

## ANALYSE FACTORIELLE (K5)

(20 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) L'**analyse factorielle** constitue une « branche » de l'**analyse des données** dont l'objet est la détermination, à partir d'un **ensemble d'observations**, de l'ensemble des **facteurs**, généralement **inobservables**, qui sont supposés structurer les observations en question. Ces facteurs sont appelés « **facteurs cachés** » ou « **facteurs sous-jacents** ».

De façon schématique, si  $\mathcal{L}(\xi)$  désigne la **loi multivariée** qui gouverne un **phénomène** donné, la **variable aléatoire**  $\xi$  n'est pas nécessairement observable (cf **variable cachée, variable latente, variable observable**). On suppose seulement observée une va  $\zeta$ , appelée « **variable manifeste** », considérée (à tort) comme décrivant le phénomène : l'analyse factorielle a alors pour objet l'**identification** de (tout ou partie de)  $\xi$  à partir de la manifeste  $\zeta$ .

C'est la va  $\xi$  qui joue ainsi le rôle de **facteur caché**.

Autrement dit, il existe une relation implicite de la forme :

$$(1) \quad \zeta = \phi(\xi),$$

dans laquelle  $\phi$  est inconnue et  $\xi$  inobservable (cf aussi **relation fonctionnelle**).

(ii) Deux conceptions d'analyses de ce type ont été distinguées :

(a) au sens large (parfois dite « acception française »), l'**analyse factorielle** recouvre l'**analyse générale des données**, ainsi que diverses méthodes de réduction qui en dérivent (**analyse des correspondances, analyse en composantes principales**, etc) ;

(b) dans un sens plus étroit (ou « acception anglo-saxonne »), elle concerne essentiellement l'**analyse en facteurs communs et spécifiques**.