

CONCLUSION DE LA TROISIÈME PARTIE

Dans cette troisième partie de l'ouvrage, nous avons abordé l'analyse de régression linéaire. L'accent a d'abord été mis sur la régression comme méthode d'opérationnalisation d'un modèle de relation entre une variable dépendante et une ou plusieurs variables indépendantes. Car ce n'est que dans le contexte de cette démarche d'opérationnalisation qu'on peut donner un sens aux résultats de la régression. À cet égard, tous les modèles d'analyse multivariée sont comme l'analyse de régression : on ne peut en interpréter les résultats qu'à la lumière du modèle conceptuel sous-jacent.

En outre, la régression linéaire est un outil passe-partout de l'analyse multivariée. Bien d'autres méthodes, plus avancées ou plus spécialisées, peuvent être approchées à partir de celle-ci. Nous avons donc exposé les rudiments du fonctionnement de l'estimateur des moindres carrés et nous avons examiné le coefficient de détermination multiple en tant que mesure d'ajustement pour évaluer la performance d'un modèle.

Les méthodes d'induction appliquées à la régression multiple ont été abordées sous trois angles complémentaires. D'abord, le rôle de l'aléatoire dans la régression multiple a été examiné, afin de démontrer que les tests d'hypothèse appliqués à la régression s'appuient sur les mêmes fondements épistémologiques qui ont été établis dans la seconde partie de cet ouvrage. Mais dans l'induction statistique faite à partir de données d'échantillon, l'aléatoire est inhérent au lien entre l'échantillon et la population. Ici, l'aléatoire reflète plutôt le caractère approximatif du modèle, qui, dans sa partie déterministe, est analogue aux Idées platoniciennes : la réalité observable – les données – n'est qu'un reflet imparfait du modèle. L'aléatoire, c'est cette imperfection.

Les méthodes d'induction appliquées à la régression multiple ont ensuite été présentées d'un point de vue pragmatique. À l'aide d'exemples, nous avons vu en particulier à quoi peut servir le test t de Student et comment l'utiliser. Enfin, les hypothèses du modèle classique de la régression linéaire ont été énoncées et leurs implications plus concrètes ont été illustrées à travers l'analyse diagnostique des résidus.