

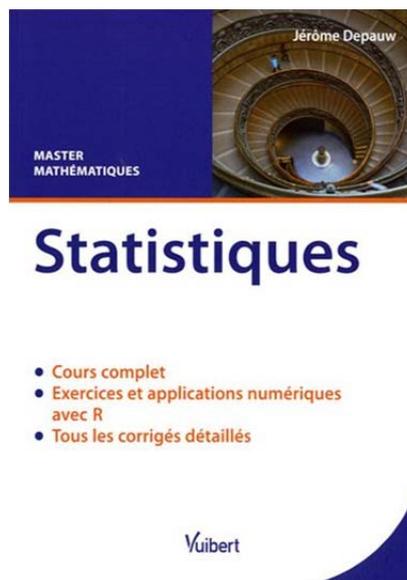
Note de lecture rédigée par Jeanne FINE¹

STATISTIQUES Cours et exercices corrigés

Jérôme DEPAUW

Livre (165 pages)

Édition : Vuibert, Master Mathématiques – 2012



L'auteur est maître de conférences à l'université François Rabelais de Tours et habilité à diriger des recherches, qu'il effectue en probabilités et théorie ergodique ; mais il s'intéresse également à l'enseignement et publie dans des revues d'enseignement de mathématiques.

L'avant-propos commence par ces mots : « Ce cours de statistiques s'adresse aux étudiants de master de mathématiques. Mais il pourra intéresser les enseignants de mathématiques qui, confrontés à l'enseignement des statistiques, cherchent une présentation à la fois rigoureuse et concise de ce domaine. Plus largement, il s'adresse à tout mathématicien désireux de découvrir les statistiques. »

Le défi que s'est proposé de relever l'auteur est le suivant : présenter de manière concise et rigoureuse quelques méthodes fondamentales de la statistique, mais aussi proposer des exercices corrigés et des applications avec le logiciel R. En effet, pour chaque chapitre, des exercices d'application directe du cours sont proposés et corrigés à l'aide de tables de quantiles de lois usuelles données en annexe, puis en utilisant le logiciel R dont une introduction est également donnée en annexe. Des exercices complémentaires permettent d'approfondir la méthode présentée dans le chapitre et d'introduire des variantes de la méthode.

¹ Professeure des Universités, Statisticienne, Toulouse, jeanne.fine@gmail.com

Note de lecture : « Statistiques » (J. Depauw, 2012)

Dans le premier chapitre, intitulé « *échantillon de la loi de Gauss* », l'auteur présente les *notions élémentaires de statistique inférentielle* : estimateurs des paramètres de cette loi et tests sur ces paramètres.

Dans le deuxième chapitre, l'auteur propose une introduction aux *méthodes bayésiennes* à partir de *l'estimation d'une proportion*. Il montre le lien entre démarche bayésienne et *théorie des jeux* et « l'intérêt de munir l'espace des paramètres d'une loi de probabilité, puis de choisir des méthodes statistiques « au hasard » ». Alors que de nombreux statisticiens ne découvrent la statistique bayésienne qu'après leurs études supérieures, l'auteur l'a placée parmi les méthodes fondamentales. C'est osé et très réussi : la présentation originale et ambitieuse qu'il propose est très intéressante.

Le troisième chapitre aborde *les tests asymptotiques* à partir du *test du chi-carré d'adéquation à une loi de probabilité finie* de Pearson et du *test du chi-carré d'indépendance* de Fisher. Ce chapitre est l'occasion de présenter la *loi gaussienne multidimensionnelle* et ses propriétés et de proposer, dans les exercices d'approfondissement, des tests asymptotiques en *analyse de variance* et en *régression linéaire*.

Le quatrième et dernier chapitre est une présentation de *l'analyse factorielle des correspondances* (AFC). Cette méthode descriptive, qui repose sur l'algèbre linéaire et la géométrie euclidienne, est présentée comme une forme optimale de représentation graphique d'un tableau de contingence en une dimension donnée. *L'analyse en composantes principales* (ACP) est proposée en exercice. C'est le seul regret que j'ai à la lecture de ce livre ; j'aurais choisi de proposer l'ACP en chapitre 4 et de présenter, dans les exercices d'approfondissement, l'AFC comme une double ACP dont on représente conjointement les individus. Le côté dissymétrique individus x variables de l'ACP permet de présenter l'interprétation géométrique des indices statistiques dans l'espace de représentation des variables : les variables sont représentées par des vecteurs, centrer une variable consiste à projeter le vecteur sur l'hyperplan orthogonal au vecteur « constant », la norme de la projection est l'écart-type de la variable et le cosinus de l'angle de deux vecteurs projetés est le coefficient de corrélation linéaire des deux variables qu'ils représentent. Toute l'interprétation géométrique des méthodes linéaires de la statistique multidimensionnelle repose sur cette lecture.

Les quatre chapitres sont courts, entre treize et vingt-huit pages, énoncés des exercices compris. Ils sont donc, vu l'ampleur des concepts abordés, extrêmement concis et denses mais, que ce soit en introduction d'un paragraphe ou d'un théorème, au cœur d'une démonstration ou dans les exercices, l'auteur propose des commentaires permettant de guider le lecteur.

L'auteur a aussi pour objectif de montrer le continuum entre les fondements mathématiques de la statistique et les méthodes classiques de statistique utilisées de façon intensive avec beaucoup d'efficacité. L'auteur y réussit grâce à la progression des types d'exercices qu'il propose : exercices d'application directe du cours, résolution des mêmes exercices en utilisant le logiciel R, variantes des méthodes présentées dans ses parties théorique et pratique.

Les contenus abordés dans ces quatre chapitres font l'objet de plusieurs unités d'enseignement dans un cursus de mathématiques appliquées ; l'idée de présenter le fondement mathématique de quatre méthodes bien choisies à des étudiants de M1 de mathématiques est très originale. Si elle était largement suivie, cette introduction à la statistique pourrait inciter davantage de mathématiciens à s'engager dans cette spécialité car,

Note de lecture : « Statistiques » (J. Depauw, 2012)

comme le dit Yves Derriennic dans la préface, « la diffusion de la culture statistique ne pourra se développer sans les professeurs de mathématiques de l'enseignement secondaire ou supérieur, même si évidemment, la spécialisation professionnelle devra ensuite s'appuyer sur des praticiens. »