

LOI (C1, C4, H, I, J)

(12 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Terme « générique » désignant :

(a) soit, de façon large, la **loi de probabilité** d'une **variable aléatoire**. La loi considérée est donc, en général, une **loi multivariée** : sa connaissance équivaut à celle de toutes les propriétés du phénomène.

Ainsi, dans le cas de **variables numériques**, on donne souvent à une **lp** le nom de la **va** qui la définit : eg **loi de POISSON** pour une variable « poissonnienne », **loi de GAUSS-LAPLACE** pour une variable « gaussienne » ou « laplacienne », **loi exponentielle** pour une variable « exponentielle », etc ;

(b) soit, de façon plus étroite, la **loi** d'un **phénomène**, ie la **relation** (supposée ou admise) entre les variables décrivant ce phénomène, ou pouvant aider à l'expliquer. Cette notion correspond donc, en général, à celle de **relation fonctionnelle**. C'est cette notion que l'on qualifie souvent de **loi scientifique**.

Ainsi, dans le cas (entièrement numérique) d'un **modèle d'interdépendance** liant une liste de **variables endogènes** η à une liste de **variables exogènes** ξ , on appelle **loi (du phénomène)** l'équation (sous forme implicite et dans l'**espace des variables**) $g(\xi, \eta) = 0$, ou encore chacune des équations élémentaires qui composent cette équation, ou encore l'équation (sous forme explicite) $\eta = f(\xi)$ obtenue par inversion partielle (cf **application inverse**). D'un point de vue probabiliste, cette notion de loi est identique à celle de **caractéristique** « légale » de la lp de (ξ, η) . Cette caractéristique peut être plus ou moins complexe : eg fonction de **régression** de η pr à ξ , ou fonction d'**interdépendance** de (ξ, η) , etc.

(ii) On doit aussi relier les deux notions précédentes avec des **espaces d'observation**. Si l'on distingue encore entre endogènes et exogènes, un procédé général consiste :

(a) à définir d'abord, d'emblée, une **représentation statistique** $(\mathcal{Z}, \mathcal{D}, \mathcal{P}^{(\xi, \eta)})$, dans laquelle :

(1) $(\mathcal{Z}, \mathcal{D})$ est un espace d'observation produit, de la forme $(\mathcal{X} \times \mathcal{Y}, \mathcal{B} \otimes \mathcal{C})$;

(2) $\mathcal{P}^{(\xi, \eta)}$ est une famille constituée des **lp** possibles pour le **couple aléatoire** (ξ, η) . On suppose que l'une d'elle, notée $P^{(\xi, \eta)}$, est la « vraie » **loi jointe** qui engendre les **donnée** (X, Y) , ie les **observations** de (ξ, η) ;

(b) puis à définir la « loi » (au sens second) reliant η à ξ , laquelle est de la forme $g(\xi, \eta) = 0$.

Le procédé en question consiste alors à poser :

$$(1) \quad E g (\xi, \eta) = 0 ,$$

où l'espérance est calculée à l'aide de la **mesure de probabilité** P dont l'image par (ξ, η) est $P^{(\xi, \eta)}$. Autrement dit, le couple aléatoire (ξ, η) est supposé lié par la relation $g (\xi, \eta) = 0$, mais seulement « en moyenne » (ou **espérance**) ou, plus généralement, en « **valeur centrale** » : **médiane**, **mode**, etc (cf **régression**).

C'est cette démarche qui permet, en plongeant une relation déterministe dans le cadre d'une « liaison » stochastique entre va , de pratiquer diverses procédures d'**inférence statistique** : **estimation**, **tests**, **prévision**, etc.

Les deux concepts précédents peuvent ainsi être « réconciliés ». Etant donné une **loi multidimensionnelle** (premier sens) portant sur une liste de K variables exogènes et de G variables endogènes, la définition d'une « liaison » entre ces variables (fonction de régression, fonction d'interdépendance, etc) revient à définir des « lois » (au second sens) à partir de caractéristiques (conditionnelles) de la loi de probabilité donnée (premier sens). Ce procédé général est central dans la méthodologie statistique.

(iv) De nombreux exemples de « lois » (au second sens) peuvent être donnés. Ainsi :

(a) en physique, l'**équation d'onde** de base (à une dimension) :

$$(2) \quad (\partial^2 u / \partial x^2) = (1 / c^2) (\partial^2 u / \partial t^2)$$

(dans laquelle u est la masse, x un point de l'espace ou une distance, t le temps et c la vitesse du son), où celle de l'attraction élémentaire :

$$(3) \quad F_{AB} = F_{BA} = g \cdot d^{-2} \cdot m_A \cdot m_B$$

(dans laquelle M_A est la masse d'un corps A , F_{AB} la force exercée par B sur A , d la distance entre les centres de gravité de A et B , et g la constante gravitationnelle dont une estimation est $\tilde{g} = 6,674 \cdot 10^{-11}$ exprimée en $N \times m^2 \times kg^{-2}$), peuvent s'interpréter comme des lois particulières ;

(b) en écologie, les relations élémentaires duales :

$$(4) \quad \begin{aligned} A(t) &= a \cdot B(t-1) + b, \\ B(t) &= c \cdot A(t-1) + d, \end{aligned}$$

dans lesquelles A désigne le nombre de proies, qui diminue lorsque le nombre B des prédateurs augmente ($a < 0$), et B le nombre de prédateurs, qui augmente lorsque A augmente ($c > 0$), sont basiques. Ces relations définissent des lois d'évolution du type « prédateur-proie » relatives à certaines populations animales ;

(c) en économie, la **fonction de consommation** des ménages :

$$(5) \quad C_t = g (C_{t-1}, R_{t-1}, L_t, p_t, \dots)$$

(dans laquelle C est la consommation, R le revenu disponible brut, L la liquidité et p le prix implicite de la consommation) ou la **fonction d'investissement** des entreprises :

$$(6) \quad I_t = h (I_{t-1} - I_{t-2}, E_{t-1}, \dots)$$

(dans laquelle I est la FBCF et E le profit courant) représentent aussi des lois macroéconomiques particulières.