

# JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

JSFS

## Bibliographie

*Journal de la société statistique de Paris*, tome 122, n° 4 (1981), p. 245-247

<[http://www.numdam.org/item?id=JSFS\\_1981\\_\\_122\\_4\\_245\\_0](http://www.numdam.org/item?id=JSFS_1981__122_4_245_0)>

© Société de statistique de Paris, 1981, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## IV

### BIBLIOGRAPHIE

Michel VOLLE, *Analyse des données* (2<sup>e</sup> édition 1981). Economica.

Le livre de Michel Volle en est déjà à sa deuxième édition alors que la première ne remonte qu'à 1978. C'est dire le succès qu'a remporté l'ouvrage. L'édition de 1981 a grossi de 60 pages et s'est enrichie de trois nouveaux chapitres. Les notations ont été rendues plus conformes à ce qu'on rencontre habituellement, la présentation matérielle s'est améliorée (la composition a remplacé l'offset). L'usage systématique d'encadrés, de passages barrés en marge rend la lecture agréable et facile.

Le livre est issu du cours enseigné par l'auteur à l'E.N.S.A.E. et au C.E.P.E. Il s'adresse donc à un public ayant une certaine habitude du calcul mais pas nécessairement une culture mathématique ou statistique très étendue. La difficulté de l'ouvrage se situe entre celle du « Que sais-je » de G. Saporta et celle des ouvrages de référence classique de J.P. Benzecri ou de F. Caillez et J.P. Pages.

Les principaux apports de cette édition — outre les chapitres nouveaux sur l'analyse des correspondances multiples et l'analyse canonique généralisée — sont l'introduction générale et les introductions de chaque chapitre, où l'auteur, à la lumière de sa pratique du « métier de statisticien » situe l'importance des différentes méthodes et la place de l'analyse des données dans la pratique du statisticien professionnel.

Le livre est divisé en trois parties : Généralités (4 chapitres en 80 pages), Méthodes factorielles (7 chapitres en 160 pages dont 50 consacrées à l'analyse des correspondances), Classification automatique (6 chapitres en 70 pages).

Dans la première partie l'auteur présente ses réflexions sur la statistique et l'analyse des données. Il donne un exemple introductif très pédagogique. Le chapitre sur la théorie de l'information est en revanche très discutable. Son objectif — manqué — était d'introduire la métrique du chi-2 de façon naturelle. En fait, cette démarche conduit naturellement, comme le prouvent les travaux récents de Küllback, à l'estimation par le maximum de vraisemblance des modèles log-linéaires. L'intérêt d'un tel chapitre pourrait donc être une comparaison de cette approche avec l'A.F.C. ou l'A.C.M. mais on sortirait alors du cadre du livre.

La troisième partie relative aux méthodes de classification, ne diffère de la première édition que par l'ordre des chapitres et les titres qui leurs sont attribués. Ceux-ci sont devenus plus généraux ce qui est justifié par les discours introductifs à chacun des chapitres, le contenu technique étant la description d'une méthode particulière au sein d'une famille de méthodes. Devant le foisonnement grandissant des algorithmes, l'auteur a sans aucun doute choisi la meilleure option en présentant des méthodes relativement anciennes, ayant prouvé leur utilité et en les situant dans un cadre plus général. L'utilisation du terme « espace métrique » dans le titre du chapitre XV est malheureux — « espace euclidien » eut été plus précis — car elle peut entraîner des confusions avec les techniques dites « d'analyse de tableaux de distance ».

La seconde partie (méthodes factorielles) débute par un exposé général de ces méthodes (recherche d'un sous-espace conservant le maximum de variance), s'enchaînant de façon logique sur un court chapitre relatif à l'analyse en composantes principales. Les deux chapitres pourraient sans difficulté être fondus en un seul. On introduit sommairement les notions de dualité.

Suit le cœur du livre, un exposé « à la Benzecri » de l'analyse factorielle des correspondances d'un tableau de contingence. La façon de voir est donc très géométrique (nuages de points); de ce fait, peut-être, l'auteur ignore la présentation de l'A.F.C. comme recherche des corrélations canoniques entre deux variables qualitatives. C'est dommage car cette façon de voir les choses — techniquement très élémentaire — permet de justifier beaucoup de choses (la métrique, l'interprétation de proximité entre « points lignes » et « points-colonnes », les rotations d'axes, les valeurs propres inférieures à 1 et permet de passer facilement aux correspondances multiples. L'application des méthodes factorielles aux séries chronologiques reste assez naïve.

Ces réticences une fois exprimées, il reste un exposé alliant la clarté à la concision. Les nombreux exemples sont bien choisis et très éclairants.

Les correspondances multiples ont droit à un chapitre autonome dans cette édition. Ce n'est que justice car les travaux de L. Lebart et N. Tabard ont démontré que l'originalité et l'utilité de cette méthode étaient au moins aussi grandes que celles de l'A.F.C.

L'auteur passe un peu rapidement sur l'analyse discriminante compte tenu des nombreuses applications de cette technique, notamment en médecine, en météorologie, en marketing ou en gestion et de l'intérêt dont elle est l'objet à l'étranger (aux U.S.A. et en U.R.S.S. tout particulièrement).

L'analyse canonique et l'analyse canonique généralisée sont traitées comme des curiosités mathématiques aux propriétés unifiantes. Il est dommage que M. Volle n'ait pas profité de ce chapitre pour examiner les rapports entre les méthodes factorielles et l'analyse multivariée « traditionnelle ».

L'ouvrage de M. Volle mérite son succès et aura de nombreux et (nouveaux) lecteurs en raison de sa clarté, de sa présentation attirante et de son volume relativement faible. C'est un bon cours d'analyse des données. Il lui manque cependant les références bibliographiques nécessaires pour approfondir la plupart des questions abordées. Les seuls travaux cités sont des livres français; les articles parus par exemple dans les *Cahiers du BUIRO*, la *Revue de statistique appliquée* ou *Statistique et analyse des données* sont « oubliés ». La littérature anglo-saxonne l'est également, ce qui est dommage, ne serait-ce que pour montrer que les francophones ne sont pas nécessairement chauvins. On pourrait, à cet égard, souhaiter que la troisième édition, sans doute prochaine, reprenne au moins la présentation de Spearman (1903) de l'« analyse en facteurs communs et spécifiques » dont les influences historiques, théoriques et pratiques se font encore sentir de nos jours.

J.-C. DEVILLE

M.B. PRIESTLEY, *Spectral Analysis and Time Series. Volume 2 : Multivariate series, prediction and control. Academic Press (1981)*.

Ce livre est la suite du traité de Priestley sur les séries chronologiques dont le premier tome était paru il y a quelques années. Son contenu était relativement classique : séries univariées, analyse spectrale, estimation dans les domaines temporels et fréquentiels. Il couvrait 653 pages non compris introductions, table des matières et index qu'on retrouve dans ce tome. Le nouveau livre comporte 237 pages (numérotées de 654 à 890) et comprend trois chapitres :

Chapitre 9 : Processus multivariés (73 pages)

Chapitre 10 : Prédiction, filtrage et contrôle (89 pages)

Chapitre 11 : Processus non stationnaires ou non linéaires (75 pages).

Beaucoup des sujets qu'aborde l'ouvrage sont proches de la recherche ou de l'expérimentation. L'optique générale est de réaliser la synthèse entre la littérature économétrique (qui privilégie les aspects temporels) et la littérature des spécialistes de la théorie du signal (qui privilégie les aspects fréquentiels). Chaque fois que cela est possible l'auteur essaie de mettre en rapport les deux approches tant en ce qui concerne les méthodes que les notations ou le vocabulaire.

Les techniques mathématiques utilisées restent axées sur l'algèbre et le calcul, l'auteur ayant souvent recours à des hypothèses simplificatrices. La plupart des sujets sont traités de façon complète pour les processus en temps discret, le cas continu faisant l'objet d'une présentation par analogie. Le point de vue géométrique (ou hilbertien) n'est presque jamais utilisé, ce qui est dommage. En particulier, dans les sections où intervient la notion de composantes principales, le recours à des « équations déterminantes » est assez agaçant.

L'ouvrage sera utile à des étudiants de 3<sup>e</sup> cycle des universités ou à des chercheurs dans le domaine des séries chronologiques. Il intéressera aussi les praticiens de l'économétrie ou de la théorie du signal dans la mesure où il rapproche ces deux disciplines.

Le chapitre 9 commence par une théorie assez rapide des processus stationnaires vectoriels. Il développe ensuite assez longuement les propriétés et l'estimation des modèles ARMA multivariés. Il se termine par quelques indications sur les processus stationnaires univariés où l'espace des « temps » est multidimensionnel.

Le chapitre 10 aborde trois sujets relativement liés. Les paragraphes sur la prévision des processus stationnaires exposent des résultats classiques dans une présentation qui l'est assez peu : la décomposition de Wold y apparaît comme un sous produit du problème de prévision et non comme un outil. Suit un exposé très clair sur la prévision dans les modèles de Box-Jenkins. On passe ensuite aux questions relatives au contrôle et à l'estimation des systèmes linéaires asservis. Quelques pages sont consacrées au filtrage des processus multidimensionnels — ARMA en particulier — qui pourront intéresser les spécialistes d'analyse des données. L'auteur termine par une dizaine de pages sur le filtre de Kalman.

Le chapitre 11 fait une synthèse des travaux de Priestley sur la notion de spectre évolutif pour les processus non stationnaires. Les approches de quelques concurrents sont rapidement présentées, mais les travaux du « groupe de Bruxelles » — Melard, Hallin, Dreesbeke, etc. — sont très injustement oubliés. La fin du livre, consacrée aux modèles non linéaires, est très expérimentale et donne de nombreuses pistes de recherche.

Ce livre contient une quantité d'information non négligeable sur des sujets qui sont l'objet de travaux très actuels. L'auteur donne une abondante bibliographie de 21 pages comportant près de 500 références. On constate cependant qu'une dizaine d'auteurs seulement — dont Priestley — fournissent environ 40 % des titres. Par ailleurs la bibliographie est anglophone uniquement à l'exception des « préhistoriens allemands » Kolmogorov (1933) et Khint-

chine (1934) et de trois titres français : Loeve (1945); Blanc-Lapierre, Fortet (1946) et Pham-Dinh (1975) pour son Ph. D of the University of Grenoble.

J.-C. DEVILLE

Christine WATTELAR, *Perspectives démographiques par sexe et par âge : les indices de mortalité et le calcul des survivants*. Département de démographie. Université catholique de Louvain. Juin 1980. Cabay, libraire-éditeur, Louvain-la-Neuve, Belgique.

L'objectif de cette étude étant essentiellement méthodologique, on a adopté comme plan de travail la séquence de la méthode des composants. En introduction, on a fait un bref rappel des caractéristiques de cette approche et l'ébauche d'un organigramme permettra de situer la place exacte qu'occupent les calculs de décès et de survie (ainsi que l'établissement des hypothèses les concernant) dans le processus perspectif classique, avant toute prise en considération de la migration.

L'étude est subdivisée en deux grandes parties : l'une est consacrée à la construction des indices perspectifs par âge (nombres relatifs), à partir des tables de mortalité existantes ou tirés de mesures isolées (quotients ou taux); l'autre partie porte sur le calcul proprement dit des effectifs survivants (nombres absolus réels) dans une situation de mortalité constante ou avec des hypothèses de mortalité variable. Les diverses formulations proposées et leurs postulants sont enfin mis à l'épreuve sur un échantillon de données concrètes que l'on soumet à un certain nombre de simulations. Une appréciation générale des véritables risques d'erreur en matière de prévision de décès et de calcul des survivants (avant l'intervention du mouvement migratoire dans le processus perspectif) est tentée à la lumière des résultats ainsi obtenus.

Josianne DUCHENE, *Un essai de modélisation de la répartition des décès selon l'âge et la cause dans les pays industrialisés*. Département de démographie. Université catholique de Louvain, mars 1980. Cabay, libraire-éditeur, Louvain-la-Neuve, Belgique.

Le premier chapitre de ce livre présente une synthèse des principaux modèles de mortalité sous trois rubriques traitant des lois et ajustements analytiques de la mortalité générale, des tables-types de mortalité et enfin de la prise en compte des composantes de la mortalité et des causes de décès. Le 2<sup>e</sup> chapitre est consacré à la présentation des données qui ont servi de base à cette étude. Le 3<sup>e</sup> chapitre présente les techniques de classification automatique appliquées par les auteurs pour réunir les causes de mortalité qui ont même distribution selon l'âge des décès qui leur sont attribués et la typologie statistique des causes de décès établie. Contrairement à la plupart des regroupements de causes de décès utilisés en démographie, les causes de décès ont été réunies dans cette étude sur la base de critères statistiques sans tenir compte de l'existence d'une proximité médicale entre ces causes.

Le chapitre 4 est consacré à la recherche d'un procédé analytique permettant de caractériser, au moyen d'un petit nombre de paramètres, le profil selon l'âge propre à chaque pays de la mortalité attribuée à chacun des groupes de causes établis précédemment.

Une dernière partie dresse un bilan des principaux résultats et pose les jalons d'une recherche destinée à permettre leur utilisation dans l'identification des principales composantes de la mortalité.

Jacques HENRIPIN, Paul-Marie HUOT, Évelyne LAPIERRE-ADAMCYK, Nicole MARCIL-GRATTON, *Les enfants qu'on n'a plus au Québec*. Département de démographie de l'Université de Montréal, 1981, 412 p. (17 × 25). Les Presses de l'Université de Montréal, 2910, boulevard Édouard-Montpetit, Montréal, Québec, Canada, H3T 1J7.

Cet ouvrage fait suite à *La Fin de la revanche des berceaux : qu'en pensent les Québécoises?*

Entre 1971 et 1976, les couples au Québec ont réduit sensiblement le nombre d'enfants qu'ils s'attendaient d'avoir et les plus jeunes d'entre eux arriveront à peine à assurer le remplacement de leur génération. Cette réduction de la fécondité s'appuie sur un usage très généralisé d'une contraception efficace, en particulier, la stérilisation, à laquelle ont eu recours le tiers des couples interrogés.

Un enfant de plus peut-il être envisagé dans certains cas et moyennant une aide de l'État? Il semble bien que non. Cet enfant supplémentaire a contre lui de sérieux obstacles : ressources économiques, santé jugée fragile et désir de liberté. Et il semble qu'aucune mesure ne puisse les vaincre dans la majorité des cas.