

MODÈLE MIXTE (G2, J2, J3)

(23 / 12 / 2019, © Monfort, Dicostat2005, 2005-2019)

L'expression **modèle mixte** peut désigner l'une des notions suivantes.

(i) Un **modèle mixte** est une **représentation statistique** dont la **variable aléatoire** associée (eg un **échantillon**) est multivariée (eg multidimensionnelle), et dont certaines de « coordonnées » sont des **variables quantitatives**, les autres étant des **variables qualitatives**.

Lorsque le modèle est défini à partir d'une « liste » de variables, il combine des variables quantitatives et des variables qualitatives de cette liste (cf **modèle qualitatif**, **modèle quantitatif**) (cf classification ci-après).

variable	exogène		
	type	simple	multiple
endogène	simple	qualitative / numérique	qualitative / numérique
	multiple	qualitative / numérique	qualitative / numérique

Dans ce sens, le **modèle d'analyse de la variance** (une variable endogène numérique, des variables exogènes indicatrices), le **modèle d'analyse de la covariance** (une variable endogène numérique, certaines variables exogènes indicatrices, les autres numériques), le **modèle Logit** ou le **modèle Probit** (une variable endogène qualitative, des variables exogènes numériques) sont des modèles mixtes.

En pratique, les variables qualitatives d'un tel modèle se présentent sous forme codée (ou codifiée) : autrement dit, ce sont des variables numériques, généralement simples, provenant de **codages** effectués sur des variables qualitatives proprement dites (ie à valeurs dans des **ensembles non numériques**). Se pose alors parfois le problème de l'influence des codages utilisés sur les **procédures statistiques** mises en oeuvre (cf eg **courbe d'influence**). En outre, dans certains cas, la signification même du modèle n'est pas toujours dénuée d'ambiguïtés.

Diverses **procédures statistiques** permettent de traiter ces modèles : eg **méthode du maximum de vraisemblance**, théorie du **codage**, théorie de la **robustesse**.

(ii) Par ailleurs, on appelle **modèle mixte** tout modèle dans lequel une variable endogène est « expliquée » (ie décomposée) en fonction de plusieurs variables (cf notamment **modèle mixte d'analyse de la variance**), dont :

(a) les unes sont des **variables certaines** (cf **variable dégénérée**) ;

(b) les autres sont des **variables aléatoires**.