2. Distributions et épreuves d'hypothèses

Objekttyp: **Chapter**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **6 (1960)**

Heft 1: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **21.07.2024**

1, 4. *Moindres carrés.*

A partir de la relation

$$\mathbf{E}\,\mathfrak{x} = \mathfrak{A}\,\mathfrak{b}_H \; ,$$

le théorème de Gauss-Markov conduit à introduire le vecteur-estimateur $\overset{*}{\mathfrak{b}}$ défini en fonction de l'observation $\mathfrak{x}$ par la condition que

$$S^2(\mathfrak{b}) \equiv (\mathfrak{x} - \mathfrak{A}\,\mathfrak{b})^T\,(\mathfrak{x} - \mathfrak{A}\,\mathfrak{b})$$

soit minimum pour $\mathfrak{b} = \overset{*}{\mathfrak{b}}$. Or, comme

$$\frac{d}{d\,b_{i,H}}\,\mathfrak{b}_H = \mathfrak{e}_{i,H} \; , \quad \frac{d}{d\,b_{i,H}}\,\mathfrak{b}_H^T = \mathfrak{e}_{i,H}^T \; ,$$

et

$$S^2(\mathfrak{b}) = \mathfrak{x}^\star\,\mathfrak{x} - \mathfrak{b}_H^T\,\mathfrak{A}^T\,\mathfrak{x} - \mathfrak{x}^\star\,\mathfrak{A}\,\mathfrak{b}_H + \mathfrak{b}_H^T\,\mathfrak{A}^T\,\mathfrak{A}\,\mathfrak{b}_H \; ,$$

on a

$$\frac{d\,S^2(\mathfrak{b})}{d\,b_{i,H}} = 2\,(\mathfrak{A}^T\,\mathfrak{A}\,\mathfrak{b}_H - \mathfrak{A}^T\,\mathfrak{x}) \; .$$

Les conditions $\dfrac{d\,S^2(\overset{*}{\mathfrak{b}})}{d\,b_{i,H}} = 0$ conduisent donc au système

$$\mathfrak{A}^T\,\mathfrak{A}\,\overset{*}{\mathfrak{b}}_H = \mathfrak{A}^T\,\mathfrak{x} \; ,$$

identique au système normal. Il en résulte que, si $r = p$, les estimateurs de moindres carrés ne sont autres que les estimateurs privilégiés. Si $r < p$, les deux méthodes conduisent aux mêmes combinaisons estimables fondamentales.

2. Distributions et épreuves d'hypothèses.

2, 1. *Sommes des carrés.*

2, 11. Soit $\mathbf{U}^\star$ un sous-espace vectoriel de $\mathbf{V}^\star$; on appelle « somme de carrés due à $\mathbf{U}^\star$ », et on note $\mathbf{SC}\,\mathbf{U}^\star$ [11]), le carré scalaire de la projection orthogonale de $\mathfrak{x}$ sur le dual $\mathbf{U}$ de $\mathbf{U}^\star$. La dimension de $\mathbf{U}^\star$ est, par définition, le « nombre de degrés de liberté » de $\mathbf{SC}\,\mathbf{U}^\star$.

Si les vecteurs $\mathfrak{v}_1^\star, \dots \mathfrak{v}_t^\star$ $(t \geqslant s)$ engendrent $\mathsf{U}^\star$, on écrit, d'ordinaire, $\{\mathfrak{v}_1^\star, \dots, \mathfrak{v}_t^\star\}$ pour $\mathsf{U}^\star$; on écrira donc aussi $\mathbf{SC}\{\mathfrak{v}_1^\star, \dots, \mathfrak{v}_t^\star\}$ pour $\mathbf{SC}\,\mathsf{U}^\star$.

2, 12. Pour calculer effectivement $\mathbf{SC}\,\mathsf{U}^\star$, on introduit une base quelconque de $\mathsf{U}^\star$, soit $\mathfrak{u}_1^\star, \dots, \mathfrak{u}_s^\star$. La projection orthogonale $\mathfrak{x}_u$ de $\mathfrak{x}$ sur U est alors définie par les relations

$$\mathfrak{x}_u = \sum \lambda_i \mathfrak{u}_i \quad , \quad < \mathfrak{x} - \mathfrak{x}_u \, , \, \mathfrak{u}_k > \; \equiv \mathfrak{u}_k^\star (\mathfrak{x} - \mathfrak{x}_u) = 0 \quad (k = 1, \dots, s)$$

d'où l'on tire

$$\sum_{i=1}^{s} \lambda_i \mathfrak{u}_i^\star \mathfrak{u}_k = \mathfrak{u}_k^\star \mathfrak{x} \; (k = 1, \dots, s) \; , \tag{7}$$

système d'équations linéaires qui détermine entièrement les λ_i (en effet, les $\mathfrak{u}_i$ formant une base de U, la matrice $\| \mathfrak{u}_i^\star \mathfrak{u}_k \|$ est de rang s). On a alors

$$\mathbf{SC}\,\mathsf{U}^\star = \mathfrak{x}_u^\star \mathfrak{x}_u = \left(\sum_1^s \lambda_i \mathfrak{u}_i^\star \right) \left(\sum_1^s \lambda_k \mathfrak{u}_k \right)$$

$$= \sum_1^s \sum_1^s \lambda_i \lambda_k \mathfrak{u}_i^\star \mathfrak{u}_k$$

moyennant (7), ce qui entraîne

$$\mathbf{SC}\,\mathsf{U}^\star = \sum_1^s \lambda_k \mathfrak{u}_k^\star \mathfrak{x} \; . \tag{8}$$

Dans le cas où $s = 1$ ($\mathsf{U}^\star$ engendré par l'unique vecteur $\mathfrak{u}^\star$), on a

$$\mathbf{SC}\{\mathfrak{u}^\star\} = (\mathfrak{u}^\star \mathfrak{x})^2 / (\mathfrak{u}^\star \mathfrak{u}) \; . \tag{9}$$

2, 13. Soient $\mathsf{U}_1^\star$ et $\mathsf{U}_2^\star$ deux sous-espaces complémentaires de $\mathsf{U}^\star$, mutuellement orthogonaux, U_1 et U_2 leurs duals; ceux-ci sont, dans U, deux sous-espaces complémentaires mutuellement orthogonaux, et on a

$$\mathfrak{x}_u^\star \mathfrak{x}_u = \mathfrak{x}_{u_1}^\star \mathfrak{x}_{u_1} + \mathfrak{x}_{u_2}^\star \mathfrak{x}_{u_2} \; ,$$

ce qui entraîne

$$\mathbf{SC}\,\mathsf{U}^\star = \mathbf{SC}\,\mathsf{U}_1^\star + \mathbf{SC}\,\mathsf{U}_2^\star \; . \tag{10}$$

ce résultat s'étend sans peine au cas de plus de deux compo santes, et on peut énoncer que

si $\mathsf{U}^\star$ *est la somme (directe) des espaces mutuellement orthogonaux* $\mathsf{U}_1^\star, \ldots, \mathsf{U}_t^\star$, *on a*

$$\mathbf{SC}\,\mathsf{U}^\star = \sum_1^t \mathbf{SC}\,\mathsf{U}_i^\star \,.$$

Il en résulte un mode de calcul des sommes de carrés qui est assez souvent plus commode que l'emploi des formules (7) et (8). On part d'une base $\mathfrak{u}_1^\star, \ldots, \mathfrak{u}_s^\star$ de $\mathsf{U}^\star$; si elle n'est pas orthogonale, on l'orthogonalise (par exemple, par le procédé pas à pas de Schmidt), ce qui fournit la base orthogonale $\mathfrak{w}_1^\star, \ldots, \mathfrak{w}_s^\star$; alors on a

$$\mathbf{SC}\,\mathsf{U}^\star = \sum_1^s \mathbf{SC}\{\, \mathfrak{w}_i^\star \,\}$$

et donc, en vertu de (9),

$$\mathbf{SC}\,\mathsf{U}^\star = \sum_1^s (\mathfrak{w}_i^\star\, \mathfrak{x})^2 / (\mathfrak{w}_i^\star\, \mathfrak{w}_i) \,. \qquad (11)$$

2, 14. On écrit, en particulier,

SCT (somme de carrés totale) pour $\mathbf{SC}\,\mathsf{V}^\star$,
SCN (somme de carrés normale) pour $\mathbf{SC}\,\mathsf{V}_+$,
SCE (somme de carrés des erreurs) pour $\mathbf{SC}\,\mathsf{V}_0$.

On notera que, V_+ et V_0 étant par définition complémentaires et orthogonaux dans $\mathsf{V}^\star$, on a toujours

$$\mathbf{SCT} = \mathbf{SCN} + \mathbf{SCE} \,.$$

D'autre part, $\mathfrak{e}_1^\star, \ldots, \mathfrak{e}_n^\star$ forment une base orthogonale de $\mathsf{U}^\star$, et $\mathfrak{e}_i^\star\, \mathfrak{x} = \mathbf{x}_i$; donc

$$\mathbf{SCT} = \sum_1^n \mathbf{x}_i^2 \,.$$

2, 2. *Distributions. Epreuves d'hypothèses.*

2, 21. Soit $\mathsf{U}^\star$ un sous-espace de $\mathsf{V}^\star$, de dimension s, $\mathfrak{w}_1^\star, \ldots, \mathfrak{w}_s^\star$ une base orthogonale de $\mathsf{U}^\star$. Chaque $\mathfrak{w}_i^\star\, \mathfrak{x}$ est une variable aléatoire normale, de moyenne $\mathfrak{w}_i^\star \mathfrak{A}\, \mathfrak{b}$ et de variance $(\mathfrak{w}_i^\star \mathfrak{w}_i)\, \sigma^2$; en outre, si $i \neq k$,

$$\operatorname{cov} (\mathfrak{w}_i^\star \mathfrak{x}, \mathfrak{w}_k^\star \mathfrak{x}) = (\mathfrak{w}_i^\star \mathfrak{w}_k)\, \sigma^2 = 0 \ ,$$

de sorte que les aléatoires $\mathfrak{w}_i^\star \mathfrak{x}$ sont les composantes non corrélées d'un vecteur multinormal, elles sont donc indépendantes.

Dès lors, **si l'on suppose** que $\mathfrak{b}$ est tel que $\mathbf{E}\, \mathfrak{w}_i^\star \mathfrak{x} = 0$ $(i = 1, ..., s)$ (c'est-à-dire sous l'hypothèse $\mathfrak{w}_1^\star \mathfrak{A}\, \mathfrak{b} = ...\ \mathfrak{w}_s^\star \mathfrak{A}\, \mathfrak{b} = 0$), les aléatoires $(\mathfrak{w}_i^\star \mathfrak{x})/\sigma$ sont gaussiennes et indépendantes, de sorte que

$$(1/\sigma^2)\ \mathbf{SC}\, \mathsf{U}^\star = \sum_1^s [(\mathfrak{w}_i^\star \mathfrak{x})/\sigma]^2$$

est une aléatoire χ_s^2.

Si, par contre, on ne suppose pas $\mathbf{E}\, \mathfrak{w}_i^\star \mathfrak{x} = 0$, $(1/\sigma^2)\ \mathbf{SC}\, \mathsf{U}^\star$ est une aléatoire χ^2 décentrée à s degrés de liberté, elle est donc, en loi, plus grande qu'une aléatoire χ_s^2:

$$\mathbf{Pr}\, [(\mathbf{SC}\, \mathsf{U}^\star)/\sigma^2 > a] \geqslant \mathbf{Pr}\, \left[\chi_s^2 > a\right] .$$

2, 22. Prenons pour $\mathsf{U}^\star$ l'espace des erreurs, V_0; alors $s = n - r$ et les conditions $\mathbf{E}\, \mathfrak{w}_i^\star \mathfrak{x} = 0$ sont identiquement [9]) satisfaites. On a donc, *indépendamment de toute hypothèse quant à* $\mathfrak{b}$,

$$\mathbf{Pr}\, [\mathbf{SC}\,\mathrm{E} > a\,\sigma^2] = \mathbf{Pr}\, \left[\chi_{n-r}^2 > a\right] ,$$

d'où, notamment,

$$\mathbf{Pr}\, [\mathbf{SC}\,\mathrm{E}/a < \sigma^2 < \mathbf{SC}\,\mathrm{E}/b] = \mathbf{Pr}\left[a < \chi_{n-r}^2 < b\right] ,$$

ce qui permet d'estimer σ.

2, 23. Supposons que $\mathfrak{f}^\star \mathfrak{b} = \mathfrak{m}_f^T \mathfrak{A}^T \mathfrak{A}\, \mathfrak{b}_H$ soit une combinaison estimable et que $\mathfrak{l}_f^\star \mathfrak{x} \equiv \mathfrak{m}_f^T \mathfrak{A}^T \mathfrak{x}$ soit son estimateur privilégié. Alors:

a) **sous l'hypothèse** $\mathfrak{f}^\star \mathfrak{b} = a$, $\left[(\mathfrak{l}_f^\star \mathfrak{x} - a)/\sigma \sqrt{(\mathfrak{l}_f^\star \mathfrak{l}_f)}\right]$ est une aléatoire gaussienne;

b) $\mathbf{SC}\,\mathrm{E}/(n - r)\, \sigma^2$ est une aléatoire χ_{n-r}^2 ;

c) $\mathbf{SC}\,\mathrm{E}$ et $\mathfrak{l}_f^\star \mathfrak{x}$ sont indépendantes (car $\mathfrak{l}_f^\star$, estimatrice, est orthogonale à tous les vecteurs de V_0).

Donc, sous l'hypothèse susdite,

$$\Delta_a = \frac{\mathfrak{l}_f^\star\, \mathfrak{x} - a}{\sigma \sqrt{(\mathfrak{l}_f^\star\, \mathfrak{l}_f)}} : \frac{\sqrt{\mathbf{SC}\,\mathrm{E}}}{\sigma \sqrt{(n-r)}} = \frac{(\mathfrak{l}_f^\star\, \mathfrak{x} - a)\sqrt{(n-r)}}{\sqrt{[(\mathfrak{l}_f^\star\, \mathfrak{l}_f)\, \mathbf{SC}\,\mathrm{E}]}}$$

est une aléatoire $\mathbf{t}_{n-r}$, ce qui permet d'éprouver l'hypothèse en question ou d'estimer $\mathfrak{f}^\star\, \mathfrak{b}$.

2, 24. Soient $\mathfrak{f}_1^\star\, \mathfrak{b}$, ..., $\mathfrak{f}_s^\star\, \mathfrak{b}$ des combinaisons estimables, linéairement indépendantes, et $\mathfrak{l}_{f_i}^\star\, \mathfrak{x}$ ($i = 1, ..., s$) leurs estimateurs privilégiés. **Sous l'hypothèse** $\mathfrak{f}_1^\star\, \mathfrak{b} = ... = \mathfrak{f}_s^\star\, \mathfrak{b} = 0$, les moyennes des $\mathfrak{l}_{f_i}^\star\, \mathfrak{x}$ sont toutes nulles, et donc $(1/\sigma^2)\ \mathbf{SC}\{\mathfrak{l}_{f1}^\star, ..., \mathfrak{l}_{f_s}^\star\}$ est une aléatoire χ_s^2; cela entraîne que

$$\mathbf{Q} = \frac{\mathbf{SC}\{\mathfrak{l}_{f1}^\star, ..., \mathfrak{l}_{f_s}^\star\}/s}{\mathbf{SC}\,\mathrm{E}/(n-r)}$$

est une aléatoire $\mathbf{F}_{s,\, n-r}$. Si l'hypothèse en question est fausse, $\mathbf{Q}$ est, en loi, plus grande que $\mathbf{F}_{s,\, n-r}$; on éprouvera donc cette hypothèse en comparant la valeur observée de $\mathbf{Q}$ à $\mathbf{F}_{s,\, n-r}$, les grandes valeurs de $\mathbf{Q}$ étant critiques.

Remarque. — Il est manifeste que, si α est un nombre certain quelconque, on a $\mathbf{SC}\{\alpha\, \mathfrak{w}^\star\} = \mathbf{SC}\{\mathfrak{w}^\star\}$. On peut donc négliger un facteur constant dans le calcul d'une somme de carrés. Il n'en est pas de même dans le calcul de l'expression Δ_a du § 2, 33.

2, 3. *Sous-espaces disjoints non orthogonaux.*

2, 31. Soient $\mathsf{U}_q^\star$ et $\mathsf{U}_{r-q}^\star$ deux sous-espaces complémentaires de V_+, de dimensions q et $r - q$: $\mathsf{V}_+ = \mathsf{U}_q^\star \oplus \mathsf{U}_{r-q}^\star$; on ne suppose pas que $\mathsf{U}_q^\star$ et $\mathsf{U}_{r-q}^\star$ sont mutuellement orthogonaux. On cherche à interpréter $\mathbf{SC}\ \mathsf{U}_q^\star$ et $\mathbf{SC}\ \mathsf{U}_{r-q}^\star$. Pour cela, on considère, outre le modèle initial, le modèle où

$$(\mathfrak{l}^\star \in \mathsf{U}_{r-q}^\star) \quad \text{implique} \quad \mathbf{E}\,\mathfrak{l}^\star\, \mathfrak{x} = 0\ , \qquad (11)$$

tandis que $(\mathfrak{l}^\star \in \mathsf{U}_q^\star)$ implique $\mathbf{E}\,\mathfrak{l}^\star\, \mathfrak{x} \neq 0$ pour une valeur au moins de $\mathfrak{b}$.

[On pourrait décrire ce modèle ainsi: soit $\mathfrak{l}_1^\star$, ..., $\mathfrak{l}_q^\star$ une base de $\mathsf{U}_q^\star$, $\mathfrak{l}_{q+1}^\star$, ..., $\mathfrak{l}_r^\star$ une base de $\mathsf{U}_{r-q}^\star$, et $\mathfrak{W}$ telle que, dans le modèle initial,

$$\hat{\mathfrak{b}}_H = \mathfrak{W}\left[\mathfrak{l}_1^\star\, \mathfrak{x}\, ..., \mathfrak{l}_r^\star\, \mathfrak{x}\right]^T, \qquad \mathbf{E}\,\mathfrak{x} = \mathfrak{A}\,\mathfrak{W}^{-1}\,\mathfrak{W}\,\mathfrak{b} = \mathfrak{A}\,\mathfrak{W}^{-1}\,\mathfrak{w}\ ;$$

le nouveau modèle est

$$\mathbf{E}\,\mathfrak{x} = \mathfrak{A}\,\mathfrak{W}^{-1}\,\mathfrak{Z}\,\mathfrak{w}\,, \qquad \mathfrak{Z} = \left\| \begin{matrix} \mathfrak{J}_q & 0 \\ 0 & 0 \end{matrix} \right\| \cdot \Big]$$

Soient $\mathbf{SCN}_q$ et $\mathbf{SCE}_{n-q}$ les sommes de carrés normale et des erreurs pour le nouveau modèle, $\mathbf{SCN}_r$ et $\mathbf{SCE}_{n-r}$ les sommes homologues du modèle initial. On a

$$\mathbf{SC}\,\mathsf{U}^{\star}_{r-q} = \mathbf{SCE}_{n-q} - \mathbf{SCE}_{n-r}\,.$$

En effet, en notant U'_q le complément orthogonal de $\mathsf{U}^{\star}_{r-q}$, on a

$$\begin{aligned} \mathbf{SCT} &= \mathbf{SC}\,\mathsf{U}'_q + \mathbf{SC}\,\mathsf{U}^{\star}_{r-q} + \mathbf{SCE}_{n-r} \\ &= \mathbf{SCN}_q + \mathbf{SCE}_{n-q}\,; \end{aligned}$$

or, de quoi se compose l'espace des erreurs du nouveau modèle, $\mathsf{V}_{0,\,n-q}$? il contient évidemment V_0, puis un sous-espace de $\mathsf{V}^{\star}$, de dimensions $r - q$, disjoint de V_0; par ailleurs, $\mathsf{U}^{\star}_{r-q}$ appartient à $\mathsf{V}_{0,\,n-q}$ en vertu de (11), et est de dimension $r - q$; donc

$$\mathsf{V}_{0,\,n-q} = \mathsf{V}_0 \oplus \mathsf{U}^{\star}_{r-q}\,;$$

en outre, V_0 et $\mathsf{U}^{\star}_{r-q} \subset \mathsf{V}_+$ sont mutuellement orthogonaux, donc

$$\begin{aligned} \mathbf{SCE}_{n-q} \equiv \mathbf{SC}\,\mathsf{V}_{0,\,n-q} &= \mathbf{SC}\,\mathsf{V}_0 + \mathbf{SC}\,\mathsf{U}^{\star}_{r-q} \\ &= \mathbf{SCE}_{n-r} + \mathbf{SC}\,\mathsf{U}^{\star}_{r-q}\,, \end{aligned}$$

d'où la thèse; on voit en outre que $\mathbf{SCN}_q = \mathbf{SC}\,\mathsf{U}'_q \neq \mathbf{SC}\,\mathsf{U}^{\star}_q$.

2, 32. Il est commode d'introduire la notation suivante [12]):

$$\mathbf{SCT} - \mathbf{SCE}_{n-q} = \mathbf{red}\left[\mathsf{U}^{\star}_q\right],$$

$$\mathbf{SCE}_{n-q} - \mathbf{SCE}_{n-r} = \mathbf{red}\left[\mathsf{U}^{\star}_{r-q} \,\middle|\, \mathsf{U}^{\star}_q\right] (\neq \mathbf{red}\left[\mathsf{U}^{\star}_{r-q}\right])\,.$$

On a alors

$$\mathbf{SCT} = \mathbf{red}\left[\mathsf{U}^{\star}_q\right] + \mathbf{red}\left[\mathsf{U}^{\star}_{r-q} \,\middle|\, \mathsf{U}^{\star}_q\right] + \mathbf{SCE}_{n-r}$$

avec

$$\mathbf{red}\left[\mathsf{U}^{\star}_{r-q} \,\middle|\, \mathsf{U}^{\star}_q\right] = \mathbf{SC}\,\mathsf{U}^{\star}_{r-q}\,,$$

$$\mathbf{red}\left[\mathsf{U}^{\star}_q\right] = \mathbf{SC}\,\mathsf{U}'_q \neq \mathbf{SC}\,\mathsf{U}^{\star}_q\,.$$

Bien entendu, les relations obtenues en permutant les rôles de $\mathsf{U}^{\star}_q$ et $\mathsf{U}^{\star}_{r-q}$ sont aussi valables; ces rôles ne sont évidemment pas symétriques, à moins que $\mathsf{U}_q^\star$ et $\mathsf{U}_{r-q}^\star$ ne soient mutuellement orthogonaux; dans ce dernier cas,

$$\mathbf{red}\left[\mathsf{U}_q^\star\right] = \mathbf{red}\left[\mathsf{U}_q^\star \,\middle|\, \mathsf{U}_{r-q}^\star\right] = \mathbf{SC}\,\mathsf{U}_q^\star \,,$$
$$\mathbf{red}\left[\mathsf{U}_{r-q}^\star\right] = \mathbf{red}\left[\mathsf{U}_{r-q}^\star \,\middle|\, \mathsf{U}_q^\star\right] = \mathbf{SC}\,\mathsf{U}_{r-q}^\star \,.$$

2, 33. Ces considérations s'étendent aisément au cas où V_+ est décomposé en plus de deux sous-espaces, suivant le schéma

$$\begin{cases} \mathsf{V}_+ = \mathsf{U}_1^\star \oplus \mathsf{U}_2^\star \oplus \dots \oplus \mathsf{U}_t^\star \,, \\ \dim \mathsf{U}_i^\star = r_i \,, \quad r_1 + \dots + r_k = \rho_k;\ \rho_t = r \,. \end{cases}$$

On doit alors considérer t modèles successifs (et l'ordre dans lequel ces modèles font intervenir les $\mathsf{U}_i^\star$ est essentiel); le $k^{\text{ème}}$ de ces modèles est caractérisé par

$$\left[\mathfrak{l}^\star \in \bigoplus_{k+1}^{t} \mathsf{U}_i^\star\right] \quad \text{implique} \quad \mathbf{E}\,\mathfrak{l}^\star \mathfrak{x} = 0 \qquad (k = 1, 2, \dots, t-1) \,,$$

le $t^{\text{ème}}$ étant le modèle initial. On note $\mathbf{SCE}_{n-\rho_k}$ la somme de carrés des erreurs attachée au $k^{\text{ème}}$ modèle, et on montre sans peine que

$$\mathbf{SC}\left(\bigoplus_{k+1}^{t} \mathsf{U}_i^\star\right) = \mathbf{SCE}_{n-\rho_k} - \mathbf{SCE}_{n-r} \,;$$

on pose alors

$$\mathbf{red}\left[\mathsf{U}_1^\star\right] = \mathbf{SCT} - \mathbf{SCE}_{n-\rho_1}$$
$$\mathbf{red}\left[\mathsf{U}_{k+1}^\star \,\middle|\, \mathsf{U}_1^\star, \dots, \mathsf{U}_k^\star\right] = \mathbf{SCE}_{n-\rho_k} - \mathbf{SCE}_{n-\rho_{k+1}} \,,$$

et on a

$$\mathbf{SCN} = \mathbf{red}\left[\mathsf{U}_1^\star\right] + \sum_{1}^{t-1} \mathbf{red}\left[\mathsf{U}_{k+1}^\star \,\middle|\, \mathsf{U}_1^\star, \dots, \mathsf{U}_k^\star\right] , \qquad (12)$$

avec

$$\mathbf{red}\left[\mathsf{U}_t^\star \,\middle|\, \mathsf{U}_1^\star, \dots, \mathsf{U}_{t-1}^\star\right] = \mathbf{SC}\,\mathsf{U}_t^\star \,,$$

cette dernière relation n'étant pas généralement vraie pour les autres $\mathsf{U}_i^\star$ (exception évidente: le cas où les $\mathsf{U}_i^\star$ sont mutuellement orthogonaux).

2, 4. *Ecarts au modèle.*

Tout ce qui précède est valide si, réellement, $\mathbf{E}\,\mathfrak{x} = \mathfrak{A}\mathfrak{b}$. S'il n'en est pas nécessairement ainsi, ce qui arrive lorsque le modèle envisagé n'est qu'un cas particulier d'un modèle plus général auquel on désire accorder aussi quelque considération, on peut modifier un peu les énoncés des hypothèses à éprouver, en disant, par exemple: « *si* $\mathbf{E}\mathfrak{x} = \mathfrak{A}\mathfrak{b}$ et *si* $\mathfrak{l}_1^\star \mathfrak{b} = \ldots = \mathfrak{l}_s^\star \mathfrak{b} = 0$, alors ... ». Sous cette nouvelle hypothèse, $\mathbf{SC}\{\mathfrak{l}_1^\star \hat{\mathfrak{b}}, \ldots, \mathfrak{l}_s^\star \hat{\mathfrak{b}}\}$ est encore distribuée comme $\sigma^2 \chi_s^2$. Mais **SCE** n'est plus distribuée comme $\sigma^2 \chi_{n-r}$, car, si $\mathbf{E}\mathfrak{x} - \mathfrak{A}\mathfrak{b}$ n'est pas identiquement nulle, les vecteurs de V_0 n'ont plus une moyenne nécessairement nulle. On est alors obligé de prendre comme espace des erreurs un sous-espace V_* de V_0, à savoir: celui des vecteurs de $\mathsf{V}^\star$ dont la moyenne est identiquement nulle dans le modèle le plus général que l'on considère. On peut dire que ce sous-espace existe dès que les observations comportent au moins une paire d'observations ayant identiquement même moyenne (dans le modèle le plus général). Nous noterons **SCint** (« somme de carrés interne ») l'expression $\mathbf{SC}\,\mathsf{V}_*$, et **SCEM** [« somme de carrés des écarts au modèle » (sous-entendu: au modèle restreint)] l'expression **SCE** — **SCint** (en désignant par V'_* le complément orthogonal de V_* dans V_0, **SCEM** n'est autre que $\mathbf{SC}\,\mathsf{V}'_*$). Dans les considérations du § 2, 2, **SCint** peut remplacer **SCE**, $n - s$ remplaçant alors $n - r$. Les composantes de **SCint** sont évidemment orthogonales à celles de **SCN**.

Remarque. — Les composantes additives de **SCT** (ou, plus exactement, leurs valeurs observées) sont le plus souvent reprises en un tableau que l'on nomme « table d'analyse de la variance ». Cette désignation n'est guère heureuse, on devrait la réserver aux études de « composantes de variance » (cfr. [V]); elle paraît néanmoins avoir reçu la sanction de l'usage, et il semble assez vain de vouloir la récuser. Une telle table se présente ainsi:

Somme de carrés	Formules	Nombre de degrés de liberté
SCT	$\sum x_i^2$	n
SCN	$SC\{\mathfrak{A}^T x\}$	r
SCE	$SCT - SCN$	$n - r$
$\quad\{ SCint$	$SC\mathsf{V}_*$	u
$\quad\{ SCEM$	$SCE - SCint$	$n - r - u$.

D'ordinaire, SCN est décomposée conformément à la formule (12). Si on utilise la table de la distribution **F**, il est utile d'adjoindre à cette table une colonne « carrés moyens », où sont repris les quotients des SC par les nombres de leurs degrés de liberté.

NOTES

1) Dans un système complet de notations, ce n-uple serait désigné, par exemple, par $\mathfrak{x}_p$.

2) Dans un système complet de notations, ce n-uple serait désigné par $\mathfrak{x}^\star_{p\star}$, ou par $\mathfrak{x}^\star_p$ si la dualité des bases va de soi. Pour des raisons de convenance typographique, nous écrirons souvent $[a_1, \ldots, a_n]^\star$ au lieu de $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$

3) Ce second usage est permis parce que $\mathfrak{P}$ et $\mathfrak{P}^\star$ sont des bases orthonormales; s'il n'en était pas ainsi, il conviendrait d'indiquer la transposition et la dualité par des signes différents (^{T} et *).

4) On notera qu'alors $\mathfrak{A}^T$ ne représente pas l'application duale (dite aussi « transposée ») de $\mathfrak{A}$; celle-ci est représentée, ici, par la même matrice $\mathfrak{A}$; mais, dans un cas, cette matrice pré-multiplie une colonne, dans l'autre elle post-multiplie une ligne.

5) « épreuve » au singulier, car il s'agit d'une abréviation de l'expression « catégorie des résultats d'épreuve », qui n'a rien à voir avec les « épreuves répétées » dont on a parfois voulu faire le fondement, sinon de la théorie des probabilités, du moins de ses applications; cfr. [II].

6) On notera que la moyenne du vecteur aléatoire $\boldsymbol{\mathfrak{b}}$ est un vecteur défini sans recours à une base (théorie de l'intégration dans les espaces vectoriels), de sorte que la notation $\mathbf{E}\,\boldsymbol{\mathfrak{b}}$ a un sens intrinsèque [il est très heureux que $(\mathbf{E}\,\boldsymbol{\mathfrak{b}})_p = \mathbf{E}\,(\boldsymbol{\mathfrak{b}}_p)$]. L'étude intrinsèque de la covariance serait un peu moins simple.

7) On dit parfois que « des variables aléatoires normales non corrélées sont indépendantes ». Cet énoncé, pris dans toute sa généralité, est faux; il est vrai pour des aléatoires (normales, nécessairement) qui sont les composantes d'une représentation (par rapport à une base certaine) d'un vecteur multinormal.

8) Plus explicitement: $\mathfrak{l}^\star \longrightarrow [\mathfrak{b} \longrightarrow \mathfrak{l}^\star\, \mathfrak{A}\, \mathfrak{b}]$.

9) « Identiquement » par rapport à la variabilité de $\mathfrak{b}$ dans B.

10) Comme on le sait, le mot « erreur » possède, en statistique, un sens très éloigné de son sens vulgaire.

11) Il s'agit là d'une variable aléatoire; la notation appropriée à ce fait est malaisée à choisir; la convention adoptée ici a, à défaut d'autre mérite, celui d'être simple.

12) Où « red » signifie « réduction » (*scil.* de la somme de carrés des erreurs).

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

(A suivre)