

## INTRODUCTION À LA QUATRIÈME PARTIE : L'ANALYSE QUANTITATIVE DE DONNÉES QUALITATIVES

Nous savons déjà qu'en un certain sens, on peut mesurer des propriétés qualitatives, comme la nationalité d'une personne (chapitre 1-1). Pourtant, même si l'on admet qu'une propriété qualitative peut être mesurée, il n'en demeure pas moins que la mesure d'une telle propriété semble imparfaite en comparaison de la mesure de propriétés comme la superficie ou le revenu. En effet, lorsqu'on considère les quatre relations qui distinguent la définition axiomatique de la mesure ( $=$ ,  $\neq$ ,  $<$ ,  $>$ ), force est de constater que pour une propriété qualitative comme la nationalité, il y a deux de ces relations qui sont indécidables. C'est pourquoi on distingue plusieurs types de variables, selon l'échelle de mesure qui leur est associée : variables catégoriques, ordinales, d'intervalle et rationnelles.

La plupart des variables d'intervalle ou rationnelles, de même que certaines variables ordinales, peuvent être considérées comme des variables continues. Mais les variables catégoriques et bon nombre de variables ordinales sont des variables discrètes. Or nous avons vu, à propos du modèle linéaire général et de l'analyse de régression, que c'est une méthode d'analyse des données qui s'applique lorsque l'on s'appuie sur un modèle théorique formalisé par une relation entre une variable dépendante *continue* et une ou plusieurs variables indépendantes *continues* ou *discrètes*. Nous examinons dans cette partie du cours des méthodes qui s'appliquent à des variables catégoriques ou ordinales discrètes, comme celles qu'on emploie généralement pour mesurer des propriétés qualitatives.

L'*analyse des tableaux de contingence* permet d'examiner les relations entre plusieurs variables catégoriques. Dans l'analyse des tableaux de contingence, aucune des variables ne joue le rôle de variable dépendante. L'*analyse de variance* s'appuie sur un modèle théorique formalisé par une relation entre une variable dépendante *continue* et une ou plusieurs variables indépendantes, *toutes discrètes*. L'analyse de variance apparaît donc comme un cas particulier de l'analyse de régression (même si le format traditionnel de présentation des résultats est différent). Nous mettrons l'accent sur la façon d'adapter le modèle de la régression multiple à l'analyse de variance. Enfin, les *modèles à variable dépendante qualitative* font appel à des méthodes comme le logit ou le probit.