

PRÉVISION DANS UN MODÈLE STATISTIQUE (G10)

(19 / 06 / 2020, © Monfort, Dicostat2005, 2005-2020)

Malgré la diversité apparente des méthodes de prévision, une présentation synthétique en est cependant possible dans le cadre général de la **théorie de la décision** statistique (cf **décision statistique**). Cette approche permet notamment de situer la notion de **prévision** par à d'autres concepts statistiques usuels (en particulier, **l'estimation**).

Soit $(\Omega, \mathcal{F}, P)$ un **modèle statistique** de base, $(\mathcal{X}, \mathcal{B})$ un **espace d'observation** et $X : \Omega \mapsto \mathcal{X}$ une **va** (**statistique** ou **échantillon**) définissant le **modèle image** $(\mathcal{X}, \mathcal{B}, \mathcal{P}^X)$.

(i) On suppose que, au vu d'une « **observation** » $x = X(\omega) \in \mathcal{X}$, une **procédure** de décision statistique adaptée permette d'approcher (estimation statistique) la « vraie » **loi** $(P^X)^*$ générant les observations dans $\mathcal{X}$: soit $(P^X)^*_{\mathcal{N}}$ l'estimateur de cette loi.

On appelle alors :

(a) **prévision ponctuelle** de X toute **caractéristique** de **centralité** $c_{\mathcal{N}}^*(X)$ (espérance, mode, médiane), supposée unique, de la **loi estimée** $(P^X)^*_{\mathcal{N}}$ (cf aussi **paramètre de position**). Comme $(P^X)^*_{\mathcal{N}}$ dépend de X , $c_{\mathcal{N}}^*(X)$ est une va (ou une **statistique**) ;

(b) **prévision ensembliste** de X la donnée d'un couple $(\alpha, S_{\mathcal{N}}^*(X))$ constitué, à la fois :

(b)₁ d'un **seuil de confiance** $1 - \alpha$, avec $\alpha \in]0, 1[$;

(b)₂ d'une **région de confiance** (aléatoire) $S_{\mathcal{N}}^*(X) \in \mathcal{B}$, tq $P(S_{\mathcal{N}}^*(X)) \geq 1 - \alpha$.

En général, α étant donné, il existe une **famille** de régions tq $S_{\mathcal{N}}^*(X)$ satisfaisant ces conditions. Des **principes de réduction** permettent de déterminer une région optimale $S_{\mathcal{N}}^{**}(X)$, unique, tq :

$$(1) \quad \mu(S_{\mathcal{N}}^{**}(X)) = \inf_{S_{\mathcal{N}}^*(X) \in \mathcal{B}} \mu(S_{\mathcal{N}}^*(X)), \quad \gamma_{\xi} \in S_{\mathcal{N}}^{**}(X),$$

où $S_{\mathcal{N}}^*(X)$ désigne par commodité $S_{\mathcal{N}}^*(X)$, μ est une **mesure positive** (**mesure σ -finie**) dominant la famille $\mathcal{P}^X$ et γ_{ξ} est une caractéristique de centralité de $P^X = X(P) \in \mathcal{P}^X$ (**loi d'échantillonnage**). Autrement dit, la **région de confiance minimale** est de μ -mesure minimum et contient le **paramètre de position** γ_{ξ} .

(ii) On suppose que, au vu d'une « observation » $x = X(\omega) \in \mathcal{X}$, une procédure de décision statistique adaptée (**test** statique) a permis de déterminer une sous-famille $\mathcal{P}_0^X$ de lois possibles $\mathcal{P}^X$ du type précédent.

Dans ce cas, on peut répéter les définitions ci-dessus pour chaque loi $P^X \in \mathcal{P}_0^X$. On obtient, en fait, une **famille de prévisions ponctuelles** $c_N^*(X)$ de X ou une **famille de régions de confiance optimales** de même niveau α donné.

En pratique, la prévision ponctuelle de X est un élément de la famille des $c_N^*(X)$ tq le **risque** associé soit minimum (cf **fonction de risque**) ; la prévision ensembliste de X est l'intersection (supposée non vide) des régions optimales $S_N^*(X)$.

(iii) Les développements précédents concernent la notion la plus élémentaire de prévision : la prévision d'une va donnée X . Elle est souvent dite **prévision inconditionnelle**, ou **prévision absolue**, car on ne fait pas d'hypothèse de liaison entre X et d'autres variables (observables). Dans ce sens, prévoir une va consiste à caractériser sa **lp**, puisque celle-ci décrit (ou « résume ») entièrement ses propriétés, ou seulement une **caractéristique de centralité** de cette loi.

Plus généralement, on peut définir des notions de prévision qui intègrent l'« **information** » apportée par :

(a) d'autres **variables aléatoires** ou d'autres **événements aléatoires** à la même date, ie de façon synchrone, simultanée ou contemporaine (cf **prévision, prévision conditionnelle**) ;

(b) ou des variables ou événements à des dates différentes, ie de façon diachrone ou décalée (cf **modèle dynamique, prévision dans un modèle d'interdépendance, prévision d'un processus**).