MODÈLE À PARAMÈTRE(S) ALÉATOIRE(S) (G2, G3, J, L5)
(07 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

Dans la **représentation statistique** d'un **phénomène**, on distingue classiquement entre la notion de **variable** et celle de **paramètre**. La première correspond souvent à la notion de **variable aléatoire**, la seconde à celle de **paramètre certain**. Dans certaines **situations statistiques**, il y a avantage à considérer un paramètre comme une variable aléatoire.

(i) Soit $(\Omega, \mathscr{T}, \mathscr{P})$ un **modèle statistique** (fondamental) écrit sous forme non paramétrée.

On dit que ce modèle est un **modèle à probabilité aléatoire** ssi il existe une **tribu** $\mathscr{A}$ de parties de $\mathscr{P}$ et une **mesure de probabilité** Q définie sur $\mathscr{A}$. Autrement dit, la « vraie » mesure de probabilité $P^* \in \mathscr{P}$ qui gouverne ce modèle est, elle-même, engendrée de façon aléatoire selon Q.

(ii) Si $\mathscr{P}$ est écrit sous forme paramétrée (ou même paramétrique) $\mathscr{P} = (P_\theta^X)_{\theta \in \Theta}$, on considère plutôt une tribu $\mathscr{B}_\Theta$ de parties de Θ et une mesure de probabilité Π définie sur $\mathscr{B}_\Theta$.

On dit alors que le modèle considéré est un **modèle à paramètre(s) aléatoire(s)** θ. On peut assimiler θ à l'**application identique** $id_\Theta : \Theta \mapsto \Theta$, considérée comme **va** pr à la tribu $\mathscr{B}_\Theta$. De même, on peut identifier Π et son image (la **loi de probabilité** de θ) par id_Θ .

(iii) En pratique, on dispose d'un **espace d'observation** $(\mathscr{X}, \mathscr{B})$ et d'une **statistique** (ou d'un **échantillon**) $X : \Omega \mapsto \mathscr{X}$. Le **modèle image** $(\mathscr{X}, \mathscr{B}, (P_\theta^X)_{\theta \in \Theta})$ ainsi obtenu est encore appelé **modèle à paramètre aléatoire** ssi le paramètre θ peut être considéré comme aléatoire.

Ceci est possible de deux façons :

(a) ou bien la loi Π est donnée à priori. Ceci est le cas de la **loi a priori** considérée en **théorie bayésienne**. Cette loi peut éventuellement posséder un **paramètre** propre : il existe alors un **ensemble** Λ et une **famille** $(P_\lambda)_{\lambda \in \Lambda}$ de lois a priori indexée par cet ensemble ;

(b) ou bien la loi Π est inconnue. On peut chercher à l'estimer, au moins partiellement : **caractéristique**, ou **paramètre d'intérêt**, de Π. On considère alors que $\Pi \in \mathbf{\Pi}$, famille de probabilités définies sur $\mathscr{B}_\Theta$, et l'on utilise essentiellement X pour baser l'**inférence statistique**. Si l'on considère P_θ^X comme **loi conditionnelle** de X sachant θ, supposée être une **loi absolument continue** pr à une **mesure positive** μ définie sur $\mathscr{B}$, la **loi inconditionnelle** de X est obtenue par

« marginalisation » de P_θ^X pr à θ, ie est définie par sa **densité** propre, qui est la suivante (cf **loi marginale**, **mélange de lois**, **probabilité de transition**, **théorème de BAYES**) :

(1) $q = \int_\Theta (dP_\theta^X / d\mu)\, d\Pi = \int_\Theta f(.\,,\theta)\, d\Pi(\theta)$,

où $f = dP_\theta^X / d\mu$ est la **densité de probabilité** (ou **vraisemblance**) de X associée à P_θ^X et q celle de la loi inconditionnelle.

Si Π dépend elle-même d'un paramètre propre $\lambda \in \Lambda$, on peut estimer P^X (qui dépend alors de λ) par la **méthode du maximum de vraisemblance**.

(iv) Des exemples de modèles statistiques avec paramètres aléatoires sont les suivants : modèle de **régression à coefficients aléatoires**, **modèle d'analyse de la variance** (ou **modèle d'analyse de la covariance**) à **effets** aléatoires.

Ainsi, on appelle parfois **modèle à paramètre aléatoire** un cas particulier du **modèle linéaire** (écrit dans un **espace d'observation $\mathbf{R}^N$**) :

(2) $y = X\,b + u$,

(en général, Z = X), où l'on suppose que la **perturbation** u elle-même est aléatoire, eg selon :

$u = Z\,c + v$,

(3) $E\,c = 0, \qquad V\,c = \Psi,$

$E\,v = 0, \qquad V\,v = \sigma_v^2\, I_N .$