

UNIVERSITE LIBRE DE KIGALI U.L.K.

B.P. 2280 KIGALI

Tél : 511526, 511527, 511531, 511533

Fax : (250) 511526

E-mail : ulk@rwandatel1.rwanda1.com



REVUE SCIENTIFIQUE DE L'UNIVERSITE

EDITIONS DE L'UNIVERSITE LIBRE DE KIGALI

TABLE DES MATIERES

EDITORIAL	5
RUSUHUZWA K. Thomas : La modélisation VAR avec application à la politique monétaire au Rwanda.	7
BUTARE Théopiste : Approches de quantification et des modélisations en économie ; l'apport de l'économétrie.....	25
KAAYA Siraje : English as a tool for correspondences.....	113
NGABOYISONGA Roch : Efficacité interne du système éducatif rwandais.....	141
KAGABIKA Boaz : Le rôle de l'éducation environnementale dans le développement durable du Rwanda.....	163
NDABATEZE Abraham : Le dual topologique de l'espace C des suites réelles convergentes.....	179
BIZIMUNGU Christophe : Le régime juridique de l'infraction d'adultère face aux impératifs du moment	201

L'Université Libre de Kigali est une institution qui est en train de prendre toute sa taille dans l'espace universitaire rwandais.

La volonté et la détermination qu'elle manifeste de devenir la première des universités rwandaises commencent à se concrétiser.

Première par sa démographie, elle ambitionne, à juste titre, d'être la première par la qualité de ses enseignements et de la recherche dont elle a fait sa priorité. Le plan quinquennal de formation de 3^e cycle des enseignants de haute qualité scientifique et morale, sa politique du bilinguisme déjà en vigueur, constituent des preuves de refus de l'auto-satisfaction et de l'immobilisme.

Sur le plan scientifique, ce dynamisme se remarque par la série ininterrompue de mémoires présentés et soutenus tout au long de cette année, les conférences tenues par diverses sommités tant du monde académique que politique, et enfin par ce dernier numéro de la revue scientifique de l'ULK qui n'est plus à présenter.

Cette revue dont la mission première est le service à la collectivité aborde une diversité d'articles dont les thèmes varient de l'économie monétaire, le droit, les mathématiques, à l'éducation et l'enseignement de l'anglais. Ce numéro touche du doigt les problèmes de la société rwandaise et propose des réponses susceptibles d'aider à relever le défi du développement au Rwanda.

Prof. Dr RWIGAMBA BALINDA



THE UNIVERSITY OF CHICAGO
LIBRARY

1950

1950

1950

1950

1950

1950

1950

1950

1950

1950

1950

1950

1950

1950

1950

**LA MODELISATION VAR AVEC APPLICATION
A LA POLITIQUE MONETAIRE
AU RWANDA**

Par Rusuhuzwa Kigabo Thomas

Enseignant à l'Université Libre de Kigali(ULK) et chercheur au GATE (Groupe d'Analyse et des Théories Economiques), un groupe de recherche du CNRS-Université Lumière Lyon2.

Introduction

Cet article vise deux principaux objectifs :

D'abord faire une présentation brève mais concise de la formulation et de l'intérêt de la modélisation VAR (Vector AutoRegressive) et en suite appliquer ce type de modélisation pour analyser les mécanismes de transmission de la politique monétaire au Rwanda. Il s'agit en d'autres termes d'identifier les canaux par les quels la politique monétaire se transmet sur les variables de l'économie réelle (la production et le prix).

Le Vecteur auto régressif est constitué d'un ensemble des variables macro économiques bien choisies pour décrire un phénomène économique, et la modélisation VAR permet d'exprimer chacune d'elles non seulement en fonction de son passé comme dans le cas d'un AR traditionnel mais également par le passé des autres variables du système. Dans le cadre de notre étude, les variables utilisées sont le PIB au prix constant, le taux d'inflation ainsi que le taux d'intérêt créditeur ainsi qu'un agrégat monétaire. Il s'agira de vérifier si les canaux de transmission de la politique monétaire au Rwanda à la sphère réelle repose sur les variations du taux d'intérêt.

Modélisation VAR

Les processus VAR, qui sont une généralisation des processus Auto Régressifs (AR), ont été introduits par SIMS(1980) après des critiques adressées aux modèles macro économétriques d'inspiration Keynésienne, avant de faire objet de plusieurs développements économétriques.

Pour Sims, ces derniers modèles sont, d'une part sujets à des restrictions a priori sur les paramètres qui sont très fortes comparativement à ce que prédit la théorie économique, et d'autre part, il y a absence des tests puissants sur la structure causale et les anticipations ne sont pas traitées de manière adéquate. Du point de vue empirique,, ces modèles ont conduit à des erreurs de prévision très importantes suite aux événements survenus au cours des années 70.

La modélisation multivariée proposée par Sims n'a pas d'autres restrictions a priori que le choix des variables et du nombre des retards, ceci compte tenu de l'hypothèse de base de la modélisation VAR selon laquelle la connaissance du comportement dynamique d'un ensemble de n vecteurs bien choisis et linéairement dépendant du passé permet de bien approcher l'évolution d'une économie. L'un des intérêts de la modélisation Var structurel, comparé aux modèles macro écono-métriques traditionnels est de cibler de manière précise la question économique que l'on veut traiter. Ce qui permet de trancher sur la question de la distinction entre les variables endogènes et exogènes.

Représentation VAR

Considérons n processus $\{X_{it}, t \in Z\}, i=1, \dots, n$ et exprimons chacune de ces n variables en fonction de ses valeurs passées mais aussi des valeurs passées et présentes des autres variables :

$$BX_t = A_0 + \sum_{i=1}^p A_i X_{t-i} \varepsilon_t \quad (1)$$

avec ε_t un bruit blanc

La formulation (1) est appelée représentation structurelle du VAR(P).

Le nombre élevé des paramètres à estimer, qui croît rapidement avec les nombres p de retards et n des variables constitue la difficulté majeure qui rend difficile l'estimation de ce modèle. Ceci fait qu'on travaille généralement à partir de la forme réduite du modèle Var obtenue en multipliant (1) par B^{-1} (inverse de la matrice B) :

$$X_t = B^{-1} A_0 + \sum_{i=1}^p B^{-1} A_i X_{t-i} + B^{-1} \varepsilon_t \quad (2)$$

ou $\phi_0 = B^{-1} A_0; \phi_i = B^{-1} A_i; v_t = B^{-1} \varepsilon_t$

Si toutes les variables sont stationnaires, v_t s'interprète comme l'innovation du processus X_t .

Dans cette dernière formulation, on observe que le niveau de X_{it} ne dépend que de ses valeurs passées, des valeurs passées de X_{jt} ; $\forall i \neq j$ ainsi que des innovations v_{it} .

Notons que les innovations v_{it} peuvent être corrélées, même si celles de la forme structurelle ne le sont pas. Le système (3) peut être estimé par les MCO. La convergence ainsi que la propriété asymptotique des estimateurs des MCO sont assurés du fait de la stationnarité de la série X_{it} . Ce qui permet de mener les tests sur les paramètres du modèle. Si Les variables ne sont pas stationnaires, ce qui est souvent le cas en économie, la distribution des estimateurs par les MCO du modèle (2) n'est plus asymptotique normale. Toutefois, si ces variables non stationnaires sont cointégrées, ce problème est résolu par utilisation des modèles à correction d'erreur.

On considère de manière générale qu'un processus Vectoriel $\{X_t, t \in Z\}$, de dimension $(n,1)$ admet une représentation VAR d'ordre p , et on note VAR(p) si :

$$X_t = \phi_0 + \sum \phi_i X_{t-i} + v_t \quad (4)$$

$$\phi_0 \in R^n$$

$$\Leftrightarrow X_t - \sum_{i=1}^p \phi_i X_{t-i} = \phi_0 + v_t$$

$$\Leftrightarrow \Phi(L)X_t = \phi_0 + v_t \quad (5)$$

$$\text{Avec } \Phi(L) = I - \sum_{i=1}^p \phi_i L^i$$

Les matrices $\Phi_i, \forall i \in [0, p]$ de dimension (n,n) satisfont

$$\phi_0 = I_n; \Phi_p \neq 0_n$$

Quelques propriétés importantes

1. Il est possible de transformer tout processus vectoriel $\{X_t, t \in Z\}$ qui satisfait une représentation VAR(p) en un processus $\{Y_t, t \in Z\}$ qui satisfait une représentation VAR(1);

2. Tout processus $\{X_t, t \in Z\}$, de dimension $(n,1)$, stationnaire, satisfaisant une représentation VAR(p):

$X_t = \sum \phi_i X_{t-i} + c + v_t, \forall t \in Z$, admet une représentation moyenne mobile convergente VMA(∞) définie par :

$$X_t = \mu + \sum_{i=0}^{\infty} \psi_i v_{t-i} = \mu + \Psi(L)v_t$$

Le polynôme matriciel $\Psi(L) = \sum_{i=0}^{\infty} \psi_i L^i$ satisfait les propriétés suivantes :

$$\psi_0 = I_n; \psi_s = \sum_{i=1}^p \phi^i \psi_{s-i}; \forall s \geq 1; \psi_s = 0, \forall s < 0.$$

Analyse des réponses impulsionnelles

Les modèles VAR sont souvent analysés au travers de leur dynamique. Cette analyse est faite en simulant des chocs aléatoires et en analysant la décomposition de la variance de ces chocs. Ceci permet par exemple de réaliser l'analyse de réponse impulsionnelle qui consiste à étudier l'effet d'un choc d'une innovation sur les valeurs courantes et futures des variables endogènes. En effet, un choc sur une variable peut affecter non seulement cette variable mais également toutes les autres variables.

La propagation des chocs est également analysée à partir de la décomposition de la variance de l'erreur de prévision des différentes séries du système (Sims, 1980). Cette décomposition permet de mesurer la contribution de chaque innovation à la variance totale de l'erreur de prévision du processus X_{it} .

$$Var(X_{t+h} - \hat{X}_{t+h}) = \sum_{j=1}^n \sum_{k=0}^{h-1} (\psi_{ij}, k) (\sigma_{\varepsilon_j})^2$$

$(\sigma_{\varepsilon_j})^2$ est la variance de la $i^{\text{ème}}$ impulsion structurelle ε_{jt} . On observe que ε_{jt} a un effet sur X_{t+h} si et seulement si l'un des multiplicateurs $\psi_{ij}, k, 0 \leq k \leq h-1$ est non nul.

Selon les principes donnés par Sims (1980, 1981), les innovations canoniques v_{it} sont estimées à partir d'une représentation VAR canonique, comme résidus des régressions qui correspondent à l'estimation, équation par équation de modèle VAR :

$$\hat{V} = X_{it} - \left[\hat{\varphi}_0 + \sum \hat{\varphi}_t X_{t-t} \right]; 1 \leq i \leq n \text{ avec } \varphi_i = \bar{\varphi}_i; i = 0, \dots, p$$

Dès que la matrice de passage B est estimée, les impulsions structurelles ε_t , économiquement interprétables ε_t , sont identifiées à partir de la relation $\hat{\varepsilon}_t = B \hat{v}_t$.

Cette analyse impulsionnelle permet alors de mesurer et d'anticiper les effets d'une **politique économique**.

Les innovations canoniques **sont associées** aux processus VAR non contraint et représente des chocs ou impulsions dont la **propagation se traduit par** les fluctuations du système dynamique étudié. L'analyse statistique correspondante suppose que les impulsions sont non corrélées.

Les réponses des X_{it} aux différentes innovations $v_{is}, s \leq t$ sont données par les multiplicateurs dynamiques :

$$\psi_{ij,t-s} = \frac{\partial X_{it}}{\partial v_{jt}}$$

$\psi_{ij,t-s}$ désigne l'effet du choc j sur la variable i , $t-s$ périodes après le choc.

On définit ainsi des fonctions de réponses impulsionnelles :

$$\forall k \geq 0; k \rightarrow \psi_{ij,k}$$

La matrice de passage B comprend n^2 paramètres inconnus. Pour pouvoir les identifier facilement, on suppose que pour une date donnée, les chocs structurels ne sont pas corrélés entre eux et qu'ils ont une variance nulle.

C'est-à-dire que : $Var(\varepsilon_t) = I$

Soulignons le fait que cette hypothèse est faite plus pour des raisons techniques. En effet, on peut se demander pour quoi un choc d'offre doit être à chaque instant non corrélé avec un choc de demande.

$$\text{On a } Var(\varepsilon_t) = I \Leftrightarrow Var(v_t) = B^{-1} var(\varepsilon_t) B^{-1} = BB' = \Xi \varepsilon$$

Comme Ξ_ε est une matrice symétrique, on impose $n(n+1)/2$ contraintes sur les éléments de la matrices B . Ces contraintes sont appelées contraintes d'orthogonalisation. Il existe plusieurs méthodes permettant d'imposer ces contraintes. La plus simple est la décomposition de Choleski de la matrice de variance Ξ , qui fournit l'unique matrice triangulaire inférieure B telle que $BB' = \Xi$.

Pour un ordre donné des composantes du VAR, B est unique. Ainsi l'ordre des séries (de la plus exogène à la plus endogène) est important et est le seul à priori.

Le choix de cet ordre des variables est justifié par le fait que le Var structurel peut être considéré comme un système d'équations simultanées récursif pour le quel les séries sont influencées par un nombre croissant des variables contemporaines et deviennent donc de plus en plus endogènes.

La principale critique formulée à l'endroit de cette décomposition de Choleski par les partisans de la méthodologie Var structurel concerne le choix de $n(n+1)/2$ contraintes. Pour eux, les contraintes identifiantes doivent être tirées de la théorie économique (Shapiro et Watson(1988), Blanchard et Quah(1989), King et alii(1991), par exemple) et portent sur les effets de court et de long terme des chocs structurels sur les différentes composantes du système.

Dans le cas de la dynamique stationnaire, l'absence d'effet instantané de certaines impulsions structurelles sur certaines séries est formulée en terme des contraintes de court terme, en considérant comme nuls certains éléments de la matrice de passage P.

En effet, les réponses instantanées aux chocs sont données par les éléments de la matrice P

Considérons la représentation VMA canonique :

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \psi(L) \varepsilon_t$$

$$\Rightarrow X_t = \sum_{j=0}^{\infty} \psi_j P P^{-1} \varepsilon_{t-j} = \sum_{j=0}^{\infty} \Omega_j W_{t-j}$$

$$\text{avec } \Omega_j = \psi_j P \text{ et } W_t = P^{-1} \varepsilon_t$$

Ω_j est la matrice des multiplicateurs dynamiques structurels et on a :

$$\Omega_{ij,t-s} = \frac{\partial X_{it}}{\partial W_{js}}$$

Les réponses instantanées (t=s) sont donc :

$$\Omega_{ij,0} = \frac{\partial X_{it}}{\partial W_{js}} = \psi_{ij,0} P = P_{ij} \text{ car } \psi_{ij,0} = I$$

Dans le cas de dynamiques non stationnaires (VAR en différences), on peut imposer des contraintes de long terme, pour exprimer le fait que certaines impulsions structurelles n'ont pas d'effet de long terme sur certaines composantes du système. Les effets de long terme des chocs structurels ne peuvent être mis en évidence que si la dynamique étudiée a une composante persistante qui est due à la présence de racines unitaires. Les multiplicateurs dynamiques de long terme définis à partir de l'écriture VMA ou de la décomposition de Wold du Var structurel en différence première :

$$\Delta X_t = \sum_{h=0}^{\infty} \Omega_h W_{t-h}$$

En considérant la composante $X_{it} = \sum_{h=0}^{t-1} \Delta X_{it-h} + X_{i0}$, on a que sa réponse ou choc W_{js} est égal au cumul des réponses des différences premières $\frac{\partial \Delta X_{it-h}}{\partial W_{js}}$; $h \leq t-s$, à ce même choc.

Comme $\frac{\partial \Delta X_{it-h}}{\partial W_{js}} = \Omega_{ij,h}$, donc la réponse de X_{it} au choc w_{js} est

$$\sum_{h=0}^{t-s} \Omega_{ij,h}$$

La réponse à long terme, notée $\Omega_{ij}(1)$ est donc

$$\Omega_{ij}(1) = \lim_{t \rightarrow \infty} \sum_{h=0}^{t-s} \Omega_{ij,h}$$

Cet effet à long terme est persistant, en d'autre terme, il est un effet sur la composante permanente de la série. Imposer une contrainte à long terme, signifie considérer l'absence de réponse de long terme d'une composante X_i à une impulsion w_j .

Ceci se traduit par $\Omega_{ij}(1) = 0$, qui est une contrainte linéaire sur les éléments p_{kj} de la matrice P , $1 \leq k \leq n$, car $\Omega_{ij}(1) = \psi(1)P$.

De cette façon, la contrainte linéaire s'écrit :

$$\Omega_{ij}(1) = (\psi(1)P)_{ij} = 0$$

$$\Leftrightarrow \sum_{k=1}^n \psi_{ik}(1) p_{kj} = 0$$

Application à la politique monétaire au Rwanda

L'objet de ce paragraphe est d'analyser les effets de la politique monétaire au Rwanda sur l'activité et sur les prix. Il faut noter ici qu'il s'agit de l'un des intérêts de la modélisation var structurel par rapport aux modèles macro-économique traditionnels. Une telle modélisation cible de manière précise la question économique que l'on souhaite étudier. Pour répondre à cette question, nous avons retenu le système suivant (PIB, π , r), comprenant le PIB constant, le taux d'inflation et le taux d'intérêt créditeur en terme nominal¹.

L'analyse des propriétés de ces différentes variables montrent qu'elles sont toutes $I(1)$ ². Les critères de Akaike, de Schwarz ainsi que la Log-Vraisemblance (LV) nous ont permis de déterminer l'ordre p du VAR. La valeur retenue est $p=1$, comme l'indique le tableau n01.

L'estimation du VAR(1) en niveau montre que le PIB dépend positivement de sa valeur retardée. Les valeurs entre crochets sont des t de Student des paramètres estimés. Le taux d'inflation quant à lui dépend négativement du PIB retardé et positivement de sa valeur retardée. Cette même estimation indique le taux d'intérêt créditeur est positivement lié au taux d'inflation retardée et à sa propre valeur retardée.

L'estimation du VAR en différence qui décrit les relations à court terme montre que le taux de croissance du PIB est positivement lié à sa valeur retardée et que le taux de variation de l'inflation est négativement liée à ce lui de la masse monétaire et celui de du PIB retardé (seulement au seuil de 10%). Ce même modèle montre également que le taux de croissance du taux d'intérêt est positivement lié à celui de l'inflation retardée d'une période (Voir tableaux 2 et 3). Le tableau n04 montre que dans l'équation de l'inflation, le coefficient correcteur est significativement négatif.

¹ Certaines études utilisent un différentiel des taux d'intérêts, pour tenir compte de l'internationalisation des actifs financiers. On peut citer par exemple l'article de Catherine Bruneau qui utilise un différentiel entre le taux d'intérêt à court terme en France et en Allemagne.

² Nous avons utilisé les tests de Dickey Fuller Augmenté (ADF) ainsi que celui de Perron

Ceci signifie que les effets d'un choc sur l'inflation sont corrigés à raison d'environ 23.5% par trimestre.

D'autres instruments d'analyse dans un tel type de modélisation sont les fonctions de réponses impulsionnelles sur les dynamiques générées par les chocs.

Les graphiques de ces fonctions montrent que l'inflation réagit instantanément à un choc monétaire et que les effets de ce choc s'estompent au bout d'environ une année. Par contre, elle ne réagit pas directement au choc sur la production, mais les effets d'un tels chocs sont plus durables (8 trimestres) que ceux des chocs de l'offre de monnaie. Ces graphiques montrent également que tout changement dans la quantité de la monnaie, dans le volume de la production et à la hauteur de l'inflation influence le taux d'intérêt créditeur.

On observe une baisse du taux d'intérêt et une hausse de la masse monétaire presque simultanément. Il s'agit là d'une impulsion de choc d'offre de monnaie et on peut penser au mécanisme de substitution à la Keynes.

En effet, si on considère que les modalités d'intervention des autorités monétaires consiste en une intervention sur le marché monétaire, cette action agit sur les taux d'intérêt qui servent de support à ces opérations. Si l'offre de monnaie s'accroît, toutes choses restant égales par ailleurs, les agents économiques ont des encaisses plus importantes que ceux qui correspondent à leurs besoins.

Le graphique 1 permet d'examiner le mécanisme par lequel la politique monétaire au Rwanda se transmet à l'économie réelle, c'est-à-dire le mécanisme par lequel les liens entre le monde de la monnaie et niveau d'activité s'établissent.

Ce graphique montre que l'accroissement de l'offre de monnaie qui a entraîné la baisse du taux d'intérêt, va encourager l'investissement et celui-ci agira positivement sur le revenu. Ceci signifie que les agents économiques rwandais trouvent plus profitable d'utiliser l'excédent des encaisses pour l'investissement que de les placer en compte bancaire.

Ces résultats sont confirmés par les fonctions de réponse impulsionnaires qui indiquent que la croissance de l'offre de monnaie entraîne celle de la production, mais pas de manière instantanée.

Ces résultats permettent de souligner l'influence des chocs de politique monétaire sur la conjoncture économique, qui se transmet probablement par la variation du taux d'intérêt. Ce qui laisserai penser que le taux d'intérêt constituerait un canal important de la transmission de la politique monétaire au Rwanda.

Annexes

Tableau n01 : Détermination de l'ordre du VAR

P	AIC	SC	LV
1	2.42	3	-80.454
2	2.28	3.34	-57.84
3	2.61	4.15	-53.83
4	2.85	4.88	-46.34
5	3.006	5.52	-34.73
6	3.30	6.32	-28.87
7	3.79	7.32	-30.09
8	4.27	8.32	-30.33
9	4.38	8.95	-16.32
10	4.73	9.84	-11.34

Tableau n02: Résultats de l'estimation du VAR(1) en niveau

	LM2	LPIBCT	TINF	TIC
LM2(-1)	0.994220 (0.01278) [77.7873]	-0.001403 (0.01487) [-0.09432]	-0.081941 (0.54146) [-0.15133]	0.261002 (0.17822) [1.46451]
LPIBCT(-1)	-0.041602 (0.04523) [-0.91988]	0.946256 (0.05262) [17.9814]	-7.024396 (1.91591) [-3.66635]	0.668543 (0.63061) [1.06015]
TINF(-1)	-0.000190 (0.00245) [-0.07769]	0.002360 (0.00285) [0.82750]	0.428600 (0.10384) [4.12741]	0.130468 (0.03418) [3.81716]

TIC(-1)	0.004991 (0.00413) [1.20827]	-0.002893 (0.00481) [-0.60191]	-0.034485 (0.17500) [-0.19705]	0.789913 (0.05760) [13.7134]
C	0.200889 (0.21105) [0.95187]	0.232568 (0.24557) [0.94705]	28.56780 (8.94065) [3.19527]	-3.775835 (2.94277) [-1.28309]
R-squared	0.989959	0.865098	0.502169	0.825997
Adj. R-squared	0.989444	0.858180	0.476639	0.817074

Tableau n03 : Résultats de l'estimation du VAR(1) en différences

	DLM2	DLPIBCT	DTINF	DTIC
DLM2(-1)	-0.170017 (0.11475) [-1.48163]	0.065785 (0.12495) [0.52648]	-11.84896 (5.69918) [-2.07906]	-2.504887 (1.70305) [-1.47082]
DLPIBCT(-1)	0.330631 (0.10142) [3.26001]	0.490249 (0.11044) [4.43910]	-9.164548 (5.03718) [-1.81938]	2.375971 (1.50523) [1.57848]
DTINF(-1)	0.002058 (0.00227) [0.90668]	-0.001581 (0.00247) [-0.63987]	-0.022755 (0.11271) [-0.20189]	0.145424 (0.03368) [4.31763]
DTIC(-1)	-0.000374 (0.00690) [-0.05413]	-0.002670 (0.00752) [-0.35506]	-0.118212 (0.34292) [-0.34472]	0.028334 (0.10247) [0.27650]
C	0.028266 (0.00712) [3.97159]	-0.003013 (0.00775) [-0.38882]	0.332710 (0.35348) [0.94126]	0.088132 (0.10563) [0.83437]
R-squared	0.125418	0.241599	0.123688	0.199703
Adj. R-squared	0.079986	0.202202	0.078166	0.158129

Tableau n04 : Estimation du VCEM

Cointegrating Eq: CointEq1				
LM2(-1)	1,000000			
LPIBCT(-1)	23,53359 (7,05430) [3,33606]			
TINF(-1)	3,044189 (0,44919) [6,77712]			
TIC(-1)	-0,793512 (0,55960) [-1,41800]			
C	-98,47980			
Error Correction: D(LM2) D(LPIBCT) D(TINF) D(TIC)				
CointEq1	-1,18E-06 (0,00100) [-0,00118]	0,001135 (0,00108) [1,05095]	-0,234832 (0,04168) [-5,63463]	0,024933 (0,01455) [1,71359]
D(LM2(-1))	-0,169986 (0,11841) [-1,43556]	0,036146 (0,12801) [0,28236]	-5,717858 (4,93918) [-1,15765]	-3,155861 (1,72440) [-1,83012]
D(LPIBCT(-1))	0,330657 (0,10434) [3,16890]	0,465713 (0,11281) [4,12843]	-4,088977 (4,35243) [-0,93947]	1,837070 (1,51955) [1,20896]
D(TINF(-1))	0,002059 (0,00269) [0,76539]	-0,003196 (0,00291) [-1,09893]	0,311376 (0,11223) [2,77448]	0,109947 (0,03918) [2,80607]

D(TIC(-1))	-0.000371 (0.00736) [-0.05042]	-0.005409 (0.00795) [-0.68011]	0.448411 (0.30684) [1.46140]	-0.031828 (0.10713) [-0.29711]
C	0.028265 (0.00721) [3.92260]	-0.002131 (0.00779) [-0.27353]	0.150151 (0.30056) [0.49957]	0.107516 (0.10493) [1.02460]
R-squared	0.125418	0.252463	0.381900	0.229473
Adj. R-squared	0.067880	0.203283	0.341235	0.178781
Sum sq. resids	0.271444	0.317254	472.2858	57.56672

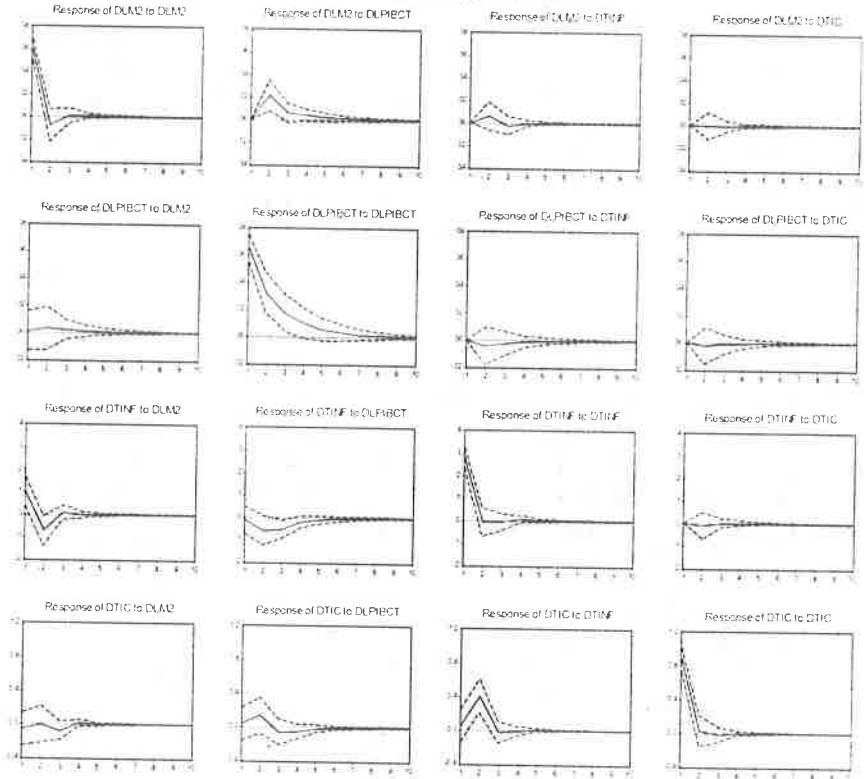
21

S.E. equation	0.059763	0.064610	2.492847	0.870320
F-statistic	2.179740	5.133453	9.391482	4.526762
Log likelihood	117.7865	111.3926	-188.1384	-101.8482
Akaike AIC	-2.726501	-2.570552	4.735083	2.630444
Schwarz SC	-2.550399	-2.394451	4.911185	2.806545
Mean dependent	0.024945	-0.000214	0.013902	0.034878
S.D. dependent	0.061901	0.072384	3.071358	0.960394

Determinant	Residual	5.63E-05
Covariance		
Log Likelihood		-51.77472
Log Likelihood (d.f. adjusted)		-64.23641
Akaike Information Criteria		2.249668
Schwarz Criteria		3.071475

Graphique1 : Analyse des réponses impulsionnelles

Response to Cholesky One S.D. Innovations ± 2 S.E.



Références bibliographiques

1. Allain de Serres, Allain Guay, Pierre St-Amant, Estimating and projecting potentiel Out put. Using structural VAR methodology. The case of Mexican Economy.
Bank of Canada, W.P. 1995
2. Carlo Giannini, Topics in structural VAR.
Econometrics, 2nd edition, Springer-Verlag 1997
3. Cathérine Bruneau, Olivier De Bandt, La modélisation Var structurel: application à la politique monétaire en France.
Banque de France, janvier 1998
4. Chris Brooks, Introductory econometrics for finance.
Cambridge University press, 2002
5. Frédérique Bec et Jean-Olivier Hairault, Taux d'intérêt, politique monétaire et activité économique en France : un examen empirique.
CREST et MAD Paris I, 1993
6. S. Johansen, Likelihood-based inference in cointegration Cector Auto-regressive models. Advanced texts in econometrics.
Oxford University Press inc. New York 1995.
7. S. Johansen, estimation and hypothesis testing of cointegration Vectors in Gaussian
Vector Auto Regression models
8. Sandrine Lardic, Valérie Mignon, Econométrie des séries temporelles macro économiques et financiers. Economica, 2002

**APPROCHES DE QUANTIFICATION
ET
DE MODELISATION EN ECONOMIE:
L'APPORT DE L'ECONOMETRIE***

*Théopiste Butare, PhD
Bureau international du Travail
Genève*

* Le texte proposé ici est une version enrichie de la conférence que j'ai présentée à l'Université Libre de Kigali (ULK) le 15 août 2000. Je voudrais remercier les étudiants de Licence Economie de l'ULK qui ont participé à cette conférence, ainsi que M. Rusuhuzwa Kigabo, Directeur académique de cette Université, pour leurs remarques très pertinentes. Evidemment, toute erreur factuelle ou de jugement qui subsisterait dans ce texte ne saurait leur être attribuée et devra sans conteste être imputée à l'auteur.

"Une compréhension du réel ne peut naître que d'une pensée claire et générale. Donc formalisée au sens où cela doit l'être aujourd'hui; et pour cela naturellement quantitative. Mais cette pensée générale, claire et cohérente, se doit d'aboutir aux faits observables, ce qui en langage conventionnel signifie qu'elle ne peut aboutir qu'à l'expérience. Expérience qui incite et soutient la réflexion de ceux qui s'inspirent d'une pensée dite opérationnelle".

(L. Solari, *De l'économie qualitative à l'économie quantitative*,
Avant-propos.)

Introduction

- **Que font les économètres?**

D'abord, et avant tout, les économètres sont des économistes, capables d'utiliser la théorie économique pour améliorer leurs analyses empiriques des problèmes auxquels ils sont confrontés. Par moments ils font un travail de mathématiciens, formulant la théorie économique de manière à la rendre analysable à l'aide de tests statistiques. Par moments ils font un travail de comptables, se préoccupant du problème de trouver et de collecter des données économiques et de faire le lien entre variables économiques théoriques et variables observables. Par moments ils font un travail de statisticiens praticiens, passant des heures sur un ordinateur en essayant d'estimer des relations économiques ou de prédire des événements du monde économique. Et par moments ils font aussi un travail de statisticiens théoriciens, utilisant leurs connaissances pour développer des techniques statistiques appropriées pour les problèmes économiques caractérisant la science économique. C'est souvent à ce dernier élément que sont consacrés les manuels de "théorie économétrique". Mais ajoutons que ce qui distingue fondamentalement les économètres des statisticiens, c'est la préoccupation des économètres à trouver des solutions aux problèmes causés par la violation des hypothèses standards posées par les statisticiens; en effet, compte tenu de la nature des relations économiques et de l'absence d'expériences contrôlées dans la vie économique réelle, ces hypothèses sont rarement vérifiées (voir Kennedy, 1994, Butare, 2002).

En postulant un ensemble de relations mathématiques et un corps d'hypothèses caractérisant ces relations, on met en place un "modèle". On peut définir tout modèle comme étant une représentation simplifiée de la réalité, le principal problème du modélisateur étant alors de trouver un bon compromis entre réalisme et facilité de traitement. Le monde étant compliqué, il est difficile de donner des règles générales permettant d'atteindre un tel compromis, et c'est pourquoi la modélisation en économie peut être considérée comme un art qui est soumis à des considérations scientifiques, sans pouvoir être pleinement déterminé par ces considérations.

Au niveau des variables, il est fréquent qu'un modèle établisse une coupure en distinguant entre variables exogènes et variables endogènes. Les variables exogènes représentent des grandeurs considérées comme fixées «en dehors du modèle», en dehors du phénomène qu'il représente. Par opposition, les variables endogènes sont considérées comme déterminées par ce phénomène.

Si on désigne par y les variables endogènes et par x les variables exogènes, et qu'on introduit des variables stochastiques désignées par u , on peut écrire la représentation formelle suivante, où t est l'indice de l'unité temporelle:

$\psi(y_t, x_t, u_t) = 0$. Plus généralement, y_t et x_t seront identifiés à des vecteurs réels.

• **Modélisation en économie: apports des théories macroéconomiquesⁱ**

Au fil du temps, les économistes se sont rendus compte qu'ils ne pouvaient pas se satisfaire de présenter des idées dans un cadre théorique cohérent, mais qu'ils devaient aussi recourir à des approches économétriques permettant de vérifier empiriquement leurs affirmations. Historiquement, on constate que l'essor de l'économétrie est étroitement lié à celui de la macroéconomie, toutes deux d'apparition relativement récente, et s'enrichissant mutuellement. Ainsi la représentation de grandeurs telles que la

production, l'emploi ou les prix, par des systèmes mathématiques formalisés et quantifiés susceptibles de résolution, a connu un développement tout à fait remarquable depuis le milieu des années trente.ⁱⁱ

Dès les années cinquante, les progrès dans l'application des techniques économétriques à la macroéconomie se sont accélérés, et c'est dès le début des années soixante qu'il y a eu une controverse restée célèbre entre keynésiens et monétaristes, ces derniers contestant les relations économétriques utilisées par les keynésiens pour expliquer notamment la fonction de consommation.ⁱⁱⁱ

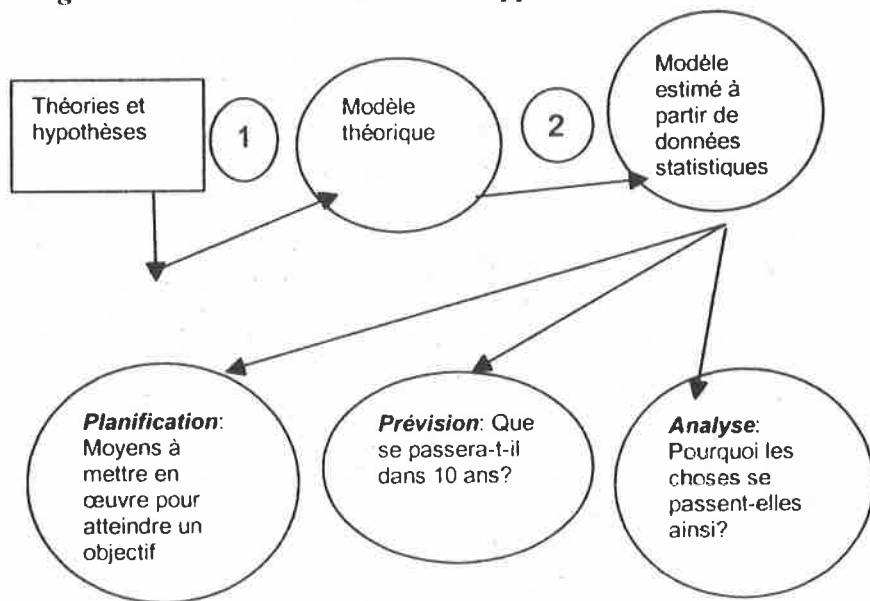
Un aspect qui a longtemps alimenté les débats porte sur la construction de modèles macroéconomiques quantitatifs de très grande taille, qui a été rendue possible par les progrès impressionnants réalisés dans le domaine de l'informatique durant les cinquante dernières années. Une des composantes de ces débats porte sur l'arbitrage à faire entre, d'une part, un niveau très poussé du détail dans la description et l'explication des systèmes à modéliser et, d'autre part, la nécessité d'éviter un éparpillement excessif qui apparaît dès lors qu'on a une infinité de relations à quantifier et qui pourrait être préjudiciable à une bonne maîtrise de l'ensemble.^{iv}

- **Econométrie et économie mathématique: quels liens?**

L'économie mathématique est la branche de l'économie qui se distingue des approches non mathématiques par le fait que les hypothèses de base et les conclusions de l'analyse sont présentées à l'aide de symboles mathématiques plutôt qu'avec simplement des mots, et à l'aide d'équations plutôt que de phrases. L'avantage avec cette façon d'appréhender les problèmes économiques est que les symboles mathématiques permettent de présenter de façon simple, concise et précise les enchaînements que propose la théorie.^v

On peut ainsi dire que, là où l'économétrie s'intéresse à appliquer aux observations économiques les méthodes d'estimation et de tests d'hypothèses, l'économie mathématique, quant à elle, cherche à appliquer les mathématiques à l'analyse économique sous un angle purement théorique. Les aspects statistiques ou d'erreurs de mesure ne sont pas du domaine de cette branche, qui se définit par l'application des mathématiques à un raisonnement déductif partant de la théorie. Le diagramme 1 ci-après fournit une schématisation des étapes successives d'une analyse économique quantitative partant de la formalisation propre à l'économie mathématique jusqu'à l'application des méthodes économétriques et aux différentes utilisations qu'on peut faire des résultats de tout l'exercice de modélisation.

Diagramme 1: Liens entre théories et applications en économie



Si on se réfère au diagramme I ci-dessus, on peut dire qu'avec l'économie mathématique on s'attachera à donner une formalisation à la partie ①; avec le passage de la partie ① à la partie ② on expliquera les liens entre le modèle et la théorie sous-jacente, et le passage de la partie ② à la partie ③ fait recours aux méthodes économétriques. Tout ce qui vient comme étapes ultérieures à la partie ③ constitue le détail des utilisations possibles des résultats de la modélisation (théorie, modélisation théorique et explorations économétriques). Comme il est extrêmement difficile de parler de tous ces points en un seul exposé, nous traiterons essentiellement de l'étude de techniques d'estimation et d'analyse faisant partie de ce que nous venons d'appeler le passage de ② à ③. Nous parlerons précisément de certaines méthodologies d'estimation, c'est-à-dire de certaines façons d'obtenir, à partir des données disponibles, des estimations pour les paramètres inconnus qui interviennent dans les théories et modèles que nous aurons mis en place.

Si on voulait considérer un cas simplifié de la relation $\psi(y_t, x_t, u_t) = 0$, qui serait par exemple la forme réduite d'un système simultané, on écrirait

$$y_t = f(x_t, u_t).$$

La vraie forme de la fonction $f(\cdot)$ et des paramètres qu'elle fait intervenir sont généralement inconnus. La théorie économique donne quelques indications sur les signes des dérivées premières de la fonction $f(\cdot)$ et éventuellement sur ses propriétés de convexité. Dans la plupart des cas, les analystes choisissent d'utiliser des formes fonctionnelles qui sont linéaires dans les paramètres. Ces formulations sont habituellement justifiées comme étant une expansion, à un certain point, de la vraie fonction. Mais il est bien connu que dans certains cas l'utilisation de formes fonctionnelles linéaires dans les paramètres peut conduire à de sérieuses erreurs de spécification. L'utilisation de spécifications non linéaires ne résout pas automatiquement ce problème d'erreur de spécification, à moins qu'elle soit fondée sur des arguments théoriques suffisants. Dans le cas contraire, elle ferait même perdre l'avantage lié à la facilité de l'interprétation des spécifications linéaires.

D'un point de vue statistique, l'objet de la régression est de modéliser et d'étudier sous tous ses aspects la relation entre variables dépendantes et variables explicatives. Choisir un estimateur revient généralement, dans le cadre des méthodes des moindres carrés, à minimiser la distance entre le modèle et les observations.

- **Portée et plan de la présentation**

Le but du présent exposé est essentiellement de présenter un certain nombre de méthodes économétriques qui, tout en ne faisant pas partie des techniques les plus enseignées dans le cadre du programme universitaire de licence en économie, n'en constituent pas moins des outils fréquemment utilisés de nos jours par les spécialistes en économétrie appliquée et par les universitaires qui se spécialisent dans cette branche après leur licence. Nous ne proposons donc pas ici un article de type classique cherchant à démontrer quelque chose de nouveau, mais plutôt un travail de sensibilisation et de diffusion de méthodes dont certaines paraissent techniquement difficiles ou basées sur des concepts peu familiers en apparence. C'est pourquoi nous nous attacherons à expliquer bien clairement la nature et l'apport de ces méthodes comparativement moins familières.

Une caractéristique de l'économétrie sur laquelle nous insisterons particulièrement est le fait que ses spécifications sont de nature aléatoire, ce qui la distingue d'autres approches de modélisation, telles que l'économie mathématique. Cette nature aléatoire s'explique par le fait que les phénomènes économiques sont influencés par un grand nombre de facteurs, dont certains ne sont pas observables ou ne sont pas explicitement présents dans le modèle utilisé, ce qui résulte en une complexité justifiant le recours à des modèles théoriques fondés sur l'idée de probabilités (voir Solari, 1977). Et les noms de certains modèles tels que les "modèles à erreurs composées" - une des spécialisations de l'auteur de cet article -, ou encore les "modèles à coefficients aléatoires", ou le critère même de "minimisation de la somme des carrés des résidus aléatoires" (à la base de la fameuse méthode des moindres

carrés ordinaires) traduisent bien le poids donné à cette caractéristique de l'économétrie, qui est assez présente dans les différentes sections de notre texte.

Le corps de l'article est structuré comme suit: La section 1 est consacrée à une présentation introductive du modèle classique de régression linéaire, ainsi qu'à ses extensions que sont les systèmes de régressions apparemment non liées, les systèmes d'équations simultanées, les modèles pour données en coupes spatio-temporelles. Dans la section 2 nous discutons de l'estimation des modèles non linéaires dans les paramètres. Les estimateurs robustes et les méthodes non paramétriques font l'objet des développements faits dans la section 3. Dans la section 4 nous proposons une introduction aux méthodes d'estimation issues de l'analyse des séries temporelles. La section 5 est consacrée aux modèles pour choix discrets et aux cas où on a des variables dépendantes limitées, et la section 6 présente les fondements et quelques illustrations de l'approche bayésienne en économétrie. Cette présentation est conclue par une section 7 consacrée à un résumé et à une brève discussion.

Deux annexes sont proposées à la fin du texte: une est consacrée aux principales fonctions de densité utilisées en analyse bayésienne, la deuxième est consacrée à une présentation synthétique de l'économétrie spatiale, qui est une branche particulière et relativement récente de l'économétrie.

1. Le modèle classique de régression linéaire

1.1. Modèle linéaire de base et méthode des moindres carrés ordinaires

L'ajustement d'une courbe entre deux variables et l'estimation de la relation entre ces deux variables par la méthode des moindres carrés ordinaires constituent un sujet bien connu depuis la naissance de la discipline économétrique - autour de l'année 1930 -

y compris pour le cas dit de régression multiple (plusieurs variables explicatives dans le modèle). Nous rappelons simplement ici les notations de base pour le modèle et la formule d'estimation du vecteur de coefficients du modèle:

Avec N observations sur la variable dépendante et sur les variables explicatives (au nombre de k), on peut donner au modèle la formulation matricielle suivante:

$$Y = X\beta + e \quad (1.1)$$

avec, en utilisant le symbole ' pour la transposition,

$$Y = (Y_1, Y_2, \dots, Y_N)' ; e = (e_1, e_2, \dots, e_N)' ; \beta = (\beta_1, \beta_2, \dots, \beta_k)'$$

et X est une matrice de dimension Nxk, dont les colonnes sont, dans l'ordre

$$(1, 1, \dots, 1)' ; (X_{21}, X_{22}, \dots, X_{2N})' ; \dots ; (X_{k1}, X_{k2}, \dots, X_{kN})'$$

On suppose que les éléments de X sont fixes et de variance finie. De plus, la matrice X est supposée être de rang k, avec k plus petit que N. Le vecteur de résidus e est de distribution normale, avec une espérance mathématique nulle (c'est-à-dire que $E(e) = 0$) et une matrice de variances-covariances scalaire (ce qui veut dire qu'on peut écrire $E(ee') = \delta^2 I$, avec I pour la matrice identité d'ordre NxN, et δ^2 étant la variance, identique pour tous les termes e_i , $i = 1, \dots, N$).

On peut montrer que l'estimateur du vecteur de paramètres β par la méthode des moindres carrés ordinaires (mco) est celui donné par la formule suivante:

$$\beta^{mco} = (X'X)^{-1} X'Y$$

Les hypothèses de base du modèle permettent de montrer que l'estimateur β^{mco} sera non biaisé et aura une distribution normale.^{vi}

1.2. Systèmes à plusieurs équations de comportement

1.2.1. Equations apparemment non liées

Considérons ici, pour prendre un exemple, un système simple avec trois équations non simultanées:^{vii}

$$Y^1_t = \beta^1_1 + \beta^1_2 X^1_{2t} + \beta^1_3 X^1_{3t} + \dots + \beta^1_k X^1_{kt} + e^1_t$$

$$Y^2_t = \beta^2_1 + \beta^2_2 X^2_{2t} + \beta^2_3 X^2_{3t} + \dots + \beta^2_k X^2_{kt} + e^2_t$$

$$Y^3_t = \beta^3_1 + \beta^3_2 X^3_{2t} + \beta^3_3 X^3_{3t} + \dots + \beta^3_k X^3_{kt} + e^3_t$$

Comme illustration de ce type de modèle on peut par exemple penser à des systèmes de fonctions de consommation linéaires telles que les courbes d'Engel. La variable Y^1 pourrait représenter la demande de biens alimentaires, la variable Y^2 la demande d'articles vestimentaires, et Y^3 la demande d'autres biens de consommation.^{viii}

Dans le cas standard les variables endogènes et les variables exogènes sont observées sur un certain nombre d'unités temporelles. La méthode d'estimation proposée par Zellner (1962) pour ces cas dits de régressions apparemment non liées part principalement de l'hypothèse que les covariances contemporaines entre les équations du système ne sont pas nécessairement nulles. Ceci implique que la structure de la matrice des variances-covariances du système sera telle que pour estimer le vecteur entier des coefficients β , on préférera recourir à une technique des moindres carrés généralisés à la place des moindres carrés ordinaires.^{ix}

1.2.2. Systèmes d'équations simultanées

D'une façon générale, les systèmes d'équations simultanées sont par nature des systèmes tels que certaines équations ont une variable endogène que l'on retrouve comme variable explicative dans d'autres équations, ce qui rend la méthode des moindres carrés ordinaires inappropriée pour l'estimation des coefficients de régression.

Un exemple de système simultané à trois équations, présenté sous forme standard et sans spécification fonctionnelle paramétrique explicite, pourrait s'écrire de la manière suivante:

$$Y^1_t = f(X^1_t, X^2_t, \dots, X^k_t; Y^2_t, Y^3_t; e^1_t)$$

$$Y^2_t = f(X^1_t, X^2_t, \dots, X^k_t; Y^1_t, Y^3_t; e^2_t)$$

$$Y^3_t = f(X^1_t, X^2_t, \dots, X^k_t; Y^1_t, Y^2_t; e^3_t)$$

Il est important de bien distinguer le cas des systèmes simultanés de celui des systèmes d'équations apparemment non liées. Dans le cas du système simultané ci-dessus, on voit que les variables endogènes Y^2 et Y^3 interviennent dans l'équation expliquant Y^1 , et avec les équations expliquant Y^2 et Y^3 on peut faire le même type d'observation. Dans le cas des régressions apparemment non liées, seules les variables exogènes interviennent dans l'explication du comportement des variables endogènes.

Les systèmes à équations simultanées ont fait l'objet de développements poussés dans la littérature spécialisée, et différentes méthodes ont été proposées permettant d'estimer les paramètres de manière convergente et sans biais de simultanéité (voir par exemple Kmenta, 2000, ou Pindyck et Rubinfeld, 1991).

1.3. Modélisation avec données en coupes spatio-temporelles

L'idée qui est à la base du recours aux modèles utilisant des données en coupes spatio-temporelles est que dans beaucoup de domaines l'information est rare, et qu'il faut utiliser au mieux toutes les données accessibles.

Grâce aux progrès dans la collecte et le traitement des données statistiques, l'information à la disposition du modélisateur acquiert souvent une double dimension, une dimension chronologique (le temps) et une dimension transversale (les individus ou entités statistiques). L'économétrie des données de panel, une branche de l'économétrie qui a connu un essor considérable depuis une trentaine d'années, s'intéresse à la combinaison optimale de ces deux sources d'information. Elle traite des problèmes d'estimation et d'induction statistique dans un contexte de panel, c'est-à-dire quand N individus sont observés sur T périodes. Les modèles fondamentaux de l'économétrie des données de panel sont le modèle de la covariance, le modèle à erreurs composées et le modèle à coefficients aléatoires.

Nous considérerons ici, en vue de la présentation des spécificités de ces méthode, un modèle linéaire général dont chacune des observations a une double dimension:

- une dimension temporelle (unité temps: année ou trimestre ou mois, etc);
- une dimension spatiale (unité: région ou firme ou individu, etc.).

Soit la spécification suivante:

$$y_{it} = \alpha_{it}^0 + \alpha_{it}^1 Z_{it}^1 + \dots + \alpha_{it}^K Z_{it}^K + \varepsilon_{it}, \quad i=1, \dots, N; \quad t=1, \dots, T \quad (1.2)$$

Les variables explicatives Z^k ($k=1, \dots, K$) sont supposées fixes, non stochastiques, et indépendantes des erreurs.

En forme matricielle condensée, si on empile toutes les observations ensemble, on peut écrire le modèle ci-dessus de la manière suivante:

$$Y_{(NT \times 1)} = X\Xi + \gamma_{(NT \times 1)} \quad \text{avec } \Xi_{(K+1) \times 1} = (\alpha^0 \quad \alpha_{(K \times 1)})\tau, \quad X = [S_{NT} \quad Z_{(NT \times K)}],$$

S_{NT} étant le vecteur unité d'ordre $NT \times 1$ (on l'appelle aussi parfois vecteur somme, car l'opération $S_{NT}\tau V$ revient en effet à faire la somme des éléments du vecteur V , en supposant bien entendu que les vecteurs S_{NT} et V ont le même nombre d'éléments NT).

Une première possibilité d'estimation pour le modèle (3.2) consiste à admettre que la réponse de Y aux variables explicatives (y compris le terme constant) est la même pour chaque individu et pour chaque période, c'est-à-dire $\alpha_{it}^k = \alpha^k$ ($k=0, 1, \dots, K$), et à appliquer la méthode d'estimation choisie, par exemple les mco, sans poser aucune autre hypothèse, ni sur les observations, ni sur les variables stochastiques. Ceci ne serait pas intéressant, car on nierait par là la présence d'un quelconque effet individuel ou temporel.

D'un autre côté, si on décompose l'ensemble des observations en N individus considérés comme indépendants et qu'on calcule, pour tout i , un estimateur des moindres carrés ordinaires $\beta_i^{mco} = (X_i' X_i)^{-1} X_i' Y_i$, on refuse ainsi toute interaction entre individus (ou régions, etc.) et on court le risque d'être confronté à une insuffisance d'observations pour l'estimation, si par exemple $T < K$ (on peut ici imaginer, sans que cela apporte une complication supplémentaire à la méthode, le cas où pour différents individus il y aurait un nombre de coefficients K_i pouvant être différent d'individu à individu, avec pour certains individus un nombre de coefficients à estimer supérieur au nombre T d'observations temporelles à disposition).

Trois spécifications intéressantes pour l'estimation de modèles avec données spatio-temporelles ont été proposées dans la littérature: le modèle de la covariance, le modèle à coefficients aléatoires, et le modèle à erreurs composées.

1.3.1. Le modèle de la covariance

La spécification la plus courante de ce modèle consiste à admettre que

$$\alpha_{it}^k = \alpha^k \quad (k=1, \dots, K); \quad \alpha_{it}^0 = \alpha_i^0 \quad (\text{existence d'un effet générique lié à l'individu}).$$

On essaie ainsi de concilier deux objectifs presque contradictoires: d'une part la spécificité de l'individu, d'autre part la parcimonie des coefficients à estimer. Ainsi, pour limiter le nombre de coefficients à estimer, on considère que la spécificité de l'individu est seulement dans la constante α^0 . Le nombre de degrés de liberté du modèle est alors $NT-N-K$. Une variante de ce modèle consiste à considérer un effet générique lié au temps, en plus ou à la place de l'effet générique lié à l'individu. Les variables explicatives correspondant à ces effets génériques (vecteurs avec des 1 et des 0) sont usuellement appelées "variables muettes". Il existe dans certains travaux un exposé assez clair sur l'estimation de ce type de modèles, notamment par moindres carrés généralisés (voir par exemple Balestra, 1980).

1.3.2. Le modèle à coefficients aléatoires

Avec la spécification générale utilisant l'hypothèse de coefficients aléatoires, on a la décomposition suivante pour tout coefficient du modèle:

$$\alpha_{it}^k = \alpha^k + \alpha_i^k + \alpha_t^k,$$

où α^k est une composante fixe, alors que α_i^k et α_t^k sont des composantes aléatoires, propres à l'individu et au temps respectivement.

Pour l'estimation, on prend les composantes aléatoires α_i^k et α_t^k et on les fait basculer dans le terme stochastique additif au modèle en même temps que les variables explicatives qui vont avec, ce qui donne pour le modèle une nouvelle variable stochastique:

$$y_{it} = (\alpha^0 + \alpha_i^0 + \alpha_t^0) + (\alpha^1 + \alpha_i^1 + \alpha_t^1)z_{it}^1 + \dots + (\alpha^K + \alpha_i^K + \alpha_t^K)z_{it}^K + \varepsilon_{it}$$

$$= \alpha^0 + \alpha^1 z_{it}^1 + \dots + \alpha^K z_{it}^K + >_{it}$$

$$\text{avec } >_{it} = \alpha_i^0 + \alpha_t^0 + (\alpha_i^1 + \alpha_t^1)z_{it}^1 + \dots + (\alpha_i^K + \alpha_t^K)z_{it}^K + \varepsilon_{it}$$

Dans cette nouvelle présentation, $>_{it}$ est la nouvelle variable stochastique du modèle. En posant des hypothèses raisonnables sur la matrice $E(>>\tau)$, on arrive, par des procédures de moindres carrés généralisés en plusieurs étapes (Anizoba, 1980), ou par des méthodes de maximum de vraisemblance (voir par exemple Swamy, 1970), à obtenir des estimateurs convergents pour les coefficients. Il faut souligner le fait qu'en général avec ce modèle les procédures de minimisation de la somme des carrés des résidus font intervenir des expressions non linéaires dans les paramètres, ce qui entraîne des difficultés d'identification et le recours à certaines approximations dans les procédures d'estimation.

1.3.3. Le modèle à erreurs composées

Avec la spécification à erreurs composées on admet que $\alpha_{it}^k = \alpha^k$ ($k=0, 1, \dots, K$), hypothèse qui permet d'avoir un nombre relativement limité de coefficients à estimer, mais on retient aussi qu'il y a une variabilité considérable des variables explicatives, qui se répercute totalement dans le terme stochastique du modèle. On a ainsi

$$\varepsilon_{it} = u_i + v_t + w_{it},$$

où u_i est la composante individuelle, v_t est la composante temporelle, et w_{it} la composante résiduelle du terme aléatoire.

Ce sont les conditions posées sur les variables stochastiques, et en particulier sur la structure de la matrice des variances-covariances $E(\varepsilon\varepsilon')$ qui permettent, grâce à l'obtention de formules analytiques relativement simples, d'estimer les coefficients du vecteur β par moindres carrés généralisés. Ces hypothèses sont les suivantes:

$$E(u_i) = E(v_t) = E(w_{it}) = 0, \quad \forall i, j \Rightarrow E(\varepsilon_{it}) = 0, \quad \forall i, j.$$

$$E(u_i u_j) = \delta_u^2 \quad \text{si } i = j; \quad E(u_i u_j) = 0 \quad \text{si } i \neq j$$

$$E(v_t v_s) = \delta_v^2 \quad \text{si } t = s; \quad E(v_t v_s) = 0 \quad \text{si } t \neq s$$

$$E(w_{it} w_{js}) = \delta_w^2 \quad \text{si } t = s \text{ et } i = j; \quad E(w_{it} w_{js}) = 0 \quad \text{autrement.}$$

On aura alors

$$E(\varepsilon_{it} \varepsilon_{js}) = \delta_u^2 + \delta_v^2 + \delta_w^2 \quad \text{si } i = j \text{ et } t = s$$

$$E(\varepsilon_{it} \varepsilon_{js}) = \delta_u^2 \quad \text{si } i = j \text{ et } t \neq s$$

$$E(\varepsilon_{it} \varepsilon_{js}) = \delta_v^2 \quad \text{si } i \neq j \text{ et } t = s$$

Ainsi les seuls paramètres à connaître de la matrice $\Omega_{\varepsilon} = E(\varepsilon\varepsilon')$ sont les variances δ_u^2 , δ_v^2 et δ_w^2 . Si ces variances sont inconnues, des procédures immédiatement applicables existent, qui permettent d'avoir pour ces variances des estimateurs BQUE ("best quadratic unbiased estimators") et d'arriver ensuite à une estimation convergente des coefficients de régression, par une méthode de moindres carrés généralisés en deux étapes (voir Balestra, 1973).

La spécification avec erreurs composées est généralement la plus utilisée dans les travaux d'économétrie exploitant des données économiques en coupes spatio-temporelles. Des extensions de cette spécification ont été élaborées, introduisant par exemple l'hypothèse d'erreurs composées dans les systèmes simultanés

(Krishnakumar, 1988), ou proposant des méthodes pour les cas où il y a des composantes spatiales hétérogènes (Butare, 1990), ou combinant l'hypothèse d'erreurs composées avec celle de régressions apparemment non liées (Avery, 1977, Butare, 1990).^x

2. Estimation des modèles non linéaires dans les paramètres

Dans tous les domaines de la modélisation et de l'analyse des données, les chercheurs sont de plus en plus confrontés à des problèmes d'ajustement pour lesquels les méthodes linéaires ne donnent pas de résultats satisfaisants. Ils doivent alors faire appel à la théorie des modèles de régression non linéaire.^{xi}

La solution d'un problème non linéaire est souvent obtenue par une méthode itérative dans laquelle chaque itération consiste à résoudre un problème linéaire. Il est donc utile de se rappeler les concepts et les résultats essentiels de la régression linéaire avant de passer à la régression non linéaire.^{xii}

On parle de modèles de régression non linéaire lorsque l'expression mathématique de la fonction liant la variable dépendante aux régresseurs dépend d'un ou plusieurs paramètres n'intervenant pas tous de manière linéaire (c'est le cas si la dérivée de la fonction par rapport à au moins un paramètre n'est pas constante)^{xiii}. Pour prendre un exemple, considérons la fonction suivante:

$$Y = \alpha_0 + \alpha_1 X_1^{\beta_1} + \alpha_2 X_1^{\beta_2} + \epsilon \quad (2.1)$$

On remarquera qu'avec l'équation (2.1) ci-dessus il y a deux variables explicatives (X_1 et X_2) et quatre coefficients à estimer (α_1 , α_2 , β_1 , β_2). C'étant un terme d'erreur additif associé à l'équation. Avec une telle spécification il n'est plus possible de résoudre directement l'équation des moindres carrés de manière à avoir une expression analytique ou une formule permettant d'obtenir, sous forme de solution unique, des estimateurs pour les coefficients du modèle.

L'approche utilisée le plus souvent pour l'estimation des paramètres de modèles économétriques non linéaires consiste à linéariser l'équation de manière itérative, jusqu'à ce que le processus ait satisfait à un critère de convergence. Différents algorithmes permettant d'obtenir cette linéarisation existent dans la littérature, mais en général les différents types de procédures suivent globalement le même principe. Nous considérerons ici une procédure utilisant des développements de Taylor de la fonction à estimer, développements dans lesquels il s'agira d'introduire, pour commencer le processus, des valeurs initiales pour le vecteur de coefficients. La méthode est succinctement décrite ci-après, à l'aide d'une fonction tout à fait générale avec k variables explicatives et p coefficients à estimer:

$$Y = f(X_1, X_2, \dots, X_k, \beta_1, \beta_2, \dots, \beta_p) + \epsilon \quad (2.2)$$

On peut faire un développement de Taylor de cette fonction au voisinage de valeurs initiales $\beta_{1,0}, \beta_{2,0}, \dots, \beta_{p,0}$ pour les paramètres $\beta_1, \beta_2, \dots, \beta_p$.

Le développement de Taylor pour la fonction ci-dessus, explicité jusqu'au second degré, aura la forme suivante:

$$Y = f(X_1, X_2, \dots, X_k, \beta_{1,0}, \beta_{2,0}, \dots, \beta_{p,0}) + \sum_{i=1, \dots, p} (\beta_i - \beta_{i,0}) f'_0(\beta_i) + (1/2) \sum_{i=1, \dots, p} \sum_{j=1, \dots, p} (\beta_i - \beta_{i,0})(\beta_j - \beta_{j,0}) f''_0(\beta_i, \beta_j) + \dots + \epsilon \quad (2.3)$$

avec $f'_0(\beta_i) = (\partial f / \partial \beta_i)_0$; $f''_0(\beta_i, \beta_j) = (\partial^2 f / \partial \beta_i \partial \beta_j)_0$

On obtient une approximation linéaire estimable en éliminant les termes du second degré et des degrés plus élevés et en réinscrivant l'expression (2.3) de la manière suivante:

$$Y - f(X_1, X_2, \dots, X_k, \beta_{1,0}, \beta_{2,0}, \dots, \beta_{p,0}) + \sum_{i=1, \dots, p} \beta_{i,0} f'_0(\beta_i) = \sum_{i=1, \dots, p} \beta_i f'_0(\beta_i) + \epsilon \quad (2.4)$$

L'expression (2.4) obtenue ci-dessus a la forme d'une équation de régression linéaire. Le côté gauche de l'équation est une variable dépendante construite. Le côté droit consiste, outre le terme d'erreur additif ϵ , d'un ensemble de coefficients inconnus qui multiplient un ensemble de variables indépendantes construites. Les coefficients peuvent donc être estimés par moindres carrés ordinaires.

Les valeurs obtenues pour les coefficients dans cette régression, qui peuvent être notées $\beta_{1,1}$, $\beta_{2,1}$, ..., $\beta_{p,1}$, sont à leur tour utilisées comme valeurs initiales pour relinéariser de nouveau l'équation non linéaire autour de ces valeurs. Cela donne lieu à une nouvelle régression linéaire:

$$Y - f(X_1, X_2, \dots, X_k, \beta_{1,1}, \beta_{2,1}, \dots, \beta_{p,1}) + \sum_{i=1, \dots, p} \beta_{i,1} f_i(\beta_i) = \sum_{i=1, \dots, p} \beta_i f_i(\beta_i) + \epsilon \quad (2.5)$$

Les moindres carrés ordinaires sont de nouveau appliqués à cette équation, et on obtient un nouvel ensemble de valeurs pour les coefficients: $\beta_{1,2}$, $\beta_{2,2}$, ..., $\beta_{p,2}$. Le processus de linéarisation itérative est répété jusqu'à ce qu'il y ait convergence, c'est-à-dire jusqu'à l'étape où l'on obtient

$$|(\beta_{i,n} - \beta_{i,n+1})/\beta_{i,n}| < \delta, \quad i=1, \dots, p,$$

δ étant un nombre assez petit choisi en fonction de la charge représentée par les calculs que nécessite cette estimation.

Si le processus ne converge pas ou s'il ne donne pas lieu à un minimum absolu pour la somme des carrés des résidus, on peut recommencer depuis le début avec de nouvelles valeurs initiales. Si cela ne produit pas le résultat escompté, on peut alors changer de méthode d'estimation.

L'approche décrite ci-dessus présente un certain nombre d'avantages: on notera en particulier qu'elle permet d'avoir recours, pour juger de la qualité de l'ajustement, aux tests usuels de la régression classique, tels que le R^2 , les statistiques de Student, etc. . . puisqu'à chaque itération on a une régression linéaire. Par ailleurs, si l'équation à estimer peut être approximée de près par une équation linéaire, le processus pourrait ne requérir qu'un nombre très limité d'itérations, ce qui peut représenter un gain non négligeable en efficacité au niveau des calculs.

3. Estimateurs robustes et méthodes non paramétriques

3.1. Notion de robustesse et estimateurs robustes

3.1.1. Notion de robustesse

Avec les méthodes paramétriques on cherche généralement à obtenir pour le problème posé une solution satisfaisant à des qualités bien précisées, les plus recherchées étant la *validité* et l'*efficacité*. Cette recherche d'une solution se fait en mettant l'accent essentiellement sur son efficacité, sa validité étant souvent imposée ou assurée asymptotiquement. Par exemple, on cherche un estimateur de variance minimum (efficacité) parmi les estimateurs sans biais (validité); on cherche le test dont la puissance (efficacité) est la plus grande parmi les tests de seuil α (validité), etc.^{XIV}

Avec les méthodes de recherche de solutions robustes, on considère tout d'abord un modèle paramétrique et on souhaite tenir compte du fait que les observations peuvent très bien, pour diverses raisons, venir, non pas d'une loi précisée par ce modèle, mais plutôt d'une loi assez proche de celles du modèle paramétrique considéré. Les raisons de cette déviation possible du modèle peuvent être dues à des erreurs d'expérimentation, des erreurs de mesure suite à une détérioration du matériel, des perturbations aléatoires non prises en compte dans le modèle, des erreurs d'arrondis dans la collecte des données, etc. Pour représenter cette déviation, on introduit un ensemble de lois "voisines" de celles spécifiées par le modèle paramétrique de la façon suivante:

Notons F_0 une loi de l'ensemble correspondant au modèle paramétrique et \mathcal{B} l'ensemble de toutes les lois sur \mathbb{R} . Alors une loi voisine de F_0 est une loi de l'ensemble

$$\{ G; G = (1-\varepsilon)F_0 + \varepsilon H, H \in \mathcal{B}, \varepsilon \in [0,1] \}.$$

L'approche robuste consiste, intuitivement, à chercher des solutions qui :

- sont valides et ont une bonne efficacité pour F_0 ou, plus généralement, dans le modèle paramétrique choisi,
- restent relativement valides et efficaces si on s'éloigne un peu (ε petit) de F_0 vers une loi "voisine",
- n'ont pas une validité et une efficacité nulle lorsqu'on est très éloigné (ε voisin de 1) de F_0 .

Cette approche recherche donc des solutions qui peuvent nous protéger, dans l'analyse des observations et dans les conclusions qu'on tire, des nombreuses petites erreurs faites dans les mesures ou des erreurs grossières, si elles sont en faible nombre. Quand une méthode satisfait aux trois conditions ci-dessus, on dit qu'elle est robuste pour F_0 du point de vue de la validité et de l'efficacité.

3.1.2. *Estimateurs robustes*

Prenons le cas où des observations sur Y (variable dépendante) ou sur des variables explicatives sont considérées comme "aberrantes" ou "excessivement influentes". Dès que celles-ci ont été identifiées, il n'est généralement pas judicieux de les éliminer avant de procéder à l'estimation. Il arrive souvent en effet que ces observations soient les plus intéressantes de l'ensemble des données: si on suppose par exemple que pendant des années les taux d'intérêt ou les prix relatifs de l'énergie n'ont pas varié, et que tout d'un coup on observe une variation, l'observation montre que la donnée qui traduit cette nouvelle tendance est précisément celle qui permet d'avoir une meilleure prévision pour les périodes suivantes. Il est clair qu'on peut aussi trouver des situations où il y a eu erreur

dans la mesure ou la classification ou le traitement informatique des données, et dans ces cas-là le mieux est de chercher à corriger les fautes commises. Si on n'y arrive pas, le fait d'éliminer des observations "aberrantes" de ce type pourrait se justifier.

Mais s'il n'est pas prouvé que ces observations résultent d'une erreur, le mieux est de les garder et de recourir à une méthode d'estimation robuste. Nous donnons ci-après quelques approches que propose la littérature économétrique sur ce sujet (voir par exemple Kennedy, 1994, ou Becker et Greene, 2001):

- Minimiser une somme pondérée des valeurs absolues des résidus, en donnant des poids différents selon que les résidus calculés sont positifs ou négatifs. Cela se fait par des méthodes dites de "*régression par quantiles*": en prenant par exemple 0,25 comme quantile de régression, on a comme programme de minimiser une somme de valeurs absolues de résidus telle que le poids 0,25 est attribué aux résidus positifs et le poids 0,75 aux résidus négatifs. D'une manière générale, avec ce type de méthode considérant deux catégories de résidus (positifs, négatifs), on associe un poids θ aux résidus positifs et un poids $(1 - \theta)$ aux résidus négatifs, avec $0 < \theta < 1$.

Une variante souvent utilisée de la régression par quantiles consiste à donner aux 50% médians de la distribution des résidus un poids égal à 0,5 et un poids égal à 0,25 à chacun des deux quartiles extrêmes de la distribution (quartile de gauche et quartile de droite). On admet qu'avec cette méthode le choix des quantiles et de leurs poids est laissé complètement à l'appréciation du chercheur.^{xv}

Moindres carrés sur échantillons tronqués. Ceci est par nature une méthode d'estimation qui implique l'élimination d'observations. En éliminant par exemple les observations correspondant aux résidus se trouvant dans la partie 5% gauche de la distribution et dans la partie 5% droite de la distribution, on calcule un estimateur des moindres carrés dits α -tronqués, α étant ici égal à 10% (proportion totale d'observations éliminées).

- Estimateurs fondés sur le critère d'influence bornée. Dans le cas des moindres carrés ordinaires, l'influence d'une observation "aberrante" sur le coefficient estimé devient d'autant plus importante que l'observation augmente en valeur absolue. Différentes techniques sont proposées pour borner l'influence de ce type d'observations, et la nature de ces techniques peut sensiblement varier en fonction de ce que l'on entend par "influence" et de ce que l'on choisit comme bornes^{xvi}.

3.2. Régressions non paramétriques

Les méthodes non paramétriques sont des méthodes de recherche de solutions qui sont applicables pour un vaste ensemble de lois spécifiées par le modèle. Ces méthodes sont appelées à fournir des solutions ayant une bonne validité dans le modèle choisi et aussi une efficacité raisonnable. On demande généralement aussi qu'elles soient simples d'application. C'est souvent le cas pour les problèmes concernant des observations univariées, mais c'est malheureusement rarement le cas pour les observations multivariées et pour l'estimation non paramétrique de densités.

Dans le domaine de la régression, les méthodes non paramétriques s'appliquent généralement à des modèles tels que la relation entre la variable dépendante et les variables explicatives ne comporte pas de paramètres à estimer, la distribution des termes d'erreur elle-même n'ayant pas une structure connue. Cela permet d'éviter le recours à des hypothèses potentiellement erronées. Les méthodes non paramétriques peuvent ainsi être considérées comme un autre type de démarche visant à obtenir des estimateurs robustes. Nous formulons ici l'idée de la démarche en nous référant à un exemple fourni par Kennedy (1994):

On suppose que l'intérêt du modèle porte sur une relation du type $y = m(x) + \varepsilon$, pour laquelle le chercheur n'entend pas produire une structure paramétrique pour ε ou pour $m(x)$, qui représente ici l'espérance mathématique conditionnelle de y . Un ingrédient central de l'approche non paramétrique est l'estimation d'une fonction de densité conjointe pour y et x , estimation pour laquelle on suppose qu'il y a N observations y_i et x_i .

La densité en un point (y_0, x_0) est estimée par la proportion de celles parmi les N observations qui peuvent être considérées comme proches de (y_0, x_0) . Dès que la distribution conjointe a été estimée, il est possible de trouver la distribution marginale de y et la distribution conditionnelle de y étant donné x . Il est important de noter que ces distributions ne sont pas "trouvées" dans le sens qu'une formule est identifiée pour l'évaluation des densités, de la façon paramétrique habituelle. Ce qu'on veut dire ici c'est que pour une valeur donnée de x , la hauteur - pour parler en termes graphiques - de la densité conditionnelle peut être estimée.^{xvii}

La distribution conditionnelle estimée de y étant donné x permet ensuite de trouver d'autres grandeurs intéressantes pour le modèle: l'espérance mathématique conditionnelle $m(x)$ peut être estimée, ainsi qu'un analogue du "coefficient de régression", donné par la variation de $m(x)$ qui résulte d'une modification unitaire de x . La variance conditionnelle de y peut aussi être estimée, donnant ainsi une estimation du terme d'erreur.^{xviii}

Le "coefficient de régression" dans de tels modèles non paramétriques est la dérivée partielle de $m(x)$ par rapport à x . Une façon de l'estimer est de faire le calcul de $m(x)$ pour $x=x_0$ et ensuite pour $x=x_0+\delta$; la différence entre les deux valeurs estimées de $m(x)$, divisée par δ , est l'estimateur qui sera retenu pour le "coefficient de régression". La valeur de cet estimateur variera avec x , à moins que $m(x)$ ne soit linéaire.

Pour résumer on peut dire que lorsque dans une régression la vraie fonction n'est pas connue, tout choix d'une forme particulière est arbitraire. Les méthodes non paramétriques permettent d'éviter cet écueil. Dans cette optique on propose des estimateurs convergents des valeurs de la fonction et de ses dérivées premières aux points d'observation, basées uniquement sur l'existence et la continuité des dérivées secondes de la fonction.^{xix}

4. Méthodes d'estimation liées à l'analyse des séries temporelles

4.1. Introduction: Séries temporelles et processus stochastiques

Les techniques d'analyse des *séries temporelles* constituent une catégorie de méthodes visant à étudier l'évolution d'une variable, non plus en se fondant sur d'autres variables considérées dans un cadre d'analyse de causalités, mais en se référant uniquement aux seuls comportements passés de la variable prise en considération. En économie on se limite le plus souvent au cas des séries discrètes, c'est-à-dire composées d'observations se rapportant à des périodes de temps bien précises et identiques.

En nous limitant aux séries à temps discret nous considérons une suite de variables aléatoires réelles du type suivant:

$$\{X_t, t \in \mathbb{Z}\}.$$

La notion de *processus stochastique* est une extension en termes dynamiques de la notion classique de variable aléatoire. Plus précisément, on peut définir un processus stochastique X_t comme une suite ordonnée de variables aléatoires réelles, dont l'ordre est établi par la variable temps t . Par ailleurs, on peut considérer un processus comme un ensemble de mécanismes similaires, dont les réalisations (c'est-à-dire les séries temporelles engendrées) présentent des différences qui peuvent s'expliquer statistiquement. On peut dire que pour un instant temporel $t=t_0$, fixé à l'avance, x_{t_0} est simplement une variable aléatoire. La définition se généralise tout naturellement au cas où x_{t_0} est élément d'un processus vectoriel à n composantes.

4.2. Séries stationnaires, séries intégrées: le diable est dans les détails

4.2.1. Notion de stationnarité

On classe les processus stochastiques en processus stationnaires et en processus non stationnaires. La signification du terme de *stationnarité* est celle d'invariance de certaines propriétés statistiques du processus pour toute translation de l'axe du temps sur lui-même, ce qui assure une certaine permanence, ou homogénéité, par rapport au temps, des phénomènes considérés. Dans la présentation que nous faisons ci-après en rapport avec le concept de stationnarité, nous nous limiterons aux considérations suivantes:

i) *Stationnarité du premier ordre:*

Un processus X_t est dit stationnaire du premier ordre si son espérance mathématique est indépendante du temps, à savoir si:

$$E\{X_t\} = \mu, \quad t = \pm 1, \pm 2, \dots$$

(4.1)

ii) *Stationnarité au sens de la covariance:*

Un processus X_t est dit stationnaire au sens de la covariance si ses covariances sont invariantes à un changement de l'origine du temps, à savoir si:

$$E\{[X_t - E\{X_t\}] \cdot [X_{t+\bar{v}} - E\{X_{t+\bar{v}}\}]\} = R(\bar{v}),$$

(4.2)

$R(\bar{v})$ étant fonction du paramètre de décalage \bar{v} seulement.

La fonction $R(\bar{v})$, qui est par construction paire et de type positif (ce qui veut dire que la covariance entre X_t et $X_{t+\bar{v}}$ est la même que celle entre X_t et $X_{t-\bar{v}}$), est appelée fonction d'autocovariance du processus: pour $\bar{v}=0$, l'expression (4.2) donne la variance (constante) du processus, tandis que

pour

$\forall = \pm 1, \pm 2, \dots, \pm K$, l'expression (4.2) exprime les covariances entre les variables du processus espacées d'une, de deux, ..., de K périodes.

iii) *Stationnarité du second ordre:*

Un processus X_t est dit stationnaire du second ordre, ou stationnaire au sens faible, si :

- a) le processus est stationnaire du premier ordre;
- b) le processus est à covariance stationnaire.

Par la suite, on dira pour simplifier que les processus stationnaires du second ordre sont stationnaires. Un exemple de processus stationnaire est le *bruit blanc*.^{xx}

L'interprétation associée aux conditions de stationnarité est que les valeurs de X_t tendent à revenir à leur moyenne dans le temps, et que l'effet d'un choc sur une série stationnaire tend à ne pas persister dans le temps.

4.2.2. Implications de la non-stationnarité sur les méthodes économétriques

On peut dire que l'intérêt accordé au comportement stochastique des séries temporelles dans les procédures d'estimation des modèles économétriques commence réellement au début des années 1980. Avant cette période les économètres avaient souvent tendance à utiliser un modèle de régression traditionnel pour faire des estimations à partir de données chronologiques et ensuite à se poser des questions sur des aspects tels que la simultanéité ou l'autocorrélation des résidus, accordant très peu d'attention à la spécification de la structure dynamique des séries temporelles. Ils supposaient implicitement que les données économiques étaient stationnaires, ce qu'on pourrait interpréter comme un "comportement stable" de la part de ces données. D'un autre côté, les statisticiens spécialistes des séries temporelles avaient pour

habitude d'ignorer le rôle des variables économétriques explicatives, et de modéliser le comportement des séries temporelles en recourant à des mécanismes d'extrapolation sophistiqués. Ils contournaient le problème de la stationnarité en travaillant avec des données transformées en différences autant de fois qu'il était nécessaire pour les rendre stationnaires.

Les deux groupes restèrent longtemps sans que l'un fasse vraiment attention à ce que faisait l'autre. Dès le début des années 1980 donc, les résultats de deux catégories de recherches attirèrent fortement l'attention des économètres: la première catégorie d'études proclamaient que les prévisions fournies par les méthodologies économétriques étaient inférieures à celles faites avec les approches des spécialistes des séries temporelles; la deuxième catégorie affirmait que les données économiques ne sont en fait pas stationnaires, et que ceci pouvait conduire à de sérieux problèmes en ce qui concerne des statistiques traditionnelles telles que le R^2 , la statistique de Durbin-Watson ou la statistique de Student. On peut noter en particulier les travaux de Nelson et Plosser (1982), qui sont arrivés à la conclusion que beaucoup de séries macroéconomiques telles que la production industrielle, les prix à la consommation, les salaires, le Produit national brut, sont non stationnaires. Ces séries se caractérisent en particulier par un trend assez marqué (la moyenne change au fil du temps).

Ces observations ont poussé les économètres à beaucoup améliorer leur traitement des données chronologiques, tout en attirant l'attention sur les avantages comparatifs respectifs des méthodes économétriques classiques et des techniques d'analyse des séries temporelles: ces dernières fournissent des résumés efficaces des dépendances temporelles des données, et à ce titre elles constituent d'utiles repères pour la prévision. Les techniques d'analyse des séries temporelles ne peuvent toutefois pas à elles seules être considérées comme satisfaisantes quand il est question d'expliquer ou de comprendre le fonctionnement d'une économie. Ces techniques peuvent être considérées comme des compléments, et non pas comme des substituts, des modèles structurels traditionnels.

4.2.3. Tests de stationnarité, séries intégrées, cointégration et estimation^{xxi}

Avant d'appliquer une méthode d'estimation, une analyse approfondie des propriétés des séries univariées est indispensable. Ce préalable a pour principal objectif de révéler la (non) stationnarité des séries. C'est en fait l'étape de la détermination de leur ordre d'intégration. Pour cela, on peut se servir des outils classiques tels que la fonction d'autocorrélation estimée (voir par exemple l'utilisation de corrélogrammes présentée par Mills, 1990): pour une variable stationnaire, le corrélogramme aurait des coefficients d'autocorrélation se réduisant considérablement au fur et à mesure que le décalage \mathcal{U} augmente. Mais pour l'analyse de la stationnarité on utilise surtout les tests de racine unitaire de Dickey et Fuller, qui ont tout spécialement été élaborés pour répondre à ce besoin.

Considérons le modèle autorégressif d'ordre un donné ci-après pour illustrer la démarche de Dickey et Fuller:

$$y_t = \lambda y_{t-1} + \varepsilon_t$$

Dans ce processus, les ε_t sont identiquement et indépendamment distribués, à variance constante δ^2 , et avec espérance mathématique nulle.

Le premier test proposé par Dickey et Fuller est un test unilatéral à gauche avec $H_0: \lambda=1$ (dans ce cas le processus suit une marche aléatoire et y_t n'est donc pas stationnaire), contre l'hypothèse alternative $H_1: |\lambda| < 1$ (cas de y_t stationnaire): dans cette situation de stationnarité, les observations présentes ont un poids plus important que les observations passées.^{xxii} Avec la statistique t de Student calculée pour l'estimateur obtenu par moindres carrés pour λ , on fait la comparaison avec les valeurs théoriques des statistiques t disponibles dans les tabulations fournies par Dickey et

Fuller (1979) pour accepter ou rejeter l'hypothèse nulle. Ce test a principalement comme objectif de nous renseigner sur la nécessité de différencier la série (pour des exemples sur ce test et d'autres types de tests de racine unitaire, voir par exemple Mills, 1990, ou Griffiths et al., 1993).

On dit qu'une série Y_t est *intégrée d'ordre d* (notée $Y_t \rightarrow I(d)$), s'il convient de la différencier d fois afin de la stationnariser. Ainsi si $d = 1$, le fait de calculer les différences premières ΔY_t (et ainsi de remplacer chaque observation Y_t par $(Y_t - Y_{t-1})$) permet d'obtenir une série stationnaire (notée $\Delta Y_t \rightarrow I(0)$).

Deux séries X_t et Y_t sont dites cointégrées si les deux conditions suivantes sont vérifiées:

- les deux séries sont affectées d'une tendance stochastique de même ordre d'intégration d ;
- une combinaison linéaire de ces séries permet de se ramener à une série d'ordre d'intégration inférieur.

Ainsi avec $X_t \rightarrow I(d)$ et $Y_t \rightarrow I(d)$, on devrait avoir une combinaison linéaire pouvant s'écrire $\alpha_1 X_t + \alpha_2 Y_t \rightarrow I(d-b)$, avec $d \geq b \geq 0$, et $[\alpha_1 \quad \alpha_2]'$ étant le vecteur de cointégration. On peut aussi écrire $(X_t ; Y_t) \rightarrow CI(d,b)$.

La généralisation à $M(>2)$ séries cointégrées se fait selon exactement les mêmes principes (voir par exemple Mills, 1990).

Il est de plus en plus admis qu'en économie beaucoup de séries temporelles sont $I(1)$, et qu'avec les procédures d'estimation on arrive à construire des combinaisons linéaires $I(0)$, satisfaisant ainsi au critère de cointégration.

En pratique, le fait que l'on part de séries à priori non stationnaires fait que le modèle initial est estimé sur des variables transformées en différences, auxquelles on ajoute un terme en niveaux permettant de tenir compte d'une propriété d'équilibre de long terme suggérée par la théorie économique. Le terme en niveaux qui donne la mesure dans laquelle le système est éloigné de l'équilibre résulte d'une combinaison linéaire $I(0)$, les termes en différences étant eux-mêmes supposés $I(0)$, après qu'on a effectué les tests appropriés. Ici on a le lien entre le concept de cointégration et le modèle dit de correction d'erreur ("Error Correction Model", ECM), lien qui a été formellement établi par Engle et Granger (1987) sous le nom de théorème de la représentation, auquel on associe les noms de ces deux auteurs.

Nous présentons ci-après un exemple simple de relation de long terme pour illustrer le déroulement de la procédure d'estimation:

$$y_t = \widehat{\alpha}_0 + \widehat{\alpha}_1 x_t + \xi_t.$$

Dans le modèle ci-dessus, $[\widehat{\alpha}_0 + \widehat{\alpha}_1]'$ est le vecteur cointégrant. On admet que la variable \triangleright_t donne une mesure de l'erreur d'équilibre, et une façon de tester la cointégration des séries y_t et x_t consiste à supposer que \triangleright_t suit un processus autorégressif d'ordre un et à recourir au même test de racine unitaire qu'on avait utilisé précédemment pour tester la stationnarité:

$$\xi_t = \rho \xi_{t-1} + u_t$$

On teste la cointégration en prenant comme hypothèse nulle $H_0: \rho=1$ (dans ce cas le processus suit une marche aléatoire et il n'y a pas cointégration), contre l'hypothèse alternative de cointégration $H_1: |\rho| < 1$ (qui correspond en fait à la stationnarité des résidus).

Dès qu'après test on s'est assuré de la cointégration, c'est-à-dire de la convergence des trajectoires des séries, on utilisera, pour la relation finale à estimer, les valeurs de ξ_t obtenues grâce à l'estimation par moindres carrés des paramètres de la relation de

long terme entre y_t et x_t (voir ci-dessus). Engle et Granger ont proposé qu'on introduise alors les valeurs décalées ξ_{t-1} dans le modèle avec variables stationnarisées, afin d'apporter la composante ECM (correction d'erreur) au modèle final à estimer. Ce modèle peut être présenté de la manière suivante:

$$\Delta Y_t = \gamma \Delta X_t - \beta \xi_{t-1} + \omega^{xxiii}$$

Ce sont précisément les valeurs calculées des ξ_{t-1} qui traduisent les ajustements de court terme, permettant ainsi au modèle de converger vers la cible de long terme.^{xxiii}

Pour synthétiser, on peut dire que les modèles et les méthodes proposées dans la démarche décrite ci-dessus résultent du fait qu'il s'est avéré judicieux d'élaborer des modèles économétriques incorporant l'information que fournit la théorie économique au sujet des forces d'équilibre de long terme, et présentant une spécification acceptable des aspects dynamiques des différentes variables. Les évolutions des prix et des salaires, celles des revenus et des dépenses des ménages, ou encore celles des importations et des exportations, constituent des cas souvent cités comme illustrations de variables cointégrées (voir par exemple Kennedy, 1994).

5. Choix discrets et cas des variables dépendantes limitées

Les modèles considérés jusqu'ici expliquaient des variables économiques continues et portant généralement sur des quantités. On a vu que dans certains cas ces modèles pouvaient avoir des variables explicatives qualitatives (ou muettes), et que cela ne posait pas de problème particulier pour ce qui est de l'estimation (Balestra, 1980). Mais quand la variable dépendante est qualitative par nature, des problèmes particuliers relatifs à l'estimation surgissent. Il s'agit de situations où le preneur de décision (l'individu i) doit par exemple choisir entre deux alternatives: la variable dépendante, appelée aussi la réponse, notée y_i , est une

variable qualitative qui prend la valeur 1 si une des deux options est choisie et 0 si c'est l'autre option qui est choisie. Il faut alors recourir à de nouvelles techniques d'estimation qui sont adaptées à ces situations. C'est le cas quand il s'agit d'expliquer le choix d'un mode de transport permettant à un individu de se rendre à son travail (voiture privée ou service public), ou encore la possession ou non d'un certain équipement ou d'une certaine technologie, l'état d'être en chômage ou pas, etc. Les méthodologies d'estimation utilisées dans ces situations dites de choix discrets fournissent une réponse statistiquement élégante et conceptuellement attrayante aux types de problèmes évoqués ci-dessus. Et on peut ajouter que la théorie des choix discrets est l'un des rares domaines où l'économétrie a apporté quelque chose de positif à l'édification de la théorie économique (Balestra, 1994).

5.1. Le modèle économique

Considérons un preneur de décision qui est amené à choisir entre différentes alternatives, et qui dérive un certain montant d'utilité de chacun des résultats possibles: on admettra que l'individu se comporte de façon à maximiser son utilité, ce qui implique qu'il choisira l'alternative qui conduit à l'utilité la plus élevée.

En prenant l'exemple d'un individu i qui doit choisir d'aller à son travail en conduisant sa voiture (alternative 1) ou en utilisant les transports publics (alternative 0), nous définissons une variable économique discrète y_i qui traduit le résultat de l'observation faite sur le choix de l'individu:

$y_i = 1$ si pour aller à son travail, l'individu i choisit de conduire sa voiture;

$y_i = 0$ si pour aller à son travail, l'individu i choisit les transports publics.

Nous convenons ainsi que l'alternative 1 sera choisie si le fait de conduire pour aller à son travail procure à l'individu plus d'utilité que s'il utilisait les transports publics.

On peut donc encore écrire:

$$y_i = 1 \text{ si } U_{i1} \geq U_{i0}; \\ = 0 \text{ si } U_{i0} > U_{i1}.$$

En tant qu'économistes nous cherchons à comprendre, expliquer, et prévoir les choix qui sont faits. Pour cela nous créerons un modèle économique de l'utilité dérivée du choix de chacune des alternatives. D'une manière générale, nous pourrions penser à l'utilité dérivée du choix d'une alternative j (avec $j=1$ ou 0) par un individu i (avec $i: 1, \dots, T$) comme étant fonction (a) des attributs (coût, convenance, qualité, etc.) de cette alternative pour l'individu, et (b) des caractéristiques (revenu, niveau d'éducation, profession, etc.) de l'individu. Pour illustrer de manière simplifiée la question des attributs associés aux alternatives, nous conviendrons que l'utilité qu'un individu dérive du fait de conduire sa propre voiture dépend du temps que lui prend ce moyen de locomotion. Si chaque mode de transport requiert le même temps de déplacement, alors le facteur temps seul n'est pas pertinent pour décider du choix. Nous allons travailler avec un modèle utilisant la différence de temps mis par les deux modes de transport. Soit la variable suivante définissant une différence de temps:

$$\zeta_i = (\text{temps de déplacement de l'individu } i \text{ par les transports publics}) - (\text{temps de déplacement de l'individu } i \text{ par sa propre voiture}).$$

Si $\zeta_i > 0$, l'utilisation des transports publics prend plus de temps que conduire et, toutes choses étant égales par ailleurs, le coût d'opportunité du recours aux transports publics est plus élevé que le coût d'opportunité associé au fait de conduire soi-même. Plus ζ_i est élevé, plus élevé est le coût relatif de l'utilisation des transports publics, ce qui augmente les chances que l'individu se serve de sa propre voiture.

Il faudra toutefois garder à l'esprit le fait que d'autres facteurs économiques affectent le choix du mode de transport: prix du billet pour les transports publics, prix de l'essence, prix des places de parking, etc. Nous commencerons avec un modèle statistique simple ayant seulement ζ_i comme variable explicative.

5.2. Le modèle statistique

Dans ce modèle nous avons une variable dépendante observable y_i qui traduit les choix des individus. Il s'agit d'une variable aléatoire discrète. Le résultat du choix est aléatoire, puisqu'on ne peut pas prédire avec certitude le choix que fera un individu tiré au hasard. Les individus font leur choix en fonction de leurs propres caractéristiques, observables ou non, et en se référant aux alternatives qui leur sont ouvertes. Pour l'individu i on aura ainsi une fonction de densité discrète pour y_i , et on s'en servira pour construire le modèle statistique des choix.

Appelons P_i la probabilité que l'individu i choisisse l'alternative 1 (conduire) et $(1-P_i)$ la probabilité que ce soit l'alternative 0 (transports publics) qui soit choisie. La fonction des densités de probabilités s'écrira alors de la manière suivante:

$$g(y_i) = P_i^{y_i} (1-P_i)^{1-y_i}, \quad y_i = 1, 0.$$

Cette fonction nous donne les expressions associées aux probabilités des deux alternatives:

$$g(1) = P[y_i = 1] = P_i = \text{probabilité de conduire};$$

$$g(0) = P[y_i = 0] = 1-P_i = \text{probabilité d'utiliser les transports publics}.$$

On peut montrer que l'espérance mathématique et la variance de la variable aléatoire discrète y_i sont respectivement données par $E[y_i] = P_i$ et $\text{Var}[y_i] = P_i(1-P_i)$.



Pour expliquer les variations de la probabilité de choix P_i , ce qui est notre but, nous devons mettre cette probabilité en relation avec les attributs des alternatives présentées à l'individu et avec les caractéristiques de l'individu. A cette fin nous considérerons trois modèles: le modèle linéaire des probabilités, le modèle probit, et le modèle logit.

5.2.1. Le modèle linéaire des probabilités

Nous considérons le problème de la modélisation de la variable aléatoire y_i et de ses variations comme un problème de régression. Spécifiquement, on admet que y_i a une composante systématique, $E[y_i] = P_i$, et une composante aléatoire e_i , à espérance mathématique nulle:

$$y_i = E[y_i] + e_i$$

Comme on l'a fait avec d'autres modèles, on admet que $E[y_i]$ est fonction linéaire (dans les paramètres) d'un ensemble de variables explicatives, y compris un terme constant:

$$E[y_i] = P_i = \beta_1 + \beta_2 \zeta_{i2} + \dots + \beta_k \zeta_{ik}$$

Le modèle pour y_i aura ainsi la formulation stochastique suivante, appelée modèle linéaire des probabilités:

$$y_i = \beta_1 + \beta_2 \zeta_{i2} + \dots + \beta_k \zeta_{ik} + e_i$$

Comme $e_i = y_i - P_i$, on aura $E[e_i] = 0$ et $\text{Var}[e_i] = P_i(1-P_i) = \delta_i^2$. On se trouve ainsi face à un cas d'hétéroscédasticité. On pourra estimer ce modèle par une méthode de moindres carrés généralisés en deux étapes. Un des principaux problèmes associés à ce modèle est que rien ne permet de garantir que les probabilités calculées resteront bien dans l'intervalle $[0,1]$.

Reprenons l'exemple où on avait une seule variable explicative ζ_i , représentant la différence de temps entre deux façons d'aller à son travail:

$$\zeta_i = \text{temps (transports publics)} - \text{temps (voiture privée)}; E[y_i] = P_i = \beta_1 + \beta_2 \zeta_i.$$

Si ζ_i augmente, on s'attendra à ce que la probabilité pour l'individu de préférer sa voiture privée augmente. On s'attendrait donc a priori à ce que β_2 soit positif. Le paramètre β_1 représente la probabilité que l'individu utilise sa voiture quand il n'y a pas de différence de temps entre les deux modes de transport (cas où ζ_i est nul). On devrait alors normalement avoir $\beta_1 \in [0,1]$. Ce serait comme un paramètre représentant le facteur "goût". Le fait que les

P_i modélisés peuvent tomber à l'extérieur de l'intervalle $[0,1]$ fait que le modèle linéaire des probabilités n'est généralement pas recommandé en pratique: pour des valeurs de ζ_i relativement proches de la médiane de la distribution ça se passe généralement bien, mais si pour calculer des prévisions on prend des valeurs de ζ_i très élevées ou très petites, on se soumet un peu plus au risque d'avoir des P_i à l'extérieur de l'intervalle $[0,1]$.

5.2.2. Le modèle probit

Le problème avec le modèle précédent c'est que sa linéarité (dans les paramètres) comporte le risque d'aboutir à des probabilités négatives ou supérieures à 1. Pour la construction d'un modèle probit, on définit tout d'abord un "indice d'utilité" μ_i pour l'individu i , de la manière suivante:

$$\begin{aligned} \mu_i &= \beta_1 + \beta_2 \zeta_{i2} + \dots + \beta_k \zeta_{ik} \\ &= \zeta_i' \beta, \end{aligned}$$

avec $\zeta_i' = (1, \zeta_{i1}, \dots, \zeta_{ik})$, et $\beta = (\beta_1, \beta_2, \dots, \beta_k)'$

L'idée avec l'indice μ_i qui varie entre $-\infty$ et $+\infty$ (sa valeur est en effet donnée par $\zeta_i' \beta$), sera de trouver une fonction reliant cet indice et P_i , étant entendu que toutes les valeurs estimées de P_i doivent être comprises dans l'intervalle $[0,1]$.

Avec ce modèle, plus la valeur de μ_i est élevée, plus grande est l'utilité que l'individu tire du choix de l'option $y_i = 1$. En effet le modèle est tel que l'alternative 1 est prise comme référence, et tous les ζ_i sont déterminés par rapport à l'alternative prise comme référence (pour cela les variables sont exprimées en différences par rapport à l'alternative de référence, comme dans l'exemple sur les modes de transport, mais ça pourrait aussi être des rapports). Ainsi, plus μ_i est élevé, et plus grande sera la probabilité P_i que l'individu choisisse l'option où $y_i = 1$.

L'étape suivante consiste à choisir une fonction strictement croissante reliant P_i et μ_i et montrant comment P_i varie entre 0 et 1 pendant que μ_i varie entre $-\infty$ et $+\infty$.

Reprenons l'exemple dans lequel l'individu i doit choisir entre conduire sa voiture et utiliser les transports publics pour aller à son travail. On notera

$$\mu_i = \beta_1 + \beta_2 \zeta_i$$

Avec $\beta_2 > 0$, quand ζ_i augmente, l'indice μ_i augmente aussi. Et plus élevée est la valeur de μ_i , plus grande sera la probabilité que l'individu choisisse de conduire. Dans ce cas-ci on peut interpréter l'indice d'utilité comme la "propension" de l'individu pour conduire.

Revenons maintenant à notre problème qui consiste à modéliser P_i de façon à ne pas avoir affaire à des probabilités négatives ou supérieures à 1. Le modèle probit constitue précisément un choix possible de transformation monotone croissante de μ_i en une probabilité (P_i). Il s'exprime de la manière suivante:

$$P_i = F(\mu_i) = F(\beta_1 + \beta_2 \zeta_{i2} + \dots + \beta_k \zeta_{iK}) = F(\zeta_i' \beta)^{\text{xxiv}},$$

et $F(\mu_i)$ aura la spécification d'une fonction de distribution cumulative associée à une variable aléatoire normale standard $N(0,1)$, évaluée au point μ_i . Cette fonction est donnée par l'intégrale suivante, ayant $-\infty$ et μ_i comme bornes:

$$P_i = F(\mu_i) = P\{z \leq \mu_i\} = \int_{-\infty}^{\mu_i} (2\pi)^{-1/2} \exp(-z^2/2) dz,$$

z étant une variable aléatoire normale standard. Avec ce modèle, les probabilités P_i sont toutes dans l'intervalle $[0,1]$; elles augmentent monotoniquement au fur et à mesure que l'indice d'utilité μ_i s'accroît de $-\infty$ à $+\infty$. On peut montrer que, comme il s'agit d'une loi normale standard, le maximum de la fonction de densité $f(\xi_i' \beta)$ est atteint pour $P_i = F(\mu_i = 0) = 0.5$.^{xxx}

Estimation des paramètres du modèle probit

Avec des observations collectées sur T individus, on utilisera la méthode du maximum de vraisemblance pour estimer β , car la variable y est discrète, et la relation fonctionnelle entre P_i et les variables explicatives ξ_{ik} est non linéaire.

Pour construire la fonction de vraisemblance on tiendra compte du fait que la fonction de densités de probabilités conjointe de l'échantillon de T observations indépendantes est le produit des T fonctions de densités de probabilités $g(y_i)$. Cela donne l'expression suivante:

$$\begin{aligned} g(y_1, y_2, \dots, y_T) &= \prod_{i=1, T} g(y_i) \\ &= \prod_{i=1, T} P_i^{y_i} (1-P_i)^{1-y_i} \\ &= \prod_{i=1, T} [F(\xi_i' \beta)]^{y_i} [1 - F(\xi_i' \beta)]^{1-y_i} \end{aligned}$$

Si les paramètres β étaient connus, cette densité conjointe pourrait être utilisée pour calculer la probabilité qu'un quelconque ensemble de résultats de choix se produise. La fonction de vraisemblance conjointe ci-dessus servira pour estimer les composantes du vecteur β de manière à maximiser la probabilité d'obtenir l'échantillon qui est effectivement observé. Pour bien montrer que cette vraisemblance dépend de β , on peut écrire

$$L(\beta) = \prod_{i=1, T} [F(\xi_i' \beta)]^{y_i} [1 - F(\xi_i' \beta)]^{1-y_i}$$

et dans cette expression, $F(\xi_i' \beta)$ est la fonction de distribution normale standard cumulative évaluée au point $\xi_i' \beta$. On transforme généralement $L(\beta)$ en vraisemblance logarithmique avant la procédure d'estimation.

Actuellement il existe des logiciels économétriques très performants pour l'optimisation numérique de fonctions de vraisemblance telles que celle présentée ci-dessus. Dès qu'on a des estimateurs pour les paramètres β , on peut calculer un estimateur de P_i pour toute valeur de ξ_i , en utilisant la fonction normale cumulative standard et en calculant

$$P_i^{\text{est}} = F(\xi_i' \beta^{\text{est}})$$

Si l'estimation obtenue pour P_i est plus élevée que 0.5, on en déduit que l'individu choisira l'alternative 1 (dans l'exemple avec les modes de transport cela voudrait dire que l'individu i préférera conduire, de sorte que dans le vecteur y on aura, pour cet individu, $y_i = 1$). On peut aussi calculer la variation de P_i qui résulte d'un accroissement unitaire dans la valeur d'une des variables explicatives (par exemple une augmentation d'une minute dans le temps que mettent les transports publics pour conduire l'individu i à son lieu de travail).^{xxvi}

5.2.3. Le modèle logit

Dans le cas du modèle logit, les probabilités de choix $P_i = F(\xi_i' \beta) = F(\mu_i)$ résulteront d'une fonction de distribution cumulative $F(\cdot)$ d'une variable aléatoire logistique, de sorte qu'on aura

$$\begin{aligned} P_i = F(\xi_i' \beta) &= 1 / (1 + \exp(-\xi_i' \beta)) \\ &= \exp(\xi_i' \beta) / (1 + \exp(\xi_i' \beta)) \end{aligned}$$

On remarquera que cette fois la fonction de distribution cumulative n'a pas la forme d'une intégrale. Cela fait que le logit est plus aisé à manier sur le plan des calculs. La probabilité que l'individu i choisisse l'alternative 0 plutôt que l'alternative 1 est donnée par

$$1 - P_i = 1 / (1 + \exp(\zeta_i' \beta))$$

On remarquera que les expressions pour P_i et $(1 - P_i)$ sont telles que

$$P_i / (1 - P_i) = \exp(\zeta_i' \beta)$$

$$\text{Log} [P_i / (1 - P_i)] = \zeta_i' \beta$$

On peut aussi montrer, par dérivation de $F(\zeta_i' \beta)$, que la fonction de densité logistique est égale à

$$\begin{aligned} f(\zeta_i' \beta) &= \exp(\zeta_i' \beta) / (1 + \exp(\zeta_i' \beta))^2 \\ &= F(\zeta_i' \beta) [1 - F(\zeta_i' \beta)] \end{aligned}$$

Cette fonction de densité est lisse, symétrique autour de zéro, avec une forme en cloche, comme la fonction de densité normale standard, mais se distingue par des extrémités un peu plus étendues.

Les formules de calcul de P_i et de $(1 - P_i)$ obtenues ci-dessus permettent de faire un groupement tel que la fonction de vraisemblance aura la forme suivante:

$$L(\beta) = \kappa_{i=1, n_1} [\exp(\zeta_i' \beta) / (1 + \exp(\zeta_i' \beta))] \kappa_{i=n_1+1, T} [1 / (1 + \exp(\zeta_i' \beta))],$$

où l'indice $\{i= 1, n_1\}$ se réfère à ceux qui ont choisi l'alternative 1, tandis que l'indice $\{i= n_1+1, T\}$ se réfère à ceux qui ont choisi l'alternative 0. Par passage aux logarithmes, on obtient l'expression suivante:

$$\begin{aligned}
\text{Log } L(\beta) &= \varphi_{i=1,n1} [\text{Log} [\exp(\zeta_i' \beta) / (1 + \exp(\zeta_i' \beta))]] + \\
&\quad \varphi_{i=n1+1,T} [\text{Log} [1 / (1 + \exp(\zeta_i' \beta))]] \\
&= \varphi_{i=1,n1} [\zeta_i' \beta - \text{Log} (1 + \exp(\zeta_i' \beta))] - \varphi_{i=n1+1,T} \text{Log} [1 + \\
&\quad \exp(\zeta_i' \beta)] \\
&= \varphi_{i=1,n1} \zeta_i' \beta - \varphi_{i=1,T} \text{Log} [1 + \exp(\zeta_i' \beta)]
\end{aligned}$$

On utilise ensuite des programmes d'ordinateur spécifiques pour calculer par maximisation de la vraisemblance l'estimateur de β dans ce modèle logit pour choix discrets (tels que le choix entre conduire et utiliser les transports publics par exemple). L'estimateur obtenu pour β permettra de calculer la probabilité P_i et d'estimer d'autres paramètres d'intérêt pour ce modèle, dans la même optique qu'avec le modèle probit.

5.2.4. Extensions

Une extension de ces modèles de choix discrets consiste à considérer les cas où on a plus que deux alternatives. Par exemple, une personne qui se rend à son travail peut choisir entre les trois alternatives suivantes: conduire sa propre voiture, avoir recours aux autobus publics, ou prendre le train. Les modèles probit et logit peuvent être généralisés pour permettre la modélisation des situations de choix multinomiaux (ou polytomiques).

□ *Logit multinomial*

On arrive à une généralisation assez pratique du logit binaire en partant de l'expression obtenue pour le ratio $P_i / (1 - P_i)$ ci-dessus et en adoptant le principe que le rapport de la probabilité de choisir la $k^{\text{ème}}$ alternative sur la probabilité de choisir une alternative de base donnée est $\exp(\zeta_i' \beta_k)$, où β_k est un vecteur de paramètres caractérisant la $k^{\text{ème}}$ alternative. Pour illustrer cette généralisation, considérons le cas où l'on a trois alternatives (possibilités de transport pour se rendre à son poste de travail) pour l'individu i :

- A = conduire sa propre voiture;
- B = avoir recours aux autobus publics;
- C = prendre le train.

En prenant C comme alternative de base, on aura ainsi^{xxvii}

$$\text{Prob}(A)/\text{Prob}(C) = \exp(\zeta_i' \beta_A)$$

$$\text{Prob}(B)/\text{Prob}(C) = \exp(\zeta_i' \beta_B)$$

Le traitement de ce cas à trois options est tel qu'il n'y a que deux ratios à déterminer, le troisième pouvant se dériver des deux autres. En utilisant les valeurs des deux ratios définis ci-dessus et le fait que $\text{Prob}(A) + \text{Prob}(B) + \text{Prob}(C) = 1$, on arrive facilement à montrer que

$$\text{Prob}(A) = \exp(\zeta_i' \beta_A) / (1 + \exp(\zeta_i' \beta_A) + \exp(\zeta_i' \beta_B))$$

$$\text{Prob}(B) = \exp(\zeta_i' \beta_B) / (1 + \exp(\zeta_i' \beta_A) + \exp(\zeta_i' \beta_B))$$

$$\text{Prob}(C) = 1 / (1 + \exp(\zeta_i' \beta_A) + \exp(\zeta_i' \beta_B))$$

Nous obtenons ainsi l'expression suivante pour la fonction de vraisemblance, où nous désignons par $D_i(\beta_A, \beta_B)$ l'expression $(1 + \exp(\zeta_i' \beta_A) + \exp(\zeta_i' \beta_B))$:

$$L(\beta_A, \beta_B) = \kappa_{i=1, n_A} [\exp(\zeta_i' \beta_A) / D_i(\beta_A, \beta_B)] \kappa_{i=n_A+1, n_B} [\exp(\zeta_i' \beta_A) / D_i(\beta_A, \beta_B)] \kappa_{i=n_B+1, T} [1 / D_i(\beta_A, \beta_B)],$$

où l'indice $\{i= 1, n_A\}$ se réfère à ceux qui ont choisi l'alternative A, l'indice $\{i= n_A+1, n_B\}$ se réfère à ceux qui ont choisi l'alternative B, tandis que l'indice $\{i= n_B+1, T\}$ se réfère à ceux qui ont choisi l'alternative C.

La maximisation de cette fonction de vraisemblance par rapport à β_A et β_B permet d'obtenir des estimateurs pour ces vecteurs de paramètres. Avec les estimations obtenues (β_A^{est} , β_B^{est}), on pourra ensuite utiliser les caractéristiques de chaque individu particulier pour estimer $\text{Prob}(A)$, $\text{Prob}(B)$, et $\text{Prob}(C)$, qui représentent les probabilités que l'individu choisisse l'une ou l'autre des trois alternatives. L'extension de cette procédure logit à plus de trois alternatives est immédiate. On peut aussi recourir au modèle probit multinomial, dont l'hypothèse de base est que les termes d'erreur de la fonction d'utilité stochastique suivent une distribution normale multivariée.^{xxviii}

Probit ordonné

Pour certaines variables polytomiques, il existe un ordre naturel entre les options. Avec l'évaluation des obligations par exemple, on peut avoir un vecteur y où l'on distingue des catégories d'obligations étiquetées A, double A, triple A, etc., selon le niveau de crédit reconnu à ces obligations. De même, des enseignants peuvent classer leurs étudiants en catégories avec mention A, B, C, etc., selon leur niveau de compréhension des enseignements pris en considération. Face à des situations de ce type, utiliser un modèle probit multinomial de type général serait inefficace car ne prenant pas en compte l'information additionnelle implicite portant sur la nature ordinale de la variable dépendante.

Pour ces cas-là on recourt souvent à un modèle du type probit ordonné, que nous exposons ci-après à travers un exemple. Supposons que y^* est un indice non observable du niveau de crédit attribué à différentes catégories d'observations, et que cet indice peut être expliqué à travers le modèle linéaire suivant:

$$y^* = \alpha + \beta \zeta + \varepsilon$$

On admet que les observations effectives sur les obligations peuvent être classées dans les catégories suivantes:

On observe	$y = B$	si $y^* < \hat{\delta}_1$.
	$y = A$	si $\hat{\delta}_1 \leq y^* \leq \hat{\delta}_2$.
	$y = AA$	si $\hat{\delta}_2 \leq y^* \leq \hat{\delta}_3$.
	$y = AAA$	$\hat{\delta}_3 \leq y^*$.

Dans une même procédure d'estimation par le maximum de vraisemblance, on estimera les paramètres α et β , mais aussi les $\hat{\delta}$, qui donnent les valeurs des bornes inconnues définissant les séparations entre les différentes catégories ordonnées de la variable y . Prenons comme exemple le cas où l'on veut estimer la probabilité d'obtenir une observation avec $y=AA$. Cela peut s'écrire de la manière suivante:

$$\text{Prob}(\hat{\delta}_2 \leq y^* = \alpha + \beta \zeta + \varepsilon \leq \hat{\delta}_3) = \text{Prob}(\hat{\delta}_2 - \alpha - \beta \zeta \leq \varepsilon \leq \hat{\delta}_3 - \alpha - \beta \zeta)$$

A partir de l'expression ci-dessus on construira, dès que la fonction de densité de ε est connue, une fonction de vraisemblance qui servira pour l'estimation. Le modèle probit ordonné résulte de l'hypothèse que ε est distribué selon une loi normale. On peut aussi postuler un modèle logit ordonné, qui résulterait alors de l'hypothèse que la densité cumulative de ε est une fonction logistique. Il semble que dans la pratique les deux formulations donnent des résultats assez proches.^{xxxix}

Variables dépendantes limitées

Assez proches des modèles de choix discrets on a aussi le cas où les variables dépendantes sont continues, mais limitées ou contraintes d'une façon ou d'une autre. Par exemple, les données sur les expériences d'impôt négatif sur le revenu sont telles que le revenu est en dessous d'un certain seuil pour toutes les

observations. De même, beaucoup d'observations sur les dépenses des ménages en achats d'automobiles sont à zéro, ceci correspondant aux ménages qui ne peuvent pas ou qui choisissent de ne pas acheter d'automobile. Dans le même ordre d'idées, si on veut construire un modèle pour expliquer le nombre de fois qu'un individu va à la pêche durant une année, on s'attend de toute évidence à ce que les réponses données par un échantillon d'individus varient entre zéro et plusieurs centaines.

Les modèles expliquant des variables soumises à la contrainte de non-négativité ont été analysés pour la première fois de manière approfondie par l'économiste américain James Tobin, et c'est ainsi que le modèle statistique qui a résulté de ses travaux a été appelé modèle tobit. On l'appelle aussi parfois modèle de régression censuré. De façon plus générale, les modèles dits censurés sont, par définition, des modèles pour lesquels les variables explicatives (par exemple le revenu des ménages) restent observables, même quand la variable dépendante (par exemple le montant qu'un ménage serait prêt à dépenser pour l'achat d'une voiture) devient inobservable. On parle de modèles tronqués quand l'information manque aussi sur les variables explicatives lorsque l'on est dans une zone où la variable dépendante n'est pas observable, contrairement au cas des modèles censurés.

Dans le cas des modèles étudiés par Tobin, si l'on considère l'exemple où un ménage achète (indice 1) ou n'achète pas (indice 0) une voiture, le critère de sélection des observations pour le vecteur y consiste à dire que y_1 est observé si $y_1 \geq 0$ ou, de manière équivalente, que l'utilité associée au choix d'acheter est supérieure à une constante donnée ($U(y_1) > K$). La méthode d'estimation devra alors tenir compte de l'implication de cette contrainte sur l'espérance mathématique de y_1 et sur celle du terme stochastique du modèle. Avec un modèle où ζ est le vecteur de variables explicatives et ε un terme d'erreur additif, on peut écrire:

$$E(y_1 \mid \zeta_{1i}, y_1 \geq 0) = \zeta_{1i}' \beta_1 + E(\varepsilon_1 \mid y_1 \geq 0)$$

Les paramètres de tels modèles sont généralement estimés par des méthodes itératives de maximisation d'une fonction de vraisemblance conditionnelle.^{xxx}

5.2.5. *Modèles de choix discrets et mesure de la qualité de l'ajustement*

Pour les tests sur le niveau de signification des coefficients, on part de l'hypothèse d'un nombre T d'individus suffisamment élevé, ce qui permet d'assurer que β^{est} aura une distribution approximativement normale avec une espérance mathématique égale à β et une matrice de variances-covariances $\Sigma_{\beta}^{\text{est}} = (\zeta' V \zeta)^{-1}$, expression dans laquelle V est une matrice diagonale ayant sur sa diagonale principale les éléments v_i définis de la manière suivante:

$$v_i = [f(\zeta_i' \beta)]^2 / [F(\zeta_i' \beta)(1 - F(\zeta_i' \beta))]$$

Ce sont les racines carrées des éléments diagonaux de la matrice $\Sigma_{\beta}^{\text{est}}$ qui fournissent des estimateurs pour les écarts-types des coefficients du modèle, et de là on calcule des statistiques t asymptotiques en divisant chaque coefficient estimé par l'écart-type correspondant.

Pour ce qui est de l'évaluation globale de la qualité de l'ajustement, il est important de faire remarquer que les coefficients de détermination R^2 calculés pour des modèles de choix discrets auront généralement des valeurs très basses, ce qui peut se comprendre compte tenu de la nature des problèmes abordés et de la modélisation à laquelle on fait usuellement recours. C'est pour cette raison qu'il est préférable d'utiliser des procédures tenant compte des spécificités de ce type de modèles.^{xxxi}

Nous présentons ci-après trois critères parmi les plus couramment utilisés pour valider les modèles de choix discrets et en mesurer globalement la qualité prédictive.

□ *"First preference recovery"*

C'est le pourcentage des individus qui choisissent réellement l'alternative à laquelle le modèle attribue la plus grande probabilité. Cette mesure peut être comparée à la proportion de personnes qui choisiraient l'alternative à laquelle le modèle attribue la probabilité $1/J$, si J est le nombre total d'alternatives.

□ *Part prédite et part réelle de chaque alternative*

Ici il s'agira de comparer la part réelle de chaque alternative j dans l'échantillon, avec la part prédite par le modèle, qu'on peut noter $\sum_{i=1, J} P_{ij}^{est}/T$, où T est le nombre total d'individus. Cette comparaison permet d'évaluer différentes spécifications du modèle. Dans le cas dichotomique, une façon d'appliquer ce critère consiste à prendre $y=1$ pour tous les cas où $P_i^{est} > 0.50$ et de compter le nombre de $y=1$ prédits correctement et le nombre de $y=0$ prédits correctement.

□ *Indice du ratio de vraisemblance (IRV)*

Cet indice est défini par la mesure $IRV = 1 - L(\beta^{est})/L(\beta_H^{est})$. Dans cette expression, $L(\beta^{est})$ est la vraisemblance logarithmique obtenue dans la maximisation de la fonction de vraisemblance (modèle non-contraint), et $L(\beta_H^{est})$ est la vraisemblance logarithmique du modèle contraint par l'hypothèse nulle (valeur de la fonction de vraisemblance logarithmique quand tous les paramètres sont égaux à zéro). Par construction, si la procédure de maximisation suggère qu'on ne gagne rien à avoir un quelconque paramètre différent de zéro (cas où $L(\beta^{est})$ deviendrait égal à $L(\beta_H^{est})$), alors IRV serait égal à zéro. Par contre, si on était suffisamment chanceux pour avoir une fonction de vraisemblance qui prédit chaque choix dans l'échantillon de manière correcte, alors la fonction de vraisemblance estimée serait égale à 1, et son logarithme égal à zéro. On aurait alors $L(\beta^{est}) = 0$, et il en

résulterait $IRV = 1$. Ainsi l'intervalle de variation de IRV va de 0 à 1. On retiendra que la valeur de IRV donne une indication sur ce qu'on gagne à ajouter une ou plusieurs nouvelles variables explicatives au modèle, mais une valeur numérique particulière de IRV est difficile à interpréter.

Pour en revenir à l'interprétation économique de la solution apportée par les modèles pour choix discrets, on peut dire que cette solution est particulièrement éclairante grâce au concept d'utilité stochastique, cher aux économistes. On suppose dans cette conceptualisation l'existence d'une fonction d'utilité indirecte dont les arguments sont les caractéristiques des individus et les attributs des différentes alternatives. Une alternative est choisie dès lors que l'utilité qu'elle procure est supérieure à celle des autres alternatives. S'agissant d'une utilité stochastique, on peut alors facilement calculer la probabilité que l'individu fasse tel ou tel choix. Ces probabilités dépendront des arguments de la fonction d'utilité, dont le poids est estimé par les méthodes probit, logit, etc.

6. Économétrie bayésienne^{xxxii}

6.1. Introduction

Il y a dans une population des paramètres inconnus sur lesquels on veut faire une induction. Deux voies principales existent en matière d'induction: l'induction classique et l'induction bayésienne. L'induction classique est caractérisée par l'expérience comme seule source d'information. L'expérience est contenue dans un échantillon sur la base duquel on induit des paramètres inconnus de la population. L'induction bayésienne part du principe qu'on peut distinguer deux sources d'information: une première qu'on pourrait appeler la connaissance a priori des phénomènes, et une deuxième qui est contenue dans les faits, l'expérience de l'échantillon. Le théorème de Bayes permet de mélanger ces deux sources et d'en tirer une connaissance qu'on appellera connaissance a posteriori.

La critique adressée par l'approche bayésienne à l'induction classique est que cette dernière a des procédures qui sont justifiées en termes de distributions d'échantillonnage: on admet que les estimateurs obtenus se caractérisent par certaines propriétés sous l'hypothèse de reproductibilité de l'expérience d'échantillonnage, alors que les tirages répétés n'existent généralement pas dans les études d'analyse économique (voir aussi Butare, 2002). On dispose le plus souvent d'un échantillon, avec lequel il faut se débrouiller le mieux possible. De plus, le coût des conséquences négatives qui pourraient résulter de conclusions incorrectes n'est généralement pas évalué dans le cadre des méthodes d'induction classique. L'approche bayésienne cherche notamment à répondre aux trois questions suivantes: (1) Pouvons-nous exprimer l'incertitude portant sur les paramètres ou les hypothèses d'un modèle avant et après qu'un échantillon a été tiré? (2) Comment pouvons-nous combiner l'information d'échantillon avec l'information a priori, qui peut être une information ne résultant pas d'un échantillon, ou alors une information issue d'expériences précédemment faites sur d'autres échantillons? (3) De quelles façons analyserions-nous les pertes résultant de décisions incorrectes consécutives à des erreurs d'estimation sur des données d'échantillon?

Les trois principales étapes de l'analyse en approche bayésienne sont les suivantes: (i) une description probabiliste des paramètres est faite avant qu'un échantillon ne soit tiré pour l'analyse; (ii) la vraisemblance associée à un échantillon spécifique tiré en vue de l'estimation est ensuite présentée; (iii) la description de l'incertitude caractérisant la valeur des paramètres après l'analyse d'échantillon - ou ce que l'on appelle aussi la fonction de densité des probabilités a posteriori -, constitue la troisième étape du travail empirique de type bayésien.

Différentes situations en économie illustrent bien l'existence de certains principes qui correspondent à une information aprioristique implicite. On sait par exemple qu'avec une équation de demande, la réaction de la quantité demandée à une variation des prix sera négative; de même, on s'attendra à ce qu'il y ait une relation négative entre l'investissement et le niveau des taux d'intérêt, etc.

6.2. Fonctions de densité et espérances mathématiques conditionnelles

Soient X et Y deux variables aléatoires. La fonction notée $f_{X,Y}(x,y)$ est la fonction de densité jointe, et elle vérifie notamment les deux propriétés suivantes:

$$f \geq 0, \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = 1$$

La fonction de densité marginale pour X est notée $f_X(x)$ et elle est définie telle que

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy.$$

On appelle fonction de densité conditionnelle $f_{X|Y}(x)$ le rapport défini par la fonction de densité jointe $f_{X,Y}(x,y)$ sur la fonction marginale $f_Y(y)$:

$$f_{X|Y}(x) = f_{X,Y}(x,y) / f_Y(y)$$

Pour les espérances mathématiques conditionnelles on a les définitions suivantes:

$$E(X | Y) = \int_{-\infty}^{\infty} x f_{X|Y}(x) dx$$

$$E(g(X) | Y) = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x) dx$$

L'espérance mathématique, si elle existe, est un nombre. S'il y a indépendance entre X et Y , on aura forcément $E(X | Y) = E(X)$. On peut aussi démontrer, en exploitant les définitions données ci-dessus, les propriétés suivantes:

$$E(E(X | Y)) = E(X).$$

$$E(Xg(Y) | Y) = g(Y) E(X | Y)$$

$$E((X_1+X_2) | Y) = E(X_1 | Y) + E(X_2 | Y)$$

La définition de la variance conditionnelle se fonde sur celle d'espérance mathématique conditionnelle. Elle est donnée par l'expression suivante:

$$V(X | Y) = \int_{-\infty}^{\infty} [x - E(X | Y)]^2 f_{X|Y}(x) dx$$

On peut démontrer, en utilisant les définitions, que

$$E\{V(X | Y)\} + V\{E(X | Y)\} = V(X).$$

6.3. Caractéristiques générales de l'analyse bayésienne

C'est de l'application du théorème de Bayes que se dérivent les résultats de l'analyse bayésienne en statistique et en économétrie. En partant des définitions des probabilités conditionnelles $f_{X|Y}(x) = f_{X,Y}(x,y) / f_Y(y)$ et $f_{Y|X}(y) = f_{X,Y}(x,y) / f_X(x)$, le théorème de Bayes nous dit que

$$f_{Y|X}(y) = [f_{X|Y}(x) \cdot f_Y(y)] / f_X(x)$$

Ainsi, si $f_{Y|\theta}(y)$ représente la fonction de densité issue de l'information d'échantillon en vue de l'estimation d'un modèle avec un vecteur de paramètres θ et une variable dépendante notée y , $f_{\theta}(\theta)$ étant la fonction de densité a priori pour les paramètres, le théorème de Bayes nous permet d'écrire

$$f_{\theta|Y}(\theta) = [f_{\theta}(\theta) \cdot f_{Y|\theta}(y)] / f_Y(y)$$

Dans la formule ci-dessus, on admet généralement que le dénominateur se réduit à une constante, sachant que y est un vecteur d'observations certaines et en supposant que les paramètres de la fonction de densité de y sont des constantes données. C'est ainsi que dans cette formule exprimant la fonction de densité a posteriori de θ , le dénominateur $f_Y(y)$ disparaît, et le signe d'égalité est remplacé par un signe de proportionnalité (qui signifie que les quantités comparées sont égales, à une constante multiplicative près):

$$f_{\theta|Y}(\theta) \propto f_{\theta}(\theta) \cdot f_{Y|\theta}(y)$$

Le passage de $f_{\theta}(\theta)$ à $f_{\theta|Y}(\theta)$ montre le changement qui s'est opéré dans les croyances du chercheur, relatives à θ , en raison de l'information apportée par l'échantillon.

Pour montrer la façon dont on combine l'information a priori et l'information d'échantillon, considérons le cas où l'on dispose de n observations indépendantes $y = \{y_i\}_{i=1, \dots, n}$ issues d'une même population normale d'espérance mathématique μ inconnue et de variance δ^2 connue. Afin de caractériser l'information a priori que l'on a sur le paramètre μ , on admet que μ est normal d'espérance mathématique μ_0 et de variance δ_0^2 .

Dans cet exemple, la fonction de vraisemblance normale de l'échantillon peut s'écrire de la manière suivante:

$$f_{Y|\mu}(y) = [2\pi\delta^2]^{-n/2} \exp\left\{-1/2\delta^2 \sum_{i=1, \dots, n} (y_i - \mu)^2\right\}$$

Quant à la fonction de densité normale univariée associée à la distribution a priori de μ , on peut l'écrire de la manière suivante:

$$f_{\mu}(\mu) = [2\pi\delta_0^2]^{-1/2} \exp\left\{-1/2\delta_0^2 (\mu - \mu_0)^2\right\}$$

La fonction de densité a posteriori peut être exprimée de la manière suivante:

$$f_{\mu|Y}(\mu) \propto f_{\mu}(\mu) \cdot f_{Y|\mu}(y)$$

En définissant $\mu^{est} = (1/n) \sum_{i=1, \dots, n} y_i$ (c'est-à-dire la moyenne d'échantillon), de simples manipulations algébriques permettent d'arriver à la fonction de densité a posteriori suivante:^{xxxiii}

$$f_{\mu|Y}(\mu) \propto \exp\left\{-\left[\frac{\delta^2_0 + \delta^2/n}{2\delta^2_0 + \delta^2/n}\right] \left[\mu - \left(\frac{\mu^{\text{est}} \delta^2_0 + \mu_0 \delta^2/n}{\delta^2_0 + \delta^2/n}\right)\right]^2\right\}$$

On peut identifier là une loi de densité normale avec espérance mathématique conditionnelle de μ et variance conditionnelle de μ , telles qu'on peut écrire

$$f_{\mu|Y}(\mu) \propto \exp\left\{\left[-1/(2 V(\mu|Y))\right] \left[\mu - E(\mu|Y)\right]\right\}$$

On aurait ainsi le noyau d'une loi de distribution normale pour μ , avec pour paramètres

$$E(\mu|Y) = (\mu^{\text{est}} \delta^2_0 + \mu_0 \delta^2/n) / (\delta^2_0 + \delta^2/n)$$

$$V(\mu|Y) = (\delta^2_0 + \delta^2/n) / (2\delta^2_0 + \delta^2/n)$$

On montre facilement, par quelques opérations de substitution, qu'on peut écrire

$$E(\mu|Y) = \alpha \mu^{\text{est}} + (1 - \alpha) \mu_0,$$

$$\text{avec } \alpha = (\delta^2/n)^{-1} / [(\delta^2/n)^{-1} + (\delta^2_0)^{-1}]$$

En effet l'espérance mathématique a posteriori peut s'interpréter comme une pondération entre la moyenne d'échantillon et celle associée à l'information aprioristique.

En analyse bayésienne, les expressions $(\delta^2/n)^{-1} = I(Y)$ et $(\delta^2_0)^{-1} = I(\mu)$ sont interprétées comme paramètres de précision associés respectivement à la donnée d'échantillon et à l'a priori. On peut écrire

$$I(\mu | Y) = [V(\mu | Y)]^{-1} = [1/\{(\delta^2/n)^{-1} + (\delta^2_0)^{-1}\}]^{-1} = I(Y) + I(\mu)$$

Avec l'exemple ci-dessus, pour des valeurs données de μ_0 , de δ^2 , de δ^2_0 , et de μ^{est} (égal à l'estimateur du maximum de vraisemblance de μ dans ce cas particulier de fonction de densité de Gauss), on peut calculer $E(\mu | Y)$ et $V(\mu | Y)$. Les résultats numériques du calcul des variances vérifient normalement la relation suivante: $V(\mu | Y) < V(\mu^{est})$. Cette relation signifie qu'en utilisant une information à priori, on gagne en précision par rapport à l'estimation d'échantillon seule.^{xxxiv}

6.4. Analyse bayésienne du modèle de régression

L'argument des bayésiens est que le processus d'estimation ne nous permet pas de déduire des valeurs pour des paramètres fixes, et que ce qu'il faut c'est de continuellement mettre à jour et affiner nos croyances subjectives concernant l'état et l'évolution du monde économique et des phénomènes qui le caractérisent.

La présentation des méthodes sera faite ici sur le modèle classique de régression linéaire standard

$$Y = X\beta + \epsilon$$

auquel nous associons les hypothèses habituelles quant aux termes d'erreurs:

$$\epsilon \sim N(0, \sigma^2 I).$$

Ce modèle comporte $k+1$ paramètres inconnus: les k coefficients et σ^2 . La fonction de vraisemblance s'écrit

$$f_{Y|\beta, \sigma^2}(y) = (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\{-1/2\sigma^2 (Y - X\beta)'(Y - X\beta)\}$$

Par quelques opérations algébriques et une décomposition de cette fonction de vraisemblance en un produit de deux fonctions, on peut écrire

$$f_{Y|\delta} (y) \propto (1/\delta^k) \cdot \exp\{(-1/2)(\beta - \beta_{mco})'[(X'X)/\delta^2](\beta - \beta_{mco})\} \cdot (1/\delta^{n-k}) \exp\{-SS/2\delta^2\},$$

avec $SS = (Y - X\beta_{mco})'(Y - X\beta_{mco})$.

On identifie dans le produit donnant cette fonction de vraisemblance les deux fonctions suivantes:

$$f_1(\beta_{mco}, \beta, \delta) = (1/\delta^k) \cdot \exp\{(-1/2)(\beta - \beta_{mco})'[(X'X)/\delta^2](\beta - \beta_{mco})\};$$

$$f_2(SS | \delta) = (1/\delta^{n-k}) \exp\{-SS/2\delta^2\}.$$

On peut noter les points suivants:

- La fonction f_1 concerne les k variables β_{mco} . La matrice de la forme quadratique que l'on trouve dans f_1 , c'est-à-dire la matrice $(X'X)/\delta^2 = [\delta^{-2}(X'X)^{-1}]^{-1} = \Omega^{-1}_{\beta_{mco}}$ est la matrice de précision associée à β_{mco} (pour δ^2 connu).
- Si δ est connu, la seule information a priori qui puisse être pertinente concerne le vecteur β . Puisque f_1 est proportionnelle à une loi normale multivariée et f_2 ne fait pas intervenir β , le conjugué naturel a priori pour β serait une loi normale.^{xxxx}

$$f_{\beta}(\beta) \sim \mathcal{N}(\beta_0, \Omega_0).$$

Il s'agit là de la loi marginale a priori pour β , si on suppose δ connu. Pour définir complètement cette fonction, il faudra déterminer les k coefficients du vecteur β_0 et les $k(k+1)/2$ variances et covariances nécessaires à la construction de Ω_0 .

- Si δ n'est pas connu, on s'intéresse aussi à $f_2(SS | \delta)$. On se rend compte que cette expression est proportionnelle à une fonction de densité gamma inverse, avec paramètres $\lambda = SS/2$, et $\tau = n - k - 1$.

Conditionnellement à δ , la densité f_1 est une loi normale,

- La détermination de la densité a posteriori dépend si l'on choisit pour β et δ une fonction a priori diffuse ou une fonction a priori informative.^{XXXX}

6.4.1. Cas d'une fonction de densité a priori non informative

On suppose que rien n'est connu a priori sur β et sur δ . On admet que les k composantes du vecteur β et $\log \delta$ sont distribués de manière uniforme et indépendante. On écrit alors la fonction de densité a priori de β et δ de la manière suivante:

$$f_{\beta, \delta}(\beta, \delta) \propto 1/\delta, \quad \text{avec } -\infty < \beta_i < \infty, \quad \delta > 0.$$

La fonction de densité a posteriori pour β et δ aura l'allure suivante:

$$f_{\beta, \delta | y}(\beta, \delta) \propto (1/\delta) \cdot f_{Y | \beta, \delta}(y)$$

En utilisant l'expression de la fonction de vraisemblance d'échantillon présentée ci-dessus, on vérifie facilement que cette densité a posteriori pour β et δ peut s'écrire de la manière suivante:

$$f_{\beta, \delta | y}(\beta, \delta) \propto (1/\delta^{n+1}) \cdot \exp\{-r/2\delta^2\}, \quad \text{avec } r = SS + (\beta - \beta_{mco})'(X'X)(\beta - \beta_{mco})$$

L'étape suivante consistera à déterminer les fonctions de densité a posteriori marginales pour β et δ

• La fonction de densité a posteriori marginale pour β se trouve en intégrant par rapport à δ . Notons que l'expression r définie ci-dessus ne dépend pas de δ . Nous pouvons écrire

$$(1/\delta^{n+1}) \cdot \exp\{-r/2 \delta^2\} = (1/\delta^{n+1}) \cdot (r/2)^{n/2} \cdot \exp\{-r/2 \delta^2\} \cdot (r/2)^{-n/2}.$$

On décompose cette expression en prenant $g(\delta) = (1/\delta^{n+1}) \cdot (r/2)^{n/2} \cdot \exp\{-r/2 \delta^2\}$, et en notant que l'expression $(r/2)^{-n/2}$ est constante par rapport à δ .

On s'aperçoit que $g(\delta)$ est proportionnelle à une loi gamma inverse de paramètres $\lambda=r/2$, et $\nu=n$. L'intégrale par rapport à δ est donc une constante. On aura ainsi la densité marginale suivante:

$$f_{\beta|Y}(\beta) \propto (r/2)^{-n/2}$$

et on peut facilement montrer qu'elle peut se mettre sous la forme suivante:

$$f_{\beta|Y}(\beta) \propto [1+(1/\nu)(\beta - \beta_{mco})' \Omega_{\beta_{mco}}^{-1} (\beta - \beta_{mco})]^{-(\nu+k)/2}.$$

Il apparaît ainsi que la fonction de densité a posteriori pour β est une Student multivariée, dont les caractéristiques sont:

$$E(\beta | Y) = \beta_{mco}, \quad \text{pour } \nu > 1; \quad V(\beta | Y) = [\nu/(\nu-2)] \Omega_{\beta_{mco}},$$

pour $\nu > 2$.

De là nous pouvons tirer les fonctions marginales pour chaque β_i :

$$f_{\beta_i|Y} \propto [1+(1/\nu)(\beta_i - \beta_{i,mco})^2 / \delta_{\beta_{i,mco}}^2]^{-(\nu+1)/2},$$

qui est une Student univariée générale.

De plus, la variable $(\beta_i - \beta_{i,mco}) / \delta_{\beta_i,mco}$ est distribuée selon une loi de Student univariée standardisée, avec $\nu = n - k$ degrés de liberté.

Nous connaissons ainsi la loi a posteriori pour β et ses caractéristiques (espérance mathématique, variance). Nous pouvons ainsi construire des énoncés de probabilités pour chaque β_i .

- Pour trouver la fonction marginale a posteriori pour δ , il faut intégrer par rapport à β . Nous écrivons

$$f_{\beta, \delta | y}(\beta, \delta) \propto (1/\delta^{\nu+1}) \cdot \exp\{-SS/2 \delta^{-2}\} \cdot (1/\delta^k) \cdot \exp\{(-1/2 \delta^{-2})(\beta - \beta_{mco})'(X'X)(\beta - \beta_{mco})\}.$$

On identifie aisément dans la fonction de densité a posteriori ci-dessus le produit des deux fonctions f_1 et f_2 suivantes:

$$f_1 = (1/\delta^{\nu+1}) \cdot \exp\{-SS/2 \delta^{-2}\}; \quad f_2 = (1/\delta^k) \cdot \exp\{(-1/2 \delta^{-2})(\beta - \beta_{mco})'(X'X)(\beta - \beta_{mco})\}.$$

On voit immédiatement que la fonction f_2 est proportionnelle à une densité normale comme fonction de distribution de β (l'intégrale est donc une constante).

L'intégration par rapport à β donne ainsi pour la fonction marginale a posteriori associée à δ une expression proportionnelle à f_1 :

$$f_{\delta | y}(\delta) \propto (1/\delta^{\nu+1}) \cdot \exp\{-SS/2 \delta^{-2}\}$$

Cela nous donne une gamma inverse avec paramètres $\lambda = SS/2$, et $\nu = n - k$.

Par les propriétés de la gamma inverse, nous obtenons

$$E(\delta | Y) = \frac{[(v-1)/2]}{[(v/2)]} \cdot (SS/2)^{1/2}; \quad E(\delta^2 | Y) = \frac{[(v-1)(v+1)]}{[(v/2)^2]} \cdot (SS/2); \\ \mathcal{N}[(v/2)-1] = 2\mathcal{N}(v-2) = SS/(n-k-2).$$

6.4.2. Cas d'une fonction de densité a priori informative

a) δ connu

Lorsque δ est connu, l'information a priori concerne uniquement le vecteur des coefficients de régression β . Pour celui-ci, nous choisissons le conjugué naturel qui, nous l'avons vu, est une loi normale multivariée.

Par conséquent

$$f_{\beta}(\beta) = \mathbf{N}(\beta_0, \Omega_0) \propto |\Omega_0|^{-1/2} \cdot \exp\{(-1/2)(\beta - \beta_0)' \Omega_0^{-1} (\beta - \beta_0)\},$$

où $|\Omega_0|$ est le déterminant de la matrice Ω_0 .

Pour la fonction de densité a posteriori nous obtenons

$$f_{\beta|Y}(\beta) \propto f_Y|_{\beta}(Y) \cdot f_{\beta}(\beta)$$

$$f_{\beta|Y}(\beta) \propto \exp\{(-1/2)[(\beta - \beta_{mco})' [(X'X)/\delta^2](\beta - \beta_{mco}) + (\beta - \beta_0)' \Omega_0^{-1} (\beta - \beta_0)]\}$$

$$f_{\beta|Y}(\beta) \propto \exp\{(-1/2)[(\beta - \beta^*)' C (\beta - \beta^*) + \text{Reste}]\}$$

Nous omettons pour l'instant de donner le détail du "Reste" qui figure dans cette expression, car il ne fait pas intervenir β . Ce reste sera défini par la suite, en vue des développements associés au cas où δ est inconnu. Nous précisons tout de suite à quoi correspondent le vecteur β^* et la matrice C .

β^* est l'espérance mathématique de la fonction de densité a posteriori pour β , et C est la matrice de précision (inverse de la matrice de variances-covariances C^{-1}):

Soit

$\Lambda_1 = (X'X)/\delta^2 = [\delta^2 (X'X)^{-1}]^{-1}$, la matrice de précision associée à β_{mco}

$\Lambda_0 = \Omega_0^{-1}$, la matrice de précision a priori.

Le développement détaillé (non reproduit ici) de la fonction de densité a posteriori $f_{\beta|y}(\beta)$ permet de constater que la matrice de précision a posteriori est égale à

$C = \Lambda_1 + \Lambda_0$ (somme des matrices de précision pour β_{mco} et a priori).

Quant à l'espérance mathématique de la fonction a posteriori, on montre qu'elle égale à

$\beta^* = C^{-1}[\Lambda_1\beta_{mco} + \Lambda_0\beta_0]$: une moyenne pondérée des deux vecteurs.

Avant d'étudier le cas où δ est inconnu, commençons par donner quelques précisions sur le "Reste" de l'exposant de la densité a posteriori présentée précédemment. On montre, par de simples développements algébriques (avec remplacement de β^* et C par leurs définitions), que

$$\begin{aligned} \text{Reste} &= -\beta^{*'} C \beta^* + \beta'_{mco} \Lambda_1 \beta_{mco} + \beta'_0 \Lambda_0 \beta_0 \\ &= \beta'_{mco} [\Lambda_1 - \Lambda_1 C^{-1} \Lambda_1] \beta_{mco} + \beta'_0 [\Lambda_0 - \Lambda_0 C^{-1} \Lambda_0] \beta_0 - 2\beta'_{mco} [\Lambda_1 C^{-1} \Lambda_0] \beta_0 \end{aligned}$$

On démontre facilement, en utilisant le fait que $CC^{-1} = C^{-1}C = \mathbf{I}$, que les trois matrices intervenant dans les trois expressions ci-dessus sont identiques:

$$A_1 - A_1 C^{-1} A_1 = A_0 - A_0 C^{-1} A_0 = A_1 C^{-1} A_0 .$$

On peut aussi montrer que ces matrices sont définies positives.

L'expression ci-dessus peut ainsi se réécrire de la manière suivante:

$$\begin{aligned} \text{Reste} &= \beta'_{mco} G \beta_{mco} + \beta_0' G \beta_0 - 2\beta'_{mco} G \beta_0 \\ &\quad (\text{avec } G = A_1 - A_1 C^{-1} A_1 = A_0 C^{-1} A_0 = \dots) \\ &= (\beta_{mco} - \beta_0)' G (\beta_{mco} - \beta_0) \end{aligned}$$

b) δ inconnu

Dans ce cas, il faut préciser une fonction a priori pour β et δ , c'est-à-dire $f_{\beta, \delta}(\beta, \delta)$. Ce qu'on a vu précédemment suggère que le conjugué naturel a priori pour β et δ peut être décomposé dans le produit d'une fonction normale pour β , conditionnelle par rapport à δ , et d'une gamma inverse pour δ . Cela signifie qu'a priori β et δ ne sont pas indépendants. On peut montrer que la fonction marginale a posteriori pour β sera une Student multivariée.

Nous posons

$$\begin{aligned} f_{\beta, \delta}(\beta, \delta) &= f_{\beta | \delta}(\beta) \cdot \mathbf{f}_{\delta}(\delta) \\ &\propto (1/\delta^k) \cdot \exp\{(-1/2\delta^2)(\beta - \beta_0)' A_0(\beta - \beta_0)\} \cdot \{(1/\delta^{v_0+1}) \cdot \exp(-SS_0/2\delta^2)\}. \end{aligned}$$

où β_0 est le vecteur d'espérances mathématiques a priori, Λ_0 est la matrice de précision a priori, à une constante δ^2 près, et $SS_0 = (Y - X\beta_0)'(Y - X\beta_0)$; ν_0 et $SS_0/2$ sont les paramètres a priori de la fonction gamma inverse.

La fonction de densité a posteriori devient

$$f_{\beta, \delta | Y}(\beta, \delta) \propto (1/\delta^{n+k+\nu_0+1}) \cdot \exp\{(-1/2 \delta^{-2})[SS+SS_0+(\beta-\beta_{mco})'A_1(\beta-\beta_{mco})+(\beta-\beta_0)'A_0(\beta-\beta_0)]\},$$

avec $A_1 = X'X$.

En introduisant l'expression "Reste" de la section précédente, nous obtenons

$$f_{\beta, \delta | Y}(\beta, \delta) \propto (1/\delta^{n+k+\nu_0+1}) \cdot \exp\{(-1/2 \delta^{-2})[SS+SS_0+(\beta-\beta^*)'C(\beta-\beta^*) + Reste]\} \\ \propto (1/\delta^k) \exp\{(-1/2 \delta^{-2})(\beta-\beta^*)'C(\beta-\beta^*)\} \cdot (1/\delta^{n+\nu_0+1}) \\ \exp\{(-SS^*/2 \delta^{-2})\},$$

avec $SS^* = SS + SS_0 + Reste$,

C et β^* et le Reste ayant les mêmes définitions qu'avant.

La fonction de densité a posteriori ci-dessus se décompose ainsi en deux fonctions:

$$f_1 = (1/\delta^k) \exp\{(-1/2 \delta^{-2})(\beta-\beta^*)'C(\beta-\beta^*)\} \sim \mathbf{N}(\beta^*, \delta^2 C^{-1}),$$

$$f_2 = (1/\delta^{n+\nu_0+1}) \exp\{(-SS^*/2 \delta^{-2})\} \sim \text{gamma inverse de paramètres } SS^*/2 \text{ et } \nu = n + \nu_0$$

L'intégration de la fonction a posteriori par rapport à β est immédiate puisque f_1 est distribuée selon une loi normale.

La fonction de densité marginale a posteriori pour δ est ainsi une gamma inverse avec paramètres $SS^*/2$ et $\nu = n + \nu_0$. On aura alors

$$E(\delta^2 | Y) = SS^*/(v-2) = SS^*/(n+v_0-2).$$

Pour trouver la fonction marginale a posteriori associée à β nous écrivons

$$f_{\beta, \delta | Y}(\beta, \delta) \propto (1/\delta^{n+k+v_0+1}) \cdot \exp\{(-r/2 \delta^2)\} \cdot [r/2]^{(n+k+v_0)/2} \cdot [r/2]^{-(n+k+v_0)/2}, \text{ avec}$$

$$r = SS^* + (\beta - \beta^*)' C (\beta - \beta^*)$$

Nous décomposons ensuite la fonction ci-dessus en un produit de fonction g_1 et g_2 , avec

$$g_1 = (1/\delta^{n+k+v_0+1}) \cdot \exp\{(-r/2 \delta^2)\} \cdot [r/2]^{(n+k+v_0)/2}$$

$$g_2 = [r/2]^{-(n+k+v_0)/2}$$

On constate que g_1 est une gamma inverse (par intégration cette partie donnera une constante), de sorte que le résultat de l'intégration par rapport à δ peut s'écrire comme suit:

$$\begin{aligned} f_{\beta | Y}(\beta) &\propto [r]^{-(n+k+v_0)/2} \\ &\propto [SS^* + (\beta - \beta^*)' C (\beta - \beta^*)]^{-(n+k+v_0)/2} \\ &\propto \{1 + [1/(n+v_0)](\beta - \beta^*)' [[1/(SS^*/(n+v_0))]C](\beta - \beta^*)\}^{-(n+k+v_0)/2} \end{aligned}$$

Il s'agit bien là d'une fonction de Student multivariée, avec

$$E(\beta | Y) = \beta^* = C^{-1} [A_1 \beta_{mco} + A_0 \beta_0], \text{ pour } n+v_0 > 1;$$

$$V(\beta | Y) = [(n+v_0)/(n+v_0-2)] [SS^*/(n+v_0)] C^{-1}, \text{ pour } n+v_0 > 2.$$

Résumons la démarche que nous venons de présenter pour ce qui concerne la régression: la fonction de densité a posteriori combine l'information a priori et la vraisemblance d'échantillon, ce qui veut dire qu'elle contient toute l'information dont le chercheur peut disposer sur les paramètres. Mais pour des raisons liées à la présentation des résultats, la densité n'est pas quelque chose de précis, et on préférerait pouvoir disposer d'un estimateur ponctuel ou par intervalles. Une approche tout à fait naturelle consiste à utiliser comme estimateur l'espérance mathématique de la densité a posteriori.^{xxxvii}

6.5. Analyse bayésienne en économétrie: avantages et limites

Un des critères de sélection d'estimateurs en analyse bayésienne consiste à se référer au coût qu'implique pour le chercheur le remplacement de la vraie valeur d'un paramètre par une valeur estimée. Un tel coût est mesuré par la quantité $L(\theta, \theta^{est})$, où θ est la vraie valeur du paramètre et θ^{est} une valeur estimée. $L(\theta, \theta^{est})$ est appelée fonction de perte. Si on sait que la vraie valeur de θ se situe dans un intervalle (a,b) , et si $f(\theta | \text{données})$ est la distribution a posteriori de θ , alors l'espérance mathématique de la perte sera donnée par

$$E[L(\theta, \theta^{est})] = \int_a^b L(\theta, \theta^{est}) f(\theta | \text{données}) d\theta.$$

La fonction de perte la plus populaire est celle associée au carré de l'erreur (fonction quadratique):

$$L(\theta, \theta^{est}) = k(\theta - \theta^{est})^2$$

où k est une constante. Cette fonction est telle que $E[L(\theta, \theta^{est})]$ atteint son minimum au point où θ^{est} est égal à l'espérance mathématique de $f(\theta | \text{données})$. L'importance de l'espérance mathématique a posteriori repose sur ce résultat fondamental de l'approche bayésienne: si on suppose que la fonction de perte

associée aux erreurs dans l'estimation de θ est quadratique, l'estimateur qui minimise l'espérance mathématique de la perte attendue est l'espérance mathématique a posteriori (pour une démonstration, voir Zellner, 1971).^{xxxviii}

De façon générale, l'avantage majeur de l'approche bayésienne est qu'elle force l'analyste à utiliser, au-delà de la simple donnée d'échantillon, toute l'information à sa disposition, d'une manière méthodique et structurée.

Un point délicat avec cette approche est que l'information a priori doit souvent prendre la forme d'une distribution de probabilités, ce qui n'est pas toujours fait d'une manière autrement justifiable que par des considérations de convenance mathématique, surtout dans le cas où le chercheur décide de faire le choix d'un conjugué naturel a priori. La justification parfois donnée pour l'utilisation de fonctions de densité a priori informatives résulte de situations où ces fonctions correspondent à des distributions a posteriori obtenues lors d'analyses précédentes. Notons que l'arbitraire existe aussi dans certains choix faits en approche classique, par exemple dans la façon d'organiser l'expérience. Enfin, une autre difficulté potentielle avec l'analyse bayésienne est que les calculs peuvent s'avérer assez lourds si le nombre de paramètres à estimer est élevé et si en plus l'espérance mathématique de la densité a posteriori ne peut être obtenue que dans une longue procédure d'intégration numérique (ce qui est généralement évité grâce à un "bon" choix de fonctions). Ce qui est rassurant pour les deux écoles en définitive, c'est que l'approche classique et l'approche bayésienne convergent dans leurs résultats si on est en présence de grands échantillons.

*Les problèmes ne viennent pas tant de ce que l'on ignore,
mais de ce que l'on sait (ARTEMUS WARD)*

7. Synthèse et discussion

L'économétrie a beaucoup évolué durant les cinquante dernières années. Jusqu'aux environs des années soixante-dix, il suffisait d'avoir une bonne connaissance de l'algèbre matricielle, des techniques de statistique multivariée et des rudiments de l'analyse des séries chronologiques pour faire de l'économétrie en utilisant les programmes d'ordinateur qui existaient à l'époque. De nos jours il faut connaître un peu plus sur des techniques appliquant le principe du maximum de vraisemblance, sur les méthodes non paramétriques et autres alternatives aux moindres carrés, sur la construction et les propriétés des modèles dynamiques, etc.

Dans notre présentation nous avons cherché à faire un exposé relativement détaillé sur un certain nombre de méthodes (telles que les modèles pour choix qualitatifs, l'approche bayésienne, les routines de régressions non linéaires, les modèles pour données de panels) qui ne sont généralement pas suffisamment abordées ou exploitées dans les programmes universitaires de niveau licence. Nous avons aussi fait en sorte d'expliquer clairement l'utilité de ces méthodes pour les applications économétriques, ainsi que l'apport spécifique de chacune d'elles dans la gamme d'outils de modélisation dont disposent de nos jours les économistes. Pour toutes ces méthodes, et en particulier pour celles sur lesquelles nous ne nous sommes pas particulièrement étendu, nous avons fourni dans la bibliographie un nombre important de références pour les lecteurs intéressés.

Lorsque l'on compare différentes spécifications économétriques, il n'y a pas de problème sérieux si les estimations vont toutes plus ou moins dans le même sens. Lorsqu'on a des estimations contradictoires, le choix entre différentes spécifications doit reposer sur l'importance relative que le décideur accorde à différents facteurs. Comme le dit si bien J. Johnston (1984), "une spécification est un tout, comme le choix d'une épouse ou d'un lieu d'habitation". Personne n'a encore trouvé la manière de constituer un ensemble parfait, même si nous essayons continuellement d'améliorer notre façon de faire. Dans le cas des

spécifications économétriques, le choix se fonde aussi bien sur des fondements théoriques vérifiables que sur l'expérience et l'intuition, puisqu'on peut tenir compte des connaissances accumulées et des critiques reçues.

Nous espérons que cette présentation constituera un stimulant significatif pour les chercheurs en économétrie dans nos universités, non pas seulement dans le but de faire des calculs ou des applications fondées sur les méthodes présentées, mais aussi dans la perspective de les pousser vers l'exploration et la découverte de nouvelles méthodes. En effet, les spécificités des situations qu'on rencontre dans le monde réel et leurs évolutions ouvrent sans cesse un champ toujours plus fertile pour les chercheurs intéressés à découvrir de nouveaux outils, de nouvelles propriétés, de nouvelles lois.

BIBLIOGRAPHIE

- Anizoba, E.K. 1980. *Modèle de régression à coefficients stochastiques*, Editions universitaires, Fribourg-CH.
- Anselin, L., 1988. *Spatial econometrics: Methods and models*, Kluwer Academic Publishers, Dordrecht (Pays-Bas).
- Antoniadis, A., Berruyer, J. et Carmona, R. 1992. *Régression non linéaire et applications*, Economica, Paris.
- Artus, P., Deleau, M. et Malgrange, P. 1986. *Modélisation macroéconomique*, Economica, Paris.
- Avery, R.B. 1977. "Error components and seemingly unrelated regressions", *Econometrica*, Vol. 45.
- Balestra, P. 1973. "Best quadratic unbiased estimators of the variance-covariance matrix in normal regression", *Journal of Econometrics*, Vol. 1.
- , 1980. "Modèles de régression avec variables muettes explicatives", *Documents de travail de l'Institut de Mathématiques économiques*, No. 39. Université de Dijon (33 pages).
- , 1994. "Statistique et analyse économique: quel mariage?", *Revue suisse d'Economie politique et de Statistique*, 130(3), 363-376.
- Balestra, P. et Nerlove, M. 1966. "Pooling cross-section and time-series data in the estimation of a dynamic model: The demand for natural gas", *Econometrica*, Vol. 34.
- Becker, W.E. et Greene, W.H. 2001. "Teaching statistics and econometrics to undergraduates", *Journal of Economic Perspectives*, Vol. 15, No 4, 169-182.

- Belsley, D.A.; Kuh, E. et Welsh, R.E. 1980. *Regression diagnostics: Identifying influential data and sources of collinearity*, John Wiley & Sons, New York.
- Bienaymé, A. 1994. "L'oeil des mathématiques et le regard des économistes", *Problèmes économiques*, No. 2.373, 1-6.
- Butare, Th. 1981. "Principes et procédures de tests d'hypothèses sur les coefficients du modèle classique de régression linéaire", *Mémoire de Licence en Sciences économiques, Université de Fribourg-CH* (68 pages).
- , 1990. "Pooling heterogeneous cross-section data for the estimation in an error components setup: Methodology and application to Engel curves", *Cahiers du Département d'Économie politique*, No. 90.12, Université de Genève (34 pages).
- , 1991. "Education et productivité dans les secteurs traditionnels: une analyse empirique", *Economie et prévision*, No 97(1), 35-48.
- , 1998. *Secteurs traditionnel et moderne dans un processus de développement*, INU Press, Genève et Cambridge MA.
- , 2002. "The statistical approach, its relevance, and data quality requirements", *Revue scientifique de l'Université Libre de Kigali*, No 2, 53-67.
- Butare, Th. et Favarger, Ph., 1992. "Analyse empirique du cas de la Suisse", in Beat Bürgenmeier, ed., *Main-d'œuvre étrangère: Une analyse de l'économie suisse*, Economica, Paris.
- Butare, Th., Favarger, Ph. et Bürgenmeier, B., 1992. "Effects of foreign labour on the production pattern: The Swiss case", *Swiss Journal of Economics and Statistics*, Vol. 128, No 2, 103-124.

- Capéraà, Ph. et Van Kutsen, B. 1988. *Méthodes et modèles en statistique non paramétrique; exposé fondamental*, Dunod, Paris.
- Chiang, A.C. 1984. *Fundamental methods of mathematical economics*, 3ème édition, McGraw-Hill, Singapour.
- Dickey, D.A. et Fuller, W.A. 1979. "Distribution of the estimates for autoregressive time series with unit root", *Journal of the American Statistical Association*, Vol. 74, 427-431.
- Engle, R.F. et Granger, C.W.J. 1987. "Cointegration and error correction: Representation, estimation and testing", *Econometrica*, Vol. 55, No 2.
- Greene, W.H. 2003. *Econometric analysis*, 5ème édition, Prentice Hall, New York.
- Griffiths, W.E.; Hill, R.C. et Judge, G.G. 1993. *Learning and practicing econometrics*, John Wiley & Sons, New York.
- Johnston, J. 1984. *Econometric methods*, 3ème édition, McGraw-Hill, New York.
- Kennedy, P. 1994. *A guide to econometrics*, 3ème édition, Blackwell Publishers, Oxford.
- Kmenta, J. 2000. *Elements of econometrics*, 2ème édition, The University of Michigan Press, Ann Arbor.
- Koenker, R. et Hallock, K.F. 2001. "Quantile regression", *Journal of Economic Perspectives*, Vol. 15, No 4, 143-156.
- Krishnakumar, J. 1988. *Estimation of simultaneous equation models with error components structure*, Springer Verlag, Berlin.

- Krishnakumar, J., 1995. "A general methodology of robust estimation for panel data models", *Cahiers du Département d'Econométrie*, No. 95.02, Université de Genève (24 pages).
- Krishnakumar, J. et Gueye, E.-H., 1998. "Maximum Likelihood Estimation of Cointegrated Systems wit Higher Order Integrated Variables and Asymptotic Equivalence with Generalized Least Squares", *Cahiers du Département d'Econométrie*, No. 98.05, Université de Genève (47 pages).
- Makatou, M. et Ronchetti, E. 1996. "Robust inference: the approach based on influence functions", *Cahiers du Département d'Econométrie*, No. 96.01, Université de Genève (47 pages).
- Mills, T.C. 1990. *Time series techniques for economists*, Cambridge University Press, Cambridge.
- Mood, A.M.; Graybill, F.A. et Boes, D.C., 1974. *Introduction to the theory of statistics*", McGraw-Hill, Kogakusha.
- Mougeot, M. 1975. *Théorie et politique économiques régionales*, Economica, Paris.
- Nelson, C.R. et Plosser, C. 1982. "Trends and random walk in macro-economic time series: some evidence and implications", *Journal of Monetary Economics*, Vol. 10, 139-162.
- Paelinck, J.H.P. et Klaassen, L.H. 1979. *Spatial econometrics*, Saxon House, Teakfield Limited (Grande-Bretagne).
- Pindyck, R.S. et Rubinfeld, D.L. 1991. *Econometric models and economic forecasts*, 3ème édition, McGraw-Hill, Singapour.

Pinheiro, M., 1989. "An econometric approach to nonparametric regression", *Cahiers du Département d'Econométrie*, No. 89.07, Université de Genève (34 pages).

Solari, L. 1977. *De l'économie qualitative à l'économie quantitative. Pour une méthode de l'approche formalisée en science économique*, Masson, Paris.

Stone, R. 1981. *Aspects of economic and social modelling*, Droz, Genève.

Swamy, P.A.V.B. 1970. "Efficient inference in a random coefficient regression model", *Econometrica*, Vol. 38.

Wonnacott, T.H. et Wonnacott, R.J. 1995. *Introductory statistics for business and economics*, 4ème édition, John Wiley & Sons, New York.

Wooldridge, J.M. 2001. "Applications of generalized method of moments estimation", *Journal of Economic Perspectives*, Vol. 15, No 4, 87-100.

Zellner, A. 1962. "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias", *Journal of the American Statistical Association*, Vol. 57, 348-368.

Zellner, A. 1971. *Introduction to bayesian inference in econometrics*, John Wiley, New York.

Annexe 1: Quelques fonctions de densité utiles en analyse bayésienne

A1.1. Fonction de densité normale, dite loi de Gauss

La densité de probabilité d'une loi normale, lorsqu'il n'y a qu'une variable, est

$$f(x | \mu, \delta^2) = [2 \pi \delta^2]^{-1/2} \exp\{(-1/2 \delta^2) (x - \mu)^2\}.$$

On note $x \sim \mathcal{N}(\mu, \delta^2)$.

Si \mathbf{X} est un vecteur à n composantes qui suit la loi normale à n dimensions, on note

$\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}_{n \times 1}, \boldsymbol{\Omega}_{n \times n})$, ce qui correspond à la fonction de densité suivante:

$$f(\mathbf{X} | \boldsymbol{\mu}_{n \times 1}, \boldsymbol{\Omega}_{n \times n}) = [2 \pi \delta^2]^{-1/2} |\boldsymbol{\Omega}|^{-1/2} \cdot \exp\{(-1/2)(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Omega}^{-1} (\mathbf{X} - \boldsymbol{\mu})\}.$$

Dans cette formulation, $\boldsymbol{\mu}$ est le vecteur d'espérances mathématiques, et $\boldsymbol{\Omega}$ est la matrice des variances-covariances.

A1.2. Fonction gamma et lois de densité gamma

- **Fonction gamma**

$$\Gamma(n) = \int_0^{\infty} z^{n-1} e^{-z} dz, \quad n > 0.$$

On montre que $\Gamma(1) = 1$; $\Gamma(1/2) = [\pi]^{1/2}$.

Et pour n entier, on montre que

$$\Gamma(n) = (n-1)!$$

$$\Gamma(n+1) = n\Gamma(n).$$

- **Loi de densité gamma**

$$f_X(x \mid \lambda, n) = \left[\frac{\lambda^n}{\Gamma(n)} \right] x^{n-1} \exp(-\lambda x), \quad 0 \leq x < \infty, \quad \lambda, n > 0$$

$$E(X) = \int X f(x) dx = n/\lambda$$

$$V(X) = \int X^2 f(x) dx = n/\lambda^2.$$

Un des principaux cas particuliers de la loi de densité gamma s'obtient avec $n = r/2$ (où r est un entier positif), et $\lambda = 1/2$.

On aura ainsi la fonction de densité chi-carré (notée Π^2):

$$f_X(x \mid r) = \left[\frac{1}{2^{r/2} \Gamma(r/2)} \right] x^{(r/2)-1} \exp(-x/2) \sim \Pi_r^2.$$

- **Loi de densité gamma inverse**

Si X est distribuée selon une loi gamma de paramètres n et λ , et que $y = (x)^{-1/2}$, on dit que Y est distribuée selon une loi gamma inverse (y est égal à l'inverse de la racine carrée positive de x).

$$f_Y(y \mid \lambda, n) = \left[\frac{2^n}{\Gamma(n)} \right] \lambda^n \left[\frac{1}{(2y^{2n+1})} \right] \exp(-\lambda y^2), \quad y > 0$$

$$E(Y) = \lambda^{-1/2} \left\{ \frac{\Gamma(n-1/2)}{\Gamma(n)} \right\}, \quad \text{si } n > (1/2)$$

$$E(Y^2) = \lambda^{-1}, \quad \text{si } n > 1$$

La fonction gamma est souvent utilisée comme fonction de densité a posteriori pour un écart-type. Soit $y = \delta$, $n = \nu/2$, $\lambda = \nu s^2/2$.
On aura

$$f_\delta(\delta \mid \nu, s^2) = \left[\frac{2^\nu}{\Gamma(\nu/2)} \right] \left(\frac{\nu s^2}{2} \right)^{\nu/2} \left[\frac{1}{\delta^{\nu+1}} \right] \exp(-\nu s^2/2 \delta^2),$$

$$\delta > 0$$

Le noyau de la distribution est composé de la partie contenant $\delta^{-\nu-1}$, c'est-à-dire les deux derniers facteurs du produit ci-dessus.

A1.2. Lois de Student

- **Student univariée standardisée**

Une variable aléatoire T est distribuée d'après la loi de Student univariée standardisée si sa fonction de densité est donnée par

$$f_T(t | n) = n^{-1/2} [(n+1)/2] / [(1/2) (n/2)] [1+(t^2/n)]^{-(n+1)/2}$$

, $-\infty \leq x \leq \infty$, $n > 0$.

Le noyau de la distribution est constitué du dernier facteur du produit, c'est-à-dire l'expression contenant t .

$$E(T) = 0, \quad \text{si } n > 1$$

$$V(T) = n/(n-2), \quad \text{si } n > 2.$$

- **Student univariée générale**

On définit $h = 1/\sigma^2$ et la transformée $T^* = [T/h^{1/2}] + \mu$. On dit que T^* suit une Student univariée générale, et sa fonction de densité est

$$f_{T^*}(t^* | n, \mu, h) = [(n+1)/2] / [(1/2) (n/2)] [h^{1/2}/n^{1/2}]$$
$$[1+hn^{-1}(t^* - \mu)^2]^{-(n+1)/2}$$

$$E(T^*) = \mu, \quad \text{si } n > 1$$

$$V(T) = [n/(n-2)] [1/h], \quad \text{si } n > 2.$$

- **Student multivariée standardisée**

Soient X_1, X_2, \dots, X_m des variables $\sim \mathbf{N}(\mathbf{0}, \mathbf{I})$ indépendantes.

Soit $Q \sim \Pi_n^2$ indépendante des X_i .

Formons la variable $T_i = X_i / (Q/n)^{1/2}$

Chaque variable T_i est une Student univariée standardisée. Le problème est alors de trouver la fonction de densité jointe des m variables T_i . On montre que

$$f_{T_1, \dots, T_m}(t_1, \dots, t_m \mid m, n) = n^{-m/2} \frac{[(n+m)/2]^{-1} [(m/2) \quad (n/2)]}{[1+(1/n) \sum t_i^2]^{-(n+m)/2}}$$

Le dernier facteur de cette expression constitue le noyau de la distribution. Les fonctions marginales pour les T_i sont des Student univariées standardisées (leurs deux premiers moments ont déjà été déterminés).

- **Student multivariée générale**

On se donne

a) un vecteur normal X à m composantes: $X \sim \mathcal{N}(\mathbf{0}, \mathbf{\Omega})$

b) une variable aléatoire $Q \sim \chi_n^2$, indépendante de X .

On construit des Student univariées générales T_i^* , éléments d'un vecteur T^* et tels que

$$T_i^* = X_i / (Q/n)^{1/2} + \mu_i,$$

et on cherche ensuite à déterminer la fonction de densité jointe des m variables aléatoires T_i^* . On peut montrer, par une intégration avec changement de variables, que

$$f_{T^*}(t^* \mid m, n, \mu, \mathbf{\Omega}) = n^{-m/2} \frac{[(n+m)/2]^{-1} |\mathbf{\Omega}|^{-1/2} [(m/2) \quad (n/2)]}{[1+(1/n) (t^* - \mu)' \mathbf{\Omega}^{-1} (t^* - \mu)]}$$

C'est la dernière partie de ce produit, c'est-à-dire $[1+(1/n) (t^* - \mu)' \mathbf{\Omega}^{-1} (t^* - \mu)]$, qui fournit le noyau de la distribution.

La fonction marginale pour T_i^* est une Student univariée générale, avec

$$E(T_i^*) = \mu_i, \text{ pour } n > 1; \quad V(T_i^*) = [n/(n-2)] \delta_{ii}, \text{ pour } n > 2.$$

Annexe 2: Brève présentation des méthodes dites de l'économétrie spatiale

Les travaux d'économétrie spatiale occupent une place particulière dans l'économétrie générale: ils se situent dans une tendance relativement récente, visant une meilleure intégration des divers développements statistiques et économétriques et de l'analyse spatiale des phénomènes économiques.

Sur le plan des fondements théoriques, les adeptes de cette discipline s'appuient sur le fait que la théorie économique habituelle est trop limitative quand elle postule un optimum supposant que les ressources et les activités sont situées en un point unique. Les recherches sur ces questions théoriques se proposent notamment de déterminer la localisation des agents et des facteurs, et de chercher ensuite l'affectation spatiale optimale des ressources caractérisées par une mobilité plus ou moins parfaite et coûteuse.

Sur la caractérisation de l'espace, différentes écoles de pensée existent: certaines considèrent l'espace comme un bien économique matériel qui fait l'objet d'une demande et d'une offre et dont l'affectation optimale dépend des conditions d'équilibre du marché ainsi défini; pour d'autres l'espace est un support, matériel ou immatériel, de l'activité économique. Dans un troisième champ de recherche, on s'intéresse à la répartition spatiale des ressources qui influencent les décisions d'implantation des activités économiques. Des liens existent bien entendu entre ces approches. Pour les spécialistes de science régionale, une théorie qui postule la maximisation de l'utilité des individus ne peut ignorer les contraintes technologiques et autres liées aux dotations régionales (Mougeot, 1975).

L'économétrie spatiale s'occupe d'interdépendances spatiales et d'hétérogénéités spatiales, qui représentent des aspects critiques des données habituellement utilisées par les spécialistes de science régionale (voir par exemple L. Anselin, 1988). On a par exemple des phénomènes économiques localisés dans un espace donné et qui doivent être expliqués par des facteurs de causalité localisés

dans d'autres espaces: pour prendre un exemple, on peut dire que les phénomènes migratoires s'expliquent par la comparaison d'avantages et d'inconvénients localisés en des endroits différents de l'espace considéré. Les caractéristiques de ces situations font que certaines techniques propres à l'économétrie classique s'avèrent inappropriées. Dans la plupart des analyses faisant usage de panels de données spatiales et temporelles, on suppose que les unités spatiales sont mutuellement indépendantes. Or si on considère que ces unités spatiales sont des régions géographiques séparées par des frontières arbitrairement tracées, comme par exemple les états composant le territoire des Etats-Unis d'Amérique, on peut s'attendre à ce que cette hypothèse ne soit pas satisfaite. Soit, pour prendre un exemple, la relation suivante:

$$y_{it} = f_{it}(x_{it}, \beta_{it}, \epsilon_{it}),$$

l'indice i étant pour l'unité spatiale observée, t pour l'unité temporelle; f_{it} est une relation fonctionnelle spécifique temps-espace qui explique la valeur de la variable dépendante y_{it} (ou un vecteur de variables dépendantes) en fonction d'un vecteur de variables indépendantes x_{it} , un vecteur de paramètres β_{it} , et un terme d'erreur ϵ_{it} . Cette formulation de base n'est pas immédiatement opérationnelle et il faudra poser des restrictions qui devront permettre d'avoir assez d'observations pour l'estimation des paramètres.

Un des aspects cruciaux de l'économétrie spatiale est le problème consistant à formaliser la structure des dépendances spatiales dont il faudrait tenir compte dans un modèle. Contrairement au domaine des séries temporelles où la notion de variable décalée ne comporte pas d'ambiguïté, en analyse spatiale les choses sont beaucoup plus complexes. La définition de décalages spatiaux de premier ordre ou d'ordres plus élevés et les méthodes qui y sont liées se caractérisent par une grande variété dans la littérature spécialisée. Les chercheurs en ce domaine recourent généralement à des notions de topologie leur permettant d'aborder les concepts de voisinage et de voisin le plus proche.

Les économètres spécialisés en analyse spatiale ont fait diverses propositions d'estimateurs ayant pour but de mieux refléter les spécificités de la science régionale et les interdépendances spatiales: estimateurs conditionnels strictement positifs, moindres carrés dynamiques simultanés, etc. (voir par exemple Paelinck et Klaassen, 1979).

Les méthodes d'estimation développées en économétrie spatiale s'avèrent prometteuses en vue d'une plus grande fiabilité des résultats notamment dans le domaine de l'analyse régionale appliquée, où les méthodes de l'économétrie classique peuvent s'avérer insuffisantes. Il y a toutefois beaucoup à faire encore (développement d'un cadre asymptotique de référence, outils d'analyse non paramétriques, extension des méthodes de l'analyse bayésienne, etc.) avant que cette discipline occupe une place significative dans le monde des économètres.

ⁱ Il est important de mentionner aussi le fait que les progrès considérables qui ont été accomplis en économie quantitative durant les récentes décennies ont beaucoup bénéficié des développements de la comptabilité nationale dans de nombreux pays, suite notamment à la parution, en 1953, du Système des comptes nationaux élaboré par les Nations-Unies.

ⁱⁱ Le professeur L.R. Klein de l'Université de Pennsylvanie a été un des premiers à vulgariser la pensée keynésienne, notamment celle véhiculée par la Théorie générale, sous forme de schémas formalisés, dans le but d'aider à éclairer la décision politique.

ⁱⁱⁱ A cette époque les échanges entre les deux groupes de protagonistes se faisaient principalement à coups de coefficients de détermination censés donner toute leur validité aux relations qui avaient leur préférence, ce qui ne manque bien sûr pas de faire sourire dans les milieux économétriques d'aujourd'hui.

^{iv} Nous donnons ici, à des fins d'illustration, quelques exemples: le modèle annuel que Tinbergen a proposé en 1936 pour les Pays-Bas avait 24 équations, le modèle annuel de Klein-Goldberger (1955) pour les Etats-Unis 25 équations, le modèle trimestriel de la Brookings Institution (1975) pour les Etats-Unis 326 équations, et le modèle annuel DMS de l'INSEE (France, 1978) a été conçu avec au départ 1300 équations. Ce débat ne porte pas uniquement sur les modèles économétriques comme ceux que nous venons d'évoquer, mais il vaut tout autant pour la catégorie de modèles dite "modèles d'équilibre général appliqués", qui se sont beaucoup développés dans le courant des vingt-cinq dernières années. Ces modèles visent à simuler des effets de politiques à moyen et long terme (politiques commerciales, réformes fiscales, etc.), en se fondant sur la théorie économique et sur les données structurelles des économies considérées. Le travail fait avec ces modèles n'incorpore généralement pas un travail d'estimation au sens économétrique du terme, les valeurs des paramètres ayant été déterminées préalablement par des méthodes de calibrage (voir par exemple la modélisation multisectorielle d'une économie en développement présentée par Butare, 1998).

^v Pour prendre une analogie dans le domaine des transports, on peut avancer que l'approche mathématique constitue un mode de transport - composé de symboles, d'équations et de théorèmes - qui permet dans beaucoup de cas d'arriver à destination - c'est-à-dire aux conclusions - d'une manière relativement rapide et claire. Toutefois ce propos mérite qu'on y apporte quelques nuances (voir Butare, 2002, en particulier les conclusions de l'article).

^{vi} Nous ne nous étendons pas plus ici sur les caractéristiques de ce modèle, que l'on trouve dans beaucoup de manuels d'économétrie, qui donnent aussi des exposés assez complets sur différentes extensions qui découlent de la violation d'une ou plusieurs hypothèses du modèle de base. Il y a par exemple les cas de l'hétéroscédasticité et de l'autocorrélation des résidus, qui donnent lieu à une matrice $E(ee')$ qui n'est plus scalaire, impliquant alors le recours à une méthode de moindres carrés généralisés, si l'on maintient le critère d'une minimisation de sommes de carrés de résidus. La formule de calcul de l'estimateur de \exists est alors donnée par l'expression

$$(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y, \text{ avec } \Sigma = E(ee'),$$

et la structure de cette matrice Σ dépendra des hypothèses postulées sur les variances et covariances des termes d'erreur. On peut montrer que cette façon d'estimer \exists revient à appliquer la méthode des moindres carrés ordinaires sur le modèle (1.1) préalablement transformé (voir par exemple Griffiths et al., 1993).

^{vii} Ceci veut dire qu'il s'agit d'un système où on n'a pas de cas où une variable endogène (c'est-à-dire expliquée par le modèle) dans une équation revient dans une autre équation en tant que variable explicative.

^{viii} Un tel système de fonctions a par exemple été estimé par Butare (1990), avec un panel de données spatiales hétérogènes.

^{ix} Pour la structure de la matrice des variances-covariances et les propriétés des estimateurs qu'on obtient dans le cadre de ce modèle, voir Zellner (1962) ou un manuel d'économétrie relativement complet tel que celui de Griffiths et al. (1993).

^x Mais les pionniers de cette approche sont incontestablement Balestra et Nerlove, avec leur article de 1966 où ils appliquent la méthodologie d'estimation avec erreurs composées à une fonction dynamique de demande de gaz naturel.

^{xi} Cette théorie est rarement enseignée dans les universités et les grandes écoles. La présentation proposée dans cet exposé porte sur une méthode d'estimation illustrée sur le cas d'une seule variable dépendante. Pour le lecteur intéressé par différentes extensions et illustrations de cette méthode, nous suggérons la consultation d'ouvrages consacrés partiellement ou entièrement au sujet (voir par exemple Antoniadis et al., 1992, ou Griffiths et al., 1993).

^{xii} L'hypothèse de linéarité facilite l'évaluation des estimateurs sous forme analytique et la détermination de leurs propriétés statistiques dans l'échantillon. Quant aux résultats des régressions paramétriques non linéaires, ils sont en général de nature asymptotique, et dans ces situations les développements et les approximations prennent une importance particulière.

^{vii} Nous n'abordons pas ici le cas de modèles initialement non linéaires pour lesquels, on arrive à obtenir, avant la spécification du terme stochastique et la procédure d'estimation, une forme linéaire découlant d'une transformation en logarithmes ou d'un autre traitement mathématique fondé par exemple sur une information spécifique sur les paramètres. Pour ces cas-là il n'y a en effet pas de difficulté majeure, puisque les méthodes de la régression linéaire classique deviennent applicables.

^{viii} Pour une analyse portant sur les principaux tests appliqués aux modèles de régression, sur les conflits potentiels entre procédures de tests et sur leur puissance, voir par exemple Butare (1981).

^{ix} De même, on peut envisager une procédure de minimisation de la somme des valeurs absolues des résidus, dans laquelle un algorithme donne des poids constants aux résidus jusqu'à un certain seuil, puis décroissants ensuite.

^x Pour d'autres précisions et quelques exemples sur ces notions, voir par exemple Markatou et Ronchetti, 1996).

^{xi} L'exposé fourni par Kennedy (1994) sur ce sujet présente en détails une méthode d'estimation de ces mesures de hauteur fondée sur la construction d'histogrammes.

^{xii} Si on pouvait par exemple soutenir que la forme fonctionnelle est de distribution normale, alors on pourrait utiliser les N observations disponibles sur y et sur x pour estimer les deux paramètres de cette spécification, c'est-à-dire la moyenne et la variance. Ainsi la densité pourrait être calculée pour toute valeur de x . L'idée avec l'approche non paramétrique est précisément d'éviter de faire des hypothèses sur la forme de la fonction de densité, raison pour laquelle il faut procéder autrement.

^{xiii} Des chercheurs ont identifié des analogies entre les résultats de ces méthodes avec certains types d'estimateurs des moindres carrés généralisés (voir par exemple Pinheiro, 1989).

^{xiv} Par définition, un bruit blanc est une suite de variables aléatoires à espérance mathématique nulle, non corrélées entre elles et de même variance. Il s'agit donc d'une suite ε_t telle que

$E\{\varepsilon_t\} = 0$ pour $t=0, \pm 1, \pm 2, \dots$; $E\{\varepsilon_t \varepsilon_{t-\bar{t}}\} = \delta^2$ pour $\bar{t}=0$; $E\{\varepsilon_t \varepsilon_{t-\bar{t}}\} = 0$ pour $\bar{t} = \pm 1, \pm 2, \dots$.

^{xxi} Les tests de racine unitaire et la cointégration ont connu quelques applications à l'Université Libre de Kigali, notamment à travers le mémoire de licence de M. Ernest Mwiza, "Impact du secteur primaire sur l'économie rwandaise et analyse de quelques programmes d'investissement public" (juillet 2000) et l'article de M. Kigabo Rusuhuzwa, "Estimation d'une fonction de demande de monnaie: le cas du Rwanda", paru dans le premier numéro de la Revue scientifique de l'Université Libre de Kigali (décembre 2001).

^{xxii} Sous l'hypothèse $|\lambda| > 1$, la variance de la série augmenterait de façon exponentielle avec t et dans ce cas on parle de processus explosif (les observations passées ont une pondération plus élevée que les observations présentes).

^{xxiii} Notons qu'avec cet exemple nous avons cherché à illustrer la méthodologie sur un cas simple. Différentes extensions de cette approche ont été développées dans la littérature, portant notamment sur des systèmes avec variables endogènes et/ou exogènes multiples, et la possibilité d'envisager un ordre d'intégration d relativement élevé pour les variables (voir par exemple Krishnakumar et Gueye, 1998).

^{xxiv} Les liens entre P_i , μ_i et $F(\zeta_i' \beta)$ sont spécifiés clairement ci-après: D'une manière générale, on part d'un modèle de choix discrets avec des utilités inobservables U_{ij} (l'indice i étant pour les individus, et l'indice j étant pour les alternatives): $U_{ij} = s_{ij}' \alpha + c_i' \lambda_j + \varepsilon_{ij}$, où les vecteurs s_{ij} et c_i représentent respectivement les attributs des alternatives j pour l'individu i et les caractéristiques individuelles de i , ε étant le terme stochastique du modèle. On admet que ε_{ij} est distribué selon une normale $N(0, \delta^2_j)$. On voit que le modèle ci-dessus est tel que l'utilité inobservable a une composante systématique \bar{U}_{ij} , qui correspond à la partie non stochastique du modèle. On peut ainsi écrire: $U_{i1} = \bar{U}_{i1} + \varepsilon_{i1}$; $U_{i0} = \bar{U}_{i0} + \varepsilon_{i0}$. On a aussi, pour reprendre les notations du début de cet exposé sur les modèles de choix discrets,

$$P_i = \text{Prob}(U_{i1} \geq U_{i0}) = \text{Prob}(\bar{U}_{i1} + \varepsilon_{i1} \geq \bar{U}_{i0} + \varepsilon_{i0}) = \text{Prob}(\varepsilon_{i0} - \varepsilon_{i1} \leq \bar{U}_{i1} - \bar{U}_{i0}) = \text{Prob}(c_i \leq \mu_i),$$

avec $c_i = \varepsilon_{i0} - \varepsilon_{i1}$, et $\mu_i = \bar{U}_{i1} - \bar{U}_{i0}$.

En partant des fondements de la théorie de l'utilité, on a défini comme indice d'utilité μ_i la différence entre les parties systématiques de l'utilité:

$\mu_i = \bar{U}_{i1} - \bar{U}_{i0} = (s_{i1}' - s_{i0}')\alpha + c_i'(\lambda_1 - \lambda_0)$. On aura ainsi

$$\begin{aligned} \mu_i &= [(s_{i1}' - s_{i0}') \quad c_i'] [\alpha \quad (\lambda_1 - \lambda_0)]' \\ &= \zeta_i' \gamma, \text{ avec } \zeta_i' = [(s_{i1}' - s_{i0}') \quad c_i'], \quad \gamma = [\alpha \quad (\lambda_1 - \lambda_0)]' \end{aligned}$$

Ensuite, en normalisant les termes d'erreur $e_i \sim \mathcal{N}(0, \sigma^2 = \sigma^2_0 + \sigma^2_1)$, on peut écrire

$$P_i = \text{Prob}(e_i \leq \zeta_i' \gamma) = \text{Prob}(e_i / \sigma \leq \zeta_i' \gamma / \sigma) = \text{Prob}(z_i \leq \zeta_i' \beta) = F(\zeta_i' \beta).$$

^{xxv} A droite et à gauche de ce point on a 50% de la surface totale couverte par la fonction cumulative, ce qui traduit aussi l'idée qu'à priori les deux alternatives ont les mêmes chances d'être choisies.

^{xxvi} Pour d'autres exemples de calculs et différents types de tests associés à ce modèle, voir par exemple Pindyck et Rubinfeld, 1991.

^{xxvii} Il est à remarquer que le ratio $\exp(\zeta_i' \beta_k)$ n'est pas affecté par la présence d'autres alternatives.

^{xxviii} Pour un exposé sur ce modèle et sur le traitement fait de la présence d'alternatives constituant de proches substituts, voir par exemple Kennedy (1994).

^{xxix} Pour la procédure d'estimation il est usuellement fait recours à une normalisation consistant à considérer que le terme d'erreur ϵ a une espérance mathématique nulle et une variance égale à 1.

^{xxx} Certaines procédures d'estimation proposent le remplacement de l'espérance mathématique conditionnelle de ϵ par une variable instrumentale et d'introduire dans le modèle un nouveau terme d'erreur ayant les propriétés habituelles. Pour éviter le coût des calculs potentiellement important des procédures de maximum de vraisemblance, on peut recourir à des procédures par moindres carrés en deux étapes, utilisant par exemple un modèle linéaire de probabilités avec une variable dépendante z prenant la valeur 1 quand l'individu possède une voiture et zéro quand il n'en possède pas. Mais on sait que les estimateurs obtenus avec ces méthodes présentent souvent des inconvénients quant aux propriétés attendues des estimateurs et au non-respect des restrictions en rapport avec l'intervalle de variation des prédictions du modèle, comme nous l'avons déjà fait remarquer.

^{xxxv} On arrive à voir assez clairement, à travers une représentation graphique, que les écarts entre les points observés et la droite d'ajustement (fournie par exemple par le modèle linéaire des probabilités) sont considérablement élevés, raison pour laquelle les coefficients R^2 auront des valeurs très basses dans le cas des modèles de choix discrets (voir par exemple Kennedy, 1994, p. 228).

^{xxxvi} Thomas Bayes (1702-1761) était mathématicien et pasteur presbytérien. C'est à lui qu'on doit la théorie de l'inférence probabiliste qui permet de calculer, en utilisant le nombre de fois qu'un événement s'est déjà produit, la probabilité qu'il se produise lors de futures expériences. Le principal ouvrage de référence en économétrie bayésienne est depuis longtemps et reste encore celui de Zellner (1971).

^{xxxvii} On montre en utilisant la fonction de vraisemblance de l'échantillon que μ^{est} est l'estimateur du maximum de vraisemblance pour μ .

^{xxxviii} Parmi les fonctions de densité utilisées en analyse bayésienne il y a notamment, en plus de la normale (dite aussi loi de Gauss), la fonction gamma et la gamma inverse, la fonction de densité beta, les fonctions de Student univariée et multivariée de type général ou de type standardisé (pour des détails sur certaines de ces fonctions, voir par exemple Mood, Graybill et Boes, 1974).

^{xxxix} Les chercheurs recourent souvent au "conjugué naturel a priori", qui est une fonction ayant la même forme que la densité conditionnelle, ce qui facilite l'intégration nécessaire pour obtenir les distributions marginales. Si on a par exemple des observations sur un vecteur X et qu'il y a un vecteur θ de paramètres à estimer, on aurait une fonction de vraisemblance du type $f_{X|\theta}(x)$. Si la fonction $f_{X|\theta}(x)$ peut être décomposée en un produit de fonctions $f_1(t|\theta) \cdot f_2(x)$, où $t = [t_1, \dots, t_k]$ sont K fonctions de X , les variables t sont appelées statistiques exhaustives pour θ . La fonction $f_\theta(\theta) \propto f_1(t|\theta)$ est appelée "conjugué naturel a priori".

xxxvi Si aucune information a priori n'est connue sur les paramètres, alors on spécifie une densité a priori non informative pour refléter cette situation. Si θ est le vecteur de paramètres, on écrit la fonction de distribution $g(\theta) \propto \text{constante}$. Si on sait que θ ne peut pas être négatif, comme c'est le cas quand il s'agit d'une variance, et si on ne sait rien d'autre sur ses propriétés, la distribution a priori de θ peut être représentée par $g(\log \theta) \propto \text{constante}$, car les logarithmes de nombres négatifs n'existent pas. Par un changement de variable on peut montrer que cela revient à écrire $g(\theta) \propto 1/\theta$. Les distributions non informatives sont aussi dites "diffuses" ou "impropres", car l'intégrale de $g(\theta)$ de $-\infty$ à $+\infty$ ou de 0 à $+\infty$ n'est pas égale à 1, comme cela serait le cas s'il s'agissait d'une vraie densité de probabilités. L'implication d'une densité a priori non informative est que sur l'intervalle de variation envisageable pour les paramètres, toutes parties égales de cet intervalle ont la même probabilité de contenir des valeurs pour les paramètres.

xxxvii Comme on l'a vu tout au long de cet exposé, l'espérance mathématique a posteriori est une moyenne pondérée de l'estimateur d'échantillon et de l'information a priori, et les poids sont fournis par les inverses des matrices de variances-covariances correspondantes, qui sont interprétés ici comme des indicateurs de précision: une information a priori relativement précise donne à l'espérance a priori un grand rôle dans la détermination de l'espérance a posteriori, et une information importante provenant de l'échantillon (grande taille et/ou faible variance) donne à l'estimateur empirique un rôle prépondérant.

xxxviii Pour illustrer l'idée de symétrie qui caractérise la fonction de perte quadratique, Griffiths et al. (1993) donnent l'exemple de l'estimation de la dépense moyenne que les ménages consacrent à un produit: pour la firme productrice, une surestimation de cette dépense moyenne conduit à des pertes (stocks superflus), et une sous-estimation conduit aussi à des pertes (manque à gagner associée aux ventes potentielles non réalisées).



**COMPETENCE ACQUISITION
IN ENGLISH
CORRESPONDENCES**

**By Lecturer KAAYA Siraje
Kigali Independent University**

July 2003

TABLE OF CONTENTS

Acknowledgement

Introduction

- The origin of correspondences
- The ways of writing letters
- The types of letters
- personal, private, social, etc.

Official letters

- Applications
- Official letters
- Public letters

Commercial correspondences

- Inquiries
- Positive reply and regret to inquiries
- Order letters
- Credit status inquiry

Concluding remarks

Bibliography

ACKNOWLEDGEMENT

Writing an article up to this level is so hard a task. A person should go through various challenges seeking for help, advice, and corrections from different intellectuals here and there.

It is in this regard that; I wish to convey my sincere thanks to all those people who rendered a hand to me in making the writing of this article a success.

Notably, the following; Prof. Dr RWIGAMBA BALINDA the Promoter and Rector of ULK. Vote of thanks also goes to Mr. RWIRASIRA Eugène, the Academic Director of Kigali Independent University, and Mr NJOROGE Timothy, who tirelessly read through all my manuscripts plus the corrections and alterations they made before the printing of this article.

KAAYA Siraje

Lecturer in Kigali Independent University

Introduction

Today, Rwanda is a first developing country where various economic, political and social changes have positively penetrated irreversibly. The economic structure of Rwanda has changed tremendously. Other organizations including Banks, Universities, Hospitals, Rwanda revenue authority, etc and other non-Government organizations including Private Institutions and enterprises have come into existence. Foreign investors are showing a great interest in joining the development of the country.

However, most of the prospective investors are English users, yet earlier, the indigenous citizens of Rwanda were committed mostly to French language as the foreign language and to a smaller extent English. I look at the students of Rwanda today, as the most privileged people in years to come. They have an open chance to turn to bilingual system, which is of a great advantage.

This article contains the historical background of letter writing, the ways of writing letters, and the various types of correspondences.

Letter writing has been vogue since the time when man learnt how to read and write. It is not only a necessity but also an essential need. Men do not and can not always live together. They can not therefore, always communicate their ideas to one another through the words of the mouth. Very often we feel the necessity of communicating some information to a person or persons who are living far away from us. If in such cases we are to go to them, and convey what we have got for them, we would really spend a lot of money plus too much wastage of time.. Letter writing saves all these. Therefore the art of letter writing came into existence out of sheer necessity.

Letter writing covers a very wide field. Whenever there is a need for any communication between two persons living apart, there is the need of letter writing or correspondence.

Letter writing thus plays a very important part in our daily life and so every educated person should know how to write a clear and a readable letter.

Sometime every one of us has to write business or private letters and may have to face the problem of writing an important letter that will vitally affect his interests in life. The art of letter writing is not therefore, a mere ornamental accomplishment, but is something that every educated person should and must acquire for practical reasons.

Types of letters

The broad divisions into which the letters can be classified are as follows:

Personal, Private, social Informal letters

Which deal with the private purpose of writers, for example: letters written to family members, relatives and friends.

Business or commercial letters

Which deal with business transactions that are meant for business matters between two or several parties.

Official letters.

Which deal with official or administrative affairs, that is, Letters which are written by an official to another, in their official capacity or between an official and an ordinary person.

Ways of writing letters

Before considering what exactly these different types of letters are, and how they should be written, we should be acquainted with the positions and the different parts of the letter itself.

Please see the examples on pages; 107-113.

THE PARTS AND POSITIONS OF THE LETTER:

1 & 2: Usually position 1. and 2. are joined together and given the name of the *HEADING OF THE LETTER*. The heading shows where and when the letter is written. It generally consists of the postal address of the writer to which the reader may reply and date of writing the letter, which may serve as the purpose of reference. The position of the heading is in the top right hand corner of the first page of the letter, when the letter is indented, or to the top left hand corner of the letter when the letter is beginning from the left margin as shown in the Figure 1.3.; and sometimes on the top center of the paper when the paper is headed. Note that the sender's address or the addresser's address should always be above and the date normally comes below the address.

It should be noted that the address of the addresser in case of official letter is written **without the name** of the sender. Names are usually written down below the signature .

For example:

***KIGALI INDEPENDENT UNIVERSITY
P.O.BOX 2280 KIGALI
RWANDA
CENTRAL AFRICA***

Unless the addresser begins his/her address with the word 'from' then the name may appear in the address.

DATE

The date may be written in full or may be abbreviated as:

May 20th , 2003, or 20th May,2003
Or 04. 05. 2003 or 04 / 06 / 2003

3. Name and address of the addressee.

Usually the names and the job title of the addressee if known to the writer are written in this position see figure 1.3. on page 129. Sometimes the addressee is referred to as the receiver or we may call it the receiver 's names and address.

4. The salutation or Greetings.

The salutation or greetings is the complimentary word or expression with which the letter begins. It comes at the left-hand side of the paper at a lower level than the receiver's address or the heading in case the writer had not included the receiver's address, and such may be written as:

My dear Brown,

The form of the salutation always depends upon the relations of the writer to the addressee or the person written to. The different forms of salutations are as follow below:

i , To the member of one's family:

Dear Father, Dear Brother, Dear uncle or my dear Father, My dear Mother, My dear Sister, etc.

ii , To very intimate friends:

My dear Mahmood, My dear Joan etc.

iii, To ordinary acquaintances: (a person whom one knows but not a friend)

Dear Mr. White, Dear Peterson, Dear Miss Recheal,

iv, To strangers:

Sir, Dear Sir, Madam, Dear Madam etc.

v, To official superiors and especially in applications . Sir, Madam, Sir/Madam

vi, To business people:

Dear Sir, Dear Sirs, Gentlemen (but not dear Gentlemen.)

The following points should be noted in this connection:

a , The salutation should always be followed by a comma.

b , It should always begin with a capital letter. If it contains two words both should begin with capital letters. Such as:

Dear Father, but when they are three, the first and the third should begin with capital letters as in: My dear Father, My dear Husband.

c, The salutation Sir is used in official letters . It is also used in writing to strangers irrespective of ranks. (Dear Sir,) is reserved for those of equal ranks, it is less formal. My dear Sir implies respectful friendliness between the parts well known to each other, but the writer is inferior in social position or in some other respect to the person addressed.

5. Reference (object / objective).

Some people tend to put reference or object to show clearly the major reason of writing such a letter. References are mostly used in official letters. A reference is usually abbreviated as :

RE: _____, and it comes just below the salutation.

6. The body of the letter.

The body of the letter, communication or message forms the main part of the letter and it contains what the writer wants to communicate to the addressee. It should begin just below the reference or where the salutation ends, but one line lower. Business and official letters should always be brief and to the point. Business letters should be very polite and the official letters should be with full of respect. Private letters to friends and relatives should be chatty, but should always be full of courtesy, respect and affection. The style in which the letter is written will depend upon the kind of letter one wishes to write.

The style of the letter to an intimate friend will, therefore, be very different from that of a business letter or an official communication.

But there are some points as stated below, which apply to all the different kinds of letters:

i, The letter, unless it is very short , should be divided into paragraphs to mark the change of subject - matter , etc. .

ii, The language used should be simple and direct and the sentences should also be short . The writer should be clear about what he wants to say and say it as directly as possible.

iii, The letter should be written neatly so that the addressee may read without being put to unnecessary trouble for bad penmanship and slovenly writing .

Vi, The writer should be very careful about his punctuation and should put in commas and semi - colons and full stops in their proper places. Incorrect punctuation may alter the whole meaning of the sentence.

7. The Subscription

The complimentary close, subscription, or courteous leave taking is a form of polite leave taking that comes at the right hand side of the last page of the letter when the letter is indented and to the left hand margin of the paper when every thing begins from the left margin, just below the last line of the communication.

THE FORMS OF SUBSCRIPTION

The forms of subscription depend upon the degree of the relationship, which the addresser has with the addressee.

It should be consistent with the salutation as shown below.

Relation	Salutation	Subscription
Father	My Dear Father,	Your affection, Yours affectionately,
Lecturer	Sir, My Dear Sir,	Your Most Obedient Student, Yours Obediently,
Friend	Dear Friend , My Dear Peter,	Yours affectionately, Yours Lovely,
Acquaintance (some body you know but he/she is not a friend)	Dear Mr. Kigabo,	Yours faithfully, Yours Sincerely, Yours Truly,
Stranger	Sir, Dear Sir,	Yours Sincerely,
Businessman or firm	Dear Sir, Dear Sirs, Dear Madam, Gentlemen,	Yours faithfully, Yours Sincerely,
Official Superior	Sir,	Yours Most Obedient Servant, Yours Obediently, Yours Respectfully ,
Intimate Official	Dear sir,	Yours Faithfully,
Newspapers Editor	Sir,	Yours Truly,
President	Your Excellence	Yours respectfully , etc
Ambassador	Your Excellence	Yours respectfully , etc
Minister	Honourable	Yours respectfully , etc
Member of parliament	Honourable	Yours faithfully, etc
King / Queen	Your Majesty	Yours obediently, etc
Mayor of the city	Your Worship	Yours faithfully, etc
The Pope	Your lordship, / Your Worship,	Yours obediently, Yours faithfully, etc
Bishop	Your Worship	Yours faithfully, etc
Educational qualifications or army ranks, they are addressed according to their qualifications or ranks	Dear Doctor, Honourable army Chief of staff, etc	Yours respectfully, Yours faithfully, Yours sincerely, etc

Table 1.1

8. The Signature

The signature of addresser or the firm he / she represents, is placed just below the complimentary close. The Signature should always be in writing. It should never be typed or made in form of a stamp. If the writer is a woman, and the letter is addressed to a stranger, she should indicate her status (that is, whether she is a Miss or a Mrs., in Brackets), beside the name (s).

The signature should be legible. The signature is written as follow below:

Yours obediently

SIGN
THE ADDRESSER'S NAME
JOB TITLE

9. Name

The name (s) of the addresser comes below the signature, unlike in French *Written letters where the signature comes after the name. The format of signing the English letter may be as shown below:

Yours Truly,

Sign

The name of the addresser

Job designation

10. This is the position where the attachments or Notice are written. (See figure 1.3, where C.C. is written. C.C.)

11. This position is where the reference Number of the addresser is written. (See figure 1.5. our ref.:)

12. This is the place where the reference Number of addressee is written.

13. The word THRU: is an abbreviation to mean through, it is used when the writer passes through another person to address the other.

OFFICIAL LETTER SPECIMEN

The following four kinds of letters are regarded as official letters:

1. Those letters that pass between the Government officials and other Government officials, or between the Government officials and ordinary people;
2. Those letters between the officials of autonomous bodies (self governing / independent), like; Universities, hospitals, schools and other non- governmental organizations etc and the citizens in general ;
3. All kinds of petitions or applications ;
4. Those letters which are addressed to News Editors for publication in public interest.

It can be deduced that the first two kinds of letters are often grouped together and given the name of Official Letters Proper and the last kinds of letters in number four are designated as Public Letters. According to the selections or groupings made above, it will be evident that official letters wear the following three distinct forms:

- A, Applications
- B, Official letters proper
- C, Public Letters

Applications

These are letters which ask, request, or demand politely for something.

In drafting an application letter for a post, the following points should be carefully noted.

- a) You should state clearly what you want and where you got the information that there is a post or a vacancy in that company or organization you're writing to.
- b) State briefly, your nationality, age, marital status and your highest academic standards.

- c) Talk briefly about your working experience if you're working.
- d) Include a paragraph to mention the photocopies of your academic Credential (Testimonials) and curriculum Vitae enclosed within the letter.

The writer should not forget to include a complimentary close in a polite way as, Your most obedient servant to be, Yours respectfully, Yours obediently, etc.

Example

Write an application letter to the General Manager J.J. Book stores for a post of junior accountant. (Post advertised in IMVAHO of 03/02/2003).

There are many ways of beginning an application letter of which some are as shown below:

1. In response to your advertisement in yesterday's IMVAHO for a post of Junior Accountant in your company, I wish to offer myself as a candidate.
2. I wish to apply for a post of Junior Accountant in your company, which was advertised in IMVAHO of 03/02/2003.
3. I humbly wish to offer myself as a candidate for the post of Junior Accountant in your company, which was advertised in IMVAHO of the 3rd February 2003.
4. Owing to the information I received from your private Secretary that you need an experienced Junior Accountant, I hereby offering myself as a candidate.
5. I hereby applying for a post of junior accountant in your company, a post which I learnt of from one of your permanent employees.

6. Basing on the information I heard over the radio last night that you need an experienced Accountant, kindly allow me to compete for the post.
7. Having perceived the information through watching a Television last evening • that you're looking for a hard working Accountant, I respectfully wish to present my candidature to you.

Here below is a fully written example of an application letter.

Kigali Independent University
P.O Box 2280
KIGALI – RWANDA

9th February,2003

The General manager
Amavubi General Stores
P.O Box 2222
KIGALI
Sir,

**RE : APPLICATION FOR A POST OF JUNIOR
COUNTANT**

I humbly wish to offer myself as a candidate for the post of Junior Accountant in your company, which was advertised in IMVAHO of the 3rd February 2001.

I am Rwandan, aged 29, unmarried and a holder of Bachelor's degree in Economics and Management which I obtained from Kigali Independent University in 1999.

Presently, I am working with Rwandacel in the same capacity, a post I have been holding since 1998. Due to the fact I am efficient, hardworking and respect my job, I hope given a chance, I will serve you to the best of your satisfaction.

Attached are photocopies of my testimonials and a curriculum vitae.

I will be very grateful if my application is kindly considered.

Yours most obedient servant to be ,



UWIMANA Mariam.
Junior Accountant Rwandacel

Figure 1.1

Specimen of the Official letters proper

1. Write a letter to your employer stating your inability to attend to your duties due to an accident.

King Fiycal Hospital
KIGALI RWANDA
9th February ,2003

The Chief Accountant
Rwandacel

Dear Sir,

RE : Application for A month's medical leave

I truly beg to state that when I was going home after the working office hours yesterday, I got an accident. As I was approaching the main round about, a truck, which was coming from the opposite direction, knocked my car. The car was smashed beyond recognition and I got injuries all over my body. Fortunately, good samaritans me rushed to the hospital where I am lying in miserable conditions.

The physician of the ward said that I should require a month to recover.

In these circumstances, I am unable to attend to my duties and would request you kindly to grant me a month's medical leave with effect from today.

Attached are photocopies of my medical admission forms as requested by the company regulations.

Yours most obedient servant ,



BOLOLEBWAMI Patrick
Junior Accountant Rwandacel

C.C.

1. General Manager
2. Company Secretary

Figure 1.2

2. Write a resignation letter to the managing director of your company, giving reasons for the cause.

Mbulabuturo
P.O Box 22
KIGALI - RWANDA

4th February, 2003
The managing Director
Amavubi General Stores
P.O Box 2222
KIGALI - RWANDA

Dear Sir,

RE : RESIGNATION

I beg to tender my resignation to the post of junior accountant in your company and would request you to kindly accept it and relieve me of my responsibilities with effect from 28th February 2001.

I respectfully wish to state that I have been ailing for the last few weeks and have failed to attend to my duties properly in the sequence of which the company has been at a great disadvantage.

I am grateful to you for the good treatment to me, but I feel, I should not continue because I will inconvenience you all. Frankly speaking it is no longer possible for me to render services to the company, which I so deeply love, due to my Doctors' advice that I should take at least a complete rest of four months.

May I respectfully once again request you to kindly accept my resignation and relieve me of my duties with effect from 28 May 2001.

Yours respectfully,


UWERA Jane

C.C.

1. Company Secretary
2. Financial Controller

Figure 1.3

PUBLIC LETTERS

All letters, which are written to the Newspaper Editors, are designated as Public letters. These letters are called public letters because they are for public interest. The only object of such letters is to draw the attention of the person(s) or authorities concerned as well as of public towards the defects, fault, neglect or cases of omissions under reference which, in the interest of the public need notice for rectification.

1. Write a public letter to the Editor of New times newspaper, expressing your discontents about the uncovered manholes in the city.

Kimicyanga market
P.O Box 222
KIGALI - RWANDA
15th February, 2003

The Editor
New times News paper
KIGALI - RWANDA

Sir,

RE : UNCOVERED MANHOLES IN THE CITY

It looks as if the authority concerned does not care about the life the people they are serving. This is due to the fact that almost on each and every street in the city now-a-days, there are uncovered manholes. These manholes are very long and on top of that they carry dirty products from the toilets in the city.

So, if any one happens to accidentally fall into such a trench, the end result would be catastrophic to the person(s) involved.

We, the pedestrians and on lookers are concerned for the above mentioned problem and have tried to address the members concerned but have got no response. It's at this juncture, on humanitarian grounds, that we have decided through your esteemed paper to address the higher authorities concerned.

We hope you will consider our article and give it the first priority.

I have the honour to be,

Yours truly,

KAPO Hassan
behalf of the citizens concerned

Figure 1.4

2. 'What is up with traffic police'¹

Sir,

Kigali police recently, recently came up an idea of barring taxis from stopping at certain places like MTN and Hotel Novotel, what is the essence of doing so? Look, where should we be disembarking when we have business around these places?

Well, am not against your programmes but we kindly request you to provide a close option for us. We are suffering.

Kamuzinzi John
Kacyiru.

¹ It's an example of a public letter which appeared in the New Times paper of December 2-4, 2002, by Kamuzinzi John of Kacyiru. Page 6 column 4

3. 'Leaders should be examples'²

Sir,

I was very happy when president Kagame was attacking bad leaders at the conference of executive leaders in Gashari, Kibungo province. In fact, when a president speaks out about that, we people who being led feel so proud because our president is concerned with our well being. So my advise is to those who belong in that category of bad leaders to change the behaviour before it's too late.

It would actually help us as we work together for the betterment of our country.

Kelane E.
Rwamagana

COMMERCIAL CORRESPONDENCES

Commercial terms :

Any dealing that involves two parties in a transfer of goods or services from one of them to another for a consideration is a transaction.

For example: a sale of Primus beer involves transfer of primus beer from the seller to the buyer in consideration of sum of money (the price of primus).

If in a transaction one or both of the parties are traders, the transaction would be a *commercial transaction*. The business of any trader consists almost entirely of transactions which includes:

- the transfer of goods or services
- the payment of them

² This also serves as a public letter which appeared in the New Times paper of December 2-4, 2002, by Kalenzi E. Of Rwamagana. Page 6 column 5

A CASH TRANSACTION:

If both parts of a transaction are executed simultaneously, the transaction is said to be *cash transactions*.

A CREDIT TRANSACTION:

If the payment is made some time after the transfer of goods or services, the transaction is called *a credit transaction*.

Some of the letters involved in Business correspondences

Within businesses various letters writing are in most cases involved of which some are as shown below:

INQUIRY LETTER

This is written when a person, a company or a businessman looks for further information about a commodity or something before placing an order. An inquiry Letter is a letter written about something. It may be a general inquiry letter or a specific inquiry letter. A general inquiry letter asks for general information whereas a specific inquiry letter clearly specifies the area of particular interest.

An inquiry letter on a headed paper

Kigali Independent University
P.O Box 2280
Kigali Rwanda

09 /02 / 2003

Our ref: ULK/6598/04/03

Your ref:.....

IKIREZI BOOK STORES

Kigali Rwanda

Dear Sir,

RE : INQUIRY

We're interested in extending the range of educational books in our library. We should be grateful if you will kindly furnish us with relevant information about any educational books offered by you.

We hope to place substantial orders and would like delivery before the commencement of the next academic session of the universities.

Your kind response will be greatly appreciated.

Yours faithfully

MUSONERA K. Andrew
General Secretary

Figure 1.5

Reply to an inquiry letter

A successful inquiry letter in most cases is replied to by a reply to Inquiry letter, accompanied by either a catalogue or a price list.

Catalogue:

This is just brochure (small pamphlet or book) in which types of commodities offered by wholesalers are listed or written.

Price List:

A pamphlet which contains the list of items with their prices respectively.

An example of a reply to inquiry letter

IKIREZI BOOK STORES
Kigali Rwanda

06/05/2003

our ref: IKBS/95/04/03

your ref: ULK/6598/04/03

The General Secretary
Kigali Independent University
P.o.box 2280
Kigali Rwanda

Dear Sir,

RE: Reply to Inquiry

We are pleased to receive your inquiry letter reference Number ULK/6598/04/03, dated 2nd April, of this year, which requested for further information about the University Education books stocked by us. Please kindly receive the enclosed Catalogue No.145 and our recent price list that we hope will avail you with all the necessary information that you require.

We look forward to working with you

Yours sincerely

Maama Mia
Sales representative

ORDER LETTER (PLACING AN ORDER)

Having received all the Necessary information, a businessman, company or retailer may now proceed to placing a formal order. In other words an order letter is given by the buyer to the seller, authorising the seller to send or to make arrangements for sending or to deliver the goods prescribed in the order letter to the buyer.

When ordering for something, the only way to receive the exactly what you want is to be as specific as possible. It would be advisable to cover most of the following points:

- Description: The type of goods.
- Catalogue number /price list, quotation.
- Size
- Colour
- Other specifications : in books ; Edition, written by ; etc
- Quantity: how many or much; The number of goods needed.
- Place to which the goods are to be delivered.
- Means of delivery: To be transported by what means (by road / air).
- Date of the order
- Signature

'It is always dangerous to assume that your correspondent will just late anything for granted .'³

Some people / businessman tend to place verbal orders but this is not **advisable**, as there will be no evidence to show in case the supplier **delivers** goods or things which do no correspond to the order. So, a written order is the most widely used document.

³ Doubleday and company, Revised edition, Letter writing (1942) Page 168.

1.The order letter

Kigali Independent University
P.O. Box 2280
Kigali Rwanda

09/01/03

Our ref: ULK/9868/05/03
Your ref:IKBS/95/03

The sales manager
Ikirezi Book Stores

Dear Sir,

RE: ORDER

Referring to your Catalogue Number 698 and the price list which you sent to us,

Please supply us the following:

- 100 copies of Business English book. K.S. Edition 1997
- 120 copies of modern accounting
- 300 reams of ruled papers (foolscap)

Kindly send these books by your vehicle. The books are needed immediately.

Yours Faithfully

MANZI A. Mubaraka
General Secretary

Figure 1.6



Credit status inquiry Letter

This letter is written if a buyer wishes to be extended with credit facilities and it is the first transaction between the two businessmen and the seller would like to make sure that the buyer is credit worthy. So the seller may write to the bankers of the buyer, other suppliers, other customers or trade associations to get confidential report on the credit worthiness of his prospective customer.

Let us assume that the seller wrote to the buyers Bankers.

KB Whole Sales Ltd
PO Box 1345
Kampala Uganda

12 /02/ 2003

Our ref: KB/3541/05/03

Your ref:

The Manager
Kigali Commercial bank
PO Box 135
Kigali Rwanda

Sir,

RE : CREDIT STATUS INQUIRY

We have been approached by JJ Book Stores PO Box 2222 Kigali for extensions of credit facilities. They intend to place orders in the range of 5 000 000 Fr to 10 000 000 per month.

The proprietor has given your name as reference we should be grateful if you kindly let us have at your earliest convenience, a report on their financial position and suitability for such credit facilities. We wish to assure you that your report will be treated with strict confidence and will not be revealed to JJ Book Stores.

A self addressed and stamped envelope is enclosed for your use

Yours faithfully

Signature

KAYIRANGA Fredrick
Sales Manager

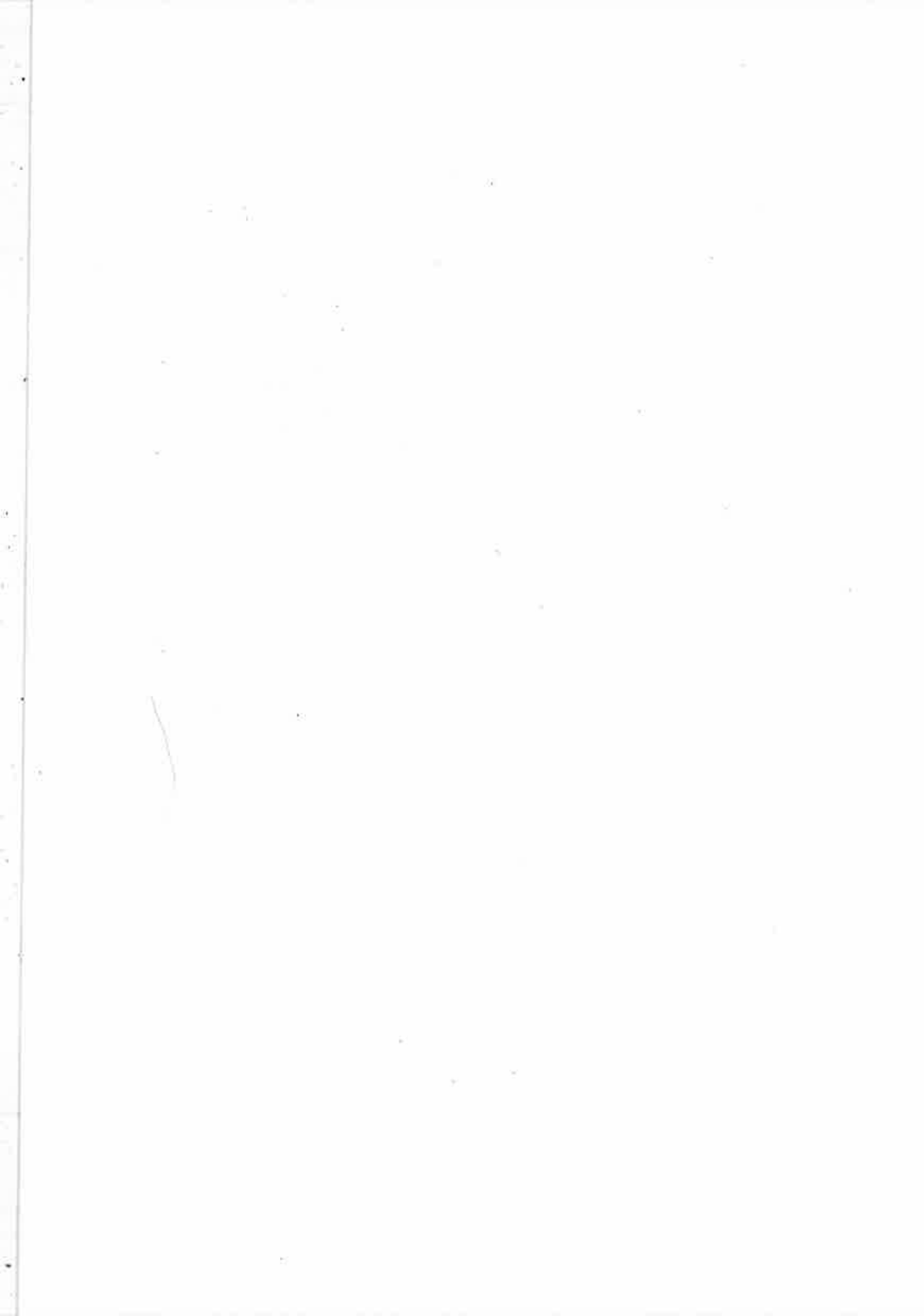
Figure 1.7

Conclusion

Correspondences are part of every educated person whereby he/she needs to carry out letter writings for the day to day life. Therefore it would be better and advantageous to our readers to try to consider the outstanding points in this article as far as letter writing is concerned

Bibliography:

1. Anna Maria Malkoc, Rationale and models for social and business correspondences
2. Doubleday and company. Letter writing revised edition, (1942)
3. New Delhi, Functional English for University students
4. The Rwandan New Times newspaper of December 2-4, 2002.



**EFFICACITE INTERNE DU SYSTEME
EDUCATIF RWANDAIS : CAS DU NIVEAU
PRIMAIRE.**

Par CCA NGABOYISONGA Roch de la Faculté
des Sciences Economiques et de Gestion de l'U.L.K.

Introduction

L'efficacité d'un système éducatif est sa capacité à réaliser les objectifs qui lui sont assignés. Ces objectifs peuvent être par exemple de scolariser à 100% la population au niveau primaire et un certain pourcentage au niveau secondaire et universitaire. On pourrait affirmer à ce sujet que le système d'éducation japonais est efficace parce que 92% de petits japonais sont au jardin d'enfants où commence le processus de scolarisation. Et puis, tout le monde au Japon suit au moins neuf ans d'enseignement. 90% au moins de la population finit le lycée, taux bien supérieur à celui des Etats Unis, de l'Angleterre et de la plupart d'autres pays. Aujourd'hui, le Japon a un taux de d'analphabétisme presque nul (0,7%)¹.

Au niveau de la nation, il est important de savoir si l'éducation a des effets sur l'économie nationale, par exemple en contribuant à l'augmentation du revenu nationale. Là, il s'agit précisément de l'efficacité externe macroéconomique. Aussi, on ne peut affirmer que les objectifs d'un système éducatif sont atteints si les éduqués n'ont pas reçu les acquisitions cognitives et non cognitives nécessaires.

Les acquisitions cognitives se subdivisent en trois catégories :

1. Capacité de mémorisation des faits scientifiquement acceptables : avoir des connaissances encyclopédiques ;
2. Capacité de réaliser des opérations simples : opérations pour lesquelles on a toutes les informations nécessaires pour trouver la réponse au problème posé ;
3. Capacité de réaliser des opérations complexes pour lesquelles on doit chercher les informations et pouvoir en comprendre la nature.

Les acquisitions non cognitives les plus importantes sont :

¹ Löwenthall P., Préparer le XXIème siècle, Editions Odile Jacob, Paris 1994, P174

1. Le développement de la personnalité pour rendre les individus plus capables d'autonomie, créativité, entrepreneurship, estime de soi, etc.
2. Préparer les citoyens à la vie en communauté, connaissance des valeurs et des lois, etc.
3. Développement du corps par l'éducation physique, la musique, la danse, etc.

La notion d'efficacité implique que l'on puisse considérer le processus éducatif comme un processus de production des services éducatifs à l'intention d'une sous population donnée. Un système de production en éducation est dit efficace si à partir d'un montant donné de ressources utilisées appelées « inputs », il permet d'obtenir un maximum d'unité de produits appelé « output ».

Dans le contexte actuel du Rwanda, il y a lieu de s'interroger sur l'efficacité de son système d'enseignement connaissant les conditions dans lesquelles fonctionnent les écoles et les universités. Cette efficacité, nous allons l'analyser en termes de taux de scolarisation et en termes d'abandons et de redoublements. En effet, si nous considérons les membres d'une cohorte comme le produit d'une chaîne de transformation, par exemple les élèves « sortants » à la fin d'un cycle peuvent être étiquetés comme produit fini, tandis que les élèves qui abandonnent sont considérés comme une perte. Plus une cohorte aura utilisé beaucoup d'années/élèves au cours de son séjour dans le système, plus il y a eu des pertes, et plus le système est inefficace.

La définition des outputs en éducation pose cependant problème étant donné la difficulté d'évaluer l'acquisition cognitive des élèves. Il convient aussi de distinguer l'efficacité qualitative de l'efficacité quantitative. L'efficacité quantitative est celle qui utilise le concept de flux d'élèves (exemple : taux de redoublement, d'abandon, de participation, de scolarisation, ...) pour mesurer l'efficacité. On s'intéresse au fait que les individus qu'on veut éduquer sont ou non présents dans le système éducatif. L'efficacité qualitative par contre consiste à vérifier si le processus d'apprentissage fonctionne de

manière à permettre aux éduqués d'acquérir les connaissances recherchées. On l'évalue par le taux de réussite aux examens et surtout aux tests standardisés, bien que ces deux indicateurs ne soient pas toujours sans subjectivité.

Tout pays, avant de penser à l'efficacité qualitative doit atteindre d'abord les objectifs quantitatifs. Le Rwanda essaie aussi de répondre à cet objectif primordial et indispensable même dans la recherche de l'efficacité externe de son système éducatif.

I. Efficacité interne au niveau macroéconomique

I.1. Approche par le taux de scolarisation

Le taux de scolarisation constitue le moyen le mieux indiqué pour se rassurer que le système éducatif d'une nation est atteint en termes d'augmentation du nombre de personnes éduquées. Toutes les nations du monde ont justement en la matière l'objectif d'augmenter sensiblement ces taux. C'est ainsi d'ailleurs que dans les années 1950 et 1960, presque tous les pays du monde, mus par deux convictions (l'une d'ordre politique et l'autre d'ordre économique) furent amenés à élargir les objectifs de l'éducation formelle. La conviction d'ordre politique était, en gros la suivante : l'expansion massive de l'éducation est le meilleur moyen, pour un Etat qui se veut démocratique, de réduire les disparités flagrantes enracinées dans les préjugés du passé et les injustices socio-économiques.

En termes de programme, cela signifiait l'enseignement primaire universel pour les enfants, la possibilité accrue pour les jeunes d'accéder à l'enseignement secondaire et supérieur, les cours d'alphabétisation pour les adultes. La conviction d'ordre économique ne faisait que corroborer la précédente et était elle-même étayée par des études réalisées dans un domaine relativement

nouveau qui était celui de l'Economie de l'Education. Cette dernière donnait une sorte de vernis scientifique à la notion selon laquelle la croissance d'un pays résulterait notamment du progrès technique et de l'augmentation de la productivité de la main d'œuvre qui à leur tour découleraient du développement de l'éducation formelle². C'est ainsi que le gouvernement rwandais s'intéresse aussi à l'éducation de sa population. La moyenne du budget alloué à l'éducation pour les dix dernières années est de 20,1%³. Ceci permet tant soit peu de réduire les taux d'analphabétisme.

La préoccupation du gouvernement rwandais en la matière est très visible surtout ces dernières années. Les taux de scolarisation sont passés entre 1998 et 2002 de 69,9% à 78,3%. Ces taux ont évolué de la manière suivante entre ces deux périodes :

Tableaux n° 1 : Evolution des taux de scolarisation de 1998 à 2002

Année 1997/1998	Année 1998/1999	Année 1999/2000	Année 2000/2001	Année 2001/2002
69,9	72,2	73,3	74,5	78,3

Source : Ministère de l'Education, Direction des Etudes et Planification de l'Education

² Coulombs Ph., La crise mondiale de l'éducation, Ed. De Boek Université, Bruxelles, 1989. P46

³ MINECOFIN, 2001

Nous voyons bien que ces taux ont augmenté d'années en années, ce qui montre bien que le gouvernement rwandais essaie, en dépit des difficultés économiques énormes de s'intéresser à la formation de la jeunesse. Si donc les taux de scolarisation constituaient le seul critère d'efficacité interne, nous dirions que le système éducatif rwandais est efficace.

I.2. Approche par l'espérance d'années d'études à 5 ans

C'est un nouvel indicateur qui permet une comparabilité globale. On appelle encore l'espérance d'années d'études à 5 ans, le nombre d'années qu'un enfant de 5 ans peut espérer passer à l'école s'il connaît un destin scolaire identique à la moyenne. L'âge de cinq ans est celui pris en considération parce que dans plusieurs pays, le préscolaire est suffisamment développé et le primaire commence à 6 ans. La formule utilisée pour le calcul de cet indicateur est la suivante :

$$Esp.5ans = \sum_{t=1}^{23} Probécole$$

Dans cette formule,

Esp. 5 ans= espérance d'années d'études à cinq ans

Probécole= probabilité d'être à l'école

La méthode utilisée par l'UNESCO consiste à exprimer en pourcentage la sommation des taux de scolarisation à différents âges.

Nous n'avons pas pu calculer cette variable pour le cas du Rwanda parce que les données sur les taux de scolarisation à différents âges ne sont pas disponibles.

Selon les données de l'UNESCO de 1995⁴, l'espérance d'années d'études en 1992 était respectivement de 8.8 années dans les pays en voie de développement et de 13.9 années dans les pays

⁴ UNESCO: World Education Report, 1995

développés. L'espérance de vie à l'école des enfants rwandais devrait toutefois être proche de celui de la moyenne pour les pays en voie de développement. Généralement, les pays où les taux de scolarisation sont les plus élevés sont ceux dans lesquels l'espérance d'années d'étude à cinq ans est le plus élevé aussi. En République démocratique du Congo, ce taux est de 9,5% pour l'année 2000. Ces moyennes cachent des fortes disparités, l'espérance d'années d'études étant par exemple dix fois plus important au Canada (17.5 années) qu'au Mali (1.5 années).

II. Efficacité interne au niveau micro économique

La mesure de l'efficacité à ce niveau se fait en analysant les taux de déperdition (abandons et redoublements). L'analyse en termes de nombre de diplômés et des acquisitions cognitives et non cognitives est aussi indispensable à ce niveau pour évaluer les performances éducatives des systèmes. En effet, fréquenter l'école ne signifie guère avoir une chance égale d'apprendre et de réussir. La conférence mondiale sur l'éducation pour tous, tenue à Jomtien (Thaïlande) en mars 1990 a remis en cause la notion même de l'éducation formelle, de la pédagogie et des processus d'apprentissage dans le but d'ouvrir des nouvelles perspectives en matière de qualité de l'éducation pour tous⁵. Pour le cas du Rwanda, la vérification des acquisitions cognitives et non cognitives qui est le seul moyen pour évaluer l'efficacité qualitative est difficile étant donné qu'il n'existe dans le pays aucune statistique dans ce sens.

Les seuls chiffres dont nous avons pu disposer et pouvant permettre de conclure sont les pourcentages de réussite aux tests standardisés organisés à la fin des études primaires.

Vinayagum C. et Alii, L'éducation de qualité pour tous, inédit, Prétoria, 2000, P16

II.1. Efficacité quantitative

La mesure de ce **type** d'efficacité utilise le concept de flux d'élèves pour mesurer l'efficacité. Ce flux concerne les abandons et les redoublements qui constituent le phénomène de déperdition. Pour Psacharopoulos et Woodhall⁶ les abandons et les redoublements dans les pays en développement sont plus fréquents parmi les enfants d'origine modeste, plus nombreux dans les zones rurales que dans les zones urbaines et chez les filles que chez les garçons. Leurs causes principales sont : l'absentéisme, la pauvreté qui engendre la maladie, la malnutrition, et l'importance des coûts d'opportunité pour les familles modestes, les facteurs culturels qui affectent les filles en particulier ; des programmes et des examens inappropriés. En Malaisie par exemple, les établissements secondaires qui obtiennent les meilleurs résultats ne dépensent pas plus que la moyenne, et les établissements les plus coûteux ne sont pas ceux qui ont les meilleurs résultats. Il est par conséquent évident qu'on ne peut utiliser les **inputs** comme des mesures de la qualité des résultats.

Dans cette étude, nous recourons au modèle de flux d'élèves. La spécification des relations entre flux d'élèves va nous permettre de reconstituer la progression d'une cohorte d'élèves à travers le système scolaire et d'évaluer les déperditions dues aux abandons et aux redoublements.

a) Les redoublements

Les redoublants sont les élèves qui reprennent une classe au bout d'une année scolaire ou académique. Le redoublement sur le plan conceptuel consiste par exemple à faire en deux ans, ce que les autres font en un an. L'année de redoublement est ignorée dans le calcul des indicateurs quantitatifs. Le redoublement est considéré comme une déperdition parce que les redoublants diminuent la

⁶ Psacharopoulos G. et Woodhall M., « L'éducation pour le développement : une analyse des choix de l'investissement », Economica, Paris, 1988

capacité d'accueil (offre de places) de leur classe, empêchant ainsi d'autres enfants de s'y faire admettre ou provoquant un surpeuplement, ce qui accroît le coût de l'enseignement.

Le calcul du taux de déperdition par redoublement soulève un problème d'informations statistiques. Le redoublement, seconde clé qui intervient dans la déperdition scolaire contribue fortement à l'augmentation du coût de l'enseignement. Durant les 6 dernières années, les redoublements ont évolué de la manière suivante :

Tableau n°2: Evolution des effectifs globaux et des redoublants d'étude du cycle primaire au Rwanda

	Année 1996 /1997	Année 1997 /1998	Année 1998 /1999	Année 1999 /2000	Année 2000/2001	Année 2001/2002	Année 2002/2003
Admis	791 615	898 174	913 635	945 754	943 693	1 066 509	1 366 225
Redou- blants	229 057	288 570	375 028	485 903	532 579	468 001	270 338
Taux	28	32,1	38,1	37,6	31,8	17,2	

Source : Ministère de l'Education, Direction des Etudes et Planification de l'éducation

Les taux de redoublement ci-dessus ne nous permettent pas de tirer une conclusion tranchée. Mais l'idéal pour le gouvernement serait d'avoir des taux de redoublement beaucoup plus réduits.

Le taux de redoublement de l'année d'étude « i » pour l'année scolaire « t » est égal au nombre d'élèves qui redoublent l'année « i » pendant l'année scolaire t+1 divisé par l'effectif total d'élèves de l'année « i » pendant l'année scolaire t¹.

¹ Matingu N., Planification des ressources humaines, Cours inédit, Deuxième Licence Démographie, UNIKIN, 1993

C'est à dire :

$$r'_i = \frac{R'_i}{E'_i}$$

r'_i = le taux de redoublement pour l'année d'étude « i » de l'année scolaire « t ».

R'_i = Effectif d'élèves qui redoublent pour l'année d'études « i » de l'année scolaire t+1

E'_i = Effectifs des élèves inscrits en l'année d'étude « i » pendant l'année scolaire t

Cette formule est applicable pour trouver les taux correspondant à toutes les années.

b) Les abandons

L'abandon est le fait de quitter un système scolaire avant la fin du cycle. L'idée derrière, c'est que quand on quitte le cycle avant d'arriver au bout, c'est à dire que les objectifs du cycle n'ont pas été atteints. Les années sont considérées comme une perte. Le cas le plus fréquent, c'est que pour le primaire, il y a risque de devenir analphabète. Il y a une forte corrélation entre le fait d'abandonner et le fait d'être faible. Ceux qui abandonnent sont en moyenne faibles. L'abandon peut avoir lieu soit pendant l'année scolaire, soit entre les années scolaires.

Les planificateurs et les statisticiens admettent en outre que dans le cheminement des élèves à travers le système scolaire, il y a une série d'étapes à franchir en un temps déterminé. Autrement dit, ils admettent, à moins que le plan national d'éducation ne prévoie le contraire, que le but des élèves admis dans un cycle déterminé est d'arriver dans les délais prescrits au bout de ce cycle. L'idée alternative c'est qu'un élève qui abandonne ses études peut avoir déjà acquis une instruction non négligeable de telle sorte que du point de vue pédagogique, il serait illogique de considérer que les

efforts accomplis pour l'instruire ont été entièrement perdus. Considérés par rapport aux intentions qui président à l'organisation et à la définition des objectifs de l'enseignement, l'abandon représente alors une déperdition. Il existe une série d'abandons, et de stagnation des élèves puisqu'une partie seulement d'élèves couvrent normalement leur scolarité.

Nous pouvons calculer les taux d'abandons connaissant le nombre d'abandons pour une ou plusieurs années. Le taux d'abandons pour l'année d'études « i » de l'année scolaire « t » est égal au nombre d'élèves ayant quitté l'école lors de l'année « i » pendant l'année scolaire « t » divisé par l'effectif total de l'année « i » pendant l'année « t » divisé par l'effectif total de l'année i pendant l'année « t » c'est à dire :

$$a_i^t = \frac{E_i^t - (P_{t+1}^{i+1} + R_i^{t+1})}{E_i^t}$$

a_i^t = le taux d'abandon pour l'année d'étude « i » de l'année scolaire « t » ?

P_{t+1}^{i+1} = effectif de nouveaux élèves de l'année i+1 pour l'année scolaire t+1

R_i^{t+1} = effectif d'élèves qui redoublent l'année d'étude « i » pendant l'année scolaire t+1

E_i^t = effectif d'élèves inscrits en l'année « i » pendant l'année scolaire « t ».

Le taux de promotion est le complément à 1 du taux de déperdition (l'ensemble du taux de redoublement et du taux d'abandon), c'est à dire :

$$P'_i + r'_i + a'_i = 1$$

Le taux de promotion de l'année « i » pour l'année scolaire « t » est le rapport entre l'effectif de nouveaux élèves de l'année d'études i+1 pour l'année scolaire t+1 sur l'effectif d'élèves de l'année d'étude i pendant l'année scolaire t. C'est à dire :

$$P'_i = \frac{P'_{i+1}}{E'_i}$$

Où P'_i = taux de promotion de l'année d'étude i pour l'année scolaire t

P'_{i+1} = effectif de nouveaux élèves de l'année scolaire t+1

E'_i = effectif des élèves inscrits en l'année d'études i pendant l'année scolaire t

Le tableau suivant donne l'évolution de ces différents taux :

Tableau n°3 : Evolution des taux de promotion, de redoublement et d'abandon pour l'enseignement primaire au Rwanda (1996 à 2002)

	1996/97	1997/98	98/2000	1999/2001	2000/2002	2002/2003
Taux de promotion	66,6 %	56,2%	50,5%	49,6%	54,0%	66,2%
Taux de redoublement	28,0%	32,1%	38,1%	37,6%	31,8%	17,2%
Taux d'abandon	5,5%	11,7%	11,4%	12,6%	14,2%	16,6%

Source: MINEDUC, Direction des Etudes et Planification de l'Education.

Nous remarquons que les taux d'abandons ont en général évolué dans un sens positif, ce qui signifie que le nombre d'élèves qui abandonnent augmente d'années en années. Des améliorations sont nécessaires à ce niveau. Il en est de même d'ailleurs pour les taux de redoublement.

Considérons les taux de promotion, d'abandon et de redoublement du tableau précédent. Nous constatons que ces taux prennent des valeurs sensiblement différentes d'une année à l'autre. Il est probable que ces taux fluctuent autour d'une moyenne pour chaque type de taux et que ces moyennes soient caractéristiques de la dynamique du système scolaire rwandais. Comme nous résumons une série de rapports entre plusieurs populations, nous utilisons la moyenne géométrique. La moyenne géométrique G d'une série $x_1; x_2; \dots; x_n$ est égale à :

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$$

Ou encore :

$$\text{Log } G = \frac{1}{n} (\log x_1 + \log x_2 + \dots + \log x_n)$$

Du tableau ci-dessus nous obtenons les moyennes géométriques des différents taux :

Tableau n° 4: Les moyennes géométriques des taux de promotion, de redoublement et d'abandon de 1996 à 2002

	Moyenne géométrique
Taux de promotion	369
Taux de redoublement	247
Taux d'abandon	205

Source : Calculs sur base du tableau n°3

Pour trouver ces moyennes il a suffi de faire la somme des log selon qu'il s'agit des redoublement, des promotions ou des abandons du tableau précédent. La somme totale trouvée est à diviser par le nombre d'années ayant fait l'objet de l'étude.

Pour trouver les moyennes géométriques des redoublements par exemple il suffit de procéder de la manière suivante :

$$\text{Log } G = \frac{1}{n} (\log 280 + \log 321 + \dots + \log 172)$$

G est la moyenne géométrique

$$G = 247$$

Ce chiffre signifie que pour la période étudiée, le taux de redoublement est de 247 pour mille.

11.2. Efficacité qualitative : évaluation des acquisitions

A ce niveau, on s'intéresse à ce que les individus ont acquis à l'école. Il s'agit de vérifier que le processus d'apprentissage fonctionne, c'est à dire que les apprenants apprennent réellement. On s'intéresse ici à ce que les individus ont appris à l'école et non au nombre d'années. Cette efficacité peut s'évaluer de deux manières différentes :

a) Le succès aux examens

Le défaut de cette mesure est que les examens ont une partie aléatoire de mesure parce qu'il peut y avoir une certaine subjectivité des enseignants qui peuvent avoir des critères de sélection propres à eux. La fixation de la barrière varie d'un enseignant à l'autre. Un autre problème, c'est qu'il n'existe pas une mesure d'appréciation stable. Il peut y avoir des pressions politiques pouvant pousser les enseignants à être beaucoup plus larges. La corruption aussi n'est pas à exclure qu'elle soit matérielle ou morale. Le taux de promotion nous donne une idée sur le succès aux examens.

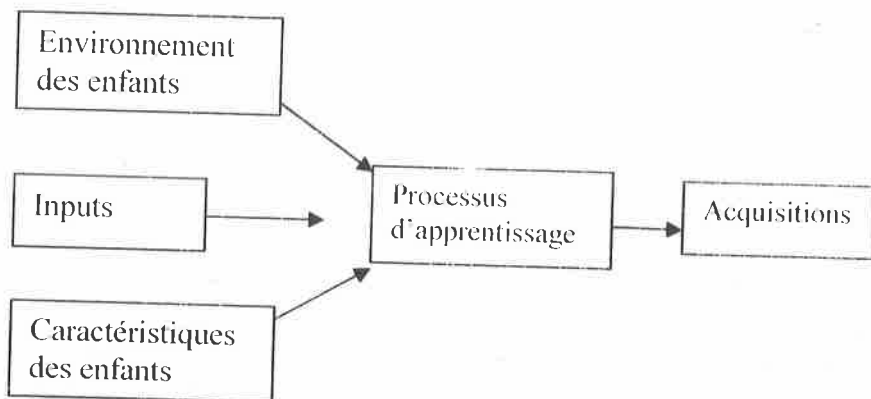
b) Les test standardisés

Ils permettent de corriger certaines erreurs des examens par :

1. Une comparaison objective des différences d'acquisition des enfants de même cohorte ;
2. Une comparaison objective de la difficulté des épreuves ;
3. Des comparaisons au niveau international.

Par ailleurs, les acquisitions non cognitives sont difficiles à évaluer par les tests standardisés. Aussi, ils ne permettent pas de suivre un raisonnement parfait. En outre, en rapport avec le processus d'apprentissage, il n'est pas facile d'expliquer les différences d'acquisition entre les élèves. Il convient de noter

d'abord que les caractéristiques des enfants ne sont pas homogènes et que l'environnement des enfants varie (au sens où ils peuvent être d'une famille dans laquelle les parents sont ou non éduqués). Les caractéristiques des enfants mesurent leurs aptitudes. Par exemple, le quotient intellectuel et les acquis antérieurs. Schématiquement, nous pouvons représenter le processus d'apprentissage comme suit :



Les inputs sont constitués des enseignants, de matériel pédagogique, des moyens financiers, des aides, de la qualité des programmes , ... Seuls les inputs ne peuvent expliquer les acquisitions des élèves. Le processus d'apprentissage est un processus cumulatif. Les connaissances antérieures ont des effets sur l'apprentissage actuel. La faiblesse en français ne peut être attribuée à une seule année d'apprentissage. Le niveau aujourd'hui dépend de ce qu'on a appris depuis les années précédentes. En conséquence, un test est nécessaire au début de chaque année pour connaître le niveau des élèves.

Ainsi, on peut éviter le piège en mesurant les acquisitions par la valeur ajoutée pour une période où les élèves ont eu les mêmes inputs. On note ici qu'aucune étude régionale n'a jusqu'ici été entreprise pour évaluer les acquis à partir d'un seuil minimum de connaissance, ni d'aptitude et de compétence des apprenants par

rapport à un niveau désiré pour tous en utilisant une approche évaluative. L'évaluation des acquis des élèves dans les trois domaines d'apprentissage (lecture, écriture, mathématique et compétence dans la vie courante) a commencé en 1928⁸

Le seul test standardisé organisé dans le pays et permettant de vérifier le niveau d'acquisition des élèves au Rwanda est l'examen national organisé chaque année pour connaître le niveau des élèves. Ainsi, on peut éviter le piège en mesurant les acquisitions par la valeur ajoutée pour une période où les élèves ont eu les mêmes inputs. Tous les tests du primaire sont coordonnés au niveau national.

Voyons comment ont évolué les résultats des examens nationaux ces dix dernières années.

Tableau n°5 : Résultats des examens nationaux pour les années 1998 à 2002

Année Scolaire	Participants	Réussites	Echecs	% des réussites
1998/1999	64.888	14.633	502555	22,55
1999/2000	59.689	16.137	43552	27,03
2000/2001	63.931	16.612	47319	26,07
2001/2002	79.226	19.061	60165	24,06

Source : Ministère de l'Education

Ces résultats sont en moyenne moins satisfaisants si l'on compare les pourcentages de réussite à ceux d'échec. Donc si l'on se base sur le critère de tests standardisés à l'école primaire, on peut dire que le système rwandais d'éducation primaire est inefficace.

8 Kennedy P., Op.Cit. P 73

En dépit des insuffisances des tests standardisés, il paraît incontournable de continuer à les organiser parce qu'ils permettent notamment d'avoir une idée générale sur le niveau des élèves au niveau national et de comparer les performances des élèves selon les écoles d'origine. Ils permettent également de comparer les performances au niveau international. C'est après ces tests d'intelligence standards qu'on est arrivé à la conclusion que le niveau moyen des enfants japonais est plus élevé par rapport à celui des américains et des européens. L'étudiant moyen japonais obtient 117 points contre 100 points pour les américains et les européens⁸.

Une notion proche de celle relative à l'efficacité qualitative est celle de la productivité du système éducatif mais cette notion se rapporte beaucoup plus à l'efficacité externe. En effet, des bonnes acquisitions cognitives devraient aboutir à une augmentation de la productivité du système des personnes éduquées. Justement, l'un des objectifs reconnus à un système éducatif, c'est l'amélioration de l'efficacité des personnes éduquées au travail. Pour mesurer l'efficacité productive de l'éducation, autrement dit son rendement (r), on mesure la différence de productivité entre un individu éduqué (Yed) et un individu non éduqué (Ynoed) et le coût de l'éducation de l'individu éduqué (C)⁹.

$$R = Y_{ed} - Y_{noed} - C$$

⁸ Kennedy P., Op. Cit., P75

⁹ Hugon P. et Alii, *Crise de l'éducation en Afrique*, trimestriel n° 172, Afrique contemporaine, octobre-décembre, 1994, P94

Une autre mesure couramment employée est celle de « taux de rendement de l'éducation » :

$$T = \frac{Y_{ed} - Y_{noed}}{C}$$

Comme la productivité est rarement observée en pratique, elle est évaluée par le salaire et les outputs du travail productif de l'individu, lequel est une mesure de la productivité si le marché du travail est à l'équilibre. La littérature fait la distinction entre le rendement social et le rendement privé de l'éducation. Les détails de cette notion de rendement de l'éducation sont analysés dans l'étude de l'efficacité externe micro économique.

Conclusion

Certaines théories économiques de l'éducation arrivent aux conclusions selon lesquelles la relation n'est pas toujours évidente entre efficacité des systèmes éducatifs et les inputs scolaires.

Hanushek¹⁰ fit remarquer que près de 400 études avaient été menées dans le temps et avaient abouti aux résultats que les inputs scolaires ne jouent pas dans le processus d'acquisition des élèves. Le rapport accablant de Coleman au congrès américain en 1966 tournait presque dans le même sens. Ce rapport affirmait que les inputs aux Etats Unis n'avaient pas d'influence sur les outputs scolaires en termes d'acquisitions et de rendement. Il ne nie pas par contre l'idée de l'existence d'inputs primaires incontournables pour que les élèves soient performants. Il cite par exemple l'influence de la famille d'origine et les caractéristiques des enfants (leur QI et leur assiduité). Hanushek ajoute à cela le fait que le processus d'apprentissage est un processus cumulatif. Les performances actuelles des apprenants sont influencées par les performances passées.

Toutefois, il existe un minimum d'inputs sans lesquels le processus d'apprentissage serait difficile. Un effort dans ce sens a été fourni par le gouvernement rwandais. C'est ce qui explique la place assez importante des dépenses d'éducation dans le budget de l'Etat. Les résultats en termes d'efficacité interne en dépendent. A ce propos justement, nous constatons que les taux de scolarisation au Rwanda sont acceptables par rapport aux ressources financières du pays. Les résultats des tests standardisés semblent pourtant médiocres. Des efforts doivent aussi être réalisés pour réduire les taux d'abandons et même les taux de redoublement. Cela permettrait de réduire le taux de déperdition en général. Des fortes déperditions signifieraient un gaspillage des ressources financières pour le pays. Par exemple, le redoublement voudrait dire dépenser deux fois pour un an.

¹⁰ Hanushek A., « Accessing the effects of school resources on students performances : an update », 1997

BIBLIOGRAPHIE

1. Coulombs P., La crise mondiale de l'éducation, Ed. De Boek Université, Bruxelles, 1989
2. Hanuchek A., « Accessing the effects of school resources on students performances : an update », 1997
3. Hugon P. et Alii, Crise de l'éducation en Afrique, trimestriel n° 172, Afrique contemporaine, octobre- décembre, 1994
4. Löwenthall P., Préparer le XXI^{ème} siècle, Editions Odile Jacob, Paris 1994
5. Matingu N., Planification des ressources humaines, Cours inédit, Deuxième Licence Démographie, UNIKIN, 1993
6. Ministère de l'éducation, Rapports de la direction des études et planification de l'éducation.
7. Psacharopoulos G. et Woodhall M., « L'éducation pour le développement : une analyse des choix de l'investissement », Edition Economica, Paris, 1988
8. Vinayagum Ch. et Alii, L'éducation de qualité pour tous, inédit, Prétoria, 2000

**LE ROLE DE L'EDUCATION
ENVIRONNEMENTALE DANS
LE DEVELOPPEMENT DURABLE
DU RWANDA**

Par BOAZ KAGABIKA

I. Introduction :

Au Rwanda comme dans la plupart des pays du monde, les problèmes d'environnement sont à la une et leurs solutions ne peuvent se concevoir sans intégrer pratiquement une connaissance capable de mettre en relation des processus biogéochimiques, des procédés pédagogiques et techniques, mais aussi des rapports socio-économiques et institutionnels. La référence à la trilogie « *population, environnement et développement* » est actuellement indispensable pour appréhender les contraintes liées à la protection de l'environnement. Une prise de conscience sur l'exploitation rationnelle des ressources naturelles est une des conditions sine qua none pour une bonne gestion durable de l'environnement. D'une manière générale, l'opinion mondiale est de plus en plus sensible à la dégradation de l'environnement et des dangers que l'humanité encourt.

Dans son article : *Mettre les connaissances des sages à l'usage des pauvres*, Martin W. Holdgate, ancien Directeur administratif du programme des Nations Unies pour l'environnement (PNUE), souligne : « *Il y a peut-être un don gratuit fait à l'humanité mais ce don est d'une valeur incalculable, et si certaines de ses parties sont brisées ou endommagées, elles s'avèrent très coûteuses à les réparer. Et les signes de cette rupture sont multiples à partir du siècle passé, tout d'abord dans les rivières salées, les ciels obscurcis et les émissions suffocantes du monde industriel et ensuite, de manière plus subtile et significative dans les progressions des déserts, l'érosion des sols, le recul des forêts, les modifications atmosphériques et la baisse de rendement de la pêche qui menacent ensemble des millions d'habitants de la planète, riches ou pauvres* »¹.

Le but poursuivi dans cette publication est de pouvoir interpeller constamment la conscience de nos concitoyens en matière

¹ HOLDGATE.M : Mettre les connaissances à l'usage des pauvres, in : *Mazigira* Vol. 8 n°5 Septembre-Octobre 1985, p7

d'environnement dans ses dimensions multiples, et plus particulièrement sociales en vue de susciter l'engouement de tout rwandais au modèle du développement durable par une éducation environnementale.

2. Qu'est-ce que l'environnement ?

Brièvement dit, l'environnement est tout ce qui nous entoure ; c'est-à-dire l'ensemble des éléments naturels et artificiels qui permettent le bon déroulement de la vie sur la Terre. Cette définition englobe trois dimensions :

- 1° La dimension physique qui comprend la lithosphère (le sol et le sous-sol), l'hydrosphère (océan, mers, lacs, cours d'eau, ...) et l'atmosphère (gaz, humidité, température, pression, ...) ;
- 2° La dimension biologique ou la biosphère c'est-à-dire la partie de la planète occupée par les êtres vivants (communautés animales et végétales) évoluant dans l'environnement physique dont l'ensemble est appelé *système écologique* « *ecology system* » ou tout simplement écosystème ;
- 3° la dimension sociale ou la sociosphère et la technosphère c'est-à-dire le monde créé par l'homme qui peut avoir un effet sur nos vies (maison, villes, machine...) ².

Avant la conférence de Stockholm en 1972, le concept d'environnement n'était considéré sous l'angle des composantes biophysiques, c'est-à-dire un réservoir des ressources que l'homme peut en tirer ; mais dans l'état actuel des choses, cette approche est incomplète et ne répond plus aux exigences de la réalité vécue. Pour avoir le sens le plus complet, il a fallu ajouter au concept d'environnement la dimension sociale qui est la plus déterminante puisqu'elle tient compte des considérations d'ordre économique,

² UNESCO-PNUE : Concepts de base de l'éducation environnementale in : *Connexion, bulletin de l'éducation relative à l'environnement*, Unesco-Pnue, Vol. XV, n°2, Juin 1990, p2

social, culturel, technologique, éthique, religieux, esthétique, etc. Comme le dit Ahmed El Hattab, ce sont les conférences de Stockholm (Suède) sur l'environnement humain en 1972 et de Tbilissi (en Géorgie en Ex-URSS) sur l'éducation environnementale en 1977 qui ont donné au concept d'environnement sa valeur exacte en lui conférant un caractère de globalité³ (approche systémique ou holistique). Etant donné qu'on ne peut parler de l'environnement sans l'homme, désormais, du moins sur le plan théorique, la relation homme → environnement allait s'inverser en homme ↔ environnement. Dans le premier cas, la relation est unilatérale : l'environnement fournit tout à l'homme sans une moindre contre partie alors que dans le second cas, l'homme ne reçoit pas seulement de l'environnement mais il l'aménage et le protège. Avec sa formation pluridisciplinaire, l'environnementaliste joue le rôle d'"alarmiste", il éveille l'attention sur tout ce que les spécialistes des autres domaines (géographes, géologues, biologistes, économistes, sociologues, juristes...) n'ont pas vu, sur tout ce qui touche à la complexité et à la globalité des interactions du réel.

3. L'état de l'environnement au Rwanda

Le Rwanda, pays enclavé, situé à cheval entre l'Afrique Centrale et l'Afrique Orientale, est une immense agglomération de 26.338 Km² dont 11% occupés par des lacs, des marais et des plaines d'inondation, 11% par des parcs nationaux, 5% par des forêts naturelles et 0,3% par des routes et pistes⁴. Il s'étend sur 1°41'et

³ AHMÉD EL HATTAB : La problématique environnement-développement in : *Actes du Séminaire sur l'environnement et le développement durable*, Barazzaville, 1989, P116.

⁴ MINITRAP : *La problématique de l'aménagement de l'espace du territoire rwandais*, étude inédite, Kigali, Sd, p8

2°51' de latitude Sud et sur 28°45' et 31°15' de longitude Est⁵. Son relief accidenté (pays de milles collines) et dominé par une chaîne volcanique, a une altitude comprise entre 1000 m et 4.500 m qui permet un étagement biogéographique des espèces aussi bien floristiques que faunistiques. C'est le cas des forêts d'altitude qui abritent les gorilles de montagne qui attirent les touristes. Le pays connaît une pluviométrie assez suffisante, un sol volcanique fertile surtout dans sa partie nord-ouest et un rythme thermique tempéré par l'altitude et restant peu variable au cours de l'année ; ce qui a attiré le dense peuplement humain.

Malgré sa beauté due à la situation géographique et aux conditions écologiques favorables, le Rwanda est confronté à la dégradation très sérieuse de l'environnement liée à la forte pression démographique. En effet, la population rwandaise connaît une très forte croissance démographique puisqu'elle est passée 1.595.500 habitants en 1934 à 7.155.391 habitants 1991⁶. En 2000, la population rwandaise était estimée à 7.700.000 habitants⁷ soit une densité moyenne de 292,6 habitants au Km². C'est le pays le plus densément peuplé d'Afrique continentale. Les analyses démographiques prévoient que la population atteindra 9.446.559 habitants en 2005 et 15.000.000 d'habitants en 2025⁸.

Cette pression démographique est donc due au fort taux d'accroissement de la population, lequel taux n'a cessé de s'accroître depuis 1950. Actuellement ce taux s'élève en moyenne à 3% alors que celui de la production agricole est à situer à 2,2%⁹. Les conséquences de cette forte pression démographique sont nombreuses et se manifestent notamment à travers les traits suivants :

⁵ SIRVEN, P& al : *Géographie du Rwanda*, Éd. A. De Bock, Bruxelles, 1974, p7.

⁶ MINITERE : *Stratégie nationale et plan d'action pour la conservation de la biodiversité au Rwanda*, Kigali, 2000, p4

⁷ MINICOFIN : *Vision 2020, étude inédite, Draft4*, Kigali, Novembre 2002, p9

⁸ MINITERE: idem, p4

⁹ MINICOFIN : idem. p9

- 1° la réduction spatiale des écosystèmes forestiers voire savanicoles causée par les défrichements et les déboisements continuels pour des fins agrico-pastorales, de constructions d'habitations, d'obtention des matières premières artisanales et médicinales et des besoins énergétiques (96,2% des ménages utilisent le bois comme sources d'énergie et 31,4% du bois comme sources d'éclairage)¹⁰. Les études menées dans ce domaine ont fait état d'une forte régression des aires protégées dont les forêts de montagne de Gishwati, de Mukura, des Volcans, de Nyungwe et la savane arborée de l'Akagera avec des taux de régression respectivement de 86,4%, de 46,7%, de 62,5% et de 66,2% de 1960 à 1999¹¹ alors qu'elles constituent des niches écologiques des espèces rares au monde, (ex. des gorilles de montagne). Les recherches récentes ont montré que certaines espèces faunistiques voire floristiques ont complètement disparu du Rwanda. Ce sont les cas du léopard du Parc national des Volcans, du lion et de l'hyène qui ont disparu de ces habitats.
- 2° la diminution de ressources en eau provoquée par la destruction du couvert végétal en amont qui favorise l'infiltration, par le drainage des marais en aval et par la pollution des rivières (cas des rivières Nyabugogo et de la Nyabarongo) et des nappes phréatiques. C'est ainsi que beaucoup d'écosystèmes humides et certains lacs comme Cyohoha Nord et Cyohoha Sud dans le Bugesera, sont en train de sécher.
- 3° la réduction des terres arables qui sont passées de 54 ares par habitant en 1964, à 28 ares en 1987 et à 23 ares par habitant en 2002. On estime qu'elles seront de 15 ares par habitant en 2013¹².

¹⁰ MINICOFIN : Les indicateurs de développement du Rwanda 2001, Kigali, 2001, p133

¹¹ MINICOFIN: Idem , p341

¹² MINITERE : *Op. Cit.* p4

4° le Rwanda étant un pays accidenté, les pluies orographiques sur les fortes pentes dénudées, provoquent l'érosion et entraînent les terres vers les vallées et les bas fonds et une grande partie est emportée en dehors du pays.

Une étude effectuée par Uwizeyimana (1986) cité par le MINECOFIN a montré que les pertes en terres dues à l'érosion sur les terres cultivées était fonction des pentes . Par exemple 34% des terres cultivées sur une pente de 0 à 5 degrés occasionnent une perte de 1,1% des terres, 16 % des terres cultivées sur une pente de 5 à 10 degrés, donnent lieu à une perte des 4,2% des terres, 23 % des terres cultivées sur une pente comprise entre 10 et 20 degrés, provoquent une perte des 11,9% des terres et 27% des terres cultivées sur une pente de plus de 20 degrés entraînent une perte des 25,7% des terres¹³.

5° la dispersion de l'habitat essentiellement rural à travers les milles collines, à la recherche des terroirs agricoles, est aussi un gaspillage des terres et une difficulté réelle pour l'aménagement du territoire car l'agriculture reste l'activité familiale, d'auto-consommation et couvre environ 72% du territoire national (avec le bananier en tête et dont 90% de la production sont utilisés pour fabriquer la bière) et occupe plus 81% de la population active. Ce qui est un record africain quand on sait qu'ailleurs ce pourcentage oscille entre 55 et 77%¹⁴ alors que le rendement reste toujours faible.

6° La population urbaine avoisine actuellement 10%¹⁵ de la population totale avec un taux de croissance que l'on peut qualifier de rapide parce qu'il était de 6% en 1996.

¹³ MINICOFIN : *Op. Cit.* p342

¹⁴ KLOTCHKOFF, J-C : *Le Rwanda d'aujourd'hui*, Éd. J.a, Paris, 1990, p64

¹⁵ MINICOFIN : *Op. Cit.* p9

7° L'insalubrité due au manque ou au mauvais assainissement des déchets solides et liquides (notamment le manque des latrines¹⁶ convenables dans les milieux ruraux, dans certains quartiers des centres urbains du pays, et même dans nombreux établissements scolaires, la bouse de la vache en divagation, le mauvais drainage des eaux usées provenant de ruissellement, des usines et des eaux-vannes, le problème des sachets qui constituent les déchets non biodégradables), ce qui entraîne les maladies environnementales (paludisme, fièvre typhoïde, verminose, infections respiratoires et cutanées...) qui sont la première cause de morbidité et de mortalité.

8° La pauvreté (220\$ US de PNB par habitant et par an en 2000) qui entraîne l'ignorance (47,64% des adultes sont sans instruction), la malnutrition (42,7% d'enfants connaissent un retard de croissance et 29% ont une insuffisance pondérale) et la propagation du VIH/SIDA dont les victimes sont estimées à 13,7% des adultes¹⁷.

Aussi, les agressions olfactives, auditives, visuelles dans beaucoup de quartiers des centres urbains du pays aggravent – elles l'état de stress des occupants.

Face à tout ce qui précède, il va sans dire que l'avenir environnemental du Rwanda est fort compromis : d'un côté, la croissance exponentielle de la population, d'autre part, la dégradation environnementale continue et incontrôlée nous pousse à nous poser cette question : quelle sera la situation environnementale du Rwanda d'ici 2025 où la population atteindra 15 millions d'habitants ? Peut être sera - ce un Rwanda sans forêt, sans pluie, sans cours d'eau, sans terres arables, sans hommes valides, etc. Or, tout un chacun aspire au bien-être qui n'est possible

¹⁶ L'étude menée par le Ministère des Finances et de la Planification Économique relative à l'assainissement et à la promotion de l'hygiène indique que 85% de la population des zones rurales dispose de latrines mais que seuls 0,8% de celles-ci sont hygiéniques.

¹⁷ MINICOFIN : Programme National de la réduction de la pauvreté. Stratégie de la réduction de la pauvreté. Kigali, Juin 2002. p16

que grâce au développement durable puisqu'il vise en premier lieu à conserver la stabilité et la productivité de l'environnement.

4. Le développement durable

La Commission mondiale sur l'environnement et le développement qu'on a appelée la **Commission Brundtland** dont le rapport paru en 1989 sous la direction de Madame Gro Halem Brundtland, à l'époque 1^{er} Ministre Norvégien, a défini le développement durable comme « *un développement qui répond aux besoins du présent sans compromettre la capacité des générations futures de répondre aux leurs* »¹⁸. Autrement dit, le développement doit se baser sur une utilisation efficiente et soucieuse de l'environnement de toutes les ressources limitées de la société, naturelles, humaines et économiques. Là encore faut-il immédiatement lever l'équivoque qui réduit le développement durable à la seule dimension écologique : il s'agit plutôt d'une approche globale qui stipule qu'un développement à long terme n'est viable qu'en conciliant trois aspects :

- 1° le respect de l'environnement ;
- 2° l'équité sociale ;
- 3° la rentabilité économique¹⁹.

Le terme durable sous-entend une gestion planifiée et signifie que le processus de développement ne doit pas produire les causes de sa propre fin et donc qu'il est important, notamment en matière d'environnement biophysique, de ne pas créer d'irréversibilité dommageable à la vie sur terre et donc de préserver des marges de manœuvre permettant de revenir sur des actions engagées. Cela revient à dire que pour toute action à entreprendre, il faudra une

¹⁸ COMMISSION BRANDTLAND : Notre avenir à nos tous, rapport de la commission mondiale sur l'environnement, 1989 in : *glossaire des outils économiques de l'environnement* de site Internet : www.igora21.org, p2

¹⁹ Le développement durable n'est pas à confondre avec l'écodéveloppement car ce dernier met l'accent sur la révalorisation des ressources endogènes en luttant contre les effets de démonstration.

étude des impacts environnementaux, c'est-à-dire une étude qui précède ou accompagne les projets ou les travaux dont la réalisation pourrait affecter l'environnement ou montrer les conséquences que ces projets risquent d'avoir sur l'environnement²⁰.

Le principe de développement durable renvoie également à une démarche globale, au sens de systémique, c'est-à-dire l'analyse des interactions entre ces différents éléments qui permettent de comprendre le fonctionnement d'un système et d'agir sur lui. Il s'agit d'une démarche transversale plutôt que sectorielle, une approche qui mesure ses implications dans les domaines multi-sectoriels.

Aussi, le développement durable repose-t-il sur la démocratie locale ou la participation populaire et l'approche citoyenne. Le développement durable désigne en fait une éthique de la responsabilité vis-à-vis de la part de l'humanité qui n'a pas accès à des conditions de vie et de culture décentes, et vis-à-vis des générations futures à qui nous nous devons de transmettre des conditions de vie acceptables tant d'un point de vue économique, que social, culturel, environnemental ou sanitaire. Cette responsabilité est donc globale, universelle, intemporelle mais renvoie à la responsabilité individuelle et locale, aux agissements de chacun. C'est l'articulation entre deux niveaux, entre le local et le global, entre l'individuel et le collectif, qui doit être réussie. Mais pour que le développement durable soit vraiment une approche citoyenne, il doit être fondé sur le principe de savoir pour agir et de savoir pour être. Or, ce savoir ne peut être acquis que grâce à une formation, une acquisition des connaissances avec conscience dont l'éducation environnementale en est le véhicule.

²⁰ CONSEIL DE L'EUROPE : *Biodiversité, questions et réponses*, Strasbourg Cedex, Décembre 1996, p32

5. Education Environnementale

L'éducation environnementale qui fait partie de « *nouvelles éducations* »²¹ à savoir l'éducation pour la paix et la compréhension internationale, l'éducation en matière de la population, l'éducation sanitaire, etc., permet de mettre en évidence la prise de conscience de problèmes de l'environnement par la nécessité d'une approche socio-pédagogique pour y répondre.

L'éducation environnementale est un ensemble des processus permanents et des procédés pédagogiques dont le contenu vise à faire acquérir des connaissances, afin de susciter des attitudes et d'inculquer aux individus et aux groupes les valeurs nécessaires et créer en eux une sensibilité pouvant leur permettre de mener un comportement correct vis-à-vis de l'environnement.

En effet, l'objectif primordial de l'éducation environnementale peut se résumer dans cette déclaration du colloque de Belgrade en 1975 : « *former une population consciente, préoccupée de l'environnement et des problèmes qui s'y rapportent, population qui par son savoir, sa compétence, son état d'esprit, sa motivation et son sens de l'engagement soit en mesure de contribuer individuellement et collectivement à résoudre les problèmes actuels et à éviter qu'ils ne s'en posent de nouveaux dans l'avenir* »²².

Mais le caractère permanent et tourné vers l'avenir assure à l'éducation environnementale, la continuité et l'efficacité des actions bénéfiques qu'il convient d'entreprendre dans l'environnement pour l'intérêt des générations présentes et futures.

En Afrique d'une façon générale et au Rwanda d'une manière particulière, l'éducation environnementale est encore au stade embryonnaire. Il ne faut pas se laisser surprendre du très bas niveau

²¹ BUREAU INTERNATIONAL DU TRAVAIL (BIT) : *L'éducation en matière de population*, Éd. de l'UNESCO, Lausanne, 1993, p20

²² BOUDZOUNOU A.: *Education et formation en matière d'environnement in : Actes du séminaire sur l'environnement et le développement durable*, Brazzaville, 1989, p209

de conscience et de l'état d'ignorance des individus et de groupes en matière d'environnement. Certes, beaucoup d'autres facteurs l'expliquent, mais l'absence d'une véritable éducation environnementale reste une des causes majeures de cette situation.

Au Rwanda comme dans la plupart des pays africains voire d'autres parties du monde, le contenu de l'éducation environnementale ne ressort pas dans le cursus de formation scolaire ou académique. On retrouve de manière isolée, les modules ou séquences d'enseignement ayant trait à la protection de l'environnement dans certaines disciplines comme la géographie, la chimie, la biologie, la sociologie, l'économie... que les apprenants gobent inconsciemment puisqu'ils n'arrivent pas à un changement des attitudes vis-à-vis de l'environnement. En matière d'environnement, il ne suffit pas de présenter une information aux élèves ou aux étudiants, il faut encore identifier leurs propres interrogations et leurs besoins, les convaincre pour qu'ils prennent conscience, les motiver pour la participation à la prise des décisions adéquates. A ce sujet, Georges Vaideanu²³ nous rappelle que les connaissances dispensées dans le cadre des différentes disciplines, ou acquises en dehors de l'école à travers les médias, n'arrivent pas à former une conception cohérente ni à donner une vision globale de problématique aussi complexe que celles de l'environnement. Une formation lacunaire et faiblement structurée ne saurait devenir une base pour la formation de comportements et de savoirs-faire appropriés car les producteurs des programmes à la télévision ou à la radio ont une tâche difficile de distinguer entre les idées et les attitudes à encourager et les préjugés à éliminer.

La nécessité d'intégrer l'éducation environnementale dans le système éducatif formel rwandais s'avère donc très indispensable, car une telle éducation contribue à la recherche de solutions efficaces et durables. C'est pourquoi si l'on dit que tout passe par l'éducation, cela veut dire qu'avant de se retrouver dans la vie sociale, bon nombre des attitudes et des comportements sont ou peuvent être modelés par l'éducation.

²³ BUREAU INTERNATIONAL DU TRAVAIL (BIT) : *Op. Cit.* p.25

Conscient de cette situation, l'Etat rwandais en collaboration avec l'UNICEF, s'est proposé d'élaborer un guide sur l'éducation environnementale pour les écoles secondaires.

Mais l'éducation environnementale de par sa conception, la structuration de son contenu éducatif, ses stratégies pédagogiques et l'organisation de ses démarches d'apprentissage fondées sur la méthode systémique et heuristique adoptée dans l'approche des questions multidimensionnelle, les conditions de son intégration dans le système éducatif rwandais doivent être réunies : à savoir la formation des enseignants, l'apport de divers auxiliaires pédagogiques, la mise en place d'un cadre institutionnel opérationnel. A ce niveau, l'éducation environnementale devient une affaire des spécialistes. Mais cela ne veut pas dire que l'éducation extra-scolaire n'a pas sa place ; elle pourra intervenir dans le cadre de formation contenue et de recyclage.

6. Conclusion

En matière d'environnement l'ignorance tue. Ce sont les citoyens informés et actifs qui feront que le Rwanda puisse se développer durablement, car l'éducation environnementale est orientée vers l'action.

La solution des problèmes environnementaux passe par une éducation qui permet aux individus et aux collectivités d'abord de bien saisir le concept d'environnement et ensuite de comprendre les relations qui les unissent à ce dernier. Puisque les problèmes de l'environnement sont créés, parfois d'une manière inconsciente, par l'homme qui en subit lui-même les conséquences, c'est à lui qu'il faut s'adresser pour changer cet état de choses.

Dans cette optique le Professeur Mamadou note : *« L'éducation environnementale apparaît comme une sorte de révolution culturelle destinée à redonner l'éducation une fonction éthique qu'elle a parfois perdue, son but ultime étant la sauvegarde de l'espèce humaine réconciliée avec lui-même dans un environnement amélioré pour les générations présentes et futures »*²⁴.

²⁴ MAMADOU, C : Éducation relative à l'environnement in : *Educafrica*, n° 7 du juin 1981, Unesco-Breda, Dakar, p168

REFERENCES BIBLIOGRAPHIQUES

1. AHMED EL HATTAB : La problématique environnement-développement in : *Actes du Séminaire sur l'environnement et le développement durable*, Barazzaville, 1989, pp 115-122.
2. BIT : *L'éducation en matière de population*, Ed. de l'UNESCO, Lausanne, 1993, 64p.
3. BOUDZOUNOU A.: Education et formation en matière d'environnement in : *Actes du séminaire sur l'environnement et le développement durable*, Brazzaville, 1989, pp 107-210.
4. HOLDGATE, M : Mettre les connaissances à l'usage des pauvres, in : *Mazigira* Vol. 8 n°5 Septembre-Octobre 1985, 37p.
5. COMMISSION BRANDTLAND : Notre avenir à nos tous, rapport de la commission mondiale sur l'environnement, les Editions de Fleurs, 1989 in : *glossaire des outils économiques de l'environnement* de site Internet : www.agora21.org, 42p.
6. CONSEIL DE L'EUROPE : *Biodiversité, questions et réponses*, Strasbourg Cedex, Décembre 1996, 32p.
7. KLOTCHKOFF, J-C : *Le Rwanda d'aujourd'hui*, Ed. J.a, Paris, 1990, 200p.
8. MAMADOU, C : Education relative à l'environnement in : *Educafrica*, n° 7 du juin 1981, Unesco-Breda, Dakar, pp 168-183.
9. MINICOFIN : *Les indicateurs de développement du Rwanda 2001*, Kigali, 2001, 1379p.
10. MINICOFIN : *Programme National de la réduction de la pauvreté, Stratégie de la réduction de la pauvreté*, Kigali, Juin 2002, 185p.

11. MINITERE : *Stratégie nationale et plan d'action pour la conservation de la biodiversité au Rwanda*, Kigali, 2000, 37p.
12. MINICOFIN : *Vision 2020*, étude inédite, Kigali, Novembre 2002, 69p.
13. MINITRAP : *La problématique de l'aménagement de l'espace du territoire rwandais*, étude inédite, Kigali, Sd, pp 8-26.
14. SIRVEN, P& al ; : *Géographie du Rwanda*, Ed. A. De Bock, Bruxelles, 1974, 175 p.
15. UNESCO-PNUE : Concepts de base de l'éducation environnementale in : *Connexion, bulletin de l'éducation relative à l'environnement*, Unesco-Pnue, Vol. XV, n°2, Juin 1990, 8p.

LE DUAL TOPOLOGIQUE DE L'ESPACE « c » DES SUITES REELLES CONVERGENTES

Par NDABATEZE Abraham
Chargé des Cours Associé

01. Résumé

Les suites et les séries réelles sont beaucoup utilisées dans les domaines divers où les mathématiques servent comme outil. Les économistes, en particulier, utilisent les suites réelles convergentes et les séries absolument convergentes ; en Analyse économique, par exemple, ces notions servent dans l'économétrie des séries temporelles macro-économiques et financières. Le lecteur peut les trouver en détails dans nos références [2], [3] et [5].

L'objectif du présent article est de montrer d'autres liens encore plus étroits qui existent entre l'ensemble " c " des suites réelles convergentes et l'ensemble « l_1 » des séries absolument convergentes ; nous montrons, en fait, que l_1 est le dual topologique de c ; c'est-à-dire que $c^* = l_1$.

02. Introduction

Dans ce travail, nous avons eu besoin des notions de base sur les suites et les séries réelles convergentes ; le lecteur peut les trouver en détails dans notre référence [2] ; aussi, les notions essentielles d'Analyse Fonctionnelle et surtout d'espaces normés nous ont été d'une grande utilité. Le lecteur peut trouver leurs détails dans nos références [1], [4], [7] et [8].

Rappelons brièvement que :

(i)- Si X et Y sont des espaces normés sur le même corps F ($F = \mathbb{R}$ ensemble des réels ou $c =$ ensemble des complexes), alors l'ensemble $LC(X, Y)$ des applications linéaires continues de X dans Y s'appelle le « dual de X ».

Et lorsque X n'est pas un singleton, la relation

$$\|T\| = \sup \left\{ \frac{\|T(x)\|}{\|x\|} : x \in X \text{ et } x \neq \theta \right\}$$
$$= \sup \{ \|T(x)\| : x \in X \text{ et } \|x\| = 1 \}$$

définit une norme sur $LC(X, Y)$.

La norme de $LC(X, Y)$ est complète si et seulement si Y est complet. Dans le cas spécial où $Y = F$, alors $LC(X, F) =$ l'espace vectoriel des formes linéaires continues sur X , s'appelle « le dual topologique de X » et est noté X^* .

(ii) - X^* est toujours un espace de Banach, même si la norme de X n'est pas complète.

(iii) - si X est un espace normé sur le corps F ; par « dual algébrique de X » noté X' , on entend l'espace vectoriel de toutes les formes linéaires sur X .

Ainsi, on a toujours que $X^* \subset X'$.

Lorsque X est de dimension finie, dans ce cas alors $X^* = X'$; dans le cas où X est de dimension infinie, c'est l'inclusion $X^* \subset X'$ qui est toujours vérifiée.

(iv) - Deux espaces normés sont identiques s'il existe une application qui est à la fois un isomorphisme et une isométrie entre eux; c'est ce que nous montrons au théorème I.5 de ce travail.

0.3 Définitions

Soient \mathbb{N} = l'ensemble des nombres entiers positifs
(ou des entiers naturels) ;

\mathbb{R} = l'ensemble des nombres réels.

- Une suite réelle est une application f de \mathbb{N} dans \mathbb{R} .

$$f : \mathbb{N} \longrightarrow \mathbb{R}$$

$$n \longrightarrow f(n) = x_n, \text{ qu'on note } (x_n).$$

- l'ensemble de toutes les suites réelles peut donc être noté $\mathbb{R}^{\mathbb{N}}$ c-à-d l'ensemble des applications de \mathbb{N} dans \mathbb{R} .

- Désignons par :

$$l_1(\mathbb{N}, \mathbb{R}) \text{ [ou simplement } l_1] = \left\{ x \in \mathbb{R}^{\mathbb{N}} : \sum_{i=1}^{+\infty} |x_i| < +\infty \right\}$$

$$l_\infty(\mathbb{N}, \mathbb{R}) \text{ [ou simplement } l_\infty] = \left\{ x \in \mathbb{R}^{\mathbb{N}} : x \text{ est bornée} \right\}$$

$$c(\mathbb{N}, \mathbb{R}) = \left\{ x \in \mathbb{R}^{\mathbb{N}} : \lim_{n \rightarrow +\infty} x_n \in \mathbb{R} \right\} \text{ [ou simplement } c]$$

$$c_0(\mathbb{N}, \mathbb{R}) \text{ [ou simplement } c_0] = \left\{ x \in \mathbb{R}^{\mathbb{N}} : \lim_{n \rightarrow +\infty} x_n = 0 \right\}$$

$$c_x(\mathbb{N}, \mathbb{R}) \text{ [ou simplement } c_x] = \left\{ \begin{array}{l} x \in \mathbb{R}^{\mathbb{N}} : \exists \text{ un nat. } K(x) \text{ tq} \\ \forall \text{ nat } n \geq K(x) \Rightarrow x_n = 0 \end{array} \right\}$$

Toutes les opérations dans les ensembles définis ci-haut seront des opérations usuelles de l'espace vectoriel $\mathbb{R}^{\mathbb{N}}$, à savoir :

$$x + y : \mathbb{N} \rightarrow \mathbb{R}$$

$$n \mapsto (x + y)(n) = x_n + y_n = (x_n) + (y_n)$$

$$tx : \mathbb{N} \rightarrow \mathbb{R}$$

$$n \mapsto (tx)(n) = tx_n = t(x_n)$$

avec $t \in \mathbb{R}$ et x et y des éléments de $\mathbb{R}^{\mathbb{N}}$.

0.4 Théorème

On a les conclusions suivantes : $c_{\infty} \subset I_1 \subset c_0 \subset c \subset I_{\infty}$.

Preuve

(a) Si $x \in c_{\infty}$, alors la série $\sum_{n=1}^{+\infty} |x_n|$ se réduira à la somme finie

$$\sum_{n=1}^K |x_n| = a, \text{ avec } K \in \mathbb{N} \text{ tq. } \forall \text{ nat } n \geq K \Rightarrow x_n = 0$$

et donc la série $\sum_{n=1}^{+\infty} |x_n| = a \in \mathbb{R}$; donc la série converge.

d'où $x \in I_1$

(b) Si $x \in I_1$; alors $\sum_{l=1}^{+\infty} |x_l|$ converge et donc la suite de ses termes généraux

$$|x_n| \text{ tend vers } 0; \text{ et donc } (|x_n|)_{n \geq 1} \text{ tend vers } 0$$

Par conséquent (x_n) converge vers 0; d'où $x \in c_0$

(c) Si $x \in c_0$; alors la suite (x_n) converge vers $0 \in \mathbb{R}$; d'où $x \in c$

(d) Si $x \in c$, alors la suite (x_n) est convergente et donc (x_n) est bornée et donc $x \in I_{\infty}$

0.5 Théorème

L'ensemble l_1 est un vectoriel réel et la relation $\|x\| = \sum_{n=1}^{+\infty} |x_n|$ définit une norme sur l_1 et cette norme est complète.

Preuve

(1) l_1 est visiblement un espace vectoriel sur \mathbb{R} .

(2) Montrons que $\|x\|_1 = \sum_{n=1}^{+\infty} |x_n|$ est une norme sur l_1

(a) $\forall x \in l_1, 0 \leq |x_n|$, pour tout nat. n

$$\text{et donc } 0 \leq \sum_{n=1}^{+\infty} |x_n| < +\infty \Leftrightarrow \mathbf{0} \leq \|x\|_1$$

(b) Pour tout réel t , et pour tout $x \in l_1$,

$$\|tx\|_1 = \sum_{n=1}^{+\infty} |tx_n| = \sum_{n=1}^{+\infty} |t| |x_n| = |t| \sum_{n=1}^{+\infty} |x_n| = |t| \|x\|_1$$

(c) Pour tout $x \in l_1$ et $y \in l_1$

$$|x_n + y_n| \leq |x_n| + |y_n|, \text{ pour tout réel nat. } n$$

$$\Rightarrow \sum_{n=1}^{+\infty} |x_n + y_n| \leq \sum_{n=1}^{+\infty} (|x_n| + |y_n|) \leq \sum_{n=1}^{+\infty} |x_n| + \sum_{n=1}^{+\infty} |y_n|$$

$$\Leftrightarrow \|x + y\|_1 \leq \|x\|_1 + \|y\|_1$$

(d) $\|x\|_1 = 0 \Leftrightarrow \sum_{n=1}^{+\infty} |x_n| = 0 \Leftrightarrow x_n = 0$, pour tout nat. n

$\Leftrightarrow x$ est la suite nulle dans l_1

(3) Pour montrer que $(l_1, \|\cdot\|_1)$ est un espace de Banach,

donnons-nous $(x^{(n)})$ une suite de Cauchy dans l_1 .

Pour tout naturel n , $x^{(n)}$ est un élément de l_1 , c-à-d une suite réelle telle que

$$x^{(n)} = x_1^{(n)}, x_2^{(n)}, \dots, x_k^{(n)}, \dots$$

Fixons un indice $k \in \mathbb{N}$: la suite réelle $(x_k^{(n)})$ étant

dans l_1 , alors $(x_k^{(n)})$ est convergente, d'après le théorème

0.4.(b) et (c) ci-haut : alors tout réel $\varepsilon > 0$,

On peut trouver un entier naturel k tq

$$\forall \text{ nat. } n, m \geq k \Rightarrow \|x^{(n)} - x^{(m)}\|_1 < \frac{\varepsilon}{2}$$

$$\Rightarrow \sum_{i=1}^{+\infty} |x_i^{(n)} - x_i^{(m)}| < \frac{\varepsilon}{2}$$

$$\Rightarrow |x_k^{(n)} - x_k^{(m)}| < \sum_{i=1}^{+\infty} |x_i^{(n)} - x_i^{(m)}| < \frac{\varepsilon}{2}$$

$$\Rightarrow |x_k^{(n)} - x_k^{(m)}| < \frac{\varepsilon}{2} \dots \dots \dots (i)$$

Résumons:

k étant dans \mathbb{N} , \forall réel $\varepsilon > 0$, on a que pour tous

$$\text{naturels } n \text{ et } m \geq k \Rightarrow |x_k^{(n)} - x_k^{(m)}| < \frac{\varepsilon}{2}.$$

Donc pour chaque entier naturel k ,

$(x_k^{(n)})$ est une suite de Cauchy dans l'espace complet \tilde{E} ; ainsi

$(x_k^{(n)})$ est convergente dans \tilde{E} . Alors pour tout $k \in \mathbb{N}$, soit

$$x = \lim_{n \rightarrow \infty} x_k^{(n)} \dots \dots \dots (ii)$$

; posons alors $x = (x_k)_{k \in \mathbb{N}}$; $x \in \tilde{E}^{\mathbb{N}}$.

Il nous faut montrer que $x \in I_1$ et que $(x^{(n)})$ converge vers x .

Pour cela , notons que la suite de Cauchy $(x^{(n)})$ est bornée ;

donc il existe un réel $\varepsilon \geq 0$ t.q. $\|x^{(n)}\|_1 \leq \frac{\varepsilon}{2}$, pour tout nat . n(iii)

A partir de la relation (ii) , on voit que

$$\begin{aligned} \lim_{m \rightarrow \infty} \sum_{k=1}^r |x_k^{(n)} - x_k^{(m)}| &= \sum_{k=1}^r \lim_{m \rightarrow \infty} |x_k^{(n)} - x_k^{(m)}| , \text{ pour tous nat . } r \text{ et } n. \\ &= \sum_{k=1}^r |x_k^{(n)} - x_k| \dots \dots \dots (iv) \end{aligned}$$

Soit un réel $\varepsilon > 0$; $(x^{(n)})$ étant une suite de cauchy dans I_1 , alors il existe un naturel k tel que

$$\begin{aligned} \forall \text{ nat } n , m \geq k &\Rightarrow \|x^{(n)} - x^{(m)}\|_1 < \frac{\varepsilon}{2} \\ &\Rightarrow \sum_{i=1}^{r+2} |x_i^{(n)} - x_i^{(m)}| < \frac{\varepsilon}{2} \dots \dots \dots (v) \end{aligned}$$

Fixons un naturel r .

Pour tout naturel $n \geq r$, on a que

$$\sum_{i=1}^r |x_i^{(n)} - x_i^{(m)}| \leq \sum_{i=1}^r |x_i^{(n)} - x_i^{(m)}| < \frac{\varepsilon}{2}$$

Retenons :

$$\text{Pour } r \in \mathbb{N} \text{ et } n \geq k \Rightarrow \sum_{i=1}^r |x_i^{(n)} - x_i| < \frac{\varepsilon}{2} \dots\dots\dots(\text{vi})$$

Fixons un entier naturel $n \geq k$;

Si $r \in \mathbb{N}$; pour tout indice $i = r$, on a que

$$x_i = x_i - x_i^{(n)} + x_i^{(n)}$$

$$\Rightarrow |x_i| \leq |x_i - x_i^{(n)}| + |x_i^{(n)}|$$

$$\Rightarrow \sum_{i=1}^r |x_i| \leq \sum_{i=1}^r |x_i - x_i^{(n)}| + \sum_{i=1}^r |x_i^{(n)}| < \frac{\varepsilon}{2} + \|x\|_1$$

$$\Rightarrow \sum_{i=1}^n |x_i| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \text{ selon (iii) } \dots\dots\dots(\text{vii})$$

comme (vii) ci – haut est vraie pour tout naturel n , alors la série réelle $\sum_{i=1}^{+\infty} |x_i|$ à termes positifs ou nuls, a ses sommes partielles majorées ; cette série est donc convergente ; et donc $x = (x_i)_{i \geq 1}$ est un élément de l_1 .

D'autre part, pour tout naturel $n \geq k$, on a, selon (vi), que

$$\sum_{i=1}^r |x_i^{(n)} - x_i| < \varepsilon, \text{ pour tout naturel } r.$$

$$\Rightarrow \sum_{i=1}^{+\infty} |x_i^{(n)} - x_i| < \varepsilon$$

$$\Leftrightarrow \|x^{(n)} - x\|_1 < \varepsilon$$

et donc $(x^{(n)})$ converge vers x . $\Leftrightarrow (l_1, \| \cdot \|_1)$ est un espace de Banach .

0.6 N.B

On démontre aisément aussi que l_p est un espace vectoriel réel et que la relation $\|x\|_p = \sup\{|x_n| : n \in \mathbb{N}\}$, $x \in l_p$, définit une norme sur l_p et que cette norme est complète. Par conséquent $(l_p, \|\cdot\|_p)$ est aussi un espace de Banach.

0.7 Théorème

Si $x \in l_1$ et $y \in l_\infty$, alors $\sum_{i=1}^{+\infty} |x_i y_i|$ converge et

$$\sum_{i=1}^{+\infty} |x_i y_i| \leq \|x\|_1 \|y\|_\infty.$$

Preuve

Pour tout naturel n , on sait que

$$\sum_{i=1}^n |x_i y_i| = \sum_{i=1}^n |x_i| |y_i|$$

$$\leq \sum_{i=1}^n |x_i| \|y\|_\infty$$

$$= \|y\|_\infty \sum_{i=1}^n |x_i|$$

$$\leq \|y\|_\infty \sum_{i=1}^{\infty} |x_i|$$

$$= \|y\|_\infty \|x\|_1$$

Donc, la série réelle $\sum_{i=1}^{\infty} |x_i y_i|$ qui est une série à termes positifs ou nuls, a ses sommes partielles majorées par le réel $\|y\|_1 \|x\|_1$. Cette série est donc convergente et sa somme $\sum_{i=1}^{\infty} |x_i y_i| \leq \|y\|_1 \|x\|_1$.

0.8 Théorème

$(E, \|\cdot\|)$ est un espace de Banach.

Preuve

Soit $(x^{(n)})$ une suite de Cauchy dans E ; alors $(x^{(n)})$ est une suite de Cauchy dans l_1 , car E est visiblement un sous espace vectoriel de l_1 . Comme l_1 est complet d'après 0.6, alors $(x^{(n)})$ converge vers un élément x de l_1 .

Montrons que ce $x \in E$.

Pour cela, soit un réel $\varepsilon > 0$; comme $(x^{(n)})$ converge vers x dans l_1 , alors il existe un entier naturel k_1 tel que

$$n \geq k_1 \Rightarrow \|x^{(n)} - x\|_1 = \|x - x^{(n)}\|_1 < \frac{\varepsilon}{4} \dots\dots\dots(i)$$

nous savons aussi que $(x^{(n)})$ est aussi une suite de Cauchy dans E , alors il existe un entier naturel k_2 tel que

$$\text{pour tous naturels } n, m \geq k_2 \Rightarrow \|x^{(n)} - x^{(m)}\|_1 < \frac{\varepsilon}{4} \dots\dots\dots(ii)$$

posons $k = k_1 + k_2$; alors de (i) et (ii) , on a que :

$$\|x^{(k)} - x\|_2 < \frac{\varepsilon}{4}$$

$$\|x^{(k+1)} - x\|_2 < \frac{\varepsilon}{4}$$

$$\|x^{(k)} - x^{(k+1)}\|_2 < \frac{\varepsilon}{4}$$

puisque $x^{(k+1)}$ est un élément de C , alors $x^{(k+1)}$ est une suite réelle convergente ; c'est donc une suite de Cauchy dans \mathbb{R} ; d'où il existe un entier naturel k_3 tel que

$$\text{Pour tous naturels } i, j \geq k_3 \Rightarrow \left| x_i^{(k+1)} - x_j^{(k+1)} \right| < \frac{\varepsilon}{4}$$

Ainsi pour $i, j \geq k_3$, nous aurons :

$$\begin{aligned} \left| x_i - x_j \right| &\leq \left| x_i - x_i^{(k)} \right| + \left| x_i^{(k)} - x_j^{(k+1)} \right| + \left| x_j^{(k+1)} - x_j^{(k+1)} \right| + \left| x_j^{(k+1)} - x_j \right| \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} \text{ selon (iii) et (iv)} \\ &= \varepsilon \end{aligned}$$

Résumons :

Pour tout réel $\varepsilon > 0$, il existe un entier naturel k_3 tel que pour tous naturels

$$i, j \geq k_3 \Rightarrow \left| x_i - x_j \right| < \varepsilon$$

Done (x_i) est une suite de Cauchy dans l'espace complet \mathbb{R} . Par conséquent (x_i) converge dans \mathbb{R} : et donc $x \in C$

I. LE DUAL TOPOLOGIQUE DE C

I.1 Lemme

Considérer $\vec{e}_k : \mathbb{N} \rightarrow \mathbb{R}$

$$n \rightarrow \vec{e}_k^{(n)} = \delta_{kn} = \begin{cases} 0 & \text{si } k \neq n \\ 1 & \text{si } k = n \end{cases}$$

Alors :

1) $\vec{e}_k \in c_{00}$

2) Pour tout y dans l_1 , on a que $y = \sum_{k=1}^{+\infty} y_k \vec{e}_k$

3) Si f est Si f est élément de l_1^* , on a

$$f(y) = \sum_{k=1}^{+\infty} y_k f(\vec{e}_k), \text{ pour tout } y$$

dans l_1 .

Preuve

(1) est trivial

(2) Si y est élément de l_1 , alors

$$\sum_{n=1}^{+\infty} |y_n| = y_n * \delta_{nn}$$

$$\sum_{n=1}^{+\infty} |y_n| = y_n * 1$$

$$= \sum_{k=1}^{+\infty} y_k \vec{e}_k$$

(3) Si f est élément de l_1^* , soit y dans l_1 .

$$\begin{aligned}
 \text{Alors } f(y) &= f\left[\sum_{k=1}^{+\infty} y_k \bar{e}_k\right] = f\left[\lim_{n \rightarrow +\infty} \sum_{k=1}^n y_k \bar{e}_k\right] = \lim_{n \rightarrow +\infty} f\left[\sum_{k=1}^n y_k \bar{e}_k\right] \\
 &= \lim_{n \rightarrow +\infty} \sum_{k=1}^n y_k f(\bar{e}_k) \text{ car } f \text{ est linéaire} \\
 &= \sum_{k=1}^{+\infty} y_k f(\bar{e}_k)
 \end{aligned}$$

I.2 Théorème

Si f est élément de c^* , alors la série $\sum_{k=1}^{+\infty} f(\bar{e}_k)$ converge absolument et de plus $\sum_{k=1}^{+\infty} f(\bar{e}_k) \leq \|f\|$

Preuve

Avant de prouver ce théorème, rappelons que pour tout réel t , on définit « signature de t », noté sgnt , par :

$$\text{Sgnt} = \begin{cases} 0 & \text{si } t=0 \\ \frac{t}{|t|} & \text{si } t \neq 0 \end{cases}$$

Il est clair que si $t > 0$, $\text{sgnt} = 1$
 si $t < 0$, $\text{sgnt} = -1 \Rightarrow t \cdot \text{sgnt} = |t|$

La preuve à présent

Fixons f dans c^* et pour tout naturel n , définissons la suite $a^{(n)} : \mathbb{N} \rightarrow \mathbb{R}$

$$k \rightarrow a_k^{(n)} = \begin{cases} \text{sgn } f(\bar{e}_k) & \text{si } k \leq n \\ 0 & \text{si } k > n \end{cases}$$



ce qui revient à dire que :

$$a^{(n)} = \sum_{k=1}^n \operatorname{sgn} f(\bar{e}_k) \bar{e}_k \dots\dots\dots(i)$$

Il est clair que $a^{(n)}$ est élément de c_∞ et donc $a^{(n)}$ est élément de l_∞ .

De plus, comme $|\operatorname{sgn} t| = 1$, pour tout réel t

$$\text{Alors } \|a^{(n)}\|_\infty \leq 1$$

$$\Rightarrow \|f(a^{(n)})\| \leq \|f\| \cdot \|a^{(n)}\|_\infty \leq \|f\| \dots\dots\dots(ii)$$

de (i), on peut également écrire que :

$$\begin{aligned} f(a^{(n)}) &= f\left(\sum_{k=1}^n \operatorname{sgn} f(\bar{e}_k) \bar{e}_k\right) \\ &= \sum_{k=1}^n f(\bar{e}_k) \operatorname{sgn} f(\bar{e}_k) \text{ car } f \text{ est linéaire} \\ &= \sum_{k=1}^n |f(\bar{e}_k)| \end{aligned}$$

$$\text{D'où } \|f(a^{(n)})\| = \sum_{k=1}^n |f(\bar{e}_k)| \dots\dots\dots(iii)$$

En combinant (iii) et (ii) ci-haut, on obtient que

$$\sum_{k=1}^n |f(\bar{e}_k)| \leq \|f\|, \text{ pour tout naturel } n.$$

Puisque $\|f\|$ est un réel, alors on a que :

$$\sum_{k=1}^\infty |f(\bar{e}_k)| \leq \sum_{k=1}^\infty |f(\bar{e}_k)| \leq \|f\|.$$

1.3 Lemme

Soit x dans l_1 et considérons l'application

$$S(x) : c \rightarrow \mathbb{R}$$

$$y \rightarrow S(x)(y) = x_1 \lim_{k \rightarrow +\infty} y_k + \sum_{k=2}^\infty x_k y_{k-1}$$

Alors :

- 1) $S(x)$ est linéaire et continue c-à-d $S(x)$ est élément de c^*
- 2) $\|S(x)\| \leq \|x\|_1$

Preuve

(1) Soit x dans l_1 . Pour voir que $S(x)$ prend ses valeurs dans \mathbb{R} , soit y dans c ; alors $\lim_{k \rightarrow +\infty} y_k$ est un nombre réel.

De plus $c \subset l_\infty$, donc y est élément de l_∞ ; d'où

$$|y_n| \leq \sup \{ |y_i| : i \in \mathbb{N} \}, \text{ pour tout naturel } n.$$

$$\Rightarrow |x_k y_{k-1}| = |x_k| |y_{k-1}| \leq |x_k| \|y\|_\infty, \text{ pour tout}$$

nat. $k \geq 2$ (i)

Puisque x est dans l_1 , alors la série $\sum_{k=1}^{\infty} |x_k|$ converge.

Nous concluons d'après les inégalités en (i) ci-haut que la série

$$\sum_{k=2}^{\infty} |x_k y_{k-1}| \text{ converge et que}$$

$$\left| \sum_{k=2}^{\infty} x_k y_{k-1} \right| \leq \sum_{k=2}^{\infty} |x_k y_{k-1}| \leq \sum_{k=2}^{\infty} |x_k| \|y\|_\infty$$

$$= \left(\sum_{k=2}^{\infty} |x_k| \right) \|y\|_\infty \dots \dots \dots \text{(ii)}$$

Par conséquent $x_1 \left(\lim_{k \rightarrow +\infty} y_k \right) + \sum_{k=2}^{\infty} x_k y_{k-1}$ est un nombre réel et donc

$S(x)$ prend bien ses valeurs dans \mathbb{R} .

Montrons que :

$S(x)$ est linéaire ; pour cela, soient y et z dans c et soient un nombre réel a . comme c est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$, posons $u = y + z$; u est évidemment élément de c ; ay est aussi un élément de c .

$$\begin{aligned}
 S(x)(u) &= x_1 \lim_{k \rightarrow +\infty} u_k + \sum_{k=2}^{\infty} x_k u_{k-1} \\
 &= x_1 \lim_{k \rightarrow +\infty} (y_k + z_k) + \sum_{k=2}^{\infty} x_k (y_{k-1} + z_{k-1}) \\
 &= x_1 \left(\lim_{k \rightarrow +\infty} y_k + \lim_{k \rightarrow +\infty} z_k \right) + \sum_{k=2}^{\infty} x_k y_{k-1} + \sum_{k=2}^{\infty} x_k z_{k-1} \\
 &= x_1 \lim_{k \rightarrow +\infty} y_k + \sum_{k=2}^{\infty} x_k y_{k-1} + x_1 \lim_{k \rightarrow +\infty} z_k + \sum_{k=2}^{\infty} x_k z_{k-1} \\
 &= S(x)(y) + S(x)(z)
 \end{aligned}$$

$$\begin{aligned}
 S(x)(ay) &= x_1 \lim_{k \rightarrow +\infty} ay_k + \sum_{k=2}^{\infty} x_k (ay_{k-1}) \\
 &= ax_1 \lim_{k \rightarrow +\infty} y_k + a \sum_{k=2}^{\infty} x_k y_{k-1} \\
 &= a \left(x_1 \lim_{k \rightarrow +\infty} y_k + \sum_{k=2}^{\infty} x_k y_{k-1} \right) \\
 &= a S(x)(y)
 \end{aligned}$$

Montrons que $S(x)$ est continue

Pour y de c , on a que $\left| \lim_{k \rightarrow +\infty} y_k \right| = \lim_{k \rightarrow +\infty} |y_k|$

Or $|y_k| \leq \|y\|_{\infty}$, pour tout naturel k

$$\Rightarrow \left| \lim_{k \rightarrow +\infty} y_k \right| \leq \|y\|_{\infty} \dots \dots \dots (i)$$

$$\begin{aligned}
\text{Ainsi } |S(x)(y)| &= \left| x_1 \lim_{k \rightarrow +\infty} y_k + \sum_{k=2}^{\infty} x_k y_{k-1} \right| \\
&\leq \left| x_1 \lim_{k \rightarrow +\infty} y_k \right| + \left| \sum_{k=2}^{\infty} x_k y_{k-1} \right| \\
&= |x_1| \left| \lim_{k \rightarrow +\infty} y_k \right| + \sum_{k=2}^{\infty} |x_k y_{k-1}| \\
&\leq |x_1| \|y\|_{\infty} + \sum_{k=2}^{\infty} |x_k| \|y\|_{\infty} \text{ selon (i)}
\end{aligned}$$

$$\begin{aligned}
|S(x)(y)| &= \left(|x_1| + \sum_{k=2}^{\infty} |x_k| \right) \|y\|_{\infty} \\
&= \|x\|_1 \|y\|_{\infty}
\end{aligned}$$

Résumé : $|S(x)(y)| \leq \|x\|_1 \|y\|_{\infty}$, pour tout y de c

.....(ii)

Comme (ii) est vraie pour tout y dans c et que $S(x)$ est linéaire, cela prouve que $\|S(x)\| \leq \|x\|_1$.

I.4. Lemme

L'application $S : l_1 \rightarrow c^*$

Définie par

$$x \mapsto S(x)(y) = x_1 \lim_{k \rightarrow +\infty} y_k + \sum_{k=2}^{\infty} x_k y_{k-1} \text{ est linéaire et}$$

Surjective

Preuve

Soient des réels a et b et soient x et z dans l_1 ; nous savons que $s(x)$ et $s(z)$ sont des éléments de c^* , d'après I.3 ci-haut. Pour un y fixé dans l_1 , nous avons, par définition de S , que

$$\begin{aligned}
s(ax + bz)(y) &= (ax_1 + bz_1) \lim_{k \rightarrow \infty} y_k + \sum_{k=2}^{\infty} (ax_k + bz_k) y_{k-1} \\
&= ax_1 \lim_{k \rightarrow \infty} y_k + bz_1 \lim_{k \rightarrow \infty} y_k + \sum_{k=2}^{\infty} ax_k y_{k-1} \\
&= a \left[x_1 \lim_{k \rightarrow \infty} y_k + \sum_{k=2}^{\infty} x_k y_{k-1} \right] + b \left[z_1 \lim_{k \rightarrow \infty} y_k + \sum_{k=2}^{\infty} z_k y_{k-1} \right] \\
&= aS(x)(y) + bS(z)(y)
\end{aligned}$$

Pour voir que S est surjective, soit f dans \mathfrak{C}^*

- Considérer la suite

$$u : \mathbb{N} \rightarrow \mathbb{R}$$

$$n \rightarrow u_n = 1, \text{ pour tout naturel } n.$$

On a la suite constante $(u_n) = (1)$ dont la limite est 1;

donc u est élément de \mathfrak{C} .

- Considérer aussi la suite :

$$x : \mathbb{N} \rightarrow \mathbb{R}$$

$$1 \rightarrow x_1 = f(u) - \sum_{k=1}^{\infty} f(e_k^{-\rightarrow})$$

$$k \rightarrow x_k = f(e_{k-1}^{-\rightarrow}), \text{ pour tout } k \geq 2$$

La suite (x_n) ainsi définie est un élément de l_1 , d'après 1.2

- Montrons alors que $f = S(x)$

Pour cela, soit y dans \mathfrak{C} et définissons la suite

$$y^{(n)} : \mathbb{N} \rightarrow \mathbb{R}$$

$$k \rightarrow y_k^{(n)} = \begin{cases} 0 & \text{si } k \leq n \\ y_k - \lim_{l \rightarrow \infty} y_l & \text{si } k > n \end{cases}$$

Cela revient à dire que

$$y^{(n)} = (0, 0, \dots, 0, y_{n+1} - a, y_{n+2} - a, \dots)$$

(avec $a = \lim_{l \rightarrow \infty} y_l$)

Puisque la suite (y_k) converge vers a , alors la suite $(y_k - a)$ converge vers 0 ;

Donc chaque $y^{(n)}$ est un élément de c_0 et donc de l_∞ .

D'où

$$\|y^{(n)}\|_\infty = \sup \{|y_k - a| : k \geq n+1\} \dots\dots\dots (i)$$

Si $t > 0$ est un réel, comme (y_k) converge vers a ,

Alors il existe un entier naturel T tq. pour tout naturel

$$k \geq T \Rightarrow |y_k - a| < \frac{t}{2}$$

alors pour tout naturel $n \geq T$, on aura aussi

$$k \geq n+1 \Rightarrow |y_k - a| < \frac{t}{2}$$

Il en découle que

$$\text{Sup} \{|y_k - a| : k \geq n+1\} \leq \frac{t}{2} < t$$

$$\Leftrightarrow \|y^{(n)}\|_\infty < t$$

Ainsi, nous aurons montré que $\lim_{n \rightarrow +\infty} y^{(n)} = 0$

D'où, d'après la définition de $y^{(n)}$, on obtient que

$$y = au + \sum_{n=1}^{\infty} (y_n - a) \bar{e}_n$$

$$\Rightarrow f(y) = a f(u) + \sum_{n=1}^{\infty} y_n f(\bar{e}_n) - a \sum_{n=1}^{\infty} f(\bar{e}_n) \quad \text{car}$$

f linéaire

$$= a \left[f(u) - \sum_{n=1}^{\infty} f(\bar{e}_n) \right] + \sum_{n=1}^{\infty} x_{n+1} y_n \quad \text{car } f(\bar{e}_n) = x_{n+1}$$

$$f(y) = x_1 \lim_{k \rightarrow +\infty} y_k + \sum_{n=2}^{\infty} x_n y_{n-1}$$

$$= S(x)(y), \text{ pour tout } y \text{ dans } c$$

$$\Rightarrow S(x) = f$$

I. 5 Théorème

L'application $S: l_1 \rightarrow c^*$ définie comme en I. 4, est à la fois un isomorphisme et une isométrie c-à-d $l_1 = c^*$.

Preuve

Nous avons montré en I. 4 que S est linéaire et surjective ; il reste alors à prouver que S est une isométrie c-à-d $\|S(x)\| = \|x\|_1$, pour tout x de l_1 .

Alors, donnons-nous un élément x dans l_1

et pour tout naturel n , définissons la suite $a^{(n)}: \mathbb{N} \rightarrow \mathbb{R}$

$$\text{par } a_k^{(n)} = \begin{cases} \operatorname{sgn} S(x)(\bar{e}_k), & \text{si } k \leq n \\ \operatorname{sgn} x_k, & \text{si } k > n \end{cases}$$

Il est évident que $\lim_{k \rightarrow +\infty} a_k^{(n)} = \operatorname{sgn} x_k$ et donc $a^{(n)}$ est dans c , pour tout naturel n

Puisque $|\operatorname{sgn} t| \leq 1$, pour tout réel t

alors $\|a^{(n)}\|_\infty \leq 1$

$$\Rightarrow \|S(x)(a^{(n)})\| \leq \|S(x)\| \|a^{(n)}\|_\infty \leq \|S(x)\| \dots \dots \dots (i)$$

par ailleurs comme x est dans l_1 , alors la série $\sum_{k=1}^{\infty} x_k \operatorname{sgn} x_k$

converge absolument et donc $\lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} x_k \operatorname{sgn} x_k = 0$
(ii)

On a ainsi que

$$S(x)(a^{(n)}) = x_1 \lim_{k \rightarrow \infty} a_k^{(n)} + \sum_{k=2}^{\infty} x_k a_{k-1}^{(n)}$$

$$= x_1 \operatorname{sgn} x_1 + \sum_{k=2}^{n+1} x_k a_{k-1}^{(n)} + \sum_{k=n+2}^{\infty} x_k a_{k-1}^{(n)}$$

$$\Rightarrow S(x)(a^{(n)}) = |x_1| + \sum_{k=2}^{n+1} x_k \operatorname{sgn} S(x)(\bar{e}_{k-1}) + \sum_{k=n+2}^{\infty} x_k \operatorname{sgn} x_1$$

$$\text{Or } S(x)(\bar{e}_{k-1}) = x_k$$

$$\text{Donc } S(x)(a^{(n)}) = |x_1| + \sum_{k=2}^{n+1} x_k \operatorname{sgn} x_k + \sum_{k=n+2}^{\infty} x_k \operatorname{sgn} x_1$$

$$= |x_1| + \sum_{k=2}^{n+1} |x_k| + \sum_{k=n+2}^{\infty} x_k \operatorname{sgn} x_1$$

$$S(x)(a^{(n)}) = \sum_{k=1}^{n+1} |x_k| + \sum_{k=n+2}^{\infty} x_k \operatorname{sgn} x_1$$

Introduisons cette valeur dans (i) ci-haut pour obtenir que :

$$\left| \sum_{k=1}^{n+1} |x_k| + \sum_{k=n+2}^{\infty} x_k \operatorname{sgn} x_1 \right| \leq \|S(x)\| \text{ pour tout naturel } n$$

Passons alors à la limite pour n tendant vers $+\infty$ pour obtenir finalement

$$\begin{aligned} & \left| \lim_{n \rightarrow +\infty} \sum_{k=1}^{n+1} |x_k| + \lim_{n \rightarrow +\infty} \sum_{k=n+2}^{\infty} x_k \operatorname{sgn} x_1 \right| \leq \|S(x)\| \\ \Rightarrow & \left| \sum_{k=1}^{+\infty} |x_k| + 0 \right| \leq \|S(x)\| \text{ selon (i)} \\ \Rightarrow & \|x\|_1 \leq \|S(x)\| \end{aligned}$$

.....(iii)

Or, nous avons déjà vu en I. 3 que $\|S(x)\| \leq \|x\|_1$

Combinons cela avec (iii) pour obtenir que $\|S(x)\| = \|x\|_1$ et donc S est une isométrie ; d'où $l_1 = c^*$

BIBLIOGRAPHIE

- [1] – CHOUQUET G. , Cours d'Analyse, Tome II, Topologie, 2^{ème} éd. Masson et Cie, Paris, 1973, 317 p.
- [2] – GOURIEROUX C. et MONFORT A., Séries temporelles et Modèles dynamiques, Ed. Economica, collection « Economie et statistiques avancées », Paris, 1999, 780 p.
- [3] – GREEN H. William, Econometric analysis, 4th ed., Prentice Hall, Inc., Upper Saddle River , New Jersey, 2000, 951p.
- [4] – KOLMOGOROV A. et FOMINE S., Eléments de la théorie des fonctions et de l'Analyse Fonctionnelle, traduction en langue française, Ed. Moscou, 1973, 536p.
- [5] – LARDIC Sandrine et MIGNON Valérie, Econométrie des séries temporelles macro-économiques et financières, Ed. Economica, Paris, 2002, 428p.
- [6] – MUBENGA Kampotu, Eléments d'Analyse Infinitésimale, vol. 1, Les fondements, P.U.Z., Kinshasa, 1976, 383p.
- [7] – MUBENGA Kampotu, Notes de cours d'Analyse Fonctionnelle, 1^{ère} Licence Mathématique, Université de Kinshasa, 1983-1984, inédit.
- [8] – MUBENGA Kampotu, Notes de cours d'Analyse supérieure, 2^{ème} Licence Mathématique, ISP-MBANDAKA, 1976-1977, inédit.

**LE REGIME JURIDIQUE DE L'INFRACTION
D'ADULTERE FACE AUX IMPERATIFS
DU MOMENT.**

Par Christophe BIZIMUNGU,

Mai 2003.

INTRODUCTION.

Les indiens considèrent l'acte sexuel comme étant un acte saint permettant à l'homme de créer à la place de Dieu. Par conséquent, il ne peut être toléré qu'un individu puisse introduire le désordre dans ce pouvoir de création. ⁽¹⁾

Les pahouins quant à eux, ⁽²⁾ considèrent la fonction sexuelle comme un mal foncier mais nécessaire par lequel il faut passer pour les besoins de la création, la valeur suprême sociale. Il convient naturellement d'en limiter les dégâts en l'ordonnant exclusivement à cette valeur positive ⁽³⁾. De façon générale en Afrique, la sexualité est une dimension fondamentale de l'humain, mais son exercice, toujours dangereux sous quelque rapport, est exclusivement ordonné à la procréation. ⁽⁴⁾

Dans ce contexte, le mariage, qui est généralement admis comme le cadre propre de procréation, consacre l'obligation de l'exclusivité sexuelle, les conjoints n'étant pas autorisés à entretenir des liens sexuels en dehors du mariage. C'est ainsi que dans beaucoup de pays, l'union sexuelle en dehors du mariage est un acte punissable.

Le législateur rwandais a ainsi prévu l'infraction d'adultère à l'article 353 du code pénal en la définissant comme étant une union sexuelle d'une personne mariée avec une personne autre que son conjoint.

Dans les dispositions qui suivent, le législateur a également précisé des peines qui diffèrent selon que le coupable est un homme ou une femme.

Il a également exigé la plainte du conjoint offensé comme préalable à toute poursuite judiciaire; tout en limitant les moyens de preuve contre le complice de l'adultère: les seules preuves qui pourront être admises contre le complice sont, outre le flagrant délit, les lettres ou autres pièces écrites par lui.

Il s'avère cependant opportun de confronter la loi aux réalités du moment: d'une part il y a l'épidémie du SIDA (Syndrome d'Immuno-Déficience Acquise) qui extermine les populations à travers le monde et au Rwanda en particulier; d'autre part il y a le grand mouvement d'émancipation de la femme et d'égalité des sexes, répandu dans le monde entier et au Rwanda en particulier, qui exige un traitement égal des personnes indépendamment de leurs sexes.

Le danger du SIDA n'exigeant pas beaucoup d'explications pour être appréhendé, nous nous limitons à préciser quelques chiffres suffisamment éloquents : En l'an 2002, 42 millions de personnes vivaient avec le virus du SIDA à travers le monde; 5 millions ont été contaminées et 3,1 millions de personnes sont mortes pendant la même année suite à cette maladie ⁽⁵⁾.

Cette épidémie affecte toutes les couches de la population sans distinction de race, de sexe ou de statut social. Quant aux personnes mariées, qui sont visées par la présente étude, le taux de prévalence est élevé chez elles : environ 13% ⁽⁶⁾.

Cet état de choses renforce l'obligation de fidélité dans le mariage, étant donné que le mode le plus important de transmission du virus est l'union sexuelle. L'adultère comporte ainsi des risques graves non pas seulement pour le coupable, mais également pour son conjoint, leur descendance, leurs familles et même la société de façon générale.

L'on pourrait garder des espoirs dans la médecine moderne qui connaît des progrès remarquables mais il y a lieu de relativiser les espérances. En effet, en 1995, Cyrille KOUPERNIK écrivait déjà que l'explosion de la pandémie à HIV avait brutalement mis fin à une ère de triomphalisme médical ⁽⁷⁾.

Il se posa ainsi des questions qui révèlent le caractère péremptoire de la situation.



<< D'autres questions peuvent venir à l'esprit : par exemple, comment se fait-il qu'une forme très inférieure de vie, un rétrovirus incapable de survivre par lui-même, se montre capable d'une extraordinaire adaptabilité qui déjoue les efforts d'innombrables savants ? D'où vient, que signifie, que présage cette << logique du vivant >> ? ... Si cette logique a un sens, il est plutôt inquiétant pour nous. Si elle n'en a pas (comme je le pense), nous ne sommes pas maîtres de l'univers, et ce n'est pas plus rassurant >> ⁽⁸⁾.

Le danger du SIDA, qui constitue une grave menace à l'existence même de la nature humaine, a amené la communauté internationale et les Etats, individuellement, à prendre d'importantes mesures dans le cadre de la lutte contre ce fléau. Les mesures prises ne sont pas négligeables, mais encore faut-il affirmer que devant une menace d'une telle envergure, aucune mesure ne doit être négligée.

Dans le domaine juridique, certains pays ont pris très peu de mesures dans ce sens, d'autres n'en ont pris aucune. Nous en voudrions pour exemple la répression de la transmission du virus : dans presque tous les pays du monde, aucune loi pénale n'a été édictée en vue de réprimer cet acte de façon spécifique. Et les juges se débrouillent en cherchant par-ci par-là des dispositions souvent inadéquates pour punir cet acte. Beaucoup reste ainsi à faire dans le domaine juridique.

Cette étude se veut donc une modeste contribution à la lutte contre le SIDA par des moyens juridiques. Nous nous limiterons cependant à l'infraction d'adultère. La question fondamentale à laquelle nous voulons répondre est celle de savoir si l'infraction d'adultère telle que prévue dans notre code pénal, cadre parfaitement avec les réalités du moment. Autrement dit, cette incrimination permet-elle de contribuer efficacement à lutter contre la pandémie du SIDA ? est-elle en harmonie avec les aspirations actuelles à l'égalité des sexes ?

La réponse à ces questions nous exige le cheminement suivant: nous examinerons dans un premier temps le régime juridique de l'adultère (chapitre I) pour en ressortir l'archaïsme et l'inadaptabilité (chapitre II) et, enfin, nous concluerons en proposant des recommandations qui permettront d'harmoniser le droit avec les impératifs du moment.

I. LE REGIME JURIDIQUE DE L'ADULTERE.

Sous ce chapitre nous examinerons les éléments constitutifs et le régime répressif de l'infraction d'adultère.

I.1. Les éléments constitutifs de l'infraction d'adultère.

L'article 353 du code pénal définit l'adultère comme étant l'union sexuelle d'une personne mariée avec une personne autre que son conjoint.

Il ressort de cette définition légale que cette infraction est constituée des trois éléments suivants :

1. L'état de mariage;
2. La conjonction sexuelle avec une personne autre que son conjoint;
3. L'intention coupable.

I.1.1. L'état de mariage.

Le prévenu doit être engagé dans un lien de mariage au moment de la commission de l'infraction. Ce mariage doit valablement et actuellement exister. C'est à dire qu'il ne doit pas être entâché de nullité, dissous ou simplement à venir.

La nullité absolue ou même relative écarte la condamnation. C'est une question préjudicielle sur laquelle la juridiction compétente doit se prononcer avant d'examiner la matière pénale.

En cas de décès du conjoint ou en cas de divorce, il va de soi que le mariage n'existe plus et que par conséquent l'adultère doit être écarté. Cependant, pour que le **divorce** soit effectif, il faut que le **jugement de divorce ait acquis l'autorité de la chose jugée**, toutes les **voies de recours étant épuisées** ou les **délais correspondants expirés**. Si le **devoir de fidélité** entre les époux disparaît après le divorce, il n'en est pas de même en cas de **séparation de corps**. En effet, la **séparation de corps** ne rompt pas le **lien de mariage** et donc l'**obligation de fidélité réciproque** qui existe entre les **deux époux séparés**. Toute **conjonction sexuelle** avec un tiers serait **condamnabile** en période de **séparation de corps**.

Le mariage futur n'a pas non plus d'effet quant à l'**incrimination de l'adultère**, même si les **fiancés** se sont promis la **fidélité** et le **droit d'exclusivité sexuelle** avant le mariage. L'**infraction d'adultère** ne pourra exister que si après la **célébration effective du mariage**, l'un des conjoints a procédé à la **conjonction sexuelle** avec un tiers.

1.1.2. La conjonction sexuelle avec une personne autre que son conjoint.

L'élément matériel de l'adultère s'identifie à celui du viol dans la mesure où la loi **exige une conjonction sexuelle**. La **conjonction sexuelle** est l'**acte de pénétration du sexe mâle** dans les parties génitales de la femme.

Cela signifie que seront exclus les autres actes proches tels que les **rapports homosexuels**, les **baisers les plus obscènes** et tous autres **actes particulièrement impudiques** quel que soit leur degré d'**obscénité**.

Toutefois, dès lors que l'**union sexuelle** aura eu lieu, la femme (ou le mari) ne pourra invoquer l'**absence de satisfaction sexuelle**. Le **mobile** importe peu également car il n'est pas nécessaire que le **coupable** vise à **blessier son conjoint** ou à **satisfaire un plaisir charnel**.

Ainsi analysé, l'élément matériel de l'adultère s'avère insuffisant à répondre aux finalités de l'incrimination. Nous y reviendrons au Chapitre II.

1.1.3. L'intention coupable.

L'infraction d'adultère ne peut être établie que si l'agent a agi en connaissance de cause. L'ignorance ou l'erreur ne suffiraient pas à caractériser l'infraction. Ainsi, une femme qui entretiendrait des relations avec un tiers croyant son mari décédé, ne serait pas coupable d'adultère.

Ne sera pas non plus punissable la femme qui aura été victime d'une ruse et se sera livrée à une autre personne se faisant passer pour son mari, ou la femme dont un homme abuserait pendant qu'elle est en sommeil.

Rappelons que l'intention étant établie, le mobile importe peu. C'est ainsi qu'une femme qui aurait des relations avec un tiers en vue d'acquérir des nouvelles techniques qu'elle préconise apprendre à son mari afin de gagner avec lui l'harmonie sexuelle, n'échappera point à la condamnation. Il en sera de même si elle s'est livrée en vue d'obtenir un emploi pour son mari. Contrairement à ce qu'affirme le Général LIKULIA ⁽⁹⁾, devra également être condamnée la femme qu'un mari indigne obligerait à se prostituer. Cette contrainte ne peut pas être considérée comme une cause de justification puisque la femme est en pareille circonstance en mesure de refuser cette humiliation. A moins de prouver que la contrainte était réellement irrésistible.

En droit coutumier comparé, les éléments constitutifs tels que examinés ci-haut ne trouvent pas une application identique.

Certaines sociétés pratiquant la coutume de purification notamment les BAYOMBE, imposaient à un mari, lors de certains événements, d'avoir impunément des relations extra-conjugales avec une femme d'autrui.

D'autres coutumes toléraient qu'un cadet puisse avoir des relations intimes avec la femme de son grand frère aîné. Dans certains cas, d'ailleurs, le grand frère pouvait intenter un procès contre son petit frère pour ne lui avoir pas procuré des enfants pendant son absence⁽¹⁰⁾.

Le droit coutumier Beti⁽¹¹⁾ voulait qu'une femme mariée n'accorde ses faveurs sexuelles à une tierce personne qu'avec l'accord de son mari.

Celui-ci pouvait donc, avec le consentement de son épouse, prêter cette dernière, selon la coutume de la prostitution hospitalière (le "Mgba"), ou la donner en location à quelqu'un de son entourage, comme le voulait également la coutume du "Mvie"⁽¹²⁾.

Les fiancés étaient également punissables d'adultère dans beaucoup de coutumes par exemple au Congo (chez les BAMBALA et AZANDE), en droit hébreu, etc..⁽¹³⁾.

Dans la corne de l'Afrique, pour prévenir l'adultère certaines coutumes permettaient de coudre les femmes mariées pendant l'absence de leurs maris.⁽¹⁴⁾

I.2. Le régime répressif de l'adultère.

Le régime répressif de l'adultère révèle une particularité à plusieurs égards :

- Au niveau de la procédure;
- Quant aux exigences de la preuve;
- Enfin, au point de vue des peines applicables aux coupables et leurs complices.

I.2.1 La procédure.

L'article 356 du code pénal institue une procédure particulière en matière de poursuite pour adultère.

1.2.1.1 La mise en mouvement de l'action publique.

L'article 356 al 1^{er} subordonne la mise en mouvement de l'action publique à la plainte préalable de l'époux offensé. La forme de la plainte n'est pas précisée. Elle peut donc être aussi bien écrite que verbale. Cette plainte ne peut provenir que de l'époux offensé et non de ses parents ou amis. Il arrive que l'un des époux "sacralise" son conjoint au point de ne plus être en mesure d'apercevoir ou de critiquer ses méfaits. Souvent, ces sont les membres de la famille qui s'en occupent à son insu ou contre son gré. Malgré la défaillance, voire la naïveté de l'un des époux, ses parents ne sont en aucune mesure autorisés à déclencher les poursuites.

Le ministère public n'est pas non plus autorisé à se saisir d'office, même si l'officier de justice a lui même directement pris connaissance du délit, à moins que les actes reprochés couvrent d'autres incriminations. De même, en cas de flagrant délit, les particuliers ne sont pas autorisés à conduire le coupable devant les instances judiciaires, à moins qu'il ait simultanément commis une autre infraction.

La plainte portée contre l'époux coupable permet au ministère public de poursuivre également son complice sans que le plaignant se prononce sur cette poursuite. Il est évident que si le complice est aussi marié, la plainte de son conjoint ne sera pas requise.

Mais l'époux offensé ne peut pas porter plainte contre le seul complice sans dénoncer son conjoint coupable . A partir de l'introduction de la plainte , le ministère public se comporte comme d'ordinaire sauf si un désistement intervient.

1.2.1.2 Le désistement.

L'article 356 al 2 dispose que le plaignant peut en tout état de cause arrêter la procédure par le retrait de la plainte. Ce désistement profite également au complice. Nous croyons que le désistement éteint l'action publique et que donc le plaignant qui s'est désisté ne peut plus porter une nouvelle plainte pour les mêmes faits. Un tel droit ouvrirait la porte au chantage et à l'insécurité juridique dont serait victime l'époux coupable.

1.2.1.3 Le "droit de grâce" de l'époux offensé.

L'article 356 al3 institue un véritable "droit de grâce" de l'époux offensé qui se voit attribuer le pouvoir d'arrêter les effets de la condamnation définitive à l'emprisonnement à la seule condition de reprendre la vie commune. Ce droit de pardon étant institué dans le strict intérêt du foyer conjugal, nous pensons qu'il ne peut pas s'étendre à d'autres coupables c'est à dire les complices.

1.2.2 La preuve.

La preuve contre l'auteur principal de l'adultère respecte les règles habituelles d'administration de la preuve c'est à dire que le ministère public établit l'existence des faits par tous les moyens à partir des aveux, des témoignages, des écrits et des indices concordants.

En revanche, l'article 355 al2 du code pénal devient étonnamment trop exigeant quant à la preuve contre le complice d'adultère, en décidant que les seules preuves qui pourront être admises contre le complice seront, outre le flagrant délit, celles qui résulteront de lettres ou autres pièces écrites par lui.

Le flagrant délit.

Le flagrant délit existe non seulement lorsque les coupables d'adultère ont été surpris la main dans le sac c'est à dire en pleine action sexuelle mais aussi lorsqu'ils sont dans des circonstances qui ne laissent pas de place au doute quant à l'acte qu'ils viennent d'accomplir. C'est le cas lorsque un homme et une femme ont été vus dans la chambre, l'homme en train de remettre la culotte et la femme nue étendue sur le lit .

“ Le retard dans l'ouverture de la chambre, empreinte de deux corps dans un même lit, tenue légère des acteurs du drame ou fuite de l'un d'eux vers un logement voisin, le désordre dans le lit ou sur l'endroit où les deux acteurs ont été surpris par exemple en brousse, les traces de spermes sur le lit, sur le pagne de la femme ou sur le sous-vêtement, le désordre dans la coiffure de la femme ou avec de la terre sur la nuque constitueraient également des preuves de l'adultère. Il en est de même évidemment de celui qui, surpris, saute par la fenêtre et rejoint, tout nu son domicile ou qui porte le sous-vêtement du partenaire”⁽¹⁵⁾ .

Ces scènes si bien décrites par le Général Likulia ne doivent cependant pas être appréciées de façon absolue car il peut en effet arriver qu'une de ces situations surviennent effectivement sans qu'il y ait eu adultère.

Les pièces écrites par le complice.

Le second et dernier moyen de preuve pouvant être fourni contre le complice est constitué des lettres ou autres pièces écrites par lui.

Une lettre dans laquelle une femme rappelle à son amant marié “les délices de la nuit inoubliable” en lui décrivant la scène amoureuse peut bien constituer la preuve de la complicité d'adultère de ladite femme.

Par contre, une lettre dans laquelle une femme mariée écrit qu'elle ne supporte plus son mari depuis qu'elle a découvert la véritable virilité de son amant, ne peut pas, paradoxalement, constituer une preuve contre ce dernier parce que la lettre n'a pas été écrite par lui. Ceci est d'autant plus aberrant que même l'aveu concordant des deux coupables ne peut pas être admis comme preuve contre le complice. Ces conclusions résultent de l'interprétation de l'article 355 al2 qui est très clair à ce sujet, lorsqu'il cite "les seules preuves qui pourront être admises contre le complice". Nous y reviendrons au chapitre II.

I.2.3. Les peines applicables.

La particularité du régime juridique de l'infraction d'adultère que nous avons remarquée au niveau de ses éléments constitutifs, des exigences de la preuve et des poursuites apparaît également au niveau des peines applicables parce que le code pénal présente une discrimination entre la femme et le mari infidèles.

En effet, l'article 354 prévoit des peines inégales en ce sens que la femme convaincue d'adultère est punie d'un emprisonnement d'un mois à un an alors que le mari est quant à lui simplement puni d'un emprisonnement d'un mois à six mois et d'une amende de mille francs ou de l'une de ces peines seulement.

Ceci pèche contre le principe de base du "gender" qui veut que l'homme et la femme soient également traités par la loi. C'est donc manifestement une disposition archaïque qui nécessite une actualisation.

II. L'ARCHAÏSME ET LA NECESSITE D'ACTUALISATION DU REGIME JURIDIQUE DE L'ADULTERE.

L'archaïsme du régime juridique de l'adultère apparaît à plusieurs points de vue : D'abord l'élément matériel requis, qui est la conjonction sexuelle, est trop restrictif (II.1). Quant à la répression, elle s'avère inégale (II.2) et semble nous renvoyer au

19°S. Les conditions de poursuite (II.3) ainsi que les modalités de preuve (II.4) ne sont pas non plus satisfaisantes car elles semblent s'efforcer à sécuriser les coupables d'adultère au moment où toutes les énergies doivent être concentrées en vue d'éradiquer cette pratique désormais plus destructrice que jamais .

II.1. Le caractère trop restrictif de l'acte matériel constitutif de l'adultère.

Comme nous l'avons remarqué précédemment, l'acte matériel de l'adultère est la pénétration du sexe mâle dans les parties génitales de la femme.

Ainsi analysé, l'élément matériel de l'adultère s'avère insuffisant à répondre à la finalité de l'incrimination : la protection de la foi conjugale.

En effet, certains actes proches de la conjonction sexuelle peuvent s'avérer parfois aussi intimes, voire plus intimes que la conjonction sexuelle elle-même, à telle enseigne qu'ils revêtent une infidélité particulièrement plus grave.

L'impunité de ce genre d'actes demeure malheureusement garantie par l'exigence de l'acte matériel propre c'est-à-dire la conjonction sexuelle. En outre, le droit pénal interdit toute interprétation extensive de la notion de conjonction sexuelle , évitant ainsi la tentation d'assimiler ces actes. Ainsi, un homme marié qui pratiquerait la fellation ou qui trouverait la jouissance sexuelle en procédant à une autre forme de relation sexuelle évitant toute pénétration, resterait totalement impuni.

Force est de constater également que nonobstant l'absence d'acte de pénétration proprement - dit, les autres formes de rapports non punis par le législateur tels que la fellation, la pénétration dans l'anus etc... sont des formes qui comportent un risque plus grave quant à la transmission du virus du SIDA.

Nous remarquons donc que des actes d'infidélité non punissables peuvent détruire la foi conjugale et la vie des conjoints au même titre sinon plus gravement que la conjonction sexuelle. Quelle marge existe-t-il entre une femme mariée qui se livrerait à une pénétration anale et une autre qui se donnerait normalement (conjonction sexuelle)? Qu'il s'agisse des effets sur son foyer conjugal, qu'il s'agisse des risques de contamination, aucune justification ne permettrait de punir la deuxième tout en laissant impunie la première.

Il faudrait donc étendre la répression aux autres rapports sexuels infidèles.

II.2. L'inégalité de la répression.

En matière d'adultère; cette inégalité de traitement des personnes de sexes différents devant la loi date de la nuit des temps. En effet, le droit ancien, surtout le droit coutumier comparé, nous laisse désarmés devant une inégalité flagrante. Nous nous arrêterons un moment ici pour examiner cet état de choses en droit ancien avant d'aborder le droit moderne, particulièrement le droit rwandais qui, au lieu de s'en démarquer, a préféré perpétuer l'inégalité.

II. 2.1. La répression de l'adultère dans les droits anciens.

La répression de l'adultère dans les droits anciens particulièrement en droit coutumier comparé était à la fois inhumaine, humiliante, cruelle, irrationnelle et inégale.

Chez les Beti, des peines physiques et très humiliantes étaient infligées à l'épouse infidèle, tandis que son complice se voyait imposer des dommages et intérêts qui pouvaient l'amener à livrer à l'époux trompé sa propre soeur. Dans quelques tribus Beti, on ne faisait pas payer une amende au complice de sa femme, mais on lui coupait les oreilles, parce qu'on estimait qu'il n'était pas bon que le mari profite de la faute de sa femme. La femme adultère, elle,

pouvait subir la clitoridectomie avec une mise en scène humiliante qui l'accompagnait, les injections vaginales au piment, l'exposition publique nue etc...⁽¹⁶⁾.

En droit hébreu, on appliquait au coupable la peine de feu et la lapidation. L'infidélité du mari n'était sanctionnée que si elle était perpétrée avec une femme mariée.

Le droit coranique punit également très sévèrement l'adultère. En effet, Mahomet prêchait ce qui suit : "si vos femmes commettent l'action infâme, appelez quatre témoins. Si leurs témoignages réunissent contre elles, enfermez-les dans des maisons jusqu'à ce que la mort les visite ou que Dieu leur procure un moyen de salut"⁽¹⁷⁾.

En droit romain, à l'origine, seul l'adultère de la femme était réprimé. La loi Julia donnait au père le droit de tuer sa fille surprise en flagrant délit d'adultère dans sa maison ou dans celle du gendre⁽¹⁸⁾.

La cruauté et l'irrationnalité du droit coutumier apparaissent aussi dans la prévention, telles les coutumes de la corne de l'Afrique qui permettaient de coudre les femmes mariées pendant l'absence de leurs maris afin qu'elles n'osent se donner à un tiers.

La description ci-haut des droits anciens démontre à plusieurs égards le caractère inhumain, cruel, irrationnel et surtout inégal de ces droits. La philosophie étant que l'adultère est un crime des femmes, comme illustré par ce proverbe cité par KARL PETIT :

"Telle est la conduite de la femme adultère;

Elle mange, puis s'essuie la bouche

En disant : << Je n'ai rien fait de mal >>"⁽¹⁹⁾

l'on peut légitimement se demander si cette femme était seule "à table", ou si l'homme marié, lui, ne fait jamais autant ! Cette conception a-t-elle persisté dans les droits modernes ?

II.2.2. Le droit moderne.

A la fin du XIX^e siècle la plupart des pays punissaient l'adultère et connaissaient encore la rigueur des droits anciens . Ce n'est qu'au début du XXe siècle qu'un double mouvement à la fois d'allègement des sanctions et de dépenalisation vit le jour ⁽²⁰⁾ .

Les partisans de cette incrimination, encore sous l'influence de la religion ou mieux de la morale religieuse ou de la morale tout court estiment que la sanction de l'adultère répond à une nécessité sociale car elle permet non seulement de sauvegarder la paix, la cohésion du foyer conjugal et l'ordre social mais aussi de discipliner le comportement sexuel des citoyens. Il s'agit généralement des pays musulmans, africains, l'Allemagne et certains pays latinistes sous l'influence du droit canon. ⁽²¹⁾ .

Les adversaires du délit d'adultère qui décriminalisent celui-ci et qui sont influencés par l'évolution des moeurs et surtout par le mouvement de l'égalité des sexes, combattent l'incrimination d'adultère pour les raisons suivantes :

- Les sanctions pénales pour adultère sont rarement appliquées et deviennent purement symboliques.
- La plainte de l'époux outragé contre l'époux infidèle a souvent pour but d'obtenir rapidement le divorce devant le juge civil;
- Cette poursuite ou cette menace de poursuites pénales constitue en général, dans l'esprit de beaucoup de plaignants, un chantage contre l'époux coupable.
- Et, enfin, dans les pays ayant dépenalisé cet acte, on ne voit aucune tendance au retour à la sanction pénale.

Ainsi, beaucoup de pays ont déjà dépenalisé l'adultère, entre autres la Norvège (1927), la Suède (1937), l'ex URSS, la France (1975), certains des Etats- Unis d'Amérique dont l'Arkansas, la Louisiane, le Nevada, le Nouveau Mexique, Tennessee etc .. ⁽²²⁾ .

Ainsi, dans ces pays comme en droit français, se trouve consacré le principe selon lequel "en mariage, trompe qui peut" ⁽²³⁾.

Le législateur rwandais quant à lui, n'a pas opté pour la dépenalisation. Mais, comme pour la plupart d'autres pays partisans de l'incrimination, si nous pouvons nous réjouir de la correction des caractères inhumain, cruel, et dégradant de la répression, l'inégalité déplorée n'a pas encore été corrigée.

En effet, l'article 354 du code pénal rwandais persiste à distinguer les peines en fonction du sexe du coupable. Ceci pêche contre le principe de base du "gender" qui veut que l'homme et la femme soient traités de façon égale devant la loi. Cette inégalité devrait disparaître, ainsi que les conditions de poursuite de l'infraction d'adultère.

II.3. Les conditions de poursuite inadéquates.

Comme déjà observé ⁽²⁴⁾, la mise en mouvement de l'action publique est subordonnée à la plainte préalable de l'époux offensé. De même, le désistement de ce dernier éteint l'action publique et le conjoint offensé dispose encore, même après la condamnation, d'un "droit de grâce" *sui generis*.

Certes, ces dispositions visaient à protéger l'intégrité du foyer conjugal. Cependant, aujourd'hui, spécialement à cause du danger que constitue le SIDA, il y a lieu de balancer la nécessité de la protection de l'intégrité du foyer conjugal avec la nécessité de protéger les vies humaines mises en danger par l'acte d'adultère sans négliger certainement la chaîne des conséquences sociales, économiques et autres.

Cette dernière nécessité devrait conduire à l'abolition des conditions précitées, afin de permettre au ministère public de protéger plus efficacement la société, sans que ses mains soient liées par des particuliers.

II.4. La rigueur des modes de preuve exigés contre le complice.

Le flagrant délit et les écrits faits par lui-même étant les seuls modes de preuve admis contre le complice d'adultère, l'impunité de ce dernier semble presque garantie dans la mesure où ces preuves sont extrêmement rares et introuvables.

En outre, cette exigence est manifestement injuste car l'auteur principal peut être puni sur base d'autres moyens de preuve réunis à la fois contre le complice et que ce dernier échappe à la condamnation. N'est-il pas injuste et même scandaleux de condamner une personne qui a avoué l'adultère en acquittant son complice qui, aussi, a avoué au seul motif qu'il y a défaut d'écrit ou de flagrant délit exigés par la loi ?

En conséquence, il faudrait permettre que tous les moyens soient utilisés pour prouver cette infraction aussi bien contre l'auteur que son complice, comme il est de principe en droit pénal.

Nous remarquons ainsi que le régime juridique de l'adultère prête le flanc à des graves critiques et peut ne plus résister aux impératifs du moment. Ce qui nous impose de formuler des recommandations précises en vue d'adapter et d'actualiser les dispositions relatives à l'adultère.

Nous avons observé le grave danger que représente le SIDA pour la survie de l'humanité tout entière et de la société rwandaise en particulier. Cette gravité impose des mesures à la hauteur du danger. Ces mesures doivent être généralisées sans négliger aucun domaine de la vie sociale.

Le domaine juridique présente un champ très vaste d'intervention et les vides à combler sont très nombreux, entre autre la réglementation de l'activité sexuelle chez les personnes mariées.

En effet, nous avons remarqué qu'en matière d'adultère, les lois en vigueur semblent dépassées par les événements. D'une part, elles lient les mains de l'autorité publique qui devrait combattre l'adultère, en exigeant des procédures trop restrictives, d'autre part elles perpétuent l'inégalité des sexes, concept révolu. La loi devrait suivre les faits.

Le législateur rwandais contribuerait ainsi à la lutte contre le désordre sexuel et par conséquent à la lutte contre le SIDA, et à la promotion de l'égalité des sexes en prenant les mesures suivantes :

1. L'extension de l'acte matériel de l'infraction d'adultère à tous les rapports sexuels c'est à dire les actes intimes impliquant le sexe, peu importe qu'il y ait pénétration ou non.
2. La suppression de l'exigence de la plainte préalable aux poursuites et des prérogatives reconnues au conjoint offensé lui permettant de mettre fin aux poursuites et à la condamnation.
3. La suppression de l'inégalité dans la répression et l'adoption des peines plus graves;
4. La suppression des restrictions en matière de preuves en adoptant le principe habituel de la liberté de preuves aussi bien pour l'auteur de l'adultère que pour son complice.

NOTES DE REFERENCES

- (1) Voir LIKULIA B, *Droit pénal spécial*, Tome 1, 2^e Ed, LGDJ, Paris, 1985, p 277.
- (2) Les pahouins sont une ethnie constituée des Beti, Fang et Boulau. Cette ethnie est installée au Cameroun, au Gabon, au Congo Brazzaville et en Guinée Equatoriale. (V.OMBOLO J.P, *Sexe et société en Afrique Noire*, Ed l'Harmattan, Paris, 1990, p 31).
- (3) Voir OMBOLO JP, *op cit*, p 369
- (4) *Idem*, p 370
- (5) ONUSIDA, *le point sur l'épidémie du SIDA*, Genève, 2002, pp 2, 34, 36 et 37.
- (6) PNLs, *Enquête de séroprévalence de l'infection à VIH dans la population générale*. Synthèse des résultats, 1997, p 6.
- (7) KOUPELNIK C, *le témoin myope. Regards sur le SIDA*, Ed l'Harmattan, Paris, 1995, p 16.
- (8) *Idem*, p 177.
- (9) LIKULIA B, *op cit*, p 282.
- (10) *Idem*, p 276
- (11) Voir note infra-paginale (2), p 1
- (12) OMBOLO JP, *op cit*, p 237.
- (13) LIKULIA B, *op cit*, p 276
- (14) *Idem*, p 277.
- (15) *idem*, p 285
- (16) OMBOLO JP, *idem*.
- (17) *Idem*, p 146
- (18) LIKULIA B, *op cit*, p 277
- (19) PETIT K, *le dictionnaire des citations du monde entier*, Ed Marabout, 1978, p 16
- (20) LIKULIA B, *op cit*, p 277.
- (21) *ibidem*.
- (22) *Idem*, p 278
- (23) VITU, A et MERLE R, *Traité de droit criminel. Droit pénal spécial*, Ed cujas, Paris, 1982, p 1665
- (24) *Voir* p 207.

BIBLIOGRAPHIE.

OUVRAGES.

1. BOUNAN (M), *le temps du SIDA*, Ed ALLIA, 1990, 151 p.
2. DAUDEL (R) et MONTAGNIER (L), *Le SIDA*, Ed Flammarion, Paris, 1994, 127p.
3. KOUPERNIC (C) *le témoin myope. Regards sur le SIDA*. Ed l'Harmattan, Paris, 1995, 188p
4. LIKULIA (B) , *Droit pénal spécial zairois*, Tome I, 2^e éd, LGDJ, Paris, 1985, 600p
5. OMBOLO (JP), *Sexe et Société en Afrique Noire* , Ed l'Harmattan, Paris, 1990, 395 p
6. PETIT (K), *le dictionnaire des citations du monde entier*, Ed Marabout, 1978.
7. VITU (A) et MERLE (R), *Traité de droit criminel. Droit pénal spécial*, Ed cujas, Paris, 1982, 2111p.

RAPPORTS DIVERS.

1. ONUSIDA, *le point sur l'épidémie du SIDA*, Genève, 2002, 37p
2. PNLS, *Enquête de séroprévalence de l'infection du VIH dans la population générale. Synthèse des résultants*. Kigali, 1997, 8p

TEXTES DE LOI.

- Décret- Loi instituant le code pénal , *in J.O*, n° 13 bis du 1^{er} juillet 1978.

*Tiré sur les presses de l'imprimerie COFICO
B.P. 439 Kigali-Rwanda
Tél./Fax : (250) 82358*



