{\mathfrak{b}}_K = \mathfrak{l}_i^\star \mathfrak{x}$, on a $\operatorname{var} (\varepsilon_i^T \hat{\mathfrak{b}}_K) = \sigma^2 \mathfrak{l}_i^\star \mathfrak{l}_i$, ce qui montre que $\mathfrak{l}_i^\star \mathfrak{l}_i = 1$; dès lors, puisque

$$\mathbf{SC} \left\{ \varepsilon_i^T \hat{\mathfrak{b}}_K \right\} = \mathbf{SC} \left\{ \mathfrak{l}_i^\star \right\} = (\mathfrak{l}_i^\star \mathfrak{x})^2 / (\mathfrak{l}_i^\star \mathfrak{l}_i) = (\mathfrak{l}_i^\star \mathfrak{x})^2 ,$$

on a

$$\mathbf{SCN} = \sum_{1}^{p} (\varepsilon_i^T \hat{\mathfrak{b}}_K)^2 . \tag{15}$$

3, 2322. $\hat{\mathfrak{b}}_H$ s'obtient immédiatement à partir de $\hat{\mathfrak{b}}_K$ par la formule

$$\mathfrak{S} \times_0 \hat{\mathfrak{b}}_H^T = \hat{\mathfrak{b}}_K ,$$

où l'on utilise successivement les lignes de $\mathfrak{S}$ en remontant de la dernière à la première.

Ici aussi, des vérifications sont possibles. On doit en effet avoir

$$(\varepsilon_i^T \mathfrak{S}^T) \hat{\mathfrak{b}}_K = \varepsilon_i^T \mathfrak{A}^T \mathfrak{x}$$

d'où

$$\mathfrak{s}_\Sigma \times_0 \hat{\mathfrak{b}}_K = \sum_i \varepsilon_i^T \mathfrak{A}^T \mathfrak{x} ,$$

ce qui vérifie le calcul de $\hat{\mathfrak{b}}_K$. Si, ensuite, on appelle $\mathfrak{S}_\Sigma^T$ la ligne des sommes des termes des colonnes de $\mathfrak{S}$, on doit avoir

$$\mathfrak{S}_\Sigma^T \times_0 \hat{\mathfrak{b}}_H^T = \sum_{1}^{p} \varepsilon_i^T \hat{\mathfrak{b}}_K ,$$

ce qui vérifie le calcul de $\hat{\mathfrak{b}}_H^T$.

3, 2323. Considérons la matrice $\mathfrak{S}_t$ formée par les t premières lignes et colonnes de $\mathfrak{S}$; le mode de calcul de $\mathfrak{S}$ montre immédiatement que $\mathfrak{S}_t$ ne dépend que de la matrice $\mathfrak{G}_t$ formée des t premières lignes et colonnes de $\mathfrak{G}$. Par contre, $(\mathfrak{G}_t)^{-1}$ ne dépend pas seulement de $\mathfrak{S}_t$ (le calcul de $\mathfrak{G}^{-1}$ commence par le coin inférieur droit). De même, $\varepsilon_1^T \hat{\mathfrak{b}}_K, \ldots, \varepsilon_t^T \hat{\mathfrak{b}}_K$ ne dépendent que de $\mathfrak{S}_t$, mais $\varepsilon_1^T \hat{\mathfrak{b}}_H, \ldots, \varepsilon_t^T \hat{\mathfrak{b}}_H$ dépendent de l'ensemble de $\mathfrak{S}$ (le calcul de $\hat{\mathfrak{b}}_H$ commence par $\varepsilon_p^T \hat{\mathfrak{b}}_H$). Dès lors, supposons que le modèle soit réduit à ses t premiers paramètres, $\varepsilon_1^T \mathfrak{b}_H, \ldots, \varepsilon_t^T \mathfrak{b}_H$, ce qui revient à remplacer $\mathfrak{G}$ par $\mathfrak{G}_t$ et $\mathfrak{b}_H$ par $[\varepsilon_1^T \mathfrak{b}_H, \ldots, \varepsilon_t^T \mathfrak{b}_H]^T$, ou encore à supposer $\varepsilon_{t+1}^T \mathfrak{b}_H = \ldots = \varepsilon_p^T \mathfrak{b}_H = 0$; ce modèle est donc caractérisé par le fait que

$$(\mathfrak{l}^\star \in \{\mathfrak{e}_{t+1}^\star, \ldots, \mathfrak{e}_p^\star\}) \quad \text{implique} \quad \mathbf{E}\, \mathfrak{l}^\star \mathfrak{x} = 0 , \tag{16}$$

relation de la forme (11). Dans ce modèle réduit, rien n'est changé au calcul de $\varepsilon_1^T \hat{\mathfrak{b}}_K, \ldots, \varepsilon_t^T \hat{\mathfrak{b}}_K$ et, si l'on note $\mathbf{SC}N_t$ et $\mathbf{SC}E_{n-t}$ res-

pectivement la somme de carrés normale et la somme de carrés de l'erreur de ce modèle, on a, en vertu de (15),

$$\mathbf{SC}N_t = \mathbf{SC}\{\mathfrak{e}_1^\star, \ldots, \mathfrak{e}_t^\star\} = \sum_1^t (\varepsilon_i^T \hat{\mathfrak{b}}_K)^2 ,$$

donc

$$\mathbf{SC}E_{n-t} = \mathbf{SC}T - \sum_1^t (\varepsilon_i^T \hat{\mathfrak{b}}_K)^2 ,$$

$$\mathbf{SC}E_{n-t-1} = \mathbf{SC}T - \sum_1^{t+1} (\varepsilon_i^T \hat{\mathfrak{b}}_K)^2 ,$$

ce qui implique

$$\mathbf{red}\left[\{\mathfrak{e}_{t+1}^\star\} \mid \{\mathfrak{e}_1^\star, \ldots, \mathfrak{e}_t^\star\}\right] = (\varepsilon_{t+1}^T \hat{\mathfrak{b}}_K)^2 .$$

La notation du premier membre de cette formule est entièrement comparable à celles du § 2, 4; elle est toutefois un peu lourde, et on la remplace le plus souvent par

$$\mathbf{red}\left[\varepsilon_{t+1}^T \mathfrak{b}_H \mid \varepsilon_1^T \mathfrak{b}_H, \ldots, \varepsilon_t^T \mathfrak{b}_H\right] \equiv \mathbf{red}\left[\{\mathfrak{e}_{t+1}^\star\} \mid \{\mathfrak{e}_1^\star, \ldots, \mathfrak{e}_t^\star\}\right] .$$

Dans ces conditions, la table d'analyse de variance s'écrit ainsi:

Sommes de carrés	Formules	Degrés de liberté
SCT	$\sum_1^n x_i^2$	n
SCN	$\sum_1^p (\varepsilon_i^T \hat{\mathfrak{b}}_K)^2$	p
$\left\{\begin{array}{l} \mathrm{red}\,[\varepsilon_1^T \mathfrak{b}_H] \\ \mathrm{red}\,[\varepsilon_2^T \mathfrak{b}_H \mid \varepsilon_1^T \mathfrak{b}_H] \\ \vdots \\ \mathrm{red}\,[\varepsilon_p^T \mathfrak{b}_H \mid \varepsilon_1^T \mathfrak{b}_H, \ldots, \varepsilon_{p-1}^T \mathfrak{b}_H] \end{array}\right.$	$\left\{\begin{array}{l} (\varepsilon_1^T \hat{\mathfrak{b}}_K)^2 \\ (\varepsilon_2^T \hat{\mathfrak{b}}_K)^2 \\ \\ (\varepsilon_p^T \hat{\mathfrak{b}}_K)^2 \end{array}\right.$	$\left\{\begin{array}{l} 1 \\ 1 \\ \vdots \\ 1 \end{array}\right.$.
SCE	$SCT - SCN$	$n - p$
$\left\{\begin{array}{l} SCint \\ SCEM \end{array}\right.$	$\left\{\begin{array}{l} ? \\ SCE - SCint \end{array}\right.$	$\left\{\begin{array}{l} u \\ n - p - u \end{array}\right.$

3, 24. La disposition matérielle des calculs a une certaine importance; le cas $p = 3$ est décrit en détail ci-dessous.

3, 241. Première méthode.

$$\mathfrak{G} = \begin{Vmatrix} a & b & c \\ b & d & e \\ c & e & f \end{Vmatrix} \qquad \mathfrak{g}_\Sigma = \begin{Vmatrix} k \\ l \\ m \end{Vmatrix}$$

$$\mathfrak{S} = \begin{Vmatrix} x & y & z \\ & u & \varphi \\ & & \varpi \end{Vmatrix} \qquad \mathfrak{s}_\Sigma = \begin{Vmatrix} t \\ r \\ s \end{Vmatrix}$$

$$\begin{matrix} \dfrac{1}{x} & \dfrac{1}{u} & \dfrac{1}{\varpi} \end{matrix}$$

$$\mathfrak{G}^{-1} = \begin{Vmatrix} \mathrm{A} & \mathrm{B} & \mathrm{C} \\ \mathrm{B} & \mathrm{D} & \mathrm{E} \\ \mathrm{C} & \mathrm{E} & \mathrm{F} \end{Vmatrix} \qquad \mathfrak{g}_\Sigma = \begin{Vmatrix} k \\ l \\ m \end{Vmatrix}$$

$$(\mathfrak{A}^T \mathfrak{x})^T = \begin{Vmatrix} n & p & q \end{Vmatrix}$$

$$\mathfrak{b}_H^T = \begin{Vmatrix} b_1 & b_2 & b_3 \end{Vmatrix}$$

On calcule successivement (les inconnues sont soulignées)

$$\underline{x}^2 = a \qquad x\underline{y} = b \qquad x\underline{z} = c \qquad xt = k$$

$$y^2 + \underline{u}^2 = d \qquad yz + u\underline{\varphi} = e \qquad yt + ur = l$$

$$z^2 + \varphi^2 + \underline{\varpi}^2 = f \qquad zt + \varphi r + \varpi s = m$$

$$\varpi\underline{F} = \frac{1}{\varpi} \qquad \varphi F + u\underline{E} = 0 \qquad zF + yE + x\underline{C} = 0 \qquad Ck + El + Fm = 1$$

$$\varphi E + u\underline{D} = \frac{1}{u} \qquad zE + yD + x\underline{B} = 0 \qquad Bk + Dl + Em = 1$$

$$zC + yB + x\underline{A} = \frac{1}{x} \qquad Ak + Bl + Cm = 1$$

$$b_1 = An + Bp + Cq\,, \quad b_2 = Bn + Dp + Eq\,, \quad b_3 = Cn + Ep + Fq\,,$$
$$SCN = b_1 n + b_2 p + b_3 q\,.$$

3, 242. Seconde méthode.

$$\mathfrak{G} = \begin{Vmatrix} a & b & c \\ b & d & e \\ c & d & f \end{Vmatrix} \qquad \mathfrak{g}_\Sigma = \begin{Vmatrix} k \\ l \\ m \end{Vmatrix}$$

$$\mathfrak{S} = \begin{Vmatrix} x & y & z \\ & u & \varphi \\ & & \varpi \end{Vmatrix} \qquad \hat{\mathfrak{b}}_K = \begin{Vmatrix} a_1 \\ a_2 \\ a_3 \end{Vmatrix} \qquad \mathfrak{A}^T \mathfrak{x} = \begin{Vmatrix} n \\ p \\ q \end{Vmatrix} \qquad \mathfrak{s}_\Sigma = \begin{Vmatrix} t \\ r \\ s \end{Vmatrix}$$

$$\hat{\mathfrak{b}}_H^T = \begin{Vmatrix} b_1 & b_2 & b_3 \end{Vmatrix}$$

$$\mathfrak{c}_\Sigma^T = \begin{Vmatrix} c_1 & c_2 & c_3 \end{Vmatrix} \qquad \mathrm{A} \qquad \mathrm{T}$$

On calcule et vérifie $\mathfrak{S}$ comme ci-dessus, puis:

$$x\underline{a}_1 = n \qquad y a_1 + u\underline{a}_2 = p \qquad z a_1 + v a_2 + w\underline{a}_3 = q$$

$$A = a_1 + a_2 + a_3 \qquad T = n + p + q \qquad t a_1 + r a_2 + s a_3 = T$$

$$c_1 = x \qquad c_2 = y + u \qquad c_3 = z + v + w$$

$$w\underline{b}_3 = a_3 \qquad v b_3 + u\underline{b}_2 = a_2 \qquad z b_3 + y b_2 + x\underline{b}_1 = a_1$$

$$b_1 c_1 + b_2 c_2 + b_3 c_3 = A$$

$$SCN = a_1^2 + a_2^2 + a_3^2 .$$

4. Exemples.

4, 1. *Les épreuves de Student.*

4, 11. Soit $\mathbf{x}_1, \ldots, \mathbf{x}_n$ un échantillon simple et fortuit d'une population normale de moyenne μ et écart-type σ. La théorie des modèles linéaires s'applique ici, avec

$$r = p = 1 , \qquad \mathfrak{A} = \| 1, \ldots, 1 \| \qquad \mathfrak{b}_H = \| \mu \| ,$$

$$\mathfrak{A}^T \mathfrak{A} = \| n \| , \qquad \mathfrak{A}^T \mathfrak{x} = \sum_1^n \mathbf{x}_i ,$$

et le système normal se réduit à

$$n \hat{\mu} = \sum_1^n \mathbf{x}_i , \qquad \hat{\mu} = \frac{1}{n} \sum_1^n \mathbf{x}_i = \mathbf{m} .$$

On a alors

$$\mathbf{SC}N = \left(\sum \mathbf{x}_i\right)^2 / n , \qquad \mathbf{SC}T = \sum_1^n \mathbf{x}_i^2 ,$$

$$\mathbf{SC}E = \sum_1^n \mathbf{x}_i^2 - \left(\sum_1^n \mathbf{x}_i\right)^2 / n = \sum_1^n (\mathbf{x}_i - \mathbf{m})^2 = \frac{n \sum_1^n - \mathbf{x}_i^2 \left(\sum_1^n \mathbf{x}_i\right)^2}{n} .$$

Si $\mu = a$, l'expression

$$\frac{(\mathbf{m} - a)\sqrt{(n-1)}}{\sqrt{(\mathbf{SC}E/n)}} = \sqrt{(n-1)}\, \frac{\sum_1^n \mathbf{x}_i - n a}{\sqrt{(n\, \mathbf{SC}E)}}$$

est une aléatoire $\mathbf{t}_{n-1}$; $\mathbf{SC}E/\sigma^2$ est une aléatoire χ^2_{n-1} .

4, 12. Soit $(\mathbf{x}_1, \ldots, \mathbf{x}_q)$ un échantillon simple et fortuit d'une population normale de moyenne μ_1 et écart-type σ, $(\mathbf{x}_{q+1}, \ldots, \mathbf{x}_n)$ un échantillon simple et fortuit d'une population normale de moyenne μ_2 et écart-type σ, les deux échantillons étant mutuellement indépendants. La théorie des modèles linéaires s'applique encore:

$$\mathbf{E} \left\| \begin{matrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_q \\ \mathbf{x}_{q+1} \\ \vdots \\ \mathbf{x}_n \end{matrix} \right\| = \left\| \begin{matrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{matrix} \right\| \left\| \begin{matrix} \mu_1 \\ \mu_2 \end{matrix} \right\| ,$$

$$r = p = 2 , \qquad \mathfrak{b}_H^T = \| \mu_1, \mu_2 \| , \qquad \mathfrak{A}^T \mathfrak{A} = \left\| \begin{matrix} q & 0 \\ 0 & n - q \end{matrix} \right\| .$$

Si on pose

$$\sum_1^q \mathbf{x}_i^s = \mathbf{S}_{1,s} , \qquad \sum_{q+1}^n \mathbf{x}_i^s = \mathbf{S}_{2,s} ,$$

on a

$$\hat{\mu}_1 = \mathbf{S}_{1,1}/q , \qquad \hat{\mu}_2 = \mathbf{S}_{2,1}/(n - q) ,$$

$$\mathbf{SC}E = \frac{q\,\mathbf{S}_{1,2} - (\mathbf{S}_{1,1})^2}{q} + \frac{(n - q)\,\mathbf{S}_{2,2} - (\mathbf{S}_{2,1})^2}{n - q} ,$$

de sorte que, sous l'hypothèse $\mu_1 - \mu_2 = \Delta$, l'expression

$$\frac{(\hat{\mu}_1 - \hat{\mu}_2 - \Delta)\sqrt{(n - 2)}}{\sqrt{\left[\left(\frac{1}{q} + \frac{1}{n - q}\right)\mathbf{SC}E\right]}}$$

est une aléatoire $\mathbf{t}_{n-2}$; $\mathbf{SC}E/\sigma^2$ est une aléatoire χ^2_{n-2}.

4, 2. *Problèmes de régression.*

4, 211. Supposons que, $u_1, \ldots, u_s$ étant des constantes certaines deux à deux distinctes, on ait $n = \Sigma_1^s k_i$ variables aléatoires $\mathbf{x}_{i,j}$ $(i = 1, \ldots, s;\ j = 1, \ldots, k_i)$, normales, de variance commune σ^2, indépendantes, avec

$$\mathbf{E}\,\mathbf{x}_{i,j} = \alpha + \beta u_i . \tag{17}$$

Ce cas rentre dans le cadre général des modèles linéaires moyennant

$$\mathfrak{b}_H = \| \alpha \quad \beta \|^T , \quad \mathfrak{A} = \left\| \begin{matrix} 1 & \dots & 1 & 1 & \dots & 1 & \dots & 1 & \dots & 1 \\ u_1 & \dots & u_1 & u_2 & \dots & u_2 & \dots & u_s & \dots & u_s \end{matrix} \right\|^T$$

($\mathfrak{A}$ composée de s groupes ayant respectivement $k_1, k_2, \dots, k_s$ lignes identiques entre elles); ici, $r = p = 2$. Si l'on procède comme au § 3, 232, en posant

$$\sum_{i=1}^{s} \sum_{j=1}^{k_i} u_i^q \, \mathbf{x}_{i,j}^t = \mathbf{S}_{q,t} \; ; \quad S_{0,0} \equiv n \; ; \quad S_{q,0} = \sum_{1}^{i} k_i \, u_i^q \; ;$$

$$L_{u,u} = \frac{n \, S_{2,0} - S_{1,0}^2}{n} \; ; \qquad \mathbf{L}_{u,x} = \frac{n \, \mathbf{S}_{1,1} - S_{1,0} \, \mathbf{S}_{0,1}}{n} \; ;$$

on a, successivement,

$$\mathfrak{G} = \left\| \begin{matrix} n & S_{1,0} \\ S_{1,0} & S_{2,0} \end{matrix} \right\|$$

$$\mathfrak{S} = \left\| \begin{matrix} \sqrt{n} & S_{1,0}/\sqrt{n} \\ 0 & \sqrt{L_{u,u}} \end{matrix} \right\| \qquad \mathfrak{A}^T \mathfrak{x} = \left\| \begin{matrix} \mathbf{S}_{0,1} \\ \mathbf{S}_{1,1} \end{matrix} \right\| \qquad \hat{\mathfrak{b}}_K = \left\| \begin{matrix} \mathbf{S}_{0,1}/\sqrt{n} \\ \mathbf{L}_{u,x} \end{matrix} \right\|$$

$$\hat{\mathfrak{b}}_H^T = \left\| \begin{matrix} \dfrac{\mathbf{S}_{0,1}}{n} - \dfrac{S_{1,0}}{n} \dfrac{\mathbf{L}_{u,x}}{L_{u,u}} & \dfrac{\mathbf{L}_{u,x}}{L_{u,u}} \end{matrix} \right\|$$

d'où la table d'analyse de variance:

$SCT = S_{0,2}$		n *d.l.*
	$\begin{cases} \text{red}\,[\alpha] = (S_{0,1})^2/n \\ \text{red}\,[\beta \mid \alpha] = (L_{u,x})^2 \end{cases}$	$\begin{cases} 1 \; d.l. \\ 1 \; d.l. \end{cases}$
$SCN = \text{red}\,[\alpha] + \text{red}\,[\beta \mid \alpha]$		2 *d.l.*
$SCE = SCT - SCN$		$(n - 2)$ *d.l.*

4, 212. On peut traiter ce même problème d'une manière un peu différente, en posant

$$u_i' = u_i - S_{1,0}/n \; ,$$

$$\alpha + \beta \, u_i \equiv \alpha' + \beta' \, u_i' \qquad (\beta' = \beta \; , \quad \alpha' = \alpha + \beta \, S_{1,0}/n) \; ;$$

ceci revient à changer de base dans B, et nous écrirons

$$\mathfrak{b}_{H'} = \| \alpha' \; \beta' \| \; .$$

On a alors, en marquant de l'apostrophe les expressions propres à la forme actuelle du modèle considéré,

$$\mathfrak{A}' = \left\| \begin{matrix} 1 & \dots & 1 & 1 & \dots & 1 & \dots & 1 & \dots & 1 \\ u'_1 & \dots & u'_1 & u'_2 & \dots & u'_2 & \dots & u'_s & \dots & u'_s \end{matrix} \right\|^T ,$$

$$S'_{1,0} = 0 \ , \quad S'_{2,0} = L_{u,u} \ , \qquad \mathbf{S}'_{1,1} = \mathbf{L}_{u,x} \ ,$$

$$\mathfrak{G}' = \operatorname{diag}(n, L_{u,u}) \ ; \quad \mathfrak{A}^T \mathfrak{x} = \| \ \mathbf{S}_{0,1} \quad \mathbf{L}_{u,x} \ \|^T$$

$$\hat{\alpha}' = \mathbf{S}_{0,1}/n \qquad \hat{\beta}' = \mathbf{L}_{u,x}/L_{u,u} \ (= \hat{\beta}) \ .$$

La table d'analyse de la variance ne change évidemment pas.

La méthode du § 3, 232 constitue, en quelque sorte, une orthogonalisation a posteriori: on part de $\mathfrak{b}_H$, les vecteurs $\varepsilon_i^T \hat{\mathfrak{b}}_H$ ($\in \mathsf{V}^\star$) ne sont pas orthogonaux, mais les calculs introduisent d'eux-même une base $\mathfrak{K}$ telle que les vecteurs $\varepsilon_i^T \hat{\mathfrak{b}}_K$ soient orthogonaux. Ici, nous venons de procéder à une orthogonalisation a priori: nous avons d'emblée introduit une base $\mathfrak{H}'$ telle que la matrice $\mathfrak{G}'$ relative à cette base soit diagonale, ce qui garantit l'orthogonalité des vecteurs $\varepsilon_i^T \hat{\mathfrak{b}}_{H'}$. Cette seconde méthode est souvent préférable à la première. C'est sur elle que reposent, notamment, les procédés de « codage linéaire » utilisés, dans les manuels d'analyse statistique, pour l'étude des plans factoriels à facteurs quantitatifs (plans factoriels « de régression »).

4, 213. Il arrive que l'on désire contrôler, par les observations elles-mêmes, la validité de la relation (17). Le modèle basé sur (17) est alors considéré comme un cas particulier du modèle défini par

$$\mathbf{E}\, \mathbf{x}_{i,j} = M_i \ ;$$

dans ce modèle plus général, l'espace des erreurs, V_*, est engendré par les fonctionnelles de $\mathfrak{x}$ qui sont de la forme $(x_{i,j} - x_{i,k})$; il admet donc la base suivante:

$$x_{i,1} - x_{i,2}, \dots, x_{i,1} - x_{i,k_i} \ , \qquad i = 1, 2, \dots, s \ ,$$

laquelle s'orthogonalise en

$$\begin{cases} \mathfrak{d}^\star_{i,t}\, \mathfrak{x} \equiv x_{i,1} + \dots + x_{i,t-1} - (t-1)\, x_{i,t} \\ \qquad i = 1, \dots, s; \quad t = 2, \dots, k_i \ . \end{cases}$$

On a alors

$$SCint \equiv SC\,\mathsf{V}_* = \sum_{i=1}^{s} \sum_{t=2}^{k_i} SC\{\mathfrak{d}^\star_{i,t}\}$$

$$= \sum_{i=1}^{s} \sum_{t=2}^{k_i} (\mathfrak{d}^\star_{i,t}\mathfrak{x})^2/(t-1)\,t$$

$$= \sum_{i=1}^{s} \frac{k_i \sum_{j=1}^{k_i} x_{i,j}^2 - \left(\sum_{j=1}^{k_i} x_{i,j}\right)^2}{k_i} \tag{18}$$

avec $\Sigma_i\,(k_i - 1) = n - s$ degrés de liberté. Le contrôle envisagé n'est donc possible que si l'un au moins des entiers k_i est > 1, et il ne présente, en pratique, quelque intérêt que si $n - s$ est, au moins, de l'ordre de s. On le fait alors en comparant $SCEM$ à $SCint$ au moyen des tables de **F**. Si l'on procède ainsi, il sied d'utiliser $SCint$, et non SCE, comme dénominateur des divers F calculés.

Remarque. — Le calcul qui a conduit à l'expression (18) de $SCint$ est valide dans des conditions extrêmement générales.

4, 214. On peut évidemment éprouver des hypothèses très diverses relativement à α et β (ou, ce qui revient au même, à α' et β') [14]). Ainsi, l'on pourrait éprouver l'hypothèse $\beta = a$, a étant un nombre donné; il suffit d'appliquer la formule du § 2, 23, en remplaçant, au besoin, SCE et $(n - r)$ par $SCint$ et $(n - s)$. Le seul point un peu délicat est le calcul de $\mathfrak{l}^\star\mathfrak{l}$; or, on a

$$\hat{\beta} = \| l_{1,1}, \ldots, l_{1,k_1}, \ldots, l_{s,k_s} \| \mathfrak{x} \equiv \mathfrak{l}^\star \mathfrak{x}$$

moyennant

$$l_{i,j} = (n\,u_i - S_{1,0})/n\,L_{u,u}\ ;$$

on a donc

$$\mathfrak{l}^\star\mathfrak{l} = \sum_i \sum_j l_{i,j}^2 = 1/L_{u,u}\ ;$$

par conséquent, sous l'hypothèse $\beta = a$, l'expression

$$\frac{(\hat{\beta} - a)\sqrt{(n-s)}}{\sqrt{(\mathfrak{l}^\star\mathfrak{l}.\ SC\,int)}} = (\hat{\beta} - a)\sqrt{\frac{(n-s)\,L_{u,u}}{SC\,int}}$$

est une valeur observée d'une aléatoire $\mathbf{t}_{n-s}$.

On éprouverait de même, par exemple, l'hypothèse que, pour des valeurs données u_0 et x_0, on a $\alpha + \beta u_0 = x_0$ (on considérerait l'expression $\hat{\alpha} + \hat{\beta}\, u_0$, qui, sous cette hypothèse, a comme moyenne x_0).

4, 22. Supposons que, u et v étant deux variables certaines, on ait

$$\mathbf{E}\, \mathbf{x}_{u,v} = \beta_0 + \beta_1 u + \beta_2 v \,, \tag{19}$$

et que les observations aient été faites aux « points » (0, 0), (2, 0), (2, 1), (1, 2), (0, 2) et (1, 1); les observations sont, ici encore, censées être des valeurs observées d'aléatoires normales, indépendantes, de même variance σ^2. La théorie générale s'applique alors, avec

$$\mathfrak{b}_H = \| \beta_0, \beta_1, \beta_2 \|^T \,, \quad \mathfrak{A}_H = \left\| \begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 & 2 & 1 \end{matrix} \right\|^T, \quad n = 6 \,, \quad p = r = 3 \,.$$

Il est commode de traiter ce problème par orthogonalisation à priori; on rapporte donc $\mathbf{B}$ à une base $\mathfrak{K}$ telle que les colonnes de $\mathfrak{A}_K$ soient deux à deux orthogonales; si l'on pose $\mathfrak{b}_K = \| \gamma_0, \gamma_1, \gamma_2 \|^T$, cela revient à chercher deux polynômes du premier degré, $\varphi(u)$ et $\psi(u, v)$, tels que

$$\gamma_0 + \gamma_1 \varphi(u) + \gamma_2 \psi(u, v) \equiv \beta_0 + \beta_1 u + \beta_2 v \,,$$

$$\sum_1^6 \varphi(u_i) = 0 \,, \quad \sum_1^6 \psi(u_i, v_i) = 0 \,, \quad \sum_1^6 \varphi(u_i)\, \psi(u_i, v_i) = 0 \,.$$

On peut prendre

$$\varphi(u) = u - 1 \,, \quad \psi(u, v) = -5 + u + 4v \,,$$

ce qui correspond à

$$\mathfrak{b}_H = \left\| \begin{matrix} 1 & -1 & -5 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{matrix} \right\| \mathfrak{b}_K \,.$$

On a alors

$$\mathbf{E} \left\| \begin{matrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \\ \mathbf{x}_5 \\ \mathbf{x}_6 \end{matrix} \right\| = \left\| \begin{matrix} 1 & -1 & -5 \\ 1 & 1 & -3 \\ 1 & 1 & 1 \\ 1 & 0 & 4 \\ 1 & -1 & 3 \\ 1 & 0 & 0 \end{matrix} \right\| \left\| \begin{matrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{matrix} \right\|$$

$$\mathfrak{A}_K^T \mathfrak{x} \equiv \left\| \begin{matrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \mathbf{A}_3 \end{matrix} \right\| = \left\| \begin{matrix} \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 + \mathbf{x}_4 + \mathbf{x}_5 + \mathbf{x}_6 \\ -\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 - \mathbf{x}_5 \\ -5\mathbf{x}_1 - 3\mathbf{x}_2 + \mathbf{x}_3 + 4\mathbf{x}_4 + \mathbf{x}_5 \end{matrix} \right\|$$

d'où

$$\mathfrak{G}_K = \begin{Vmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 60 \end{Vmatrix}, \qquad \begin{aligned} \hat{\gamma}_0 &= \mathbf{A}_1/6 \\ \hat{\gamma}_1 &= \mathbf{A}_2/4 \\ \hat{\gamma}_2 &= \mathbf{A}_3/60 \end{aligned}$$

$$\mathbf{SC}\{\hat{\gamma}_0\} = A_1^2/6\,, \qquad \mathbf{SC}\{\hat{\gamma}_1\} = \mathbf{A}_2^2/4\,, \qquad \mathbf{SC}\{\hat{\gamma}_2\} = \mathbf{A}_3^2/60\,.$$

$$\mathbf{SC}\{\hat{\gamma}_0\} = \mathbf{red}[\beta_0]\,, \quad \mathbf{SC}\{\hat{\gamma}_1\} = \mathbf{red}[\beta_1 \mid \beta_0]\,, \quad \mathbf{SC}\{\hat{\gamma}_2\} = \mathbf{red}[\beta_2 \mid \beta_0, \beta_1]$$

$$\mathbf{SC}\,N = \mathbf{SC}\{\hat{\gamma}_0\} + \mathbf{SC}\{\hat{\gamma}_1\} + \mathbf{SC}\{\hat{\gamma}_2\}\,,$$

$$\mathbf{SC}\,E = \mathbf{SC}\,T - \mathbf{SC}\,N\,.$$

4, 3. *Problèmes de classification.*

4, 31. Supposons que l'on dispose des valeurs observées de douze aléatoires normales, indépendantes, de même variance σ^2, classées suivant deux critères: « lignes », de « valeurs » L_1 et L_2, et « colonnes », de « valeurs » C_1, C_2, C_3, suivant le schéma

	C_1	C_2	C_3
L_1	x_1, x_2	x_3, x_4	x_5, x_6
L_2	x_7, x_8	x_9, x_{10}	x_{11}, x_{12} .

On suppose a priori qu'il y a additivité, c'est-à-dire qu'il existe cinq nombres λ_1, λ_2, γ_1, γ_2, γ_3 tels que la valeur moyenne d'une observation de la ligne L_i et de la colonne C_k soit $\lambda_i + \gamma_k$ ($i = 1, 2$; $k = 1, 2, 3$). On a donc, par hypothèse,

$$\mathbf{E}\begin{Vmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \\ \mathbf{x}_5 \\ \mathbf{x}_6 \\ \mathbf{x}_7 \\ \mathbf{x}_8 \\ \mathbf{x}_9 \\ \mathbf{x}_{10} \\ \mathbf{x}_{11} \\ \mathbf{x}_{12} \end{Vmatrix} = \begin{Vmatrix} 1 & & 1 & & \\ 1 & & 1 & & \\ 1 & & & 1 & \\ 1 & & & 1 & \\ 1 & & & & 1 \\ 1 & & & & 1 \\ & 1 & 1 & & \\ & 1 & 1 & & \\ & 1 & & 1 & \\ & 1 & & 1 & \\ & 1 & & & 1 \\ & 1 & & & 1 \end{Vmatrix} \begin{Vmatrix} \lambda_1 \\ \lambda_2 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{Vmatrix}.$$

Si on appelle $S_{i,-}$, la somme des observations de la ligne L_i, et $S_{-,j}$ celle des observations de la colonne C_k, les équations normales s'écrivent:

$$6\hat{\lambda}_1 + 2\hat{\gamma}_1 + 2\hat{\gamma}_2 + 2\hat{\gamma}_3 = \mathbf{S}_{1,-} \qquad (a)$$

$$6\hat{\lambda}_2 + 2\hat{\gamma}_1 + 2\hat{\gamma}_2 + 2\hat{\gamma}_3 = \mathbf{S}_{2,-} \qquad (b)$$

$$2\hat{\lambda}_1 + 2\hat{\lambda}_2 + 4\hat{\gamma}_1 = \mathbf{S}_{-,1} \qquad (c)$$

$$2\hat{\lambda}_1 + 2\hat{\lambda}_2 + 4\hat{\gamma}_2 = \mathbf{S}_{-,2} \qquad (d)$$

$$2\hat{\lambda}_1 + 2\hat{\lambda}_2 + 4\hat{\gamma}_3 = \mathbf{S}_{-,3} \qquad (e)$$

Ici, manifestement, $p = 5$, $r = 4$ [en effet, $(a) + (b) \equiv (c) + (d) + (e)$]. On est donc amené à mettre en évidence quatre combinaisons estimables fondamentales; on peut prendre

$$\mu = 6(\lambda_1 + \lambda_2) + 4(\gamma_1 + \gamma_2 + \gamma_3)$$
$$\Delta\lambda = \lambda_1 - \lambda_2$$
$$\Delta_1\gamma = \gamma_1 - \gamma_2 , \qquad \Delta'\gamma = \gamma_1 - \gamma_3 .$$

On constate immédiatement que les estimateurs privilégiés de ces quatre combinaisons sont deux à deux orthogonaux, à l'exception près de la dernière paire; l'orthogonalité complète est atteinte en remplaçant $\Delta'\gamma$ par

$$\Delta_2\gamma = (\lambda_1 + \lambda_2 - 2\lambda_3)/2 = (\lambda_1 + \lambda_2)/2 - \lambda_3 .$$

Alors:

$$\hat{\mu} = \mathbf{S}_{1,-} + \mathbf{S}_{2,-} \; (= \mathbf{S}_{-,1} + \mathbf{S}_{-,2} + \mathbf{S}_{-,3}) ,$$
$$\widehat{\Delta\lambda} = (\mathbf{S}_{1,-} - \mathbf{S}_{2,-})/6$$
$$\widehat{\Delta_1\gamma} = (\mathbf{S}_{-,1} - \mathbf{S}_{-,2})/4$$
$$\widehat{\Delta_2\gamma} = (\mathbf{S}_{-,1} + \mathbf{S}_{-,2} - 2\mathbf{S}_{-,3})/8 .$$

Ici, on peut calculer

$$\mathbf{SC}\,int = (1/2)\,[(\mathbf{x}_1 - \mathbf{x}_2)^2 + \ldots + (\mathbf{x}_{11} - \mathbf{x}_{12})^2] ,$$

avec six degrés de liberté, puis, avec deux degrés de liberté,

$$\mathbf{SC}\,EM = \mathbf{SC}\,E - \mathbf{SC}\,int .$$

On notera que, ici, on a

$$\mathbf{SC}\{\hat{\mu}\} = \mathbf{red}\,[\mu] = \mathbf{red}\,[\mu \mid \Delta\lambda] = \ldots = \mathbf{red}\,[\mu \mid \Delta\lambda, \Delta_1\gamma, \Delta_2\gamma] ,$$

et des relations analogues pour les autres paramètres; ceci en vertu de l'orthogonalité de leurs estimateurs privilégiés.

SC*EM* peut servir à éprouver l'hypothèse d'additivité, mais on ne le voit clairement qu'en étudiant le modèle non additif.

4, 32. Ne supposons donc plus a priori qu'il y ait additivité; admettons que

$$\mathbf{E}\,\mathbf{x}_{2i-1} = \mathbf{E}\,\mathbf{x}_{2i} = M_i \ ,$$

de sorte que $p = 6$, $\mathfrak{b}_H = \| M_1\, M_2\, M_3\, M_4\, M_5\, M_6 \|^T$. On voit aisément que

$$\mathfrak{S} = \mathrm{diag}\,(2, \ldots, 2)$$

d'où

$$\hat{\mathbf{M}}_i = (\mathbf{x}_{2i-1} + \mathbf{x}_{2i})/2 \ ,$$

$$SC\left\{ \hat{\mathbf{M}}_i \right\} = (\mathbf{x}_{2i-1} + \mathbf{x}_{2i})^2/2 \ ,$$

$$\mathbf{SC}N = \sum_1^6 SC\left\{ \hat{\mathbf{M}}_i \right\} \ ,$$

$$\mathbf{SC}E = \sum_1^{12} \mathbf{x}_i^2 - \mathbf{SC}N = (1/2) \sum_1^6 (\mathbf{x}_{2i} - \mathbf{x}_{2i-1})^2 \ .$$

Donc, pour le modèle général actuel, $\mathbf{SC}E$ vaut l'expression $\mathbf{SC}int$ du modèle additif.

L'hypothèse d'additivité (c'est-à-dire, répétons-le, l'hypothèse qu'il existe cinq nombres λ_1, λ_2, γ_1, γ_2, γ_3 tels que

$$M_1 = \lambda_1 + \gamma_1 \ , \quad M_2 = \lambda_1 + \gamma_2 \ , \ldots, \quad M_6 = \lambda_2 + \gamma_3)$$

est satisfaite si et seulement si

$$\theta_1 \equiv M_1 - M_2 - M_4 + M_5 = 0 \ , \quad \theta_2 \equiv M_1 - M_3 - M_4 + M_6 = 0 \ .$$

On doit donc former, pour éprouver cette hypothèse,

$$\hat{\theta}_1 = \hat{\mathbf{M}}_1 - \hat{\mathbf{M}}_2 - \hat{\mathbf{M}}_4 + \hat{\mathbf{M}}_5 \ , \qquad \hat{\theta}_2 = \hat{\mathbf{M}}_1 - \hat{\mathbf{M}}_3 - \hat{\mathbf{M}}_4 + \hat{\mathbf{M}}_6 \ ,$$

puis $\mathbf{SC}\left\{ \hat{\theta}_1, \hat{\theta}_2 \right\}$, et éprouver si $\mathbf{SC}\left\{ \hat{\theta}_1, \hat{\theta}_2 \right\}$ est, ou non, significativement plus grande que $\mathbf{SC}E$ ($\mathbf{SC}int$ du modèle additif). $\hat{\theta}_2$ n'est pas orthogonal à $\hat{\theta}_1$, mais bien

$$\hat{\theta}_3 = 2\,\hat{\theta}_2 - \hat{\theta}_1 = \mathbf{M}_1 + \mathbf{M}_2 - 2\mathbf{M}_3 + \mathbf{M}_4 + \mathbf{M}_5 - 2\,\mathbf{M}_6 \ ;$$

or,

$$\mathbf{SC}\left\{ \hat{\theta}_1 \right\} = \hat{\theta}_1^2/8 \ , \qquad \mathbf{SC}\left\{ \hat{\theta}_3 \right\} = \hat{\theta}_3^2/24 \ ,$$

$$\mathbf{SC}\left\{ \hat{\theta}_1, \hat{\theta}_2 \right\} = \mathbf{SC}\left\{ \hat{\theta}_1, \hat{\theta}_3 \right\} = \mathbf{SC}\left\{ \hat{\theta}_1 \right\} + \mathbf{SC}\left\{ \hat{\theta}_3 \right\} ,$$

ce qui permet d'éprouver l'additivité.

On montre aisément que $\mathbf{SC}\{\hat{\theta}_1, \hat{\theta}_2\}$ n'est autre que $\mathbf{SC}EM$ du modèle additif. Chacune de ces deux sommes de carrés est en effet due à un sous-espace de $\mathbf{V}^*$ ayant deux dimensions et orthogonal tant à $\hat{\mu}$, $\hat{\Delta}\lambda$, $\hat{\Delta}_1\gamma$, $\hat{\Delta}_2\gamma$ qu'aux différences « internes » $x_{2i} - x_{2i-1}$, et un tel espace est unique.

$\hat{\theta}_1$ et $\hat{\theta}_2$ (et leurs combinaisons linéaires, notamment $\hat{\theta}_3$) portent le nom de « contrastes de non-additivité » (le terme « contrastes d'interaction » n'est guère heureux). On remarquera que, si les M_i déterminent entièrement les paramètres « orthogonalement estimables » μ, $\Delta\lambda$, $\Delta_1\gamma$, $\Delta_2\gamma$, θ_1, θ_2, ceux-ci à leur tour déterminent entièrement les M_i. C'est en le posant au moyen des paramètres μ, ..., θ_2 que le problème de la classification (2×3) (avec un nombre quelconque d'observations par cellule) se traite le plus aisément. Toutefois, ces paramètres ne restent orthogonalement estimables que si toutes les cellules renferment un même nombre d'observations.

Remarque. — Ces considérations s'étendent immédiatement aux autres problèmes de classification.

4, 4. *Covariance* [15]).

4, 41. Supposons que l'on dispose d'une observation de chacune de n variables aléatoires indépendantes, normales, de même variance σ^2, réparties en s groupes, le $i^{ème}$ groupe étant formé de n_i éléments, avec

$$\begin{cases} \mathbf{E}\,\mathbf{x}_{i,j} = T_i + \beta\, v_{i,j} \\ i = 1, \ldots, s\ ; \quad j = 1, \ldots, n_i\ ; \quad n_1 + \ldots + n_s = n\ ; \end{cases}$$

les $v_{i,j}$ étant des nombres certains [16]). On a ici

$$\mathfrak{b}_H = \| T_1, \ldots, T_s, \beta \|^T ,$$

et, en supposant que dans chaque groupe il y a au moins deux valeurs $v_{i,j}$ distinctes,

$$p = r = s + 1 .$$

On se convainc aisément que la matrice $\mathfrak{G}_H$ n'est pas diagonale. Toutefois, il est aisé d'orthogonaliser le problème; il suffit de poser

$$\begin{aligned} v_{i,-} &= (1/n_i) \sum_j v_{i,j} \\ u_{i,j} &= v_{i,j} - v_{i,} \quad , \end{aligned}$$

ce qui entraîne

$$\begin{cases} \mathbf{E}\, x_{i,j} = A_i + \beta\, u_{i,j} \\ i = 1, \ldots, s \; ; \quad j = 1, \ldots, n_i \; ; \quad \sum_j u_{i,j} = 0 \\ A_i \equiv T_i - \beta\, v_{i,-} \; . \end{cases}$$

En fait, cela revient à référer $\mathbf{B}$ à une base $\mathfrak{K}$ telle que

$$\mathfrak{b}_K = \left\| \begin{matrix} 1 & 0 & \ldots & 0 & -v_{1,-} \\ 0 & 1 & & & \\ \vdots & & \ddots & & \\ 0 & & & 1 & -v_{s,-} \\ 0 & \ldots & & 0 & 1 \end{matrix} \right\| \mathfrak{b}_H$$

On a alors

$$\mathfrak{A}_K = \left\| \begin{matrix} 1 & & & & u_{1,1} \\ \vdots & & & & \vdots \\ 1 & & & & u_{1,n_1} \\ & 1 & & & u_{2,1} \\ & \vdots & & & \vdots \\ & 1 & \ldots & & u_{2,n_2} \\ & & \ldots & & \vdots \\ & & & 1 & u_{s,1} \\ & & & \vdots & \vdots \\ & & & 1 & u_{s,n_s} \end{matrix} \right\|$$

(les termes non écrits étant nuls).

4, 42. Il est alors utile de poser

$$\begin{aligned} S_{i;t,q} &= \sum_j (u_{i,j})^t \, (x_{i,j})^q \, , \\ S_{;t,q} &= \sum_i S_{i;t,q} \, , \\ S'_{i;t,q} &= \sum_j (v_{i,j})^t \, (x_{i,j})^q \, , \\ S'_{;t,q} &= \sum_i S'_{i;t,q} \, , \end{aligned}$$

ce qui conduit à

$$\begin{aligned} \mathfrak{A}_K^T \mathfrak{x} &= \| S_{1;0,1}, \ldots, S_{s;0,1}, S_{;1,1} \| \, , \\ \mathfrak{G}_K &= \operatorname{diag} (n_1, \ldots, n_s, S_{;2,0}) \, , \end{aligned}$$

d'où

$$\begin{cases} \hat{\mathbf{A}}_i = \mathbf{S}_{i;0,1}/n_i \\ \hat{\beta} = \mathbf{S}_{;1,1}/S_{;2,0} \,, \end{cases}$$

$\mathbf{SC}T = \mathbf{S}_{;0,2}$ (avec n degrés de liberté),

$\mathbf{SC}\{\hat{\mathbf{A}}_i\} = \mathbf{red}\,[A_i] = (\mathbf{S}_{i;0,1})^2/n_i$ (avec 1 degré de liberté),

$\mathbf{SC}\{\hat{\beta}\} = \mathbf{red}\,[\beta] = (\mathbf{S}_{;1,1})^2/S_{;2,0}$ (avec 1 degré de liberté),

$\mathbf{SC}N = \sum_i SC\{\hat{\mathbf{A}}_i\} + SC\{\hat{\beta}\}$ (avec $s+1$ degrés de liberté),

$\mathbf{SC}E = \mathbf{SC}T - \mathbf{SC}N$

$$= \sum_{i=1}^{s} \frac{n_i\,\mathbf{S}_{i;0,2} - (\mathbf{S}_{i;0,1})^2}{n_i} - \frac{\sum_i \left[n_i\,\mathbf{S}'_{i;1,1} - S'_{i;1,0}\,\mathbf{S}_{i;0,1}\right]}{\sum_i \left[n_i\,S'_{i;2,0} - (S'_{i;1,0})^2\right]}$$

(avec $n - s - 1$ degrés de liberté).

σ^2 est estimé par $\mathbf{SCE}/(n - s - 1)$.

4, 43. En fait, l'intérêt se porte généralement sur l'estimation des différences $T_i - T_k$, les paramètres A_i n'ayant pas d'intérêt propre. On a, en posant $\varphi_{i,-} - \varphi_{k,-} = d_{i,k}$,

$$\begin{aligned} (\mathbf{T}_i - \mathbf{T}_k)^\wedge &= \left[(\hat{\mathbf{A}}_i + \hat{\beta}\,\varphi_{i,-}) - (\hat{\mathbf{A}}_k + \hat{\beta}\,\varphi_{k,-})\right]^\wedge \\ &= \hat{\mathbf{A}}_i - \hat{\mathbf{A}}_k + \hat{\beta}\,d_{i,k} \\ &= \mathfrak{l}^\star\,\mathfrak{x} \equiv \| l_{1,1}, \ldots, l_{s,n_s} \|\,\mathfrak{x} \end{aligned}$$

moyennant

$$\begin{aligned} l_{t,q} &= d_{i,k}\,u_{t,q}/S_{;2,0} \qquad \text{si } t \neq i,\ t \neq k\,, \\ l_{i,q} &= d_{i,k}\,u_{i,q}/S_{;2,0} + 1/n_i\,, \\ l_{k,q} &= d_{i,k}\,u_{k,q}/S_{;2,0} - 1/n_k\,, \end{aligned}$$

d'où

$$\mathfrak{l}^\star\,\mathfrak{l} = (d_{i,k})^2/S_{;2,0} + 1/n_i + 1/n_k\,.$$

Il résulte de là que l'expression

$$\Delta_{i,k} = \frac{\left[(\hat{\mathbf{T}}_i - \hat{\mathbf{T}}_k) - (T_i - T_k)\right]}{\sqrt{[\mathbf{SC}E/(n-s-1)}\,.\,\sqrt{[(d_{i,k})^2/S_{;2,0} + 1/n_i + 1/n_k]}}$$

est une aléatoire $\mathbf{t}_{n-s-1}$.

On voit que l'erreur-type de $\Delta_{i,k}$, produit de σ par

$$\varepsilon_{i,k} = \left[\frac{(d_{i,k})^2}{S_{;2,0}} + \frac{1}{n_i} + \frac{1}{n_k}\right]^{1/2},$$

varie, en général, avec la paire (i, k), même si les n_i sont égaux entre eux ($n_1 = \ldots = n_s = \nu$). Toutefois, dans ce dernier cas, on utilise d'ordinaire une expression approchée de $\varepsilon_{i,k}$, obtenue en considérant les $\varphi_{i,j}$ comme n observations indépendantes d'une même aléatoire, de moyenne φ et écart-type θ [17]). En désignant par $\mathbf{M}$ la valeur moyenne prise par rapport à la distribution de cette aléatoire (fictive), on a

$$\mathbf{M}\,(\varphi_{i,-} - \varphi_{k,-})^2 = \mathbf{M}\,[(\varphi_{i,-} - \varphi) - (\varphi_{k,-} - \varphi)]^2$$
$$= 2\,\mathbf{M}\,(\varphi_{i,-} - \varphi)^2 = 2\,\theta^2/\nu\ ;$$

$$\mathbf{M}\,S_{;2,0} = \mathbf{M}\sum_i \sum_j (\varphi_{i,j} - \varphi_{i,-})^2 = \sum_i (\nu - 1)\,\theta = k\,(\nu - 1)\,\theta^2\ ;$$

$$\left\{\frac{\mathbf{M}\,(\varphi_{i,-} - \varphi_{k,-})^2}{\mathbf{M}\,S_{;2,0}} + \frac{1}{n_i}\,\frac{1}{n_k}\right\}^{1/2} = \left\{\frac{2}{\nu}\left[1 + \frac{1}{k\,(\nu - 1)}\right]\right\}^{1/2}\ ;$$

c'est cette dernière expression que l'on utilise comme valeur approchée de $\varepsilon_{i,k}$.

4, 44. Il est utile aussi de calculer la somme de carrés due au sous-espace $\mathsf{U}^\star$ engendré par les estimateurs des contrastes $T_i - T_k$, afin de pouvoir construire une épreuve globale de la nullité de ceux-ci. On pourrait évidemment construire une base orthogonale de $\mathsf{U}^\star$, mais les calculs nécessaires sont d'une grande complication. Il est beaucoup plus simple d'appliquer le théorème du § 2, 31, en procédant comme suit: on considère le modèle

$$\mathbf{E}\,\mathbf{x}_{i,j} = T + \beta\,\varphi_{i,j}$$

et on calcule la somme de carrés de l'erreur qui lui est associée, c'est-à-dire (voir § 4, 211)

$$(\mathbf{SC}E)_{T_i = T} = \mathbf{S}_{;0,2} - \frac{(\mathbf{S}_{;0,1})^2}{n} - \frac{(\mathbf{S}_{;1,1})^2}{S_{;2,0}}\ ;$$

or ce modèle s'obtient à partir du modèle initial en supposant

que, pour tout vecteur $\mathfrak{l}^\star$ de $\mathsf{U}^\star$, on a $\mathfrak{E}\,\mathfrak{l}^\star = 0$; par conséquent

$$\begin{aligned}\mathbf{SC}\,\mathsf{U}^\star &= (\mathbf{SC}E)_{T_i=T} - \mathbf{SC}E \\ &= \sum_{i=1}^{n} \frac{(\mathbf{S}_{i;0,1})^2}{n_i} - \frac{(\mathbf{S}_{;0,1})^2}{n} \,. \qquad (19)\end{aligned}$$

Si on pose

$$\mathbf{x}_{i,-} = \frac{1}{n}\sum\nolimits_j \mathbf{x}_{i,j}\,, \quad \mathbf{x}_{-,-} = \frac{1}{n}\sum\nolimits_i \sum\nolimits_j \mathbf{x}_{i,j} = \sum\nolimits_i \left(\frac{n_i}{n}\right) \mathbf{x}_{i,-}\,,$$

la formule (19) peut s'écrire

$$\begin{aligned}\mathbf{SC}\,\mathsf{U}^\star &= \sum_{i=1}^{s} \frac{\left(\sum_j \mathbf{x}_{i,j}\right)^2}{n_i} - \frac{1}{n}\left(\sum_{i=1}^{n} \sum\nolimits_j \mathbf{x}_{i,j}\right)^2 \\ &= n \sum_{i=1}^{s} \frac{n_i}{n} (\mathbf{x}_{i,-})^2 - n\left(\sum_{i=1}^{s} \frac{n_i}{n}\mathbf{x}_{i,-}\right)^2 \\ &= \sum_{i=1}^{s} n_i\,(\mathbf{x}_{i,-} - \mathbf{x}_{-,-})^2 \,. \qquad (20)\end{aligned}$$

L'expression (20), moins aisée à mettre en nombres que l'expression (19), est plus parlante qu'elle.

L'hypothèse $T_1 = T_2 = \ldots = T_s$ s'éprouve en comparant

$$\frac{\mathbf{SC}\,\mathsf{U}^\star/(s-1)}{\mathbf{SC}E/(n-s-1)}$$

à la distribution de $\mathbf{F}$ à $(s-1)$ et $(n-s-1)$ degrés de liberté. La table d'analyse de variance se présente ainsi:

$$\begin{array}{lll}
SCT = S_{;0,2} & & n \text{ d.l.} \\
SCN \begin{cases} \text{red}\,[T] = (S_{;0,1})^2/n \\ \text{red}\,[\beta] = (S_{;1,1})^2/S_{;2,0} \\ \text{red}\,[T_1 - T_2, \ldots, T_1 - T_s \mid T, \beta] = \mathrm{SC}\,\mathsf{U}^\star \end{cases} & (s+1) & \begin{cases} 1 \text{ d.l.} \\ 1 \text{ d.l.} \\ (s-1) \text{ d.l.} \end{cases} \\
SCE = SCT - SCN & (n-s-1) \text{ d.l.} &
\end{array}$$

Si plusieurs paires (i, j) correspondent à une même valeur de $v_{i,j}$, on peut introduire, en outre, la décomposition habituelle de SCE en SCint et $SCEM$.

NOTES

[13]) Dans ce paragraphe, les notations ε_i et ε_i^T désignent des vecteurs à p (et non n) composantes, toutes nulles, sauf la $i^{ème}$, qui vaut 1.

[14]) Rappelons qu'il convient que ces hypothèses soient formulées indépendamment des observations.

[15]) On trouvera dans [IX], chap. 4, un exposé non technique particulièrement lucide des principes statistiques de l'analyse de la covariance.

[16]) Si les $v_{i,j}$ sont des valeurs observées de certaines aléatoires, les résultats subsistent inchangés, sous la condition nécessaire et suffisante que la distribution conjointe de ces aléatoires ne dépende pas des valeurs des paramètres $T_1, \ldots, T_s$, β. Cela résulte immédiatement du théorème des probabilités composées (cfr. [X], §§ I.7 ss.).

[17]) Cette hypothèse apparemment farfelue est, en fait, assez souvent justifiée par les techniques d'attributions fortuites (« randomization ») mises en jeu; en tout état de cause, on peut la considérer comme un procédé d'approximation numérique basé sur le sentiment qu'« il n'existe pas de différence *systématique* entre les $v_{i,j}$ correspondant à diverses valeurs de i ».

RÉFÉRENCES BIBLIOGRAPHIQUES

I. W. Graeub, *Lineare Algebra.* Berlin (Springer), 1958.

II. H. Breny, Sur les « fondements » de la théorie des probabilités. *Revue des questions scientifiques*, CXXIX (1958), 161.

III. H. Cramér, *Mathematical Methods of Statistics.* Princeton (Univ. Press), 1946.

IV. K. Pearson, *Tables of the Incomplete Beta Function.* London (Biometrika Office), 1948.

V. S. Crump, The present status of Variance Components Analysis. *Biometrics*, 7 (1951), 1.

VI. A. M. Turing, Rounding-off errors in matrix processes. *Quarterly Journal of Mechanics and Applied Mathematics*, 1 (1948), 287.

VII. L. Fox, Practical methods for the solution of linear equations... *Journal of the Royal Statistical Society*, B, XII/1 (1950), 120.

VIII. S. Rushton, On least squares fitting... using the Choleski method. *Journal of the Royal Statistical Society*, B, XIII/1 (1951), 92.

IX. D. R. Cox, *Planning of Experiments.* New-York (Wiley), 1958.

X. J. L. Doob, *Stochastic Processes.* New-York (Wiley), 1954.

Manuscrit reçu le 30.10.1959.

H. Breny
Centre interdisciplinaire d'analyse stochastique
et de recherche opérationelle,
Université de Liège.