

Note de lecture rédigée par Karen LEFFONDRE¹

LE LOGICIEL R
MAÎTRISER LE LANGAGE
EFFECTUER DES ANALYSES STATISTIQUES

Pierre LAFAYE DE MICHEAUX, Rémy DROUILHET et Benoît LIQUET

Livre (488 pages)

Édition : Springer-Verlag France – Collection Statistique et probabilités appliquées – 2011

Le logiciel R est un livre d'introduction et de perfectionnement de l'utilisation du logiciel R pour la mise en œuvre des méthodes statistiques de base. Il est destiné aux étudiants, enseignants, chercheurs, techniciens et ingénieurs voulant s'initier ou se perfectionner à R. Il peut aussi servir d'aide mémoire pratique et détaillé pour les utilisateurs réguliers. Conçu de manière très pédagogique, il est aussi un véritable support de cours sur le logiciel. Ce livre est en effet fondé sur les notes d'un cours enseigné pendant plusieurs années au Département de statistique et informatique décisionnelle de l'Institut universitaire de Grenoble 2. Il comprend ainsi des exercices à la fin de chaque chapitre et des sujets de travaux pratiques sur différents jeux de données épidémiologiques réelles disponibles sur le site du livre², qui servent également de fil conducteur tout au long de l'ouvrage.

Le livre est composé de deux parties : les bases du logiciel R et les mathématiques et statistiques élémentaires. La partie I comprend sept chapitres sur les concepts de base du logiciel : l'importation, l'exportation et la création de données ; leur manipulation ; les moyens d'accéder à de l'aide sur le logiciel ; les graphiques ; une initiation à la programmation en R pour le développement de nouvelles fonctions ; et la gestion des sessions sous R. La partie II comprend six chapitres : les fonctions mathématiques de base ; la statistique descriptive (incluant la mise en œuvre des fonctions graphiques de base) ; la simulation de données à partir de lois usuelles et moins usuelles ; les estimations par intervalles de confiance et les tests d'hypothèses ; la régression linéaire simple et multiple ; et l'analyse de la variance élémentaire.

Chaque notion mathématique ou statistique est généralement d'abord brièvement introduite avant l'illustration de sa mise en œuvre avec R. Les lecteurs qui maîtrisent déjà toutes ces notions pourront trouver ces rappels superflus. Les lecteurs peu à l'aise avec les notations mathématiques pourront les trouver trop théoriques. Cependant, de manière générale, ces rappels permettent d'introduire naturellement de nombreuses fonctions R, et peuvent être utiles et suffisants pour une majorité de lecteurs ayant déjà acquis les concepts de base en statistique. Ainsi, ces rappels peuvent éviter la consultation parallèle d'un ouvrage de statistique lors de la mise en œuvre des fonctions R décrites dans ce livre.

¹ Maître de Conférences, ISPED, Université Bordeaux Segalen, INSERM U897, karen.leffondre@isped.u-bordeaux2.fr

² <http://www.biostatisticien.eu/springeR>

Note de lecture : « Le logiciel R » (P. Lafaye de Micheaux et al., 2011)

Un des points forts du livre est son mode de présentation très visuel qui rend sa lecture facile et extrêmement agréable. De nombreux encadrés soulignent en effet des astuces, des points importants et des conseils très utiles qui sont indiqués par des icônes spécifiques. Le livre comprend de nombreux tableaux récapitulatifs extrêmement utiles tels que ceux qui récapitulent les paramètres pour la gestion des graphiques, les fonctions R pour chaque type de lois et tests d'hypothèses. En plus d'un index des commandes et symboles R et d'un index des packages R mentionnés dans le livre, chaque fin de chapitre présente un encadré qui récapitule les fonctions R vues au cours du chapitre.

En conclusion, *Le logiciel R* est un excellent livre d'auto-apprentissage ou de support de cours d'initiation, voire de perfectionnement, au logiciel R et à la mise en œuvre des méthodes statistiques de base avec ce logiciel. Il est également un aide mémoire utile, pratique et détaillé pour les utilisateurs de ce logiciel.