

JSFS

Comptes rendus de lecture

Journal de la société française de statistique, tome 145, n° 3 (2004),
p. 89-91

http://www.numdam.org/item?id=JSFS_2004__145_3_89_0

© Société française de statistique, 2004, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<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/>

COMPTES RENDUS DE LECTURE

Statistique La théorie et ses applications

Michel Lejeune
1 vol., 339 pages, Springer-Verlag France, Paris, 2004,
ISBN 2-287-21241-8

L'ouvrage présente les bases de la statistique inférentielle depuis l'introduction des outils probabilistes nécessaires jusqu'à la régression (linéaire, logistique, non paramétrique), en passant par les grands principes de l'estimation et les principales procédures de test. De nombreux ouvrages répondent certes à cette description sommaire, mais celui de Michel Lejeune vient à son heure occuper une place tout à fait originale et répondre à un besoin certain. Essayons d'expliquer en quoi et pourquoi. En premier lieu il y a le style adopté qui permet plusieurs niveaux de lecture. Une certaine dose de mathématiques est nécessaire pour comprendre les mécanismes profonds de la modélisation et de l'analyse statistiques, mais un bon baccalauréat scientifique – et quand même une certaine motivation – seront ici suffisants à qui se contentera de l'essentiel, un niveau de première ou deuxième année d'université permettant d'entrer dans un certain nombre de raffinements, la présentation de l'essentiel étant complétée par des notes plus techniques et par des exercices. Pour tous cependant les grands principes méthodologiques sont expliqués de façon précise et très accessible, quoique sans concession. Et c'est, je crois, le point important de l'ouvrage : alors que les logiciels statistiques véhiculent les techniques, on explique ici clairement leur portée et leurs limites. À maints égards d'ailleurs l'ouvrage est visiblement écrit en pensant à l'utilisateur des outils informatiques ; par exemple dans sa présentation détaillée des techniques fondées sur la vraisemblance (estimation ou tests), par la place accordée aux méthodes modernes exigeant des calculs intensifs (*jackknife*, *bootstrap*), par tout ce qui concerne l'estimation fonctionnelle et les divers usages que l'on peut en faire, par l'introduction de la régression logistique, etc. Bref, comme son titre l'indique, l'ouvrage fait le lien entre théorie (j'aurais peut-être dit méthodologie) et application, sans pédanterie inutile dans l'exposé mathématique mais sans masquer l'essentiel ; ainsi, seules les grandes lignes des démonstrations sont souvent données et toutes les hypothèses techniques nécessaires ne sont pas listées en détail mais, en revanche, les hypothèses fondamentales pour l'utilisateur sont soigneusement soulignées et expliquées, en insistant sur leur degré d'importance (quel test peut être considéré comme

robuste à la non normalité et quel ne l'est pas, à partir de quelle valeur de n peut-on se considérer dans la situation asymptotique?...).

On l'aura compris, j'ai pris grand plaisir à lire cet ouvrage et je le recommande à un très large public. Il constituera un excellent manuel pour nombre d'enseignements de base, mais aussi je suis certain que les collègues, aussi bien enseignants que praticiens, auront à leur tour plaisir à le recommander autour d'eux tant il correspond à bien des attentes qu'ils ont certainement entendu s'exprimer... sans compter les besoins non conscients que nous devrions tous nous employer à mettre au jour !

Henri Caussin

Du groupe à l'individu Synthèse multiniveau

Daniel Courgeau
1 vol., 242 pages, Éditions de l'Institut National d'Études
Démographiques, Paris, 2004,
ISBN 2 7332-2016-0

Notre collègue Daniel Courgeau a récemment édité un livre collectif, en anglais, sur l'approche multiniveau dans les sciences sociales (un compte rendu de lecture a été fait dans ce Journal, vol. 144, n° 4, 2003). Son nouvel ouvrage approfondit la question, en français, dans le domaine privilégié de l'auteur : la démographie. La première partie du livre est avant tout historique et épistémologique. On y décrit et analyse les étapes successives de la pensée (et des moyens mis en œuvre) : l'analyse transversale instantanée (chapitre 1), puis l'analyse longitudinale, travaillant sur les générations, qui introduit l'aspect temporel mais relève encore de l'histoire collective (chapitre 2), enfin l'analyse des données individuelles (chapitre 3) qui change totalement de point de vue et fait d'ailleurs appel à une méthodologie statistique toute nouvelle. L'examen de ces diverses approches fait ressortir leurs faiblesses respectives (erreurs écologiques ou atomistes liées respectivement aux points de vue holistique ou individualiste) et conduit à introduire l'analyse « multiniveau » (chapitre 4). La seconde partie de l'ouvrage présente les aspects techniques de l'analyse statistique multiniveau, les modèles utilisés, les techniques d'estimation correspondantes et leur mise en œuvre sur des exemples variés. Après un chapitre consacré à la définition et la discussion des divers types de niveaux, le chapitre 6 concerne les données continues et l'analyse linéaire ; on voit que les modèles linéaires mixtes, même s'ils n'ont

pas ici l'importance qu'ils peuvent trouver dans d'autres applications, ont néanmoins toute leur place en démographie. Le chapitre 7 traite ensuite de caractéristiques discrètes, les plus fréquentes en démographie (données binaires ou polytomiques, décomptes) et des modèles correspondants, qui ressortissent essentiellement au modèle linéaire généralisé à effets mixtes, fixes et aléatoires. Le chapitre 8 est enfin consacré à l'analyse biographique multiniveau dans laquelle l'unité statistique est «l'événement»; l'individu subissant une histoire événementielle est alors une unité de niveau deux. Tous les modèles présentés sont mis en œuvre sur des exemples très détaillés afin d'en montrer utilité et limites. En plus de deux index, l'ouvrage comporte un glossaire des termes épistémologiques et une liste des principaux logiciels utilisables pour la mise en œuvre des modèles présentés. Ces deux derniers points soulignent bien le désir de l'auteur d'être aussi complet que possible à la fois sur les aspects méthodologiques fondamentaux et sur les aspects techniques.

Je recommande particulièrement cet ouvrage à tous les statisticiens intéressés par l'articulation entre problèmes concrets et techniques mathématiques (en espérant qu'ils sont nombreux...), d'autant que le domaine d'application nous concerne évidemment tous au plus haut point.

Henri Caussinus