FAMILLE DE LOIS (passim, C, J, N)
(16 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Dans un **modèle statistique**, ou un **problème de décision** statistique, un « élément » de base est la famille de ses **lois de probabilité**.

En effet, cette **famille** $\mathscr{P}^Z$ en constitue la donnée fondamentale puisque :

(a) sa connaissance implique, ou suppose, celle de l'ensemble $\mathcal{Z}$ des valeurs possibles de l'**échantillon aléatoire** ou de la **statistique** Z. Par suite, la considération d'un **espace mesurable** $(\mathcal{Z}, \mathcal{D})$ conduit à définir une **représentation statistique** $(\mathcal{Z}, \mathcal{D}, \mathscr{P}^Z)$;

(b) on suppose généralement que l'une des lois $P^Z \in \mathscr{P}^Z$ est la « vraie » loi P_v^Z qui engendre le **phénomène** analysé (cf aussi **vraie valeur d'un paramètre**) : autrement dit, on suppose que $P_v^Z \in \mathscr{P}^Z$. Alternativement, le fait de ne pas imposer de **spécification** « étroite » à $\mathscr{P}^Z$ permet toujours d'inclure, implicitement, la vraie loi P_v^Z dans cette famille ;

(c) la représentation $(\mathcal{Z}, \mathcal{D}, \mathscr{P}^Z)$ peut aussi être conçue comme une famille d'**espaces probabilisés** $(\mathcal{Z}, \mathcal{D}, P^Z)$, famille elle-même indexée par $\mathscr{P}^Z$. On peut, en effet, écrire aussi bien $(\mathcal{Z}, \mathcal{D}, \mathscr{P}^Z)$ que $(\mathcal{Z}, \mathcal{D}, P^Z)_{PZ \in \mathscr{P}Z}$ (en notant, par commodité, PZ pour représenter l' « **indice** » P^Z et $\mathscr{P}Z$ pour représenter la famille $\mathscr{P}^Z$). Dans le premier symbolisme, l'espace probabilisable $(\mathcal{Z}, \mathcal{D})$ est chargé par une famille $\mathscr{P}^Z$, tandis que, dans le second formalisme, l'espace probabilisé $(\mathcal{Z}, \mathcal{D}, P^Z)$ est indexé par $P^Z \in \mathscr{P}^Z$;

(d) si la « liste » des variables définies (ou « résumées ») par Z est exhaustive, s'il s'agit de **variables observables**, observées sans **erreurs**, etc, la connaissance de P_v^Z équivaut à une connaissance complète du phénomène : structure, fonctionnement et évolution (cf aussi **niveau, répartition, évolution**). En pratique, cependant, on se contente généralement de chercher seulement à connaître une **caractéristique légale** de P_v^Z.

(ii) Les considération précédentes se transposent, mutatis mutandis, au cas où il est possible d'expliciter la **paramétrisation** de la famille $\mathscr{P}^Z$ des lois P^Z à l'aide d'un **paramètre** $\theta \in \Theta$, ce qui conduit à une représentation sous forme paramétrée d'un modèle (cf aussi **paramètre d'intérêt**).

On formalise alors ce qui précède selon la **forme paramétrée** $(\mathcal{Z}, \mathcal{D}, (P_\theta^Z)_{\theta \in \Theta})$ ou selon $(\mathcal{Z}, \mathcal{D}, P_\theta^Z)_{\theta \in \Theta}$, avec les mêmes interprétations.

(iii) Lorsqu'on partitionne Z selon (X, Y) afin de distinguer des **variables exogènes** et des **variables endogènes**, les notations s'adaptent directement.

Diverses caractéristiques de P_v^Z sont alors, notamment, de nature fonctionnelle : cf **relation fonctionnelle**, **fonction de régression** ou **fonction d'interdépendance**.

(iv) Une famille de loi doit généralement posséder des propriétés « intéressantes » avant de procéder à l'**inférence statistique**. Ces propriétés peuvent être :

(a) de nature « intrinsèque » au **domaine de connaissance** ou au phénomène analysé ;

(b) ou encore de nature à permettre une inférence optimale.

Ainsi, le **statisticien** est conduit à définir les notions suivantes :

(a) **famille de lois dominée**. Cette notion est très importante, dans la mesure où elle définit celle de **modèle dominé**, donc conduit aux notions très utilisées de **vraisemblance** ou de **fonction de vraisemblance**. Ces concepts sont à l'origine de la **méthode du maximum de vraisemblance**, qui permet de définir **estimateur du mv**, ainsi que de diverses méthodes dérivées. On peut rattacher à ces notions celle de **famille de lois homogène** ;

(b) **identifiabilité** (cf **famille de lois identifiable**). Cette propriété permet de repérer, sans ambiguïté, la loi P_θ^Z associée à une valeur θ du paramètre ;

(c) **famille de lois complète**.

(v) Il existe des familles de **lois « privilégiées »** : l'une des plus importantes est la **famille exponentielle**, dont se déduisent de nombreuses lois particulières (cf eg **changement de variable aléatoire**, **transformation des données**).

(vi) Une famille assez générale est constituée de la **classe $S_K(\mu, \Sigma)$** définie comme suit.

Soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé** et $\xi : \Omega \mapsto \mathbf{R}^K$ un **vecteur aléatoire** réel de loi P^ξ tq $\xi \in L^2_{\mathbf{R}K}(\Omega, \mathcal{T}, P)$. Soit $\mu \in \mathbf{R}^K$ et $\Sigma \in M_n^+(\mathbf{R}) \cap D_n^0(\mathbf{R})$ une **matrice définie positive** qui est aussi une **matrice symétrique**, et α : $\mathbf{R}_+ \mapsto \mathbf{R}$ une fonction donnée.

On dit que ξ appartient à la classe S_K (μ, Σ) ssi sa **fonction caractéristique** est de la forme :

(1) $\varphi(t) = \alpha(t' S t)\, e^{i t' \mu}$.

On note alors $\xi \in S_K(\mu, \Sigma)$.

Les lois de la classe $S_K(\mu, \Sigma)$ vérifient les propriétés suivantes :

(a) si $\xi \sim \mathcal{N}_K(\mu, \Sigma)$ (**loi normale multidimensionnelle**), alors $x \in S_K(\mu, \Sigma)$;

(b) si f est la densité de P^ξ pr à la **mesure de LEBESGUE** λ_K , ie si $dP^\xi / d\lambda_K = f$, alors il existe une fonction $\beta : \mathbf{R}_+ \mapsto \mathbf{R}_+$ tq :

(2) $f(x) = c_K \cdot |\Sigma|^{-1/2} \cdot \beta\{(\xi - \mu)' \Sigma^{-1} (x - \mu)\}, \qquad \forall x \in \mathbf{R}^K,$

où c_K est une **constante** de **normalisation**.

Par ailleurs, si $\int_{\mathbf{R}+} r^{K-1} \beta(r^2)\, dr < +\infty$ (où **R+** désigne $\mathbf{R}_+$), alors f n'est pas dégénérée dans $\mathbf{R}^K$. Si $\int_{\mathbf{R}+} r^{p+K-1} \beta(r^2)\, dr < +\infty$, cette densité possède des moments jusqu'à l'ordre $p \in \mathbf{N}^*$.

Ainsi, lorsque $p \geq 1$, on a $E\,\xi = \mu$; si $p \geq 2$, on a $V\,\xi = \lambda\,\Sigma$ (où λ est un scalaire réel positif indépendant de (μ, Σ)).

En particulier, si $\mu = 0$ et $\Sigma = I_K$ (matrice unité), alors λ est la **variance** commune aux **lois marginales** à 1 dimension, ie aux lois $P^{\xi(k)} = \xi_k(P) = (pr_k\, \xi) P$ (où $\xi(k)$ désigne la va ξ_k) ;

(c) si $\xi \in S_K(\mu, \Sigma)$ et si $A \in M_{JK}(\mathbf{R})$, avec $J \leq K$ et $rg\, A = J$, alors $A \in S_J(A\,\mu, A\,\Sigma\,A')$;

(d) si $\xi \in S_K(\mu, \Sigma)$ (**matrice symétrique**) et si toutes ses marginales sont gaussiennes, ie :

(3) $(\xi_{k(1)}, ..., \xi_{k(L)})(P) = \mathcal{N}_L(\mu_{k(1)...k(L)}, \Sigma_{k(1)...k(L)}), \quad \forall \{k_1, ..., k_L\} \subset \{1, ..., K\},$

pour tout L tq $1 \leq L \leq K - 1$, alors ξ est gaussienne : $\xi \sim \mathcal{N}_K(\mu, \Sigma)$;

(e) si $\xi \in S_K(\mu, \Sigma)$ et si $\Sigma = \Lambda \in D_K(\mathbf{R})$ (**matrice diagonale**), alors les va $\{\xi_1, ..., \xi_K\}$ sont indépendantes ssi $\xi \sim \mathcal{N}_K(\mu, I_K)$;

(f) soit $\xi \in S_K\ (0,\ I_K)$ et $P^{\xi} = f\ .\ \lambda_K$ (ie P^{ξ} admet une **densité de probabilité** f pr à λ_K). On partitionne ξ en posant $\xi = (\xi_1\ ///\ \xi_2)$, avec $\xi_i : \Omega \mapsto \mathbf{R}^{K(i)}$, et en notant K(i) pour désigner K_i (i = 1, 2), avec $K_1 + K_2 = K$. Par suite :

(4) $\tau = ||\xi_1||^2 / ||\xi||^2 \sim \beta\ (K_1 / 2\ ,\ K_2 / 2)$ (**loi beta**),

et :

(5) $f = (K_2 / K_1)\ .\ ||\xi_1||^2 / ||\xi_2||^2 \sim \mathcal{F}\ (K_1\ ,\ K_2)$

(**loi de FISHER-SNEDECOR** à K_1 et K_2 **degré de liberté**).