CONDITION DE RÉGULARITÉ (A16)
(28 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

Une **condition de régularité** est une hypothèse, ou un prérequis, mathématique, ou « technique », nécessaire à l'établissement ou à la vérification d'une propriété mathématique ou d'une propriété statistique (théorème, lemme, propriété, etc).

En effet, la plupart de ces propriétés exige un **contexte** suffisamment « régulier » pour pouvoir être démontrées (ie établies, prouvées, etc). Autrement dit, un corpus minimum d'hypothèses est nécessaire pour obtenir des résultats mathématiques ou statistiques (*« pas d'hypothèse, pas de conclusion »*).

(i) La nature d'une condition de régularité peut varier. On peut ainsi souhaiter que les « entités » ou objets mathématiques suivant(e)s possèdent certaines propriétés importantes :

(a) pour un **ensemble** : être doté d'une **structure** algébrique, d'une structure topologique, ou d'une structure mesurable, etc ;

(b) pour une **partie ensembliste** : vérifier une hypothèse de **convexité**, de compacité (cf **partie compacte**), etc ;

(c) cas d'une **suite** d'éléments (resp de **parties**) : posséder une propriété de convergence (cf aussi **mode de convergence** pour des suites de **va**), etc ;

(d) cas d'une **application** ou d'une **fonction** : être doté d'une propriété de **continuité**, de **différentiabilité**, de **convexité**, etc.

(ii) Ces deux types d'hypothèses, mathématique ou statistique, s'expriment toujours sous forme mathématique : en effet, et de ce point de vue, la **Statistique**, comme le **calcul des probabilités**, n'est qu'une branche des mathématiques.

On peut cependant distinguer entre :

(a) **hypothèses « proprement » mathématiques** ;

(b) **hypothèses de nature probabiliste** ;

(c) **hypothèses de nature statistique**.

Ainsi, dans un **modèle de régression linéaire** standard observé dans un **espace d'observation** $\mathbf{R}^N$ selon :

(1) $y = X b + u,$

et tq :

$(2)_a$ rg X = K (plein rang),

$(2)_b$ $E\, y / X = 0$ et $V\, y / U = \sigma^2 I_N$ (deux premiers moments conditionnels),

$(2)_c$ $u / X \sim \mathcal{N}(0, \sigma^2 I_N)$ (**normalité** conditionnelle),

on peut considérer que :

(a) l'hypothèse $(2)_a$ rg X = K (plein rang) est de nature purement mathématique et permet d'utiliser commodément des **matrices inverses** simples (au lieu des **matrices inverses généralisées**) ;

(b) l'hypothèse $(2)_b$ sur les moments et l'hypothèse $(2)_c$ sur la **loi** de la **perturbation** sont de nature probabiliste : il s'agit de **caractéristiques conditionnelles** associées à une **loi de probabilité** (cf **loi conditionnelle**) ;

(c) que l'hypothèse (1) (spécification linéaire) est plutôt de nature statistique, dans la mesure où elle se conforme suffisamment aux données observables.