0. Préliminaires.

Objekttyp: **Chapter**

Zeitschrift: **L'Enseignement Mathématique**

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

Heft 1: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **24.05.2024**

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

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

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

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, $\mathbf{x}$, $\mathbf{b}$, $\mathbf{A}$ 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, $\mathbf{A}$, de noter $\mathbf{E}\,\mathbf{A}$ la matrice dont l'élément (i, j) est la valeur moyenne de l'élément (i, j) de $\mathbf{A}$:

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

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 $\mathbf{b}$ un vecteur aléatoire, $\mathbf{b}_P$ 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

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

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

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_p^2);

b) si $\mathbf{x}$ est une aléatoire gaussienne et $\mathbf{u}$ une aléatoire χ_p^2 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:

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