

Les différentes variables

Lorsque l'on mène des études scientifiques, différents types de données sont recueillis (ex : poids, taille, tension artérielle, présence ou non d'un diabète gestationnel, glycémie, ressenti des mères, vécu de l'accouchement, etc.).

Afin de pouvoir réaliser des analyses pertinentes et de pouvoir appliquer, si nécessaire, les bons tests statistiques, il est nécessaire de connaître la nature des variables que l'on étudie.

Les différents types de variables existants sont :

1. Les variables quantitatives

Elles expriment une quantité et sont exprimées par des chiffres (valeurs numériques). Il s'agit donc de variables numériques ou encore mesurables. Il existe 2 sortes de variables quantitatives :

- **Les variables discrètes**

Ce sont des variables qui ne peuvent prendre qu'un nombre fini de valeurs. On les appelle aussi parfois les variables « de comptage ».

Exemple : le nombre d'enfants (0, 1, 2, ...).

- **Les variables continues**

Ce sont des variables qui peuvent prendre un nombre infini de valeurs c'est-à-dire qu'elles peuvent prendre toutes les valeurs possibles dans un intervalle donné (valeurs décimales la plupart du temps). Il s'agit de la majorité des variables étudiées en recherche scientifique.

Exemple : l'âge (en années), la tension artérielle (en mmHg), la glycémie (en g/l), le poids de naissance (en grammes), etc.

ATTENTION ! En pratique, une variable continue peut être considérée comme discrète alors qu'elle est en fait continue. Il faut s'assurer de sa nature afin de bien comprendre la variable. La meilleure façon de s'en assurer est de regarder la nature et l'unité de mesure de la variable.

2. Les variables qualitatives

Elles expriment une qualité et sont exprimées par des « modalités » (= valeurs que peuvent prendre les variables qualitatives exprimées par des noms). On les appelle aussi les variables nominales.

Exemple : l'état civil (célibataire, marié, divorcé, veuf = 4 modalités), le groupe sanguin (A, B, AB, O = 4 modalités), la couleur des yeux (brun, bleu, vert = 3 modalités), etc.

Ce sont des variables sur lesquelles il est interdit de faire des calculs. Il sera juste possible de compter le nombre de réponses par modalités.

Il existe des cas particuliers des variables qualitatives :

- **Les variables ordinales**

On parle de variable ordinale lorsqu'il est possible d'établir une relation d'ordre au sein des modalités composant la variable qualitative.

Exemple : le grade obtenu aux examens (ajourné < réussite < satisfaction < distinction < grande distinction < plus grande distinction).

Dans ce cas précis, il est alors possible de numéroter les modalités de la variable et d'y effectuer des calculs comme sur les variables discrètes.

- **Les variables catégorisées**

Dans certains questionnaires, les chercheurs font parfois le choix de catégoriser des variables quantitatives pour des raisons de facilité de réponse : il s'agit alors de variables catégorisées. Le traitement statistique de ces variables n'est pas toujours facile.

Exemple : les tranches d'âge de la population des femmes en âge de procréer (< 18 ans, 18-25 ans, 25-35 ans et > 35 ans : 4 modalités de réponse).

3. Les variables binaires

Entre les variables qualitatives et les variables quantitatives, on retrouve les variables binaires qui sont très souvent utilisées en recherche scientifique. Une variable binaire peut être vue, d'une part, comme une variable qualitative à 2 modalités (que l'on peut numéroter (la plupart du temps : 0 = non et 1 = oui) ; ou, d'autre part, comme une variable discrète prenant 2 valeurs (0 ou 1) et sur laquelle on peut alors faire des calculs.

Exemple : le sexe (0 = homme, 1 = femme), le statut tabagique (0 = non-fumeur, 1 = fumeur), présence d'un diabète gestationnel (0 = non, 1 = oui), etc.

Source : Albert A. Biostatistique. Les Éditions de la Province de Liège et les Editions de l'Université de Liège, Liège, 2005.