CONDITIONNEMENT (D1, G, H, I)
(07 / 01 / 2023, © Monfort, Dicostat2005, 2005-2023)

Dans un **problème statistique**, on suppose souvent que certains **événements** (resp certaines **variables aléatoires**) influent sur d'autres. Les concepts de base mis en oeuvre sont alors ceux de **probabilité conditionnelle** ou de **loi conditionnelle** (cf **inférence conditionnelle**).

Les **décisions** que le **statisticien** adopte à partir des événements influencés (resp à partir des **va** influencées), en principe **observables**, dépendent donc des premiers, ie des variables d'influence.

(i) Le **principe de conditionnement** consiste à préférer conditionner explicitement de telles décisions pr aux évènements influents (resp pr aux variables influentes). Par leur influence, ces événements (resp variables) peuvent, en effet, apporter une **information supplémentaire utilisable**.

Ceci suppose qu'ils (elles) possèdent un minimum de **pertinence**, ie que le modèle aléatoire considéré est bien spécifié (absence d' « effets pervers », notamment) (cf **spécification**).

(ii) Les concepts centraux des **méthodes de conditionnement** sont ceux de **probabilité conditionnelle**, de **loi conditionnelle** (ou **loi de probabilité conditionnelle**) ou de **probabilité de transition**.

Diverses propriétés de ces lois sont ainsi utilisées, notamment celles de leurs **caractéristiques conditionnelles**. En particulier, les notions de **relation fonctionnelle** ou de **régression** sont souvent associées à celle d'**espérance conditionnelle**.

(iii) Le conditionnement intervient de façon quasi constante, eg dans l'étude de l'**exhaustivité** et de la **liberté**, du **modèle de régression** ou du **modèle d'interdépendance**, dans l'étude des **lois a posteriori** (cf **école bayésienne**) (cf aussi **théorème de BLACKWELL-RAO**).

Les décisions prises par le statisticien sont alors, explicitement ou non, qualifiées de conditionnelles, ie conditionnées par les évènements (resp variables) en question. Elles peuvent porter sur des **problèmes de décision** variés : **estimation** (conditionnelle), **tests d'hypothèses** (conditionnels), **prévision conditionnelle**, etc.

(iv) On considère un **phénomène** relevant d'un **domaine de connaissance** donné, et décrit à l'aide d'une liste de variables ζ observées sur des **unités statistiques**.

Son **étude scientifique** (donc statistique) se réfère, en général, plus ou moins implicitement à la **loi multivariée** $\mathscr{L}(\zeta)$ qui régit ζ.

Cependant, en général, la « vraie » liste des variables, ζ^*, qui devrait entièrement décrire le phénomène considéré, diffère de ζ. Autrement dit, ces deux listes peuvent être différentes, eg (symboliquement) : $\zeta \cap \zeta^* \neq \varnothing$ (seules certaines variables sont communes), ou encore $\zeta \subset \zeta^*$ (ie ζ^* « englobe » ζ) (certaines variables sont absentes). En effet :

(a) les variables non incluses (symboliquement $\zeta^* \setminus \zeta$) peuvent relever d'autres phénomènes, que ceux-ci soient propres au domaine de connaissance en question, ou qu'ils lui soient extérieurs (autres domaines) ;

(b) certaines variables de ζ^* sont **inobservables** pour diverses raisons : **censure** ou « effets de masque », contrainte budgétaire, etc ;

(c) certaines autres variables peuvent être tronquées (cf **troncature**), donc ne pas correspondre aux variables attendues.

Par suite, la loi $\mathscr{L}(\zeta)$ étudiée diffère de la « vraie » loi $\mathscr{L}(\zeta^*)$ qui devrait être étudiée. Ceci revient, implicitement, à admettre :

(a) soit que $\mathscr{L}(\zeta)$ est une loi marginale de $\mathscr{L}(\zeta^*)$. Cette **situation statistique** (ou attitude) s'appelle **marginalisation** (d'un phénomène) (cf **loi multivariée**, in fine) ;

(b) soit que $\mathscr{L}(\zeta)$ est une loi de $\mathscr{L}(\zeta^*)$ conditionnelle à la liste des variables non prises en compte (symboliquement, la liste $\zeta^* \setminus \zeta$). Cette attitude alternative s'appelle **conditionnement** (d'un phénomène).