

## RÉSISTANCE (G, H)

(21 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) La notion de **résistance** (W.F. MOSTELLER - J.W. TUKEY, 1977) exprime la propriété d'une **procédure statistique** d'être « insensible » à un changement ou à une altération des **données** traitées par cette procédure (eg **échantillon** utilisé, **matériel expérimental**).

Cette notion est donc relative, et elle est, en principe, non probabiliste.

(ii) Le changement de données consiste :

(a) soit en une **modification** de l'**ensemble** des données, ou **transformation des données** : eg arrondissement ou troncature numériques (cf **partie entière**), **agrégation**, etc ;

(b) soit en une altération suffisamment importante d'une partie des données (cf **aberration**) ;

(c) soit en un changement de l'échantillon lui-même (remplacement par un autre).

(iii) La fonction  $\phi$  qui relie les données au « résultat » de la procédure est appelée **fonction de résistance**. Cette fonction est, en général, une **fonction numérique** : ainsi, si  $X$  est un « **jeu de données** » et  $R$  un résultat procédural d'un certain type (eg **estimateur**, **test** statistique, **classification**, **prévision**, etc), la fonction de résistance peut s'écrire  $R = \phi(X)$ .

A titre d'exemple, si  $X'$  et  $X''$  désignent deux jeux de données, et si  $S'$  et  $S''$  désignent les **statistiques** correspondantes (exprimant les résultats précédents), et ici supposées scalaires, la résistance peut se mesurer par l'**écart** absolu :

$$(1) \quad \Delta S = \Delta \phi(X) = \phi(X'') - \phi(X'),$$

ou par l'écart relatif :

$$(2) \quad \Delta S / S' = S'^{-1} \cdot \Delta S,$$

avec  $\Delta S = S'' - S'$ ,  $S' = \phi(X')$  et  $S'' = \phi(X'')$ .

Des extensions plus complexes peuvent se concevoir.

(iv) La notion de résistance se distingue de celle de **robustesse**.

Elle compare eg  $\phi(X'')$  et  $\phi(X')$ , tandis que celle de robustesse compare  $\phi''(X)$  et  $\phi'(X)$ , où  $\phi'$  et  $\phi''$  sont deux fonctions, a priori différentes, reliant les données aux résultats. Ces fonctions correspondent, en général, à des **règles de décision** (estimateurs, tests, etc) différentes. De plus, divers **tests** permettent de valider ou non l'importance des variations de résultat en fonction de celles des fonctions de décision.