Actes des journées en Statistique Fonctionnelle et Opératorielle

10-11 Juin 2002, Toulouse

Recueillis par
H. Cardot, F. Ferraty, A. Mas, Y. Romain, P. Sarda, P. Vieu et A. Yousfate

STATISTIQUE FONCTIONNELLE ET OPÉRATORIELLE
Journées des 10-11 Juin 2002, Toulouse

Présentation des journées

Hervé CARDOT, Frédéric FERRATY, André MAS,
Yves ROMAIN, Pascal SARDA, Philippe VIEU et Abderrahmane YOUSFATE

cardot@toulouse.inra.fr, ferraty@cict.fr, mas@cict.fr, romain@cict.fr
sarda@cict.fr, vieu@cict.fr, yousfate_a@yahoo.com

Le thème général choisi pour ces rencontres est celui de la **Statistique Fonctionnelle**. Il s'agit d'un thème important de la Statistique moderne suscitant l'intérêt tant sur le plan théorique que sur celui des applications. Il est très présent au sein du **Laboratoire de Statistique et Probabilités** de l'Université Paul Sabatier au travers des activités du groupe de travail **STAPH**. Les problématiques abordées par ce groupe de travail sont entre autres l'estimation fonctionnelle, la modélisation pour variables fonctionnelles, la statistique opératorielle ou encore la statistique des processus, qui sont autant de thèmes privilégiés de ces rencontres (voir les références ci-dessous pour des accès aux activités de ce groupe de travail).

La Statistique Fonctionnelle est également un des centres d'intérêt des Statisticiens de **l'Université de Sidi-Bel-Abbes**. Ainsi, une collaboration entre des chercheurs de l'université de Sidi-Bel-Abbes d'une part et ceux du groupe STAPH d'autre part a été amorcée depuis deux ans. Ces rencontres s'inscrivent naturellement dans cette collaboration qui devrait s'amplifier dans le futur. Ces journées ont aussi été organisées en lien avec le Laboratoire de Biométrie et Intelligence Artificielle de **l'INRA de Toulouse** dont la collaboration avec le LSP sur ces thèmes liés à la Statistique Fonctionnelle est déjà effective depuis quelques années.

En nous ouvrant sur la communauté statistique, toulousaine, nationale et internationale, nous avons souhaité donner à ces journées un impact dépassant largement le contexte initial du groupe de travail STAPH. Nous tenons à remercier tous les orateurs, et plus généralement tous les participants, qu'ils soient venus de loin (Algérie, Espagne, Italie) de moins loin (Angers, Bordeaux, Montpellier, Paris) ou en simples voisins toulousains. Leur présence, ajoutée à la diversité des aspects fonctionnels de la Statistique que nous avons pu aborder durant ces deux jours, a grandement contribué au rayonnement de notre initiative et ne peut que nous encourager dans le futur.

Nous tenons enfin à souligner que sans la logistique apportée par le **LSP de Toulouse**, par **l'Université de Sidi-Bel-Abbes** et par **l'INRA de Tou-**

louse, et sans l'efficacité de Françoise MICHEL dans la gestion des divers aspects matériels inhérents à ce type de manifestation, ces journées n'auraient certainement pas pu se dérouler dans de si bonnes conditions.

Il ne nous reste plus qu'à souhaiter que les divers contacts amorcés ou renforcés durant ces journées s'avèrent fructueux, afin de ne faiore de celles-ci que le point de départ d'une dynamique durable. Nous sommes bien entendu preneurs de toute initiative et/ou proposition de nature à faire que la prochaine édition de cette manifestation n'ait rien à envier à celle-ci ...

Références

[1] Staph, (2001a). Statistique Fonctionnelle I : Groupe de Travail STAPH, Résumés des Exposés, Années 1999-2000. *Publication du laboratoire de Statistique et Probabilités, LSP-2001-05, Toulouse.*

[2] Staph, (2001b). Statistique Fonctionnelle II : Groupe de Travail STAPH, Résumés des Exposés, Années 2000-2001. *Publication du laboratoire de Statistique et Probabilités, LSP-2001-07, Toulouse.*

[3] Staph, (2002). Statistique Fonctionnelle III : Groupe de Travail STAPH, Résumés des Exposés, Années 2001-2002. *Publication du laboratoire de Statistique et Probabilités, en cours d'impression, Toulouse.*

[4] (*http ://www.lsp.ups-tlse2.fr/Fp/Ferraty/staph.html*)

Méthode hongroise pour les accroissements limites d'un processus de Wiener

BAHRAM Abdelkkader et YOUSFATE Abderrahmane
Sidi Bel Abbès, ALGÉRIE.

e-mail : bahram@univ-sba.dz, yousfate_a@yahoo.com

Pour contrôler la vitesse de convergence de certaines sommes partielles de variables aléatoires i.i.d., on utilise l'approximation wienerienne. Tenant compte des propriétés des accroissements limites d'un processus de Wiener standard, on utilise la loi du logarithme itéré et on introduit des paramètres de normalisation pour cerner ces accroissements limites.

Cette étude fait suite aux travaux de Csörgö et Révész ainsi que ceux de Book et Shore sur les accroissements limites d'un processus de Wiener standard et le résultat obtenu généralise les deux approches.

Références

[1] Book S. A., Shore T. R. (1978) On large intervals in the Csörgö-Révész theorem on increment of a Wiener process. *Z. Wahrsch. Verw. Geb.*, **46**, 1-11.
[2] Csörgö M., Révész P. (1981) *Strong approximation in probability and statistics.* Academic Press. New York.
[3] Gut, A., Spataru, A. (2000) Precise asymptotics in the law of the iterated logarithm. *Ann. Probability*, **4**, 1870-1883.

ACP conditionnelle

Mohamed E. BAOUCHE, Tawfik, BENCHIKH
Fatiha RACHEDI et Aberrhamane YOUSFATE
Sidi Bel Abbès, ALGÉRIE.

e-mail : baouche@univ-sba.dz, tbenchikh@univ-sba.dz,
frachedi@univ-tlemcen.dz yousfate_a@yahoo.com

Dans ce travail nous présentons l'*A.C.P.* conditionnelle "pas à pas" d'un opérateur linéaire continu U sous l'aspect d'une optimisation sous contrainte linéaire itérée ou globale. Ces résultats sont déjà connus (voir Dauxois et Pousse, Rachedi, ...). On présente ensuite l'*A.C.P.* conditionnelle d'un opérateur U qui présente deux types de solutions distinctes selon que la contrainte soit dans l'espace de définition de U ou dans l'espace image. Il est à noter que l'*A.C.P.* conditionnelle a connu des développements selon plusieurs axes : choix de métriques (Tebboune), variables instrumentales (Sabatier), variables Hilbertiennes (Smadhi), filtrage (Yousfate), transformations splines (Besse), (Durand), *A.C.P.C.L.* sur les L^p transformés en espaces de Hilbert (Rachedi), ... Une extension de l'*A.C.P.* conditionnelle dans des espaces de Banach réflexifs est proposée.

Références

[1] Benchikh T. (1999) Analyses factorielles dans un espace de Banach sous contraintes linéaires. Magister. UDL. Sidi-Bel-Abbès.
[2] Dauxois J., Pousse A. (1976) Les analyses factorielles en calcul des probabilité et en statistique : essai d'etude synthétique. Thèse es-sciences, Toulouse.
[3] Durand J. (1993) Generalized principal component analysis with respect to instrumental variables via univariate spline transformations. *Computational statistics and data analysis*, **16**, 423-440.
[4] Rachedi, F. (1996) ACPCL dans un espace de Hilbert et essai d'extension dans un espace de Banach. Magister. Sidi-Bel-Abbès.

Estimation localement suroptimale et adaptative de la densité

Denis BOSQ
Paris 6, FRANCE.

e-mail : bosq@ccr.jussieu.fr

On étudie une version tronquée de l'estimateur par projection dans un cadre général. On montre que cet estimateur atteint une vitesse suroptimale sur un ensemble dense dans la classe des densités à estimer et une vitesse quasi-optimale ailleurs. Cet ensemble peut être choisi par le statisticien et la vitesse suroptimale est atteinte pour l'erreur quadratique intégrée et la convergence uniforme presque sûre. Une version adaptative de l'estimateur est egalement considérée.

Références

www.ccr.jussieu.fr/LSTA (2002-5)

ACP dans le domaine des fréquences : applications

Alain BOUDOU* et Sylvie VIGUIER-PLA**
* Toulouse 3 et ** Perpignan, FRANCE.

e-mail : boudou@cict.fr, viguier@cict.fr

Etant donnée une série stationnaire $p-$dimensionnelle (c'est-à-dire d'une famille $(X_n)_{n\in\mathbb{Z}}$ d'éléments de $L^2_p(\Omega, \mathcal{A}, \mathcal{P})$, telle que $EX_n^t\overline{X_m} = EX_{n-m}^t\overline{X_0}$), nous nous proposons de la résumer par une série $q-$dimensionnelle ($q < p$) du type $\left(\sum_{k\in\mathbb{Z}} A_k X_{n-k}\right)_{n\in\mathbb{Z}}$.

Pour cela, nous effectuons une analyse en composantes principales (ACP) dans le domaine des fréquences. D'une façon plus précise, le résumé optimal est obtenu par l'ACP de la mesure aléatoire (m.a.) Z associée au processus, telle que cette analyse est définie en Boudou and Dauxois (1994).

Cela nécessite la diagonalisation d'une infinité de matrices et ne peut donc être envisagé concrètement. Nous tournons cette difficulté grâce à une discrétisation du spectre. On approche ainsi le "résumé idéal" par l'ACP d'une m.a. image de Z. Cette m.a. se réduisant à une somme finie de Dirac, son analyse se réduit à la diagonalisation d'un nombre fini de matrices. Nous montrons que, moyennant une hypothèse de régularité, lorsque le nombre de pas tend vers l'infini, il y a de bonnes propriétés de convergence. Une hypothèse d'ergodicité nous permet d'estimer les opérateurs qu'il est nécessaire de diagonaliser.

Enfin, nous examinons un exemple concernant les températures moyennes mensuelles de certaines villes de France, durant 50 années.

Références

[1] Boudou, A. and Dauxois, J. (1994) Principal Component Analysis for a stationary Random Function Defined on a Locally Compact Group. *J. Multivariate Anal.*, **51**, pp1 − 16

Test d'additivité en régression non paramétrique sous des conditions de β-mélange

Christine CAMLONG-VIOT
Paris Sud, FRANCE.

e-mail : christine.viot@cep.u-psud.fr

Dans la pratique, lorsqu'il s'agit d'étudier une fonction de régression, on a souvent affaire à une fonction multivariée, c'est-à-dire que l'on explique une variable aléatoire réelle Y par un vecteur aléatoire $\mathbf{X}$ de $\mathbb{R}^d$. Dans le cadre général des modèles de régression non paramétriques purs, les vitesses de convergence des estimateurs décroissent lorsque la dimension d des covariables augmente : c'est le "fléau de la dimension". Cet effet de la dimension peut être éliminé par l'utilisation de modèles de régression non paramétriques additifs sous lesquels il s'agit non plus d'estimer une seule fonction multivariée mais d fonctions univariées. Sous ces modèles de régression il existe plusieurs méthodes d'estimation des composantes additives. Celle que nous avons choisi d'utiliser est celle de l'intégration marginale. Le cas de variables dépendantes a été largement étudié dans le passé ; l'étude que nous présentons ici considère le cas où les variables sont β-mélangeantes. Il faut remarquer que lorsque les observations utilisées présentent des structures de dépendance, l'estimation d'une fonction de régression permet aussi de faire de la prévision en séries chronologiques.

Mais quand peut-on raisonnablement considérer qu'un modèle de régression présente une structure additive ? Afin de répondre à cette question, nous proposons un test statistique reposant sur les déviations pondérées entre les observations et la fonction de régression additive estimée pour lequel, après avoir montré sa normalité asymptotique, nous étudions sa puissance au moyen de simulations numériques.

Références

[1] Camlong-Viot Ch. & Louani D. (2002). Additivity test for nonparametric regression models under β-mixing conditions. *Preprint.*

On functional linear models and anova tests

Antonio CUEVAS
Univ. Autonoma, Madrid, ESPAGNE.

e-mail : antonio.cuevas@uam.es

After a short partial review of the theory of linear models with functional data, the talk will focus on the problem of testing the equality of several functional means. More precisely, given k independent samples of trajectories drawn from k stochastic processes, with respective mean functions $m_1(t), \ldots, m_k(t)$, the problem of testing $H_0 : m_i = \ldots = m_k$ is considered. So the setting is quite similar to that of the classical one-way anova model but the k samples under study consist of functional data. In this talk we consider a simple natural test for this problem which can be seen as a "functional analog" of the well-known F-anova test. The asymptotic validity of the method will be shown. A simulation study will be presented and a real-data example in experimental cardiology will be considered in some detail.

This is joint work with Manuel Febrero and Ricardo Fraiman.

Références

[1] Cuevas, A., Febrero, M. and Fraiman, R. (2002). An anova test for functional data. *en péparation.*
[2] Cuevas, A., Febrero, M. and Fraiman, R. (2002). Linear functional regression . *Cand. J. of Statist.*, in print.

Modèles de régression sur variables fonctionnelles

Frédéric FERRATY
Toulouse 2 et Toulouse 3, FRANCE.

e-mail : ferraty@univ-tlse2.fr ou ferraty@cict.fr
Travaux en lien étroit avec le groupe de travail STAPH
(*http ://www.lsp.ups-tlse2.fr/Fp/Ferraty/staph.html*)

Les appareillages de mesures de plus en plus sophistiqués ainsi que l'augmentation des capacités de stockage de l'information mettent à la disposition des statisticiens des données de type fonctionnel. L'observation d'un individu peut être alors considéré comme une courbe, voire un objet mathématique plus complexe. On parle généralement de variables aléatoires fonctionnelles (v.a.f.) car ces variables sont à valeurs dans un espace de fonctions $\mathcal{F}$ de dimension infinie. Frank et Friedman (1993) ainsi que Ramsay et Silverman (1997) proposent un tour d'horizon des différentes méthodes existantes autour des données fonctionnelles. Staph (2001a) et Staph (2001b) fournissent un large éventail à la fois méthodologique et théorique sur les avancées les plus récentes dans ce domaine. Par ailleurs, les exemples de données fonctionnelles ainsi que leurs champs d'applications sont de plus en plus nombreux. Qu'il s'agisse de courbes spectrométriques dans le proche infra-rouge en chimie quantitative (Ferraty et Vieu, 2002), de courbes digitalisées en reconnaissance vocale (Hastie et *al.*, 1995, Marx et Eilers, 1999) ou encore de trajectoires de phénomènes évoluant dans le temps à savoir consommation d'électricité (Ferraty et *al.* 2002), valeurs boursières, indices météorologiques, ..., dans tous ces exemples on est en présence de variables aléatoires fonctionnelles appelées aussi données fonctionnelles voire données longitudinales, bien que la terminologie puisse sensiblement différer d'un auteur à un autre.

Concernant cet exposé, on s'intéresse plus particulièrement aux modèles de régression d'une variable aléatoire réelle (v.a.r.) Y sur une v.a.f. X. On pose alors le modèle de régression

$$Y = R(X) + \epsilon$$

où ϵ est une v.a.r. centrée et indépendante de la v.a.f. X. Dans un premier temps, on considère une approche linéaire en supposant que R est un opérateur linéaire continu (Cardot et *al.*, 1999, Cardot et *al.*, 2002). Puis, dans un second temps, on relache cette contrainte de linéarité en proposant un cadre nonparmétrique. Enfin, dans chacune de ces deux situations (linéaire/nonparamétrique), un estimateur $\hat{R}$ de R est proposé pour lequel on donne des propriétés asymptotiques. Finalement, au cours de cette présentation, nous verrons comment des aspects mathématiques peuvent justifier ces développements statistiques d'un point de vue théorique alors que le bon comportement de ces méthodes sur des exemples concrets valide aussi leurs pratiques.

Références

[1] H. Cardot, F. Ferraty, A. Mas et P. Sarda, (2002). Test of nullity in the Functional Linear Model. *À paraître dans Scand. J. Stat.*.
[2] H. Cardot, F. Ferraty, P. Sarda, (1999). Linear Functional Model. *Statist. & Prob. Letters*, **45**, 11-22.
[3] F. Ferraty, A. Goïa, P. Vieu, (2002). Functional Nonparametric Model for Time Series : a Fractal Approach for Dimension Reduction. *Test*, in print.
[4] F. Ferraty, P. Vieu, (2002). Functional Nonparametric Model and Application to Spectrometric Data. *Computational Statistics*, in print.
[5] F. Ferraty, P. Vieu, (2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *C. R. Acad. Sci. Paris*, t. 330, Série I, 139-142.
[6] Frank, I.E., J.H. Friedman, (1993). A statistical view of some chemometrics regression tools. *Technometrics*, **35**, 109-148.
[7] T. Hastie, A. Buja, R. Tibshirani, (1995). Penalized Discriminant Analysis. *Ann. Stat.*, **23**, 73-102.
[8] B. Marx, P. Eilers, (1999). Generalized Linear Regression on Sampled Signals and Curves : A P-Spline Approach. *J.A.S.A.*, **41**, 1-13.
[9] J. Ramsay, B. Silverman, (1997) *Functional Data Analysis*. Springer-Verlag.
[10] Staph, (2001b). Statistique Fonctionnelle II : Groupe de Travail STAPH, Résumés des Exposés, Années 2000-2001. *Publication du laboratoire de Statistique et Probabilités, LSP-2001-07, Toulouse.*
[11] Staph, (2001a). Statistique Fonctionnelle I : Groupe de Travail STAPH, Résumés des Exposés, Années 1999-2000. *Publication du laboratoire de Statistique et Probabilités, LSP-2001-05, Toulouse.*

Un modèle semi-paramétrique Hilbertien

Louis FERRÉ
Toulouse 2, FRANCE.

e-mail : loferre@univ-tlse2.fr

Nous nous intéressons à un modèle semi-paramétrique pour décrire les liaisons existant entre deux variables X et Y à valeurs dans des espaces de Hilbert. Ce modèle s'écrit $Y = f(g(X), e)$ où g est une application linéaire continue, f une application mesurable et e une variable aléatoire, chacune étant définie sur des espaces de Hilbert. L'application g joue ici le rôle de la partie paramétrique par analogie avec le cadre multivarié. Nous démontrons alors que, sous certaines hypothèses, l'image de g est contenue dans le sous-espace propre engendré par les valeurs propres non-nulles de $IG(X)G(E(X|Y))$ où $IG(X)$ est l'inverse de l'opérateur de covariance de X et $G(E(X|Y))$ est l'opérateur de covariance de l'espérance conditionnelle à Y. Ces résultats généralisent donc ceux présentés par Li dans le cas multivarié et conduisant à la méthode SIR lorsque Y est univariée. Nous proposons alors une version fonctionnelle (cas où X prend ses valeurs dans un espace de fonctions) de la méthode SIR en utilisant soit une méthode de filtrage pour estimer $IG(X)$, soit en estimant $(IG(X)G(E(X|Y)))+$ l'inverse généralisé de $IG(X)G(E(X|Y))$. Puis nous montrons comment ces résultats conduisent à définir une règle d'affectation dans le cadre de l'analyse discriminante fonctionnelle en utilisant ici une méthode de régularisation. Plusieurs exemples, réels ou simulés, étayent notre propos.

Estimation fonctionnelle en Ψ-régression. Vitesse de convergence et normalité asymptotique

Ali LAKSACI
Sidi Bel Abbès, ALGÉRIE.

e-mail : alilak@yahoo.fr

La Ψ-régression a déjà été étudiée par beaucoup d'auteurs (Collomb, Härdle, ...) pour montrer la convergence uniforme de l'estimateur pour des processus mélangeants. Dans ce travail, on considère un processus stationnaire au sens strict à valeurs dans R^{d+1} et on se propose d'étudier la normalité asymptotique d'un estimateur fonctionnel de la Ψ -régression qui consiste à substituer l'estimateur formel par un estimateur robuste que l'on déterminera. Sous certaines hypothèses classiques, en considérant les principaux types de mélange, on montre que cet estimateur est asymptotiquement normal et on en donne une vitesse de convergence en moyenne quadratique. Cet estimateur peut être exploité également sur certains processus de Markov.

Mots clés : Ergodicité, mélange, normalité asymptotique, Ψ-régression.

Références

[1] Collomb G. , Härdle W. (1986) A strong uniform convergence rates in robust nonparametric time series analysis and prediction : Kernel regression estimation from dependent observations. *Stoch. Proc. and Appl.*, **23**, pp 77-89.
[2] Laib N., Ould-Saïd (2000) A robust nonparametric estimation of the autoregression function under an ergodic hypothesis. *The Canadian Journal of Statistics*, **28**, no. 4, pp 817-828

Un test d'homoscedasticité conditionnelle dans les modèles de régression avec erreur ARCH

Djamal LOUANI
Paris 6, FRANCE.
e-mail : louani@ccr.jussieu.fr

On considère un modèle $AR(1)$ avec des erreurs $ARCH$ défini par

$$X_t = \theta X_{t-1} + \epsilon_t$$

$$\epsilon_t = e_t \left(\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \ldots + \alpha_p \epsilon_{t-p}^2\right)^{1/2},$$

où X_t est un processus ergodique stationnaire et les α_j sont des réels positifs. La suite (e_t) est une différence de martingale par rapport à la tribu $\mathcal{F}_t$ engendrée par $(X_0, \epsilon_{-p+1}, \ldots, \epsilon_0, \epsilon_1, \epsilon_t)$ stationnaire et vérifiant $E(e_t|\mathcal{F}_{t-1} = 0)$ p.s., et $E(e_t^2|\mathcal{F}_{t-1} = 1)$ p.s.

Nous proposons un test pour tester l'homoscedasticité conditionnelle des erreurs ϵ_t. Notre test est basé sur la loi fonctionnelle limite du processus des résumés estimés cumulés donnée, pour tout $s \in \mathbb{R}$ par

$$\hat{D}_n(s) = \sqrt{\frac{1}{n}} \sum_{t=1}^{n} U_{t,n} I_{\{X_{t-1} < s\}},$$

$U_{t,n}$ désignant le processus des résidus estimés. Sous l'hypothèse nulle, nous montrons que le processus limite $D(s)$ est un processus de Wiener changé de temps plus une v.a. gaussienne. Une décomposition de type Karhunen-Loève nous permet alors de construire un test du χ^2 pour l'homoscedasticité conditionnelle du modèle.

Une étude numérique nous permet de voir que notre test est largement plus performant que celui de Kolmogorov-Smirnov proposé par Chen-An (1997) et qu'il reste compétitif par rapport au test du multiplicateur de Lagrange en dehors du fait qu'il n'impose aucune condition de connaissance de loi a priori.

Référence

[1] Chen, P.M. and An, H.Z. (1997). A Kolmogorov-Smirnov test for conditional heteroscedasticité in time series. *Statist. & Proba. Letters*, **33**, 321-331.
[2] Laïb, N. and Louani, D. (2002). On the conditional homoscedasticity test in autoregressive model with ARCH errors. *Comm. in Statist., Theory and Methods*, **33**, 1179-1202.

Prédiction dans le modèle linéaire fonctionnel

André MAS
Toulouse 3, FRANCE.

e-mail : mas@cict.fr

Le modèle de régression linéaire fonctionelle (les inputs sont des fonctions ou des courbes et les outputs sont des observations fini-dimensionnelles) a été étudié en detail par Cardot, Ferraty et Sarda. Des théorèmes limites de convergence en probabilité et presque sûre ont été obtenus pour l'estimateur de l'opérateur de régression. Les hypothèses portent sur la vitesse de décroissance du spectre de l'opérateur de covariance des observations fonctionnelles. L'objet de l'exposé a pour but de compléter ces résultats asymptotiques par un TCL pour un prédicteur statistique. Le resultat de convergence en loi est particulier puisque la vitesse n'est pas en racine carrée de n mais dépend également des valeurs propres de l'opérateur de covariance. Ce travail est en cours mais il pourrait permettre de déterminer des boules ou des ensembles de confiance dans le cadre de la prévision par le modèle linéaire fonctionnel.

Références

[1] H. Cardot, F. Ferraty, A. Mas et P. Sarda, (2002). Test of nullity in the Functional Linear Model. *À paraître dans Scand. J. Stat.*.
[2] H. Cardot, F. Ferraty, P. Sarda, (1999). Linear Functional Model. *Statist. & Prob. Letters*, **45**, 11-22.

On the (intradaily) seasonality and dynamics of a financial point process : a semi-parametric approach

Juan M. RODRIGUEZ POO
Univ. Cantabria, Santander, ESPAGNE.

e-mail : rodrigjm@ccaix3.unican.es

A new method of estimating a component model for the analysis of financial durations is proposed. The components are long-run dynamics and seasonality. The latter is left unspecified and the former is assumed to fall within the class of a certain family of parametric functions. The proposed estimation procedure is based on a generalized profile likelihood approach and requires the assumption either of a likelihood function for the model errors or, at least, that the error density belongs to the class of exponential densities. Its main interest is twofold : first, consistent and asymptotically normal estimators for both the parameters of the long-run stochastic component and the nonparametric curve that approximates the deterministic seasonal component are provided. Hence, it is possible to derive correct inference for both parametric and nonparametric components. Second, the method is computationally very appealing since the resulting nonparametric estimator of the seasonal curve has an explicit form that turns out to be a transformation of the Nadaraya-Watson estimator. The method is applied to price and volume durations of a stock traded at the NYSE, and compared to estimation with splines and with adjustment methods. It is shown that the proposed method outperforms the other methods.

This talk is about a joint work with V. VEREDAS from CIRANO (Canada) and A. ESPOSA from Universidad Carlos III in Madrid.

Le produit tensoriel saurait-il mieux la Statistique que le statisticien ?

Yves ROMAIN
Toulouse 3, FRANCE.

e-mail : romain@cict.fr

Le titre est d'abord commenté rapidement : outils et formalismes (tensoriels en l'occurence) sont générateurs de nouvelles applications (comme cela a déjà été démontré en Physique[1] notamment).

Ensuite, à partir de plusieurs exemples issus de travaux récents, on développe le potentiel prometteur de la tensorisation de la Statistique multidimensionnelle :
- perturbation d'éléments tensoriels fonctionnels (avec application aux opérateurs de covariance) ;
- produits tensoriels de mesures aléatoires (avec application aux processus stationnaires associés) ;
- formalisme tensoriel opératoriel (algébrique) de la Statistique non commutative ;
- applications aux analyses multivariées tensorielles ou multiplicatives.

[1] Jeu-concours : qui a déclaré "Le Calcul Tensoriel sait mieux la Physique que le physicien lui-même" ?

Une approche semi-paramétrique pour l'estimation de courbes de références

Jérôme SARACCO
USTL Montpellier, FRANCE.

e-mail : saracco@math.univ-montp2.fr

La construction de courbes de référence est un sujet fréquemment rencontré dans les domaines biomédicaux ou industriels. Ces courbes sont classiquement estimées par des méthodes paramétriques nécessitant des hypothèses contraignantes. Pour pallier cet inconvénient, des méthodes non paramétriques, fondées sur l'estimation de quantiles conditionnels, ont été développées. Lorsque la covariable est multidimensionnelle, ces méthodes non paramétriques souffrent du "fléau de la dimension". Nous proposons une nouvelle méthodologie basée sur une étape de réduction de la dimension et sur une étape d'estimation non paramétrique de quantiles conditionnels. Cette approche semiparamétrique combine la méthode SIR et l'estimation à noyau de quantiles conditionnels. Nous montrons la convergence de l'estimateur ainsi construit. A partir de simulations, nous comparons, pour différentes dimensions de la covariable, cette procédure d'estimation avec l'estimateur non paramétrique à noyau sans réduction de dimension préalable. Nous illustrons aussi cette méthode sur un jeu de données réelles afin d'établir des courbes de référence pour des propriétés biophysiques de la peau de femmes françaises. Pour plus de détails, on pourra se reporter à [1].

Référence

[1] Gannoun, A., Girard, S., Guinot, C., Saracco, J. (2001). Dimension reduction in reference curves estimation, Rapport de Recherche, Unité de Biométrie, ENSAM-INRA-UMII, n^{o}01-06.

Sur l'estimation fonctionnelle des opérateurs de transition des processus U-markoviens.

Abderrahmane YOUSFATE
Sidi Bel Abbès, ALGÉRIE.

e-mail : yousfate_a@yahoo.com

Un processus est dit u-markovien ($u \in R^+$) si le conditionnement par rapport à son passé équivaut à son conditionnement sur l'intervalle de temps le plus proche de longueur u. On vérifie que tout processus markovien (au sens classique) est aussi u-markovien pour tout $u \in R^+$. En mettant des hypothèses peu restrictives sur les trajectoires du processus considéré, on montre que les techniques d'estimation fontionnelle peuvent être exploitées pour faire de la prévision en utilisant un estimateur de Nadaraya-Watson.

Références

[1] Chung K. L. (1982) *Lectures from Markov Processes to brownian Motion.* Springer-Verlag. New York.
[2] Ferraty F., Vieu P. (2001) Statistique fonctionnelle : Modèle de régression pour variables aléatoires uni, multi et infiniment dimensionnées. *Publications du LSP. Toulouse.*
[3] Laksaci A., Yousfate A. (2002) Estimation fonctionnelle de la densité de l'opérateur de transition d'un processus de Markov à temps discret. *C. R. Acad. Sc. Paris*, (sous presse).
[4] Roussas G. (1969) Nonparametric estimation of the transition distribution function of a Markov process. *Annals of Math. Stat.*, **40**, 1386-1400.

Pourquoi les scores de second ordre sont des opérateurs à signe non constant pour les distributions générant une variété à courbure négative ?

Abdelghani ALI-ZAZOU et Abderrahmane YOUSFATE
Sidi Bel Abbès, ALGÉRIE.

e-mail : zazou@yahoo.fr, yousfate_a@yahoo.com

La fonction score (gradient du logarithme de la vraissemblance) est très utilisée en théorie de l'information. Quand la fonction score génère une variété à courbure négative, en présence d'échantillons de taille faible, l'opérateur associé au score de second ordre (hessien du logarithme de la vraissemblance) possède une valeur propre négative. Cette valeur propre est souvent annulée pour se rapprocher (au sens des moindres carrés) de la matrice d'information (au sens de Fisher) dans un espace euclidien. En utilisant l'outil de géométrie différentielle, on arrive à donner un sens à la valeur propre négative qui est très gênante pour un utilisateur qui la considère comme valeur propre d'un opérateur linéaire associé à une métrique dans un espace euclidien, alors qu'elle est, en fait, valeur propre d'un opérateur linéaire associé à une pseudo-métrique (dans un espace pseudo-euclidien).

Références

[1] Amari S. I. (1990) *Differential geometrical methods in statistics.* Springer verlag (2nd edition). Berlin.
[2] Barndorff-Nielsen O. E., Cox D. R., Reid N. (1986) The role of differential geometry in statistical theory. *I.S.T. Review*, **54**, 83-96.

Liste et adresse des particpants

A. Ali-Zazou Univ. Sidi bel Abbes, alizazou@yahoo.fr

JM. Azaïs Univ. Toulouse 3, azais@cict.fr

W. Bahram Univ. Sidi bel Abbes, walid-bahramdz@yahoo.com

W. A. Baouche Univ. Sidi bel Abbes, elouard@univ-sba.dz

JM. Bardet Univ. Toulouse 3, bardet@cict.fr

T. Benchikh Univ. Sidi bel Abbes, tbenchikh@univ-sba.dz

P. Besse Univ. Toulouse 3, besse@math.ups-tlse.fr

D. Bosq Univ. Paris 6, bosq@ccr.jussieu.fr

A. Boudou Univ. Toulouse 3, boudou@cict.fr

R. Boumaza Inst. Horticulture, Angers, rachid.boumaza@inh.fr

C. Camlong-Viot Fac. Pharm., Univ. Paris Sud, christine.viot@cep.u-psud.fr

H. Cardot INRA Toulouse, cardot@toulouse.inra.fr

V. Couallier Univ. Bordeaux 2, couallier@sm.u-bordeaux2.fr

C. Crambes Univ. Toulouse 3, crambes.christophe@wanadoo.fr

A. Cuevas Univ. Auton. Madrid, cuevas@uam.es

A. Daouia Univ. Toulouse 3, daouia@cict.fr

P. Dubois Lyon 1.

F. Ferraty Univ. Toulouse 2, ferraty@univ-tlse2.fr

L. Ferré Univ. Toulouse 2, loferre@univ-tlse2.fr

J. Fine Univ. Toulouse 3, fine@cict.fr

F. Gamboa Univ. Toulouse 3, gamboa@cict.fr

H. Gharib Univ. Sidi bel Abbes, gharib-h@yahoo.com

A. Goïa Univ. Novara, Italie, goia@cict.fr

R. Lafosse Univ. Toulouse 3, lafosse@cict.fr

A. Laksaci Univ. Sidi bel Abbes, alilak@yahoo.fr

D. Louani Univ. Paris 6, louani@ccr.jussieu.fr

A. Mas Univ. Toulouse 3, mas@cict.fr

O. Mestre Météo Toulouse, olivier.mestre@meteo.fr

E. Molinari Univ. Montpellier 1 et INRA toulouse, molinari@helios.ensam.inra.fr

A. Rabhi Univ. Sidi bel Abbes, Rabhi-abbes@yahoo.fr

J. Rodriguez-Poo, Univ. Santander, Espagne, rodrigjm@ccaix3.unican.es

Y. Romain Univ. Toulouse 3, romain@cict.fr

J. Saracco Univ. Montpelleir 2, saracco@math.univ-montp2.fr

P. Sarda Univ. Toulouse 3, sarda@cict.fr

P. Vieu Univ. Toulouse 3, vieu@cict.fr

S. Viguier-Pla Univ. Perpignan, viguier@cict.fr

AF. Yao Univ. Toulouse 1, yao@cict.fr

A. Yousfate Univ. Sidi bel Abbes, yousfate-a@yahoo.com