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.
Kapitel: 4, 3. Problèmes de classification.
DOI: https://doi.org/10.5169/seals-36341

d'où

$$\mathfrak{S}_K = \left\| \begin{matrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 60 \end{matrix} \right\| , \qquad \begin{aligned} \hat{\gamma}_0 &= \mathbf{A}_1/6 \\ \hat{\gamma}_1 &= \mathbf{A}_2/4 \\ \hat{\gamma}_2 &= \mathbf{A}_3/60 \end{aligned}$$

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

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

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

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

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,

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

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:

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

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

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

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

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

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

$$\begin{aligned} \mu &= 6(\lambda_1 + \lambda_2) + 4(\gamma_1 + \gamma_2 + \gamma_3) \\ \Delta\lambda &= \lambda_1 - \lambda_2 \\ \Delta_1\gamma &= \gamma_1 - \gamma_2 \,, \qquad \Delta'\gamma = \gamma_1 - \gamma_3 \,. \end{aligned}$$

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:

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

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

$$\mathfrak{G} = \text{diag}\,(2, \ldots, 2)$$

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, **SC**E vaut l'expression **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

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

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 **SC**E (**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 $\mathsf{V}^\star$ 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

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

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

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

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

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