

## PARTICULARITES DE LA QUATRIEME EDITION

La troisième édition avait bénéficié de profondes modifications. L'élan est poursuivi dans cette nouvelle édition avec les trois objectifs récurrents (1) d'apporter des corrections, notamment celles qui étaient disponibles sur le site <https://c3bi.pasteur.fr/gael-millot-livres/>, (2) d'insérer des ajouts afin de produire un ouvrage complet aussi bien sur les tests statistiques que sur l'utilisation de R et (3) de mettre à jour l'ensemble du code décrit dans le livre pour qu'il soit compatible avec une version récente de R. Au total, 58% des pages de la troisième édition ont été retouchées et 152 pages ont été ajoutées, incluant 14 nouvelles figures.

Concernant la biostatistique, le problème des facteurs de confusion est introduit dans le chapitre 3 sur la démarche scientifique et les erreurs associées. La méthode de Benjamini & Hochberg (BH) est proposée dans les méthodes de correction de *p value*. L'annexe sur l'estimateur aborde maintenant l'intervalle de proportion, de pari et la notion de biais de l'estimateur. Une nouvelle annexe présente la possibilité de rendre non paramétrique un test paramétrique grâce au test de permutation. Enfin, une nouvelle annexe développe la problématique autour des faux positifs, faux négatifs, courbe ROC et notions associées.

Concernant R, l'ensemble du code et des explications relatives aux attributs de R a été mis à jour pour profiter de la version 3.3.3 du logiciel. Par ailleurs, six nouvelles annexes sont proposées. Elles concernent (1) les objets de mode "expression", "call" et "name" pour manipuler du code sans l'exécuter, (2) les symboles mathématiques et formats particuliers dans les graphiques, (3) la comparaison de chaînes de caractères, (4) l'identification de chaînes de caractères à l'aide d'expressions régulières, (5) la modification du format d'affichage des nombres et (6) les problèmes de virgule flottante et d'entier long. En plus de ces annexes, le nombre de fonctions décrites a été augmenté (24 pages au total pour l'annexe 26). Les 7 tests exacts ont été améliorés et bénéficient maintenant de meilleurs temps de calcul. Pour finir, le chapitre 1 sur la présentation de R a été profondément modifié et remanié. Il comporte 110 pages supplémentaires. Il inclut maintenant l'utilisation de l'interface RStudio. Il bénéficie d'exercices corrigés. Il aborde la notion d'indice hors limite. Dans le paragraphe de la manipulation des couleurs, les graphiques insérés sont en couleur. Le paragraphe de la manipulation des polices de caractères est augmenté de 4 pages. Enfin, une section de 30 pages décrit en profondeur la création de graphiques de type *ggplot2*.

Au final, j'espère que vous trouverez dans cette quatrième édition un contenu exhaustif sur la compréhension et l'utilisation des tests statistiques, ainsi que toutes les informations nécessaires à l'utilisation de R.