LES MODÈLES LINÉAIRES EN ANALYSE STATISTIQUE

Autor(en): **Breny, H.**

Objekttyp: **Article**

Zeitschrift: **L'Enseignement Mathématique**

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

Heft 1: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **25.04.2024**

Persistenter Link: https://doi.org/10.5169/seals-36336

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.
Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.
Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Ein Dienst der *ETH-Bibliothek*
ETH Zürich, Rämistrasse 101, 8092 Zürich, Schweiz, www.library.ethz.ch

http://www.e-periodica.ch

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 $\mathbf{b}$ est un vecteur défini sans recours à une base (théorie de l'intégration dans les espaces vectoriels), de sorte que la notation $\mathbf{E}\,\mathbf{b}$ a un sens intrinsèque [il est très heureux que $(\mathbf{E}\,\mathbf{b})_p = \mathbf{E}\,(\mathbf{b}_p)$]. 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 $\mathbf{B}$.

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)

LES MODÈLES LINÉAIRES EN ANALYSE STATISTIQUE

par H. BRENY [1])

0. PRÉLIMINAIRES.

0, 1. *Introduction.*

La présente note a pour but d'exposer, d'une manière rigoureusement déductive, l'état actuel (bien proche, semble-t-il, de la perfection) de la théorie statistique des modèles linéaires (avec « erreurs » indépendantes, normales, et homoscédastiques), en insistant particulièrement, d'une part, sur la définition intrinsèque des « sommes de carrés » et, d'autre part, sur l'interprétation des sommes de carrés attachées à des groupes non orthogonaux. L'exposé utilise les notions élémentaires courantes d'algèbre linéaire (cfr. [I]), avec des notations précisées ci-dessous.

0, 2. *Notations.*

0, 21. Dans une question d'algèbre linéaire, il est, en général, essentiel d'utiliser un système de notations qui fasse nettement apparaître la distinction entre un vecteur (élément d'un espace vectoriel) et la représentation de ce vecteur par rapport à une base déterminée; il est clair, d'autre part, qu'un tel système doit, pour être complet, multiplier les signes diacritiques ($\mathfrak{b}$ est un vecteur, $\mathfrak{b}_P$ sa représentation par rapport à la base $\mathfrak{P}$, etc.). Or si, dans un contexte donné, la base à laquelle on se réfère est définie sans ambiguïté, il n'y a, somme toute, pas d'inconvénient à omettre ces signes et à utiliser, pour la représentation d'un vecteur par rapport à *cette* base, le même symbole que pour le vecteur lui-même.

---

[1]) L'auteur est « Associé » du Fonds national de la Recherche scientifique (Belgique). Le texte des notes et les références bibliographiques se trouvent en fin d'article.

0, 22. D'une manière précise, soit V un espace vectoriel réel, euclidien, à n dimensions, $\mathfrak{x}$, $\mathfrak{y}$, $\mathfrak{z}$, ... des vecteurs de V, $<\mathfrak{x}, \mathfrak{y}>$ le produit scalaire de $\mathfrak{x}$ et $\mathfrak{y}$; soit $\mathsf{V}^\star$ l'espace dual de V (ensemble des fonctionnelles linéaires définies sur V), $\mathfrak{x}^\star$ le dual de $\mathfrak{x}$ (c'est-à-dire l'élément de $\mathsf{V}^\star$ défini par

$$\mathfrak{x}^\star : \quad \mathfrak{y} \in \mathsf{V} \to <\mathfrak{x}, \mathfrak{y}>) \ .$$

S'il résulte clairement du contexte que V est rapporté à une base $\mathfrak{P}$ *bien déterminée*, que $\mathsf{V}^\star$ est rapporté à la base duale $\mathfrak{P}^\star$, et que *ces bases sont orthonormales*, la notation $\mathfrak{x}$ désignera, dans ce contexte, non seulement le vecteur $\mathfrak{x}$ de V mais aussi le n-uple de ses composantes par rapport à $\mathfrak{P}$, n-uple écrit d'ailleurs sous forme de colonne [1]); en outre, $\mathfrak{e}_i$ désigne le i-ème vecteur de $\mathfrak{P}$, de sorte que

$$\mathfrak{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathfrak{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad \mathfrak{x} = \begin{bmatrix} <\mathfrak{x}, \mathfrak{e}_1> \\ \vdots \\ <\mathfrak{x}, \mathfrak{e}_n> \end{bmatrix} .$$

De même, et toujours en supposant que le contexte empêche toute ambiguïté, $\mathfrak{x}^\star$ désigne non seulement le dual du vecteur $\mathfrak{x}$ mais aussi le n-uple de ses composantes par rapport à $\mathfrak{P}^\star$, n-uple écrit d'ailleurs sous forme de ligne [2]). Cette ambivalence des notations, combinée à l'usage du signe * pour désigner à la fois le passage d'un vecteur à son dual et le passage d'une matrice à sa transposée [3]) permet d'appliquer sans peine les règles usuelles du calcul matriciel, et d'écrire, par exemple

$$<\mathfrak{x}, \mathfrak{y}> = \mathfrak{x}^\star \mathfrak{y} = \mathfrak{y}^\star \mathfrak{x}$$

tant pour les vecteurs que pour leurs représentations.

De même, si $\mathfrak{A}$ est une application linéaire de V dans un espace vectoriel U, et si les bases auxquelles sont rapportés ces espaces satisfont aux trois conditions énoncées ci-dessus, la notation $\mathfrak{A}$ sera aussi utilisée pour la matrice qui représente l'application $\mathfrak{A}$ par rapport à ces bases [4]).

0, 23. La définition usuelle du produit de deux matrices introduit une multiplication lignes par colonnes; cette définition est évidemment nécessaire aux applications les plus courantes des

notations matricielles, et il n'est nullement question de la modifier. Toutefois, en ce qui concerne les calculs numériques, cette disposition n'est guère heureuse; on est ainsi conduit à introduire une multiplication lignes par lignes, notée $\times_0$, et une multiplication colonnes par colonnes, notée $\times^0$, et à utiliser les relations

$$\mathfrak{A}\mathfrak{B} = \mathfrak{A} \times_0 \mathfrak{B}^{\mathrm{T}} = \mathfrak{A}^{\mathrm{T}} \times^0 \mathfrak{B} .$$

0, 24. Considérons une variable aléatoire, c'est-à-dire une application mesurable d'une catégorie d'épreuve [5]) (espace probabilisé) dans l'ensemble des nombres réels; considérons d'autre part la valeur attachée par cette application à un élément déterminé (ou censé tel) de la catégorie d'épreuve (le plus souvent, une « valeur observée » — ou « réalisation » — de la variable aléatoire). Il importe de pouvoir distinguer nettement ces deux êtres mathématiques; c'est pourquoi nous désignerons la variable aléatoire par un symbole en caractère gras, et une valeur observée de cette variable par le même symbole en caractères ordinaires.

Cette convention s'étend immédiatement aux éléments aléatoires plus généraux, tels que vecteurs ou matrices; ainsi, $\mathbf{x}$, $\mathfrak{b}$, $\mathfrak{A}$ représentent respectivement une variable, un vecteur, et une matrice aléatoires, x, $\mathfrak{b}$, $\mathfrak{A}$, représentent des valeurs observées de ces mêmes éléments aléatoires.

Il sera commode, lorsque nous considérerons une matrice aléatoire, $\mathfrak{A}$, de noter $\mathbf{E}\,\mathfrak{A}$ la matrice dont l'élément (i, j) est la valeur moyenne de l'élément (i, j) de $\mathfrak{A}$:

$$\mathbf{a}_{i,j} = \mathrm{e}_i^{\star}\ \mathfrak{A}\,\mathrm{e}_j,\ \mathbf{E}\,\mathfrak{A} = \sum_i \sum_j (\mathbf{E}\,\mathbf{a}_{i,j})\ \mathrm{e}_i\ \mathrm{e}_j^{\star} .$$

Cette définition, appliquée à la représentation par une matrice $1 \times n$ ou $n \times 1$ d'un vecteur aléatoire, est en parfait accord avec la définition directe de la valeur moyenne d'un vecteur aléatoire considéré comme application mesurable d'une catégorie d'épreuve dans un espace vectoriel, pour autant que la représentation se fasse par rapport à une base certaine de cet espace, ce qui, pour nous, sera toujours le cas. Soit donc $\mathfrak{b}$ un vecteur aléatoire, $\mathfrak{b}_P$ la matrice qui le représente par rapport à la base $\mathfrak{P}$, de sorte que

$$\mathbf{E}\,\mathfrak{b}_P = (\mathbf{E}\,\mathfrak{b})_P ;$$

nous poserons

$$\mathbf{C}\,\mathfrak{b}_P = \mathbf{E}[\mathfrak{b}_P - \mathbf{E}\,\mathfrak{b}_P][\mathfrak{b}_P - \mathbf{E}\,\mathfrak{b}_P]^{\mathrm{T}} .$$

$\mathbf{C}\,\mathfrak{b}_P$ est une matrice carrée, symétrique; elle est appelée «matrice des covariances de $\mathfrak{b}_P$» [6]).

0, 3. *Distributions multinormales.*

0, 31. Rappelons que, si $\mathfrak{b}$ est un vecteur aléatoire multinormal, non dégénéré, de dimension n, si $\mathfrak{b}_P$ est la matrice $n \times 1$, de composantes $\mathbf{b}_1, \ldots, \mathbf{b}_n$, qui le représente par rapport à une base certaine $\mathfrak{P}$, et si l'on pose

$$\mathbf{E}\,\mathfrak{b}_P = \mathfrak{m}_P, \quad \mathbf{C}\,\mathfrak{b}_P = \mathfrak{S}_P ,$$

on a $\operatorname{rg} \mathfrak{S}_P = n$ et

$$\mathbf{Pr}\,[\mathbf{b}_i \leqslant u_i ,\ i = 1, \ldots, n] =$$

$$= (2\pi)^{-n/2} \mid \mathfrak{S} \mid^{-1/2} \int_{-\infty}^{u_1} \ldots \int_{-\infty}^{u_n} \exp\left[-\frac{1}{2}(\mathfrak{x} - \mathfrak{m}_P)^T \mathfrak{S}_P^{-1} (\mathfrak{x} - \mathfrak{m}_P)\right] d\mathfrak{x} .$$

0, 32. Par ailleurs, si $\mathfrak{b}$ est le vecteur décrit ci-dessus, et si $\mathfrak{a}$ est un vecteur aléatoire lié à $\mathfrak{b}$ par une transformation linéaire régulière et certaine:

$$\mathfrak{a} = \mathfrak{A}\,\mathfrak{b} \quad (\operatorname{rg} \mathfrak{A} = n) ,$$

$\mathfrak{a}$ est aussi un vecteur aléatoire multinormal non dégénéré, de dimension n, et on a

$$\mathbf{E}\,\mathfrak{a}_P = \mathfrak{A}_P\,\mathbf{E}\,\mathfrak{b}_P , \qquad \mathbf{C}\,\mathfrak{a}_P = \mathfrak{A}_P\,(\mathbf{C}\,\mathfrak{b}_P)\,\mathfrak{A}_P^T .$$

0, 33. Rappelons encore que, si $\mathfrak{b}$ est comme ci-dessus, les composantes $\mathbf{b}_1, \ldots, \mathbf{b}_n$ de $\mathfrak{b}_P$ sont des variables aléatoires normales; celles-ci sont indépendantes si, et seulement si, $\mathbf{C}\,\mathfrak{b}_P$ est une matrice diagonale [ou, ce qui revient au même, si $\operatorname{cov}(\mathbf{b}_i, \mathbf{b}_j) = 0$, $i \neq j$ [7])]. La réalisation de cette condition, pour un $\mathfrak{b}$ donné, dépend essentiellement du choix de $\mathfrak{P}$. En fait, il est toujours possible de rapporter un vecteur multinormal à une base (certaine) telle que ses composantes soient des aléatoires indépendantes.

0, 34. Rappelons enfin que, si l'on nomme « gaussienne » toute aléatoire normale de moyenne nulle et de variance égale à 1, on a les énoncés suivants (cfr. [III], chap. 18):

a) la somme des carrés de p aléatoires gaussiennes indépendantes est une aléatoire χ^2 à p degrés de liberté (en abrégé, χ^2_p);

b) si $\mathbf{x}$ est une aléatoire gaussienne et $\mathbf{u}$ une aléatoire χ^2_p indépendante de $\mathbf{x}$, le quotient $\mathbf{x}/\sqrt{(\mathbf{u}/p)}$ est une aléatoire de Student à p degrés de liberté (en abrégé, $\mathbf{t}_p$);

c) si $\mathbf{u}$ et $\mathbf{v}$ sont des aléatoires χ^2, respectivement à m et n degrés de liberté, indépendantes, le quotient $(\mathbf{u}/m):(\mathbf{v}/n)$ est une aléatoire $\mathbf{F}$ de Snedecor à (m, n) degrés de liberté (en abrégé, $\mathbf{F}_{m,n}$); il est souvent plus commode d'utiliser alors le fait que $\mathbf{v}/(\mathbf{u}+\mathbf{v})$ est une aléatoire $\beta_{p,q}$ avec $p = n/2$, $q = m/2$, et de se référer aux tables de la distribution β (en général, en effet, $n \geqslant m$, donc $p \geqslant q$, comme dans les tables de Pearson [IV]); on notera que les grandes valeurs de $\mathbf{F}$ correspondent aux petites valeurs de β.

1. Modèles linéaires. Estimateurs.

1, 1. *Définitions.*

1, 11. Considérons une expérience aléatoire dont le résultat est un n-uple ordonné de nombres réels, toutes les valeurs a priori possibles, de $-\infty$ à $+\infty$, étant en effet à prendre en considération. Structurons l'ensemble des observations possibles en un espace vectoriel euclidien sur le corps des réels en postulant que, si $\alpha = (a_1, \ldots, a_n)$ et $\beta = (b_1, \ldots, b_n)$ sont deux observations, on a

pour la somme:

$$\alpha + \beta = (a_1 + b_1, \ldots, a_n + b_n) \ , \tag{1}$$

pour le produit par un scalaire:

$$p\,\alpha = (p\,a_1, \ldots, p\,a_n) \ , \tag{2}$$

pour le produit scalaire:

$$< \alpha, \beta > = a_1\,b_1 + \ldots + a_n\,b_n \ . \tag{3}$$

L'espace des observations, ainsi structuré, sera désigné par V, et son dual par $\mathsf{V}^\star$; V et $\mathsf{V}^\star$ sont évidemment de dimension n.

Les n observations (1, 0, ..., 0), ..., (0, ..., 0, 1) forment évidemment une base de V; elle sera désignée par $\mathfrak{P}$, et sa duale par $\mathfrak{P}^\star$. La base $\mathfrak{P}$ n'a, bien entendu, aucun privilège de droit; son lien particulièrement étroit avec la forme même des observations lui confère cependant un privilège de fait, qui se traduit notamment en ceci, que c'est par rapport à $\mathfrak{P}$ que les relations (1), (2), (3) explicitent la définition des opérations fondamentales de V. Sauf mention expresse du contraire, les représentations de vecteurs de V ou $\mathsf{V}^\star$ seront toujours censées être faites par rapport à $\mathfrak{P}$ ou $\mathfrak{P}^\star$; comme la relation (3) implique que $\mathfrak{P}$ est orthonormale, les conventions du § 0, 2 seront appliquées.

1, 12. A côté de l'espace des observations, nous considérerons un autre espace vectoriel, B, de dimension $p \leqslant n$, dit « espace des paramètres », et son dual, $\mathsf{B}^\star$, « espace des combinaisons paramétriques ». Il n'existe pas, en général, de base « naturelle » qui soit à B ce que $\mathfrak{P}$ est à V; aussi les représentations des vecteurs de B seront-elles notées par un symbole rappelant la base utilisée (la dualité des bases de B et $\mathsf{B}^\star$ allant de soi).

1, 13. Le caractère aléatoire des observations a pour conséquence l'existence d'une catégorie d'épreuve C, munie d'une mesure probabiliste $\mathbf{Pr}$; en tant qu'éléments aléatoires, les observations constituent un vecteur aléatoire de V (application mesurable de C dans V). La mesure $\mathbf{Pr}$ est décrite d'une façon explicite par référence à la base $\mathfrak{P}$, par le postulat que les composantes par rapport à $\mathfrak{P}$ du vecteur aléatoire $\mathfrak{x}$ représentant les observations sont n variables aléatoires normales, indépendantes, de même variance σ^2, dont les moyennes sont n éléments bien déterminés de $\mathsf{B}^\star$:

$$\mathbf{E}\, e_i^\star\, \mathfrak{x} \equiv e_i^\star \int_{\mathsf{C}} \mathfrak{x}(\lambda)\, \mathrm{Pr}\,(d\lambda) = e_i^\star\, \mathfrak{A}\, \mathfrak{b} \quad (\mathfrak{x} \in \mathsf{V},\ \mathfrak{b} \in \mathsf{B}) \ ,$$

l'application $\mathfrak{A}$ étant de rang r ($\leqslant p$).

On voit donc que l'opérateur « valeur moyenne dans C » induit une application linéaire de $\mathsf{V}^\star$ dans $\mathsf{B}^\star$; si on la désigne par $\mathfrak{E}$, on a, pour tout $\mathfrak{l}^\star \in \mathsf{V}^\star$,

$$\mathbf{E}\,(\mathfrak{l}^{\star}\,\mathfrak{x}) = \mathfrak{l}^{\star}\,\mathbf{E}\,\mathfrak{x} = \mathfrak{l}^{\star}\,\mathfrak{A}\,\mathfrak{b}\ , \qquad (4)$$

d'où

$$\mathfrak{E}:\ \mathfrak{l}^{\star} \to \mathfrak{l}^{\star}\,\mathfrak{A}\,\mathfrak{b}\ .\ ^{8})$$

$\mathfrak{E}$ est évidemment de rang r.

1, 14. Le noyau de $\mathfrak{E}$ (sous-espace V_0 de $\mathsf{V}^{\star}$ formé des vecteurs $\mathfrak{l}^{\star}$ tels que $\mathfrak{E}\,\mathfrak{l}^{\star}$ soit identiquement nul [9]) est appelé «espace des erreurs» [10]); il est de dimension n — rg $\mathfrak{E} = n - r$.

Le complément orthogonal de V_0 (sous-espace V_+ de $\mathsf{V}^{\star}$ formé des vecteurs $\mathfrak{m}^{\star}$ pour lesquels $\mathfrak{m}^{\star}\,\mathfrak{l} = 0$ pour tout $\mathfrak{l}^{\star} \in \mathsf{V}_0$) est appelé « espace des estimatrices ». Une estimatrice est donc, par définition, une fonctionnelle linéaire des observations, orthogonale à toute fonctionnelle dont la moyenne est identiquement [9]) nulle. On remarquera que, $\mathsf{V}^{\star}$ étant somme directe de V_0 et V_+, à toute fonctionnelle linéaire des observations dont la moyenne n'est pas identiquement [9]) nulle correspond une et une seule fonctionnelle de V_+ ayant identiquement [9]) même moyenne qu'elle.

Enfin, l'image de $\mathfrak{E}$ (sous-espace B_+ de $\mathsf{B}^{\star}$ formé des combinaisons paramétriques $\mathfrak{b}^{\star}$ pour lesquelles il existe un vecteur $\mathfrak{l}^{\star} \in \mathsf{V}^{\star}$ tel que $\mathbf{E}\,\mathfrak{l}^{\star}\,\mathfrak{x} = \mathfrak{b}^{\star}\,\mathfrak{b}$) s'appelle « espace des combinaisons (paramétriques) estimables ». Nous noterons B_0 un complément quelconque de B_+.

On sait que la restriction de $\mathfrak{E}$ à V_+ est un isomorphisme de V_+ sur B_+; on peut donc énoncer que

toute combinaison estimable est la moyenne d'une et une seule estimatrice, et réciproquement.

1, 2. *Estimateurs privilégiés.*

1, 21. Si l'on a, pour $\mathfrak{l}^{\star} \in \mathsf{V}^{\star}$, $\mathbf{E}\,\mathfrak{l}^{\star}\,\mathfrak{x} = \mathfrak{f}^{\star}\,\mathfrak{b}$, $\mathfrak{f}^{\star}\,\mathfrak{b}$ est évidemment une combinaison estimable, et $\mathfrak{l}^{\star}\,\mathfrak{x}$ en est un estimateur fidèle (au sens de la théorie statistique de l'estimation); si $\mathfrak{m}^{\star} \in \mathsf{V}_0$, $(\mathfrak{l}^{\star} + \mathfrak{m}^{\star})\,\mathfrak{x}$ est aussi un estimateur fidèle de $\mathfrak{f}^{\star}\,\mathfrak{b}$; pour distinguer, parmi tous ces estimateurs fidèles de $\mathfrak{f}^{\star}\,\mathfrak{b}$, l'unique estimatrice, celle-ci est dite « estimateur privilégié de $\mathfrak{f}^{\star}\,\mathfrak{b}$ », et désignée par $\mathfrak{l}^{\star}_{f}\,\mathfrak{x}$ (donc, par définition, $\mathbf{E}\,\mathfrak{l}^{\star}_{f}\,\mathfrak{x} = \mathfrak{f}^{\star}\,\mathfrak{b}$ et $\mathfrak{l}^{\star}_{f} \in \mathsf{V}_+$).

1, 22. Théorème. *Parmi tous les estimateurs fidèles de la combinaison estimable* $\mathfrak{f}^\star\,\mathfrak{b}$, *l'estimateur privilégié a la variance minimum.*

Soit en effet $\mathfrak{m}^\star$ tel que $\mathbf{E}\,\mathfrak{m}^\star\,\mathfrak{x} = \mathfrak{f}^\star\,\mathfrak{b}$. On a, en posant $\tilde{\mathfrak{x}} = \mathfrak{x} - \mathbf{E}\,\mathfrak{x}$,

$$\begin{aligned}\operatorname{var}\mathfrak{m}^\star\,\mathfrak{x} &= \mathbf{E}\,[\mathfrak{m}^\star\,\mathfrak{x} - \mathbf{E}\,\mathfrak{m}^\star\,\mathfrak{x}]\,[\mathfrak{m}^\star\,\mathfrak{x} - \mathbf{E}\,\mathfrak{m}^\star\,\mathfrak{x}] \\ &= \mathbf{E}\,(\mathfrak{m}^\star\,\tilde{\mathfrak{x}})\,(\tilde{\mathfrak{x}}^\star\,\mathfrak{m})^\star = \mathfrak{m}^\star\,(\mathbf{E}\,\tilde{\mathfrak{x}}\,\tilde{\mathfrak{x}}^\star)\,\mathfrak{m}\,. \qquad (5)\end{aligned}$$

La définition des propriétés distributionnelles de $\mathfrak{x}$ faisant intervenir la base $\mathfrak{P}$, introduisons cette base pour un calcul explicite:

$$\mathbf{E}\,\tilde{\mathfrak{x}}\,\tilde{\mathfrak{x}}^\star = \mathbf{E}\,\tilde{\mathfrak{x}}_P\,(\tilde{\mathfrak{x}}_P)^{\mathrm{T}} = \mathbf{E}\begin{Vmatrix} \tilde{\mathfrak{x}}_1^2 & \tilde{\mathfrak{x}}_1\,\tilde{\mathfrak{x}}_2 \ldots & \tilde{\mathfrak{x}}_1\,\tilde{\mathfrak{x}}_n \\ \vdots & & \\ \tilde{\mathfrak{x}}_n\,\tilde{\mathfrak{x}}_1 \ldots & & \tilde{\mathfrak{x}}_n^2 \end{Vmatrix}$$

$$= \mathfrak{I}_n\,\sigma^2$$

d'où

$$\operatorname{var}\mathfrak{m}^\star\,\mathfrak{x} = (\mathfrak{m}^\star\,\mathfrak{m})\,\sigma^2\,.$$

Soit alors

$$\mathfrak{m}^\star = \mathfrak{l}_f^\star + \mathfrak{m}_0^\star\,,$$

de sorte que

$$\mathfrak{l}_f^\star\,\mathfrak{m}_0 = \mathfrak{m}_0^\star\,\mathfrak{l}_f = 0\,;$$

on a

$$\operatorname{var}\mathfrak{m}^\star\,\mathfrak{x} = (\mathfrak{l}_f^\star\,\mathfrak{l}_f)\,\sigma^2 + (\mathfrak{m}_0^\star\,\mathfrak{m}_0)\,\sigma^2 \geqslant (\mathfrak{l}_f^\star\,\mathfrak{l}_f)\,\sigma^2 = \operatorname{var}\mathfrak{l}_f^\star\,\mathfrak{x},$$

l'inégalité étant d'ailleurs stricte si $\mathfrak{m}_0^\star \neq 0$.

1, 3. *Exécution des calculs.*

1, 31. Pour l'exécution effective des calculs, il importe d'introduire une base dans chacun des espaces considérés; dans ce paragraphe, V et $\mathsf{V}^\star$ sont rapportés aux bases $\mathfrak{P}$ et $\mathfrak{P}^\star$, B et $\mathsf{B}^\star$ sont rapportés à des bases déterminées $\mathfrak{H}$ et $\mathfrak{H}^\star$.

1, 32. Pour que $\mathfrak{l}^\star \in \mathsf{V}_0$, il est nécessaire et suffisant que $\mathfrak{l}^\star\,\mathfrak{A} = 0$; donc

l'espace des erreurs est engendré par ceux des vecteurs de $\mathsf{V}^\star$ *qui sont orthogonaux aux colonnes de la matrice* $\mathfrak{A}$ *(plus explicitement:* $\mathfrak{A}_{P,H}$*).*

Il en résulte immédiatement que l'espace des estimatrices est engendré par les lignes de la matrice $\mathfrak{A}^T$, donc

toute estimatrice est de la forme $\mathfrak{l}^\star\, \mathfrak{A}^T\, \mathfrak{x}$.

Comme $\mathbf{E}\, \mathfrak{l}^\star\, \mathfrak{A}^T\, \mathfrak{x} = \mathfrak{l}^\star\, \mathfrak{A}^T\, \mathbf{E}\, \mathfrak{x} = \mathfrak{l}^\star\, \mathfrak{A}^T\, \mathfrak{A}\, \mathfrak{b}_H$, cet énoncé, à son tour, entraîne celui-ci:

toute combinaison estimable est de la forme $\mathfrak{l}^\star\, \mathfrak{A}^T\, \mathfrak{A}\, \mathfrak{b}_H$ *et réciproquement.*

Ainsi la correspondance biunivoque entre estimatrices et combinaisons estimables est clairement mise en évidence:

a) si $\mathfrak{f}^T\, \mathfrak{b}_H$ est estimable, il existe nécessairement un vecteur $\mathfrak{m}_f^T \in \mathbf{B}^\star$ tel que $\mathfrak{f}^T = \mathfrak{m}_f^T\, \mathfrak{A}^T\, \mathfrak{A}$;

b) l'estimateur privilégié de $\mathfrak{f}^T\, \mathfrak{b}_H$ est alors $\mathfrak{m}_f^T\, \mathfrak{A}^T\, \mathfrak{x}$ (de sorte que $\mathfrak{l}_f^\star = \mathfrak{m}_f^T\, \mathfrak{A}^T$);

c) la variance de cet estimateur vaut $[\mathfrak{m}^T\, (\mathfrak{A}^T\, \mathfrak{A})\, \mathfrak{m}]\, \sigma^2$.

1, 33. Supposons que $r = p$. $\mathfrak{A}$, $\mathfrak{A}^T$, et $\mathfrak{A}^T\, \mathfrak{A}$ sont alors des matrices de rang p, et le système de dimension p en l'inconnue $\hat{\mathfrak{b}}_H$

$$\mathfrak{A}^T\, \mathfrak{A}\, \hat{\mathfrak{b}}_H = \mathfrak{A}^T\, \mathfrak{x} \qquad (6)$$

détermine entièrement cette inconnue. Celle-ci jouit de la précieuse propriété que voici:

l'estimateur privilégié de $\mathfrak{f}^T\, \mathfrak{b}_H$ *n'est autre que* $\mathfrak{f}^T\, \hat{\mathfrak{b}}_H$.

En effet, soit $\mathfrak{f}^T = \mathfrak{m}_f^T\, \mathfrak{A}^T\, \mathfrak{A}$; on a

$$\mathfrak{l}_f^\star\, \mathfrak{x} = \mathfrak{m}_f^T\, \mathfrak{A}^T\, \mathfrak{x} = (\mathfrak{m}_f^T\, \mathfrak{A}^T\, \mathfrak{A})\, \hat{\mathfrak{b}}_H = \mathfrak{f}^T\, \hat{\mathfrak{b}}_H\,, \quad \text{q.e.d.}$$

1, 34. Si $r < p$, le système (6) ne détermine pas univoquement l'inconnue $\hat{\mathfrak{b}}_H$. Pourtant, il reste vrai que, quelle que soit la détermination choisie pour $\hat{\mathfrak{b}}_H$, l'estimateur privilégié de la combinaison estimable $\mathfrak{f}^T\, \mathfrak{b}_H$ est $\mathfrak{f}^T\, \hat{\mathfrak{b}}_H$; en d'autres termes, les premiers membres de (6) mettent en évidence r combinaisons estimables particulières qui constituent une base de $\mathbf{B}_+$.

Le système (6) est dit « système normal »; il faut évidemment se garder d'y vouloir introduire l'inverse de $\mathfrak{A}^T\, \mathfrak{A}$ lorsque $r < p$ (cfr. § 3).

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

Si les vecteurs $v_1^\star, \ldots v_t^\star$ $(t \geqslant s)$ engendrent $\mathsf{U}^\star$, on écrit, d'ordinaire, $\{ v_1^\star, \ldots, v_t^\star \}$ pour $\mathsf{U}^\star$; on écrira donc aussi $\mathbf{SC}\{ v_1^\star, \ldots, v_t^\star \}$ 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 $u_1^\star, \ldots, u_s^\star$. La projection orthogonale x_u de x sur U est alors définie par les relations

$$x_u = \sum \lambda_i u_i \ , \ < x - x_u \ , \ u_k > \equiv u_k^\star (x - x_u) = 0 \quad (k = 1, \ldots, s)$$

d'où l'on tire

$$\sum_{i=1}^{s} \lambda_i u_i^\star u_k = u_k^\star x \quad (k = 1, \ldots, s) \ , \tag{7}$$

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

$$\mathbf{SC}\,\mathsf{U}^\star = x_u^\star x_u = \left(\sum_1^s \lambda_i u_i^\star \right) \left(\sum_1^s \lambda_k u_k \right)$$

$$= \sum_1^s \sum_1^s \lambda_i \lambda_k u_i^\star u_k$$

moyennant (7), ce qui entraîne

$$\mathbf{SC}\,\mathsf{U}^\star = \sum_1^s \lambda_k u_k^\star x \ . \tag{8}$$

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

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

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

$$x_u^\star x_u = x_{u_1}^\star x_{u_1} + x_{u_2}^\star x_{u_2} \ ,$$

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 composantes, 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 $\mathbf{E}\,\mathfrak{w}_i^\star \mathfrak{x} = 0$ $(i = 1, \ldots, s)$ (c'est-à-dire sous l'hypothèse $\mathfrak{w}_1^\star \mathfrak{A} \mathfrak{b} = \ldots \mathfrak{w}_s^\star \mathfrak{A} \mathfrak{b} = 0$), 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** $\mathfrak{f}^\star \mathfrak{b} = a$, $\left[(\mathfrak{l}_f^\star \mathfrak{x} - a)/\sigma \,\middle/ \sqrt{(\mathfrak{l}_f^\star \mathfrak{l}_f)}\right]$ 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 $(1/\sigma^2)\,\mathbf{SC}\{\mathfrak{l}_{f_1}^\star, ..., \mathfrak{l}_{f_s}^\star\}$ est une aléatoire χ_s^2; cela entraîne que

$$\mathbf{Q} = \frac{\mathbf{SC}\{\mathfrak{l}_{f_1}^\star, ..., \mathfrak{l}_{f_s}^\star\}/s}{\mathbf{SC}\,\mathrm{E}/(n-r)}$$

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

$$\mathbf{E}\,\mathfrak{x} = \mathfrak{A}\,\mathfrak{W}^{-1}\,\mathfrak{Z}\,\mathfrak{w}\,, \qquad \mathfrak{Z} = \begin{Vmatrix} \mathfrak{I}_q & 0 \\ 0 & 0 \end{Vmatrix}\,. \Big]$$

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]):

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

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

On a alors

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

avec

$$\begin{aligned} \mathbf{red}\,\big[\mathsf{U}^{\star}_{r-q} \,\big|\, \mathsf{U}^{\star}_q\big] &= \mathbf{SC}\,\mathsf{U}^{\star}_{r-q}\,, \\ \mathbf{red}\,\big[\mathsf{U}^{\star}_q\big] &= \mathbf{SC}\,\mathsf{U}'_q \neq \mathbf{SC}\,\mathsf{U}^{\star}_q\,. \end{aligned}$$

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,

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

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

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

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

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

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

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

et on a

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

avec

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

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, $\mathbf{E}\,\mathfrak{y} = \mathfrak{A}\mathfrak{b}$. 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 $\sigma^2\,\gamma_s^2$. Mais **SCE** n'est plus distribuée comme $\sigma^2\,\gamma_{n-r}$, 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 **SC**int (« somme de carrés interne ») l'expression $\mathbf{SC}\,\mathsf{V}_*$, et **SC**EM [« somme de carrés des écarts au modèle » (sous-entendu: au modèle restreint)] l'expression **SC**E — **SC**int (en désignant par V'_* le complément orthogonal de V_* dans V_0, **SC**EM n'est autre que $\mathbf{SC}\,\mathsf{V}'_*$). Dans les considérations du § 2, 2, **SC**int peut remplacer **SC**E, $n - s$ remplaçant alors $n - r$. Les composantes de **SC**int sont évidemment orthogonales à celles de **SC**N.

Remarque. — Les composantes additives de **SC**T (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	$SCT - SCN$	$n - r$
SCint	$SC\,\mathsf{V}_*$	u
SCEM	$SCE - SCint$	$n - r - u$.