La modélisation Econométrique des Transitions Individuelles sur le Marché du Travail*

by

Jean-Pierre Florens, Denis Fougere***, Thierry Kamionka**** and Michel Mouchart*****

February 1995

Abstract

This paper presents recent econometric models of individual mobility on the labor market. The general framework is given by the theory of point processes which are continuous time finite state space processes. We keep the mathematical formalism at the lowest possible level and a specific attention is paid to the main basic concepts. The first part of the paper is devoted to specification problems and the inference of the introduced models is not treated. We present some structural models leading to point processes and some applications to french labor market are summarized. The last part of the paper consider the introduction of unobservable heterogeneity and of endogeneous sampling in specification and inference analysis.

* Remerciements : Cet article synthétise deux textes antérieurs, l'un préparé par J.P. Florens et T. Kamionka pour une communication invitée aux Journées de l'AFSE (Dijon, Mai 1993), et l'autre par D. Fougère et M. Mouchart pour une Journée d'Etude du GDR 995 du CNRS (Toulouse, Mars 1992). Cette synthèse est essentiellement fondée sur un ensemble de travaux théoriques et appliqués réalisés au sein d'une équipe qui, outre les quatre auteurs de cet article, comprend Liliane Bonnal, Catherine Cazals, Sergio Pastorello et Anne Sérandon. Il nous appartient de les remercier vivement au moment de puiser dans des résultats auxquels ils ont collaboré. Enfin, des commentaires de R. Gouet nous ont été précieux au moment de la révision finale de cet article. Le quatrième auteur remercie le CEME, Université de Paris I, pour l'accueil offert au cours de l'élaboration de ce travail, et le Pôle d'Action Interuniversitaire (PAI).

** IDEI, GREMAQ, Université des Sciences Sociales de Toulouse, Toulouse, France.
*** CNRS, CREST-ENSAE, Paris, France.
**** CNRS et GREMAQ, Université des Sciences Sociales de Toulouse, Toulouse, France.
***** Institut de Statistique and CORE, Université Catholique de Louvain, Voie du Roman Pays 34, 1348 Louvain-la-Neuve, Belgium.

1 INTRODUCTION

De nombreux phénomènes micro-économiques, notamment sur le marché de l'emploi, présentent les caractéristiques suivantes :

(i) la variable d'intérêt, ou variable endogène, est "qualitative" : elle peut prendre un nombre fini de "valeurs" et ces valeurs appartiennent à un espace "abstrait", c'est-à-dire sans structure mathématique ; ce sont par exemple les "états" possibles d'un individu sur le marché du travail ;

(ii) l'individu peut changer d'état "n'importe quand", à l'opposé, par exemple, de phénomènes purement saisonniers où des changements d'état se font à des périodes (et donc des périodicités) prédéterminées.

Il apparaît dès lors naturel de modéliser l'évolution (ou la mobilité) d'un individu dans un tel système par un processus aléatoire (X_t) où, pour tout t, X_t prend sa valeur dans un espace fini et où t varie, de façon continue, dans un intervalle I; en général, on considèrera $I = [0, a]$ ou $I = [0, \infty)$. D'un point de vue probabiliste, on parlera alors de "*processus ponctuel*".

Au-delà du marché de l'emploi, il existe de nombreux exemples qui se modélisent naturellement comme des processus ponctuels en micro-économie. Citons, en particulier, des problèmes en économie financière (état de solvabilité d'une entreprise), en assurance (fréquence des accidents d'automobile), en gestion commerciale (problème de fidélité à une marque), etc. Cet article n'a pas pour ambition de couvrir l'ensemble des domaines d'application des processus ponctuels. Son objet est de montrer leur particulière utilité pour l'analyse statistique et économétrique des mobilités individuelles sur le marché du travail . Notre espoir est que le lecteur puisse, à partir de notre présentation, imaginer sans trop de difficultés comment appliquer ce type de modélisation à d'autres situations micro-économiques.

Dans la formalisation envisagée ici, une mobilité vécue par un travailleur à l'instant t correspond à une modification de la valeur du processus aléatoire (X_t) à cet instant ; l'individu réalise alors une transition d'un état vers un autre, par exemple du chômage vers l'emploi. Ainsi donc, les modèles économiques de transitions sur le marché du travail sont destinés à décrire et à analyser des observations essentiellement constituées de trajectoires d'individus circulant à l'intérieur d'un ensemble fini d'états. Plus précisément, la situation d'un individu vis à vis du marché du travail est résumée par le fait qu'il se trouve dans l'un des états d'un ensemble $E = \{1, ..., k, ..., K\}$ (par exemple,1≡ emploi stable, 2≡ emploi précaire, 3≡ chômage, 4≡ inactivité). Les observations sont donc des trajectoires $X_t \in$ E, t $\in I$ pour tout individu i = 1, ..., n ; $X_t = k$ signifie que l'individu est dans l'état k au temps t.

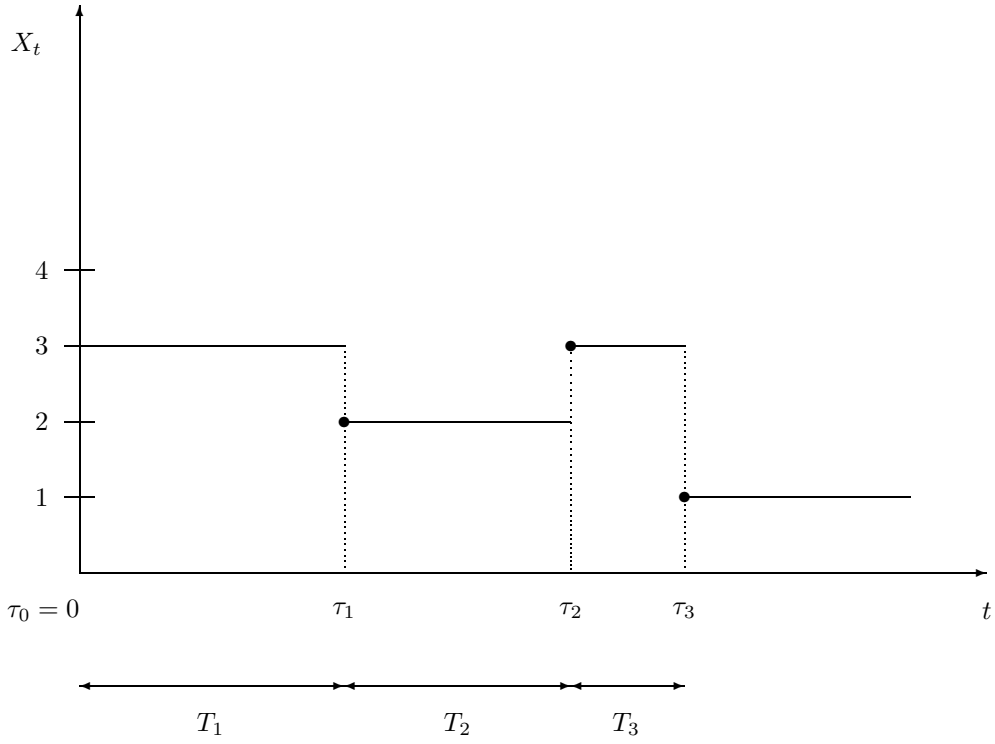
Pour illustrer notre propos, imaginons une réalisation particulière d'un processus (X_t) dans l'ensemble d'états E préalablement défini. Cette réalisation, représentée sur la Figure 1, comprend tout d'abord un séjour dans l'état de

chômage (indicé par 3), d'une durée égale à $T_{1,}$ commençant à la date τ_0 (égale à 0 par convention et correspondant à la première entrée du travailleur sur le marché de l'emploi) et se terminant en τ_1. Cette période de chômage est suivie d'une période d'emploi précaire (état indicé 2), débutant en τ_1 et finissant en τ_2, donc de durée égale à $T_2 = \tau_2 - \tau_1$. A la date τ_2 , l'individu retourne en chômage jusqu'à la date τ_3 , où il est embauché dans un emploi sur CDI (état indicé 1), et ainsi de suite.

Cette trajectoire peut en particulier être totalement décrite par la séquence $(\tau_p, x_p)_{p\geq o}, (\tau_0 = 0)$, des temps de passage et des états atteints : l'individu est en $x_o \in E$ en $t = 0$, puis passe en $x_1 \in E$ en $t = \tau_1$, puis en x_2 en $t = \tau_2$, etc....Bien que le modèle décrit soit naturellement en temps continu, on voit donc que les observations sont représentables par des suites discrètes de variables aléatoires.

L'observation des trajectoires individuelles constituent une source d'information très importante et très riche dans la mesure où elles permettent d'analyser un phénomène à la fois dans sa dimension individuelle (seront introduites des caractéristiques individuelles observées ou non observées) et dans sa dimension temporelle, dans laquelle il faut s'efforcer de séparer la dynamique spécifique de chaque individu de l'évolution générale de son environnement.

Figure 1 : Une réalisation d'un processus de transition sur le marché du travail X_t



Cette double dimension des observations les situe dans le cadre général des "données de panel". L'intérêt dans ces modèles est toutefois fortement concentré sur la dynamique individuelle des phénomènes : on cherche à savoir par exemple si la probabilité de sortir du chômage vers une forme donnée d'emploi dépend de la durée passée en chômage, mais aussi de la totalité de l'histoire de l'individu (nombre de périodes de chômage, durée totale du chômage et de l'emploi, âge, etc...), ou encore l'on cherche à mettre en évidence la succession des transitions par l'emploi partiel entre l'emploi à temps complet et la retraite. L'observation de la dynamique permet d'identifier des décalages chronologiques entre événements (perte des indemnisations et sortie du chômage, mariage et passage à l'inactivité)

pouvant s'interpréter comme des causalités, du moins au sens statistique utilisé en économétrie.

Le microéconomètre du marché du travail rencontre de plus en plus fréquemment des bases de données contenant des observations de trajectoires individuelles. Ces observations sont souvent partielles (observation du phénomène en certains instants seulement) ou entachées de biais de sélection (on suit la trajectoire d'une fraction d'individus chômeurs en août d'une certaine année, ou l'on observe des individus ayant connu une période de chômage de plus d'un an), mais la référence à un modèle latent de transitions en temps continu est nécessaire pour un traitement correct de ces observations incomplètes.

Les instituts de statistiques nationaux collectent de plus en plus souvent des données de panel en réinterrogeant plusieurs fois des individus. L'enquête Emploi de l'INSEE permet ainsi de suivre pendant plusieurs périodes un échantillon représentatif de la population française. Des enquêtes comme "Suivi des Chômeurs" ont permis de retracer en partie au moyen d'interrogations successives des trajectoires de sous-populations. Dans de nombreux pays, on dispose de données de panel portant sur la consommation et la situation relative au marché du travail.

D'autres sources de données sont des enquêtes biographiques où lors d'une seule interrogation, l'on s'efforce de retracer rétrospectivement la trajectoire d'un individu (enquête "Situations défavorisées" de l'INSEE, enquêtes de l'INED (voir Courgeau et Lelièvre (1989)). Ces données ont en particulier l'intérêt de retracer des historiques longs, mais les défauts de mémoire des individus engendrent des erreurs d'observation ou interdisent des observations très précises.

Des observations de trajectoires individuelles proviennent souvent des fichiers constitués à des fins administratives par des organismes tels l'ANPE, l'UNEDIC, des Caisses de retraite, des mutuelles. Ces organismes gèrent des clients dont on garde l'histoire de la situation. Le développement de l'informatisation a considérablement accru le nombre de tels fichiers qui constituent pour les économistes une source nouvelle et extrêmement importante d'informations.

Les défauts de ces différentes sources d'informations ne sont pas identiques. Les enquêtes des instituts de statistiques sont en général représentatives mais n'observent les individus que durant une courte période et de manière discontinue. Les reconstitutions de biographies sont entachées d'erreurs d'observation alors que les fichiers administratifs observent des individus pendant de longues périodes et avec précision mais présentent par définition des biais de sélection importants. De plus, les variables retenues ne sont pas toujours les grandeurs les plus pertinentes en regard de la théorie économique.

La démarche théorique d'analyse des données de transition consiste d'abord à modéliser les trajectoires, c'est-à-dire à décrire leur loi de probabilité, en général

conditionnelle à des variables explicatives. Ces lois de probabilité dépendent d'éléments inconnus (paramètres appartenant à un espace euclidien de dimension finie ou paramètres fonctionnels) que l'on estimera à partir des observations. Nous présenterons dans la section suivante l'essentiel de ces modèles statistiques en centrant notre présentation sur le rôle des intensités stochastiques de transition comme outil de spécification et d'analyse des propriétés dynamiques des modèles statistiques de transitions. La troisième section esquisse brièvement le type de raisonnement économique conduisant à des modèles de la classe évoquée dans la section antérieure. Cette approche plus structurelle des mobilités individuelles décompose essentiellement les probabilités instantanées de transition en diverses composantes, s'interprétant comme résultant de choix individuels et de chocs sur le marché du travail. La quatrième section aborde une des principales difficultés rencontrées en pratique : la prise en compte dans les modèles du caractère incomplet ou tronqué des trajectoires effectivement disponibles, qui souvent ne sont observées qu'à des instants précis et non de manière continue (en suivi ou en rétrospectif). C'est le cas dans certains panels extraits des Enquêtes Emploi de l'INSEE. Enfin, les deux dernières sections de cet article de synthèse seront consacrées à l'examen de deux aspects importants des modèles de transition : le mode de sélection de l'échantillon et l'hétérogénéité non observée des individus.

2 Modélisation Mathématique de Processus Ponctuels

2.1 Définition et Hypothèses de base

L'exposé qui suit sera simplifié en s'intéressant, à un seul individu; dans cette section on supprimera donc l'indice i des individus.

Considérons un ensemble *fini* d'états $E = \{1, 2, ..., K\}$ et un espace des temps sous forme d'un intervalle I du type $I = [0, a]$ ou $I = [0, \infty)$. Une fonction définie sur I et à valeurs dans E sera appelée une *Trajectoire*. La figure 1 ci-dessus représente un exemple de trajectoire dans les "bons" cas.

Un *processus ponctuel* est une fonction (mesurable) définie sur un espace probabilisé et à valeurs dans un espace de trajectoires. De façon plus spécifique, soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé qui représente ici un environnement économique (aléatoire) et soit X une fonction mesurable définie sur Ω et à valeurs dans E^I :

$$X : \Omega \to E^I, \omega \to (X_t(\omega) : t \in I) \ \ avec\ X_t(\omega) \in E,$$

où E est muni de sa tribu des parties (car E est finie) et I de sa tribu borélienne (car il sera plus tard nécessaire de pouvoir intégrer des trajectoires sur l'espace des temps). En d'autres termes, pour tout Borélien B de I et tout état $k \in E$, la fonction X est telle que :

$$\{\omega \,|X_t(\omega) = k, \forall t \in B\} \in \mathcal{A}.$$

Cette condition de mesurabilité n'est cependant pas suffisante car l'analyse qui va suivre demandera que soient de "bonnes variables aléatoires" (c'est-à-dire des fonctions mesurables) des fonctions définies sur Ω et à valeurs réelles positives telles que :

(i) les instants d'accès aux états, par exemple : $\inf \{t \,|X_t(\omega) = k\}$,

(ii) les proportions de temps passé dans un état donné, par exemple :

$$(b-a)^{-1} \int_{[a,b]} \mathbb{1}_{\{X_t(\omega)=k\}} \, dt.$$

Une condition suffisante pour la mesurabilité de ces fonctions est que les trajectoires $(X_t(\omega) : t \in I)$ soient des fonctions (de t, vu que E est fini !) continues à droite (c'est-à-dire $X_{t_+}(\omega)$ existe et $X_{t_+}(\omega) = X_t(\omega), \forall t \in I$ et $\forall \omega \in \Omega$) avec limite à gauche (c'est-à-dire $X_{t_-}(\omega)$ existe, $\forall t \in I$ et $\forall \omega \in \Omega$). On parlera alors de processus à trajectoires CADLAG ; on ne discutera pas ici de modifications négligeables de trajectoires. Le "bon" cas de trajectoire de la Figure 1 représente précisément cette propriété. Dans la suite, on ne s'intéressera qu'à de tels processus.

2.2 Caractérisation et processus dérivés (ou marginaux) de processus ponctuels

2.2.1 Remarque introductive

L'aspect "temps continu" des processus ponctuels pose des difficultés pratiques à la fois au niveau de la représentation des trajectoires (Comment "résumer" numériquement la figure 1, par exemple ?) et de la modélisation. Dans cette section, nous allons montrer comment caractériser l'information fournie par l'observation d'un processus ponctuel de différentes façons équivalentes. Dans la suite, nous verrons que ces différentes représentations offrent des facilités différentes de modélisation, particulièrement en vue de tenir compte du fait qu'on n'a pas souvent la possibilité d'observer "complètement" les trajectoires d'un processus ponctuel.

2.2.2 Représentations discrètes

Une conséquence importante de la propriété CADLAG est que les changements d'états auront lieu au plus un nombre dénombrable de fois et que les durées de permanence dans un même état seront des variables aléatoires presque sûrement strictement positives. Dès lors, le processus en temps continu $(X_t : t \in I)$ peut être représenté de façon équivalente du point de vue aléatoire, par des processus bivariés en temps discret. De façon plus précise, définissons à partir du processus $X = (X_t : t \in I)$ les processus, en temps discret, suivants :

(i) un processus d'instants (aléatoires) de changement d'état, défini récursivement comme suit :

$$\tau_o = 0,\ \tau_p = \inf\left\{t \mid t > \tau_{p-1},\ X_t \neq X_{\tau_{p-1}}\right\},$$

et donc, $\tau_p : \Omega \to I\!R_+, \forall p \in I\!N^*$, avec $\tau_p > \tau_{p-1}$ presque sûrement.

(ii) un processus de durée du séjour dans un même état, défini récursivement comme suit :
$T_0 = 0,\ T_p = \tau_p - \tau_{p-1}$.

En d'autres termes, T_p représente la durée du séjour dans l'état $X_{\tau_{p-1}}$ entre la transition (p-1) et la transition p. Il est utile de remarquer que :
$\tau_p = \sum_{0 \leq i \leq p} T_i$

Dès lors, l'information fournie par l'observation (d'une trajectoire complète) du processus X est équivalente à celle fournie par les (trajectoires complètes de) processus en temps discret :

$$\textit{soit} \qquad \left\{\left(\tau_p, X_{\tau_p}\right) : p \in I\!N\right\}, \qquad \begin{array}{ll} \textit{où} & \left(\tau_p, X_{\tau_p}\right) \in I\!R_+ \times E \\ \textit{avec} & \tau_p > \tau_{p-1} \textit{ presque sûrement,} \end{array}$$

$$\textit{soit} \quad \left\{\left(T_p, X_{\sum\limits_{0 \leq i \leq p} T_i}\right) : p \in I\!N\right\}, \quad \begin{array}{ll} \textit{où} & \left(T_p, X_{\sum_{0 \leq i \leq p} T_i}\right) \in I\!R_+ \times E \\ \textit{avec} & T_p > 0 \textit{ presque sûrement.} \end{array}$$

2.2.3 Représentations par processus de dénombrement

A un processus ponctuel X, on peut aussi associer un processus multivarié de dénombrement qui lui est équivalent (au point de vue de l'information fournie). Rappelons qu'un processus univarié de dénombrement est un processus aléatoire en temps continu et espace d'états égal à $I\!N$ tel que ses trajectoires, supposées CADLAG, sont non décroissantes et de sauts égaux à un. On peut interpréter de tels processus comme processus consistant à compter, en accumulant, le nombre de réalisations d'un "événement ponctuel typique", tel que, par exemple, le nombre d'accidents à un carrefour donné comme fonction du temps, en supposant qu'il n'y a jamais deux, ou plus, accidents simultanément. Un processus multivarié de dénombrement est simplement une famille finie de processus de dénombrement définis sur un même espace de probabilité.

En vue de la représentation de processus ponctuels, considérons comme "événements ponctuels typiques" les changements d'état. L'ensemble des transitions (j, k) de probabilité non nulle sera écrit C:

$$(j,k) \in C \subset (E \times E)$$

Dans $E \times E$, on excluera, au moins, la diagonale principale : $\{(k,k)|k \in E\}$ mais aussi, dans certaines applications, d'autres transitions de probabilité nulle.

On peut alors représenter le processus ponctuel X par le processus multivarié de dénombrement N :

$$N = (N_{j,k}(t) : (j,k) \in C, t \in I)$$

où $N_{j,k}(t)$ représente le dénombrement des transitions de j vers k au fil du temps et réalise donc un saut d'une unité aux instants où se réalise une transition de j vers k.

2.3 Typologie de processus ponctuels

Au point de vue de la modélisation, il convient de distinguer les processus ponctuels selon le nombre de transitions possibles et selon le nombre d'états possibles.

Certains processus, appelés parfois "processus de survie", ou encore "modèle simple de durée", admettent un seul changement d'état (la fin de la "vie"). Ils sont caractérisés par une durée univariée et un seul état de sortie. D'autres processus admettent plusieurs changements d'états. Ils sont caractérisés par des durées multivariées et des états de sortie à la fin de chaque durée.

De même, au point de vue du nombre des états possibles, certains processus n'ont que deux états possibles (par exemple,"chômer" ou "ne pas chômer") tandis que d'autres processus admettent plus de deux états (par exemple, en distinguant le chômage, l'inactivité et l'emploi). Dans ce cas, une classe particulièrement importante de modèles est celle des *modèles à risques concurrents* : il y a une concurrence possible entre la réalisation de plusieurs événements et à chacun de ces événements est associée une sortie vers un état différent. Par exemple, un chômeur peut sortir du chômage dès qu'il reçoit et accepte une offre d'emploi ou dès qu'il reçoit et accepte une possibilité de formation complémentaire. La sortie observée est alors celle qui est associée au premier de ces événements qui se réalise.

La suite de cette section aura pour objectif d'esquisser quelques modèles relatifs à des processus ponctuels de différents niveaux de complexité et de montrer, en particulier, que négliger la possibilité de plus d'une transition ou de plus d'un état de sortie peut faire perdre des propriétés essentielles d'un modèle économique et entrainer ainsi de graves erreurs de spécification. Nous considérons ainsi trois classes de processus, de complexité croissante : les processus de survie (ou modèles élémentaires de durée) fréquemment utilisés en micro-économétrie du marché du travail, les processus de dénombrement univariés qui comptent typiquement le nombre d'occurrences d'un même événement au sein d'un intervalle donné, et permettent d'introduire des concepts fondamentaux, tels que ceux d'intensité stochas-

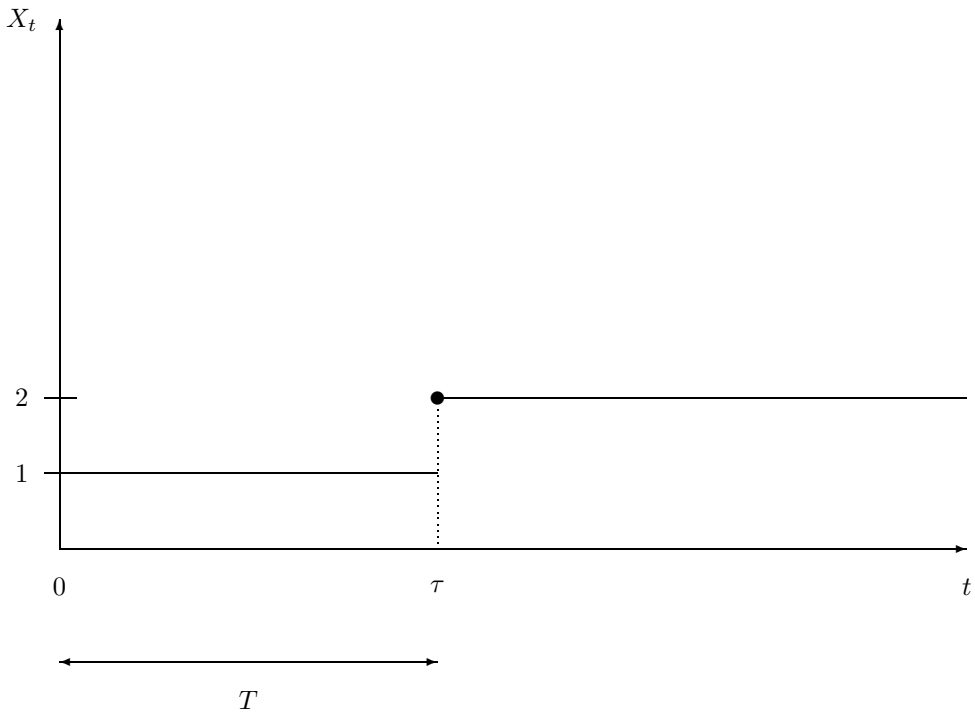
tique et de compensateur, et enfin les processus de transition qui permettent de distinguer plusieurs états possibles et plusieurs sauts.

2.4 Modèles de durée

2.4.1 Définition

La situation la plus simple se rencontre dans des processus de survie , c'est-à-dire des processus ponctuels caractérisés par deux états {1,2} avec $X_o = 1$ et une seule transition. Dans les commentaires qui suivent, on interprètera 1 comme "la vie" et 2 comme "la mort". Une trajectoire d'un tel processus est alors caractérisée par une seule variable aléatoire positive : la durée de permanence dans l'état initial (de vie). On parlera alors de "modèle de durée" sans nécessairement se référer explicitement à un processus ponctuel sous-jacent. La figure 2 représente une trajectoire d'un tel processus.

Figure 2 : Trajectoire typique d'un processus de survie



2.4.2 Caractérisation de distributions de durées

Les modèles de durée sont aujourd'hui devenus d'utilisation courante en économétrie du marché de l'emploi (voir, par exemple, Heckman et Singer [1984, 1985], Kiefer [1988], Lancaster [1990]). On ne s'y étendra donc pas, mais on en rappelera seulement quelques concepts essentiels pour introduire les notations et dans la mesure où celles-ci se retrouvent dans la modélisation de processus plus complexes.

Rappelons d'abord que la distribution d'une durée T, c'est-à-dire d'une variable aléatoire non-négative, peut être caractérisée par plusieurs fonctions. Pour

plus de brièveté, rappelons dans le seul cas continu leur notation, définition et quelques unes des relations qui les relient :

Fonction de distribution :

$$F_T(t) \quad = \quad prob\,(T \leq t)$$

Fonction de survie :

$$S_T(t) \quad = \quad 1 - F_T(t) \quad = \quad prob\,(T > t) = \int_t^{\infty} f_T(u)\, du$$

$$= \quad \exp\{-\Lambda_T(t)\} \quad = \quad \exp\left\{-\int_0^t \lambda_T(u)\, du\right\}$$

Fonction de densité :

$$f_T(t) \quad = \quad \frac{d}{dt} F_T(t) \quad = \quad -\frac{d}{dt} S_T(t)$$

$$= \quad \lambda_T(t)\ S_T(t) \quad = \quad \lambda_T(t) \exp\left\{-\int_0^t \lambda_T(u)\ du\right\}$$

Fonction de hasard :

$$\lambda_T(t) \quad = \quad \frac{f_T(t)}{S_T(t)}$$

Fonction de hasard intégré :

$$\Lambda_T(t) \quad = \quad -\ln S_T(t) \quad = \quad \int_0^t \lambda_T(u)\, du.$$

Rappelons aussi qu'une fonction de densité peut s'interpréter comme une "probabilité instantanée de décès", c'est-à-dire :

$$f_T(t) = \lim_{\delta \downarrow 0} \frac{prob\,[t < T \leq t + \delta]}{\delta}$$

et qu'une fonction de hasard peut s'interpréter comme une "probabilité instantanée de décès" *conditionnelle* au fait d'avoir survécu jusqu'à l'instant immédiatement antérieur, c'est-à-dire :

$$\lambda_T(t) = \lim_{\delta \downarrow 0} \frac{prob\,[t < T \leq t + \delta \mid T > t]}{\delta}$$

Un intérêt majeur des fonctions de hasard est donc de fournir un outil naturel pour une approche dynamique de la modélisation, à savoir décrire en probabilité l'évolution d'un système conditionnellement à son histoire. Souvent, la modélisation économique donnera des informations sur l'allure générale de la fonction de hasard : par exemple, certains modèles économiques de recherche d'emploi impliquent que la probabilité de sortie du chômage dépend d'un salaire de réserve

qui serait fonction, entre autres facteurs, du temps déjà passé dans l'état de chômage ; dès lors, on peut s'attendre à une fonction de hasard non-décroissante, à moins que des éléments de signaux n'incitent les entreprises à préférer l'embauche de chômeurs d'ancienneté moindre. En ce cas, on s'attendrait à une fonction de hasard décroissante et une combinaison de ces deux facteurs pourrait conduire à une fonction de hasard unimodale. Une caractéristique importante d'une distribution de durée sera donc la forme de sa fonction de hasard : on s'intéressera plus particulièrement au fait que cette fonction soit constante (distribution exponentielle), monotone (distribution gamma ou Weibull) ou unimodale (distribution log-normale, log-logistique ou des généralisations des distributions antérieures).

On trouvera des estimations de modèles de durée du chômage sur données individuelles françaises dans les articles de Bonnal et Fougère (1990), Joutard et Werquin (1992), Ruggiero (1992) ou Cases et Lollivier (1993a).

2.5 Modélisation de processus de dénombrement univariés

2.5.1 Introduction

Une fois que l'on considère des processus plus complexes que le modèle simple de durée (c'est-à-dire, plus de deux états et/ou plus d'une seule transition), les problèmes de modélisation gagnent beaucoup en substance vu, notamment, les multiplicités de représentations évoquées en section 2.2. Bien que n'étant pas encore d'un emploi très courant en économétrie du travail, ces modélisations plus complexes ont connu récemment un intérêt considérable (voir, par exemple, Cazals [1992], Fougère et Kamionka [1992d], Kamionka [1992], Visser [1992], Bonnal, Fougère et Sérandon [1994]).

Sur le plan économique, l'intérêt porte essentiellement sur l'aspect dynamique de la modélisation, c'est-à-dire sur la spécification du rôle de l'histoire du processus quant à son évolution aléatoire. Ce problème général de spécification dynamique est plus délicat en temps continu qu'en temps discret. Par exemple, en temps discret, le modèle le plus simple est celui de l'indépendance complète ; mais un processus où les X_t seraient mutuellement indépendants en temps continu donnerait, avec probabilité un, des trajectoires qui ne seraient pas CADLAG (ni même mesurables en t) car, sur tout intervalle ouvert de I, il y aurait une infinité non-dénombrable de changements d'état : un modèle clairement peu intéressant pour décrire le comportement d'un individu !

Deux outils vont se révéler particulièrement performants pour affronter ces difficultés. D'une part, le théorème fondamental en théorie des processus, dû à Kolmogorov, montre que la loi d'un processus en temps continu (ou à espace de temps infini dénombrable) est caractérisée par l'ensemble de toutes ses distributions marginales fini-dimensionnelles, c'est-à-dire par l'ensemble de toutes les distributions multivariées de $(X_{t_1}, X_{t_2}, ..., X_{t_p})$ pour tout $t_1 < t_2 < ... < t_p$ et tout p

fini (on les appelle un "système projectif" car cet ensemble de distributions est contraint par une condition évidente de cohérence vis–à-vis de la marginalisation). On pourra donc caractériser certaines propriétés structurelles d'un processus ponctuel, telle que la markovianité, à partir de ces seules distributions fini-dimensionnelles.

Le second outil, fourni par la décomposition de Doob-Meyer, consiste heuristiquement à étendre au cas continu la décomposition usuelle des processus en temps discret (du type ARMA, par exemple) en une somme de l'espérance conditionnelle au passé et de l'innovation. L'objet de cette section est d'esquisser les grandes lignes de cet outil lorsqu'il est appliqué à un processus de dénombrement univarié, comptant le nombre d'occurrences d'un certain événement au cours d'un intervalle de temps. Nous envisageons dans la section suivante l'utilisation de cet outil dans un cas plus général, mais plus compliqué, celui des processus de transition à plusieurs états.

2.5.2 Intensité stochastique et compensateur

a) Introduction. Considérons un processus de dénombrement univarié $N = (N_t : t \in I)$ et remarquons que, semblablement à un processus ponctuel, il peut être représenté par le processus en temps discret des instants de saut
$\tau = (\tau_p : p \in \mathbb{N})$, avec $\tau_p \in \mathbb{R}_+, \tau_p > \tau_{p-1}$ presque sûrement, et :

$$\tau_p = \inf \{t \mid N_t = p\}, \text{avec } \tau_0 = 0.$$

En effet, le processus N_t peut être représenté de la façon suivante :

$$N_t = \sum_{p \geq 1} \mathbb{1}_{\{\tau_p \leq t\}} = \sum_{\tau_p \leq t} \delta_{\tau_p},$$

où δ_{τ_p} représente la mesure de probabilité qui donne la masse 1 au point τ_p. On pourra donc considérer un processus de dénombrement comme une mesure aléatoire (à support discret). Dans la suite, on écrira indifféremment N_t ou $N(t)$. Il peut aussi être représenté par le processus en temps discret des intervalles entre les sauts : $T = (T_p : p \in \mathbb{N})$, avec $T_p \in \mathbb{R}_+$, défini par :

$T_p = \tau_p - \tau_{p-1}$, avec $T_0 = 0$ et $\tau_p = \sum_{1 \leq i \leq p} T_i$.

Les lois de ces représentations en temps discret sont caractérisées par la suite des distributions conditionnelles générant $(\tau_{p+1} | \tau_1, \tau_2, ..., \tau_p)$ ou $(T_{p+1} | T_1, T_2, ..., T_p)$ pour $p \geq 0$, celles-ci étant reliées entre elles par les relations :

$$S_{\tau,p+1}\left(t \mid \tau_1,...,\tau_p\right) = S_{T,p+1}\left(t - \textstyle\sum_{1\leq i\leq p} T_i \mid T_1,T_2,...,T_p\right), t > \tau_p = \textstyle\sum_{1\leq i\leq p} T_i,$$

$$\lambda_{\tau,p+1}\left(t \mid \tau_1,...,\tau_p\right) = \lambda_{T,p+1}\left(t - \textstyle\sum_{1\leq i\leq p} T_i \mid T_1,T_2,...,T_p\right)$$

$$S_{T,p+1}\left(t \mid T_1,T_2,...,T_p\right) = S_{\tau,p+1}\left(t+\tau_p \mid \tau_1,...,\tau_p\right), t > 0,$$

$$\lambda_{T,p+1}\left(t \mid T_1,T_2,...,T_p\right) = \lambda_{\tau,p+1}\left(t+\tau_p \mid \tau_1,...,\tau_p\right).$$

b) Décomposition de Doob-Meyer. Nous allons reconsidérer la fonction de hasard, qui caractérisait la distribution d'une variable aléatoire de durée à la section 2.4, du point de vue du processus de dénombrement sous-jacent. Dans ce but, notons $\mathcal{F}_N(t)$ l'"histoire" du processus N_u jusqu'au temps t, c'est-à-dire l'information fournie par l'observation du processus N_u jusqu'au temps t (plus précisément, il s'agit de la tribu engendrée par les variables aléatoires $\{N_u : 0 \leq u \leq t\}$); l'ensemble $\{\mathcal{F}_N(t) : t \in I\}$ constitue une famille croissante de tribus, c'est-à-dire une filtration ; $\mathcal{F}_N(t)$ représente aussi l'information fournie par l'observation de :

$$(\tau_1 < \tau_2 < ... < \tau_p \leq t) \cap \{\tau_{p+1} > t\}$$

ou par l'observation de :

$$\left(T_1,T_2,...,T_p : \sum_{1\leq i\leq p} T_i \leq t\right) \cap \left(T_{p+1} > t - \sum_{1\leq i\leq p} T_i\right).$$

Si on écrit $\sigma(X)$ pour la tribu engendrée par une variable aléatoire X, on a alors de façon un peu plus précise :

$$\mathcal{F}_N(t) = \sigma\left(\tau_1,\tau_2,...,\tau_p,\mathbf{1}_{\{\tau_{p+1}>t\}}\right) = \sigma\left(T_1,T_2,...,T_p,\mathbf{1}_{\left\{T_{p+1}>t-\sum_{1\leq i\leq p} T_i\right\}}\right).$$

On notera aussi N_{t_-} le processus de limite à gauche, c'est-à-dire :

$$N_{t_-} = \lim_{u\uparrow t} N_u.$$

De même, on notera $\mathcal{F}_N(t_-)$ l'histoire qui lui est associée : on pourrait aussi écrire $\mathcal{F}_N([0,t))$ mais ce serait plus lourd. L'*intensité stochastique* $h_N(t)$ du processus N_t est la "probabilité instantanée" de saut, qui ne peut être qu'unitaire, conditionnellement à son histoire ; plus précisément :

$$h_N(t) = \lim_{\delta\downarrow 0} \frac{prob\left[N_{t+\delta} - N_t = 1 \mid \mathcal{F}_N(t_-)\right]}{\delta}$$

et le *compensateur* $H_N(t)$ du processus N_t est l'intégrale de son intensité stochastique :

$$H_N(t) = \int_{[0,t]} h_N(u)\, du.$$

La décomposition de Doob-Meyer consiste à décomposer le processus N_t en la somme

$$N_t = H_N(t) + M_t$$

où le processus M_t, défini par $M_t = N_t - H_N(t)$, est une martingale par rapport à $\mathcal{F}_N(t)$, c'est-à-dire $E[M_t \mid \mathcal{F}_N(s)] = M_s, \forall s \leq t$. Cette décomposition est l'analogue, en temps continu, de la décomposition d'un processus en temps discret en une somme de son processus d'espérance conditionnelle par rapport au passé et d'un processus d'innovation. La martingale M_t joue le rôle de processus d'innovation, tandis que le compensateur $H_N(t)$ est la partie prévisible de N_t et est un processus non-décroissant tel que $H_N(0) = 0$. Il faut cependant noter que M_t n'est pas l'innovation du processus mais bien la somme des innovations jusqu'au temps t ; en termes différentiels, on a en effet :

$$dN_t = h_N(t)\, dt + dM_t$$

où dM_t est le processus d'innovation, qui est donc une différentielle de martingale.

Un exemple simple de processus en temps discret aidera à mieux comprendre le sens de cette dernière équation. Le processus auto-régressif d'ordre 1 :

$$y_t = \alpha y_{t-1} + \varepsilon_t, \qquad t = 1, 2, ...$$

peut en effet s'écrire, de façon équivalente :

$$\triangle y_t = (\alpha - 1)\, y_{t-1} + \varepsilon_t \quad , \quad t = 1, 2, ...$$

On reconnaîtra en $\triangle y_t$ l'analogue de dN_t , $(\alpha - 1)\, y_{t-1}$ est l'espérance conditionnelle au passé de $\triangle y_t$ et est donc une fonction du passé du processus comme $h_N(t)\, dt$, et ε_t est une innovation comme dM_t. En sommant cette dernière équation, on obtient :

$$y_t - y_0 = (\alpha - 1) \sum_{0 \leq i \leq t-1} y_i + \sum_{1 \leq i \leq t} \varepsilon_i$$

ou encore :

$$y_t = \left[\alpha y_0 + (\alpha - 1) \sum_{1 \leq i \leq t-1} y_i \right] + \sum_{1 \leq i \leq t} \varepsilon_i.$$

On observe que le premier terme de droite est une fonction du passé mais n'est pas l'espérance conditionnelle au passé du terme de gauche, comme $H_N(t)$, et que le

second terme de droite est une somme d'innovations comme $M_{t.}$

Il y a cependant une spécificité importante de l'aspect dénombrement. Une des caractéristiques des processus ponctuels est en effet que les modélisations du compensateur et de l'innovation sont liées. Il y a en effet bijection entre l'espace des compensateurs et l'espace des processus ponctuels et le processus d'innovation est dérivé de l'un comme de l'autre. Plus spécifiquement, on a la situation suivante. Conditionnellement à l'histoire $\mathcal{F}_N(t_-)$, $h_N(t)\,dt$ est de variance nulle et dMt est d'espérance nulle et de variance égale à $h_N(t)\,dt$; en d'autres termes, contionnellement à l'histoire $\mathcal{F}_N(t_-)$, dN_t a même espérance et variance, chacune égale à $h_N(t)\,dt$: on dira que ces processus sont "localement poissonniens" (où "localement" s'entend au sens du temps). Cette précision mathématique renforce l'idée que la modélisation d'un processus ponctuel se fonde naturellement sur la modélisation de son compensateur ou de son intensité stochastique.

Remarque. Nous avons passé sous silence les conditions qui assurent l'existence de la limite définissant l'intensité stochastique et son intégrabilité, qui permet de définir le compensateur. Au point de vue mathématique, il est plus naturel de commencer par définir le compensateur au moyen de la décomposition de Doob-Meyer et d'en dériver les conditions de dérivabilité qui permettent de définir l'intensité stochastique. Cette décomposition est l'outil adéquat pour la définition du concept de non causalité instantanée dans les processus ponctuels (voir Florens et Fougère, 1991).

c) Quelques exemples. Dans le cas d'un processus de survie, le processus de dénombrement qui lui est associé est un processus élémentaire univarié et avec un seul saut. Son intensité stochastique s'écrit alors :

$$h_N(t) = \lambda_T(t)\ \left(1 - N_{t_-}\right) = \lambda_T(t)\,\mathbb{1}_{]0,T]}(t)$$

où $\lambda_T(t)$ est la fonction de hasard de la distribution de la seule durée T caractérisant le processus, et où le second terme vaut 1 avant le saut et vaut 0 après ; ce terme indique donc si l'individu est encore "à risque" au temps t pour la seule transition possible. En effet, dans ce cas, il y une seule transition sur toute trajectoire complète du processus, et nous avons clairement :

$$\begin{aligned} \{N_{t+\delta} - N_t = 1\} &= \{t < T \leq t + \delta\}, \\ \{N_{t_-} = 0\} &= \{T > t\}, \\ et\ \{N_t = 1\} &= \{T \leq t\}. \end{aligned}$$

De même, dans le cas d'un processus de dénombrement général mais univarié, l'intensité stochastique s'écrit :

$$(*) \qquad h_N(t) = \sum_{p \geq 1} \lambda_{T,p}(t - \tau_{p-1} | T_1, \cdots, T_{p-1})\,\mathbb{1}_{]\tau_{p-1}, \tau_p]}(t)$$

En particulier, au point aléatoire $t = \tau_p$, on obtient

$$h_N(\tau_p) = \lambda_{T,p}(T_p|T_1, \cdots, T_{p-1})$$

et $\lambda_{T,p}$ est la fonction de hasard de la distribution de la durée s'écoulant entre le $(p-1)$-ième et le $p-$ième saut.

L'exemple des *processus poissonniens* peut aussi être éclairant pour montrer comment le compensateur, ou l'intensité stochastique, permet de caractériser la loi d'un processus ponctuel. Rappelons qu'en général, un processus poissonnien est un processus à accroissements indépendants (i.e. $N_s - N_r \perp\!\!\!\perp N_u - N_t$ pour tout $r < s < t < u$) tel que ses accroissements $N_s - N_r$ sont distribués selon une loi de Poisson de paramètre $\mu(]r, s])$ où μ est une mesure quelconque sur $\mathbb{R}_+$. Un processus poissonnien est *homogène* lorsque la mesure μ est proportionnelle à la mesure de Lebesgue sur $\mathbb{R}_+$: $\mu(]r, s]) = (s - r)\,\lambda$ où λ est une constante strictement positive ; dans ce cas, on peut vérifier que :

$$H_N(t) = \lambda t \quad \text{et} \quad h_N(t) = \lambda\,.$$

Si la mesure μ est diffuse de densité $m(t)$ par rapport à la mesure de Lebesgue, $h_N(t) = m(t)$ et $H_N(t) = \mu([0, t])$. Ainsi donc, le problème de l'existence d'une intensité stochastique est de même nature que celui de l'existence d'une densité de probabilité.

d) loi de probabilité d'un processus de dénombrement univarié. Considérons maintenant une observation d'un processus univarié de dénombrement N entre 0 et t : $(N_u : 0 \leq u \leq t)$ Comme mentionné précédemment, il est équivalent d'observer :

$$\left(T_1, T_2, ..., T_p : \sum_{1 \leq i \leq p} T_i \leq t\right) \cap \left\{T_{p+1} > t - \sum_{1 \leq i \leq p} T_i\right\},$$

où $p = N_t$. On peut donc décrire la loi de $(N_u : 0 \leq u \leq t)$ au moyen d'une fonction $l(t)$ composée de la densité conjointe des durées, et de la probabilité de l'événement caractérisant l'arrêt de l'observation. Cette représentation est particulièrement importante car elle conduit directement à l'écriture des fonctions de vraisemblance dans les modèles statistiques impliquant des processus ponctuels. Dans la suite, on ne s'intéressera qu'aux lois de processus conditionnelles à X_0, l'état initial.

On obtient alors

$$l(t) = \left\{\prod_{1 \leq i \leq p} f_{T,i}(T_i|T_1 \cdots T_{i-1})\right\} S_{T,p}(t - \tau_p|T_1 \cdots T_p)\, \mathbb{1}_{[\tau_p, \tau_{p+1}[}(t)$$

La classe des modèles envisagés étant très étendue, il sera utile d'en considérer différentes écritures équivalentes. En termes des fonctions de hasard, et en remarquant que la dernière fonction vaut 1 vu que $N_t = p$, on peut écrire

$$l(t) = \left\{ \prod_{1 \leq i \leq N_t} \lambda_{T,i}(T_i | T_1 \cdots T_{i-1}) \ \exp - \int_0^{T_i} \lambda_{T,i}(u | T_1 \cdots T_{i-1}) du \right\}$$

$$\exp - \int_0^{t-\tau_p} \lambda_{T,p+1}(u | T_1 \cdots T_p) du$$

En utilisant la relation $(*)$ aux points $t = \tau_i$ et en utilisant les changements de variables $s = u + \tau_{i-1}$, on obtient successivement :

$$\begin{aligned} l(t) \quad &= \left\{ \prod_{1 \leq i \leq N_t} h_N(\tau_i) \right\} \exp - \left[\sum_{1 \leq i \leq N_t} \int_{\tau_{i-1}}^{\tau_i} \lambda_{T,i}(s - \tau_{i-1} | T_1 \cdots T_{i-1}) ds \right. \\ &\qquad \left. + \int_{\tau_p}^{t} \lambda_{T,p+1}(s - \tau_p | T_1 \cdots T_p) ds \right] \\ &= \left\{ \prod_{1 \leq i \leq N_t} h_N(\tau_i) \right\} exp - \int_0^t h_N(s) ds \\ &= \left\{ \prod_{1 \leq i \leq N_t} h_N(\tau_i) \right\} exp - H_N(t) \end{aligned}$$

ou encore :

$$ln \ l(t) = \int_0^t ln \ h_N(s) dN_s - H_N(t)$$

où l'intégrale stochastique est obtenue en considérant le processus de dénombrement N_t comme une mesure aléatoire $N_t = \sum_{1 \leq i < \infty} \delta_{\{\tau_i\}}$: ses trajectoires sont en effet non-décroissantes, constantes par morceaux et de sauts unitaires.

Ce développement montre, de nouveau, le rôle naturel du compensateur, de l'intensité stochastique et de sa représentation par les fonctions de hasard conditionnelles $\lambda_{T,i+1}$ dans la modélisation d'un processus ponctuel ou du processus de dénombrement qui lui est associé.

Les processus de dénombrement univariés sont utilisés en économétrie, par exemple, pour l'étude des déterminants des durées séparant les naissances successives d'enfants au sein d'un ménage (cf. Newman et Mac Cullough [1984], Heckman et Walker [1990]) ou pour étudier le processus de dépôt des brevets par une firme (voir Pakes [1986]).

2.6 Processus de transition : caractérisation et exemples

Dans le cas de processus de dénombrement multivariés associés à des processus ponctuels plus complexes que le modèle simple de durée, le résultat final est semblable, dans son essence, au cas univarié mais sensiblement plus compliqué au point de vue analytique. La difficulté provient du fait que la description de l'histoire du processus jusqu'à l'instant t, $\mathcal{F}_N(t)$, doit maintenant tenir compte de toutes les durées (ou instants de saut) passées pour toutes les coordonnées du processus multivarié, c'est-à-dire pour toutes les transitions possibles du processus ponctuel.

2.6.1 Intensités stochastiques de transition

Plaçons–nous à nouveau dans le cas simple de données de transition (X_t), $t \in I$, d'espace d'états E, observées continûment sans biais de sélection. Différentes stratégies de modélisation peuvent être suivies et l'une des plus naturelles serait par exemple de décrire la probabilité de la suite (τ_p, X_p) déjà introduite, ou celle des (T_p, X_p) avec $T_p = \tau_p - \tau_{p-1}$.

Cette dernière suite est donc constituée des durées de séjour dans un état et des états atteints à chaque transition. Nous préférons toutefois centrer notre présentation sur la notion d'intensité stochastique de transition, qui vient d'être introduite pour des processus ponctuels généraux.

La famille des intensités stochastiques des transitions (j,k) est la famille :

$$h_{jk}(t) = q_{jk}(t) Y_j(t_-)$$

où $Y_j(t_-) = \mathbb{1}_{\{X(t_-)=j\}}$ indique si l'individu est à risque pour la transition de j vers k au temps t et où $q_{jk}(t)$ représente la probabilité instantanée de passage de l'état j à l'état k entre t et $t+dt$, conditionnellement à toute la trajectoire antérieure au temps t, et au fait d'être dans l'état j juste avant t, et, éventuellement, à des variables explicatives. Plus formellement :

$$q_{jk}(t) = \lim_{\delta \downarrow 0} prob(X_{t+\delta} = k | X_t = j, \mathcal{F}_N(t_-), Z)/\delta$$

où $\mathcal{F}_n(t_-)$ représente toute l'information contenue dans l'observation de l'individu de l'instant 0 à l'instant immédiatement antérieur à t (formellement $\mathcal{F}_N(t_-) = \sigma(X_s : 0 \leq s < t)$) et Z représente un vecteur de variables caractérisant l'individu. On devrait donc écrire cette probabilité instantanée de transition sous forme d'une fonction $q_{jk}((x_s : 0 \leq s < t), Z)$ mais les arguments de cette fonction seront synthétisés par "t" en vue de simplifier l'écriture. Notons enfin que si $Y_j(t_-) = 0$, l'intensité stochastique de la transition (j,k) est évidemment nulle puisque cette transition là est impossible au temps t.

Associons maintenant au processus de transition (X_t) le processus de dénombrement multivarié $(N_{jk}(t)), (j,k) \in C$, tel que, pour tout couple $(j,k) \in C$, $N_{jk}(t)$ s'accroît de 1 chaque fois que l'individu effectue une transition de j vers k. Plus formellement:

$$
\begin{aligned}
N_{jk}(t) &= \sum_{p\geq 1} \mathbb{1}_{(t\geq \tau_p)} \times \mathbb{1}_{\left(X_{\tau_{p-1}}=j\right)} \times \mathbb{1}_{\left(X_{\tau_p}=k\right)} \\
&= \sum_{s\leq t} \mathbb{1}_{(X_{s_}=j, X_s=k)} \cdot
\end{aligned}
$$

Remarquons que nous pouvons aussi écrire :

$$\mathcal{F}_N(t_-) = \sigma(N_{jk}(s) : (j,k) \in C, 0 \leq s < t)$$

A l'aide de la décomposition de Doob-Meyer de ce processus de dénombrement multivarié, on obtient la formule différentielle suivante :

$$dN_{jk}(t) = h_{jk}(t)dt + dM_{jk}(t),$$

où $M_{jk}(t)$ est une martingale centrée et $h_{jk}(t)$ est l'intensité stochastique de transition.

Par un raisonnement analogue à celui développé dans le cas d'un processus de dénombrement univarié, on peut décrire la loi de probabilité de $(X_s)_{s\in I}$, avec $I = [0,t]$, ou de façon équivalente celle de $\{(N_{jk}(s)),\ (j,k) \in C, 0 \leq s \leq t\}$ par la densité conjointe des durées composant la trajectoire $(X_s)_{s\in I}$, des probabilités des transitions composant cette trajectoire et de la probabilité de l'événement caractérisant l'arrêt de l'observation au temps t. De façon analogue, cette densité des observations, arrêtées au temps $t, l(t)$, s'écrira en fonction de la famille des intensités stochastiques comme suit :

$$\ln l(t) = \sum_{(j,k)\in C} \left[\int_0^t \ln h_{jk}(s)\ dN_{jk}(s) - \int_0^t h_{jk}(s)\ ds \right]$$

où la première intégrale est une intégrale stochastique qui peut aussi s'écrire $\sum_{\tau\leq t} \ln h_{jk}(\tau)$ où τ décrit l'ensemble des temps auxquels l'individu a réalisé une transition de j vers k.

Ce dernier point montre que la probabilité d'une trajectoire $(X_t)_{t\in I}$ est entièrement caractérisée par les fonction $h_{jk}(t)$ qui contiennent toutes les propriétés dynamiques du processus de transition. Ces fonctions représentent la généralisation à des modèles de transition généraux des fonctions de hasard des modèles de durée ordinaires. On écrit donc en pratique des fonctions $h_{jk}(t)$ dépendantes de paramètres (éventuellement fonctionnels) et l'on peut par exemple en déduire une écriture très manipulable de la fonction de vraisemblance ou une estimation non paramétrique des paramètres fonctionnels. Dans des modèles avec variables explicatives, celles-ci sont introduites dans la spécification des intensités stochastiques de transition. Nous nous limiterons tout d'abord à des exemples des modèles sans variables explicatives. Pour ces modèles, les intensités de transition sont identiques pour tous les individus.

2.6.2 Modèles markoviens

Dans les processus en temps discret et espace d'états continu, la classe des modèles auto-régressifs est une façon naturelle d'aborder l'aspect dynamique de la modélisation, car elle se base sur la suite des modèles conditionnels à une suite croissante d'histoires et explicite au départ la "longueur de la mémoire" au moyen de l'ordre d'auto-régressivité. Dans les processus en temps continu, la propriété markovienne est équivalente à la propriété d'autorégressivité du temps discret: le processus ne retient que le passé "le plus proche". Pour définir de manière plus précise le caractère Markovien d'un processus ponctuel de transition, on utilise la caractérisation de sa loi par le système projectif de toutes ses distributions fini-dimensionnelles de $\left(X_{t_1}, X_{t_2}, ..., X_{t_p}\right)$. On dira que le processus ponctuel $X = (X_t : t \in I)$ est *Markovien* si pour tout $t_1 < t_2 < t_p$ et tout p fini :

$$X_{t_p} \perp\!\!\!\perp \left(X_{t_1}, X_{t_2}, ..., X_{t_{p-1}}\right) \mid X_{t_{p-1}}$$

ou encore

$$prob\left(X_{t_p} \mid X_{t_1}, X_{t_2}, ..., X_{t_{p-1}}\right) = prob\left(X_{t_p} \mid X_{t_{p-1}}\right)$$

En d'autres termes, la loi de probabilité de X_t à un instant quelconque t_p, conditionnellement à un résumé fini de son histoire $\left(X_{t_1}, X_{t_2}, ..., X_{t_{p-1}}\right)$ ne dépend que de sa réalisation la plus proche dans cette histoire finie $\left(X_{t_{p-1}}\right)$.

La loi d'un processus Markovien est alors caractérisée par une fonction définie sur $\{(s;t) \in I \times I \mid s < t)$ à valeurs dans l'espace des matrices stochastiques d'ordre $K \times K$ dont les éléments sont définis par :

$P(s,t) = [p_{jk}(s,t)]$, $s < t$,

$p_{jk}(s,t) = prob(X_t = k \mid X_s = j)$.

On dénommera $P(s,t)$ *la matrice des probabilités de transition.* La propriété CADLAG est alors équivalente à ce que pour tout (j,k) :

$$\lim_{t \downarrow s} p_{j,k}(s,t) = \mathbb{1}_{\{k=j\}}.$$

La proprieté markovienne permet de simplifier la probabilité instantanée de transition, des processus de dénombrement associés, de la façon suivante :

$$
\begin{aligned}
q_{jk}(t) &= \lim_{\delta \downarrow 0} \frac{1}{\delta} prob(X_{t+\delta} = k | \mathcal{F}_N(t_-), X_t = j, Z) \quad \text{(par définition)} \\
&= \lim_{\delta \downarrow 0} \frac{1}{\delta} prob(X_{t+\delta} = k | X_t = j, Z) \quad \text{(par markovianité)} \\
&= \lim_{\delta \downarrow 0} \frac{1}{\delta} p_{jk}(t, t+\delta) \\
&= \frac{d}{dt} p_{jk}(s,t) \mid_{s=t} , \qquad j \neq k.
\end{aligned}
$$

L'intensité stochastique de transition devient :

$$h_{jk}(t) = q_{jk}(t) Y_j(t_-)$$

et ne dépend donc de l'histoire que par le fait d'être à risque pour la transition de j vers k.

(la condition $k \neq j$ s'interprète comme suit : on exclut la possibilité d'une transition d'un état vers lui-même). En posant :

$$q_{jj}(t) = -\sum_{k \neq j} q_{jk}(t) ,$$

on peut alors définir la *matrice d'intensité des transitions* :

$$Q(t) = [q_{jk}(t)] \ .$$

On dira que le processus markovien est *homogène* (vis-à-vis du temps) si ses matrices de transition ne dépendent pas de l'origine des temps, c'est-à-dire si

$$p_{j,k}(s,t) = p_{j,k}(0, t-s) \ .$$

Dans ce cas, les *matrices de probabilités de transition* s'écriront simplement $P(t) = [p_{j,k}(t)]$. La propriété CADLAG est alors équivalente à ce que pour tout (j,k) :

$$\lim_{t \downarrow 0} p_{j,k}(t) = \mathbf{1}_{\{k=j\}} ,$$

et les intensités de transition ne dépendent plus de t . En effet :

$$\lim_{\delta \downarrow 0} \frac{p_{j,k}(t, t+\delta)}{\delta} = \lim_{\delta \downarrow 0} \frac{p_{j,k}(0, \delta)}{\delta} = \left[\frac{d}{dt} p_{j,k}(t) \right]_{t=0} = q_{jk} \geq 0, \ k \neq j.$$

Dès lors, la matrice d'intensité de transition devient simplement : $Q = [q_{jk}]$. En définissant la fonction exponentielle matricielle comme de coutume :

$$e^{Qt} = \sum_{0 \leq l < \infty} \frac{Q^l t^l}{l!},$$

la relation entre les deux matrices $P(t)$ et Q est donnée par :

$$\begin{array}{lcl} \frac{d}{dt}P(t) & = & QP(t) \\ \text{soit} & & \\ P(t) & = & exp(Qt). \end{array}$$

Dans un processus markovien homogène de transition, on notera que l'intensité stochastique de transition $h_{jk}(t)$ ne dépend de t que par le fait d'être à risque pour la transition de j vers k . Sachant que $X_{t_} = j$, le passé du processus n'apparaît plus car $q_{jk}(t)$ est une constante. Les processus markoviens homogènes représentent donc les modèles de transition sans mémoire. Dans de tels modèles, il est naturel de prendre comme paramètres les éléments $(q_{jk}), (j,k) \in E \times E, j \neq k$. On peut bien entendu restreindre l'espace paramétrique en posant $q_{jk} = q_{jk}(\theta)$ pour un vecteur $\theta \in \Theta$ de paramètres. La fonction de vraisemblance d'un processus markovien homogène est alors :

$$\begin{array}{rcl} \ln l(t) & = & \sum\limits_{\{(j,k)|j\neq k\}} \left[\int_0^t ln(q_{jk}Y_j(s_))dN_{jk}(s) - \int_0^t q_{jk}Y_j(s_)ds\right] \\ & = & \sum\limits_{\{(j,k)|j\neq k\}} [N_{jk}(t)ln\ q_{jk} - q_{jk}T_{jt}] + c^{te} \end{array}$$

où

$$T_{jt} = \int_0^t Y_j(s_)ds = \int_0^t I\!\!1_{\{X_{s_}=j\}}ds$$

est la durée de séjour dans l'état j pendant l'intervalle $[0,t]$.

Les processus ponctuels markoviens homogènes jouissent dès lors d'une grande simplicité quant leur structure aléatoire. En effet, comme tout processus ponctuel, leur trajectoire peut être décrite par la suite $\left\{\left(T_p, X_{\sum_{0\leq i\leq p} T_i}\right) : p \in I\!N\right\}$ des durées de séjour et des états atteints à chaque transition. Pour simplifier l'écriture, on notera X_p au lieu de X_{τ_p} pour désigner l'état atteint lors de la $p-ième$ transition.

La loi de ces trajectoires a les propriétés suivantes :

(i) le processus$\{(T_p, X_p) : p \in I\!N\}$ est markovien, c'est-à-dire :

$$(T_p, X_p) \perp\!\!\!\perp \{(T_m, X_m) : 0 < m < p-1\} \mid T_{p-1}, X_{p-1};$$

(ii) de plus :

$$(T_p, X_p) \perp\!\!\!\perp T_{p-1} \mid X_{p-1} \quad et \quad X_p \perp\!\!\!\perp T_p \mid X_{p-1}\ ;$$

(iii) le processus $(X_p : p \in I\!N)$ est markovien et sa loi est caractérisée par :

$$prob\,(X_p = k \mid X_{p-1} = j) = \phi_{j,k}$$

où $\phi_{j,k}$ est défini par :

$$\begin{aligned} \phi_{j,k} &= \frac{q_{jk}}{\sum_{k' \neq j} q_{jk'}} \quad \text{si } k \neq j, \\ &= 0 \qquad\quad \text{si } k = j. \end{aligned}$$

(iv) Les distributions conditionnelles $(T_p \mid X_{p-1} = j)$ des durées sont exponentielles de paramètre

$$(-q_{jj}) = \sum_{k \neq j} q_{jk} \ .$$

Une propriété très importante des chaînes de Markov homogènes est que, sous des conditions assez générales, les limites des probabilités de transition $p_{j,k}\,(0,t) = p_{j,k}\,(t)$ existent et sont indépendantes de l'état initial j de la chaîne en t = 0. Plus exactement, une chaîne markovienne homogène vérifie :

$$\lim_{t \uparrow \infty} p_{jk}\,(t) = \Pi_k \qquad \text{avec} \qquad \sum_{k=1}^{K} \Pi_k = 1.$$

où la distribution limite $\Pi = (\Pi_1, ..., \Pi_K)$ est invariante ou stationnaire puisque, elle vérifie : $\Pi = \Pi.P\,(t)\,,\ \forall t \in I\!R_+$.

Le calcul du vecteur Π est facilité dès lors que l'on remarque qu'après avoir différencié le système précédent, ce vecteur Π est également l'unique solution du sytème : $\Pi.Q = 0$.

A titre d'illustration, considérons maintenant deux exemples de processus markoviens, le premier homogène dans le temps, le second non-homogène.

Exemple 1. Soit un processus markovien homogène à deux états (emploi, chômage), c'est-à-dire avec K = 2, et pour lequel on note respectivement :

$$\begin{aligned} q_{12}(t) &= q_{12} = \alpha \quad , \\ q_{21}(t) &= q_{21} = \beta \quad , \quad \forall \quad t \in I\!R_+ \ , \end{aligned}$$

les intensités de transition de l'emploi vers le chômage, et du chômage vers l'emploi. Les durées de séjour en emploi (respectivement, en chômage) sont indépendantes et identiquement distribuées selon une loi exponentielle de paramètre α (respectivement, de paramètre β). Si p_0 et $1 - p_0$ sont les probabilités qu'un individu soit dans l'état 1 (emploi) ou 2 (chômage) à l'instant initial $t_0 = 0$, alors il découle

de la relation fondamentale entre les matrices $\mathrm{P}(t)$ et Q que les probabilités que l'individu soit en emploi ou en chômage à l'instant $t\,(t>0)$ s'écrivent comme :

$$\begin{aligned} p_1(t) &= \frac{\beta}{\alpha+\beta} + \left\{ p_0 - \frac{\beta}{\alpha+\beta} \right\} e^{-(\alpha+\beta)t}, \\ p_2(t) &= \frac{\alpha}{\alpha+\beta} + \left\{ (1-p_0) - \frac{\alpha}{\alpha+\beta} \right\} e^{-(\alpha+\beta)t}, \\ \textit{avec} \qquad & p_1(t) + p_2(t) = 1, \ \forall t \geq 0. \end{aligned}$$

Notons maintenant (Π_1, Π_2) la distribution limite (qui est également la distribution stationnaire) de ce processus. Des deux équations qui précèdent, on déduit aisément que :

$$\begin{aligned} \Pi_1 &= \lim_{t\uparrow\infty} p_1(t) = \frac{\beta}{\alpha+\beta} \\ \text{et} & \\ \Pi_2 &= \lim_{t\uparrow\infty} p_2(t) = \frac{\alpha}{\alpha+\beta} \quad . \end{aligned}$$

Des extensions à des modèles à trois états (emploi, chômage et inactivité) ont été considérées, par exemple, par Flinn et Heckman (1982), et Burdett, Kiefer, Mortensen et Neuman (1984). Des modèles markoviens homogènes de mobilité ont été estimés par Tuma et Robins (1980), Flinn et Heckman (1983), Mortensen et Neuman (1984) et Cazals (1992).

Exemple 2. Modèles Markoviens non-homogènes :

Considérons maintenant l'exemple d'une chaîne de Markov non-homogène à 2 états (emploi, chômage). La non-homogénéité implique que les intensités de transition $q_{12}(t)$ et $q_{21}(t)$ dépendent effectivement du temps (t). En ce cas, la distribution des probabilités d'occupation des états 1 et 2 à l'instant t s'écrivent maintenant :

$$\begin{aligned} p_1(t) = & \exp\left\{ -\textstyle\int_0^t \left[q_{12}(s) + q_{21}(s) \right] ds \right\} \\ & \times \left\{ p_0 + \textstyle\int_0^t q_{21}(s) \exp\left\{ \int_0^s \left(q_{12}(u) + q_{21}(u) \right) du \right\} ds \right\} \\ p_2(t) = & \, 1 - p_1(t) \end{aligned}$$

(voir Chesher et Lancaster [1983]).

Les modèles markoviens non-homogènes sont souvent utilisés pour des processus de mobilité principalement influencés par l'âge atteint au moment d'effectuer la transition. Par exemple, considérons un processus $\{X_t\}_{t\geq 0}$ dont l'espace d'états

est E = {1,2,3} (1 pour emploi, 2 pour chômage, 3 pour inactivité} et pour lequel l'échelle temporelle est fournie par l'âge de l'individu à la date t, soit A_t. Pour la commodité $t = 0$ au début du processus et A_0 est donc l'âge de l'individu au début du processus.

Supposons maintenant que les intensités de transition, à l'instant t dépendent seulement de l'âge atteint à cet instant, c'est à dire $A_t = t + A_0$:

$$q_{jk}(t, A_0, \mathcal{F}_N(t_-)) = q_{jk}(A_t) = q_{jk}(t + A_0)$$

A titre d'exemple, supposons une spécification log-linéaire :

$$q_{jk}(A_t) = exp\{\alpha_{jk} + \beta_{jk}(t + A_0)\}$$

On obtient ainsi un modèle markovien non homogène (vu que l'intensité de transition dépend de t) avec A_0 comme variable explicative.

Considérons maintenant une trajectoire commençant en X_0 et observée durant un intervalle $[0, t]$: $(X_s \ : 0 \leq s \leq t)$ ou, de façon équivalente $(X_0, (T_p, X_{\tau_p}) \ : 1 \leq p \leq N_\cdot(t))$ où τ_p sont les instants de transition, dont le nombre est $N_\cdot(t) = \sum_{j,k} N_{jk}(t), X_{\tau_p}$ est l'état atteint à la p-ième transition et T_p est la durée du p-ième épisode. Notons que l'observation de cette trajectoire permet, entre autres, de déterminer les âges à chacune des transitions : $A_p = A_0 + \tau_p$.

La fonction de hasard de T_p s'écrit :

$$\begin{aligned}
&\lambda_{T_p}(u|X_{\tau_{p-1}} = j, \tau_{p-1}, A_0) \\
&= \lim_{\delta \downarrow 0} \frac{1}{\delta} prob\left[u < T_p \leq u + \delta | T_p \geq u, X_{\tau_{p-1}} = j, \tau_{p-1}, A_0\right] \text{ (par définition)} \\
&= \lim_{\delta \downarrow 0} \frac{1}{\delta} \sum_{k \neq j} prob\left[u < T_p \leq u + \delta, X_{\tau_{p-1}+u+\delta} = k | T_p \geq u, X_{\tau_{p-1}} = j, \tau_{p-1}, A_0\right] \\
&= \sum_{k \neq j} \lim_{\delta \downarrow 0} \frac{1}{\delta} \ prob\left[X_{\tau_{p-1}+u+\delta} = k | X_{\tau_{p-1}+u} = j, \tau_{p-1}, A_0\right] \text{ (par markovianité)} \\
&= \sum_{k \neq j} \lim_{\delta \downarrow 0} \frac{1}{\delta} \ p_{jk}(\tau_{p-1} + u, \tau_{p-1} + u + \delta) \\
&= \sum_{k \neq j} q_{jk}(A_0 + \tau_{p-1} + u) \\
&= \sum_{k \neq j} exp\left[\alpha_{jk} + \beta_{jk}(A_0 + \tau_{p-1} + u)\right] \\
&= \sum_{k \neq j} exp\left[\alpha_{jk} + \beta_{jk}(A_{p-1} + u)\right] \quad \text{vu que } A_{p-1} = A_0 + \tau_{p-1}
\end{aligned}$$

De même, la fonction de survie de T_p s'écrit :

$$prob\left[T_p \geq u | X_{\tau_{p-1}} = j, \tau_{p-1}, A_0\right]$$

$$= \exp - \int_0^u \sum_{k \neq j} \exp\left[\alpha_{jk} + \beta_{jk}(A_{p-1} + s)\right] ds$$

$$= \exp\left\{ - \sum_{k \neq j} \frac{exp\left(\alpha_{jk} + \beta_{jk} A_{p-1}\right)}{\beta_{jk}} \left[\exp \beta_{jk} u - 1\right] \right\}$$

dès que $\beta_{jk} \neq 0$. La vraisemblance de la trajectoire $(X_s, 0 \leq s \leq t)$, conditionnellement à X_0, est dans ce cas :

$$\ln l(t) = \sum_{\{(j,k) \in E \times E | j \neq k\}} \left[\int_0^t \ln q_{jk}(s)\, dN_{jk}(s) - \int_0^t q_{jk}(s) Y_j(s_-)\, ds \right] + c^{te}$$

où $c^{te} = \int_0^t \sum_{\substack{j,k \\ j \neq k}} ln Y_j(s_-) dN_{jk}(s)$.

$$\begin{aligned} ln\ l(t) &= \sum_{\substack{j,k \\ j \neq k}} \left\{ \int_0^t \left[\alpha_{jk} + \beta_{jk}(s + A_0)\right] dN_{jk}(s) \right. \\ &\qquad \left. - \int_0^t exp\left[\alpha_{jk} + \beta_{jk}(s + A_0)\right] Y_j(s_-) ds \right\} + c^{te} \\ &= \sum_{\substack{j,k \\ j \neq k}} \left[\alpha_{jk} + \beta_{jk} A_0\right] N_{jk}(t) + \sum_{\substack{j,k \\ j \neq k}} \beta_{jk} W_{jk}(t) \\ &\quad - \sum_{\substack{j,k \\ j \neq k}} \frac{e^{\alpha_{jk}}}{\beta_{jk}} \Big\{ \sum_{1 \leq p \leq N_\cdot(t)} \mathbb{1}_{\{X_{\tau_{p-1}} = j\}} \left(e^{\beta_{jk} A_p} - e^{\beta_{jk} A_{p-1}}\right) \\ &\quad + \mathbb{1}_{\{X_{\tau_{N_\cdot(t)}} = j\}} \left(e^{\beta_{jk}(A_0 + t)} - e^{\beta_{jk} A_{N_\cdot(t)}}\right) \Big\} + c^{te} \end{aligned}$$

où

$$W_{jk}(t) = \sum_{1 \leq p \leq N_\cdot(t)} \tau_p \mathbb{1}_{\{X_{\tau_{p-1}} = j, X_{\tau_p} = k\}}$$

est la somme des dates où il y a eu transition de j vers k.

On trouvera, par exemple, dans l'article de Ridder (1986) l'estimation d'un modèle markovien non-homogène de transition entre emploi et chômage.

Les modèles markoviens non-homogènes peuvent être traités non-paramétriquement, et on estime alors les fonctions $q_{jk}(t)$ ou certaines de leurs transformations

(par exemple, l'intensité cumulée $\int_o^t q_{jk}(u)du$). Les intensités cumulées peuvent être estimées à partir des observations grâce à l'estimateur de Nelson-Aalen (voir Fougère et Tahar [1987] pour une application de cet estimateur à des données longitudinales de participation féminine au marché du travail), alors que les intensités nécessitent une technique de lissage (estimations par noyau ou fonctions spline, par exemple).

2.6.3 Modèles semi-markoviens

Les processus ponctuels semi-markoviens peuvent être considérés comme les extensions directes des processus markoviens : leurs propriétés sont celles des processus markoviens, à l'exception des distributions conditionnelles des durées de séjour qui, dans le cas semi-markovien, ne sont plus exponentielles mais quelconques. Plus précisément, si $(X_p)_{p\in\mathbb{N}}$ et $(T_p)_{p\in\mathbb{N}}$ sont respectivement les suites des états visités et des durées de séjour pour un processus ponctuel en temps continu $\{X_t\}_{t\geq 0}$ à espace d'états discret $E = \{1, ...K\}$, on dira que $\{X_t\}_{t\geq 0}$ est un processus *semi-markovien* si

$$(X_p, T_p) \perp\!\!\!\perp (X_0^{p-1}, T_0^{p-1}) | X_{p-1}$$

c'est à dire

$$prob\left[X_p{=}k, T_p \geq t | (X_0^{p-1}, T_0^{p-1})\right] = prob\left[X_p{=}k, T_p \geq t | X_{p-1}\right], \quad \forall k{\in}E, \forall p{\in}\mathbb{N}^*, \forall t{\in}\mathbb{R}_+$$

Remarquons que tout processus markovien est semi-markovien mais que l'inverse est faux. La différence essentielle est la suivante. Un processus semi-markovien est markovien *entre* les périodes, c'est à dire la suite $(X_p, T_p)_{p\in\mathbb{N}}$ est markovienne, mais n'est pas markovien *au sein* de la période dès que une des distributions de durée n'est pas exponentielle. Remarquons aussi que dans un processus semi-markovien, la suite $(X_p)_{p\in\mathbb{N}}$ est markovienne.

La loi d'un processus semi-markovien est par conséquent caractérisée par l'ensemble des probabilités :

$$\mathcal{P}_{jk}^p(t) = prob\left[X_p = k, T_p \geq t | X_{p-1} = j\right].$$

On les appellera des *"sous-distributions"* car $\mathcal{P}_{jk}^p(0){=}prob\left[X_p{=}k | X_{p-1}{=}j\right] \leq 1$.

Un processus semi-markovien est homogène dans le temps si les sous-distributions sont les mêmes pour tous les épisodes p :

$$\mathcal{P}_{jk}^p(t) = \mathcal{P}_{jk}(t)$$

en d'autres termes, dans le cas homogène, les probabilités de transition ci-dessus ne dépendent pas du nombre de mobilités préalablement vécues par l'individu. La

fonction $\mathcal{P}$ est appelée le noyau du processus semi-markovien $\{X_t\}_{t\geq 0}$. La suite $(X_p)_{p\in I\!N}$ est une chaîne de Markov ayant pour matrice de transition :

$$r_{j,k} = \mathcal{P}_{jk}(0),\ (j,k) \in E \times E,\ k \neq j$$

et les durées de séjour dans les états visités $T_1, T_2, ...$ sont conditionnellement indépendantes, étant donnée la suite X :

$$\underset{p}{\perp\!\!\!\perp} T_p \mid X$$

On remarquera que si le noyau $\mathcal{P}$ est de la forme

$$\mathcal{P}_{jk}(t) \quad = \quad r_{jk} e^{-\lambda_j\, t},\ (j,k) \in E \times E,$$

où $\lambda_j > 0,\ r_{jj} = 0\ et\ \sum_{k\in E} r_{jk} = 1, \forall j \in E,$
alors $\{X_t\}_{t\geq 0}$ est un processus ponctuel markovien homogène dans le temps, ayant pour matrice d'intensités de transition :

$$Q = [q_{jk}]$$

où $q_{jk} = \lambda_j\, r_{jk},\ \ j \neq k$, et $q_{jj} = -\sum_{k\neq j} q_{jk},\ \forall j \in E.$

On peut remarquer que si λ_j est remplacé par λ_{jk} on obtient un processus qui n'est plus markovien.

Dans un processus semi-markovien homogène, les distributions de survie conditionnelles à l'état de sortie, s'écrivent alors :

$$\begin{aligned} S_{jk}(t) \quad &= \quad prob\,[T_p \geq t | X_{p-1} = j, X_p = k] \\ &= \quad \frac{\mathcal{P}_{jk}(t)}{\mathcal{P}_{jk}(0)} = \frac{\mathcal{P}_{jk}(t)}{r_{jk}} \end{aligned}$$

Les noyaux $\mathcal{P}_{jk}(t)$ peuvent donc être spécifiés par les matrices de probabilité de transition $[r_{jk}]$ et de survie conditionnelle $[S_{jk}(t)]$:

$$\mathcal{P}_{jk}(t) = r_{jk} S_{jk}(t)$$

Et les survies non conditionnelles à l'état de sortie sont :

$$\begin{aligned} S_j(t) \quad &= \quad prob(T_p \geq t | X_{p-1} = j) \\ &= \quad \sum_{k\neq j} \mathcal{P}_{jk}(t) \end{aligned}$$

Les processus semi-markoviens homogènes jouissent des deux premières propriétés fondamentales du cas markovien mais les distributions des durées conditionnelles à l'état initial de l'épisode $(T_p | X_{p-1} = j)$ ne sont généralement pas exponentielles, bien que dans le cas homogène elles soient les mêmes pour tous les épisodes p; de plus, la distribution des durées $(T_p | X_{p-1}, X_p)$ dépend de X_p.

Lorsque le noyau $\mathcal{P}_{jk}(t)$ est différentiable, on peut définir :

$$\tilde{P}_{jk}(t) = -\frac{d}{dt}\mathcal{P}_{jk}(t)$$

et écrire la vraisemblance de $(X_s \ : 0 \leq s \leq t)$ sous la forme (conditionnellement à X_0, en supposant que $N_{\cdot}(t) \geq 1$)

$$l(t) = \left\{ \prod_{1 \leq p \leq N_{\cdot}(t)} \tilde{P}_{X_{p-1},X_p}(T_p) \right\} S_{N_{\cdot}(t)} \left(t - \sum_{1 \leq p \leq N_{\cdot}(t)} T_p \right)$$

Une façon de modéliser un processus semi-markovien homogène est de s'inspirer des modèles à risques concurrents. Associons au processus $(X_t)_{t \geq 0}$, un processus latent (repéré par une étoile) de survie conditionnelle :

$$S^*_{jk}(t) = prob\left[T^*_p \geq t | X^*_{p-1} = j, X^*_p = k\right]$$

Le modèle à risques concurrents construit les probabilités de transitions de la façon suivante :

$$r^*_{jk} = \int \prod_{\substack{l \neq k \\ l \neq j}} S^*_{jl}(u) f^*_{jk}(u) du$$

où $f^*_{jk}(t)$ est la densité correspondant à $S^*_{jk}(t)$. On peut alors construire le noyau sous la forme :

$$\mathcal{P}_{jk}(t) = r^*_{jk} S^*_{jk}(t)$$

La vraisemblance peut alors s'écrire :

$$l(t) = \prod_{1 \leq p \leq N_{\cdot}(t)} f^*_{X_{p-1},X_p}(T_p)$$

$$\prod_{\substack{k \neq X_{p-1} \\ k \neq X_p}} S^*_{X_{p-1},k}(T_p)$$

$$\prod_{k \neq X_{N_{\cdot}(t)}} S^*_{X_{N_{\cdot}(t)},k} \left(t - \sum_{1 \leq p \leq N_{\cdot}(t)} T_p \right)$$

Les modèles à intensités multiplicatives peuvent apparaître comme une généralisation des modèles markoviens homogènes. Jusqu'à présent, nous avons présenté des modèles économétriques sans nous préoccuper d'inférence statistique. La motivation des modèles à intensité multiplicative est essentiellement statistique dans la mesure où le rôle des paramètres inconnus de l'économètre devient essentiel pour la structure du modèle.

Réécrivons d'un point de vue statistique l'intensité stochastique d'un modèle markovien :

$$h_{jk}(t|\theta) = q_{jk}(t,\theta)Y_j(t_-).$$

L'élément essentiel de cette factorisation est que le second terme de droite ne dépend que des observations. La généralisation des modèles à intensité stochastique est obtenue en structurant l'intensité stochastique de la façon suivante :

$$h_{jk}(t|\theta) = \mu_{jk}(t|\theta)A_{jk}(t_-)\ Y_j(t_-)$$

où $A_{jk}(t_-)$ est un processus ne dépendant que du passé de X_t et $\mu_{jk}(t|\theta)$ est une fonction non aléatoire de t (et de θ).

En particulier, le processus est markovien dès que $A_{jk}(t_-) \equiv 1$ et il sera markovien homogène s'il est markovien et $\mu_{jk}(t|\theta) = \mu_{jk}(\theta)$.

Plus généralement, l'économètre pourra spécifier librement les fonctions $A_{jk}(t_-)$ en choisissant le mode de dépendance au passé qu'il souhaite. Par exemple, on peut choisir $A_{jk}(t_-)$ comme une fonction du temps total passé dans l'état j ou du nombre de transitions entre j et k antérieures à t. A partir de la formule générale de la loi d'un processus ponctuel, on dispose d'une écriture de la fonction de vraisemblance permettant une estimation relativement simple des paramètres du modèle. En revanche, une spécification des intensités stochastiques de transition hors des classes de modèles standards peut conduire à une dynamique complexe du processus, difficile à maîtriser conceptuellement. Une bonne stratégie est alors de recourir à une simulation des trajectoires. Un moyen de mesurer la pertinence du modèle est de calculer les probabilités d'occupation des états à chaque instant et de les comparer avec les fréquences empiriques d'occupation déduites de l'échantillon.

Le traitement des variables explicatives doit maintenant être précisé. Nous avons noté Z_i ces variables que nous supposons ici complètement observables. Il faut souligner que parmi les composantes de Z_i, figurent des variables fixes dans le temps (sexe, formation initiale,...) et des processus dépendant du temps (nombre d'enfants, situation matrimoniale, salaire, indemnités de chômage...).

La formalisation la plus courante de la dépendance de la probabilité instantanée de transition envers les caractéristiques individuelles est le modèle de Cox. On suppose dans ce modèle que :

$$h_{jk}(t|Z,\theta) = \mu_{jk}(t|\theta_1)\psi_{jk}(Z|\theta_2)A_{jk}(t_-,Z)Y_j(t_-)$$

où $\theta = (\theta_1, \theta_2)$.

Lorsque les caractéristiques individuelles Z dépendent du temps, une hypothèse de non causalité (de X vis à vis de Z) impliquera que Z n'intervient que par ses valeurs passées ou présentes et cette dépendance peut être spécifiée de multiples façons : par exemple, $\psi_{jk}(Z|\theta_2)$ ou $A_{jk}(t_-,Z)$ peuvent dépendre de la moyenne temporelle de Z_t pour des instants antérieurs à t, ou peuvent dépendre de$Z_t, Z_{t-r}, Z_{t-2r}, \cdots$ en discrétisant le passé en périodes de longueur r.

Cette modélisation est particulièrement utile lorsque θ_1 est un paramètre fonctionnel tandis que θ_2 est un paramètre de dimension finie. Souvent dans ce cas θ_1 est un paramètre de nuisance, alors que θ_2 est un paramètre d'intérêt. Ce modèle est connu comme le modèle de Cox.

Références bibliographiques de la Section 2 :

Il existe de nombreux ouvrages décrivant les propriétés probabilistes et les modélisations stochastiques des processus de transition. Nous citerons, par exemple, Cox et Lewis (1966), Daley et Vere Jones (1988), Jacobsen (1982), Karr (1986). A côté de ces traités fondamentaux, une présentation tournée vers la démographie est donnée par Courgeau et Lelièvre (1989) et un exposé tourné vers l'économétrie est dû à Lancaster (1990). On trouvera une synthèse plus formalisée que le présent article dans Florens et Fougère (1992). Un très grand nombre d'articles fondamentaux et appliqués ont été publiés dans ce domaine. On trouvera dans Fougère et Kamionka (1992d), outre une analyse détaillée des processus makoviens, une bibliograhie d'articles d'économétie des modèles de transition.

3 Modélisation structurelle des mobilités

De manière tout à fait habituelle en économétrie, deux approches peuvent être adoptées face à des données individuelles de transition. On peut adopter une démarche essentiellement statistique consistant à rechercher le meilleur modèle explicatif des observations et en particulier retraçant au mieux les dynamiques individuelles. Cela conduit le plus souvent à choisir une modélisation statistique non paramétrique ou semi paramétrique, dont le danger est de conduire à estimer des vecteurs ou des fonctions difficilement interprétables en regard de la théorie économique. A côté de cette approche en forme réduite existent des approches plus structurelles dans lesquelles sont définies des règles de comportement individuel et des processus d'arrivée des perturbations. Les réactions individuelles aux perturbations déterminent alors les transitions entre états. Le processus de transition est modélisé en fonction des paramètres des éléments des choix individuels (utilités) et des paramètres de processus d'arrivée des perturbations. Ces hypothèses de comportement peuvent être vues comme déterminant des contraintes sur les modèles réduits déterminés sur une base purement statistique. L'une des difficultés de cette démarche réside dans la complexité des calculs dans le modèle économique (en général, un modèle de choix discret dynamique), difficulté contournée grâce à des hypothèses simplificatrices aboutissant à un appauvrissement de la dynamique. C'est ainsi que de nombreux modèles structurels conduisent à des modèles markoviens homogènes dont on a déjà souligné le caractère rudimentaire des interactions intertemporelles qu'ils déterminent.

Le principe général des modèles structurels de transition est de construire un modèle de choix discret dynamique. Un tel modèle se caractérise par un ensemble de choix E = {1,...,K} possibles pour un individu. Si un individu se trouve dans

l'état j en t, il reçoit un flux d'utilité instantané $u_j(t)$. Son utilité pendant une certaine période sera l'intégrale dans le temps de ce flux d'utilité. Ces flux d'utilité sont des processus stochastiques dépendant de l'histoire de l'individu et de composantes aléatoires. En chaque instant, l'individu choisit l'état dans lequel il se trouve de manière à maximiser l'espérance conditionnelle au passé de l'utilité totale future. Cette procédure de choix conduit donc l'individu à décrire une trajectoire en temps continu entre les états de E, et cette trajectoire est le résultat observable par l'économètre des mécanismes de décision individuels. Techniquement, la procédure d'optimisation est construite à partir du principe de Bellman consistant à attribuer à chaque décision une valeur en t, et à utiliser un principe de récurrence "arrière" : la valeur en t d'une décision dépendra de la valeur immédiatement après t des décisions possibles.

La construction des fonctions d'utilité sépare en général l'impact des décisions antérieures de l'individu des phénomènes résultant de choix lui étant extérieurs (arrivée des offres d'emploi, variation sur le salaire...).

A titre d'exemple, on rappellera ici un modèle de comportement engendrant des transitions sur le marché du travail. Il s'agit du modèle publié dans Flinn et Heckman (1982) (voir Lancaster (1990), chapitre 6). L'espace E = {e,c,i} des états est de cardinalité 3 : emploi, chômage et inactivité, et les trois processus d'utilité sont les suivants :

$$\begin{aligned} u_e(t) &= w, \\ u_c(t) &= -c, \quad c > 0, \\ u_i(t) &= n, \end{aligned}$$

où w est le taux instantané de salaire qui est tiré dans une loi F, c est le coût fixé instantané d'une recherche d'emploi en situation de chômage, et n est la valeur instantanée du loisir (ou d'un travail domestique) en situation d'inactivité, elle-même tirée dans une loi G. Les transitions sont rendues possibles par le fait que, dans chaque état, l'individu reçoit des opportunités de transition vers d'autres états en des instants générés par des processus de Poisson homogènes. Par exemple, lorsqu'il est employé au taux de salaire instantané w , un travailleur peut recevoir une nouvelle offre de salaire engendrée par un processus de Poisson de taux λ_e, ou une opportunité de transition vers l'inactivité à un taux poissonnien λ. On supposera, pour simplifier, qu'un travailleur en emploi ne peut retourner en chômage et que le taux d'arrivée des opportunités de transition vers l'inactivité est encore égal à λ lorsque l'individu est en chômage. Le taux poissonnien d'arrivée des offres de salaire est égal à λ_i lorsqu'il est inactif et à λ_c lorsqu'il est au chômage (on suppose généralement que $\lambda_c > \lambda_i$).

A titre d'exemple, nous donnons l'équation de Bellman relative à la valeur de l'état d'emploi pendant une durée infinitésimale h, en supposant le taux d'actualisa-

tion instantané égal à ρ :

$$\begin{aligned} V_e(w) &= \frac{1}{1+\rho h}[wh + (1-\lambda_e-\lambda)\,hV_e(w) \\ &+ \lambda_e\, h\, E_W \max(V_e(W), V_e(w)) \\ &+ \lambda\, h\, E_N \max(V_e(w), V_i(N))] \qquad +o(h). \end{aligned}$$

Le premier terme du crochet représente l'accroissement d'utilité pendant la période h, et les trois suivants la valeur espérée du choix futur qui peut être $V_e(w)$ si rien ne change (avec la probabilité $(1-\lambda_e-\lambda)\,h), \max(V_e(W), V_e(w))$ si un nouveau salaire W est proposé (avec la probabilité approchée $\lambda_e h$) ou $\max(V_e(w), V_i(N))$ si une nouvelle utilité du loisir est engendrée (avec une probabilité approchée λh).

En passant à la limite quand $h \to 0$ et après quelques manipulations probabilistes, on obtient une équation du type :

$$\begin{aligned} V_e(w) = \frac{w}{\rho} &+ \frac{\lambda_e}{\rho}\int_w^\infty (V_e(W) - V_e(w))\; F(dW) \\ &+ \frac{\lambda}{\rho}\int_{\xi_{ei}(w)}^\infty (V_i(N) - V_e(w))\; G(dN). \end{aligned}$$

Cette équation signale qu'un travailleur actuellement employé au salaire w est prêt à accepter, en l'absence de coûts de mobilité, toute offre de salaire supérieure à w, et passera en situation d'inactivité dans le cas où celle-ci lui procure une utilité du loisir (ou un revenu non salarial) supérieure au salaire de réservation $\xi_{ei}(w)$.

Des écritures du même type sont obtenues pour les autres états. Un chômeur acceptera un emploi si le salaire proposé est supérieur à un salaire de réservation ξ_{ce},ou ira vers l'inactivité si son utilité instantanée devient supérieure à ξ_{ci}. Un inactif ira vers l'emploi s'il obtient une offre de salaire supérieure à $\xi_{ie}(n)$, ou vers le chômage si son utilité du loisir devient inférieure à $\xi_{ic.}$

On obtient finalement un processus de transition dont la matrice des intensités de transition s'écrit :

		état de destination e	c	i
état inital	e	$\lambda_e \overline{F}(w)$	0	$\lambda\, \overline{G}(\xi_{ei}(w))$
	c	$\lambda_c\, \overline{F}(\xi_{ce})$	$-$	$\lambda\, \overline{G}(\xi_{ci})$
	i	$\lambda_i\, \overline{F}(\xi_{ie}(n))$	$\lambda\, G(\xi_{ic})$	$-$

Ce processus n'est pas markovien car les intensités de transition dépendent des valeurs passées de w et n. Il est toutefois markovien conditionnellement aux processus des valeurs de réservation.

Le seul problème statistique posé par ce modèle est l'inobservabilité de l'utilité du loisir qui doit être traitée comme une composante d'hétérogénéité non observable.

De multiples généralisations de ce type de modèles existent dans la littérature récente. Les contributions les plus intéressantes concernent l'introduction de mécanismes de rencontre entre l'offre et la demande de travail.

Références bibliographiques de la Section 3 :

Le premier modèle structurel de transition sur le marché du travail est sans doute le modèle de recherche d'emploi (job search) initialement développé par Mc Call (1970) et Mortensen (1970). De nombreuses applications de ce modèle ont été publiées et on trouvera dans Bonnal (1992) une synthèse de l'ensemble de ces travaux. Bonnal et Fougère (1992), Moreau et Visser (1992) et Cases et Lollivier (1993b) présentent des estimations sur données individuelles françaises de modèles structurels de job-search. Le modèle de transition exposé dans cette section est celui proposé par Flinn et Heckman (1982). Depuis cette date, la littérature économique sur les modèles structurels de transition a rapidement progressé. Dans cette littérature, on distingue principalement deux grandes familles de modèles : les modèles dynamiques avec utilité aléatoire (Burdett, Kiefer, Mortensen et Neuman (1984), Olsen, Smith et Farkas (1986)), et les modèles de programmation dynamique (Miller (1984), Rust (1987), Wolpin (1987), Eckstein et Wolpin (1989a), et Hotz et Miller (1993)). Cette seconde classe de modèles, qui est de fait la plus générale mais débouche sur des méthodes d'estimation plus difficiles à mettre en oeuvre, est présentée de manière extensive dans les surveys d'Eckstein et Wolpin (1989 b) et de Rust (1993).

4 Modélisation de données en temps discret : les panels des Enquêtes-Emploi

Nous présenterons, à titre d'exemple d'application des modèles de transitions, une synthèse de quelques modèles réalisés à partir d'un panel de l'Enquête Emploi de l'INSEE (1986-1988) *.

Ce panel permet d'observer un processus de transitions sur le marché du travail et les quatre états possibles retenus sont : emploi sur CDI, emploi sur CDD, chômage et inactivité. Pour chaque individu i, le processus X_{it} n'est observé qu'en trois dates : mars 1986, mars 1987 et mars 1988. Le temps retenu dans ces modèles est l'âge des individus. On dispose donc d'un moyen d'examiner les propriétés dynamiques, au cours de la vie des individus, des transitions sur le marché du travail

*Notre synthèse est directement tirée des articles de Fougère et Kamionka (1992 a,b,c), Kamionka et Pastorello (1992) et Kamionka (1994).

aux conditions des années 86 à 88. Divers échantillons ont été utilisés, de taille variant entre 21000 et 33000 individus. L'Enquête Emploi ne présente a priori pas de biais de sélection, mais les procédures de sélection de sous-échantillons (cylindrage, sélection sur la disponibilité des observations des variables explicatives) ont éventuellement un caractère endogène qui sera négligé dans les modèles présentés ici.

Les spécifications retenues sont essentiellement des modèles markoviens homogènes par strate d'échantillon, et le problème statistique ici rencontré est donc celui de l'estimation d'un tel processus à partir de données collectées en temps discret et non en temps continu. Les modèles estimés ainsi que les principaux résultats empiriques obtenus seront brièvement rappelés ici.

1) Le premier type de modèle est markovien homogène par sexe et classe d'âge. En d'autres termes, le modèle est markovien pour chaque sexe et les intensités de transition sont constantes à l'intérieur de chaque classe d'âge. L'estimation de ce modèle permet en particulier d'estimer des probabilités d'occupation des états en fonction de l'âge, des nombres moyens de passage dans chaque état, des durées moyennes de séjour (cf. Fougère et Kamionka (1992 a)).

2) Le second type de modèle introduit plus de variables explicatives. Il s'agit là encore de modèles markoviens homogènes par tranche d'âge, dont les intensités de transition s'écrivent :

$$q_{jk}^{u}\left(Z_i\right) = \exp\left(Z_i \beta_{jk}^{u}\right)$$

où u est la tranche d'âge. Afin de pouvoir traiter le problème de la discrétisation temporelle des observations, les variables explicatives ont été contraintes à être uniquement qualitatives ou discrétisées. Les variables explicatives retenues sont le sexe, le fait d'être marié, être titulaire d'un CAP ou d'un BEP, d'un diplôme supérieur au baccalauréat et la classe de salaire (cf. Kamionka et Pastorello (1992)).

3) Le troisième type de modèle est plus structurel dans la mesure où il se fonde sur une décomposition de l'intensité de transition :

$$q_{jk}^{u}\left(Z_i\right) = \lambda_{j}^{u}\left(Z_i\right)\ r_{k}^{u}\left(Z_i\right)$$

dont les composantes s'interprètent comme un taux d'arrivée des chocs et une probabilité de choix entre diverses modalités. Ces deux fonctions de Z_i sont paramétrées de la façon suivante :

$$\lambda_{j}^{u}\left(Z_i\right) = \exp\left(\beta_{j}^{u} Z_i\right),$$

$$r_{k}^{u}\left(Z_i\right) = \frac{\exp\left(\gamma_{k}^{u} Z_i\right)}{1 + \sum_{l \neq K} \exp\left(\gamma_{l}^{u} Z_i\right)}\ .$$

L'indice u désigne toujours la tranche d'âge et l'état K est l'état de référence (inactivité). L'ensemble des états dans ce modèle se réduit à trois éléments: em-

ploi, chômage et inactivité (cf. Kamionka et Pastorello (1992)).

Nous allons résumer ici les principaux résultats obtenus de l'estimation de modèles de transitions à partir du panel de l'Enquête Emploi de l'INSEE (1986-1988). Les modèles retenus sont définis en temps discret en utilisant les mêmes données : modèle markovien ou mélange de modèles de ce type, comme le modèle dit du "mover-stayer" (individu mobile, individu stable), par exemple (cf. Fougère et Kamionka (1992 b,c)).

1) Un modèle markovien homogène a été estimé sur l'échantillon stratifié selon le sexe et l'âge. Huit classes d'âges ont été constituées : 15-20, 21-25, 26-35, 36-45, 46-55, 56-60, 61-65, 66 ans et plus.

On y observe que l'intensité de transition de l'emploi précaire sur CDD vers l'emploi sur CDI est généralement supérieure ou égale à l'intensité de transition de l'emploi précaire vers le chômage. Ainsi, entre 26 et 35 ans, pour les hommes, trois-quarts des séjours en emploi précaire sont suivis d'un passage en emploi stable, alors que, pour les femmes, cette proportion est plus proche des deux-tiers. L'emploi précaire présente un rôle insertionnel, surtout pour les hommes, puisque la majorité des épisodes d'emploi précaire sont suivis d'un séjour en emploi sur CDI. Cette constatation doit être nuancée puisque des études complémentaires (cf. Fougère et Kamionka (1992 b) et Kamionka (1994)) ont montré que l'emploi précaire peut avoir un double rôle : insertionnel pour une partie des individus qui sont pris dans des trajectoires qualifiées d'"ascendantes", ou précarisant pour ceux qui, confinés dans l'alternance d'épisodes de chômage, d'inactivité et d'emploi précaire, ne peuvent pas passer en emploi stable.

Au sortir du chômage, pour les hommes, l'intensité de transition vers l'emploi précaire sur CDD est plus forte en général que le taux de transition vers l'emploi sur CDI. Pour les femmes, l'intensité de transition du chômage vers l'emploi sur CDI est voisine de celle des hommes alors que celle du chômage vers l'emploi précaire est en général deux fois plus faible. Aussi, au sortir du chômage, les hommes trouvent plus souvent un emploi que les femmes. D'autre part, si c'est à partir du chômage que les individus se retirent de l'activité le plus fréquemment, ces retraits sont plus intenses pour les femmes (à partir de 21 ans).

A partir des estimations qui sont faites des intensités de transition, il est possible d'obtenir un certain nombre d'indicateurs pertinents pour l'analyse des mobilités individuelles : nombres moyens de récurrences, durées moyennes de séjour, probabilités limites d'occupation.

A partir des proportions stationnaires d'individus occupant chaque état, on peut faire quelques remarques relatives au chômage des jeunes. Les différences observées entre les taux de chômage des jeunes (21-25 ans) et des adultes (36-45 ans) peuvent se répartir en une composante de "sélection" mesurée par la différence en-

tre les probabilités stationnaires (estimées) des jeunes et des adultes et une autre dite d'"insertion" mesurée par la différence entre la proportion observée et la probabilité stationnaire (estimée) des jeunes dans l'état de chômage (pour plus de détails, voir Fougère et Kamionka (1992 a)). Pour les hommes, la première est égale à 4,38 % et la seconde à 3,95 % ; pour les femmes, la composante de "sélection" est beaucoup plus importante (9,1 %), alors que la composante d'"insertion" n'est que de 1,3 %.

2) Le second modèle markovien homogène estimé combine une stratification de l'échantillon et une paramétrisation des intensités de transition. Cette approche a permis de tenir compte de plus de variables explicatives. Pour estimer cette spécification en forme réduite, cinq classes d'âge ont été constituées : 19-22, 23-29, 30-40, 41-54, 55-59 ans. L'espérance du taux de salaire horaire a été estimée à partir des données de l'Enquête Emploi de l'INSEE en utilisant un modèle Tobit généralisé.

On constate une forte dépendance du taux de salaire horaire en fonction de l'âge. Les signes des paramètres associés à la variable "âge" sont conformes au type d'évolution que l'on s'attend à trouver, c'est-à-dire un taux de salaire croissant avec l'âge mais à une vitesse décroissante. Les effets du diplôme et de l'âge que l'on trouve ici, sont semblables à ceux mis en évidence par Murphy et Welch (1992) sur les données américaines du CPS (Current Population Survey) pour la période 1963 à 1989. Ces deux auteurs montrent que les taux de salaire moyens s'accroissent à un taux décroissant au cours de la carrière des individus, jusqu'à ce qu'ils se mettent à décroître légèrement.

Les signes des paramètres relatifs aux diplômes confirment l'idée qu'une augmentation du capital humain individuel a un effet positif sur l'espérance du taux de salaire. On peut noter que le gain apporté par l'obtention d'un baccalauréat (ou plus) sur cette espérance est particulièrement important.

L'effet de la situation matrimoniale apparaît différent selon le sexe. En particulier, lorsque les femmes sont mariées, elles ont une espérance de taux de salaire significativement plus faible ; au contraire, lorsqu'ils sont mariés, les hommes ont une espérance de taux de salaire significativement plus élevée.

On peut se demander si l'effet négatif constaté pour les femmes mariées est la conséquence d'une discrimination à leur égard : sachant qu'elles ont une plus forte probabilité d'avoir des enfants, les entreprises vont-elles sanctionner leur plus grande indisponibilité par une moins grande rémunération moyenne ? Le modèle, pour l'instant, ne nous permet pas de répondre à cette question.

3) Le modèle de comportement de Burdett et alii (1984) a été estimé sur les cinq classes d'âge que nous venons de présenter. Ce modèle, très proche de celui exposé dans la Section précédente, décompose chaque intensité de transition en un produit du taux d'arrivée d'un choc par la probabilité que l'individu, qui maximise son espérance d'utilité intertemporelle, choisisse un état donné du marché du travail lorsque ce choc est survenu. Dans le cadre de ce modèle structurel, on peut s'attendre à trouver les résultats suivants :

a) Les individus qui ont une espérance de taux de salaire élevée doivent avoir des intensités de transition vers l'emploi plus élevée et des intensités de sortie de l'emploi plus faibles.

b) Plus l'espérance du taux de salaire sera forte, plus les intensités de transition hors de l'inactivité seront fortes et plus les intensités de transition vers cet état seront faibles.

En considérant les signes des paramètres et les valeurs de la statistique de Student qui leur sont associées, on remarque que tous les paramètres significatifs pour une erreur de première espèce de 10 % ont le signe qui convient. En fait, les paramètres de l'espérance du taux de salaire sont souvent peu significatifs et, lorsqu'ils le sont, c'est souvent pour les intensités de transition au sortir de l'inactivité.

Pour plus de détails concernant les résultats des estimations de ce modèle sur les données de l'Enquête Emploi, on pourra se référer à Kamionka (1992). Kamionka et Pastorello (1992) ont effectué une comparaison des processus de transitions sur les marchés du travail français et italien à partir des estimations du modèle de Burdett et alii (1984). Ce modèle sert également de cadre théorique de référence à la recherche menée par Magnac et Robin (1991) sur les transitions entre emploi salarié et professions indépendantes.

Références bibliographiques de la section 4 :

L'analyse du panel de l'Enquête Emploi de l'INSEE au moyen des processus de transitions a été développée par Fougère et Kamionka (1992 a,b,c) et Kamionka (1994). L'extension de ces travaux aux modèles à variables explicatives sous formes réduite ou structurelle a été réalisée par Kamionka et Pastorello (1992) et une présentation générale des résultats se trouve dans Kamionka (1992). Lollivier (1994), prolongeant l'analyse développée dans Fougère et Kamionka (1992 a), a étudié les transitions individuelles sur le marché français du travail à partir de plusieurs panels des Enquêtes Emploi de l'INSEE. Plus généralement, le traitement des processus de Markov observés en temps discret a fait l'objet de nombreux articles en statistique et en économétrie. Parmi ceux-ci, citons les travaux de Singer et Spilerman (1976), Singer (1981), Kalbfleisch et Lawless (1985), Geweke, Marshall et Zarkin (1986 a et b).

5 Hétérogénéité non observée

Les modèles économétriques de transition considèrent les individus comme indépendants mais différents, et les probabilités engendrant leurs trajectoires personnelles sont conditionnelles à des variables caractéristiques. Nous avons déjà introduit celles-ci dans le cas où elles sont observables et allons examiner dans ce para-

graphe le cas de variables caractéristiques inobservables. On supposera donc qu'à chaque individu est attribuée une variable (ou un vecteur) ς, dont les différentes valeurs sont tirées indépendamment dans une même loi de probabilité Φ (dans le cas continu, de densité $\varphi(\varsigma)$). Cette distribution peut dépendre des caractéristiques observables.

Ces facteurs d'hétérogénéité non observés résument l'incidence sur les trajectoires de variables éventuellement observables mais non disponibles au statisticien, mais aussi celle d'éléments individuels relatifs aux préférences ou aux capacités des individus, qui sont par nature inobservables.

On se limitera au cas où les variables d'hétérogénéité non observables sont fixes dans le temps. Intuitivement l'observation d'un individu pendant une période suffisamment longue permettra d'apprendre quelque chose sur la valeur ς qui lui est propre. C'est ainsi par exemple qu'une compagnie d'assurance automobile discriminera les "bons" et "mauvais" conducteurs en observant la séquence des sinistres dans lesquels ils sont impliqués.

Le principe général de traitement des modèles à variables non observables est de spécifier le modèle conditionnel à celles-ci ainsi que leur distribution et de construire le modèle marginal (ou mélange) dans lequel les grandeurs inobservables n'interviennent plus, une fois faite l'intégration relativement à leur probabilité. On peut alors considérer ce type de construction comme une façon d'obtenir des modélisations plus complexes et donc mieux adaptées à la description des données.

Inversement, les observations ne permettent que l'identification et l'estimation du modèle marginal, et sa décomposition en un modèle à variable inobservable et en une distribution de celle-ci est en général contingente à un ensemble d'hypothèses de spécifications qui limitent la robustesse du modèle (pour une discussion de ce problème, on consultera Heckman [1991]).

Plus formellement, si $(X_t)_{t\in I}$ est la trajectoire d'un individu de caractéristiques observables Z et inobservables ς, on aura :

$$\overline{l}(t \mid Z, \theta, \varphi) = \int l(t \mid Z, \varsigma, \theta)\, g(\varsigma \mid Z, \varphi)\, d\varsigma$$

où l désigne la densité des observations $(X_s : 0 \leq s < t)$ conditionnelle à l'ensemble des caractéristiques et aux paramètres, $\overline{l}$ la densité conditionnelle aux seules caractéristiques observables et aux paramètres, g la densité de ς et φ ses paramètres. Notons que dans la mesure où les ς sont indépendamment distribués, le modèle de densité $\overline{l}$ conserve la propriété d'indépendance entre individus.

L'étude des modèles avec hétérogénéité peut être conduite sous deux points de vue. Dans une analyse de la modélisation, on s'intéressera aux propriétés du modèle marginal en relation avec les propriétés du modèle conditionnel aux ς et

à la distribution de celles-ci. Dans une analyse statistique, on examinera les propriétés d'identification de θ et de φ en l'absence d'observation des ς.

Nous nous concentrerons essentiellement ici sur le premier de ces problèmes. Nous ferons d'abord une remarque très générale mais dont l'utilité n'apparaîtra qu'en considérant des constructions particulières manipulables. Soit $(X_t)_{t\in I}$ un processus de transitions entre états de l'ensemble $E = \{1, ..., K\}$ et $h_{jk}(t)$ l'intensité stochastique de transition entre j et k. Pour préciser notre résultat, nous écrivons $h_{jk}(t, \varsigma, \mathcal{I}_{t_-}, \theta)$, où ς demeure le vecteur d'hétérogénéité inobservable et $\mathcal{I}_{t_-}$ l'ensemble des observations intervenant dans l'intensité de transition, c'est-à-dire les caractéristiques observables et le passé de la trajectoire antérieur à t, $\mathcal{I}_{t_-} = \left((X_s)_{0\leq s<t}, Z\right)$. Les paramètres (éventuellement fonctionnels) sont toujours notés θ. On montre alors que le processus de transition marginal possède des intensités stochastiques de transition vérifiant la relation :

$$\overline{h}_{jk}\left(t, \mathcal{I}_{t_-}, \theta\right) = \int h_{jk}\left(t, \varsigma, \mathcal{I}_{t_-}, \theta\right)\ g\left(\varsigma \mid \mathcal{I}_{t_-}, Z, \varphi, \theta\right) d\varsigma$$

Pour obtenir les intensités de transition marginales, il faut donc calculer en chaque instant t la distribution $g\left(\varsigma \mid \mathcal{I}_{t_-}, Z, \varphi, \theta\right)$ de l'hétérogénéité conditionnellement à l'observation du processus de transition jusqu'en t et conditionnellement aux caractéristiques observables et aux paramètres. Ce calcul s'effectue en général grâce à la règle de Bayes :

$$g\left(\varsigma \mid (X_u)_{0\leq u<t}, Z, \theta, \varphi\right) = \frac{g\left(\varsigma \mid Z, \varphi\right)\ l\left(t \mid Z, \varsigma, \theta\right)}{\int g\left(\varsigma \mid Z, \varphi\right)\ l\left(t \mid Z, \varsigma, \theta\right) d\varsigma}\ .$$

Illustrons ce résultat dans un cas particulier. Supposons le processus de transition markovien homogène[†] à deux états, d'intensités $q_{12}(\varsigma)$ et $q_{21}(\varsigma)$. Ce modèle ne contient pas de caractéristique individuelle observable et l'on supposera que ς ne peut prendre que deux valeurs, 0 et 1, avec les probabilités $\Phi(0)$ et $\Phi(1) = 1 - \Phi(0)$. Les paramètres sont alors les quatre nombres positifs $q_{12}(0), q_{12}(1), q_{21}(0)$ et $q_{21}(1)$, ainsi que $\Phi(0)$.

En d'autres termes, l'échantillon observé est constitué d'individus de deux types déterminant deux processus de Markov différents. Le type de chaque individu est inconnu du statisticien. Rappelons que les intensités de transition conditionnelles à ς s'écrivent ici indépendamment du passé du processus de transition :

$$h_{12}(t, \varsigma) = q_{12}(\varsigma)\ Y_1(t_-) \text{ et } h_{21}(t, \varsigma) = q_{21}(\varsigma)\ Y_2(t_-)\ .$$

Sous ces hypothèses, on calcule la probabilité pour qu'un individu soit du type 0 après observation du processus de transition jusqu'en t_-. En notant T_1 et T_2 les durées totales de séjour dans les états 1 et 2, on trouve que cette probabilité est

[†] "L'homogénéité" du processus markovien renvoie à une stabilité temporelle des intensités et n'est pas contradictoire avec "l'hétérogénéité" individuelle considérée ici.

égale à :

$$\Phi\left(0|X_s : 0{\leq}s{\leq}t\right) = \frac{\Phi(0)q_{12}(0)^{N_{12}(t_-)}q_{21}(0)^{N_{21}(t_-)}\exp\left\{-\left(q_{12}(0)T_1 + q_{21}(0)T_2\right)\right\}}{\sum_{l=0,1}\Phi(l)q_{12}(l)^{N_{12}(t_-)}q_{21}(l)^{N_{21}(t_-)}\exp\left\{-(q_{12}(l)T_1 + q_{21}(l)T_2)\right\}}$$

Les intensités du processus marginal sont égales à :

$$\overline{h}_{jk}(t, X_s \ : 0 \leq s < t) = Y_j(t_-)\sum_{l=0,1} q_{jk}(l)\Phi(l \mid X_s \ : 0 \leq s < t)$$

Il apparaît clairement que le processus marginal a une intensité stochastique dépendante du passé du processus (le nombre de transitions de chaque sorte et les temps de séjour dans chaque état apparaissent) et que le processus marginal n'est donc plus markovien. Ce processus n'appartient pas non plus à la classe des processus à intensité multiplicative. Le mélange par l'hétérogénéité engendre alors une dynamique beaucoup plus complexe que le modèle conditionnel à ς. Cette dynamique est portée par la probabilité "a posteriori" du type d'individu conditionnelle au passé du processus.

Complétons cette approche par quelques mots sur l'analyse statistique de ce mélange. Si les trajectoires sont observées en temps continu, le modèle sera en général identifiable. Le problème est plus complexe si le phénomène n'est observé que le long d'une suite de temps discrets. Un cas particulier intéressant a fait l'objet d'importants développements dans la littérature sour le nom de modèle de "mover-stayer". On suppose que le type 0 d'individu est soumis à un processus markovien homogène comme précédemment, alors que les individus de type 1 sont contraints à rester dans un état (les stayers). Le second processus est donc dégénéré. Dans ce cas, les intensités du processus des "movers" et la probabilité de chaque type sont identifiables, même dans le cas d'observations en temps discret.

Le modèle du "mover-stayer" peut être étendu au cas où l'ensemble des états est de cardinalité supérieure à 2 et où le nombre de types d'individus est aussi supérieur à 2. L'identification du modèle observé en temps discret est alors obtenue en supposant que chaque type d'individu ne peut circuler qu'à l'intérieur d'un sous-ensemble spécifique d'états, éventuellement inclus les uns dans les autres.

Le modèle précédent peut aussi être étendu en considérant des distributions continues d'hétérogénéité. Dans un modèle markovien homogène, on écrira par exemple :

$$q_{jk}\left(\varsigma\right) = \varsigma \ q_{jk}$$

où ς est un nombre aléatoire positif et q_{jk} est une intensité constante. En imposant à ς une distribution gamma de moyenne 1, on vérifie alors que les intensités q_{jk} et la distribution d'hétérogénéité sont identifiables dans le modèle marginal observé continûment.

Références bibliographiques de la Section 5 :

La mise en évidence du phénomène du "mover-stayer" est due à Blumen, Kogan et McCarthy (1955) et a donné lieu à de nombreux développements. Par ailleurs, la prise en compte de l'hétérogénéité dans les modèles de durée a fait l'objet d'une importante littérature (voir, en particulier, l'article d' Elbers et Ridder (1982)). L'approche générale de l'hétérogénéité dans les modèles de transition est due à Aalen (1987). Le modèle de mélange de processus markoviens d'espaces d'états différents a été étudié par Kamionka (1994). Plus généralement, l'analyse des mélanges est un thème important de la statistique mathématique (voir, par exemple, Florens, Mouchart et Rolin (1992)).

6 Sélection endogène de l'échantillon

La constitution d'un échantillon à partir de critères reliés au phénomène étudié est extrêmement courante en microéconométrie, principalement dans le cas où la source de données n'a pas été définie avec des objectifs scientifiques, mais avec des objectifs administratifs : par définition, les fichiers gérés par l'ANPE ou l'UNEDIC (fichiers GIDE ou FNA) ne contiennent que des individus ayant connu une période de chômage. L'économètre estimant un modèle de transition à partir de ces données doit donc s'interroger sur la méthode de traitement permettant de redresser un éventuel biais de sélection ou de traiter ces données avec le maximum d'efficacité.

De manière formellement semblable au problème consistant à analyser convenablement un échantillon donné, peut se poser le problème de la détermination optimale d'un échantillon. Il est bien connu qu'un échantillonnage stratifié, éventuellement sur des critères partiellement endogènes, peut aboutir, après utilisation des procédures statistiques adaptées, à une information meilleure que celle résultant d'un échantillonnage aléatoire. Ceci est particulièrement vrai dans le cas de modalités des observations de très petite probabilité, que l'on a donc intérêt à sur-représenter dans l'échantillon.

Nous n'examinerons pas ici le problème de l'échantillonnage optimal et nous nous limiterons à présenter un exemple simple d'échantillonnage endogène.

Soit $\{1, ..., N\}$ une population d'individus. A tout $i \in \{1, ..., N\}$ est associé un vecteur de variables caractéristiques constantes dans le temps, Z_i , et un processus markovien homogène stationnaire à deux états, caractérisé par ses deux intensités de transition $q_{12}(Z_i, \theta)$ et $q_{21}(Z_i, \theta)$, dépendantes de Z_i et de paramètre θ. Par exemple, les deux états peuvent être pour des individus actifs, l'emploi et le chômage. On désignera par X_{it} la trajectoire de l'individu i entre ces deux états. Le processus a été supposé stationnaire pour simplifier : les individus étudiés sont présents sur le marché du travail depuis un temps suffisamment long pour que les conditions initiales puissent être négligées. Définissons enfin le mécanisme d'échantillonnage : on suppose que, en t = 0, ne sont observés que les individus

se trouvant dans l'état 1. Ces individus seront suivis jusqu'à un temps t donné. Le nombre total N d'individus est connu.

Pour i = 1, ..., N, on notera $\delta_i = 0$ le fait que i n'est pas observé et $\delta_i = 1$ quand i est observé. Les caractéristiques Z_i ne seront pas connues.

Pour écrire la fonction de vraisemblance de l'échantillon, la connaissance de la distribution des caractéristiques est nécessaire et l'on supposera que les Z_i sont indépendants de densité $b(.)$. La probabilité que $\delta_i = 1$, conditionnellement à Z_i, est égale à :

$$\alpha_1(Z_i, \theta) = \frac{q_{21}(Z_i, \theta)}{q_{12}(Z_i, \theta) + q_{21}(Z_i, \theta)}.$$

La densité d'une trajectoire $(X_{is} : 0 \leq s \leq t)$ étant donnés $Z_i, \delta_i = 1$ et θ, vérifie :

$$\begin{aligned} \ln l_i(t \mid Z_i, \delta_i, \theta) &= N_{12}^i(t) \ln q_{12}(Z_i, \theta) - q_{12}(Z_i, \theta) T_1 \\ &\quad + N_{21}^i(t) \ln q_{21}(Z_i, \theta) - q_{21}(Z_i, \theta) T_2 \end{aligned}$$

où $N_{12}^i(t)$ (resp. $N_{21}^i(t)$) est le nombre de transitions observées de i entre 1 et 2 (resp. 2 et 1) dans l'intervalle $[0, t]$, et T_1 et T_2 sont les temps de séjour en 1 et 2. Remarquons que δ_i n'intervient pas dans cette densité conditionnelle.

On calcule alors la probabilité marginale d'occuper l'état 1 :

$$\alpha_1(\theta) = \int \alpha_1(Z_i, \theta)\, b(Z_i)\, dZ_i$$

et la densité conditionnelle des variables explicatives étant donné δ_i :

$$b(Z_i \mid \delta_i = 1, \theta) = \frac{b(Z_i)\, \alpha_1(Z_i, \theta)}{\alpha_1(\theta)}$$

On obtient alors la fonction de vraisemblance :

$$L(\theta) = \prod_{i=1}^{N} \alpha_2(\theta)^{1-\delta_i}\ \alpha_1(\theta)^{\delta_i}\ b(Z_i \mid \delta_i = 1, \theta)^{\delta_i}\ \ l(t \mid Z_i, \delta_i = 1, \theta)^{\delta_i}$$

où $\alpha_2(\theta) = 1 - \alpha_1(\theta)$.

La maximisation de $\ln L(\theta)$ conduira à l'estimateur convergent asymptotiquement normal le plus efficace.

On peut remarquer que dans cet exemple, négliger le mécanisme de sélection et maximiser l'expression :

$$\sum_{i=1}^{N} \delta_i \ln l\left(t \mid Z_i, \delta_i, \theta\right)$$

conduit à un estimateur convergent asymptotiquement normal, mais de précision asymptotique moindre que le maximum de vraisemblance. Dans ce modèle markovien stationnaire, cela revient à ne pas retenir l'information contenue dans les conditions initiales du phénomène. Toutefois, le maximum de vraisemblance nécessite la connaissance de la distribution des variables explicatives de toute la population et son résultat sera sensible à une éventuelle mauvaise spécification de celle-ci.

Dans des processus de transition plus complexes, la non-prise en compte des procédures de sélection peut engendrer des biais d'estimation : ce serait par exemple le cas si l'on n'observe que des individus ayant séjourné plus d'un certain temps dans un état (étude des chômeurs de longue durée) dans un processus non markovien homogène.

Références bibliographiques de la section 6 .

Le biais de sélection endogène dans les modèles de transition sur le marché du travail a été mis en évidence par Chesher et Lancaster (1983) et Ridder (1984). Depuis cette date, son traitement statistique a été particulièrement étudié par Lancaster et Imbens (1991) et Gouriéroux et Monfort (1991). Cases et Lollivier (1993a) et Bonnal, Fougère et Sérandon (1994) ont appliqué ces techniques de redressement de biais au panel de l'Enquête "Suivi des Chômeurs" recueilli par l'INSEE : le premier de ces papiers concerne un seul épisode de chômage avec une composante univariée d'hétérogénéité non observable, le second, plus général, étudie l'impact de passages par des mesures "jeunes" sur le devenir des chômeurs bénéficiaires au sein d'un modèle de transition avec cinq états et un vecteur de composantes d'hétérogénéité non observée.

Résumé

Cet article propose une présentation générale, et assez peu technique, des modèles économétriques actuellement utilisés pour analyser les mobilités individuelles sur le marché du travail. Le point de vue général est de présenter ces modèles dans le cadre des processus ponctuels, c'est-à-dire des processus aléatoires en temps continu avec espace d'états fini. Le formalisme mathématique est maintenu à un minimum compatible avec une compréhension raisonnable des difficultés du sujet, et une attention particulière est donnée à la motivation des principaux concepts de base. La première partie de l'exposé est consacrée aux problèmes relatifs à la modélisation et les problèmes d'inférence sont à peine évoqués. Par la suite, sont rappelées quelques modélisations structurelles conduisant aux spécifications économétriques présentées au début de l'exposé, et nous insistons sur des exemples d'utilisation relatifs au marché français du travail. Enfin, parmi les divers problèmes de spécification et de traitement statistique, nous avons choisi de présenter ceux découlant de la prise en compte de l'hétérogénéité non observée et ceux résultant d'un mécanisme d'échantillonnage endogène.

Bibliographie

Aalen O. (1987). "Mixing Distributions on a Markov Chain", *Scandinavian Journal of Statistics*, vol. 14, pp. 281–289.

Blumen I., Kogan M., Mac Carthy P.J. (1955). "The Industrial Mobility of Labor as a Probability Process", in *"Cornell Studies of Industrial and Labor Relations"*, Ithaca N.Y., Cornell University Press, vol. VI.

Bonnal L. (1992). "Microéconométrie de la durée de chômage et de la recherche d'emploi", thèse de doctorat en sciences économiques, Université des Sciences Sociales de Toulouse.

Bonnal L., Fougère D. (1990). "Les déterminants individuels des durées de chômage", *Economie et Prévision*, n°5, pp. 45–82.

Bonnal L., Fougère D. (1992). "Estimating the Structural Effect of Unemployment Insurance on Job Search Costs", *Cahier Gremaq*, n°92.20.267, Université des Sciences Sociales de Toulouse.

Bonnal L., Fougère D., Sérandon A. (1994). "L'impact des dispositifs d'emploi sur le devenir des jeunes chômeurs: une évaluation économétrique sur données longitudinales", *Economie et Prévision*, n°4, pp. 1-28.

Burdett K., Kiefer N., Mortensen D.T., Neuman G. (1984). "Earning, Unemployment and the Allocation of Time over Time", *The Review of Economic Studies*, vol. 51, pp. 559–578.

Cazals C. (1992). "Transitions emploi-retraite : étude économétrique sur données individuelles", thèse de doctorat en sciences économiques, Université des Sciences Sociales de Toulouse.

Cases C., Lollivier S. (1993 a). "Estimation d'un modèle de sortie de chômage à destinations multiples", *Document de travail* 9327, Crest–Insee, Paris.

Cases C., Lollivier S. (1993 b). "A Structural Model of Transition from Unemployment with Multiple Issues", Crest, *Document de travail* 9332,Crest–Insee, Paris.

Chesher A., Lancaster T. (1983). "The estimation of Models of Labour Market Behavior", *The Review of Economic Studies*, vol. 50, pp. 609–624.

Courgeau D., Lelievre E. (1989). *"Analyse démographique des biographies"*, Editions de l'Ined, Paris.

Cox D.R., Isham V. (1980). *"Point Processes"*, Chapman and Hall, London.

Cox D.R., Lewis P.A.W. (1966). *"The Statistical Analysis of Series of Events"*, Chapman and Hall, London.

Daley D.J., Vere-Jones D. (1988). *"An Introduction to the Theory of Point Processes"*, Springer-Verlag, New-York.

Eckstein Z., Wolpin K. (1989 a). "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience", *Review of Economic Studies*, vol. 56, pp. 375–390.

Eckstein Z., Wolpin K. (1989 b). "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey", *The Journal of Human Resources*, vol. 24, pp. 562–598.

Elbers C. et Ridder G. (1982). "True and Spurious Duration-Dependence : the Identifiability of the Proportional Hazard Model", *The Review of Economic Studies*, vol. 49, pp. 403–409.

Flinn C., Heckman J. (1982). "New Methods for Analyzing Structural Models of Labor Force Dynamics", *Journal of Econometrics*, Annals 1, vol. 18, pp. 115–168.

Flinn C., Heckman J. (1983). "Models of the Analysis of Labor Force Dynamics", in *Advances in Econometrics*, vol. I, édité par R. Basmann and G. Rhodes. JAI Press, Greenwich, pp. 35–95.

Florens J.P., Fougère D. (1991). "Non Causality in Continuous Time: Application to Counting Processes", *Cahier Gremaq 91b*, Université des Sciences Sociales de Toulouse.

Florens J.P., Fougère D. (1992). "Point Processes", in *The Econometrics of Panel Data*, édité par L. Matyas et P. Sevestre, Kluwer Academic Press, pp. 316–352.

Florens J.P., Mouchart M., Rolin J.M. 1992. "Bayesian Analysis of Mixtures : Some Results on Exact Estimability and Identification", in *Bayesian Statistics 4*, édité par J.M. Bernardo et al., Oxford Science Publication, 127-145.

Florens J.P., Mouchart M., Rolin J.M. (1993). "Non Causality and Marginalization of Markov Processes", *Econometric Theory*, vol. 9, pp. 241–262.

Fougère D., Kamionka T. (1992 a). "Un modèle markovien du marché du travail", *Annales d'Economie et de Statistique*, n° 27, pp. 149–189.

Fougère D., Kamionka T. (1992 b). "Mobilité et précarisation sur le marché français du travail : une analyse longitudinale pour les années 1986 à 1988", *Economie et Prévision*, n° 1-2, pp. 157–178.

Fougère D., Kamionka T. (1992 c). "Bayesian Inference for the Mover Stayer in Continuous Time", *Cahier Gremaq* n° 92.38.285, Université des Sciences Sociales de Toulouse.

Fougère D., Kamionka T. (1992 d). "Individual Labour Market Transitions", in *The Econometrics of Panel Data*, édité par L. Mátyás et P. Sevestre, Kluwer Academic Press, pp. 470–508.

Fougère D., Tahar G. (1987). "Women's Labour Force Participation and Marriage: a Study of Interdependences Using Cohort Data", *Document de Travail du Cejee*, Université des Sciences Sociales de Toulouse.

Geweke J., Marshall R., Zarkin G. (1986 a). "Exact Inference for Continuous Time Markov Chains", *The Review of Economic Studies*, vol. 53, pp. 653–669.

Geweke J., Marshall R., Zarkin G. (1986 b). "Mobility Indices in Continuous Time Markov Chains", *Econometrica*, vol. 54, pp. 1407–1423.

Gourieroux C., Monfort A. (1991). "Modèles de durées et effets de génération", *Document de Travail*, Crest–Insee, Paris.

Heckman J. (1991). "Indentifying the Hand of Past: Distinguishing State Dependence from Heterogeneity", *The American Economic Review*, vol. 81, pp. 75–79.

Heckman J., Singer B. (1984). "Econometric Duration Analysis", *Journal of Econometrics*, vol. 24, pp. 63–132.

Heckman J., Singer B. (1985). *"Longitudinal Analysis of Labor Market Data"*, Econometric Society Monographs, Cambridge University Press, New-York, J. Heckman et B. Singer eds.

Heckman J., Walker J. (1990). "The Relationship between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data", *Econometrica*, vol. 58, pp. 1411–1442.

Hotz V.J., Miller R. (1993). "Conditional Choice Probabilities and the Estimation of Dynamic Models", *The Review of Economic Studies*, vol. 60, pp. 497–529.

Jacobsen M. (1982). *"Statistical Analysis of Counting Processes"*, Springer-Verlag, Berlin.

Joutard X., Werquin P. (1992). "Les déterminants individuels de la durée du chômage : de l'intérêt de distinguer les emplois stables des emplois précaires", *Economie et Prévision*, n° 1-2, pp-143–156.

Kalbfleisch J.D., Lawless J.F. (1985). "The Analysis of Panel Data under a Markov Assumption", *Journal of the American Statiscal Association*, vol. 80, pp. 863–871.

Kamionka T. (1992). "Analyse microéconométrique des transitions sur le marché du travail", thèse de doctorat en sciences économiques, Université des Sciences Sociales de Toulouse.

Kamionka T. (1994). "Heterogeneity and Transition Models: an Analysis of Several Confinement Patterns on the French Labor Market", *Cahier Gremaq* 94.08.334, Université des Sciences Sociales de Toulouse.

Kamionka T., Pastorello S. (1992). "Structural and Reduced Form Analysis of Labour Mobility Processes with Discrete Time Observations: a Comparison France/Italy", *Cahier Gremaq* 92.36.283, Université des Sciences Sociales de Toulouse.

Karr A.F. (1986). *"Point Processes and their Statistical Inference"*, Marcel Dekker, New-York.

Kiefer N. (1988). "Economic Duration Data and Hazard Functions", *The Journal of Economic Literature*, vol. 26, pp. 646–679.

Lancaster T. (1990). *"The Econometric Analysis of Transition Data"*, Econometric Society Monographs, Cambridge University Press, New-York.

Lancaster T., Imbens G. (1991). "Choice-Based Sampling of Dynamic Populations", *Working Paper*, Brown University.

Lollivier S. (1994). "L'évolution du marché du travail dans les années 1980", *Revue Economique*, vol. 45, n° 3, pp. 429–441.

Mac Call J. (1970). "Economics of Information and Job Search", *The Quarterly Journal of Economics*, vol. 84, pp. 113–126.

Magnac T., Robin J.M. (1991). "An Econometric Analysis of Labour Market Transitions Using Discrete and Tenure Data", *Document de travail* 9112, Crest–Insee, Paris.

Miller R. (1984). "Job Matching and Occupational Choice", *Journal of Political Economy*, vol. 92, pp. 1086–1120.

Mortensen D. (1970). "Job Search, the Duration of Unemployment, and the Phillips Curve", *The American Economic Review*, vol. 60, pp. 847–862.

Mortensen D., Neumann G.R. (1984). "Choice or Chance ? A Structural Interpretation of Individual Labor Market Histories", in *Studies in Labor Market Dynamics*, édité par G.R. Neumann et N. Westergard-Nielsen, Springer-Verlag, Berlin.

Moreau A., Visser M. (1992). "Duration of Youth Unemployment in France", *Annales d'Economie et de Statistique*, n° 20-21.

Murphy K.M., Welch F. (1992). "The Structure of Wages", *The Quarterly Journal of Economics*, vol. 57, pp.285–326.

Newman J., Mac Cullough C. (1984). "A Hazard Rate Approach to the Timing of Births", *Econometrica*, vol. 52, pp. 939–962.

Olsen R., Smith D., Farkas G. (1986). "Structural and Reduced Form Models of Choice among Alternatives in Continuous Time: Youth Employment under a Guaranteed Jobs Program", *Econometrica*, vol. 54, pp. 375–394.

Pakes A. (1986). "Patents as Options: some Estimates of the Value of Holding European Patent Stocks", *Econometrica*, vol. 54, pp. 755–784.

Ridder G. (1984). "The Distribution of Single-Spell Duration Data", in *Studies in Labor Market Dynamics*, édité par G.R. Neumann et N. Westergard-Nielsen, Springer-Verlag, Berlin.

Ridder G. (1986). "An Event History Approach to the Evaluation of Training Recruitment and Employment Programs", *Journal of Applied Econometrics*, vol. 1, pp. 109–126.

Ruggiero M. (1992). "Ancienneté au chômage et principaux facteurs associés", *Economie et Prévision*, n° 4, p. 35–53.

Rust J. (1987). "Optimal Replacement of Bus Engines: an Empirical Model of Harold Zurcher", *Econometrica*, vol. 55, pp. 999–1034.

Rust J. (1993). "Estimation of Dynamic Structural Models : Problems and Prospects. Part 1: Discrete Decision Processes", in *Advances in Econometrics: Proceedings of the Sixth World Congress of the Econometric Society*, édité par C. Sims, Cambridge University Press, New-York.

Singer B. (1982). "Aspects of Non-Stationarity", *Journal of Econometrics*, Annals of Applied Econometrics 1982-1, vol. 18, pp. 169–190.

Singer B., Spilerman S. (1976). "The Representation of Social Processes by Markov Models", *American Journal of Sociology*, vol. 82, pp. 1–54.

Tuma N.B., Robins P.K. (1980). "A Dynamic Model of Employment Behavior: An Application of the Seattle and Denver Income Maintenance Experiments", *Econometrica*, vol. 48, pp. 1031–1052.

Visser M. (1992). "Analysis of Labour Market Histories with Panel Data", *Document de Travail* 9209, Insee, Paris.

Wolpin K. (1987). "Estimating a Structural Search Model: the Transition from School to Work", *Econometrica*, vol. 55, pp. 801–817.