MATHÉMATIQUES ET SCIENCES HUMAINES

J. VORANGER

Pour une économétrie « bayésienne ». Processus économétriques linéaires récursifs de Bayes avec erreurs sur les variables exogènes, et décisions collectives de politique économique

Mathématiques et sciences humaines, tome 8 (1964), p. 9-14

<http://www.numdam.org/item?id=MSH_1964__8__9_0>

J. VORANGER *

POUR UNE ECONOMETRIE "BAYESIENNE"

PROCESSUS ECONOMETRIQUES LINEAIRES RECURSIFS DE BAYES AVEC ERREURS SUR LES VARIABLES EXOGENES, ET DECISIONS COLLECTIVES DE POLITIQUE ECONOMIQUE

I – L'APPROCHE "BAYESIENNE" EN STATISTIQUE

Analytiquement formulée par P.S. de LAPLACE à partir du mémoire original de Th. BAYES, l'approche "bayésienne" en Statistique a été progressivement abandonnée vers la fin du 19ème siècle en raison non pas de la fragilité de ces concepts mais à cause des dangers qu'elle présentait au niveau des applications lorsque les hypothèses de travail et leur probabilisation se trouvaient par trop fantaisistes. C'est ainsi que les travaux de CONDORCET et de POISSON intéressant le domaine judiciaire sont tombés prématurément dans le discrédit. La Statistique objective classique s'est développée sur des bases apparemment plus assurées depuis. A l'heure actuelle, cependant, un revirement est en train de s'opérer sous l'influence de divers ~~penseurs anglo-saxons~~, de L.J. SAVAGE en particulier. Cet auteur, en fondant la Statistique sur une théorie personne de la Décision, montre que toute prise de décision (dans une classe étendue de situations; voir par exemple (6)) demande d'attribuer des probabilités-poids aux hypothèses jugées personnellement possibles par l'agent responsable. Il réhabilite ainsi l'approche "bayésienne" en montrant par ailleurs la continuité qui existe dans la vie réelle entre probabilité-poids et probabilité-fréquence objective.

On peut saisir l'essentiel du débat en considérant l'exemple illustre du dé pipé. La Statistique classique ne formule pas d'hypothèses à priori sur la forme, la densité, etc... pourtant observables en partie. Un nombre élevé d'épreuves indépendantes révèle, seul, la probabilité objective de telle face particulière. Lorsqu'on dispose d'un nombre limité d'épreuves indépendantes (avec le même dé), une estimation de la "vraie" probabilité inconnue est formée par le rapport du nombre des arrivées de l'évènement favorable au nombre total des épreuves. Cette conception est exigeante car la seule observation attentive et objective du dé permettrait déjà d'avancer un ordre de grandeur vraisemblable pour la probabilité de l'évènement en question. Si l'on devine bien, c'est le rôle de l'expert, il se peut même qu'un nombre limité d'épreuves révèle presqu'exactement, joint à l'information à priori, la probabilité fréquentielle inconnue en question. Le fait d'être responsable d'une décision basée sur ces probabilités moins rigides garantit le sérieux avec lequel elles ont pu être attribuées aux évènements. En résumé, à condition d'être manipulées avec soin, ces probabilités au sens large, éclaireront l'action sans qu'il soit besoin d'attendre une information statistique abondante.

* Laboratoire de Statistique, Faculté de Droit et Sciences Economiques d'Aix-en-Provence.

II - H. RAIFFA ET SA THEORIE DES PROCESSUS LINEAIRES "BAYESIENS"

A la fin de son remarquable traité de Statistique appliquée, H. RAIFFA (1) présente une théorie très complète des processus "bayésiens" de régression linéaire multiple. Les valeurs possibles des paramètres de ces processus sont dotées au départ, comme dans l'exemple du dé pipé, d'une distribution de probabilité personnelle, ensuite révisée en fonction des statistiques disponibles. Le traité ouvre ainsi la voie à des applications à grand nombre de problèmes de décisions de la firme.

Il s'agit essentiellement de dégager des distributions de demande conditionnelles, compte tenu des valeurs exactement prévisibles des variables explicatives de la demande.

III - INTERET DE CETTE THEORIE POUR L'ESTIMATION DES STRUCTURES MACRO-ECONOMIQUES

Il semble que l'on ne puisse trouver que des avantages à recourir à cette théorie "bayésienne" des processus linéaires pour l'élaboration des structures ou modèles économétriques plus ou moins globaux. L'idée d'une telle transposition s'est fait jour dans l'esprit de quelques économètres. Dans la préface de son traité d'économétrie, E. MALINVAUD (2) dit: "Ainsi verrai-je grand intérêt à l'établissement d'une Statistique subjectiviste qui reposerait sur le principe de BAYES. Mais les recherches dans cette voie sont encore trop peu avancées pour permettre une application systématique de ces principes". Par ailleurs une théorie voisine a été indépendemment présentée par H. THEIL (3) dans un article récent du Journal de la Société Américaine de Statistique.

Grosso modo, on peut dire que cette conception "bayésienne" de l'économétrie que l'on préconise éviterait aux modélistes de demeurer uniquement tributaires de l'information statique disponible. Comme dans l'exemple du dé, on peut se prononcer à l'avance. Peut-être en économie plus qu'ailleurs, le jugement et l'expérience personnelle constituent une information très valable qu'il serait souvent déraisonnable d'ignorer sous prétexte qu'elle ne se présente pas sous forme de séries. L'information statistique n'est qu'une partie de l'information. Les calculateurs n'ont d'ailleurs pas attendu les développements de la Statistique formelle de BAYES pour suppléer tant bien que mal à l'insuffisance des sources. La Statistique de BAYES leur demande quelques indications supplémentaires toutefois afin d'éviter de donner d'illusoires certitudes sur la valeur d'un paramètre non accompagné de sa variance. On peut penser que cette conception trop large ouvre la porte à des abus. C'est exact. Mais si la connaissance des experts est considérée à priori comme insuffisamment objective, comment faut-il faire si d'une part l'information numérique fait partiellement ou totalement défaut et si d'autre part des décisions de politique économique doivent être prises? Le caractère personnel de l'estimation peut sembler un inconvénient. Plusieurs experts peuvent en venir à un accord. S'il existe des contradictions elles se trouveront pour le moins identifiées. A mesure que des statistiques appropriées deviennent disponibles, les estimations à priori jointes au "résumé exhaustif" de l'échantillon permet d'obtenir des estimations à postériori plus solides. Le modèle économétrique devient de plus en plus objectif. Voyons sur un exemple simple les étapes principales de la méthode.

IV – APPROCHE "BAYESIENNE" D'UN MODELE DE REGRESSION SIMPLE $\tilde{y}_t = \tilde{\beta}_1 x_{1t} + \tilde{\varepsilon}_t$

Si on ne dispose d'aucune information statistique, tous les paramètres doivent faire l'objet d'une estimation personnelle. La théorie est élaborée dans un cadre laplacien, aussi toute variable aléatoire simple d'un paramètre indépendant est entièrement définie par une espérance et une variance spécifiques de cette variable. Soit b'_1 et $\sigma^2_o(\tilde{\beta}_1)$ l'espérance et la variance à priori du paramètre β_1 inconnu (propension à consommer, par exemple), zéro et $\sigma^2(\tilde{\varepsilon}_t)$ l'espérance et la variance de la variable aléatoire résiduelle $\tilde{\varepsilon}_t$ représentant l'effet résultant des autres facteurs explicatifs de la variance conditionnelle du phénomène $\tilde{y}_t$ (consommation, par exemple). La valeur prochaine de la variable exogène $x_{1,t+1}$ (Revenu national, par exemple) est supposée prévisible sans erreur. Dans ces conditions, le modèle a priori s'écrit:

$$\tilde{y}_{t+1} = E_o\left[\tilde{\beta}_1\right] \cdot x_{1,t+1} + \tilde{\varepsilon}_{t+1} \qquad b'_1 = E_o\left[\tilde{\beta}_1\right], \; E\left[\tilde{\varepsilon}_{t+1}\right] = 0 \qquad \text{et } \sigma(\tilde{\varepsilon}_{t+1})$$

est certain pour simplifier. La distribution conditionnelle de $\tilde{y}_{t+1}$ est de Laplace d'espérance $b'_1 . x_{1,t+1}$ et de variance $\sigma^2_o(\tilde{\beta}_1) + \sigma^2(\tilde{\varepsilon}_{t+1})$. Mais le plus souvent des séries de réalisations antérieures, plus ou moins limitées, sont disponibles. Pour le processus actuel, on disposerait de couples (y_t, x_{1t}) pour $t = 1, 2, \ldots, n\ldots$ Le théorème de BAYES montre que la distribution de probabilité à postériori de $\tilde{\beta}_1$ à pour espérance une moyenne pondérée de l'espérance a priori et du "résumé exhaustif" de l'échantillon, ce dernier ayant comme valeur $\sum_t y_t x_{1t} / \sum x^2_{1t}$; pour autant que la valeur du paramètre soit restée constante au cours de la période de réalisation des épreuves. Dès que l'on dispose d'un échantillon même petit, l'hypothèse précédente concernant la certitude du paramètre $\sigma^2(\tilde{\varepsilon}_{t+1})$ est assez valable. Le modèle révisé ou a postériori s'écrit:

$$\tilde{y}_{t+1} = b''_1 \cdot x_{1,t+1} + \tilde{\varepsilon}_{t+1}$$

Toute chose égale, plus les observations sont nombreuses, plus la variance a postériori du paramètre inconnu est petite. A partir d'un certain moment, la variance est presque entièrement indépendante des donnés a priori, il en est de même de l'espérance.

La démarche générale est la même pour des modèles plus complexes, comprenant plusieurs variables exogènes exactement prévisibles. Considérons par exemple le cas $\tilde{y}_t = \tilde{\beta}_1 x_{1,t} + \tilde{\beta}_2 x_{2,t} + \tilde{\varepsilon}_t$, il se peut que l'on ne dispose pas de statistiques pour la variable x_{2t} mais seulement pour (y_t, x_{1t}). La théorie peut tenir compte de cette situation sans obliger de faire semblant d'ignorer l'effet approximativement connu du facteur x_{2t} sur y_t. L'ouvrage de H. RAIFFA renferme l'exposé détaillé pour un nombre quelconque de variables exogènes, $\sigma^2(\tilde{\varepsilon}_t)$ étant connue ou non

$$\left[\tilde{y}_t = \sum_j \tilde{\beta}_j . x_{tj} + \tilde{\varepsilon}_t\right]$$

V - L'INCERTITUDE DES VARIABLES EXOGENES PRISE EN COMPTE

On reproche parfois aux constructions économétriques d'être seulement capables de vérifier le passé, les variables exogènes n'étant généralement pas prévisibles. Là encore l'approche "bayésienne" permet d'introduire plus de nuance et de réalisme. Si le revenu national $x_{1,t+1}$ n'est pas exactement prévisible pour l'an prochain, on est conduit à introduire $\tilde{x}_{1,t+1}$ comme variable aléatoire dotée d'une espérance propre et d'une variance. Le modèle s'écrit:

$$\tilde{y}_{t+1} = \tilde{\beta}_1 . \tilde{x}_{1,t+1} + \tilde{\varepsilon}_{t+1}$$

Il est raisonnable de poser $\tilde{x}_{1,t+1} = 1 . x^*_{1,t+1} + \tilde{\eta}_{t+1}$ avec $E\left[\tilde{\eta}_{t+1}\right] = 0$ et $V(\tilde{\eta}_{t+1})=$ constante, par exemple.

L'agent de la décision prédit en toute certitude à une erreur près $\tilde{\eta}_{t+1}$ nulle en moyenne. C'est ainsi que le paramètre multiplicateur de $x_{1,t+1}$ est posé égal à l'unité en toute certitude. Dans ces conditions:

$$\tilde{y}_{t+1} = \tilde{\beta}_1 . \tilde{x}_{1,t+1} + \tilde{\varepsilon}_{t+1} = \tilde{\beta}_1 . (x^*_{1,t+1}) + \tilde{\beta}_1 . \tilde{\eta}_{t+1} + \tilde{\varepsilon}_{t+1}$$

En posant l'indépendance en probabilité de $\tilde{\beta}_1$ et $\tilde{\eta}_t$, on voit que

$$E\left[\tilde{y}_{t+1}\right] = E\left[\tilde{\beta}_1\right] . x_{1,t+1}$$

$$V\left[\tilde{y}_{t+1}\right] = V\left[\tilde{\beta}_1\right] . x^2_{1,t+1} + V\left[\tilde{\beta}_1\right] . V\left[\tilde{\eta}_{t+1}\right] + V\left[\tilde{\varepsilon}_{t+1}\right]$$

A noter que les espérances et les variances en β peuvent être des quantités a priori ou a postériori. La généralisation pour un nombre quelconque de variables exogènes est immédiate. L'incertitude concernant les variables exogènes n'est pas prise en considération dans le traité de H. RAIFFA.

VI - PROCESSUS RECURSIFS

Ces brèves remarques touchant l'incertitude qui affecte les variables exogènes conduisent à une généralisation possible pour le cas où le modèle, comprenant plusieurs équations, serait de type récursif dans le sens défini par H. WOLD (4):

a) si le développement des variables exogènes est connu jusqu'au temps t-1, le système ou modèle donne les valeurs des variables au temps t,

b) les variables au temps t s'obtiennent une par une de chaque équation par ailleurs interprétable comme une relation de dépendance causale unilatérale. Soit par exemple le système suivant:

$$\tilde{y}^1_{t+1} = \tilde{\alpha}^1_1 . y^1_t + \tilde{\beta}^1_1 . x^1_{1,t} + \tilde{\varepsilon}^1_{t+1}$$

$$\tilde{y}^2_{t,+1} = \tilde{\alpha}^2_1 . y^1_{t+1} + \tilde{\beta}^2_1 . x^2_{1,t-1} + \tilde{\varepsilon}^2_{t+1}$$

Le traitement statistique de la première équation ne pose pas de questions nouvelles. Par contre, du point de vue prévisionnel, la seconde équation, en faisant dépendre y^2_{t+1} de y^1_{t+1} présente une structure aléatoire plus compliquée mais précisément de même nature que celle définie en V où des erreurs affectaient les variables exogènes. Ainsi

$$\tilde{y}^2_{t+1} = \tilde{\alpha}^2_1 . \tilde{y}^1_{t+1} + \tilde{\beta}^2_1 . x^2_{1,t-1} + \tilde{\varepsilon}^2_{t+1} \qquad \text{où } E\left[\tilde{\varepsilon}^2_{t+1}\right] = 0$$

se traite sensiblement comme plus haut.

Le traitement "bayésien", en terme de moments, des modèles récursifs (avec ou sans erreur sur les variables) ne semble pas poser de difficultés spéciales. A noter cependant que la distribution de probabilité de $\tilde{y}_{t+1}$ n'est pas de Laplace. Il conviendrait d'étudier dans quelle mesure elle s'en écarte par rapport à une distribution de Laplace ayant les mêmes moments.

VII - LES TRAVAUX DE H. AUJAC

Ils montrent qu'en pratique les structures de W. LEONTIEF se rangent dans la classe des structures récursives. La triangulation constitue selon l'auteur une bonne approximation de la réalité. La valeur ajoutée d'un secteur s'exprime alors comme une fonction linéaire des valeurs ajoutées des autres secteurs d'amont.

L'estimation des matrices est ainsi un cas spécial de VI et ne présente pas non plus de difficultés particulières. L'approche "bayésienne" offre un cadre formel souple qui consolide la procédure d'estimation informelle cependant très réaliste suggérée par H. AUJAC. Ce ne serait pas un luxe que d'expliciter l'incertitude qui affecte les éléments des matrices de W. LEONTIEF. La formulation pour une matrice triangulaire à 3 lignes et colonnes serait:

$$\tilde{x}_{1,t} = \tilde{\beta}_2 . x_{2,t} + \tilde{\beta}_3 . x_{3,T} + \text{variables exogènes} + \tilde{\varepsilon}_t$$

$$\tilde{x}_{2,t} = \tilde{y}_3 . x_{3,t} + \dots + \tilde{\xi}_t$$

On a toute latitude pour introduire des variables exogènes spécifiques au secteur, que ces variables aient joué ou non dans le passé, que l'on dispose ou non à leur propos de statistiques. La démarche générale pour l'estimation consisterait à se donner (en passant par les commissions compétentes) une distribution de probabilité multiple a priori des paramètres β_2, β_3, γ_3, ... puis à dégager éventuellement à partir d'une suite même très limitée de matrices expérimentales des estimations, c'est-à-dire, une matrice de W. LEONTIEF a posteriori.

VIII - DECISIONS COLLECTIVES DE POLITIQUE ECONOMIQUE

Chaque parti politique pourrait définir son échelle d'utilité-probabilité au sens de von NEUMANN et MORGENSTERN. Toute politique économique revient à décider en faveur de la combinaison optimum, pour la collectivité, de certaines variables exogènes comme les impôts, les droits, tarifs, taxes, etc... A chaque combinaison de valeurs de ces variables exogènes de décision correspond pour chaque parti politique et pour chaque combinaison des variables endogènes résultantes du modèle (combinaison affectée de sa probabilité), une utilité-probabilité.

A chaque parti est associé un jeu contre la Chance dont les éléments sont indiqués, sous une autre forme, dans le tableau ci-joint:

PARTI A

		Actions: combinaisons diverses des modalités des variables de décision (impôt, etc...)
	Prob.	"
Evènements:		"
Combinaisons diverses des	"	"
modalités des variables	"	"
endogènes du modèle.	"	"
(niveaux possibles des	"	"
invest., revenus, etc.)	P(i,r,..)	u(i,r,...)
Utilités espérées	1	E(u)

S'il y a désaccord aucun optimum ne peut être défini, s'il y a accord au contraire sur les probabilités des évènements, l'introduction de "poids sociaux" (6) permet de dégager une politique optimum pour la collectivité, celle pour laquelle la moyenne pondérée des utilités espérées de chaque groupe est la plus grande (ou le "regret" le moindre). Il conviendrait de spécifier une distribution de probabilité multiple pour les évènements et une fonction continue pour les utilités. Il serait intéressant d'évaluer le prix d'une politique sans "regret".

BIBLIOGRAPHIE

(1) - H. RAIFFA - R. SCHLAIFER Applied Statistical Decision Theory, Harvard, 1961.

(2) - E. MALINVAUD Méthodes Statistiques de l'Econométrie, DUNOD, 1964.

(3) - H. THEIL On the Use of Incomplete Information in Regression analysis, J.A.S.A., June 1963.

(4) - H. WOLD Demand Analysis, Wiley, 1953.

(5) - H. AUJAC La Hiérarchie des Industries dans un Tableau des Echanges Industriels. R. Economique N° 2 - 1960.

(6) - G. MORLAT Des Poids et des Choix, Mathématiques et Sciences Humaines, 1963.

- - - - - - - - - - - - - - -