CATASTROPHE DANS UN PROCESSUS (N10)

La **théorie des catastrophes** se traduit, en **Statistique**, par la **spécification** d'un **modèle**, ie d'une **représentation statistique** décrivant l'existence de changements possibles, rapides ou de forte amplitude, relatifs à un **phénomène observable** donné, qu'il s'agisse d'un phénomène physique, écologique, biologique, psychologique ou sociologique.

Ce changement est supposé attribuable à des variables, ou **facteurs** appelés **facteurs cachés** (car le plus souvent **inobservables**), ou parfois même (improprement) qualifiés de **« paramètres » cachés**.

Ainsi, un modèle de **régression à plusieurs régimes** traduit-il l'idée selon laquelle les observations disponibles relèvent de **structures** sous-jacentes différentes selon les périodes d'**observation** (cf aussi **aberration**, **classification**, **mélange de lois**). De même, une catastrophe peut s'analyser comme résultat d'un phénomène aléatoire (temporel) dont la **loi de probabilité** serait un **mélange de lois** dont la forme est multimodale (L. COBB) (cf **loi multimodale**).

(i) Le **système** aléatoire, associé au phénomène étudié, est souvent représenté par un **processus stochastique** de la forme $Z = \{(\Omega, \mathscr{T}, P), (\mathscr{Z}, \mathscr{D}), (Z_t)_{t \in T}\}$. On peut alors modéliser de plusieurs façons la notion de catastrophe.

On suppose que :

(a) $\mathscr{Z}$ est un **espace topologique** produit ($\mathscr{X} \times \mathscr{Y}$, $\mathscr{O}$), muni de sa **tribu borélienne** $\sigma(\mathscr{O})$. L'espace $\mathscr{Y}$ est dit **espace des états**, ou **espace des phases** ;

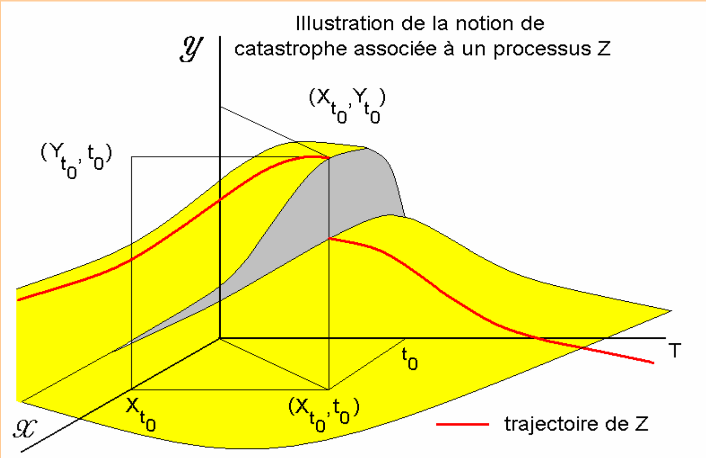
(b) l'ensemble des **temps** T est un espace topologique (T, $\mathscr{O}$(T)) muni de sa tribu borélienne $\mathscr{B}_T$;

(c) la famille $Z = (Z_t)_{t \in T}$ des **va** du processus s'explicite selon $Z_t = (X_t, Y_t) : \Omega \mapsto \mathscr{X} \times \mathscr{Y}$. L'espace $\mathscr{X}$ des valeurs de X_t est parfois dit **espace de(s) paramètre(s)**, ou **espace de(s) contrôle(s)** : il peut traduire l'existence de facteurs cachés. Cet espace est muni de sa topologie et de la tribu borélienne associée. En pratique, les variables X_t peuvent dépendre, à leur tour, d'un certain nombre de facteurs (ou de variables) inobservables (ou seulement partiellement observables) ;

(d) la loi P^Z du processus Z est définie, de façon usuelle, à partir de ses **projections** finies (cf conditions de compatibilité dans un **système projectif de probabilités**).

Il se produit une **catastrophe**, ou une **singularité**, dans le système Z à l'instant $t_0 \in T$ (ou à partir de l'état Y_{t0} correspondant) ssi il n'existe pas d'**homéomorphisme** (ou, selon une autre définition possible, d'**homotopie**) entre (X_t , Y_t) et (X_{t0} , Y_{t0}), où $X_t \in$

$\mathcal{V}_0$ (voisinage donné de $X_{t0} \in \mathcal{X}$) (en notant commodément t0 pour désigner t_0). Autrement dit, lorsque X_t varie, la variable d'état Y_t varie de façon non continue (cf graphique ci-dessous).



(ii) Dans la **théorie de R. THOM**, la définition d'une catastrophe fait appel à la notion de **fonction potentiel**, ou **fonction d'énergie** : v : $\mathcal{X}$ x $\mathcal{Y} \mapsto$ **R**. On définit alors une **catastrophe** ssi il n'existe pas d'homotopie entre les fonctions partielles v (X_t , .) et v (X_{t0} , .).

Plus précisément, on considère :

(a) un **espace d'états** ou **espace de réponses** (comportements, etc) $\mathcal{X}$, supposé être un **espace euclidien** de dimension Dim $\mathcal{X}$ = K ;

(b) un **espace de contrôles** $\mathcal{Y}$, supposé euclidien de dimension Dim $\mathcal{Y}$ = L ;

(c) une **fonction potentiel** v : $\mathcal{X}$ x $\mathcal{Y} \mapsto$ **R** décrivant un phénomène dynamique.

On suppose que le système décrivant le phénomène analysé cherche à tendre vers un minimum d'énergie, ie qu'il vise à minimiser (localement) l'énergie v (x,y), quel que soit le contrôle y $\in \mathcal{Y}$. Si l'on note $\mathcal{V}$ la **partie** de $\mathcal{X}$ x $\mathcal{Y}$ tq $Grad_x$ v (x,y) = 0 (cf **gradient**) et p : $\mathcal{V} \mapsto \mathcal{Y}$ la **projection** de $\mathcal{V}$ sur $\mathcal{Y}$, on dit qu'il existe une **singularité** en un point a $\in \mathcal{X}$ ssi (∂^2 v / ∂ x_k^2) v (a , y) = 0, $\forall$ y $\in \mathcal{Y}$, $\forall$ k $\in$ N^*_K . La projection p, restreinte aux singularités tq a, est appelée **catastrophe**, ou **ensemble des bifurcations**, de v sur $\mathcal{Y}$.

(iii) La **théorie des catastrophes** peut aussi se formaliser selon d'autres conceptions :

(a) on peut supposer que la **loi** gouvernant le processus est « différente », au voisinage de $t_0 \in T$, de celle qui gouverne le processus en dehors de ce voisinage. Ainsi, Z peut avoir pour loi P^Z un **mélange de lois** dépendant du temps, eg :

(1) $P^Z = \mathbf{1}_A(t) \,.\, P_1^z + \mathbf{1}_B(t) \,.\, P_2^z$,

où il existe $T_0 \subset T$ tq (1) soit vérifiée, où $A(t) = T_0$, $B(t) = T_0^c$, où Z est supposé iid selon la loi P^ζ d'une **va** ζ, et où P_1^z et P_2^z sont deux lois composant le mélange (leurs poids resp variant en fonction du temps).

(b) une autre version suppose que Z vérifie, en outre, une **équation d'évolution** représentée par une **équation différentielle stochastique**, en général non linéaire, de la forme :

$$(2) \quad \begin{cases} Z_\tau = z_0 \text{ (donné, éventuellement aléatoire)}, \\ dZ_t / dt = f(t, Z_t) + (dW_t / dt), \quad \text{si } \tau \neq t, \end{cases}$$

où $W = (W_t)_{t \in T}$ est un processus « perturbateur » (famille de **perturbations aléatoires** indexée par T) tq celui du **mouvement brownien** ou celui du **bruit blanc**. Les solutions possibles de (2) peuvent consister en des trajectoires comportant des « bifurcations ».

On peut étendre l'hypothèse d'une **mesure de LEBESGUE** dt précédente à celle d'une **mesure positive** ν définie sur $\mathcal{B}_T$.

(iv) Un problème fondamental est celui de la **prévision** (donc, le plus souvent, de la **prévention**) d'une catastrophe. Ce problème fait partie des problèmes généraux d'**estimation** et de prévision (cf **problème d'estimation**, **problème de prévision**), notamment dans le cadre de la **théorie des processus** (cf **prévision d'un processus**).

Des difficultés viennent, notamment :

(a) du caractère (partiellement) **inobservable** du processus $X = (X_t)_{t \in T}$;

(b) des **erreurs** de mesures qui peuvent porter sur les états Y_t ($t \in T$) : **erreurs** d'observation, instrument de mesure ou **système statistique** trop « grossier » pr à l'échelle d'observation, ou **dispositif expérimental** inadapté, etc.

Des tests, plus ou moins sensibles, peuvent être développés pour détecter des changements d'états catastrophiques. Certains sont issus de l'analyse des **séries temporelles** (eg **test des retournements**, **test des signes**, etc). D'autres sont liés à des **modèles** (de **régression** ou autres) spécifiques portant sur des séries temporelles : eg **modèle à structure variable** ou **régression à plusieurs régimes** et tests associés, **test de changement structurel**, ou de **robustesse** structurelle (eg **test de CHOW**), etc.