LISSAGE (C2, C5, K2, N3
(24 / 09 / 2020, © Monfort, Dicostat2005, 2005-2020)

Un **lissage** est un procédé de « régularisation » portant sur des fonctions ou sur des concepts probabilistes ou statistiques (processus, séries temporelles).

(iii) **Premier sens**. On définit un **lissage** comme une **méthode de « régularisation »**, ou **méthode par régularisation**, portant sur des classes de fonctions importantes : **densité de probabilité**, spectre de **fréquences** (cf aussi **ajustement**, **approximation**, **filtrage**, **interpolation**).

L'exemple classique est le lissage du **spectre** d'un processus, qui consiste à transformer sa **densité spectrale** f selon (cf **méthode du noyau**) :

(1) $\quad I(\omega) = \int f(\omega)\, w(\omega)\, d\omega,$

où w est une **fonction de poids**, appelée **fenêtre spectrale**, choisie de façon adéquate pour simplifier l'étude de f (ie celle du processus considéré). Il existe diverses formes de fenêtres, eg : **fenêtre de R.B. BLACKMAN - J.W. TUKEY** (cf aussi **régressogramme de TUKEY**), **fenêtre de M.S. BARTLETT** (cf aussi **formule de BARTLETT**), **fenêtre de R.W. HAMMING - J.W. TUKEY** (cf aussi **lissage de HAMMING**), **fenêtre de J. von HANN** (cf **lissage de HANN**) ou **fenêtre de E. PARZEN**.

(ii) **Deuxième sens**. On appelle parfois (plus ou moins improprement) **lissage** toute méthode visant à l'**estimation de données** à un « niveau d'observation » plus (resp moins) « fin » que celui auquel elles sont disponibles.

Un exemple typique consiste à estimer en des points a_i (régulièrement) espacés (avec $i \in N_I^*$) les valeurs d'une variable ou d'une fonction f dont on ne connaît que les valeurs « moyennes » f_c sur des intervalles de classes $[x_{c-1}, x_c[$ recouvrant ces points (avec $c \in N_C^*$, et éventuellement $x_0 = -\infty$ et $x_C = +\infty$) (cf **histogramme**, **estimateur de la densité**, **estimateur spline de la densité**).

Une méthode peut consister à minimiser la somme des carrés des variations (eg absolues) :

$(2)_a \quad \Sigma_{i=0}^{I} (f_i - f_{i-1})^2$

sous les C contraintes de cohérence

$(2)_b \quad \Sigma_{a(i) \in [x(c)-1,\, x(c)[} \; w_{a(i)} \,.\, f_{a(i)} = f_c \,.\, \Sigma_{a(i) \in [x(c)-1,\, x(c)[} \; w_{a(i)} \,, \quad \forall c \in N_C^*,$

où les $w_{a(i)}$ sont d'éventuels **« poids »** associés aux données détaillées (en notant par commodité resp a(i) pour désigner a_i et x(c) pour désigner x_c).

(iii) **Troisième sens**. Un **lissage** est une **méthode de régularisation** d'un **processus stochastique** $X = (X_t)_{t \in T}$ ou d'une **série temporelle** $x = (x_t)_{t \in T}$ afin d'en dégager une **tendance**. Parmi les méthodes de ce type, on peut distinguer :

(a) les méthodes « empiriques » : **méthode des moyennes cycliques**, **méthode des moyennes groupées** (ou des moyennes échelonnées), **méthode des moyennes mobiles** ou **méthode des médianes mobiles**, etc. Ces méthodes, qui sont plutôt de nature descriptive, peuvent cependant constituer un moyen de visualisation rapide d'une tendance ;

(b) les méthodes « statistiques » ou méthodes « inférentielles » : méthodes de **régression** de la tendance $\mu = (\mu_t)_{t \in T}$ en fonction du temps $t \in T$ ou en fonction d'autres variables (exogènes), etc.

(iv) **Quatrième sens**. On appelle aussi **lissage** une **méthode d'évaluation** (ie une **méthode de prévision**) d'une série temporelle x au-delà de la **période** d'**observation**. Si x est en **temps** discret (ie si $T = N_T^*$), on cherche à prévoir le processus sous-jacent à x à un instant « futur » T + H (où $H \in \mathbf{N}^*$).

A titre d'exemple, on peut citer le **lissage exponentiel simple**, le **lissage exponentiel double** ou le **lissage exponentiel généralisé**.

Des lissages de ce type sont notamment fondés sur l'hypothèse que les observations anciennes exercent moins d'influence sur l'observation courante que les observations récentes : **hystérésis**, ou **estompage** (phénomène de mémoire qui s'atténue). Ceci est le cas de certaines séries temporelles (eg en économie).