IDENTIFIABILITÉ (B1, C4, G2)
(05 / 09 / 2020, © Monfort, Dicostat2005, 2005-2020)

La propriété d'**identifiabilité** permet de « remonter » d'un **modèle transformé** vers un **modèle initial** de façon non ambigüe (cf **inférence statistique**). Elle est donc liée au problème de l'**identification** des **paramètres** d'un **modèle statistique** (cf **modèle identifiable**).

Un exemple de modèle pour lequel cette notion intervient de façon centrale est celui du **modèle d'interdépendance**.

(i) Soit $(\Omega, \mathcal{T}, \mathcal{P})$ un **modèle statistique** fondamental, $(\mathcal{X}, \mathcal{B})$ un **espace d'observation** portant sur des **variables exogènes** $X = [x_1, ..., x_K]$ (vecteurs colonnes), $(\mathcal{Y}, \mathcal{C})$ un espace d'observation portant sur des **variables endogènes** $Y = [y_1, ..., y_G]$ et $(\mathcal{U}, \mathcal{B}_{\mathcal{U}})$ un **espace latent**, ie un **espace probabilisable** associé à une **va inobservable** $U = [u_1, ..., u_G]$, appelée « **perturbation** » du modèle. On observe donc un triplet aléatoire $(X, Y, U) : \Omega \mapsto \mathcal{X} \times \mathcal{Y} \times \mathcal{U}$ (cf **couple aléatoire**) et l'on suppose donnée une **application mesurable** $f : \mathcal{X} \times \mathcal{Y} \mapsto \mathcal{U}$.

Un **modèle d'interdépendance** peut être considéré, de façon équivalente :

(a) comme le **modèle image** par (X, Y, U) du modèle fondamental précédent, modèle qui doit, en outre, satisfaire la contrainte suivante :

(1) $\quad f(X, Y) = U,$

appelée **forme structurelle** du modèle ;

(b) soit comme le modèle défini à partir de la **relation fonctionnelle** (1) considérée comme **caractéristique conditionnelle** de l'une quelconque des **lois de probabilité** $P^{(X, Y, U)}$, images des **mesures de probabilité** $P \in \mathcal{P}$ par le triplet (X, Y, U).

L'équation (1) correspond à une équation implicite dans laquelle f est souvent une **donnée** (eg fournie par l'**homme de l'art** au **statisticien**) et U une **va** dont la **valeur centrale** est supposée « négligeable » (eg une **espérance mathématique** nulle si $\mathcal{U}$ est un espace numérique et si U est intégrable).

Dans ce modèle, il existe a priori deux types de **paramètres d'intérêt** :

(a) ceux de la famille $\mathcal{P}$, si celle-ci est paramétrable (ie peut s'écrire sous la forme $(P_\theta)_{\theta \in \Theta}$) ;

(b) l'application f elle-même, ou encore les paramètres θ dont f peut dépendre (lorsque sa forme analytique est donnée et dépend de paramètres inconnus θ).

Cependant, f étant une caractéristique généralement associée à P (cf **relation fonctionnelle**), elle comporte aussi certains des paramètres θ pris en compte dans la formalisation de $(P_\theta)_{\theta \in \Theta}$.

S'il existe une application mesurable g : $\mathcal{X} \times \mathcal{U} \mapsto \mathcal{Y}$ déduite de f et tq :

(2) $Y = g(X, U)$,

(ie si f est partiellement inversible pr à Y), on dit que (2) est la **forme réduite** du modèle. Elle exprime donc les variables endogènes en fonction des variables exogènes et des variables inobservables.

Soit $P \in \mathcal{P}$ l'une des probabilités susceptibles de gouverner le **phénomène** aléatoire décrit par le modèle initial, $P^{(X,Y,U)}$ la **loi** de (X, Y, U), P^Y la **loi marginale** pr à Y déduite de $P^{(X,Y,U)}$ et $P^{g\,(X,U)}$ la **lp** de la va composée g o (X, U).

On dit que la forme (1) du modèle structurel est une **forme identifiable** ssi, connaissant la loi $P^{g(X,U)}$ qui a engendré la variable Y définie en (2), on peut lui associer de manière biunivoque la loi $P^{(X,Y,U)}$ qui a engendré le triplet (X, Y, U) lié par la relation (1), ie la **correspondance** :

(3) $\phi : P^{g\,(X,U)} \mapsto P^{(X,Y,U)}$

est une **application bijective** (l'injectivité de ϕ suffit) (cf **application injective**).

(ii) L'identifiabilité ainsi définie est une **indentifiabilité globale** (ie d'**ensemble**). Dans certains cas, il suffit qu'elle soit réalisée pour certaines lp seulement (**identifiabilité locale**) ou pour certaines valeurs de (X, Y, U).

(iii) Si le modèle précédent est un **modèle linéaire**, ie :

(4) $f(X, Y) = Y B + X C = U$,

la forme réduite s'écrit :

(5) $Y = A X + V$, avec $A = - C B^{-1}$ et $V = U B^{-1}$.

Dans ce cas, on dit que (B, C) est un **paramètre identifiable** ssi il est possible de passer, de façon unique, du paramètre (matriciel) A de la forme réduite au paramètre (B, C) de la forme structurelle.

La forme (4) étant donnée, on appelle **passage à la forme réduite**, ou simplement **réduction**, l'opération r : $M_G(\mathbf{R}) \times M_{GK}(\mathbf{R}) \mapsto M_{GK}(\mathbf{R})$ définie par :

(6) $(B, C) \mapsto r(B, C) = A = - C B^{-1}$.

Alors, dire que (B, C) est identifiable équivaut à dire que r est une **application injective**.

(iv) Dans le cas général, il n'y a pas, le plus souvent, identifiabilité. Il est nécessaire de faire des hypothèses sur f pour rendre identifiable le modèle. Ainsi, pour un modèle linéaire (4), diverses **conditions d'identification** ont été définies, qui portent sur (B, C), afin que r soit injective.