JOURNAL DE LA SOCIÉTÉ FRANÇAISE DE STATISTIQUE

THIERRY KAMIONKA

La modélisation des données haute fréquence

Journal de la société française de statistique, tome 141, n° 1-2 (2000), p. 167-211

<http://www.numdam.org/item?id=JSFS_2000__141_1-2_167_0>

© Société française de statistique, 2000, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (http://publications-sfds.math.cnrs.fr/index.php/J-SFdS) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

LA MODÉLISATION DES DONNÉES HAUTE FRÉQUENCE

Thierry KAMIONKA [1]

RÉSUMÉ

La mise en place des systèmes automatisés de cotation autorise l'enregistrement en continu des transactions et des ordres. Aussi, les instants d'arrivée des événements analysés doivent être considérés comme aléatoires. Pour étudier les motifs de placement des ordres et les implications des théories de la microstructure des marchés financiers, on a été amené à considérer les aspects liés à la modélisation de durées issues de ces données haute fréquence. Nous traitons aussi les cas où on doit modéliser de façon jointe une durée et une caractéristique qui peut être discrète, continue ou encore vectorielle. On peut avoir à modéliser des durées multiples. Ainsi, on peut être amené à étudier la durée entre transactions successives et la durée entre les modifications des cotations. Nous écrivons, alors, la vraisemblance en terme des fonctions de hasard des durées conditionnelles à l'ensemble du passé du processus et aux caractéristiques du marché. Ce conditionnement peut incorporer l'ensemble des événements passés affectant les distributions conditionnelles des temps en avant.

INTRODUCTION

La mise en place de systèmes automatisés de cotation (Supercac à la Bourse de Paris ou SuperDOT pour le NYSE) autorise l'enregistrement en continu des transactions, des modifications des cotations à la vente ou à l'achat, ou encore, l'ensemble des ordres placés sur un marché au cours d'une période de temps donnée. Ces systèmes permettent aussi d'enregistrer l'état du carnet d'ordres et son évolution au cours du temps (c'est le cas à la Bourse de Paris). La disponibilité de telles données nécessite de considérer les temps d'arrivée des ordres ou des transactions comme aléatoires. Ainsi, on peut s'attendre qu'à la suite d'une série de transactions rapprochées dans le temps, l'espérance de la durée d'attente pour la prochaine transaction soit plus faible.

---

1. CNRS, CREST-INSEE (Paris) et IDEI (Toulouse).
Adresse : Thierry Kamionka; CREST; Timbre J390; 15, Bd Gabriel Péri, 92245 Malakoff Cedex; FRANCE. e-mail : kamionka@ensae.fr. Je remercie Bruno Biais, Christian Gouriéroux, Gaëlle Le Fol et deux rapporteurs anonymes pour leurs commentaires. Les participants de la journée « Risque de modèle » de la SFdS, des séminaires Paris-Berlin à Garchy, Finance-Assurance du Crest à Paris, d'économétrie du CORE à Louvain-la-Neuve, du GREQAM à Marseille et du GREMARS à Lille ont fait de nombreuses suggestions utiles. Les précautions habituelles s'appliquent. Cet article a été écrit, en grande partie, lorsque j'étais membre du GREMAQ, UMR du CNRS 5604, Université Toulouse I.

Biais, Hillion et Spatt [1995], à partir de données haute fréquence de la bourse de Paris, étudient la relation existant entre le flux d'arrivée des ordres et la composition du carnet d'ordre dans le but d'analyser certaines propositions issues des théories de la microstructure des marchés financiers (*cf.* Glosten et Milgom 85, Kyle 85, Easley et O'hara 92). Dans cet article, les auteurs utilisent une classification des ordres selon leur direction (achat ou vente) et selon la volonté que l'investisseur a de voir son ordre exécuté plus ou moins rapidement. Ces auteurs étudient la durée qui s'écoule entre deux ordres successifs et ses relations avec les caractéristiques du carnet d'ordres. Par exemple, dans la mesure où il est relativement coûteux de consommer la liquidité, une fourchette de prix grande doit favoriser l'arrivée d'ordres à cours limites et défavoriser l'apparition d'ordres à exécution immédiate. En ce sens, à la fois la fourchette de prix, les quantités disponibles de chaque côté du carnet d'ordre, le type des ordres qui ont été soumis (quantité associée, situation du prix par rapport à la composition du carnet d'ordres) vont influencer les réactions des autres investisseurs et avoir un effet sur la distribution conditionnelle du processus d'arrivée des ordres.

Par la suite, Engle [2000], Engle et Russel [1998] ont proposé une classe de modèles particulièrement adaptés à la modélisation des données haute fréquence. Il s'agit des modèles de durée ACD. Ils supposent que l'espérance conditionnelle de la durée a une structure auto-régressive. Engle et Russel [1998] ont utilisé ce type de modélisation pour étudier les durées entre transactions sur le titre IBM négocié à la bourse de New-York (NYSE).

D'autres auteurs ont ensuite complété cette classe de modèles. C'est le cas de Bauwens et Giot [1997] qui proposent un modèle dans lequel le logarithme de l'espérance conditionnelle de la durée entre transactions va avoir une forme auto-régressive : il s'agit du modèle Log-ACD. Bauwens et Veradas [1999] ont introduit le modèle de durée SCD (stochastic conditional duration model). Dans le modèle SCD, les durées sont générées par une distribution qui dépend d'un facteur latent stochastique. La distribution de la durée est déterminée par un mélange de distributions et ce modèle est utilisé pour étudier les durées entre deux modifications des prix.

Dans certains travaux réalisés sur des données haute fréquence, les auteurs ont tenu compte de l'information contenue dans les durées et dans les prix. Ainsi, Engle [2000] modélise les durées entre transactions en utilisant un modèle ACD et un modèle conditionnel de type GARCH pour les prix. Dans ce modèle, la volatilité conditionnelle dépend de l'information passée et des durées. Ghysels et Jasiak [1996] introduisent une classe de modèles appelée ACD-GARCH dans lesquels une caractéristique de la distribution des durées entre transactions va déterminer un coefficient du modèle conditionnel de type GARCH. Russel et Engle [1998] modélisent de manière jointe la probabilité de transition du prix et le temps d'arrivée des transactions.

Bauwens et Giot [1998] modélisent les durées entre les modifications des cotations et caractérisent ces modifications par une variable dichotomique (augmentation, diminution). Ils utilisent un modèle de transition dans lequel la durée latente pour une catégorie de la variation des cotations dépend

des caractéristiques passées du marché. Ainsi, Bisière et Kamionka [2000], considèrent un modèle de transition pour étudier les durées qui séparent deux ordres successifs à la bourse de Paris et le type du prochain ordre conditionnellement à l'état du marché au moment ou le dernier ordre a été placé. Les valeurs estimées des paramètres de ce modèle sont utilisées pour considérer un certain nombre d'implications des théories de la microstructure des marchés financiers.

En général, l'ensemble d'information des agents peut varier à chaque arrivée d'un événement sur le marché et, en particulier, lorsque un ordre limite est placé. Aussi, lorsque l'on modélise une durée – par exemple, entre transactions – on doit s'attacher à tenir compte de l'effet sur le hasard conditionnel de l'arrivée de ces événements intermédiaires.

Dans cet article, nous proposons une revue de la littérature relative à la modélisation des données haute fréquence. Ces modèles ont en commun le caractère aléatoire des instants d'arrivée des événements étudiés : les transactions, par exemple. Aussi, lorsque l'on utilise de telles données, on doit modéliser, en particulier, la durée qui sépare deux événements successifs.

Dans cet article, nous proposons, aussi, un cadre permettant de modéliser un ensemble de durées financières qui peuvent se dérouler simultanément.

Dans la section 2, nous traitons le cas de la modélisation marginale des durées en présentant les classes de modèles les plus couramment utilisées. Dans la section 3, nous considérons le cas de la modélisation jointe d'une durée et d'une caractéristique (discrète, continue ou vectorielle). La section 4 est consacrée à la présentation de modèles bivariés à épisodes multiples. La section 5 contient le cas de la modélisation des durées d'exécution des ordres limites. Dans la dernière section nous concluons.

2. MODÉLISATION DES DURÉES

On observe une séquence de durées financières. Il peut s'agir de durées entre transactions, de durées entre les instants de modification des cotations, de durées entre des instants de modification du prix, de durées pour qu'une quantité minimale soit échangée [*cf.* Giot, (2000)], une séquence de durées entre ordres ou, encore, de durées pour que des ordres limites soient exécutés partiellement ou complètement.

Soit $\tau_{\ell-1}$ le temps d'arrivée de l'événement (transaction ou ordre) $\ell-1$ dans la journée et τ_ℓ le temps d'arrivée de l'événement ℓ. $u_\ell = \tau_\ell - \tau_{\ell-1}$ est le temps qui s'écoule entre les événements $\ell-1$ et ℓ (*cf.* graphique 1).

2.1. Prise en compte de l'activité intra-journalière

Plusieurs auteurs ont montré l'existence d'un effet correspondant au niveau de l'activité intra-journalière. Gourieroux, Jasiak et Le Fol [1999] mettent en évidence, à partir d'une estimation non paramétrique de la densité des

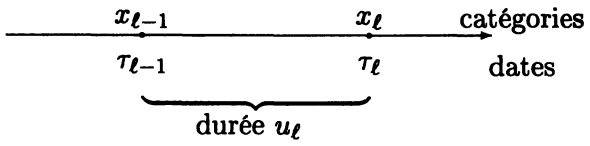

FIG 1. — Episode ℓ

durées entre transactions à la Bourse de Paris, la forme en U usuelle lorsqu'ils incorporent les ouvertures. Ainsi, l'activité du marché est maximale en début et en fin de journée et minimale au milieu de la journée (entre 13h00 et 13h30). Cette structure en U caractéristique est retrouvée par Bisière et Kamionka [2000], pour Alcatel, sur les mêmes données de la Bourse de Paris.

Bisière et Kamionka [2000] proposent d'incorporer le traitement de cette composante à l'intérieur de la journée à partir de la spécification des fonctions de hasard. L'intérêt d'un tel traitement est de pouvoir distinguer différents schémas de saisonnalité selon, par exemple, la direction des ordres qui sont placés.

Engle et Russel [1998] proposent d'estimer cette composante spécifique intra-journalière de la distribution des durées de façon à travailler ensuite sur les données auxquelles on a enlevé l'effet intra-journalier.

Dans ce cas, on va considérer la séquence des durées :

$$u'_\ell = \frac{u_\ell}{\phi(\tau_{\ell-1})},$$

où ϕ est une fonction de la date de démarrage de l'épisode ℓ [3].

Dans le cas où ϕ est une spline cubique (voir Green et Silverman, 1994), on a

$$\phi(t) = \xi_{k,3}(t-t_k)^3 + \xi_{k,2}(t-t_k)^2 + \xi_{k,1}(t-t_k) + \xi_{k,0}, \text{ pour } t_k \leqslant t \leqslant t_{k+1},$$

où $k = 0, \ldots, k_0$ ($\xi_{k,j}$ est un paramètre réel).

Les paramètres de la fonction ϕ sont contraints par des conditions de continuité et de dérivabilité.

Engle [2000] propose de régresser les durées sur une spline linéaire par morceaux. On va découper la journée en intervalles disjoints $[t_{k-1}; t_k[$ pour $k = 1, \ldots, k_0$ et $[t_{k_0}; t_{k_0+1}]$.

3. Par la suite, nous supposerons que nous travaillons sur les données une fois enlevée cette composante et nous omettrons le prime.

Dans ce cas, on pose

$$\phi(t) = \xi_0 + \sum_{m=0}^{k-1} \xi_{m,1}(t - t_m), \text{ si } t_{k-1} \leqslant t < t_k,$$

où ξ_0 et $\xi_{m,1}$ $(m = 0, \ldots, k_0)$ sont des paramètres réels.

Considérons les données concernant le titre IBM échangé sur le NYSE de septembre à novembre 1998 (63 jours ouvrables). Ces données proviennent de la base TAQ du NYSE (Trades and Quotes). Construisons des intervalles d'une demie heure de 9h30 à 16h00 ($k_0 = 12$). Nous n'avons considéré que les transactions passées entre 9h30 et 16h00. Comme Engle [2000] nous ne tenons pas compte des durées égales à zéro et celles associées à des épisodes situés sur plusieurs journées. Nous avons ainsi retenu $N = 93489$ durées entre transactions.

Pour obtenir le tracé de la fonction $\phi(t)$, nous considérons le modèle de regression linéaire dans lequel :

$$E(u_\ell \mid I_{\ell-1}) = \phi(\tau_{\ell-1}),$$

où $I_{\ell-1}$ est l'ensemble d'information disponible à la date $\tau_{\ell-1}$.

Le tracé de la fonction ϕ est alors donné au graphique 2. L'activité du marché pour ce titre est bien minimale en milieu de journée [4].

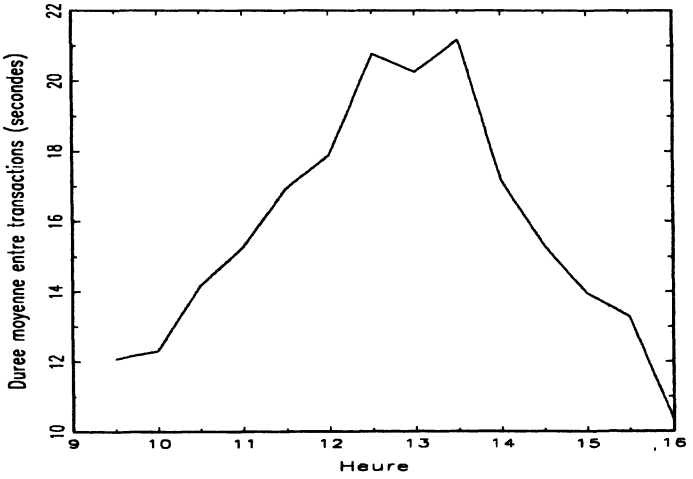

FIG 2. — Activité (IBM, sep-oct 98, NYSE)

4. Ce graphique a été réalisé, en particulier, à partir d'une adaptation des programmes EViews 3.1 proposés par Engle sur son site Web [voir Engle (2000)].

Bauwens et Giot [1997] déterminent $\phi(t)$ en faisant la moyenne des durées observées sur des intervalles de 30 minutes[5]. Soit N le nombre total d'observations présentes dans l'échantillon. Ainsi, on a

$$\phi(t) = \frac{1}{n_k} \sum_{j \in A_k} u_j, \text{ si } t_{k-1} \leqslant t < t_k,$$

où

$$A_k = \{\ell \in \{1, 2, \ldots, N\} : t_{k-1} \leqslant \tau_{\ell-1} < t_k\} :$$
$$n_k = \text{Card}(A_k) \ (k = 1, \ldots, k_0)$$

et

$$\phi(t) = \frac{1}{n_{k_0+1}} \sum_{j \in A_{k_0+1}} u_j, \text{ si } t_{k_0} \leqslant t \leqslant t_{k_0+1},$$

où $A_{k_0+1} = \{\ell \in \{1, 2, \ldots, N\} : t_{k_0} \leqslant \tau_{\ell-1} \leqslant t_{k_0+1}\}$, $n_{k_0+1} = \text{Card}(A_{k_0+1})$.
Le tracé de la fonction $\phi(t)$ de 9h30 à 16h00 est donné au graphique 3 pour le titre IBM. Au milieu de chaque intervalle d'une demie heure, on a placé la durée moyenne calculée sur l'intervalle. Les points sont reliés par des segments de droites.

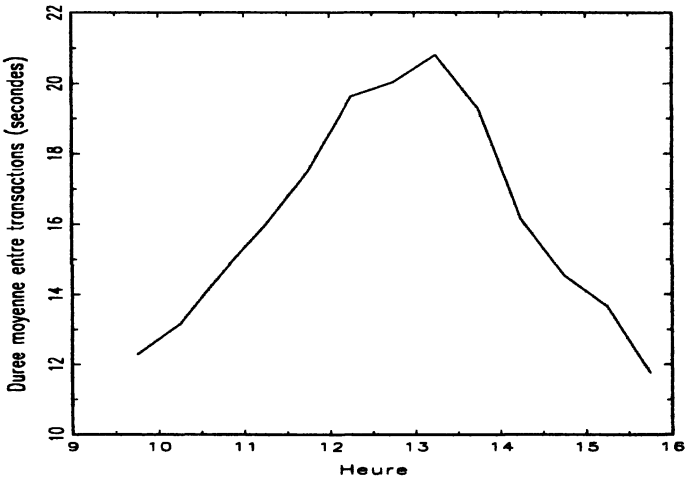

FIG 3. — Activité (IBM, sep-oct-nov 98, NYSE)

5. Bauwens et Giot [1997] lissent, ensuite, ces moyennes par des splines.

2.2. Modélisation

On doit modéliser les durées entre transactions, ou entre ordres successifs, ou jusqu'à ce que les ordres soient exécutés [6], ou pour que la variation du prix dépasse une valeur fixée, ou pour que l'on échange un quantité minimale [*cf.* Giot, (1999) ou Gouriéroux, Jasiak et Le Fol, (1999)]. Soit U une variable aléatoire continue non négative. U représentera la durée qui est étudiée.

La distribution de U peut être représentée par sa fonction de densité de probabilité $f(u;\theta)$, sa fonction de répartition $F(u;\theta)$, sa fonction de survie $S(u;\theta) = 1 - F(u;\theta)$ ou par sa fonction de hasard $h(u;\theta) = f(u;\theta)/S(u;\theta)$. θ est un vecteur de paramètres ($\theta \in \mathbb{R}^p$).

On peut considérer différentes spécifications paramétriques pour la distribution de la durée. Deux spécifications sont particulièrement générales.

Si la variable U est distribuée selon une loi Singh-Maddala de paramètres $\alpha > 0$, $\lambda > 0$ et $\beta > 0$ alors

$$f(u;\theta) = \lambda^\alpha \ \alpha \ u^{\alpha-1}[1 + (\beta \ u)^\alpha]^{-(\frac{\lambda}{\beta})^\alpha - 1},$$

et

$$h(u;\theta) = \lambda^\alpha \ \alpha \ u^{\alpha-1}[1 + (\beta \ u)^\alpha]^{-1}.$$

La distribution Singh-Maddala englobe la loi log-logistique lorsque $\beta = \lambda$, la loi Weibull lorsque $\beta = 0$, la loi exponentielle lorsque $\beta = 0$ et $\alpha = 1$.

Par exemple, pour $0 < \alpha < 1$, la fonction de hasard est décroissante. Pour $\alpha > 1$ et $\beta \neq 0$, la fonction de hasard est d'abord croissante puis décroissante.

Lorsque la variable U est distribuée selon une loi Gamma généralisée, on a que

$$f(u;\theta) = \frac{1}{\Gamma(\alpha)} \ \beta \ \lambda \ (\lambda \ u)^{\beta\alpha-1} \exp(-(\lambda \ u)^\beta),$$

$$h(u;\theta) = \frac{(\lambda \ u)^{\beta\alpha-1} \exp(-(\lambda \ u)^\beta)}{\displaystyle\int_u^{+\infty} (\lambda \ s)^{\beta\alpha-1} \exp(-(\lambda \ s)^\beta) \ ds},$$

où $\alpha > 0$, $\lambda > 0$ et $\beta > 0$.

La loi Weibull (lorsque $\alpha = 1$) est un cas particulier de distribution Gamma généralisée.

Supposons que l'on observe un échantillon $((u_1, \delta_1), \ldots, (u_N, \delta_N))$ i.i.d. où [7]

$$\delta_i = \begin{cases} 1, & \text{si l'observation i n'est pas censurée à droite,} \\ 0, & \text{sinon.} \end{cases}$$

---

6. Pour une application empirique, voir Lo, MacKinlay et Zhang, 1997.

7. Dans le cas de durées financières, l'hypothèse d'indépendance n'est pas raisonnable. Cependant, comme on va le voir dans ce qui suit, on peut se ramener à ce cas de figure. En effet, on considèrera des variables $\epsilon_1, \ldots, \epsilon_N$ indépendantes telles que $\epsilon_\ell = \frac{u_\ell}{\Phi_\ell}$, où Φ_ℓ ne dépend que des réalisations des durées situées avant l'épisode ℓ ($\ell = 1, \ldots, N$).

La fonction de vraisemblance associée à l'échantillon s'écrit :

$$L(\theta) = \prod_{i=1}^{N} h(u_i;\theta)^{\delta_i} \exp\{-\int_0^{u_i} h(s;\theta)\, ds\},$$

où θ est un vecteur de paramètres ($\theta \in \mathbb{R}^p$).

Sous les conditions de régularité usuelles [voir, par exemple, Gouriéroux et Monfort, (1989)], l'estimateur du maximum de vraisemblance de θ existe et est convergent. La distribution asymptotique de l'estimateur maximum de vraisemblance est celle d'une loi normale.

Il existe deux catégories principales de modèles de durée conditionnels : les modèles à vie accélérée et les modèles à hasards proportionnels.

Dans le cas des modèles de durée à vie accélérée, on suppose que les variables explicatives ont un effet multiplicatif sur la durée. La fonction de survie de la durée conditionnelle aux variables explicatives z est telle que

$$S(u \mid z;\theta) = S_0(u\ \varphi(z;\theta_1);\theta_2),$$

où $\varphi(z;\theta_1)$ est une fonction positive de la variable z. θ, θ_1 et θ_2 sont des vecteurs de paramètres ($\theta' = (\theta_1', \theta_2')$) et S_0 est la fonction de survie de base.

Ainsi, on obtient

$$F(u \mid z;\theta) = 1 - S(u \mid z;\theta) = 1 - S_0(u\ \varphi(z;\theta_1);\theta_1) = F_0(u\ \varphi(z;\theta_1);\theta_2),$$

$$f(u \mid z;\theta) = \frac{\partial F_0(u\ \varphi(z;\theta_1);\theta_2)}{\partial u} = \varphi(z;\theta_1)\ f_0(u\ \varphi(z;\theta_1);\theta_2),$$

$$h(u \mid z;\theta) = \frac{f(u \mid z;\theta)}{S(u \mid z;\theta)} = \varphi(z;\theta_1)\ h_0(u\ \varphi(z;\theta_1);\theta_2),$$

Par exemple, on peut prendre $\varphi(z;\theta_1) = \exp(-z'\theta_1)$.

Dans le cadre d'un modèle de durée à hasards proportionnels, la fonction de hasard conditionnelle est donnée par

$$h(u \mid z;\theta) = \varphi(z;\theta_1)\ h_0(u;\theta_2),$$

où θ_1 et θ_2 sont des paramètres et φ est une fonction positive ($\theta' = (\theta_1', \theta_2')$).

Dans ce modèle, les variables explicatives ont un effet multiplicatif sur le hasard de base h_0.

La fonction de survie dans ce modèle devient

$$S(u \mid z;\theta) = \exp\{-\int_0^u h(s \mid z;\theta)\ ds\} = S_0(u;\theta_2)^{\varphi(z;\theta_1)}.$$

La fonction de densité de probabilité dans un modèle de durée à hasards proportionnels est donnée par

$$f(u \mid z;\theta) = h(u \mid z;\theta)\ S(u \mid z;\theta) = \varphi(z;\theta_1) f_0(u;\theta_2)\ S_0(u;\theta_2)^{\varphi(z;\theta_1)-1}.$$

Les paramètres de ce modèle peuvent être estimés par maximum de vraisemblance. On peut, par ailleurs, obtenir une estimation non paramétrique de la fonction de survie de la durée sur une strate définie par une combinaison des variables explicatives.

Soient $u_{(1)} < u_{(2)} < \ldots < u_{(N)}$ les durées observées après avoir été ordonnées. N est le nombre total de durées de l'échantillon. On suppose ici qu'il n'y a ni censure, ni ex aequo.

$R_i = \{(i), (i+1), \ldots, (N)\}$ défini un ensemble d'observations : celles dont les valeurs sont supérieures ou égales à $u_{(i)}$. Soit r_i le nombre de durées dont les valeurs sont supérieures ou égales à $u_{(i)}$.

L'estimateur nonparamétrique de la fonction de survie à pour expression

$$\hat{S}(u) = \prod_{i:u_{(i)}<u} \frac{r_i - 1}{r_i}.$$

Il s'agit de l'estimateur Kaplan-Meier de la fonction de survie (cf. Cox et Oakes, 1984).

Dans le cas de censure ou d'ex aequo, l'expression de l'estimateur devient

$$\hat{S}(u) = \prod_{i:u_{(i)}<u} \frac{r_i - e_i}{r_i},$$

où e_i est le nombre de durées dont les valeurs sont exactement égales à $u_{(i)}$ et non censurées.

2.3. Le Modèle ACD

Ce modèle a été introduit par Engle [2000] et Engle et Russel [1998] pour étudier des durées entre transactions. Soit $z_\ell = (u_1, \ldots, u_{\ell-1})$ la séquence des réalisations passées des durées. Soit $\Phi_\ell > 0$ une fonction de z_ℓ et de θ ($\theta \in \mathbb{R}^s$).

Le modèle :

$$\begin{cases} u_\ell &= \Phi_\ell \, \epsilon_\ell, \\ E[u_\ell \mid z_\ell] &= \psi_\ell, \\ \psi_\ell &= \omega + \sum_{k=1}^{p} \alpha_k u_{\ell-k} + \sum_{k=1}^{q} \beta_k \psi_{\ell-k}, \end{cases}$$

où les ϵ_ℓ sont i.i.d. et ont une distribution à support sur $\mathbb{R}^+$ qui ne dépend pas de z_ℓ.

Dans le modèle ACD, l'espérance conditionnelle de la durée a une forme autorégressive.

Remarque : Dans le cas où $E[\epsilon_\ell \mid z_\ell] = 1$, alors on a $\Phi_\ell = \psi_\ell$.

Soit $f_0(\epsilon, \theta_1)$ la densité de ϵ_ℓ ($\theta_1 \in \Theta_1$). La fonction de densité de U_ℓ va s'écrire

$$f(u \mid z_\ell; \theta) = \frac{1}{\Phi_\ell} f_0(\frac{u}{\Phi_\ell}; \theta_1).$$

La fonction de hasard conditionnelle de la durée entre deux événements successifs [8] a pour expression

$$h(u \mid z_\ell) = \frac{1}{\Phi_\ell} h_0(\frac{u}{\Phi_\ell}, \theta_1).$$

où $h_0(\epsilon; \theta_1)$ est la fonction de hasard de ϵ_ℓ.

Le modèle ACD est donc un modèle de durée à vie accélérée.

Exemple 1 : On va supposer que $\epsilon_\ell \sim$ *i.i.d.* $Weibull(1, \gamma)$, alors

$$f_0(\epsilon; \gamma) = \gamma\, \epsilon^{\gamma-1} \exp(-\epsilon^\gamma),\ \gamma \in \mathbb{R}^{+*},$$

$$f(u \mid z_\ell; \theta) = \gamma\, u^{\gamma-1} \frac{1}{\Phi_\ell^\gamma} \exp\{-\frac{u^\gamma}{\Phi_\ell^\gamma}\}.$$

Dans ce cas, on obtient que $E[u_\ell \mid z_\ell] = \Phi_\ell\, \Gamma(1 + \frac{1}{\gamma}) = \psi_\ell$.

Exemple 2 : On va supposer que $\epsilon_\ell \sim$ *i.i.d.* $Weibull(b, \gamma)$, où $b = \Gamma(1 + \frac{1}{\gamma})^\gamma$, on obtient [9] que

$$h(u \mid z_\ell) = \frac{1}{\psi_\ell} h_0(\frac{u}{\psi_\ell}, \gamma),$$

où $h_0(\epsilon, \gamma) = \gamma \epsilon^{\gamma-1} \Gamma(1 + \frac{1}{\gamma})^\gamma$.

Comme $E(\epsilon) = \dfrac{\Gamma(1 + \frac{1}{\gamma})}{b^{1/\gamma}}$, on peut vérifier que, dans ce cas, $E[u_\ell \mid z_\ell] = \psi_\ell = \Phi_\ell$.

Remarque : Pour $\gamma = 1$, on obtient le modèle $EACD$.

Exemple 3 : Le modèle Burr-ACD [*cf.* Grammig et Maurer, (1999)]
On suppose que $\epsilon_\ell \sim Burr(1, \alpha, \beta^\alpha)$. Alors, la loi conditionnelle de u_ℓ est une $Burr(\frac{1}{\Phi_\ell^\alpha}, \alpha, (\beta\, \Phi_\ell^{-1})^\alpha)$.

8. Il peut s'agir de deux transactions, par exemple.
9. De fonction de densité $f_0(\epsilon, \gamma) = b\, \gamma\, \epsilon^{\gamma-1} \exp(-b\, \epsilon^\gamma)$, où $b = \Gamma(1 + \frac{1}{\gamma})^\gamma$.

La fonction de hasard la loi $Burr(\mu, \alpha, \sigma^2)$ est égale à

$$h_0(\epsilon) = \frac{\alpha\ \mu\ \epsilon^{\alpha-1}}{(1+\sigma^2\ \mu\ \epsilon^{\alpha})}.$$

Aussi, la fonction de hasard conditionnelle a pour expression

$$h(u \mid z_\ell) = \frac{\alpha\ u^{\alpha-1}}{\Phi_\ell^{\alpha}\ (1+\beta^{\alpha}\Phi_\ell^{-\alpha}\ u^{\alpha})}.$$

Il s'agit d'un modèle Singh-Maddala-ACD (de paramètres $(\alpha, \Phi_\ell^{-1}, \beta\Phi_\ell^{-1})$).

Estimation : Considérons l'exemple 2. On va maximiser la quasi vraisemblance [10]

$$Q(\theta) = \prod_{\ell=1}^{N} \frac{1}{\psi_\ell} \exp(-\frac{u_\ell}{\psi_\ell}),$$

où N est le nombre total de durées considérées dans l'échantillon.

On peut montrer (voir Engle 2000, Engle et Russel 1998) que l'estimateur du quasi maximum de vraisemblance est convergent et asymptotiquement normal. $V^{-\frac{1}{2}}\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow{d} N(0, I)$ et la matrice V peut être estimée par $\hat{V} = \hat{J}^{-1}\hat{I}\hat{J}^{-1}$ où $\hat{J} = \frac{1}{N}\sum_{\ell=1}^{N} d_\ell(\hat{\theta}_N)$, $\hat{I} = \frac{1}{N}\sum_{\ell=1}^{N} s_\ell(\hat{\theta}_N)\, s_\ell(\hat{\theta}_N)'$, $d_\ell(\theta) = -\frac{1}{\psi_\ell^2}\frac{\partial\psi_\ell}{\partial\theta}\frac{\partial\psi_\ell}{\partial\theta'}$ et $s_\ell(\theta) = -\frac{(\psi_\ell - u_\ell)}{\psi_\ell^2}\frac{\partial\psi_\ell}{\partial\theta}$.

Nous avons estimé deux modèles à partir des données TAQ du NYSE (*cf.* section 2.1). Un modèle ACD(1,1) et un modèle ACD(2,2) qui vérifient respectivement :

$$\psi_\ell = \omega + \alpha_1\ u_{\ell-1} + \beta_1\ \psi_{\ell-1}, \tag{1}$$

et

$$\psi_\ell = \omega + \alpha_1\ u_{\ell-1} + \alpha_2\ u_{\ell-2} + \beta_1\ \psi_{\ell-1} + \beta_2\ \psi_{\ell-2}. \tag{2}$$

Les modèles sont estimés en considérant des pseudo lois normales. Nous obtenons un estimateur du quasi maximum de vraisemblance qui est convergent quelque soit la vraie densité dans le cas où la vraie espérance conditionnelle est donnée par l'équation (1) (respectivement, l'équation (2)). Les

10. En général, les durées étudiées sont observées sur un ensemble de journées. Par exemple, Engle [2000] considère les transactions pour IBM de novembre 90 à janvier 91. L'espérance conditionnelle de la durée calculée en fin de journée et la dernière durée observée de la journée peuvent être prises pour initialiser le processus pour la journée suivante. De plus, pour le modèle ACD(1,1), on pourra fixer, par exemple, $\psi_0 = \frac{\omega}{1-\beta}$ et $u_0 = 0$.

TABLEAU 1. — Estimations de modèles ACD, durée entre transactions IBM, données TAQ (NYSE, sep-oct-nov 98)(Ecarts-types méthode de Bollerslev-Wooldridge)

	Equation 1			*Equation 2*		
Variable	*Coef*	*E. Type*	*Z-stat*	*Coef*	*E.Type*	*Z-stat*
ω	0,0142	0,0012	11,9279	0,0238	0,0039	6,1121
ARCH(1)	0,0310	0,0012	26,7757	0,0186	0,0045	4,1533
ARCH(2)				0,0292	0,0059	4,9242
GARCH(1)	0,9548	0,0019	497,1361	0,5661	0,1952	2,9005
GARCH(2)				0,3623	0,1855	1,9529
Log Vrais		-131797,0			-131790,5	

estimations sont réalisées en considérant $\sqrt{u'_\ell}$ (où $u'_\ell = \dfrac{u_\ell}{\phi(\tau_{\ell-1})}$, *cf.* section 2.1) comme variable dépendante à partir du logiciel EViews 3.1 dans le cadre d'un modèle GARCH sans constante dans l'équation de la moyenne.

La matrice des variances-covariances est celle de l'estimateur QML [voir Bollerslev et Wooldridge (1992)].

Les estimations des paramètres sont significatives à 5 % sauf pour β_2 dans le modèle ACD(2,2). On peut remarquer que le paramètre β_1, dans le modèle ACD(1,1), à une valeur estimée très élevée.

2.4. Le Modèle T-ACD

Il s'agit d'une extension du modèle ACD dans laquelle on permet aux paramètres de la distribution de varier d'un épisode à l'autre. La valeur du vecteur des paramètres relatif à l'épisode ℓ va dépendre de la valeur prise par la variable z_ℓ. Par exemple, supposons que $\theta_\ell = \theta^{(1)}$ si $u_{\ell-1} \in [0; a_1]$, $\theta_\ell = \theta^{(2)}$ si $u_{\ell-1} \in [a_1; a_2]$ $(0 < a_1 < a_2)$ et $\theta_\ell = \theta^{(3)}$ si $u_{\ell-1} \in [a_2; +\infty]$. Il s'agit ici d'un modèle à 3 régimes.

Pour chaque épisode ℓ, la distribution conditionnelle de la durée U_ℓ est obtenue à partir d'un modèle ACD dont le paramètre θ_ℓ est déterminé selon la procédure que l'on vient de décrire suivant la valeur de la durée observée à l'épisode précédent.

Le modèle :

$$\begin{cases} u_\ell &= \psi_\ell \, \epsilon_\ell, \\ E[u_\ell \mid z_\ell] &= \psi_\ell, \\ \psi_\ell &= \omega^{(\ell)} + \displaystyle\sum_{k=1}^{p} \alpha_k^{(\ell)} \, u_{\ell-k} + \sum_{k=1}^{q} \beta_k^{(\ell)} \, \psi_{\ell-k}, \end{cases}$$

où les ϵ_ℓ sont i.i.d. et ont une distribution à support sur $\mathbb{R}^+$ de paramètre γ_ℓ.

Zhang *et alii* [1999] maximisent la log-vraisemblance en utilisant une grille de recherche sur les paramètres a_1 et a_2. Ils utilisent ce modèle pour estimer la distribution des durées entre transactions pour IBM sur le NYSE dans le but de rechercher d'éventuels changements structurels. Zhang *et alii* [1999] donnent des conditions suffisantes pour que le processus soit ergodique et pour que ses moments existent.

2.5 Le Modèle Log-ACD

Dans ce modèle on fait l'hypothèse que :

$$u_\ell = e^{\Phi_\ell}\ \epsilon_\ell,$$

où $\Phi_\ell = \omega + \alpha\ g(u_{\ell-1}, \epsilon_{\ell-1}) + \beta\ \Phi_{\ell-1}$ et les ϵ_ℓ sont *i.i.d.* selon un loi de support positif. Dans ce qui suit on supposera que $\epsilon_\ell \sim Weibull(1, \gamma)$.

Bauwens et Giot [1997] proposent plusieurs choix pour la fonction g. Par exemple :

$g(u_{\ell-1}, \epsilon_{\ell-1}) = \ln(u_{\ell-1})$. Dans ce cas [11], $\Phi_\ell = \omega + (\alpha + \beta)\ \Phi_{\ell-1} + \alpha\ \ln(\epsilon_{\ell-1})$.

On peut se ramener à un modèle de régression linéaire puisque l'on a

$$\ln(u_\ell) = \Phi_\ell + \mu_0 + v_\ell,$$

où $E[v_\ell] = E[\ln(\epsilon_\ell)] - \mu_0 = 0$.

La fonction de densité de la durée u_ℓ :

$$f(u_\ell \mid z_\ell; \theta) = \gamma\ (\frac{u_\ell}{e^{\Phi_\ell}})^{\gamma-1}\ \exp\{-(\frac{u_\ell}{e^{\Phi_\ell}})^\gamma\}\ \frac{1}{e^{\Phi_\ell}}.$$

Le modèle Log-ACD est aussi un modèle de durée à vie accélérée. Conditionnellement à $Z_\ell = z_\ell$, la durée U_ℓ suit une loi $Weibull(\frac{1}{e^{\Phi_\ell \gamma}}, \gamma)$.

La log-vraisemblance s'écrit [12] :

$$L_N(\theta) = \sum_{\ell=1}^{N} [\ln(\gamma) + (\gamma - 1)\ln(u_\ell) - \gamma\Phi_\ell - (\frac{u_\ell}{e^{\Phi_\ell}})^\gamma]$$

Dans le cas où γ=1, on peut remarquer que, dans le cadre de ce modèle, on a

$$E[u_\ell \mid z_\ell] = \exp\{\omega + \alpha\ g(u_{\ell-1}, \epsilon_{\ell-1}) + \beta\ \Phi_{\ell-1}\},$$

11. On doit avoir $\mid \alpha + \beta \mid < 1$ (*cf.* Bauwens et Giot 1997).
12. *cf.* note 9.

Aussi, on peut estimer $\theta = (\omega, \alpha, \beta)$ par moindres carrés non linéaires en minimisant

$$Q(\theta) = \sum_{\ell=1}^{N} (u_\ell - m_\ell(\theta))^2,$$

où $m_\ell(\theta) = E[u_\ell \mid z_\ell] = \exp\{\omega + \alpha\ g(u_{\ell-1}, \epsilon_{\ell-1}) + \beta\ \Phi_{\ell-1}\}$.

Giot [2000] a utilisé les modèles Log-ACD pour étudier le temps qu'il faut attendre pour qu'une quantité minimale soit échangée (« volume duration ») et la durée qui s'écoule jusqu'à la prochaine modification minimale des prix (« price duration ») pour IBM sur le NYSE.

2.6 Le modèle PHARMA

Il s'agit d'un modèle dans lequel on introduit une dépendance sérielle des durées étudiées et une structure à hasards proportionnels. La version présentée ici est paramétrique mais Gerhard et Hautsch [2000] considèrent une version semi-paramétrique dans laquelle la fonction de hasard de base n'est pas spécifiée et peut être estimée.

Considérons, tout d'abord, la fonction de hasard de la durée étudiée :

$$h(u_\ell \mid z_\ell) = h_0(u_\ell)\exp(-z'_\ell\gamma),$$

et

$$h_0(u_\ell) = \beta\ u_\ell^{\beta-1},$$

où $\beta > 0$, γ est un vecteur de paramètres et z_ℓ un vecteur de variables explicatives calculées à la date $\tau_{\ell-1}$.

Kiefer (1988) a montré que l'on peut obtenir une représentation linéaire de ce modèle :

$$\tilde{u}_\ell = -\ln(\Lambda_0(u_\ell)) = z'_\ell\gamma + v_\ell, \tag{3}$$

où $\Lambda_0(u) = \int_0^u h_0(s)\ ds$ est la fonction de hasard de base intégrée et v_ℓ est distribué selon une loi extreme value de type 1.

Pour introduire une structure autorégressive, Gerhard et Hautsch [2000] notent qu'il suffit que la distribution conditionnelle de $\tilde{u}_\ell$ soit de type extreme value.

Aussi, il considèrent le modèle de durée correspondant à l'équation (3) dans lequel

$$v_\ell = \sum_{k=1}^{p} a_k v_{\ell-k} + \sum_{k=1}^{q} b_k(\epsilon_{\ell-k} + \xi'_{\ell-k}\alpha) + \xi'_\ell\alpha + \epsilon_\ell$$

où a_j et b_j sont des paramètres. ξ_ℓ est un vecteur de variables explicatives à la date $\tau_{\ell-1}$. α est un vecteur de paramètres.

Dans le cas de la version semi-paramétrique du modèle, Gerhard et Hautsch [2000] proposent de discrétiser la durée et de maximiser la vraisemblance conditionnelle correspondante en estimant les variables latentes du modèle à partir de résidus généralisés [*cf.* Gouriéroux, Monfort et Trognon, (1985)].

2.7 Le modèle FIACD

Le modèle FIACD(p,d,q) a été introduit par Jasiak [1996] dans le but d'étudier la dépendance à long terme des durées entre transactions.
Le modèle :

$$\begin{cases} U_\ell = \psi_\ell \, \epsilon_\ell, \\ \epsilon_\ell \text{ i.i.d. de densité } f_0(\epsilon;\theta_1), \\ (1-\gamma(L)) \, (1-L)^d \, u_\ell = \omega + (1-\beta(L)) \, v_\ell, \end{cases}$$

où f_0 est la densité d'une variable positive d'espérance 1, L est l'opérateur retard, $0 < d < 1$, $v_\ell = u_\ell - \psi_\ell$, $\beta(L) = \beta_1 L + \ldots + \beta_p L^p$, $\alpha(L) = \alpha_1 L + \ldots + \alpha_q L^q$, $\gamma(L) = \alpha(L) + \beta(L)$ et

$$(1-L)^d = 1 + \sum_{k=1}^{\infty} \frac{d(d-1)\ldots(d-k+1)}{k!}(-1)^k L^k.$$

On supposera que $\beta_k \geqslant 0$, $\alpha_k \geqslant 0$ et $\sum_{k=1}^{q} \gamma_k < 1$.
Jasiak [1996] estime ce modèle sur Alcatel à la bourse de Paris. Elle trouve une persistance importante des durées entre transactions (de 16 à 32 heures).

2.8 Terme non observable

Dans le modèle SCD proposé par Bauwens et Veredas [1999], les auteurs introduisent un terme latent stochastique dans la distribution conditionnelle des durées (entre transactions, par exemple). La modélisation conditionnelle des durées sachant le terme non observable est identique à celle sur laquelle repose un modèle Log-ACD. La variable non observable est distribuée selon une loi normale.
Le modèle :

$$\begin{cases} u_\ell &= e^{\Phi_\ell} \, \epsilon_\ell, \\ \Phi_\ell &= \omega + \nu_\ell + \beta \Phi_{\ell-1}. \end{cases}$$

où les variables ϵ_ℓ sont i.i.d. et ont une distribution à support sur $\mathbb{R}^+$ de densité de probabilité $f(\epsilon;\theta_1)$. θ_1 est un vecteur de paramètres.

Conditionnellement à l'ensemble d'information à la date $\tau_{\ell-1}$, la distribution de ϵ_ℓ à un support positif (ne dépendant pas des paramètres), ν_ℓ sachant $I_{\ell-1}$ est distribué selon une loi normale centrée et de variance σ^2, ν_ℓ est indépendant de $\epsilon_{\ell'}$ $(\forall \ell, \ell')$.

On suppose que $\mid \beta \mid < 1$ (*cf.* section 2.5). Bauwens et Veredas supposent que la valeur initiale Φ_1 est tirée dans la distribution stationnaire de Φ_ℓ.

La densité de $U_1^N = (U_1, U_2, \ldots, U_N)$ conditionnelle à Φ_1 à pour expression

$$f(u_1^{\mathrm{N}}|\Phi_1;\theta) = \int_{\mathbb{R}^{\mathrm{N}-1}} f(u_1|\Phi_1;\theta_1) \, f(u_2|\Phi_1, v_2^2;\theta_1) \ldots f(u_{\mathrm{N}}|\Phi_1, v_2^{\mathrm{N}};\theta_1) \, q(v_2^{\mathrm{N}};\sigma^2) \, dv_2^{\mathrm{N}},$$

où $v_2^\ell = (\nu_2, \ldots, \nu_\ell)$, $\ell = 2, \ldots, N$. $q(v_2^N; \sigma^2) = \prod_{j=2}^N p(\nu_j; \sigma^2)$ est la densité du terme non observé [13]. θ, θ_1 sont des vecteurs de paramètres et σ^2 est un scalaire positif.

Aussi, la fonction de vraisemblance, pour un jour donné, comporte une intégrale multiple. Ce modèle peut être estimé par maximisation de la vraisemblance simulée [*cf.* Kamionka, (1998)] ou par quasi maximum de vraisemblance en utilisant un filtre de Kalman [*cf.* Bauwens et Veredas, (1999)]. Bauwens et Veredas [1999] indiquent que ce modèle peut être estimé aussi par GMM et par d'autres techniques d'estimation par simulation (inférence indirecte, quasi-maximum de vraisemblance simulé, MCMC).

L'idée d'introduire un terme aléatoire non observable dans l'expression de la distribution conditionnelle de la durée a déjà été envisagé aussi par Ghysels, Gouriéroux et Jasiak [1997] et Meddahi, Renault et Werker [1998] .

2.9 Choix de modèle

Le processus ayant engendré les durées financières peut être représenté par un des modèles décrit plus haut (ACD, Log-ACD, PHARMA, T-ACD, FIACD, SCD). Bauwens *et alii* [2000] proposent de comparer ces modèles en utilisant les prévisions de densité qu'ils permettent d'obtenir.

Soit $\{f_\ell(u_\ell \mid F_{\ell-1})\}_{\ell=1}^L$ une séquence de prévisions de densités une période en avant et $\{f_\ell^0(u_\ell \mid F_{\ell-1})\}_{\ell=1}^L$ la séquence des densités qui correspondent au processus ayant engendré les données. $F_{\ell-1} = \{u_{\ell-1}, u_{\ell-2}, \ldots\}$ est l'ensemble d'information engendré par les durées.

On veut tester :

$$H_0 : \{f_\ell(u_\ell \mid F_{\ell-1})\}_{\ell=1}^L = \{f_\ell^0(u_\ell \mid F_{\ell-1})\}_{\ell=1}^L$$

Soit $\bar{Z}_\ell = \int_0^{U_\ell} f_\ell(s \mid F_{\ell-1})\, ds$. Sous H_0, les variables $\bar{Z}_\ell$ $(\ell = 1, \ldots, L)$ sont i.i.d. selon une loi uniforme sur $[0, 1]$.

On utilise la séquence observée $\{\bar{z}_\ell\}_{\ell=1}^L$ pour la comparer avec les caractéristiques, comme l'histogramme, d'une séquence qui serait issue d'un tirage i.i.d. $U(0, 1)$.

Par exemple, pour le modèle ACD, $f_\ell(u \mid F_{\ell-1}) = f_0(\frac{u}{\Phi_\ell}; \theta_1)\frac{1}{\Phi_\ell}$, où f_0 est la densité d'une loi de moyenne égale à 1.

Fernandes et Grammig [2000] considèrent des tests de spécification non paramétriques pour les modèles de durée conditionnels.

Ils veulent tester l'hypothèse

$$H_0 : \exists \theta \in \Theta \text{ tel que } f(u, \theta) = f^0(u),$$

où f^0 est la vraie densité ayant engendrée les données.

---

13. Pour la distribution de ϵ_ℓ Bauwens et Veredas considèrent la distribution Weibull $W(1, \gamma)$ et la distribution Gamma $G(\gamma, 1)$. $p(\nu, \sigma^2)$ est la fonction de densité de probabilité de la loi $N(0, \sigma^2)$.

Soit $\hat{f}^0(u)$ un estimateur non-paramétrique par la méthode du noyau de la densité $f^0(u)$.

Fernandes et Grammig [2000] proposent de tester H_0 en utilisant la statistique :

$$\Psi_{\hat{f}} = \frac{1}{N} \sum_{\ell=1}^{N} \mathbb{1}(u_\ell \in S) \; [f(u_\ell; \hat{\theta}) - \hat{f}^0(u_\ell)]^2,$$

où $\hat{\theta}$ est une estimation convergente de θ et S est un sous ensemble compact.

L'hypothèse nulle est rejetée lorsque Ψ_f est suffisamment grande. Fernandes et Grammig [2000] considèrent la distribution asymptotique de la statistique de test. On peut remarquer qu'un M-estimateur de θ peut être obtenu en minimisant cette fonction objectif par rapport à $\hat{\theta}$.

2.10 Transformation des durées

Lorsque l'on étudie le temps qui s'écoule pour que l'on ait une modification du prix, il peut s'avérer intéressant de ne retenir que les changements du prix d'une certaine importance. *A priori*, les estimations des paramètres du modèle vont dépendre du choix fait pour la valeur minimale de la modification de prix.

Lorsque l'on étudie la durée qu'il faut attendre pour que le prix se modifie d'au moins une certaine valeur (on la notera c_u), on peut mettre en relation l'intensité conditionnelle de cette durée et la volatilité conditionnelle instantanée [voir Engle et Russell (1998) ou Giot (2000)].

Ainsi, l'espérance de la volatilité conditionnelle par seconde est égale à

$$\sigma^2(t \mid I_{N(t)}) = \left(\frac{c_u}{p(t)} \right)^2 h_{c_u}(t - \tau_{N(t)} \mid I_{N(t)}),$$

où $p(t)$ est le prix, $N(t)$ le nombre de modifications du prix au moins égales à c_u et $h_{c_u}(u \mid I_{N(t)})$ le hasard conditionnel de la durée pour la modification du prix.

Chaque fois qu'un événement va modifier la fonction de hasard, la volatilité instantanée va changer. Dans la pratique, si les durées sont distribuées selon une loi de hasard conditionnel décroissant avec la durée, on va observer des pics de l'espérance de la volatilité chaque fois qu'une modification suffisante du prix est enregistrée.

De la même façon, on peut s'intéresser à modéliser le temps qu'il faut attendre pour qu'une quantité minimale soit échangée [*cf.* Gouriéroux *et alii*, (1999)] afin, par exemple, d'analyser la liquidité de l'actif étudié.

3. MODÉLISATION D'UNE CARACTÉRISTIQUE ET DE LA DURÉE

La caractéristique peut être discrète ou continue. Dans le cas discret, on peut chercher, par exemple, à modéliser la direction des ordres (achat, vente), la direction et l'importance de la variation des prix [Russel et Engle, (1998)], la direction et l'agressivité des ordres [Bisière et Kamionka (2000)]. Dans ce cas, on peut parler de modèle de transition. Dans le cas continu, il peut s'agir d'une modélisation jointe de la durée entre transactions et de la volatilité des prix [Engle (2000), Ghysels et Jasiak (1996), Grammig et Wellner (1999)]. Dans ce cas, on peut parler de processus ponctuel marqué.

3.1 Caractéristique discrète

La première étude des durées entre ordres et du type des ordres est due à Biais, Hillion et Spatt [1995]. Bisière et Kamionka [2000] ont proposé une modélisation jointe des durées et des types des ordres pour étudier les stratégies de placement des ordres à la bourse de Paris. Bauwens et Giot [1998] ont adapté le modèle ACD pour tenir compte du signe des changements de prix. Russell et Engle [1998] proposent une modélisation jointe la durée qui s'écoule entre deux transactions et de la catégorie de la variation du prix d'une transaction à l'autre. Ces catégories du prix sont construites en référence au signe et à l'importance de sa variation. Darolles *et alii* [2000] modélisent les probabilités de transition entre les catégories de la variation du prix et la distribution de la durée entre transactions conditionnellement à la catégorie de la variation du prix.

Considérons le cas de la bourse de Paris. Il s'agit d'un marché gouverné par les ordres, en continu et informatisé (voir Biais, Foucault, Hillion 97).

Lorsqu'un ordre d'achat est placé au mieux, il va attaquer le carnet d'ordre et provoquer une ou plusieurs transactions (*cf.* graphique 4). Il sera exécuté en totalité à moins que la quantité totale disponible de l'autre côté du carnet ne soit pas suffisante. P_A représente le meilleur prix à la vente et P_B le meilleur prix à l'achat.

Dans le cas d'un ordre d'achat au prix du marché, si la quantité spécifiée pour cet ordre est inférieure à la quantité totale disponible au meilleur prix de l'autre côté du carnet, la fourchette de prix n'est pas modifiée (graphique 6), sinon la fourchette de prix se déplace vers le haut (graphique 5). La quantité qui n'a pas été satisfaite est transformée en un ordre d'achat à prix limite pour le prix qui correspond à l'ancien meilleur prix de l'autre côté du carnet d'ordre (graphique 5 encore).

Dans le cas d'un ordre d'achat à prix limite pour un prix spécifié à l'intérieur de la fourchette de prix, il n'y a pas de transaction puisque cet ordre ne peut rencontrer une offre de vente pour un prix inférieur ou égal (graphique 7). La fourchette de prix se resserre par le bas (le dernier ordre placé gagne la priorité prix).

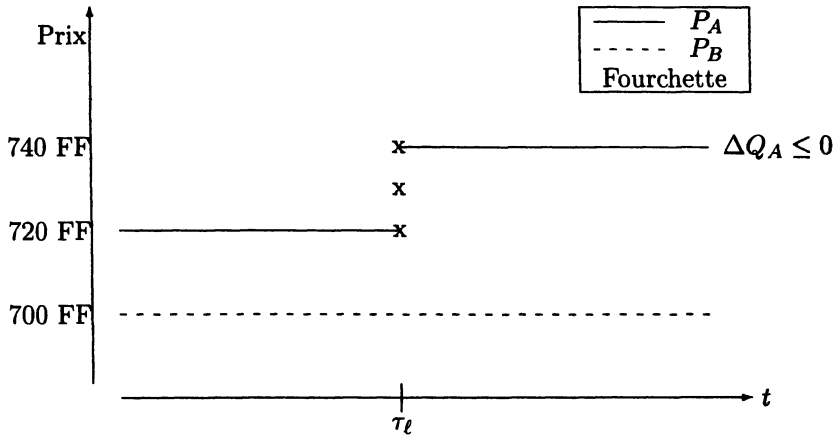

FIG 4. — Un achat au mieux

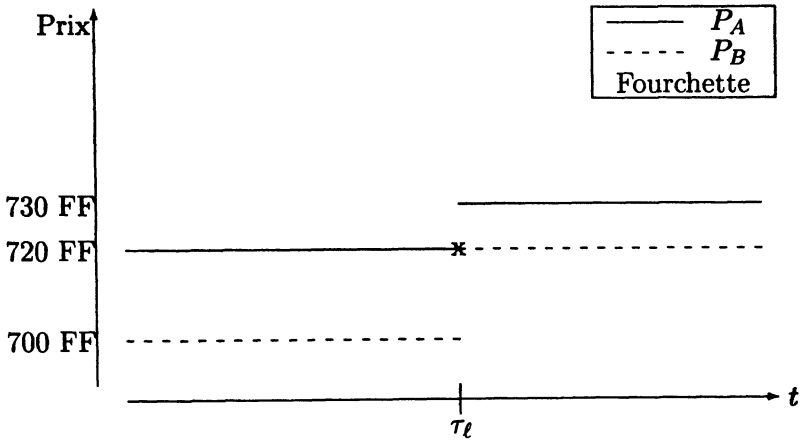

FIG 5. — Un achat au prix du marché (Quantité importante)

Si l'ordre qui est placé est un ordre d'achat situé sur la fourchette de prix, il n'y a pas de transaction (graphique 8). La fourchette de prix n'est pas modifiée. Seule la quantité totale disponible au meilleur prix à l'achat augmente de la quantité associée au dernier ordre d'achat ($\Delta Q_B > 0$). Cet ordre ne gagne pas la priorité temps (il sera exécuté en dernier parmi les ordres placés au meilleur prix à l'achat).

Lorsque l'ordre qui est placé est un ordre d'achat sous la fourchette il n'obtient ni la priorité prix, ni la priorité temps. La fourchette de prix n'est pas modifiée (graphique 9).

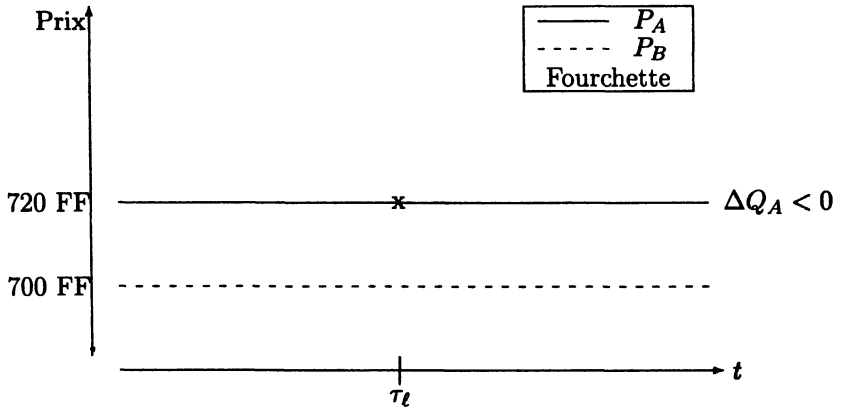

FIG 6. — Un achat au prix du marché (Quantité faible)

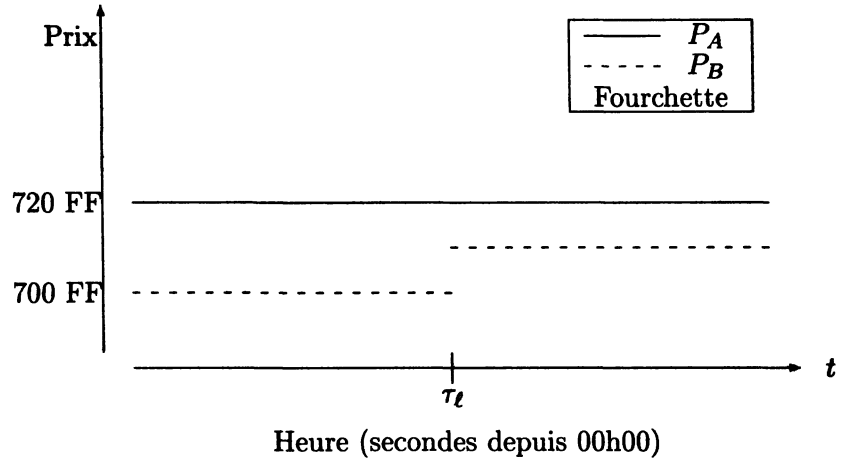

FIG 7. — Un achat à prix limité (dans la fourchette)

Des ordres à prix limites peuvent aussi être passés sur le NYSE [voir Hasbrouck, Sofianos et Sosebee, (1993)]. Harris et Hasbrouck, 1996, notent qu'environ 45 % des ordres passés par l'intermédiaire du système SuperDOT sur le NYSE sont des ordres à prix limites. Cependant, le fonctionnement du NYSE diffère de celui de la bourse de Paris notamment parce que des spécialistes interviennent pour assurer la liquidité [voir Madhavan et Sofianos, (1998)].

On remarquera que, dans ce cas, tous les ordres n'entraînent pas de transaction immédiate et que entre deux transactions successives peuvent s'intercaler des ordres à prix limites qui véhiculent une information sur la valeur de l'actif échangé.

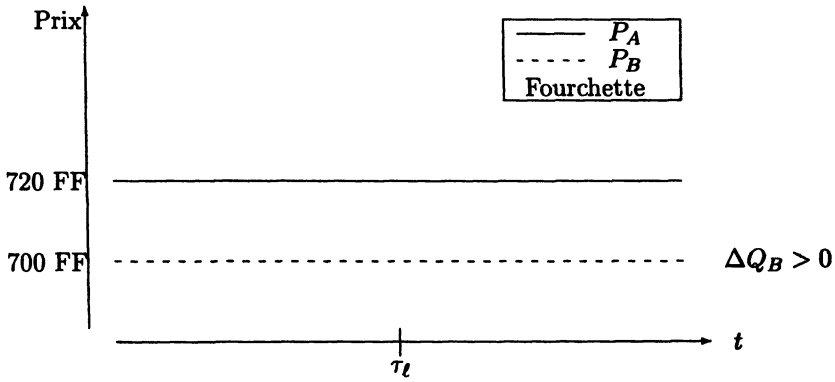

FIG 8. — Un achat à prix limité (sur la fourchette)

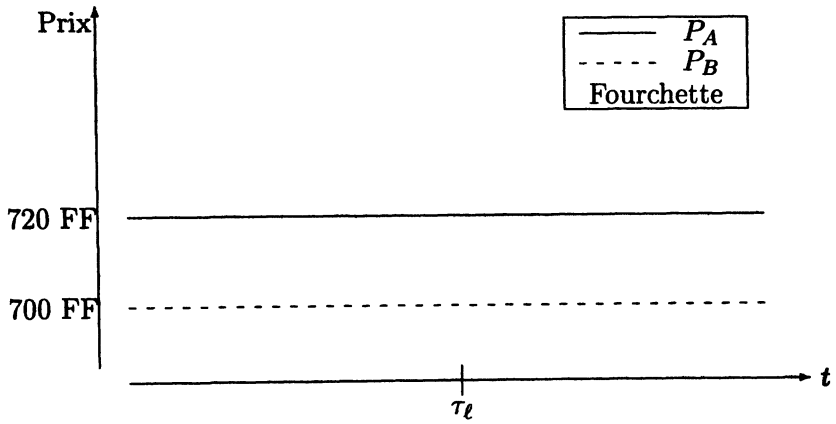

FIG 9. — Un achat à prix limité (sous la fourchette)

On peut ainsi utiliser l'information disponible sur les caractéristiques des ordres qui sont placés pour modéliser de manière jointe la durée entre les ordres et le type de l'ordre qui va suivre [*cf.* Bisière et Kamionka, (2000)]. Un modèle de transition peut aussi être utilisé pour modéliser et caractériser d'autres types de durées. Par exemple, Bauwens et Giot [1998], modélisent des durées entre cotations ainsi que le sens de variation de la cotation.

3.1.1 Le processus étudié

Dans ce qui suit, nous considérerons que nous étudions des durées entre ordres successifs. On va considérer K catégories d'ordres différentes [$K = 12$ pour Bisière et Kamionka (2000)].

Le processus étudié va être noté X_t ($t \in \mathbb{R}^+$) et il prend ses valeurs dans un espace d'états discret, noté E (aussi, $E = \{1, \ldots, K\}$). La réalisation du processus à l'instant t est notée x_t. x_t représente le type du dernier ordre arrivé lorsque l'on se place à l'instant t. Le processus X_t est constant entre deux transitions. Nous allons noter τ_0 l'instant où débute l'observation du processus et τ^* la date de fin de l'observation. Une réalisation du processus, pour une journée, est constituée de la séquence des catégories des ordres qui ont été successivement passés et des instants où ces ordres ont été placés sur le marché. Il s'agit d'un modèle de transition [*cf.* Flinn et Heckman (1983) ou Lancaster (1990)].

Considérons un exemple de réalisation du processus pour une journée (*cf.* graphique 10). On s'intéresse à la durée qui s'écoule entre deux ordres successifs. Le premier ordre est passé à l'instant τ_0, il s'agit d'un ordre de vente. Puis on observe un ordre de d'achat à l'instant τ_1. A la date τ_2, un ordre d'achat est encore passé. Enfin, à la date τ_3, un ordre de vente est placé et à l'instant de clôture on arrête l'observation (en τ^*).

Un épisode va représenter une période de temps délimitée par deux ordres successifs. Pendant un épisode, le processus garde la même valeur, celle qu'il a pris à la date du premier des deux ordres qui délimitent cet épisode. À chaque épisode est associée une durée : la longueur de l'épisode. L'épisode ℓ ($1 \leqslant \ell \leqslant n+1$) débute à la date $\tau_{\ell-1}$ et se termine à la date τ_ℓ ($\tau_\ell > \tau_{\ell-1}$). Soit u_ℓ la durée de l'épisode ℓ. On a que $u_\ell = \tau_\ell - \tau_{\ell-1}$ (*cf.* graphique 1).

La variable x_t est constante au cours de l'épisode ℓ. Pour simplifier, on posera $x_\ell = x_{\tau_\ell}$. On va supposer, aussi, que les variables explicatives sont constantes au cours de chaque épisode (elles sont regroupées dans le vecteur z). n est le nombre de transitions observées entre les dates τ_0 et τ^* pour une journée fixée. $n+1$ est le nombre d'épisodes observés sur cette même période d'observation. τ_ℓ est la date d'arrivée de la transition ℓ ($\ell \geqslant 1$). $x_{n+1} = 0$ pour indiquer la censure due à la clôture en fin de journée.

Le dernier épisode (l'épisode $n+1$) est censuré à droite. On sait que la réalisation de la durée U_{n+1} est au moins égale à $\tau^* - \tau_n$ (*cf.* graphique 10).

Considérons une réalisation du processus entre les dates τ_0 et τ^* :

$$r(\tau_0, \tau^*) = (y_0, y_1, \ldots, y_\ell, \ldots, y_n, y_{n+1}),$$

$$\text{où} \quad y_k = \begin{cases} (\tau_0', x_{\tau_0'}), & \text{si } k = 0, \\ (u_k, x_{\tau_k}), & \text{si } 1 \leqslant k \leqslant n, \\ (\tau^* - \tau_n, 0), & \text{si } k = n+1, \end{cases}$$

où τ_0' ($\tau_0' \geqslant \tau_0$) est la date où le premier ordre est placé au cours de la période d'observation.

La première composante de la réalisation (*i.e.* y_0) contient l'information relative à la date τ_0' de placement du premier ordre enregistré et à la catégorie

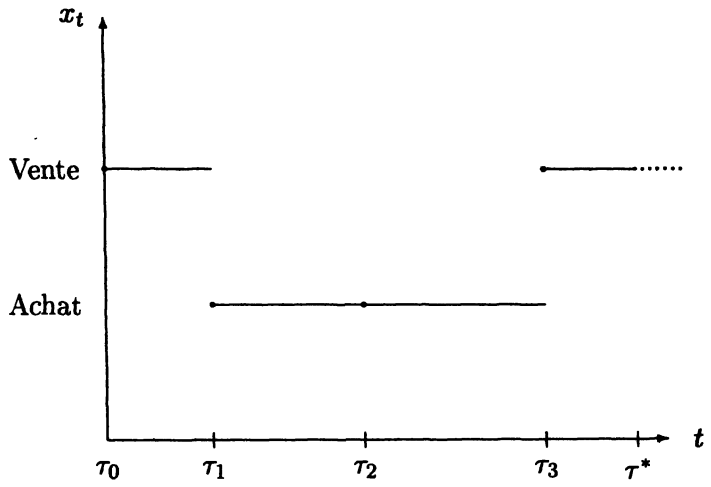

FIG 10. — Une réalisation du processus (durée entre ordres)

de ce premier ordre $x_{\tau_0'}$. La dernière composante de la réalisation y_{n+1} inclut l'information que la durée du dernier épisode est au moins égale à la valeur u_{n+1} $(u_{n+1}=\tau^* - \tau_n)$.

3.1.2. Découpage Marginal - Conditionnel

Dans ce qui suit, nous allons modéliser une réalisation du processus pour une journée et considérer que les réalisations pour les différentes journées sont indépendantes.

Il nous faut écrire la densité de la réalisation journalière $(y_0, y_1, \ldots, y_{n+1})$ conditionnellement à la première observation y_0 et au vecteur de variables explicatives z. Nous allons modéliser chaque réalisation en décomposant sa densité comme un produit de densités marginales et de densités conditionnelles. Soit $f(y_1, \ldots, y_{n+1} \mid y_0, z; \theta)$ cette densité.

En appliquant un raisonnement par récurrence on peut montrer que

$$f(y_1, \ldots, y_{n+1} | y_0, z; \theta) = \prod_{k=1}^{n+1} f(y_k | y_0, \ldots, y_{k-1}, z; \theta),$$

où $f(y_k | y_0, y_1, \ldots, y_{k-1}, z; \theta)$ est la densité de Y_k sachant les valeurs prises par $Y_0, \ldots, Y_{k-1}$ et Z.

Dans ce qui suit nous allons supposer, en particulier, que la distribution de la variable y_ℓ ne dépend pas de l'ordre d'arrivée de l'épisode mais seulement de l'histoire passée du processus et des variables explicatives. Dans ces variables explicatives ont peut avoir, par exemple, les quantités totales disponibles aux bornes du carnet d'ordres à la date $\tau_{\ell-1}$.

3.1.3. Modélisation au sein d'un épisode

On cherche à écrire la loi de $y_\ell = (u_\ell, x_{\tau_\ell})$ conditionnellement à l'histoire passée du processus $(y_0, \ldots, y_{\ell-1})$ et aux variables explicatives.

Soit $U^*_{\ell,k}$ la durée qu'il faudrait attendre pour que l'on observe le placement d'un ordre appartenant à la catégorie k en fin de l'épisode ℓ. On supposera que ces durées latentes sont indépendantes.

Soit $g_k(u \mid r(\tau_0, \tau_{\ell-1}), z; \theta)$ la fonction de densité de la durée latente $U^*_{\ell,k}$.

La durée de séjour dans l'épisode ℓ est simplement le minimum de ces durées latentes, soit

$$U_\ell = \inf_{k \in E} U^*_{\ell,k}.$$

Soit $h_k(u \mid y_0, \ldots, y_{\ell-1}, z; \theta)$ la fonction de hasard associée à la durée latente $U^*_{\ell,k}$ et $S_k(u \mid y_0, \ldots, y_{\ell-1}, z; \theta)$ la fonction de survie correspondante.

La densité de probabilité de la durée de séjour dans l'épisode ℓ lorsqu'il y a sortie vers l'état k sachant l'histoire du processus jusqu'à la date $\tau_{\ell-1}$ d'entrée dans l'épisode ℓ et $Z{=}z$ est égale à

$$f(u, k \mid y_0, \ldots, y_{\ell-1}, z; \theta) = h_k\ (u \mid y_0, \ldots, y_{\ell-1}, z; \theta) \prod_{k=1}^{K} S_k(u \mid y_0, \ldots, y_{\ell-1}, z; \theta).$$

3.1.4 Fonction de vraisemblance

La contribution à la fonction de vraisemblance conditionnelle d'une réalisation journalière donnée va s'écrire

$$L(\theta) = \prod_{\ell=1}^{n+1} f(y_\ell \mid y_0, y_1, \ldots, y_{\ell-1}, z; \theta),$$

où $y_\ell = (u_\ell, x_{\tau_\ell})$.

Aussi, une contribution journalière à la fonction de vraisemblance est donnée par l'expression (le dernier épisode est censuré)

$$L(\theta) = \prod_{\ell=1}^{n+1} h_{x_{\tau_\ell}}(u_\ell \mid y_0, \ldots, y_{\ell-1}, z; \theta)^{\delta_\ell} \prod_{k=1}^{K} S_k(u_\ell \mid y_0, \ldots, y_{\ell-1}, z; \theta),$$

où u_ℓ est la longueur observée, pour la journée considérée, de l'épisode ℓ et

$$\delta_\ell = \begin{cases} 1, \text{ si } \ell \leqslant n, \\ 0, \text{ sinon,} \end{cases}$$

est un indicateur de censure.

On peut encore réécrire la contribution journalière à la fonction de vraisemblance comme suit

$$L(\theta) = \prod_{k=1}^{K} L_k(\theta),$$

où $L_k(\theta) = \prod_{\ell=1}^{n+1} h_k(u_\ell \mid y_0, \ldots, y_{\ell-1}, z; \theta)^{\delta_{\ell,k}} \; S_k(u_\ell \mid y_0, \ldots, y_{\ell-1}, z; \theta)$ et

$$\delta_{\ell,k} = \begin{cases} 1, \text{ si } x_{\tau_\ell} = k, \\ 0, \text{ sinon,} \end{cases}$$

$\ell = 1, \ldots, n{+}1$.

Dans le cas où la fonction de hasard $h_k(u \mid y_0, \ldots, y_{\ell-1}, z; \theta)$, $k = 1, \ldots, K$, ne dépend que d'un sous vecteur de paramètres θ_k et si ces paramètres n'apparaissent pas dans les autres fonctions de risques, alors il est équivalent de maximiser la fonction de vraisemblance par rapport à l'ensemble des paramètres ou de maximiser chaque partie de la vraisemblance relative à un état par rapport au sous vecteur de paramètres correspondant ($\theta{=}(\theta_1, \ldots, \theta_K)$). Cette remarque est utile lorsque le modèle comporte un grand nombre de paramètres.

3.1.5 Spécification des fonctions de risque

On doit choisir une spécification paramétrique pour la distribution des durées latentes de séjour dans un épisode. Cette distribution peut alternativement être caractérisée par sa fonction de risque, sa fonction de densité de probabilité, sa fonction de répartition ou sa fonction de survie. Nous présenterons ici une spécification de type Singh-Maddala car elle est suffisamment générale.

La fonction de risque associée à l'épisode ℓ a pour expression :

$$h_k(u_\ell \mid y_0, \ldots, y_{\ell-1}, z; \theta) = \frac{\beta_{x_{\ell-1},k,\ell} \;\; \alpha_{x_{\ell-1},k} \;\; u_\ell^{\alpha_{x_{\ell-1},k}-1}}{(1 + \gamma_{x_{\ell-1},k} \;\; u_\ell^{\alpha_{x_{\ell-1},k}})},$$

où $\alpha_{x_{\ell-1},k}$ et $\gamma_{x_{\ell-1},k}$ sont des paramètres réels ($\alpha_{x_{\ell-1},k} > 0$, $\gamma_{x_{\ell-1},k} > 0$). $x_\ell = x_{\tau_\ell}$ est la catégorie enregistrée à la date τ_ℓ. $\beta_{j,k,\ell}$ est une fonction positive du vecteur des variables explicatives $z_{\tau_{\ell-1}}$. On supposera, par exemple, que $\beta_{j,k,\ell} = \exp(z_{\tau_{\ell-1}} \, \lambda_{j,k})$, où $z_{\tau_{\ell-1}}$ représente l'état du marché à l'instant $\tau_{\ell-1}$ et $\lambda_{x_{\ell-1},k}$ est un vecteur de paramètres.

3.1.6 Le Modèle ACD asymétrique

Ce modèle a été introduit pour la première fois par Bauwens et Giot, [1998]. Ils modélisent la durée qui sépare deux cotations successives et le sens de variation de la cotation moyenne.

Considérons les couples

$$(u_\ell, x_\ell), \; \ell = 1, \ldots, n,$$

où u_ℓ est la durée qui sépare la cotation $\ell - 1$ de la cotation ℓ et x_ℓ est le sens de variation de la cotation ℓ,

$$x_\ell = \begin{cases} 1, & \text{si la cotation moyenne augmente,} \\ 2, & \text{si la cotation moyenne diminue.} \end{cases}$$

Il s'agit d'un modèle de transition.

On va supposer que l'on a deux durées latentes : l'une pour l'arrivée d'une augmentation de la cotation moyenne et l'autre pour l'arrivée d'une diminution de cette cotation.

Soit $h^0_{\ell,j}$ la fonction de hazard de la durée latente pour l'arrivée d'une cotation de catégorie j au cours de l'épisode ℓ ($j \in \{1,2\}$, et $h^0_{\ell,j} \in \mathbb{R}^{+*}$).

La durée que l'on va observer entre les cotations $\ell - 1$ et ℓ est le minimum de deux durées latentes (une pour chaque catégorie envisageable pour la cotation ℓ). Soit $U^*_{\ell,j}$ la durée latente pour l'arrivée d'une cotation de catégorie j. On va considérer, ici, le modèle :

$$\begin{cases} U^*_{\ell,j} = (h^0_{\ell,j})^{-1}\ \epsilon_{\ell,j}, \\ (h^0_{\ell,j})^{-1} = \omega_{x_{\ell-1},j} + \alpha_{x_{\ell-1},j}\ u_{\ell-1} + \beta_{x_{\ell-1},j}\ (h^0_{\ell-1,j})^{-1}, \end{cases}$$

où $x_{\ell-1}$ est la catégorie de la cotation $\ell - 1$ et les $\epsilon_{\ell,j}$ sont des variables indépendantes et distribuées selon des lois exponentielles d'espérance 1. On a donc un modèle ACD pour chaque durée latente.

La contribution d'une observation journalière à la vraisemblance a pour expression :

$$L(\theta) = \prod_{\ell=1}^{n} h^0_{\ell,x_\ell}\ \exp(-(h^0_{\ell,1} + h^0_{\ell,2})\ u_\ell),$$

où x_ℓ est le type de la cotation ℓ. On suppose ici qu'il n'y a pas de censure à droite : on a omit la dernière observation de la journée.

On peut déduire la probabilité de l'arrivée d'une cotation de catégorie j au cours de l'épisode ℓ (à la suite d'une cotation de catégorie $x_{\ell-1}$) :

$$\pi_{\ell,j} = \frac{h^0_{\ell,j}}{h^0_{\ell,1} + h^0_{\ell,2}},$$

où $j = 1, 2$.

On peut aussi considérer, par exemple, que les variables $\epsilon_{\ell,j}$ sont indépendantes et distribuées selon une loi $Weibull(1, \gamma_j)$. On parlera, alors, d'un modèle WACD asymétrique.

3.1.7. Le Modèle Log-ACD asymétrique

La durée latente pour qu'il y ait une cotation de type j placée ($j \in \{1,2\}$) est donnée par l'expression :

$$U^*_{\ell,j} = e^{\Phi_{\ell,j}}\ \epsilon_{\ell,j},$$

où les variables $\epsilon_{\ell,j}$ sont i.i.d. suivant une $Weibull(1,\gamma_j)$ et [14]

$$\Phi_{\ell,j} = \begin{cases} \omega_1 + \alpha_1\ u_{\ell-1} + \beta_1 \Phi_{\ell-1,j}, & \text{si } j = 1, \\ \omega_2 + \alpha_2\ u_{\ell-1} + \beta_2 \Phi_{\ell-1,j}, & \text{si } j = 2. \end{cases}$$

La durée observée pour l'épisode ℓ est le minimum des réalisations des deux variables latentes indépendantes $U^*_{\ell,1}$ et $U^*_{\ell,2}$. Il s'agit d'un modèle de transition.

Soit $h_{x_{\ell-1},j}(u)$ la fonction de hasard pour l'arrivée d'une cotation appartenant à la catégorie j ($j \in \{1,2\}$). On obtient que

$$h_{x_{\ell-1},j}(u) = \gamma_j\ \exp(\Phi_{\ell,j})^{-\gamma_j}\ u^{\gamma_j - 1},$$

où $j \in \{1,2\}$.

La contribution de l'observation journalière à la vraisemblance :

$$L(\theta) = \prod_{\ell=1}^{n} h_{x_{\ell-1},x_\ell}(u_\ell)\ \exp\{-\int_0^{u_\ell} [h_{x_{\ell-1},1}(s) + h_{x_{\ell-1},2}(s)]\ ds\ \},$$

où $x_{\ell-1}$ est le type de la cotation $\ell-1$ et x_ℓ est la catégorie de la cotation ℓ. On suppose ici qu'il n'y a pas de censure à droite.

Dans ce modèle, le logarithme de l'espérance de la durée latente à une forme linéaire.

3.1.8 Le modèle ACM

Russel et Engle [1998] , considèrent un modèle à cinq états selon le sens et l'importance de la variation du prix (importante et négative, faible et négative, nulle, faible et positive, importante et positive). Il s'agit d'un modèle de transition. Prigent, Renault et Scaillet [1999] envisagent une version à deux états de ce modèle (modèle ACB) dans laquelle l'état sera égal à 1 lorsque le prix augmente de plus de x % (x fixé) et l'état sera égal à -1 lorsque le prix baisse de plus de x %.

La densité de la durée et du type de la variation des prix :

$$f(u,k \mid I_{\ell-1};\theta) = p_k(u, I_{\ell-1};\theta)\, f(u \mid I_{\ell-1};\theta),$$

où p_k est la probabilité conditionnelle que la variation du prix appartienne à la catégorie k lorsque la transaction a lieu à la date $\tau_{\ell-1} + u$, $f(u \mid I_{\ell-1};\theta)$

14. La spécification de $\Phi_{\ell,j}$ introduite par Bauwens et Giot [1997] est plus générale.

est la densité de probabilité de la distribution de la durée entre transactions conditionnelle à l'ensemble d'information de la date $\tau_{\ell-1}$.

Pour la distribution marginale de la durée entre transactions, Engle et Russel [1998] considèrent une modèle ACD (*cf.* section 2).

Soit $\bar{p}_\ell$ le vecteur des probabilités conditionnelles de transition à la date τ_ℓ. On a que $\bar{p}_\ell = (p_1(u_\ell, I_{\ell-1}, \theta), \ldots, p_{K-1}(u_\ell, I_{\ell-1}, \theta))'$. On fait l'hypothèse que :

$$g(\bar{p}_\ell) = \sum_{i=1}^{p} A_i M_{\ell-i}(v_{\ell-i} - \bar{p}_{\ell-i}) + \sum_{i=1}^{q} B_i v_{\ell-i} + \sum_{i=1}^{r} C_i g(\bar{p}_{\ell-i}) + G z_\ell \quad (4)$$

g est une fonction[15] de $[0,1]^{K-1} \longrightarrow \mathbb{R}^{K-1}$. La composante k de v_ℓ ($1 \leqslant k \leqslant K-1$) est égale à 1 si la variation du prix est de type k à la date de transaction τ_ℓ (elle vaut 0 sinon). A_ℓ, B_ℓ, C_ℓ et M_ℓ sont des matrices carrées ($K-1$ lignes). G est une matrice comportant $K-1$ lignes et s colonnes. z_ℓ est un vecteur colonne de s éléments. La matrice M_ℓ peut être prise égale à la matrice identité ou encore à la puissance -1/2 de la matrice des variances-covariances conditionnelles associée au vecteur v_ℓ.

Comme la distribution de la catégorie de la variation du prix à la date τ_ℓ dépend de ses valeurs passées et des réalisations passées des variations du prix, Engle et Russel ont baptisé ce modèle « Autoregressive Conditional Multinomial model » (modèle ACM(p,q,r)).

Etant donné le nombre important de paramètres dans le modèle, Russel et Engle [1998] proposent de considérer des matrices A_ℓ, B_ℓ et C_ℓ symétriques. Ils estiment ce modèle sur des durées issues de transactions sur IBM (données TORQ).

3.1.9 Modélisation alternative au sein d'un épisode

Une formalisation alternative consisterait à modéliser la probabilité de voir apparaître la catégorie k conditionnellement au passé et la distribution de la durée de l'épisode sachant le passé du processus et la catégorie k. C'est ce type d'approche qui a été adopté par Darolles *et alii* [2000] pour modéliser la catégorie du mouvement du prix de transaction et la durée qui s'écoule entre deux transactions. Le prix de transaction ne peut varier que d'un tick au maximum par rapport au dernier prix enregistré. Pour ce faire, les auteurs distinguent les transactions qui ont été initiées par un ordre d'achat de celles qui sont issues d'un ordre de vente. Ils sont finalement amenés à considérer trois distributions de la durée entre transactions selon le sens de la variation du prix de transaction.

On considère, dans ce cas, un modèle à trois états : $x_{\tau_\ell} \in \{-1; 0; 1\}$ (on posera $x_{\tau_0} = 0$). $\sum_{k=1}^{\ell} x_{\tau_k}$ représente la variation du prix depuis l'ouverture (calculée à la date τ_ℓ). Soit $y_\ell = (u_\ell, x_{\tau_\ell})$, le couple à modéliser pour l'épisode ℓ. Un épisode représente une période de temps qui s'écoule entre deux transactions successives initiées, par exemple, par des ordres de vente.

---

15. Par exemple, g est telle que sa composante k est égale à $g_k(\bar{p}_\ell) = ln(p_k(u_\ell, I_{\ell-1}; \theta) / P_K(u_\ell, I_{\ell-1}; \theta))$.

La densité conditionnelle de Y_ℓ va s'écrire

$$f(y_\ell \mid y_0, y_1, \ldots, y_{\ell-1}; \theta) = p(x_{\tau_\ell} \mid \sum_{k=1}^{\ell-1} x_{\tau_k})\ g(u_\ell \mid \sum_{k=1}^{\ell-1} x_{\tau_k}; \sum_{k=1}^{\ell} x_{\tau_k}),$$

où $p(x_{\tau_\ell} \mid \sum_{k=1}^{\ell-1} x_{\tau_k})$ est la probabilité conditionnelle de transition vers la catégorie x_{τ_ℓ}. $g(u \mid \sum_{k=1}^{\ell-1} x_{\tau_k}; \sum_{k=1}^{\ell} x_{\tau_k})$ est la densité de la durée de l'épisode ℓ conditionnelle à la dernière variation du prix et à la suivante.

Darolles *et alii* [2000] supposent que la probabilité de transition entre catégories de mouvement de prix ne dépend de la dernière réalisation de la catégorie qu'à travers la somme de toutes les réalisations passées des types de mouvement de prix (c'est-à-dire la variation du prix entre τ_0 et $\tau_{\ell-1}$). De plus, ces auteurs supposent que la distribution conditionnelle de la durée de l'épisode ℓ ne dépend des catégories passées du mouvement du prix et de la suivante qu'à travers les variations du prix depuis l'ouverture calculées aux dates $\tau_{\ell-1}$ et τ_ℓ.

LEMME : *Soit U une variable aléatoire continue sur $\mathbb{R}^+$. Soit $f(u)$ sa fonction de densité de probabilité ($u \in \mathbb{R}^+$). Alors,*

$$S(u) = \text{Prob}[U \geqslant u] = \exp\left\{-\int_0^u h(s)\, ds\right\},$$

où $h(u) = f(u) / \int_u^{+\infty} f(s)\, ds$ est la fonction de hasard associée à U.

Démonstration : voir Gouriéroux [1989], propriété 11.2.

PROPRIÉTÉ 1 : *Soit X une variable aléatoire discrète à K modalités. La loi conditionnelle de X a pour probabilité $p_x(y_0, \ldots, y_{\ell-1})$ (où $x \in E = \{1, \ldots, K\}$). Soit U une variable aléatoire positive et continue. La loi conditionnelle de U admet pour densité de probabilité $g(u \mid y_0, \ldots, y_{\ell-1}, x)$. On notera $h(u \mid y_0, \ldots, y_{\ell-1}, x)$, la fonction de hasard correspondante. Alors,*

$$h(u \mid y_0, \ldots, y_{\ell-1}) = \sum_{j \in E} h_j(u \mid y_0, \ldots, y_{\ell-1}),$$

$$g(u \mid y_0, \ldots, y_{\ell-1}) = h(u \mid y_0, \ldots, y_{\ell-1}) \exp\left\{-\int_0^u h(s \mid y_0, \ldots, y_{\ell-1})\, ds\right\},$$

$$p_x(y_0, \ldots, y_{\ell-1}, u) = \frac{h_x(u \mid y_0, \ldots, y_{\ell-1})}{h(u \mid y_0, \ldots, y_{\ell-1})},$$

$$\text{où } h_x(u \mid y_0, \ldots, y_{\ell-1}) = \frac{p_x(y_0, \ldots, y_{\ell-1})\ g(u \mid y_0, \ldots, y_{\ell-1}, x)}{\sum_{k \in E} \int_u^{+\infty} p_k(y_0, \ldots, y_{\ell-1})\ g(s \mid y_0, \ldots, y_{\ell-1}, k)\, ds}.$$

Démonstration : voir Annexe.

PROPRIÉTÉ 2 : *Soit U une variable aléatoire positive et continue. La loi de U conditionnelle à $Y_0, \ldots, Y_{\ell-1}$ admet pour densité de probabilité $g(u \mid y_0, \ldots, y_{\ell-1})$. On notera $h(u \mid y_0, \ldots, y_{\ell-1})$, la fonction de hasard correspondante. Soit X une variable aléatoire discrète à K modalités. La loi de X conditionnelle à $Y_0, \ldots, Y_{\ell-1}$ et U admet pour probabilité $p_x(y_0, \ldots, y_{\ell-1}, u)$ (où $x \in E = \{1, \ldots, K\}$). Alors,*

$$f(u, x|y_0, \ldots, y_{\ell-1}) = g(u|y_0, \ldots, y_{\ell-1})\ p_x(y_0, \ldots, y_{\ell-1}, u),$$

$$= h_x(u|y_0, \ldots, y_{\ell-1}) \exp\{-\int_0^u \sum_{k \in E} h_k(s|y_0, \ldots, y_{\ell-1})\ ds\},$$

où $h_k(u|y_0, \ldots, y_{\ell-1}) = h(u|y_0, \ldots, y_{\ell-1})\ p_x(y_0, \ldots, y_{\ell-1}, u)$.

Démonstration : voir Annexe.

De la propriété 1 ou 2, on déduit que :

$$f(u, x|y_0, \ldots, y_{\ell-1}) = h_x(u|y_0, \ldots, y_{\ell-1})\ S(u \mid y_0, \ldots, y_{\ell-1}),$$

$$= f_x(u|y_0, \ldots, y_{\ell-1}) \prod_{\substack{k \in E \\ k \neq x}} S_k(u|y_0, \ldots, y_{\ell-1}),$$

où $S_k(u|y_0, \ldots, y_{\ell-1}) = \exp\{-\int_0^u h_k(s|y_0, \ldots, y_{\ell-1})\ ds\}$ et $f_k(u|y_0, \ldots, y_{\ell-1}) = h_k(u|y_0, \ldots, y_{\ell-1})\ S_k(u|y_0, \ldots, y_{\ell-1})$.

Dans tous les cas, le couple (u, x) peut être vu comme le résultat du tirage de K durées latentes indépendantes de fonctions de hasard $h_k(u|y_0, \ldots, y_{\ell-1})$ $(k \in E)$. La valeur du minimum de ces K variables latentes va déterminer la valeur de la durée u de l'épisode et, l'indice de la variable latente qui correspond à ce minimum, la catégorie x.

Darolles *et alii* [2000] considèrent des probabilités $p_x(y_0, \ldots, y_{\ell-1})$ qui ont une forme logistique et, pour la densité $g(u|y_0, \ldots, y_{\ell-1}, x)$, celle d'une loi exponentielle. Aussi, la fonction de hasard correspondante $h(u \mid y_0, \ldots, y_{\ell-1}, x)$ est constante par rapport à la durée (pas de dépendance temporelle). On pourra vérifier, dans ce cas et par application de la propriété 1, que la fonction de hasard non conditionnelle à l'état futur $g(u \mid y_0, \ldots, y_{\ell-1})$ est, *a priori*, non monotone en u même si $h(u \mid y_0, \ldots, y_{\ell-1}, x)$ est constante.

On constatera, aussi, que lorsque Darolles *et alii* [2000] supposent que les distributions considérées ne dépendent du passé qu'à travers la somme des catégories observées (*i.e.*, la valeur de la variation du prix depuis l'ouverture), cette hypothèse va se retrouver dans les expressions des fonctions de hasard des durées latentes.

3.2 Caractéristique continue et durée

On étudie les liens existants entre la durée qui sépare deux transactions et la volatilité du prix (ou de son logarithme). Ainsi la durée [*cf.* Engle, (2000)] ou l'espérance de la durée (*cf.* Ghysels et Jasiak [1997] , Grammig et Wellner [1999]) peuvent avoir une influence sur la volatilité conditionnelle. De même, la volatilité passée et le carré de la variation passée du prix peuvent avoir un impact sur l'espérance de la durée entre transactions (Grammig et Wellner [1999]).

3.2.1 Le modèle UHF-GARCH

En ce qui concerne la distribution conditionnelle de la durée, on va poser

$$\begin{cases} U_\ell = \psi_\ell \, \epsilon_\ell, \\ \epsilon_\ell \sim \text{ i.i.d. de densité } f_0(\epsilon; \theta_g), \\ \psi_\ell = \omega^u + \alpha^u \, u_{\ell-1} + \beta^u \psi_{\ell-1}, \end{cases} \tag{5}$$

où $\psi_\ell = E[U_\ell \mid u_1, \ldots, u_{\ell-1}]$. ω^u, α^u et β^u sont des paramètres. θ_g est le vecteur des paramètres de la densité f_0 de ϵ_ℓ.

Soit y_{τ_ℓ} le logarithme de la moyenne des cotations (à la vente et à l'achat) à la date de transaction τ_ℓ et $r_\ell = \Delta y_{\tau_\ell} = \log(p(\tau_\ell)) - \log(p(\tau_{\ell-1}))$.

On va supposer que [16]

$$\frac{r_\ell}{\sqrt{u_\ell}} = a_0 \, u_\ell + a_1 \frac{r_{\ell-1}}{\sqrt{u_{\ell-1}}} + e_\ell + a_2 \, e_{\ell-1},$$

$$V\left(\frac{r_\ell}{\sqrt{u_\ell}} \mid I_{\ell-1}\right) = \sigma_\ell^2 = c_0 + c_1 \, e_{\ell-1}^2 + c_2 \, \sigma_{\ell-1}^2$$

où $c_0 > 0$, $c_1 \geqslant 0$ et $c_2 \geqslant 0$ et e_ℓ est une innovation.

La variation de la cotation sur la racine carrée de la durée entre transaction suit un ARMA(1,1). Les paramètres de la variance conditionnelle de $\frac{r_\ell}{\sqrt{u_\ell}}$ peuvent être estimés comme ceux d'un modèle GARCH(1,1).

3.2.2. Le modèle ACD-GARCH

Ghysels et Jasiak [1997] considèrent un modèle dans lequel les durées entre transactions sont représentées par un modèle ACD (*cf.* système (5)).

Soit $x_\ell \equiv x_{\tau_\ell} = \log(p(\tau_\ell))$, le log du prix de transaction enregistré à la date τ_ℓ et $r_\ell = x_\ell - x_{\ell-1}$. Ghysels et Jasiak [1997] indiquent que l'espérance conditionnelle de la durée peut dépendre aussi de variables pertinentes relativement aux théories de la microstructure des marchés financiers.

Dans le modèle ACD-GARCH, la volatilité conditionnelle va vérifier [17] :

$$\sigma_\ell^2 = \omega_\ell^r + \alpha_\ell^r \, r_{\ell-1}^2 + \beta_\ell^r \sigma_{\ell-1}^2.$$

16. Engle [2000] suppose que la première différence r_ℓ par unité de temps suit un ARMA(1,1) pour tenir compte du caractère discret du prix observé.

17. Ghysels et Jasiak [1997] envisagent aussi le cas où la moyenne conditionnelle de r_ℓ incorpore une constante et des variables indicatrices pour tenir compte d'effets saisonniers.

Les paramètres ω_ℓ^r, α_ℓ^r et β_ℓ^r sont obtenus à partir des paramètres ω^r, α^r, β^r et κ en utilisant la formule d'agrégation temporelle de Drost et Nijman [1993] pour un Garch(1,1) faible (*cf.* Annexe). Cette expression fait aussi intervenir l'espérance conditionnelle de la durée entre transactions. ω^r, α^r, β^r caractérisent la volatilité conditionnelle du processus sous jacent.

Exemple : Un modèle à durée et volatilité interdépendantes

Grammig et Wellner [1999] considèrent un modèle dans lequel les durées entre transactions sont représentées par un modèle ACD. L'espérance conditionnelle de la durée va dépendre, en particulier, de la volatilité passée. Soit le modèle

$$\begin{cases} U_\ell = \psi_\ell \, \epsilon_\ell, \\ \epsilon_\ell \sim \text{ i.i.d. de densité } f_0(\epsilon;\theta_g), \\ \psi_\ell = \omega^u + \alpha_1^u u_{\ell-1} + ... + \alpha_p^u u_{\ell-p} + \beta_1^u \psi_{\ell-1} + ... + \beta_q^u \psi_{\ell-q} + \xi_1 \sigma_{\ell-1}^2 + \xi_2 \dfrac{r_{\ell-1}^2}{\sigma_{\ell-1}^2}, \end{cases}$$

où $\psi_\ell = E[U_\ell \mid u_1, \ldots, u_{\ell-1}]$ et $r_\ell = p(\tau_\ell) - p(\tau_{\ell-1})$. ω^u, α_k^u $(k = 1, ..., p)$, β_j^u $(j = 1, ..., q)$, ξ_1 et ξ_2 sont des paramètres.

Dans ce modèle, la volatilité conditionnelle vérifie :

$$\sigma_\ell^2 = \omega_\ell^r + \alpha_\ell^r \, r_{\ell-1}^2 + \beta_\ell^r \sigma_{\ell-1}^2$$

Ils supposent que le kurtosis κ est fixé. Les expressions des paramètres ω_ℓ^r, α_ℓ^r et β_ℓ^r sont données par la formule d'agrégation temporelle (voir Annexe) à partir des paramètres ω^r, α^r et β^r.

3.3. Caractéristiques multiples et durée

Dufour et Engle [2000] proposent un modèle dans lequel les durées entre transactions suivent un processus ACD et, conditionnellement à ces durées, les changements dans les cotations et les caractéristiques des transactions suivent une modélisation VAR (auto-regressive et vectorielle). Le modèle proposé par Dufour et Engle [2000] est construit sur la base d'un modèle VAR introduit par Hasbrouck [1991].

Considérons, ici, le modèle suivant :

Soit τ_ℓ la date de la transaction ℓ (en secondes). Soit U_ℓ la durée qui sépare les transactions $\ell - 1$ et ℓ[18]. On supposera que $\epsilon_\ell = (\dfrac{U_\ell}{\psi_\ell})^\theta$ suit une loi exponentielle de paramètre 1 et que

$$\psi_\ell = w + \sum_{k=1}^{p} \alpha_k \, u_{\ell-k} + \sum_{k=1}^{q} \beta_k \, \psi_{\ell-k},$$

où ω, α_k et β_k sont des paramètres.

18. Dufour et Engle [2000] ajoutent une seconde à cette durée pour que son logarithme soit au moins égal à 0.

Soit q_ℓ la cotation moyenne juste avant la transaction ℓ et $r_\ell = (\log(q_{\ell+1}) - \log(q_\ell))$. Soit $v_\ell = 1$ si la transaction ℓ est initiée par un ordre d'achat ($p_\ell > q_\ell$), $v_\ell = -1$ si la transaction a été initiée par un ordre de vente ($p_\ell < q_\ell$) et $v_\ell = 0$ sinon ($p_\ell = q_\ell$), p_ℓ étant le prix de transaction. Soit $\bar{p}_\ell$ la probabilité conditionnelle de transition associée à la variable donnant le sens de la transaction. Soit g une fonction de $[0,1]^2 \longrightarrow \mathbb{R}^2$ qui donne la dynamique de cette probabilité conditionnelle.

L'expression de la probabilité $\bar{p}_\ell$ est similaire à celle donnée par l'équation (4) dans le cas où le nombre de modalités K est égal à 3 :

$$g(\bar{p}_\ell) = \sum_{k=1}^{p'} a_k (v^0_{\ell-k} - \bar{p}_{\ell-k}) + \sum_{k=1}^{q'} b_k \, v^0_{\ell-k} + \sum_{k=1}^{r'} c_k \, g(\bar{p}_{\ell-k}),$$

où a_k, b_k et c_k sont des paramètres réels. La composante 1 du vecteur v^0_ℓ est égale à 1 si $v_\ell = 1$, 0 sinon. La composante 2 du vecteur v^0_ℓ vaut 1 lorsque $v_\ell = -1$ et 0 sinon.

La révision des cotations est donnée par l'équation :

$$r_\ell = \sum_{k=1}^{s''} a'_k \, r_{\ell-k} + \sum_{k=0}^{s''} b'_k \, v_{\ell-k} + \sum_{k=0}^{s''} c'_k \, v_{\ell-k} \ln(u_{\ell-k}) + \nu_\ell,$$

où les ν_ℓ sont centrés et non corrélés entre eux. a'_k, b'_k et c'_k sont des paramètres réels.

Dans ce modèle, on modélise la durée entre transactions, une caractéristique discrète (le sens de l'ordre à l'origine de la transaction) et une caractéristique continue (la variation de la cotation moyenne). Par rapport à Dufour et Engle [2000] , nous avons modifié l'équation gouvernant la probabilité conditionnelle de transition. De plus, Dufour et Engel [2000] introduisent, par l'intermédiaire de variables indicatrices, des corrections pour la première observation de chaque journée. On travaille sur les durées entre transactions une fois enlevé le facteur journalier (*cf.* section 2.1). Dufour et Engle [2000] utilisent le fichier TORQ sur le New York Stock Exchange en retenant 18 titres parmi les plus échangés pour étudier l'effet de la durée entre transactions sur le mécanisme de formation du prix.

4. UN MODÈLE DE DURÉES BIVARIÉ À ÉPISODES MULTIPLES

4.1. Introduction

On va observer deux durées. Par exemple, les durées entre transactions et les durées entre cotations.

Soit τ'_ℓ la date d'arrivée de la transaction ℓ et τ''_ℓ la date de survenance de la modification ℓ des cotations. En fait, les dates d'arrivée des événements pour

BIBLIOTHÈQUE DÉPARTEMENT DE MATHÉMATIQUES

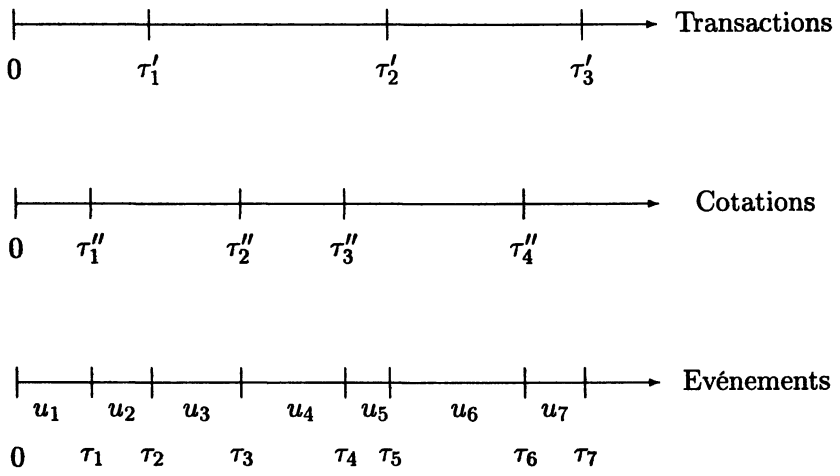

FIG 11. — Processus Multivarié

les deux processus ponctuels peuvent ne pas être les mêmes. Sur une période de temps donnée, le nombre de transactions ou de cotations reçues peuvent être différents (*cf.* graphique 11).

Soit v_ℓ la durée qui s'écoule entre la transaction $\ell-1$ et la transaction ℓ. Soit w_ℓ, la durée entre la cotation $\ell-1$ et la cotation ℓ. On a les relations suivantes (*cf.* graphique 11) :

$$v_\ell = \tau'_\ell - \tau'_{\ell-1},$$

pour les durées entre transactions et

$$w_\ell = \tau''_\ell - \tau''_{\ell-1},$$

pour les cotations.

Les temps de départ pour ces deux durées ne sont pas les mêmes. Aussi, Engle et Lunde [1999] proposent de se ramener à des temps de démarrage identiques en considérant le temps qui s'écoule jusqu'à la date d'arrivée de la première modification des cotations après chaque transaction. Ce qui revient à modéliser des «temps en avant» pour les cotations. Une autre façon de procéder consiste à définir les durées qui séparent deux événements successifs : transactions ou modifications des cotations.

4.2. Modélisation

Soient τ_ℓ ($\ell = 1, \ldots, n$) les dates d'arrivée des événements (modification des cotations ou transactions), $I_{\ell-1}$ l'ensemble d'information à la date $\tau_{\ell-1}$ et $u_\ell = \tau_\ell - \tau_{\ell-1}$ la durée de l'épisode ℓ.

Soient

$$h_1(u \mid I_{\ell-1}; \theta) = \varphi_{\ell-1}^{-1},$$

et

$$h_0(u \mid I_{\ell-1}; \theta) = \psi_{\ell-1}^{-1},$$

où $\tau_{\ell-1} < \tau_{\ell-1}+u \leqslant \tau_\ell$. h_1 est la fonction de hasard pour qu'une modification des cotations survienne et h_0 est la fonction de hasard pour la durée jusqu'à la prochaine transaction.

On passe ainsi de deux séquences de durées à une seule séquence de durées marquées.

On suppose ici que ces fonctions de hasard sont constantes par intervalle (sur chaque épisode). On va considérer que les espérances conditionnelles des durées latentes correspondantes ont pour expressions :

$$\varphi_{\ell+1} = \exp\{\omega_1 + \alpha_1 \ln(\varphi_\ell) + \beta_{1,1}\frac{u_\ell}{\psi_\ell}\,(1-d_\ell) + \beta_{2,1}\frac{u_\ell}{\varphi_\ell}\,d_\ell + \gamma_1 z_\ell\},$$

et

$$\psi_{\ell+1} = \exp\{\omega_0 + \alpha_0 \ln(\psi_\ell) + \beta_{1,0}\frac{u_\ell}{\psi_\ell}\,(1-d_\ell) + \beta_{2,0}\frac{u_\ell}{\varphi_\ell}\,d_\ell + \gamma_0 z_\ell\},$$

où

$$d_\ell = \begin{cases} 1, & \text{si en } \tau_\ell \text{ on a observé une modification des cotations,} \\ 0, & \text{sinon,} \end{cases}$$

z_ℓ est un vecteur de caractéristiques du marché observées à la date τ_ℓ. On suppose que l'ensemble d'information ne change pas entre deux dates successives d'arrivée des événements.

La contribution d'une observation journalière à la vraisemblance va s'écrire :

$$L(\theta) = \prod_{\ell=1}^{n} h_{d_\ell}(u_\ell \mid I_{\ell-1};\theta)\, \exp\{-u_\ell \sum_{k\in\{0,1\}} h_k(u_\ell \mid I_{\ell-1};\theta)\}$$
$$\exp\{-u_{n+1} \sum_{k\in\{0,1\}} h_k(u_{n+1} \mid I_n;\theta)\}, \tag{6}$$

où n est le nombre d'événements considérés. Le dernier épisode est censuré à droite.

Cette contribution à la fonction de vraisemblance peut se réécrire :

$$L(\theta) = \prod_{\ell=1}^{n+1} h_0(u_\ell \mid I_{\ell-1};\theta)^{\delta_{0,\ell}}\, \exp\{-u_\ell\, h_0(u_\ell \mid I_{\ell-1};\theta)\}$$
$$\prod_{\ell=1}^{n+1} h_1(u_\ell \mid I_{\ell-1};\theta)^{\delta_{1,\ell}}\, \exp\{-u_\ell\, h_1(u_\ell \mid I_{\ell-1};\theta)\},$$

où

$$\delta_{1,\ell} = \begin{cases} 1, & \text{si en } \tau_\ell \text{ est observée une modification des cotations,} \\ 0, & \text{sinon,} \end{cases}$$

et

$$\delta_{0,\ell} = \begin{cases} 1, & \text{si en } \tau_\ell \text{ est observée une transaction,} \\ 0, & \text{sinon.} \end{cases}$$

La contribution à la vraisemblance se réécrit comme une produit des densités des durées entre transactions successives et des densités des durées entre modifications successives des cotations. En particulier, cette écriture de la contribution à la vraisemblance des durées entre transactions intègre les arrivées d'événements intermédiaires susceptibles d'avoir un impact sur la distribution de la durée entre transactions.

Ceci peut être généralisé en considérant des fonctions de hasard qui ne sont pas constantes sur chaque épisode.

4.3 Le modèle ACI

Dans le modèle à intensités conditionnelles autoregressives (ACI) proposé par Russell [1999] , on considère un processus ponctuel multivarié. Ici, nous décrivons le cas d'un processus bivarié. Soit $N_k(t)$ le nombre d'événements relatifs au processus k qui sont survenus à la date t ($t \in \mathbb{R}^+$, $k = 1, 2$). Soient τ_ℓ^k, $\ell = 0, \ldots, N_k(t)$, les dates d'arrivée des événements relatifs au processus k ($\tau_\ell^k > \tau_{\ell-1}^k$) et τ_ℓ celles relatives à l'ensemble des événements ($\ell = 0, \ldots, N(t)$). Soit F_t la filtration du processus joint. On supposera que $N_k(t)$ est adapté à F_t. Soit $\Phi_{N(t)}^k$ une fonction mesurable. Soit $z_{N_k(t)}^k$, un vecteur de variables explicatives. Soit x_{τ_ℓ} le type d'événement survenu à la date τ_ℓ ($x_{\tau_\ell} \in \{1, 2\}$). Russell [1999] modélise le processus conditionnellement aux marques. De plus, $z_{N_k(t)}^k$ ne peux pas varier entre deux événements successifs relatifs au processus étudié. On posera $z_{N(t)} = z_{N_1(t)}^1$ si le dernier événement survenu appartient à la catégorie 1 et $z_{N(t)} = z_{N_2(t)}^2$, sinon.

La fonction d'intensité conditionnelle relative au processus k aura pour expression

$$h_k(t \mid F_t) = \exp\{\Phi_{N(t)}^k\}\ h_k(t - \tau_{N_1(t)}^1, t - \tau_{N_2(t)}^2),$$

où

$$\Phi_{N(t)} = \begin{pmatrix} \Phi_{N(t)}^1 \\ \Phi_{N(t)}^2 \end{pmatrix} = \begin{cases} \epsilon_{N(t)-1}\ A_1 + B_1\ \Phi_{N(t)-1} + z'_{N_1(t)-1}\ C_1, & \text{si } x_{\tau_{N(t)-1}} = 1, \\ \epsilon_{N(t)-1}\ A_2 + B_2\ \Phi_{N(t)-1} + z'_{N_2(t)-1}\ C_2, & \text{si } x_{\tau_{N(t)-1}} = 2, \end{cases}$$

$$\epsilon_{N(t)} = \begin{cases} 1 - \displaystyle\int_{\tau_{N_1(t)-1}^1}^{\tau_{N_1(t)}^1} h_1(s \mid F_s)\ ds, & \text{si } x_{\tau_{N(t)}} = 1, \\ 1 - \displaystyle\int_{\tau_{N_2(t)-1}^2}^{\tau_{N_2(t)}^2} h_2(s \mid F_s)\ ds, & \text{si } x_{\tau_{N(t)}} = 2. \end{cases}$$

A_k est un vecteur colonne à deux éléments et B_k est une matrice carrée (deux lignes, deux colonnes). C_k est un vecteur colonne ($C_k \in \mathbb{R}^s$).

Les fonctions h_k ($h_k > 0$) permettent aux intensités de varier entre deux événements successifs. En particulier, h_k peut être constante entre deux événements successifs ou, encore, elle peut ne dépendre que de $t - \tau^k_{N_k(t)}$.

L'intérêt de ce modèle est qu'il permet d'étudier les relations de causalité entre les deux processus (causalité dans un sens ou dans les deux sens, absence de causalité entre les processus).

Supposons que l'observation ait lieu, pour une journée, de la date 0 (ouverture du marché) à la date T (heure de clôture).

La contribution à la vraisemblance va s'écrire

$$L = \prod_{j=1}^{N_1(T)} h_1(\tau^1_j \mid F_{\tau^1_j}) \prod_{k=1}^{N_2(T)} h_2(\tau^2_k \mid F_{\tau^2_k}) \exp\{-\sum_{k=1}^{2} \int_0^T h_k(s \mid F_s)\, ds\},$$

On retrouve une contribution à la vraisemblance dont l'expression est similaire aux expressions précédentes (*cf.*, par exemple, équation (6)). Ce modèle se généralise au cas où la fonction d'intensité conditionnelle dépend de plus de valeurs retardées de Φ_k ($k = N(t) - 1, \ldots, N(t) - p$) et de ϵ_j ($j = N(t)-1, \ldots, N(t)-q$). On parlera, alors, de modèle ACI(p,q). De plus, on peut facilement généraliser ce modèle au cas multivarié de dimension supérieure à 2.

5. MODÉLISATION DES DURÉES D'EXÉCUTION DES ORDRES LIMITES

Dans le premier exemple de trajectoire du processus pour une journée (*cf.* graphique 12), on peut avoir des ordres de vente ou d'achat à prix limite et on s'intéresse à la durée jusqu'à ce qu'ils soient exécutés. Ces ordres peuvent ne pas être exécutés immédiatement (*cf.* section 3.1). Le premier ordre est passé à l'instant τ_1, il s'agit d'un ordre de vente. Il est annulé à la date τ_2. A la date τ_3, un ordre d'achat est placé. A la date τ_4, un ordre de vente est passé et à l'instant de clôture on arrête l'observation.

Dans le second exemple de trajectoire du processus (*cf.* graphique 13), le premier ordre est placé à l'instant τ_1, il s'agit d'un ordre de vente. Il est exécuté à la date τ_4. A la date τ_2, un ordre d'achat est placé ($\tau_2 > \tau_1$). A la date τ_3, un ordre de vente est encore placé.

Ici, le modèle peut être considéré comme un modèle de durées multivarié en ce sens qu'à un instant donné plusieurs durées peuvent ne pas être encore complétées et la réalisation d'une de ces durées va avoir un impact sur les hasards conditionnels des autres durées.

Les dates τ_ℓ ($\ell = 1, \ldots, n$) sont définies comme les instants auxquels arrivent un événement concernant le titre (le placement d'un ordre à prix limite, l'exécution d'un ordre à prix limite, le placement d'un ordre à exécution immédiate ou une annulation).

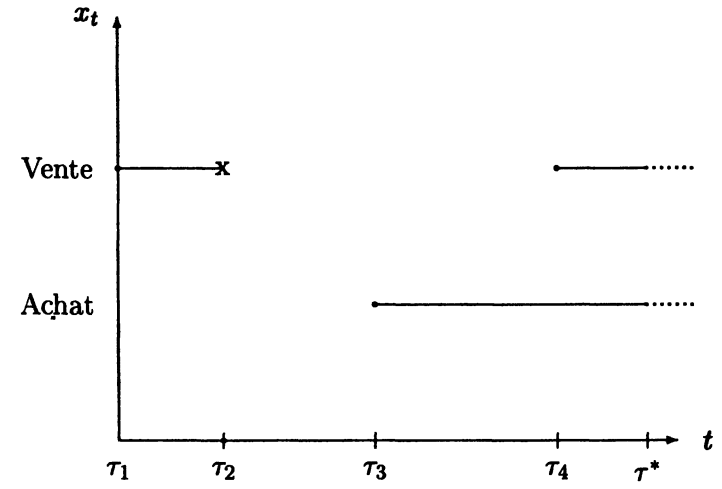

FIG 12. — Une réalisation du processus (durée jusqu'à exécution)

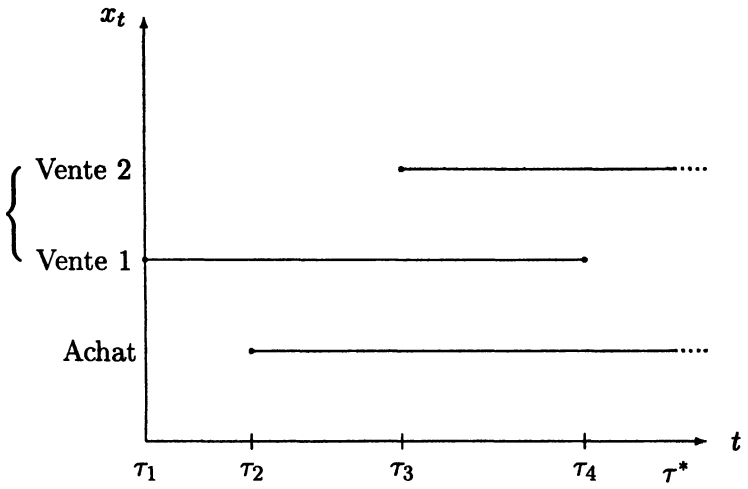

FIG 13. — Une réalisation du processus (durée jusqu'à exécution)

Soit t_k la date de placement de l'ordre k sur le marché ($k = 1, \ldots, \bar{n}$) et t'_k la date d'exécution de l'ordre k.

On supposera que l'ensemble d'information ne peut changer qu'aux dates τ_ℓ, $\ell = 1, \ldots, n$. Soit $I(t^-)$ l'ensemble d'information juste avant la date t.

$h(u \mid I((t+u)^-); z; \theta)$ est la fonction de hasard pour la durée d'exécution d'un ordre à prix limite lorsque cet ordre a été placé à la date t et qu'il s'est écoulé depuis u unités de temps. z est un vecteur de caractéristiques de l'ordre.

On supposera que l'expression de la fonction de hasard $h(u|I((t+u)^-); z; \theta)$ à la date $t+u$ ne dépend du carnet d'ordre qu'à travers sa composition à la date du dernier événement qui est survenu (placement d'un ordre, exécution ou annulation) juste avant la date $t+u$. Soit $\bar{u}_k = t'_k - t_k$ la durée d'exécution de l'ordre k.

La contribution à la vraisemblance d'une observation journalière a pour expression

$$L(\theta) = \prod_{k=1}^{\bar{n}} h(\bar{u}_k \mid I(t'^-_k); z_k; \theta)^{\delta_k} \exp\{-\int_0^{\bar{u}_k} h(s \mid I((t_k+s)^-); z_k; \theta)\, ds\},$$

où

$$\delta_k = \begin{cases} 1, & \text{si la durée d'exécution de l'ordre } k \text{ n'est pas censurée,} \\ 0, & \text{sinon.} \end{cases}$$

On peut avoir plusieurs ordres qui ne sont pas exécutés ou qui ont été annulés.

Si la fonction de hasard conditionnelle est constante entre deux événements, on a

$$L(\theta) = \prod_{k=1}^{\bar{n}} h(\bar{u}_k \mid I(t'^-_k); z_k; \theta)^{\delta_k} \exp\{-\sum_{\ell\,:\, t_k < \tau_\ell \leqslant t'_k} u_\ell \;\; h(u_\ell \mid I(\tau_\ell^-); z_k; \theta)\},$$

où $\bar{u}_k = t'_k - t_k$ et $u_\ell = \tau_\ell - \tau_{\ell-1}$.

Lorsqu'il y a annulation d'un ordre à prix limite présent dans le carnet, cette annulation va affecter la fonction de hasard. De même, si un ordre intermédiaire est placé, comme au graphique 13, en τ_2, cela va affecter l'expression de la fonction de hasard h. Ainsi, si un ordre d'achat a été placé à l'intérieur de la fourchette à cette date, cette modification de l'ensemble d'information va affecter les valeurs des fonctions de hasard à partir de la date τ_2 car des agents peuvent réagir au placement de cet ordre. Cela peut accélérer l'arrivée d'un ordre de vente au prix du marché car il devient relativement moins coûteux de consommer la liquidité.

Lo, MacKinlay et Zhang [1997] étudient la durée d'exécution, partielle ou totale, des ordres limites.

D'une façon générale, la prise en compte de la présence d'ordres limites, même lorsque l'on étudie des durées entre transactions, peut être importante. Plusieurs marchés sont gouvernés par les ordres. C'est le cas, par exemple, du NYSE avec intervention d'intermédiaires (les « market makers »), de la bourse de Paris, du TAIEX (Taiwan) [19]. La présence de ces ordres limites va avoir un impact sur la fonction de hasard de la durée entre transactions lorsque ces ordres s'intercalent entre deux transactions successives. Aussi, en général, la distribution du temps en avant pour la durée entre transactions est modifiée par l'arrivée d'un ordre limite.

19. Pour une typologie des procédures d'appariement des ordres, on pourra se reporter à Le Fol [1998] ou, encore, à Gouriéroux et Jasiak [2000] .

6. CONCLUSION

Dans cet article, nous modélisons des données haute fréquence pour lesquelles les instants d'arrivée des ordres peuvent être considérés comme aléatoires. La classe de modèle proposée par Engle [2000] et Engle et Russel [1998] repose sur des hypothèses fortes concernant la distribution des durées étudiées. En effet, un modèle ACD est un modèle de durée à vie accélérée dans lequel l'espérance conditionnelle a une forme autorégressive. Le modèle Log-ACD proposé par Bauwens et Giot [1997] ne nécessite pas d'estimer le modèle sous la contrainte que l'espérance conditionnelle soit positive. Dans le cadre de ces deux modèles, les durées ne sont pas indépendantes. Aussi, les propriétés des estimateurs (QML) doivent être déterminées en conséquence.

D'une façon générale, plusieurs auteurs [Biais, Hillion et Spatt (1995), Bisière et Kamionka (2000), Lo, MacKinlay et Zhang (1997)] ont souligné le rôle joué par les ordres à prix limites à la bourse de Paris ou encore sur le NYSE. Ainsi, chaque ordre, même lorsqu'il n'entraîne pas une exécution immédiate, va modifier la composition du carnet des ordres qui ont été placés et avoir un impact sur l'ensemble d'information des agents. Aussi, la fonction de hasard conditionnelle de la durée étudiée, puisque elle dépend de l'ensemble d'information, n'est pas connue parfaitement *ex ante*. En effet, la distribution de la durée entre deux événements – deux transactions, par exemple – n'aura pas la même expression selon qu'un ordre intermédiaire aura été placé ou pas. Lorsque l'on étudie les implications des théories de la microstructure des marchés financiers, on peut être amené à modéliser de manière jointe une durée et une caractéristique de cette durée. La caractéristique étudiée peut être discrète [*cf.* Bisière et Kamionka (2000), Russell et Engle (1998), Coppejans et Domowitz (1999)]. Cette caractéristique peut être continue [voir Engle (2000)] ou être vectorielle [Dufour et Engle (2000)]. En général, l'estimation de ces modèles ne peut être faite que de façon jointe. L'étude de ces mêmes implications stratégiques peut nous amener à considérer des modèle multivariés. Dans ce cas, il est intéressant de pouvoir faire dépendre les fonctions de hasard relatives aux durées de l'ensemble de l'information disponible à chaque instant et, en particulier, des réalisations passées de toutes les durées. Aussi, dans le cadre de cet article, nous proposons un moyen de traiter toutes les durées de façon symétrique.

ANNEXE

Démonstration de la propriété 1 :

$$p_x(y_0,\ldots,y_{\ell-1})\; g(u \mid y_0,\ldots,y_{\ell-1},x) = g(u \mid y_0,\ldots,y_{\ell-1})\; p_x(y_0,\ldots,y_{\ell-1},u),$$

$$\text{où } g(u \mid y_0,\ldots,y_{\ell-1}) = \sum_{k\in E} p_k(y_0,\ldots,y_{\ell-1})\; g(u \mid y_0,\ldots,y_{\ell-1},k),$$

$$S(u \mid y_0, \ldots, y_{\ell-1}) = \int_u^{+\infty} \sum_{k \in E} p_k(y_0, \ldots, y_{\ell-1})\ g(s \mid y_0, \ldots, y_{\ell-1}, k) ds \text{ et}$$

$$\begin{aligned} h(u \mid y_0, \ldots, y_{\ell-1}) &= \frac{\sum_{k \in E} p_k(y_0, \ldots, y_{\ell-1})\ g(u \mid y_0, \ldots, y_{\ell-1}, k)}{\int_u^{+\infty} \sum_{k \in E} p_k(y_0, \ldots, y_{\ell-1})\ g(s \mid y_0, \ldots, y_{\ell-1}, k)\ ds} \\ &= \sum_{k \in E} h_k(u \mid y_0, \ldots, y_{\ell-1}), \end{aligned}$$

où $h_j(u \mid y_0, \ldots, y_{\ell-1}) = \dfrac{p_j(y_0, \ldots, y_{\ell-1})\ g(u \mid y_0, \ldots, y_{\ell-1}, j)}{\int_u^{+\infty} \sum_{k \in E} p_k(y_0, \ldots, y_{\ell-1})\ g(s \mid y_0, \ldots, y_{\ell-1}, k)\ ds}$.

En utilisant le lemme, on obtient

$$g(u|y_0, \ldots, y_{\ell-1}) = \sum_{k \in E} h_k(u|y_0, \ldots, y_{\ell-1}) \exp\{-\int_0^u \sum_{k \in E} h_k(s|y_0, \ldots, y_{\ell-1})\ ds\},$$

De plus,

$$\begin{aligned} p_x(y_0, \ldots, y_{\ell-1}, u) &= \frac{p_x(y_0, \ldots, y_{\ell-1})\ g(u \mid y_0, \ldots, y_{\ell-1}, x)}{g(u \mid y_0, \ldots, y_{\ell-1})}, \\ &= \frac{p_x(y_0, \ldots, y_{\ell-1})\ g(u \mid y_0, \ldots, y_{\ell-1}, x)}{h(u \mid y_0, \ldots, y_{\ell-1}) \exp\{-\int_0^u h(s \mid y_0, \ldots, y_{\ell-1})\ ds\}}, \\ &= \frac{\dfrac{p_x(y_0, \ldots, y_{\ell-1})\ g(u \mid y_0, \ldots, y_{\ell-1}, x)}{\int_u^{+\infty} \sum_{k \in E} p_k(y_0, \ldots, y_{\ell-1})\ g(s \mid y_0, \ldots, y_{\ell-1}, k)\ ds}}{h(u \mid y_0, \ldots, y_{\ell-1})}, \\ &= \frac{h_x(u \mid y_0, \ldots, y_{\ell-1})}{h(u \mid y_0, \ldots, y_{\ell-1})}. \end{aligned}$$

Démonstration de la propriété 2 :

$$g(u \mid y_0, \ldots, y_{\ell-1}) = h(u \mid y_0, \ldots, y_{\ell-1}) \exp\{-\int_0^u h(s \mid y_0, \ldots, y_{\ell-1})\ ds\},$$

où $h(s \mid y_0, \ldots, y_{\ell-1})$ est la fonction de hasard conditionnelle (cf. Lemme).

$$\begin{aligned} f(u, x \mid y_0, \ldots, y_{\ell-1}) &= h(u|y_0, \ldots, y_{\ell-1}) \exp\{-\textstyle\int_0^u h(s|y_0, \ldots, y_{\ell-1})\, ds\} \\ &\quad p_x(y_0, \ldots, y_{\ell-1}, u), \\ &= h_x(u|y_0, \ldots, y_{\ell-1}) \exp\{-\textstyle\int_0^u h(s|y_0, \ldots, y_{\ell-1})\, ds\}, \end{aligned}$$

où $h_x(u \mid y_0, \ldots, y_{\ell-1}) = h(u \mid y_0, \ldots, y_{\ell-1})\, p_x(y_0, \ldots, y_{\ell-1}, u)$.

De plus, on a que

$$h(u \mid y_0, \ldots, y_{\ell-1}) = \sum_{k \in E} h_k(u \mid y_0, \ldots, y_{\ell-1}),$$

d'où

$$f(u, x|y_0, \ldots, y_{\ell-1}) = h_x(u|y_0, \ldots, y_{\ell-1}) \exp\{-\int_0^u \sum_{k \in E} h_k(s|y_0, \ldots, y_{\ell-1})\, ds\}.$$

Valeurs des paramètres d'un GARCH(1,1) faible d'après la formule d'agrégation temporelle de Drost et Nijman [1993] :

$$d = \psi_\ell,$$

$$\omega_\ell^r = d\, \omega^r\, \frac{1 - (\alpha^r + \beta^r)^d}{1 - (\alpha^r + \beta^r)},$$

$$\alpha_\ell^r = (\alpha^r + \beta^r)^d - \beta_\ell^r,$$

où β_ℓ^r ($| \beta_\ell^r | < 1$) est solution de

$$\frac{\beta_\ell^r}{(1 + \beta_\ell^{r\,2})} = \frac{a(\alpha^r, \beta^r, d, \kappa)(\alpha^r + \beta^r)^d - b(\alpha^r, \beta^r, d)}{a(\alpha^r, \beta^r, d, \kappa)[1 + (\alpha^r + \beta^r)^{2d}] - 2\, b(\alpha^r, \beta^r, d)},$$

$$\begin{aligned} a(\alpha^r, \beta^r, d, \kappa) &= d(1 - \beta^r)^2 + \frac{2d(d-1)(1 - \alpha^r - \beta^r)^2(1 - \beta^{r\,2} - 2\alpha^r\beta^r)}{(\kappa - 1)(1 - (\alpha^r + \beta^r)^2)} \\ &\quad + \frac{4\, c(\alpha^r, \beta^r, d)}{(1 - (\alpha^r + \beta^r)^2)}, \end{aligned}$$

$$b(\alpha^r, \beta^r, d) = [\alpha^r(1 - \beta^r(\alpha^r + \beta^r))][1 - (\alpha^r + \beta^r)^{2d}]/[1 - (\alpha^r + \beta^r)^2],$$

$$c(\alpha^r, \beta^r, d) = [d(1 - \alpha^r - \beta^r) - 1 + (\alpha^r + \beta^r)^d][\alpha^r(1 - \beta^r(\alpha^r + \beta^r))].$$

RÉFÉRENCES BIBLIOGRAPHIQUES

BAUWENS, L., et P. GIOT, [2000], « The Logarithmic ACD Models : An Application To The Bid/Ask Quote Process of Three NYSE Stocks », *Annales d'Economie et de Statistique*, 60, pp. 117–149.

BAUWENS, L., et P. GIOT, [1998], « Asymetric ACD Models : Introducing Price Information In ACD Models With A Two State Transition Model », *Core Discussion Paper* 9844, Core, Université Catholique de Louvain (Version révisée : Janvier 99).

BAUWENS, L., GIOT, P., GRAMMIG, J. et D. VEREDAS, [2000], « A Comparison of Financial Duration Models Via Density Forecast », *mimeo*, CORE, Louvain-La-Neuve.

BAUWENS, L., et D. VERADAS, [1999], « The Stochastic Conditional Duration Model : A Latent Factor Model For The Analysis Of Financial Durations », *Core Discussion Paper* 9958, Core, Université Catholique de Louvain.

BIAIS, B., FOUCAULT, T. et C. HILLION, [1997], *Microstructure des marchés financiers. Institutions, modèles et test empiriques*, Collection Finance, puf.

BIAIS, B., HILLION, P. et C. SPATT, [1995], « An Empirical Analysis Of The Limit Order Book And The Order Flow In The Paris Bourse », *Journal of Finance*, 50, pp. 1655–1689.

BISIèRE, C. et T. KAMIONKA, [2000], « Timing of Orders, Orders Aggressiveness and Order Book at the Paris Bourse », *Annales d'économie et de Statistique*, 60, pp. 43–72.

BOLLERSLEV, T. et J. WOLLDRIDGE, [1992], « Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances », *Econometric Reviews*, 11, pp. 143–172.

COPPEJANS, M. et I. DOMOWITZ, [1999], « Screen Information, Trader Activity and Bid-Ask Spreads in a Limit Order Market », *mimeo*, Duke University, Durham.

COX, D.R. et D. OAKES, [1984], *Analysis of Survival Data*, Chapman and Hall.

DAROLLES, S., GOURIéROUX C. et G. LE FOL [2000], « Intraday Transaction Price Dynamics », *Annales d'Economie et de Statistique*, 60, pp. 207–238.

DUFOUR, A. et R. F. ENGLE, [2000], « Time and Price Impact of a Trade », 55, *Journal of Finance*, pp. 2467–2498.

DROST, F. C. et T. E. NIJMAN, [1993], « Temporal Aggregation of Garch Processes », *Econometrica*, Vol. 61, pp. 909–927.

EASLEY, D. et M. O'HARA, [1992], « Time and the Process of Security Price Adjustment », *Journal of Finance*, pp. 577–605.

ENGLE, R. F., [2000], « The Econometrics Of Ultra-High Frequency Data », *Econometrica*, vol. 68, 1, pp. 1–22.

ENGLE, R. F. et A. LUNDE, [1999], « Trades and Quote : A Bivariate Point Process », *mimeo*, University of California, San Diego.

ENGLE, R. F., et J. R. RUSSELL, [1998], « Autoregressive Conditional Duration : A New Model for Irregularly Spaced Transaction Data », *Econometrica*, pp. 1127–1162.

FERNANDES, M., et J. GRAMMIG, [2000], « Non-parametric Specification Tests for Conditional Duration Models », *mimeo*, European University Institute et Frankfurt University.

FLINN, C. J., et J.J. HECKMAN, [1983], « The Likelihood Function for the Multistate-Multiepisode Model », In *Advances in Econometrics*, 3, Ed. R. Bassman et G. Rhodes, Greenwich, JAI Press, pp. 225–231.

GERHARD, F. et N. HAUTSCH, [2000], « Determinants of Inter-Trade Durations and Hazard Rates Using Proportional Hazard ARMA Models », *COFE Discussion Paper*, Université de Konstanz, Allemagne.

GHYSELS, E., GOURIÉROUX C. et J. JASIAK, [1997], « Stochastic Volatility Duration Models », *Cahier de Recherche* 9746, CREST-INSEE, Paris.

GHYSELS, E. et J. JASIAK, [1996], « GARCH for Irregularly Spaced Financial Data : The ACD-GARCH Model », *mimeo*, Cirano, Montréal.

GIOT, P., [2000], « Time Transformations, Intraday Data And Volatility Models », *mimeo*, version révisée du *cahier de recherche* du CORE 9944, Louvain-La-Neuve. A paraître dans *Journal of Computational Finance.*

GLOSTEN, L. R. et P. MILGROM, [1985], « Bid, Ask and Transaction Prices In A Specialist Market With Heterogeneously Informed Traders », *Journal of Financial Economics*, pp. 71–100.

GOURIÉROUX, C. , [1989], *Econométrie des variables qualitatives*, Economica, Paris.

GOURIÉROUX, C. et J. JASIAK, [2000], *Econometrics of Finance - Textbook*, *mimeo*, York University, Canada.

GOURIÉROUX, C., JASIAK J. et G. LE FOL, [1999], « Intra-Day Market Activity », *Journal of Financial Markets*, volume 2, issue 3, pp. 193–226.

GOURIÉROUX C. et G. LE FOL, [1998], « Effets des Modes de Négociation sur les échanges », *Revue Economique*, vol 49, 3, pp. 795–808.

GOURIÉROUX, C. et A. MONFORT, [1989], *Statistique et Modèles Econométriques*, Economica, Paris.

GOURIÉROUX, C., MONFORT A. et A. TROGNON, [1985], « A General Approach to Serial Correlation », *Econometric Theory*, pp. 315–340.

GRAMMIG J. et K.-O. MAURER, [1999], « Non-monotonic Hazard Functions and the Autoregressive Conditional Duration Model », *Humboldt-Universität zu Berlin Discussion paper* 50-1999, Université Humboldt, Berlin.

GRAMMIG J. et M. WELLNER, [1999], « Modeling the Interdependence of Volatility and Inter-transaction Duration Processes », *Humboldt-Universität zu Berlin Discussion paper* 21-1999, Université Humboldt, Berlin.

GREEN P. J. et B. W. SILVERMAN, [1994], *Nonparametric Regression and Generalized Linear Models. A Roughness Penalty Approach*, Chapman & Hall.

HARRIS, L. et J. HASBROUCK, [1996], « Market vs. Limit Orders : The SuperDOT Evidence on Order Submission Strategy », *Journal of Financial and Quantitative Analysis*, 31, pp. 213–231.

HASBROUCK, J. [1991], « Measuring the Information Content of Stock Trades », *The Journal of Finance*, Vol. XLVI, 1, pp. 179–207.

HASBROUCK, J., SOFIANOS, G. et D. SOSEBEE, [1993], « New York Stock Exchange Systems and Trading Procedures », *NYSE Working Paper 9301*, New York, NY.

JASIAK, J., [1996], « Persistence in Intertrade Durations », *Finance*, Vol. 19, pp. 166–195.

KAMIONKA, T., [1998], « Simulated Maximum Likelihood Estimation in Transition Models », *Econometrics Journal*, 1, pp. C129–153.

KIEFER, N.M., [1988], « Economic Duration Data and Hazard Functions », *Journal of Economic Literature*, 26, pp. 646–679.

KYLE, A., [1985], « Continuous Auction Markets and Insider Trading », *Econometrica*, pp. 1315–1335.

LANCASTER, T., [1990], *The Econometric Analysis of Transition Data*, Cambridge University Press.

LE FOL, G., [1998], Microstructure et données haute fréquence : une étude du marché français des actions, thèse de doctorat, Université de Paris I - Panthéon - Sorbonne.

LO, A. W., MACKINLAY, A. C. et J. ZHANG, [1997], « Econometric Models of Limit-Order Executions », *NBER Working Paper* 6257, Cambridge, MA. A paraître dans le *Journal of Financial Economics.*

MADHAVAN, A. et G. SOFIANOS, [1998], « An Empirical Analysis of NYSE Specialist Trading », *Journal of Financial Economics*, 48, pp. 189–210.

MEDDAHI, N., RENAULT E. et B. WERKER, [1998], « Modelling High-Frequency Data in Continuous Time », *mimeo*, Université de Montréal.

PRIGENT, J.-L., RENAULT, O. et O. SCAILLET, [1999], « An Autoregressive Conditional Binomial Option Pricing Model », *mimeo*, Thema, Université Cergy-Pontoise.

RUSSELL, J. R., [1999], « Econometric Modeling of Multivariate Irregularly-Spaced High-Frequency Data », *mimeo*, University of Chicago.

RUSSELL, J. R. et R. F. ENGLE, [1998], « Econometric Analisys Of Discrete-Valued Irregularly-Spaced Financial Transactions Data Using A New Autoregressive Conditional Multinomial Model », *Working Paper* 98-10, University of California, San Diego.

ZHANG, M. Y., RUSSELL J.R. et R. S. TSAY, [1999], « A Nonlinear Autoregressive Conditional Duration Model With Applications to Financial Transaction Data », *mimeo*, University of Chicago, Graduate School of Business. A paraître dans le *Journal of Econometrics.*