CONTRAINTE A PRIORI (J)
(12 / 10 / 2019, © Monfort, Dicostat2005, 2005-2019)

Dans un **modèle statistique**, ou dans un **problème statistique**, on peut connaître au préalable (ie avant **inférence statistique**) une **information** portant sur le(s) **paramètre**(s) du modèle ou du problème.

Ceci ne nécessite pas nécessairement une complète **spécification** de la **loi de probabilité** du (des) paramètre(s) (comme en **théorie bayésienne**).

L'information disponible peut se traduire par des **contraintes** auxquelles doivent satisfaire ce(s) paramètre(s), et qui traduisent, d'une certaine manière, la représentation formelle d'une théorie, notamment celle élaborée par l'**homme de l'art**.

(i) Soit ($\mathcal{X}$, $\mathcal{B}$, $(P_\theta^X)_{\theta \in \Theta}$) un **modèle statistique** et $\mathcal{B}_\Theta$ une **tribu de parties** de Θ munie d'une **loi a priori** Π.

Une **contrainte a priori** peut être définie :

(a) soit somme une partie stricte $T \subset \Theta$ à laquelle θ doit appartenir. Ceci est le cas lorsqu'il existe une fonction g : $\Theta \mapsto \Gamma$ tq τ = g (θ), etc. On dit que g (ou Γ) est une contrainte a priori sur le paramètre (ou sur la loi a priori, ou sur le modèle, ou encore sur le problème) ;

(b) soit, plus généralement, par des lois P_θ^X devant satisfaire des propriétés données. Ainsi, lorsque $\Theta = \mathbf{R}^Q$, la contrainte sur P_θ^X peut prendre la forme d'une **régression**, tq g (X, θ) = u (où la **fonction de régression** g est spécifiée d'avance).

On peut interpréter ces situations comme des propriétés selon lesquelles la loi a priori Π ne « charge » qu'une partie de Θ.

(ii) Deux exemples classiques concernent le **modèle de régression** ou le **modèle d'interdépendance**. Ainsi, dans le cas d'un modèle d'interdépendance sous forme implicite f (ξ, η, θ) = ε, l'ensemble des contraintes a priori portant sur les **coefficients** θ peut s'écrire sous la forme g (θ) = 0, avec f : $\mathbf{R}^K \times \mathbf{R}^G \times \mathbf{R}^Q \mapsto \mathbf{R}^G$, g : $\mathbf{R}^Q \mapsto \mathbf{R}^S$, G étant le nombre de **variables endogènes**, K le nombre de **variables exogènes** et L le nombre de contraintes sur θ.

La contrainte est souvent une simple **relations d'exclusion** (ie une variable donnée ne figure pas dans une équation de la **forme structurelle**) : eg $\eta_g = 0$ (resp $\xi_k = 0$) signifie que, dans l'équation g, la variable endogène η_g (resp la variable exogène ξ_k) ne figure pas.