VARIABILITÉ (C5)

(13 / 06 / 2020, © Monfort, Dicostat2005, 2005-2020)

« L'ennui naquit un jour de l'uniformité »

(Antoine Houdar de la Motte, *Fables nouvelles*, 1719)

On appelle **variabilité**, ou parfois **incertitude**, **versatilité**, **fluctuation** ou encore **instabilité**, la propriété (ou la qualification) de **non uniformité**, ou de **non homogénéité**, d'un **phénomène** observable (cf **observabilité**).

Souvent irréductible et propre au phénomène, elle constitue pour le **statisticien** une source d'**information** utile (cf **information statistique**, **dispersion**, **entropie**). On peut l'observer selon divers niveaux d'analyse.

(i) Variabilité interne

Au sein d'un **domaine de connaissance** donné (eg physique, biologie, écologie, psychologie ou sociologie), on distingue des phénomènes spécifiques du domaine. Chacun d'eux se décrit à l'aide de diverses **variables** : par définition, les « valeurs » de ces variables, aussi appelées **observations**, diffèrent généralement d'une **unité statistique** (**unité expérimentale**, **unité de sondage**) à une autre. Ceci explique l'intérêt de l'étude de leur répartition (cf **histogramme**, **distribution de probabilité**).

Autrement dit, le **matériel expérimental** (cf **dispositif expérimental**) ou l'**échantillon** obtenu par **sondage** sont constitués d'unités :

(a) distinctes, ie identifiables sans ambiguïté (eg numéros d'ordre ou identifiants différents) ;

(b) diverses, ou diversifiées : les observations (ou « mesures observées ») sur ces unités sont généralement différentes (existence d'**écarts**).

On parle, dans ce cas, de **variabilité « intrinsèque »**, ou de **variabilité « interne »**.

Une hypothèse scientifique fondamentale (cf aussi **loi scientifique**) est celle selon laquelle tout phénomène est « gouverné » par une **loi de probabilité** reliant l'ensemble des variables qui peuvent s'y relier (pertinence). Cette loi incorpore notamment, en tant que **caractéristiques légales** (d'ensemble, conditionnelles ou marginales), des paramètres de **dispersion** spécifiques du phénomène (cf aussi **paramètre d'échelle** ou différents types d'écarts, dont **écart-type**).

Cette variabilité fait donc partie intégrante ou constitutive du phénomène, et elle doit être admise (ou parfois subie). Par suite, on peut observer l'enchaînement suivant :

(a) le **schéma** explicatif (hypothétique) du phénomène, qui possède une **validité** plus ou moins importante. Ce schéma résume une « théorie du phénomène », ie une hypothèse relative à ce dernier : en effet, une théorie n'est qu'une hypothèse reconnue comme acceptable tant qu'une nouvelle ne parvient pas à la supplanter ;

(b) le **modèle** associé au schéma précédent comporte divers **paramètres** : ceux-ci sont estimés avec une certaine variabilité ;

(c) après validation (ie adéquation reconnue du modèle à l'observation du phénomène), l'utilisation du schéma pour la **décision** (actions diverses : prévision, prévention, etc) révèle le degré d'**efficacité** de ce schéma ou du modèle initial.

(ii) Variabilité externe au sein d'un même domaine

Dans un domaine de connaissance donné, les phénomènes composants peuvent posséder entre eux une variabilité « variable », notamment en fonction de l'**échelle** d'observation (micro-, méso- ou macro-scopique). Ainsi (physique), la variabilité d'un ensemble de corps célestes (macroscopie) n'a aucune raison a priori d'être identique à celle d'un ensemble de particules (microscopie) : un objet physique d'apparence « stable » (variabilité « nulle ») recouvre des molécules et atomes dotés de mouvements (variabilité non nulle).

(iii) Variabilité externe entre domaines

Enfin, entre domaines de connaissance différents, les phénomènes observables (de nature différente, par définition) peuvent aussi posséder une variabilité différente.

La variabilité est souvent supposée augmenter entre sciences dites « dures » (physique, biologie) et sciences dites « molles » (psychologie, sociologie) :

(a) ainsi, une prévision météorologique (physique) possède une probabilité plus grande d'être réalisée (avec un degré de confiance donné, notamment aux solstices) qu'une thérapie individuelle (psychologie) (eg récidives de psychopathes) ;

(b) mais cette supposition (ou « croyance ») n'est pas absolue : dans les deux domaines précédents, il peut être plus difficile de prévoir un cyclone (cf **catastrophe**) que d'empêcher un individu de passer à un acte contre lui-même (auto-destruction) (probabilités d'occurrences différentes).

(iv) Réduction de variabilité

Dans certains cas, la variabilité dépend, non seulement des unités observées, mais aussi de la **liste des variables** retenue dans la **représentation statistique** considérée (cf aussi **spécification**).

Un **jeu de variables** différent (mais de même pertinence), qu'elles soient observées sur le même ensemble d'unités ou sur un ensemble différent, peut conduire à une réduction de variabilité. Ceci est notamment le cas de **variables omises** lorsque celles-ci sont pertinentes.

Ainsi, dans un **modèle de régression**, le changement de la liste des **variables exogènes** (ajouts, suppressions) peut réduire la variabilité de la **variable endogène**.