INVARIANCE (G11)
(05 / 08 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Le terme d'**invariance** désigne une propriété qui demeure inchangée par une **opération** (ou transformation) donnée (cf **invariant**, **invariant maximal**, **principe d'invariance**, **propriété d'invariance**).

Cette notion intervient notamment en **théorie de la décision** : **modèle statistique invariant**, **règle de décision invariante**, **test invariant**, **estimateur invariant**, etc.

(ii) Ainsi, la notion est utilisée lorsque le **statisticien** procède à une **transformation de données** tq un changement d'**unité de mesure** (cf **échelle**, **échelle de mesure**). Il est alors préférable que la décision (estimation, test, prévision, classification) dépendant de la **procédure statistique** mise en oeuvre ne dépende pas de cette transformation.

A titre d'exemple, soit $z = (x_n , y_n)_{n=1,...,N}$ un jeu de N **observations** d'un **couple aléatoire** $\zeta = (\xi, \eta)$ dans lequel ξ désigne la température (en degrés Celsius) et η la longueur (en cm) d'une barre métallique. On suppose que le **phénomène** de dilatation étudié relève d'une loi affine de la forme suivante (**régression** « simple »), exprimée dans l'**espace des variables** (ξ, η) :

$$(1) \qquad \eta = a\,\xi + b + \varepsilon,$$

et que la **méthode des mco** est admissible pour estimer $(a, b) \in \mathbf{R}^2$. L'**estimateur des mco** de (a, b) s'explicite selon :

$$(2) \qquad \begin{aligned} a^{\wedge} &= \{N^{-1}\, \Sigma_{n=1}^{N} (x_n - \overline{x}_N)\, (y_n - \overline{y}_N)\} / \{N^{-1}\, \Sigma_{n=1}^{N} (x_n - \overline{x}_N)^2\}, \\ b^{\wedge} &= \overline{y}_N - a^{\wedge}\, \overline{x}_N , \end{aligned}$$

en notant vectoriellement $x = (x_1 , ..., x_N)'$, $y = (y_1 , ..., y_N)'$, et où $\overline{z}_N$ désigne la **moyenne arithmétique** empirique (cf **moyenne empirique**) d'un vecteur quelconque $z = (z_1 , ..., z_N)' \in \mathbf{R}^N$.

Autrement dit, $a^{\wedge} = C_N (\xi, \eta) / V_N\, \xi$ (rapport entre la covariance empirique de ξ et η et la **variance empirique** de ξ).

Avec des unités de mesure différentes (eg température en pouces et longueur en degrés Fahrenheit), le changement d'**unités de mesure** $(\xi, \eta) \mapsto (\xi^*, \eta^*)$ s'explicite selon :

$$(3) \qquad \begin{aligned} \xi^* &= (9 / 5) \,.\, \xi + 32 && \text{(relation affine),} \\ \eta^* &= 2{,}54 \,.\, \eta && \text{(relation linéaire).} \end{aligned}$$

Par suite, en notant $x^* = (x_1^* , ..., x_N^*)'$ et $y^* = (y_1^* , ..., y_N^*)'$ les observations correspondantes, on peut :

(a) soit remplacer dans (2) (x, y) par (x*, y*) et estimer (a, b) par les mco ;

(b) soit définir le **modèle avec variables transformées** :

(4) $\eta^* = a^* \xi^* + b^* + \varepsilon^*$

et estimer (a, b) à partir de (x*, y*) avec des formules analogues à (2).

Dans les deux cas, l'estimateur ($a^{\wedge *}$, $b^{\wedge *}$) obtenu n'est pas égal à l'original donné en (2).

(iii) La notion d'invariance se rencontre encore dans les **contextes statistiques** suivants :

(a) **méthode du maximum de vraisemblance** (cf **invariance du maximum de vraisemblance**) ;

(b) étude des **suites** de **variables aléatoires** ou en **théorie des processus**.