TEST DE CHANGEMENT STRUCTUREL (I2, J, H)
(27 / 09 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) L'analyse d'un **phénomène**, suppose un repérage des **unités** concernées par (ou intervenant dans) ce phénomène. Les **observations** effectuées sont les « valeurs » prises par des **variables** que l'on peut associer à ces unités : ces variables peuvent être des **variables qualitatives** ou des **variables numériques**.

Pour le **statisticien**, ces dernières peuvent toujours être considérées comme résultant d'une **expérience aléatoire**, donc « régies » ou « gouvernées » par une **loi de probabilité** (cf **loi multivariée**). Il peut alors être nécessaire :

(a) de tester si un (ou plusieurs) changement(s) dans la **loi** qui a engendré ces observations s'est (se sont) produit(s) : autrement dit, on cherche à détecter l'apparition d'un **changement de régime**, voire de plusieurs, au cours de l'évolution du phénomène considéré (cf **régression à plusieurs régimes**, **catastrophe**, **test de CHOW**) ;

(b) de discriminer entre les observations, ie de détecter la loi qui a engendré chacune d'elles. Ces observations peuvent, parfois aussi, être considérées comme des « **aberrations** » les unes relativement aux autres. Par ailleurs, le changement structurel peut n'être qu'apparent : la liste des variables considérée peut ne pas être la « vraie » liste à prendre en compte, ce qui entraîne des influences changeantes, selon que l'on prend en compte ou non certaines variables (cf **spécification**, **test de spécification**).

(ii) Un **contexte statistique** de cette nature peut notamment survenir au cours du temps, eg dans l'étude du **modèle de régression** sur données temporelles (cf **série temporelle**). Dans ce cadre, une **variable endogène** η est liée (eg linéairement) à K **variables exogènes** ξ_1 ,…, ξ_K . Toutes ces variables sont numériques, et l'on dispose des observations y de η et X de $\xi = (\xi_1 ,\dots, \xi_K)$.

On est en présence d'un **changement de structure** (de la partie certaine) du modèle, à un instant $t^* \in \{1 ,\dots, T\}$, ssi ce modèle peut s'écrire sous la forme :

$$(1) \qquad y_n = \begin{cases} X_t\, b^1 + u_t , & \forall\, t \in \{1 ,\dots, t^*\}, \\ X_t\, b^2 + u_t , & \forall\, t \in \{t^* + 1 ,\dots, T\}, \end{cases}$$

où b^1 et b^2 sont des paramètres à valeurs (chacun) dans $\mathbf{R}^K$.

Si $t^* \in \{1 ,\dots, T - 1\}$, alors le changement de structure se traduit par l'hypothèse (à tester) $H : b^2 \neq b^1$.

Si $t^* = T$, aucun changement n'est intervenu sur la période $N_T{}^*$. Un test d'absence de changement structurel consiste, dans ce cas, à tester l'hypothèse $H_0 : t^* = T$ contre l'hypothèse $H_1 : t^* \in \{1 ,\dots, T - 1\}$.

Le cas de plusieurs changements structurels s'analyse de façon similaire.

Un changement de structure peut aussi se rencontrer dans l'**espace**, avec des données transversales (cf **coupe instantanée**) : il correspond généralement à l'existence d'une **hétérogénéité**, ou d'une « rupture » de liens selon les zones prises en compte dans cet espace.

(iii) Plus généralement, on peut considérer que un (ou plusieurs) changement(s) structurel(s) affecte(nt) la loi de probabilité gouvernant le phénomène, donc la structure et l'évolution de celui-ci (cf aussi **mélange de lois**, **catastrophe**, **catastrophe dans un processus**).

Par suite, la **relation fonctionnelle**, qui fait l'objet de l'**inférence statistique**, doit être modifiée ou adaptée.