PROBLÈME D'ESTIMATION ENSEMBLISTE (H2)

(18 / 10 / 2019, © Monfort, Dicostat2005, 2005-2019)

L'estimation d'une **caractéristique légale** ou d'un **paramètre** d'une **lp** par une **méthode ensembliste** est fondée sur la notion de **région de confiance** (cf **estimateur ensembliste**).

(i) Etant donné un **modèle statistique** $(\mathcal{X}, \mathcal{B}, (P_{\theta}^X)_{\theta \in \Theta})$ et une **famille** de parties de Θ , notée $\mathcal{S}$, l'**estimation ensembliste** est fondée sur l'idée selon laquelle un « élément » $S \in \mathcal{S}$ est à l'origine de la loi qui a engendré l'**observation** $x = X(\omega) \in \mathcal{X}$. Autrement dit, l'un des éléments $P_{\theta^*}^X$ de la sous-famille $(P_{\theta}^X)_{\theta \in \mathcal{S}}$ est la « véritable » loi de la **va** (ou **statistique**) $X : \Omega \mapsto \mathcal{X}$; par suite, la « valeur » θ^* de son **paramètre** est la « **vraie** » **valeur** de θ .

(ii) La région S dépend en général de X , soit $S(X)$, et il existe plusieurs telles régions possibles correspondant à un degré de confiance donné. On appelle **problème d'estimation ensembliste** le problème du choix d'une région (de confiance) S qui contienne la vraie valeur θ^* du paramètre $\theta \in \Theta$ avec une **probabilité** importante (ou « suffisante »), eg avec une probabilité au moins égale à $1 - \alpha$, où $\alpha \in]0, 1[$ est un **risque de non inclusion** du type précédent.

Ce choix s'effectue au vu de l'observation (eg **échantillon**) X .

En pratique, l'existence de classes trop vastes de régions de confiance optimales conduit à définir des **principes de réduction** pour ces ensembles : on définit ainsi les notions de **région de confiance la plus précise**, de **région de confiance minimale**, de **région de confiance symétrique** pr à θ^* , etc.