Curriculum Vitae de Paul Deheuvels – 2011

Table des matières

---

1. Renseignements Personnels

1.a. État Civil

Deheuvels, Paul, René, Louis
né le 11 mars 1948 à Istanbul, Turquie
nationalité française[1]
marié[2], 4 enfants[3]

1.b. Adresses

Professionnelle:
L.S.T.A., Tour 15-25, 2ème étage, B213, Université Pierre et Marie Curie (Paris VI)
4 place Jussieu, 75252 Paris Cedex 15
Tél.: Sec. 01 44 27 85 62 Tel: 01 44 27 33 51 Fax: 01 44 27 33 42

Privée:
7 avenue du Château, 92340 Bourg-la-Reine
Téléphone & Fax : Tel. Privé: 01 46 61 20 61 Fax Privé: 01 46 61 66 85

Courrier électronique:
paul.deheuvels@upmc.fr, paul.deheuvels@sfr.fr

2. Carrière

2.a. Diplômes & Études

École Normale Supérieure (rue d'Ulm), 1967-70[4]
Maîtrise de Mathématiques, Université Paris VI, 1968
DEA de Mathématiques, Université Paris VI, 1969
Agrégation de Mathématiques, 1969
Thèse d'État, Université Paris VI, 1974

2.b. Emplois & Fonctions universitaires

Assistant agrégé, Université Paris VI, 1969-72
Maître-Assistant, Université Paris VI, 1972-74

---

[1] Par filiation (père : René Deheuvels, professeur émérite à l'Université Pierre et Marie Curie, mère : France Lagarde, sans profession)

[2] A Joële Cormerais, le 4 mars 1971.

[3] Fleur Deheuvels, ancienne élève de l'Ecole Polytechnique, mariée, 4 enfants, Sophie Deheuvels, ancienne élève de l'ESTP, mariée, 3 enfants, Camille Deheuvels, sage-femme diplômée, mariée, 2 enfants, Aurore Deheuvels, opticienne diplômée, célibataire.

[4] Après 2 années de préparation, Math. Sup. (1965-66), Math. Spé. (1966-67) au Lycée Louis-le-Grand, Paris, reçu à l'Ecole Polytechnique en 1967 (démission).

Maître de Conférences (ancien régime)[5], Université Pierre et Marie Curie, Paris VI, 1974-1978[6]
Professeur, Université Pierre et Marie Curie, Paris VI
2ème classe, 1978-1984, 1ère classe, 1985-1995,
classe exceptionnelle, 1er échelon, 1995-1999, 2ème échelon, 1999-
Directeur du Laboratoire de Statistique Théorique et Appliquée (L.S.T.A.), Université Pierre et Marie Curie, Paris VI, 1982-
Responsable du DEA de Statistique (puis Master M2, Spécialité Statistique) de l'Université Pierre et Marie Curie, Paris VI, 1980-
Président de la Commission des Thèses de Mathématiques, Université Pierre et Marie Curie, Paris VI, 1990-1997

2.c. Emplois & Fonctions industrielles

Conseiller de la Direction de la Compagnie Française des Pétroles – TOTAL, 1974-1994
Conseiller de la Direction – ELF – 1978-1992
Conseiller du groupe SANOFI – Recherche & Dévelopement – 1980-2010

2.d Congés sabbatiques & Invitations

Visiting Professor, Columbia University, New York, Semestre de printemps 1989
Visiting Professor, Columbia University, New York, Semestre de printemps 1999
Visiting Professor University of Rotterdam, 1993
Visiting Professor KUL (Katholieke Universiteit te Leuven), 1987
Professeur Invité, Universita degli Studi di Torino, 1985-1995

2.e. Sociétés savantes

Membre de l'Académie des Sciences, 2000- (Correspondant, 1996-2000)
Membre correspondant étranger de la "Real Academia de Ciencias Exactas, Físicas y Naturales" (Académie Royale d'Espagne), 2002-
Fellow of the Institute of Mathematical Statistics [IMS], 1985-
Membre de l'Institut International de la Statistique [IIS-ISI, International Statistical Institute], 1978-
Membre de la Société Bernoulli, 1978-

2.f. Récompenses Académiques

Prix Gegner, Académie des Sciences, 1988
Prix Pierre-Simon de Laplace, Société de Statistique de France, 2007

---

[5] Corps assimilé aux professeurs d'université de 2ème classe, en 1978.
[6] Service militaire de juillet 1975 à juin 1976, comme sous-lieutenant (statut IMO) d'artillerie. Affecté au 45ème Régiment de Transmissions (Montélimar), avant d'être affecté au S.R.O.A.T., Service de Recherche Opérationnelle de l'Armée de Terre, de septembre 1975 à juin 1976. Lieutenant de Réserve.

2.g. Autres Activités d'Intérêt Collectif

2.g.a. Fonctions éditoriales (liste partielle)

Directeur de la collection "Mathématiques", Presses Universitaires de France, 1980-2000
Éditeur Associé de "Mathematical Methods in Statistics", 1991-2001
Éditeur Associé de "Statistics and Probability Letters", 1994-2010
Membre du Comité Éditorial des "Comptes Rendus de l'Académie des Sciences", 1995-
Éditeur Associé de "Extremes", 1997-
Éditeur Associé de "Statistical Inference for Stochastic Processes", 1997-
Éditeur Associé de "M.C.A.P." 2002-

2.g.b. Organisation de congrès

Co-organisateur du congrès "Stochastics", Oberwolfach, 1993

2.g.c. Activités d'expertise éditoriale (liste partielle)

Annals of Probability, Annals of Statistics, Probability Theory and Related Fields, Annales de l'Institut Henri Poincaré, Journal of Applied Probability, Advances in Applied Probability, Stochastic Processes and their Applications, Journal of Multivariate Analysis, Journal of Statistical Planning and Inference, Statistics and Probability Letters, Statistics and Decisions, Scandinavian Actuarial Journal, Comptes Rendus de l'Académie des Sciences, Statistics, ESAIM, Metrika.

2.g.d. Autres responsabilités administratives, Jurys & Concours

Examinateur (oral de mathématiques) du Concours de l'Ecole Spéciale Militaire de Saint-Cyr-Coëtquidan (1972-82).
Examinateur (écrit de mathématiques) de l'Ecole de l'Air de Salon de Provence (1980-1993).
Examinateur (écrit de mathématiques) du Concours Commun, Mines-Ponts (1976-1994).
Membre du Conseil de l'UER de Mathématiques - UFR 920 de Mathématiques, puis UFR 929, Université Paris VI, 1974-1997, 1997-2009.
Membre du CNU, Section 23.4 & 26, 1987-1995, 2003-
Membre du Comité National de la Recherche Scientifique, Section 03, 1987-1991
Directeur de l'I.S.U.P. [Institut de Statistique de l'Université de Paris], 1981-82.
Directeur de l'ESILV - Ecole Supérieure d'Ingénieurs Léonard de Vinci, 2009-2010.
Membre du Jury de l'Agrégation de Mathématiques, 1971, 1986, 1987

2.h. Anciens élèves de thèse (liste partielle)

Adrian Raftery, full professor, University of Seattle, Washington (Membre de la National Academy of Sciences des USA depuis 2009).
Michel Broniatowski, professeur, Université Pierre et Marie Curie (Paris VI)
George Haiman, professeur, Université de Lille
Armelle Guillou, professeur, Université Louis Pasteur (Strasbourg)
Jean Diebolt, directeur de recherches au C.N.R.S.
Jean-Noël Bacro, professeur, Université de Montpellier
Zhan Shi, professeur, Université Pierre et Marice Curie (Paris VI)
Alexandre Berred, professeur, Université du Hâvre
Marie-France Kratz, professeur, ESSEC
Charles El Nouty, professeur, Université Paris XIII
Philippe Berthet, professeur, Université Paul Sabatier, Toulouse
Margarida Brito, professeur, Université de Porto
Gane Samb Lo, professeur, Université de Saint Louis (Sénégal)
Abdelhakim Necir, professeur, Université de Biskra (Algérie)
Abdelouahid Imlahi, professeur, Université de Tanger (Maroc)
Gratiane Ennadifi, maître de conférences, Université de Lyon II
Sergio Alvarez-Andrade, maître de conférences, Université de Compiègne
Zohra Cherfi, maître de conférences, Université de Compiègne
Ludovic Menneteau, maître de conférences, Université de Montpellier
Alain Lucas, maître de conférences, IUT de Caen
Myriam Maumy, maître de conférences, Université Louis Pasteur (Strasbourg)
Jean-Renaud Pycke, maître de conférences, Université d'Evry
Davit Varron, maître de conférences, Université de Besançon
Pierre Ribereau, maître de conférences, Université de Montpellier
Vivian Viallon, maître de conférences, Université de Lyon 1
Salim Bouzebda, maitre de conférences, Université de Compiègne
Julien Cornebise, Research Associate, University College, London
Mamadou Kone, maître de conferences, CHU de Caen
Sarah Ouadah, maître de conférences, AGRO Paris-Tech

3. Publications Principales.

3.a. Articles de recherche publiés

[1] Sur la convergence de sommes de minimums de variables aléatoires (1973). *C. R. Acad. Sci. Paris* Ser. A-B **276** 309-312 [MR 48 #5156a]

[2] Sur la convergence de certaines suites de variables aléatoires (1973). *C. R. Acad. Sci. Paris* Ser. A-B **276** 641-644 [MR 48 #5156b]

[3] Sur une application de la théorie des processus de renouvellement à l'estimation de la densité d'une variable aléatoire (1973). *C. R. Acad. Sci. Paris* Ser. A-B **276** 943-946 [MR 48 #12710]

[4] Sur une famille d'estimateurs de la densité d'une variable aléatoire (1973). *C. R. Acad. Sci. Paris* Ser. A-B **276** 1013-1015 [MR 48 #12711]

[5] Sur l'estimation séquentielle de la densité (1973). *C. R. Acad. Sci. Paris* Ser. A-B **276** 1119-1121 [MR 48 #12712]

[6] Valeurs extrémales d'échantillons croissants d'une variable aléatoire réelle (1974). *Annales de l'Institut Henri Poincaré* Ser. B **10**, f.1, 89-114 [MR 50 #11404]

[7] Majoration et minoration presque sûre des extrémums de suites de variables aléatoires indépendantes de même loi (1974). *C. R. Acad. Sci. Paris* Ser.A **278** 513-516 [MR 50 #3317]

[8] Majoration et minoration presque sûre des extrema de processus Gaussiens (1974). *C. R. Acad. Sci. Paris* Ser. A **278** 989-992 [MR 51 #1930]

[9] Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité (1974). *C. R. Acad. Sci. Paris* Ser. A **278** 1217-1220 [MR 49 #10032]

[10] Majoration et minoration presque sûre optimale des éléments de la statistique ordonnée d'un échantillon croissant de variables aléatoires indépendantes (1974). *Rendi Conti della Academia Nazionale dei Lincei* 8, **56**, f.5, 707-719 [MR 52 #15625]

[11] Estimation non paramétrique de la densité par histogrammes généralisés (1977). *Publications de l'Institut de Statistique de l'Université de* Paris **22**, f.1, 1-24 [MR #81h:62071]

[12] Estimation non paramétrique de la densité par histogrammes généralisés (II) (1977). *Revue de Statistique Appliquée* **25**, f.3, 5-42 [MR 58 #18876]

[13] Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes (1978). *Publications de l'Institut de Statistique de l'Université de* Paris **23**, f.3, 1-36.

[14] Propriétés d'existence et propriétés topologiques des fonctions de dépendance (1979). *C. R. Acad. Sci. Paris*, Ser. A **288** 217-220.

[15] Détermination complète du comportement asymptotique en loi des valeurs extrêmes multivariées d'un échantillon de vecteurs aléatoires indépendants (1979). C. R. Acad. Sci. Paris, Ser. A **288**, f.3, 631-634 [MR #80c:60038]

[16] Détermination des lois limites jointes de l'ensemble des points extrêmes d'un échantillon multivarié (1979). *C. R. Acad. Sci. Paris*, Ser. A **288** 631-644 [MR #80c:60034]

[17] Estimation non paramétrique de la densité compte tenu d'informations sur le support (1979). *Revue de Statistique Appliquée* **27**, f.3, 47-68 (avec P. Hominal) [MR #81e:62040]

[18] Estimation séquentielle de la densité (1979). Dans: *Contribuciones en Probabilidad y Estadistica Matematica Enseñanza de la Matematica y Analysis*. 156-168, Grindley, Granada, Espagne [MR #81h:62072]

[19] La fonction de dépendance empirique et ses propriétés, un test non paramétrique d'indépendance (1979). *Bulletin de l'Académie Royale de Belgique*, Classe des Sciences (5) **65**, f.6, 274-292 [MR #81h:62073]

[20] Non-parametric tests of independence (1980). Dans: *Statistique Non Paramétrique Asymptotique* (J. P. Raoult, Edit.) 95-107, Lecture Notes in Mathematics **821**, Springer Verlag, Berlin [MR #82c:62061]

[21] Estimation automatique de la densité (1980). *Revue de Statistique Appliquée* **28**, f.1, 23-55 (avec P. Hominal) [MR #82j:62026]
[22] Some applications of dependence functions: nonparametric estimates of extreme value distributions and a Kiefer-type bound for the uniform test of independence (1980). Dans: *Nonparametric Statistical Inference*. Colloquia Math. Soc. János Bolyai **32**, 183-201, North Holland, Amsterdam [MR #85g:62064]
[23] The decomposition of infinite order and extreme multivariate distributions (1980). Dans: *Asymptotic Theory of Statistical Tests and Estimation*. (I. M. Chakravarti, Edit.), 259-286, Academic Press, New York [MR #82j:62032]
[24] A Kolmogorov-Smirnov test for independence (1981). *Revue Roumaine de Mathématiques Pures et Appliquées* **26**, f.2, 213-226 [MR #83c:62064]
[25] La prévision des séries économiques, une technique subjective (1981). *Archives de l'I.S.M.E.A.* **34**, f.4, 729-748.
[26] An asymptotic decomposition for multivariate distribution-free tests of independence (1981). *Journal of Multivariate Analysis* **11** 102-113 [MR #82g:62067]
[27] Multivariate tests of independence (1981). Dans: *Analytical Methods in Probability Theory*. (D. Dugué, E. Lukacs et V. K. Rohatgi, Edit.), 102-113, Lecture Notes in Mathematics **861**, Springer Verlag, Berlin [MR #83g:62074]
[28] A non-parametric test for independence (1981). *Publications de l'Institut de Statistique de l'Université de Paris* **26**, f.2, 29-50.
[29] The strong approximation of extremal processes (1981). *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **58** 1-6.
[30] Univariate extreme values - Theory and applications (1981). *Proceedings of the 43d Session of the International Statistical Institute*. **49**, f.2, 837-858.
[31] Strong limiting bounds for maximal uniform spacings (1982). *Annals of Probability* 1**0** 1058-1065.
[32] Spacings, record times and extremal processes (1982). Dans: *Exchangeability in Probability and Statistics* (G. Koch et F. Spizzichino, Edit.), North Holland, Amsterdam, 223-243.
[33] A construction of extremal processes (1982). Dans: *Probability and Statistical Inference* (W. Grossmann, G. C. Pflug et W. Wertz, Edit.), 53-58, Reidel, Dordrecht.
[34] Sur des tests d'ajustement indépendants des paramètres (1982). Dans: *Actas, II Coloquio de Estatistica - A Estatistica nos Processos Estocasticos*, Departamente de Matematica, Universidade de Coimbra, 7-18.
[35] On record times associated with *k*-th extremes (1982). *Proceedings of the 3rd Pannonian Symposium on Mathematical Statistics* (J. Mogyoródi, I. Vincze et W. Wertz, Edit.), Akadémiai Kiadó, Budapest.
[36] Invariance of Wiener processes and Brownian bridges by integral transforms and applications (1982). Stochastic Processes and their Applications. **13**, f.3, 311-318.
[37] L'encadrement asymptotique des éléments de la série d'Engel d'un nombre réel (1982). *C. R. Acad. Sci. Paris*, Ser. A **295** 21-24.
[38] Point processes and multivariate extreme values (1983). *Journal of Multivariate Analysis.* **13** 257-272.

[39] The complete characterization of the upper and lower class of the record and inter-record times of an i.i.d. sequence (1983). *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete.* **62**, 1-6.
[40] The strong approximation of extremal processes (II) (1983). *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete.* **62**, 7-15.
[41] Upper bounds for *k*-th maximal spacings (1983). *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete.* **62**, 465-474.
[42] Strong bounds for multidimensional spacings (1983). *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete.* **64**, 411-424.
[43] Indépendance multivariée partielle et inégalités de Fréchet (1983). Dans: *Studies in Probabilities and Related Topics* (Papers in Honour of Octav Ionicescu on his 90th Birthday) (M. Demetrescu et M. Iosifescu, Edit.), 145-155, Nagard, Bucarest.
[44] Strong bounds for the maximal *k*-spacing when $k \leq c \log n$ (1984).). *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **66**, 315-334 (avec L. Devroye).
[45] Strong limit theorems for maximal spacings from a general univariate distribution (1984). *Annals of Probability* **12**, 1181--1193.
[46] How to analyze bioequivalence studies - The right use of confidence intervals (1984). *Journal of Organizational Behavior and Statistics* **1**, f.1, 1-15.
[47] Asymptotic results for the pseudo-prime sequence generated by Hawkins's random sieve: twin primes and Riemann's hypothesis (1984). Dans: *Proceedings of the 7th Conference in Probability Theory*, Brasov (M. Iosifescu, Edit.), 109-115, Editura Academiei, Bucarest.
[48] Probabilistic aspects of multivariate extremes (1984). Dans: *Statistical Extremes and Applications* (J. Tiago de Oliveira, Edit.), 117-130, D. Reidel, Dordrecht.
[49] Strong approximations of records and record times (1984). Dans: *Statistical Extremes and Applications* (J. Tiago de Oliveira, Edit.), 491-496, D. Reidel, Dordrecht.
[50] Strong approximation in extreme values, theory and applications (1984). Dans: *Limit Theorems in Probability and Statistics* (P. Révész, Edit.) Vol. 1, 369-404, Colloquia Math János Bolyai **36**, North Holland, Amsterdam.
[51] The characterization of distributions by order statistics and record values - A unified approach (1984). *Journal of Applied Probability.* **21** 326-334 (Corr. (1985). **22** 997).
[52] Point processes and multivariateextreme values (II) (1985). Dans: *Multivariate Analysis VI* (P. R. Krishnaiah, Edit.), 145-164, North Holland, Amsterdam.
[53] On the Erdős-Rényi theorem for random fields and sequences and its relationships with the theory of runs and spacings (1985). *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete.* **70**, 91-115.
[54] Lois de type Pareto et applications à la théorie mathématique du risque (1985). *Rend. Sem. Mat. Univers. Politecn. Torino.* **43**, f.1, 25-41.
[55] Kernel estimates of the tail index of a distribution (1985). *Annals of Statistics.* **13** 1050-1077 (avec S. Csörgő et D. M. Mason).
[56] Spacings and applications (1985). Dans: *Probability and Statistical Decision Theory* (F. Konecny, J. Mogyoródi et W. Wertz, Edit.), Vol. A, 1-30, Reidel, Dordrecht.
[57] The limiting behavior of the maximal spacing generated by an i.i.d. sequence of Gaussian random variables (1985). *Journal of Applied Probability.* **22**, 816-827.

[58] Exact convergence rate in the limit theorems of Erdős-Rényi and Shepp (1986). *Annals of Probability*. **14**, 209-223 (avec L. Devroye et J. Lynch).
[59] Exact convergence rate of an Erdős-Rényi strong law for moving quantiles (1986). *Journal of Applied Probability*. **23**, 355-369 (avec J. Steinebach).
[60] Strong laws for the k-th order statistic when $k \leq c \log n$ (1986). *Probability Theory and Related Fields*. **72**, 179-186.
[61] On the influence of the extreme values on the maximal spacing (1986). *Annals of Probability*. **14**, 194-208.
[62] A semigroup approach to Poisson approximation (1986). *Annals of Probability*. **14**, 663-676 (avec D. Pfeifer).
[63] Simple random walk on the line in random environment (1986). *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*. **72**, 215-230 (avec P. Révész).
[64] Operator semigroups and Poisson convergence in selected metrics (1986). *Semigroup Forum*. **34**, 203-224 (avec D. Pfeifer).
[65] Many heads in a short block (1987). Dans: *Mathematical Statistics and Probability Theory* (M. L. Puri, P. Révész et W. Wertz, Edit.), 53-67, D. Reidel, Dordrecht (avec P. Erdős[7], K. Grill et P. Révész).
[66] Weak laws for the increments of Wiener processes and Brownian bridges and applications (1987). Dans: *Mathematical Statistics and Probability Theory* (M. L. Puri, P. Révész et W. Wertz, Edit.), 69-87, D. Reidel, Dordrecht (avec P. Révész).
[67] Limit laws of the Erdős-Rényi-Shepp type (1987). *Annals of Probability*. **15**, 1363-1386 (avec L. Devroye).
[68] Exact convergence rates in strong approximation laws for large increments of partial sums (1987). *Probability Theory and Related Fields*. **76**, 369-393 (avec J. Steinebach).
[69] Semigroups and Poisson Approximation (1987). Dans: *New Perspectives in Theoretical and Applied Statistics* (M.Puri, J. Vilaplana et W. Wertz, Edit.) 439-448, Wiley, New York (avec D. Pfeifer).
[70] Exact convergence rates in Erdős-Rényi-type theorems for renewal processes (1987). *Annales de l'Institut Henri Poincaré*. **23**, 195-207 (avec J. N. Bacro et J. Steinebach).
[71] An approximation of stopped sums with applications in queuing theory (1987). *Advances in Applied Probability*. **19**, 674-690 (avec M. Csörgő et L. Horváth).
[72] The asymptotic behavior of sums of exponential extreme values (1988). *Bulletin des Sciences Mathématiques*. **112**, 211-233 (avec D. M. Mason).
[73] Strong approximations of k-th records and k-th record times by Wiener processes (1988). *Probability Theory and Related Fields*. **77**, 195-209.
[74] Limit laws for the modulus of continuity of the partial sum process and for the Shepp statistic (1988). *Stochastic Processes Appl*. **29**, 223-245 (avec J. Steinebach).
[75] The almost sure behavior of maximal and minimal k_n-spacings when $k_n = O(\log n)$ (1988). *Journal of Multivariate Analysis*. **24**, 155-176 (avec J. H. J. Einmahl, D. M. Mason et F. Ruymgaart).
[76] Almost sure convergence of the Hill estimator (1988). *Mathematical Proceedings of the Cambridge Philosophical Society*. **104**, 371--384 (avec E. Haeusler et D. M. Mason).

[7] Cet article me vaut l'insigne honneur d'être co-auteur direct du célèbre mathématicien hongrois Pal Erdős, et d'avoir ainsi un « nombre d'Erdős » égal à 1.

[77] Poisson approximations in selected metrics by coupling and semigroup methods with applications (1988). *Journal of Statistical Planning and Inference.* **20**, 1-22 (avec A. Karr, D. Pfeifer et R. Serfling).
[78] Poisson approximations of multinomial distributions and point processes (1988). *Journal of Multivariate Analysis.* **25**, 65-89 (avec D. Pfeifer).
[79] A new semigroup technique in Poisson approximation (1988). *Semigroup Forum.* **201**, 1-13 (avec D. Pfeifer et M. L. Puri).
[80] On the relationship between Uspensky's theorem and Poisson approximations (1988). *Annals of the Institute of Mathematical Statistics.* **40**, 671-681 (avec D. Pfeifer).
[81] Strong laws for the k-th order statistic when $k \leq c \log n$ (II) (1989). Dans: *Extreme Value Theory* (J. Hüsler et R. Reiss, Edit.), 21-35, Lecture Notes in Statistics **51**, Springer-Verlag, Berlin.
[82] A Bahadur-Kiefer-type two sample statistic with applications to tests of goodness of fit (1989). Dans: *Limit Theorems in Probability and Statistics* (P. Révész, Edit.), Colloquia Mathematica János Bolyai **57**, 157-172 (avec D. M. Mason).
[83] Sharp rates for the increments of renewal processes (1989). *Annals of Probability.* **17**, 700-722 (avec J. Steinebach).
[84] Asymptotic expansions of sums of non identically distributed binomial random variables (1989). *Journal of Multivariate Analysis.* **25**, 65-89 (avec M. L. Puri et S. Ralescu).
[85] On the relationship between stability of the extreme order statistics and convergence of the maximum likelihood kernel density estimate (1989). *Annals of Statistics.* **17**, 1070-1086 (avec M. Broniatowski et L. Devroye).
[86] On the non-parametric estimation of the bivariate extreme value distributions (1989). *Statistics and Probability Letters.* **9**, 241-251 (avec J. Tiago de Oliveira).
[87] On the almost sure behavior of sums of extreme values from a distribution in the domain of attraction of a Gumbel law (1990). *Bulletin des Sciences Mathématiques*, Ser. 2, **114**, 61-95 (avec E. Haeusler et D. M. Mason).
[88] Bahadur-Kiefer-type processes (1990). *Annals of Probability.* **18**, 669-697 (avec D. M. Mason).
[89] On the approximation of *P-P* and *Q-Q* plot processes by Brownian bridges (1990). *Statistics and Probability Letters.* **9**, 241-251 (avec J. Beirlant).
[90] On the sample path behavior of the first passage time process of a Brownian motion with drift (1990). *Annales de l'Institut Henri Poincaré.* **26**, 145-179 (avec J. Steinebach).
[91] Nonstandard functional laws of the iterated logarithm for tail empirical and quantile processes (1990). *Annals of Probability.* **18**, 1693-1722 (avec D. M. Mason).
[92] On some alternative estimates of the adjustment coefficient in risk theory (1990). *Scandinavian Actuarial Journal.* 135-159 (avec J. Steinebach).
[93] Bahadur-Kiefer theorems for uniform spacings processes (1991). *Teoria Verojatnosti i ee Primienienia.* **36**, 724-743 (avec J. Beirlant, J. H. J. Einmahl et D.M. Mason).
[94] Laws of the iterated logarithm for density estimators (1991). Dans: *Nonparametric Functional Estimation and Related Topics* (G. Roussas, Edit.), 19-29, Kluwer, Dordrecht.
[95] Estimating the quantile-density function (1991). Dans: *Nonparametric Functional Estimation and Related Topics* (G. Roussas, Edit.), 213-224, Kluwer, Dordrecht (avec M. Csörgő et L. Horváth).

[96] A tail empirical process approach to some nonstandard laws of the iterated logarithm (1991). Journal of Theoretical Probability. **4**, 53-85 (avec D. M. Mason).

[97] On the limiting behavior of the Pickands estimator for bivariate extreme-value distributions (1991). *Statistics and Probability Letters.* **12**, 429-439.

[98] Statistical analysis and modelling of slug lengths (1991). Dans: Multi-Phase Production (A. P. Burns, Edit), 80-112, *Proceedings of the 5-th International Conference on Multi-Phase Production*, Elsevier, London avec H. Dhulesia et M. Bernicot).

[99] Pointwise Bahadur-Kiefer-Type theorems (II) (1992). Dans: *Nonparametric Statistics and Related Topics* (A. K. Md. E. Saleh, Edit.), 331-346, North Holland, Amsterdam.

[100] On the limiting behavior of the Bahadur-Kiefer statistic for partial sums and renewal processes when the fourth moment does not exist (1992). *Statistics and Probability Letters*. **13**, 179-188 (avec J.Steinebach).

[101] Functional Erdős-Rényi laws (1991). St*udia Scientiarum Math. Hungarica*. **26**, f.2-3 261-295.

[102] Deux exemples d'estimation de la fiabilité de matériels à partir d'informations fournies par les maintenances corrective et préventive. (1992). *Revue de Statistique Appliquée.* **40**, f.2 77-90 (avec J. Lannoy and C. Chevalier).

[103] Large deviations by Poisson approximations (1992). *Journal of Statistical Planning and Inference.* **32**, 75-88.

[104] A functional law of the iterated logarithm for tail quantile processes (1992). *Journal of Statistical Planning and Inference.* **32**, 63-74.

[105] Pointwise Bahadur-Kiefer-type theorems. I. (1992). *Probability Theory and Applications.* Essays to the Memory of Joszef Mogyoródi. (J. Galambos et I. Katai edit.) 235-256, Math. Appl. **80**, Kluwer, Dordrecht [MR #94a:60028]

[106] A functional LIL approach to pointwise Bahadur-Kiefer theorems. (1993). *Probability in Banach Spaces.* **8**, Proceedings of the Eight International Conference. R. M. Dudley, M. G. Hahn et J. Kuelbs edit. 255-266. Progress in Probability **30** Birkhäuser, Boston (avec D. M .Mason).

[107] Functional laws of the iterated logarithm for the increments of empirical and quantile processes (1992). *Annals of Probability.* **20**, 1248-1287 (avec D. M. Mason).

[108] Functional laws of the iterated logarithm for large increments of empirical and quantile processes (1992). *Stochastic Processes and Applications.* **43**, 133-163.

[109] Approximations and two-sample tests based on *P-P* and *Q-Q* plots of the Kaplan-Meier estimators of lifetime distributions. (1992) Journal of Multivariate Analysis. **43**, 200-217. (avec J. H. J. Einmahl).

[110] Some results on the influence of extremes on the bootstrap (1993). *Annales de l'Institut Henri Poincaré.* **29**, 83-103. (avec D. M. Mason and G. R. Shorack).

[111] On fractal dimension and modelling of bubble and slug flow processes. (1993). *Multi Phase Production.* BHR Group Conference Series **4**, 95--16, A. Wilson Edit, Mechanical Engineering Publications Ltd, London (avec M. Bernicot and H. Dhulesia).

[112] Records in the F^{α}-scheme. I. Martingale properties. *Zapiski Nauchnykh Seminarov POMI* **207**, 19-36. (en Russe) (avec V. B. Nevzorov).

[113] On the coverage of Strassen-type sets by sequences of Wiener processes (1993). Journal of Theoretical Probability. **6**, 427-449. (avec P. Révész).

[114] Strassen-type functional laws for strong topologies (1993). Probability Theory and Related Fields. **97**, 151-156 (avec M. A. Lifshits).
[115] Limit laws for *k*-record times (1994). *Journal of Statistical Planning and Inference*. **38**, 279-307 (avec V. B.Nevzorov).
[116] Random fractals generated by oscillations of processes with stationary and independent increments. (1994). Pr*obability in Banach Spaces* **9** J. Hoffman-Jørgensen, J. Kuelbs, M. B. Marcus edit., Birkhäuser, Boston, 73-89 (avec D.M.Mason).
[117] Functional laws of the iterated logarithm for local empirical processes indexed by sets (1994). *Annals of Probability* **22**, 1619-1661 (avec D. M. Mason) [MR #96e:60048]
[118] Necessary and sufficient conditions for the Strassen law of the iterated logarithm in non-uniform topologies. (1994). *Annals of Probability.* **22**, 1838-1856 (avec M. A. Lifshits) [MR #96d:60051]
[119] A unified model for slug flow generation (1995). *Revue de l'Institut Français du Pétrole.* **50**, 219-236 (avec M. Bernicot).
[120] Nonstandard local empirical processes indexed by sets (1995). *Journal of Statistical Planning and Inference*. **45**, 91-112 (avec D.M.Mason [MR #97c:60011]
[121] On the fractal nature of empirical increments (1995). *Annals of Probability.* **23**, 355- -387 (avec D. M. Mason) [MR #96e:60064]
[122] A new law of the iterated logarithm for arrays (1996). Studia Scientiarum Mathematicarum Hungarica. **31**, 145-185 (avec H. Teicher) [MR #96k:60073]
[123] Strong laws for extreme values of sequences of partial sums (1996). *Studia Scientiarum Mathematicarum Hungarica*. **31**, 121-144 (avec J. Steinebach).
[124] Cramér-Von Mises-type tests with applications to tests of independence for multivariate extreme--value distributions (1996). *Communications in Statistics - Theory and Methods.* **25**, n°4, 871-908 (avec G. V. Martynov) [MR #97d:62087]
[125] On the strong limiting behavior of local functionals of empirical processes based upon censored data (1996). *Annals of Probability.* **24**, 504-525 (avec J. H. J. Einmahl) [MR #98f:62088]
[126] Functional laws for small increments of empirical processes (1996). *Statistica Neerlandica*. **50**, 261-280.
[127] On the Hausdorff dimension of the set generated by exceptional oscillations of a Wiener process (1997). *Studia Scientiarum Mathematicarum Hungarica*. **33**, 75-110 (avec M. A. Lifshits).
[128] On random fractals (1997). *Supplemento ai Rendiconti del Circolo Matematico di Palermo*, Ser. II **50**, 111-132.
[129] Strong laws for local quantile processes (1997). *Annals of Probability.* **25**, 2007-2054 [MR #99c:60064]
[130] Protuberance effect in the generalized functional Strassen-Révész law (1998). *Journal of Mathematical Sciences.* **88**, 22-28 (avec M.A.Lifshits) [traduit de *Zapiski Nauchnykh Seminarov St. Petersburg. Otdel. Mat. inst. Steklov. POMI.* **216**, 33-41 (1994)] [MR #96m:60071]
[131] Records in the F^{α}-scheme. II. Limit laws (1998). Journal of Mathematical Sciences. **88**, 29-35 (avec V. B. Nevzorov) [traduit du Russe de *Zapiski Nauchnykh Seminarov St. Petersburg. Otdel. Mat. inst. Steklov. POMI.* **216**, 42-51 (1994)] [MR #96e:60091]

[132] Random fractals and Chung--type functional laws of the iterated logarithm (1998). *Studia Scientiarum Mathematicarum Hungarica*. **34**, 89-106 (avec D. M. Mason).

[133] On local oscillations of empirical and quantile processes (1998). *Asymptotic Methods in Probability and Statistics*. 127-134. Edit. B. Szyszkowicz, North Holland, Amsterdam [MR #99i:60076]

[134] On the approximation of quantile processes by Kiefer processes (1998). *Journal of Theoretical Probability.* **11**, 997-1018 [MR #2000f:60045]

[135] Fluctuations du processus de risque et coefficient d'ajustement (1998). *Atti del XXI Convegno Annuale A.M.A.S.E.S. Appendice.* 41-60.

[136] Bootstrap for maxima and records (1999). *Zapiski nauchny seminarov POMI* (Notes of Sci. Semin. of Peterburg Branch of Math. Inst.). **260**, 119-129 (en Russe) [Traduction dans: J. Math. Sci. (New York) **109**, (2002), n°6, 2115-2121] [MR #2001h:62077]

[137] Chung-type functional laws of the iterated logarithm for tail empirical processes. (2000). *Annales de l'Institut Henri Poincaré*. **36**, 583-616 [MR #2001m:60065]

[138] Functional Limit laws for the Increments of Kaplan-Meier Product-Limit Processes and Applications (2000). *Annals of Probability.* **28**, 1301-1335 (avec J. H. J. Einmahl) [MR #2001k:62039]

[139] Nonstandard Strong Laws for Local Quantile Processes (2000). *Journal of Statistical Planning and Inference.* **91**, 239-266 [MR #2001m:62067]

[140] Une approche nouvelle pour l'exploitation des empreintes chromatographiques. (2000). *Ann. Fals. Exp. Chim*. **92**, 455-470 (avec Richard, J. P. et Pierre-Loti-Viaud, D.).

[141] Uniform limit laws for kernel density estimators on possibly unbounded intervals. (2000). *Recent Advances in Reliability Theory: Methodology, Practice and Inference*. N. Limnios et M. Nikulin, Edit. 477-492, *Stat. Ind. Technol*. Birkhäuser, Boston [MR #2001j:62042]

[142] Strong Approximation of Quantile Processes by Iterated Kiefer Processes (2000). *Annals of Probability.* **28**, 909-945 [MR #2001j:60038]

[143] Asymptotic independence of the local empirical process indexed by functions (2001). *High dimensional Probability II* (Seattle, WA, 1999). 183-206. Progress in Probability, Vol. **47** (E. Giné, D. M. Mason, J. A. Wellner edit.), Birkhäuser, Boston (avec D.M.Mason et U.Einmahl) [MR #2002f:60057]

[144] Limit Laws for kernel density estimators for kernels with unbounded supports (2001). *Asymptotics in Statistics and Probability.* Papers in Honor of George Gregory Roussas. 117-131. M.L. Puri Edit. VSP International Science Publishers, Pays Bas.

[145] Probabilities of hitting shifted small balls by a centered Poisson processes (2001). *Zapiski Nauchn. Seminarov St. Petersburg. Otdel. Mat. Inst. Steklov POMI.* (en russe) **278**, 63-85 [Verojatn. i Stat. **4**, 310-311] (en russe) [MR #2002i:60100]

[146] Estimation non-paramétrique de la régression dichotomique - Application biomédicale (2002). *C. R. Acad. Sci. Paris*, Ser. I. **334**, 59-63 [MR #2002m:62054]

[147] Tests of Independence with Exponential Marginals (2002). *Goodness-of-Fit Tests and Model Validity*. 463-476 (C. Huber-Carol, N. Balakrishnan, M. S. Nikulin et M. Mesbah edit.), Birkhäuser Verlag, Basel [MR1901856].

[148] Exact laws for sums of logarithms of uniform Spacings (2003). A*ustrian Journal of Statistics*. **32**, n°1 & 2, 29-47 (avec G.Derzko).

[149] Karhunen-Loeve expansions for weighted Wiener processes and Brownian bridges via Bessel functions. (2003). *High dimensional Probability III.* Progress in Probability, 57--93 (M. Marcus et J. A. Wellner edit.), Birkhäuser, Boston (avec G. V. Martynov) [MR2033881#2005i:62074]

[150] General confidence bounds for nonparametric functional estimators (2004). *Statistical Inference for Stochastic Processes*. Vol. **7**, 225-277 (avec D. M. Mason) [MR2111291#2005k:62131]

[151] An extension of Lévy's formula to weighted Wiener processes (2004). Dans: *Parametric and Semiparametric Models wit Applications to Reliability, Survival Analysis, and Quality of Life.* 519-531. *Stat. Ind. Technol.*, Birkhäuser, Boston [MR2091632#2006f:60079]

[152] Tests of fit based on products of spacings (2006). (avec G. Derzko). *Probability, statistics and modelling in public health,* 119-135, *Springer, New York* [MR2230727#2007h:62042]

[153] Weighted multivariate Cramér-von Mises type statistics (2005). Afrika Statistika **1** n°1, 1-14 [MR2298871#2008k:62103]

[154] Limit laws for kernel smoothed histograms (avec G. Derzko et L. Devroye). Proceedings of the St.Petersburg Conference on limit theorems in Probability.

[155] On quadratic functionals of the Brownian sheet and related functionals (2006) (avec G. Peccati et M. Yor). *Stochastic Processes Applications*. Vol. **116**, 493-538 [MR2199561#2007k:60064]

[156] Karhunen-Loève expansions of mean-centered Wiener processes (2006). *IMS Lecture Notes Monograph Series High Dimensional Probability.* Vol. **51**, 62-76, Institute of Mathematical Statistics [MR2387761#2009e:62184]

[157] Weighted multivariate tests of independence (2007). *Comm. Statist. Theory Meth.* Vol. **36**, 1-15 [MR2413589];

[158] A Karhunen-Loève expansion for a mean-centered Brownian bridge (2007). *Statist. Probab. Letters.* Vol. **77**, 1190-1200 [MR2392790#2009b:62089].

[159] A Karhunen-Loeve decomposition of a Gaussian process generated by independent pairs of exponential random variables (2008). *Journal of Functional Analysis*. Vol. **255**, n°9, 2363-2394. (avec G.Martynov) [MR2473261].

[160] Topics on empirical processes (2007). *EMS Ser. Lect. Math,* 93-190.*, Eur. Math. Soc.,* Zürich [MR2347004].

[161] Asymptotic certainty bands for kernel density estimators based upon a bootstrap resampling scheme (2008). *Statistical models and methods for biomedical and technical systems.* 171-186, Stat. Ind. Technol., Birkhäuser, Boston [MR2462913].

[162] A Glivenko-Cantelli-type theorem for the spacing-ratio empirical distribution functions (2008). *Ann. I.S.U.P.* Vol. **52**, 25-38 [MR2435038#2010a:62037].

[163] A multivariate Bahadur-Kiefer representation for the empirical copula process (2009). *Journal of Mathematical Sciences.* Vol. **163**, n°4, 382-398.

[164] Spacings -ratio empirical processes (avec G. Derzko) (2010). *Periodica Mathematica Hungarica.* Vol. **61**, 121-164.

[165] Non-uniform spacings processes (2011). *Statistical Inference Stoch. Processes*. Vol.**14**, n°2, 141-175.

[166] One bootstrap suffices to generate sharp uniform bounds in functional estimation (2011). *Kybernetika*, Vol. **47**, n°6, 855-865.
[167] Uniform in bandwidth functional limit laws (avec S. Ouadah) (2013). *J. Theoretical Probab.* Vol. 26, 697-721.
[168] On Testing Stochastic Dominance by Exceedance, Precedence and Other Distribution-Free Tests, with Applications (2014). In: *Statistical Models and Methods for Reliability and Survival Analysis* (eds V. Couallier, L. Gerville-Réache, C. Huber-Carol, N. Limnios and M. Mesbah), John Wiley & Sons, Inc., Hoboken, USA, 145-159.

4. Synthèse des Travaux de Recherche Principaux

4.a. Estimation Fonctionnelle

Soit une suite d'observations $X_1, X_2, \ldots,$ composée de répliques aléatoires indépendantes, de même loi de probabilité, d'une variable générique X. En notant $F(x) = \mathbb{P}(X \leq x)$, la fonction de répartition de X, nous nous plaçons sous l'hypothèse d'existence d'une densité (au sens de Lebesgue)

$$f(x) = \frac{d}{dx} F(x) \text{ pour } x \in J \subset \mathbb{R},$$

où $J =]a', b'[$ désigne un intervalle ouvert non vide de $\mathbb{R}$. Par la suite, nous nous limiterons essentiellement au cas où (il existe une version de) f (qui est) continue sur J, et désignerons par $I = [a, b] \subset J$ un intervalle tel que $-\infty < a' < a < b < b' < \infty$.

Une partie importante de mes travaux concerne l'*estimation* de f à partir de l'échantillon de taille $n \geq 1$ de X composé des n premières observations $X_1, \ldots, X_n$ de la suite. Il s'agit ici d'*estimation fonctionnelle non-paramétrique*, du fait qu'on construit des statistiques :

$$f_n(x) = f_n(x; X_1, \ldots, X_n),$$

dont l'intérêt est de converger vers la densité inconnue $f(x)$ lorsque la taille n de l'échantillon tend vers l'infini, et ceci, sous des hypothèses très générales sur f (comme celle, par exemple, consistant à supposer que f est continue), et ne requiérant pas d'admettre a priori que f est donnée par un *modèle paramétrique* du type $f(x) = f(x;\theta)$, où $\theta \in \mathbb{R}^p$, et où la forme fonctionnelle $f(.;.)$ est connue.

L'une des statistiques non-paramétrique les plus utilisées pour estimer f est l'*estimateur à noyau* (Rosenblatt (1956), Parzen (1962)), défini, à partir d'un *noyau* $K(\bullet)$, et d'un *paramètre de lissage* h_n, *par*

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right) \text{ pour } x \in \mathbb{R}.$$

- Le *noyau* $K(\bullet)$ est, par hypothèse, une fonction réelle de variable réelle, vérifiant des propriétés minimales de régularité (on supposera typiquement que $K(\bullet)$ est à variation bornée, et s'annulant en dehors d'une partie compacte de $\mathbb{R}$). On suppose également que $K(\bullet)$ vérifie l'égalité

$$\int_{-\infty}^{\infty} K(t)dt = 1,$$

cette dernière condition impliquant que

$$\int_{-\infty}^{\infty} f_n(t)dt = 1.$$

La classe des estimateurs à noyaux comprend, comme cas particulier l'*histogramme*, obtenu lorsque $K(.)$ est défini par

$$K(u) = 1_{\left\{u \in \left[-\frac{1}{2}, \frac{1}{2}\right]\right\}}.$$

- Le *paramètre de lissage* $h_n > 0$ est à choisir de manière que

$$h_n \to 0 \text{ et } nh_n \to \infty, \text{ lorsque } n \to \infty.$$

Ces conditions sont nécessaires et suffisantes pour la convergence en norme *L₁* de f_n vers f lorsque $n \to \infty$, pour tout choix de f, ceci signifiant que, pour tout intervalle compact

$$I = [a,b] \subset J \subset \mathbb{R},$$

et pour toute densité de probabilité f, on a, lorsque $n \to \infty$,

$$\int_I | f_n(x) - f(x) | dx \to 0.$$

Par la suite, nous supposerons, plus spécifiquement que

$$h_n \downarrow 0 \text{ et } nh_n \uparrow \infty, \text{ lorsque } n \uparrow \infty.$$

Mes travaux sur ce type d'estimateur ont débuté en 1972-74, et ont d'abord cherché à résoudre le problème (alors ouvert) de caractériser la convergence presque sûre *uniforme*, ainsi que la convergence presque sûre *ponctuelle* de f_n vers f lorsque $n \to \infty$. Ces deux questions recueillaient un intérêt constant de la part de nombreux chercheurs depuis le debut des années 1960, et on ne disposait alors que de résultats très partiels. J'ai résolu ce problème dans [9], en montrant que, indépendamment de K, il est *nécessaire et suffisant* pour la convergence presque sûre [p.s.] de f_n vers f lorsque lorsque $n \to \infty$, $\forall f$ continue, que (en *notant* $x_0 \in$ R un point fixé à l'avance),

$$nh_n / \log\log n \to \infty,$$

pour la convergence (ponctuelle) p.s. de $f_n(x_0)$ vers $f(x_0)$ lorsque $n \to \infty$, et

$$nh_n / \log n \to \infty,$$

pour la convergence p.s., uniforme sur tout compact de f_n vers f lorsque lorsque $n \to \infty$.

Ces recherches se sont poursuivies dans une série d'articles, portant, entre autres, sur le choix de h_n ([11], [12], [17], [21]), les estimateurs séquentiels et apparentés ([3], [4], [5], [18]), et, plus récemment, pour l'obtention de propriétés du deuxième ordre, consistant à déterminer les vitesses de convergence exactes de f_n vers f. Ces derniers résultats ont été basés sur la technique nouvelle des *lois limites fonctionnelles locales* pour le processus empirique, technique que j'ai développée dans une série d'articles à partir de 1990. Ces lois seront discutées plus en détail au §5. Les *lois limites fonctionnelles locales* sont inspirées de la loi fonctionnelle du logarithme itéré démontrée par Strassen (1964) pour le processus de Wiener, et sont exposées, notamment dans mes articles [91], [94], [96], [107], [108], [117], [120], [125]. J'y établis, entre autres, les théorèmes suivants, cités dans (A)-(B) ci-dessous. On suppose que $a < b$ sont des constantes réelles fixées.

(A) Pour tout $x_0 \in \mathbb{R}$ fixé, sous l'hypothèse que f est continue en x_0, et en supposant que

$$h_n \downarrow 0, \quad nh_n \uparrow \text{ et } nh_n / \log\log n \to \infty \text{ lorsque } n \to \infty,$$

on , presque sûrement,

$$\liminf_{n\to\infty} \left(\frac{nh_n}{2\log\log n} \right)^{1/2} \left\{ \pm \left(f_n(x_0) - \mathbb{E} f_n(x_0) \right) \right\} = \left(f(x_0) \int_{-\infty}^{\infty} K^2(t)dt \right)^{1/2} \text{ p.s.}$$

Ici, "$\pm$" signifie que la formule est vraie aussi bien dans le cas "+" que dans le cas "-".

Ce résultat a été généralisé dans [125] au cas d'une estimation de la densité pour des *lois de survie*. On considère alors le cas où les observations sont des *données censurées*, traitées par la méthode de Kaplan-Meier. Ce problème sera discuté plus loin dans ce même paragraphe. La loi limite présentée dans ce dernier article [125] permet de traiter le cas où la densité f est *discontinue* en x_0, sous réserve qu'elle ait des limites à gauche et à droite en x_0, (éventuellement distinctes). La version multivariée de ce théorème, correspondant à des observations $X_1, \ldots, X_n$ à valeurs dans $\mathbb{R}^p$ pour $p \geq 1$ quelconque, a été obtenue dans [117]. Le comportement de f_n dans le cas où $nh_n / \log\log n \to c < \infty$ a été également traité dans les articles [91], [120], [126] et [138].

(B) Sous l'hypothèse que f est continue sur un voisinage ouvert de $[a,b]$, et que, pour une constante (éventuellement infinie) $c \geq 0$,

$$h_n \downarrow 0,\ nh_n \uparrow,\ nh_n / \log n \to \infty,\ \left(\log(1/h_n) \right) / \log\log n \to c \in [0, +\infty] \text{ lorsque } n \uparrow \infty,$$

alors, on a (cf. [108], [138])

$$\limsup_{n\to\infty} \left(\frac{nh_n}{2\{\log(1/h_n) + \log\log n\}} \right)^{1/2} \sup_{a \leq x \leq b} \left\{ \pm \left(f_n(x) - \mathbb{E} f_n(x) \right) \right\}$$

$$= \left(\sup_{a \le x \le b} f(x) \int_{-\infty}^{\infty} K^2(t)dt \right)^{1/2} \quad \text{p.s.}$$

et

$$\liminf_{n\to\infty} \left(\frac{nh_n}{2\left\{\log\left(1/h_n\right)+\log\log n\right\}} \right)^{1/2} \sup_{a\le x\le b} \left\{\pm\left(f_n(x) - \mathbb{E} f_n(x)\right)\right\}$$
$$= \left(\frac{c}{c+1} \sup_{a \le x \le b} f(x) \int_{-\infty}^{\infty} K^2(t)dt \right)^{1/2} \quad \text{p.s.}$$

Récemment (cf. [141], [144]) j'ai montré que ces propriétés demeurent valides lorsqu'on remplace $[a,b]$ par un intervalle non borné au voisinage duquel *f* est uniformément continue, le noyau *K* pouvant, quant à lui, être une fonction à variation bornée quelconque sur $\mathbb{R}$, au support non nécessairement borné. Ces derniers travaux fournissent des extensions, ne pouvant plus être améliorées, de résultats anciens de Hall (1991) (ces derniers, pour permettre l'utilisation de noyaux à supports non bornés, établissaient des versions sous-optimales de ces théorèmes, imposant, notamment l'existence de constantes $0 < c_1 < c_2 < 1$, telles que $n^{-c_1} < h_n < n^{-c_2}$ pour les grandes valeurs de *n*.

Les *lois limites fonctionnelles locales* qui m'ont permis d'établir ces résultats seront exposées plus en détail au §5 ci-dessous.

Plusieurs de mes travaux, de 1973 à 1979, ont concerné les *estimateurs séquentiels* de la densité, c'est à dire, pouvant être calculés sous la forme récursive

$$\widehat{f}_n(x) = R_n\left(\widehat{f}_{n-1}(x), X_n\right).$$

En particulier, j'ai établi, dans [5], l'optimalité asymptotique des estimateurs de la forme :

$$\widehat{f}_n(x) = \left(\sum_{i=1}^{n} h_i H(h_i) \right)^{-1} \sum_{i=1}^{n} H(h_i) K\left(\frac{x - X_i}{h_n} \right),$$

pour le choix de $H(u) = 1/u$ (estimateur de Yamato-Wolverton-Wagner), dans le cas du critère d'optimalité du IMSE (Integrated Mean Square Error), et de $H(u) = 1$, pour le critère d'optimalité de la variance minimale asymptotique (voir aussi [4] et [18]).

J'ai établi dans [85] (en collaboration avec L. Devroye et M. Broniatowski) que le choix de h_n fourni par la *méthode de la validation croisée* donnait un estimateur de la densité convergeant dans le seul cas où les extrêmes de l'échantillon étaient *stables* [une suite aléatoire Z_n est dite *stable* s'il existe des constantes a_n telles que

$$Z_n - a_n \overset{\mathbb{P}}{\to} 0$$

(convergence en probabilité vers 0 lorsque $n \to \infty$)]. Ce résultat, qui lie le comportement limite d'estimateurs fonctionnels aux proprietés asymptotiques des e*xtrêmes* et des *espacements*, s'inscrit dans une logique conceptuelle expliquant les raisons qui m'ont poussé à traiter des questions en apparence éloignées les unes des autres. Leur impact a été important dans la mesure où la méthode de la validation croisée était, à l'époque de ces travaux, utilisée de plus en plus par les praticiens, sur des bases empiriques non fondées, et sans qu'on sache pour autant caractériser les situations où elle était inadaptée.

Depuis 1992, j'ai entrepris l'étude d'estimateurs de la densité pour des données de survie censurées (qui permet d'estimer le *taux de mortalité*, appellé aussi *taux de panne* ou *taux de hasard* suivant les applications). Le modèle de censure aléatoire considéré est le suivant. Etant donné une suite de durées de vie $X_1,...,X_n$, indépendantes de même loi, et une suite de *temps de censure* $Y_1,...,Y_n$ indépendants de même loi, on n'observe, pour *i*=1,…,*n*, que les valeurs de $Z_i = \min\{X_i, Y_i\}$ et de l'indicatrice $\delta_i = 1_{\{X_i \le Y_i\}}$, permettant de savoir si l'observation X_i est *censurée* ou non par la variable Y_i. À titre d'exemple, si on étudie le temps de survie de patients hospitalisés à la suite d'accidents graves, la *censure* correspond au cas où l'observation se termine à la fin de l'hospitatisation, et lorsque le patient est alors encore en vie. Dans ce cas, pour le $i^{\text{ème}}$ patient, X_i désigne la durée de vie du patient (non observée), et Y_i, sa durée d'hospitalisation (observée).

Le problème principal à résoudre est celui d'estimer la loi de *probabilité de survie* $1-\mathbb{F}(x) = \mathbb{P}(X_i > x)$, la *densité de survie* $f(x) = \mathbb{F}'(x)$, et le *taux de mortalité* (ou *taux de panne*) $f(x)/(1-\mathbb{F}(x))$, indépendamment de la loi commune inconnue $\mathbb{G}(y) = \mathbb{P}(Y_i \le y)$ des temps de censure. On raisonne en général en supposant *f* et $\mathbb{G}$ continues.

L'estimateur le plus classique de $\mathbb{F}(x)$, dû à Kaplan et Meier (1958), est défini par

$$\mathbb{F}_n^*(x) = 1 - \prod_{i:1\le i\le n, Z_i \le x} \left[1 - \frac{1}{N_n(Z_i)}\right]^{\delta_i},$$

où $\delta_i = 1_{\{X_i \le Y_i\}}$, et

$$N_n(z) = \sum_{i=1}^{n} 1_{\{Z_i \ge z\}}.$$

On construit à partir de $\mathbb{F}_n^*$ un *estimateur à noyau* de *f*, en posant

$$f_n^*(x) = \int_{-\infty}^{\infty} h_n^{-1} K\left(\frac{x-t}{h_n}\right) d\mathbb{F}_n^*(t),$$

où $K(.)$ est un noyau du même type que ceux qui sont utilisés pour l'estimation de la densité "classique", comme dans la définition de f_n donnée plus haut (au début de ce même §1). On notera d'ailleurs que, dans le cas d'observations non censurées, correspondant à des temps de censure infinis $Y_1 = ... = Y_n = \infty$, on a, presque sûrement, $\mathbb{F}_n^* = \mathbb{F}_n$ et $f_n^* = f_n$, où $\mathbb{F}_n$ désigne la fonction de répartition empirique de $X_1,...,X_n$, définie par

$$\mathbb{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \le x\}}.$$

De ce fait, les résultats de convergence obtenus pour les données censurées contiennent comme cas particuliers les propriétés analogues du cas classique non censuré.

Récemment, j'ai pu établir dans [125] (en collaboration avec J. H. J. Einmahl), que si

$$h_n \downarrow 0,\ nh_n \uparrow,\ nh_n / \text{loglog}\, n \to \infty, \text{ lorsque } n \uparrow \infty,$$

alors, pour tout $x_0 \in \mathbb{R}$ tel que $1 - G(x_0) \neq 0$,

$$\limsup_{n\to\infty} \left(\frac{nh_n}{2\text{loglog}\, n} \right)^{1/2} \left\{ \pm \left(f_n^*(x_0) - \widehat{\mathbb{E}}\, f_n^*(x_0) \right) \right\} = \left(\frac{f(x_0)}{1-G(x_0)} \int_{-\infty}^{\infty} K^2(t)dt \right)^{1/2} \text{ p.s.}$$

où

$$\widehat{\mathbb{E}}\, f_n^*(x_0) = \int_{-\infty}^{\infty} h_n^{-1} K\left(\frac{x_0 - t}{h_n} \right) dF(t).$$

Comme mentionné plus haut, ces résultats couvrent également le cas des estimateurs de la densité usuels, dans la mesure où $f_n(x) = f_n^*(x)\ \forall x$, lorsque $G(x) = 1\ \forall x$, choix qui correspond à des données non censurées. La version uniforme de ce résultat,décrivant le comportement de $f_n^*(x)$ lorsque $x \in [a,b]$ varie sur un intervalle non réduit à un point, est la suivante (cf. [138]).

-- Sous l'hypothèse que f est continue sur un voisinage ouvert de $[a,b]$, avec $G(b) < 1$, et supposant que

$$h_n \downarrow 0,\ nh_n \uparrow,\ nh_n / \log n \to \infty,\ \left(\log(1/h_n)\right)/\text{loglog}\, n \to c \in [0,\infty], \text{ lorsque } n \uparrow \infty,$$

on a

$$\limsup_{n\to\infty} \left\{ \frac{nh_n}{2\left\{\log(1/h_n) + \text{loglog}\, n\right\}} \right\}^{1/2} \left\{ \sup_{a \le x \le b} \pm \left(f_n^*(x) - \widehat{\mathbb{E}}\, f_n^*(x) \right) \right\}$$

$$= \left(\left\{ \sup_{a \le x \le b} \frac{f(x)}{1-G(x)} \right\} \int_{-\infty}^{\infty} K^2(t)dt \right)^{1/2} \text{ p.s.},$$

et

$$\liminf_{n\to\infty} \left\{ \frac{nh_n}{2\left\{\log(1/h_n) + \text{loglog}\, n\right\}} \right\}^{1/2} \left\{ \sup_{a \le x \le b} \pm \left(f_n^*(x) - \widehat{\mathbb{E}}\, f_n^*(x) \right) \right\}$$

$$= \left(\frac{c}{c+1} \left\{ \sup_{a \le x \le b} \frac{f(x)}{1-G(x)} \right\} \int_{-\infty}^{\infty} K^2(t)dt \right)^{1/2} \text{ p.s.},$$

Dans [110], j'ai également construit des tests d'homogénéité basés sur plusieurs échantillons censurés, et établi les propriétés en permettant l'application pratique. Ces tests permettent de comparer entre elles les distributions de survie de chacun des groupes, *indépendamment des lois des temps de censure*. Ils sont aujourd'hui régulièrement utilisés sur des ensembles de données biologiques et médicales, afin de vérifier si ces derniers peuvent être agrégés pour constituer des échantillons plus importants, permettant des estimations plus précises des paramètres d'intérêt.

L'étude des estimateurs fonctionnels non paramétriques m'a amené à faire des recherches dans plusieurs domaines connexes: les *valeurs extrêmes*, les *processus empiriques*, les *approximations fortes* et les *espacements*. De manière un peu inattendue, ces domaines se trouvent être inextricablement liés sur le plan méthodologique. Ce fait, déjà mentionné en liaison avec [85], sera illustré plus loin par d'autres exemples.

4.b. Valeurs extrêmes et lois multivariées.

4.b.a. Lois fortes pour les statistiques d'ordre.

Soient $X_{1,n} \leq \ldots \leq X_{n,n}$ les *statistiques d'ordre*, obtenues en rangeant par ordre croissant les $n \geq 1$ premières observations d'une suite $X_1, \ldots, X_n$ de variables aléatoires indépendantes de même loi. Plusieurs de mes travaux ont décrit le comportement limite de telles statistiques d'ordre, ou concernent l'étude de processus aléatoires qui leur sont liés.

En particulier, mes premiers travaux de recherche ont été consacrés à l'étude des trajectoires du *premier temps de passage* $\tau_t = \inf\{n \geq 1 : X_{n,n} \geq t\}$, à un niveau *t* donné de la suite des *maxima partiels* $\{X_{n,n} : n \geq 1\}$ de $X_1, X_2, \ldots$, (on remarquera que le maximum des *n* premières observations est $X_{n,n} = \max\{X_1, \ldots, X_n\}$). Dans [1] et [2], j'ai établi la propriété que le processus $\{\tau_t : -\infty < t < \infty\}$ était *à accroissements indépendants*.

En faisant usage de ce résultat, j'ai pu établir, dans [1], en 1971, le théorème limite suivant, qui résoud un problème ouvert, posé par Grenander en 1965.

Si $U_1, U_2, \ldots$, désigne une suite de variables aléatoires indépendantes de même loi uniforme sur (0,1), alors

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^{n} \min\{U_1, \ldots, U_n\} = 1 \text{ p.s.}$$

J'ai ensuite généralisé ces résultats au processus $\tau_t^{(k)} = \inf\{n \geq k \ : \ X_{n-k+1,n} \geq t\}$, des premiers temps de passage à un niveau *t* donné, de la suite des $k^{\text{èmes}}$ maxima partiels $\{X_{n-k+1,n} : n \geq k\}$ de $X_1, X_2, \ldots$, lorsque $k \geq 1$ est un entier fixé. Ces travaux m'ont mené à l'obtention d'encadrements presque sûrs pour les processus de premier temps de passage $\{\tau_t^{(k)} : -\infty < t < \infty\}$. Par un procédé simple d'inversion, il est possible d'en déduire des résultats analogues pour chacune des suites $\{X_{n-k+1,n} : n \geq k\}$ correspondant à des valeurs fixées de $k = 1, 2, \ldots$. Cependant, cette dernière partie de mes recherches fut rendue partiellement caduque par des résultats de Kiefer (1972) et Robbins et Siegmund (1972) pour *k*=1, puis par ceux de Shorack et Wellner (1978), pour $k \geq 1$ fixé. Ces derniers auteurs ont, en effet, obtenu, par des techniques différentes, des conditions nécessaires et suffisantes portant sur une suite $\{c_n : n \geq 1\}$, pour que

$$\mathbb{P}\left(X_{n-k+1,n} \geq c_n \ \text{ i.s.}\right) = 0 \ \text{(resp. 1), et } \mathbb{P}\left(X_{n-k+1,n} \leq c_n \ \text{ i.s.}\right) = 0 \ \text{(resp. 1).}$$

Ici l'abréviation $\{A_n \ \text{ i.s.}\}$ signifie que la suite d'événements A_n a lieu « *infiniment souvent* ». Il est possible de montrer (loi du 0 ou 1) que, pour des événements A_n tels ceux qui sont considérés ici, $\mathbb{P}(A_n \ \text{ i.s.})$ ne peut prendre que la valeur 0 ou 1.

Dans le but d'obtenir des caractérisations plus puissantes que celles qui avaient déjà été obtenues, je me suis alors attaqué au problème, plus complexe, de l'encadrement presque sûr de $\{X_{n-k_n+1,n} : n \geq 1\}$ lorsque $\{k_n : n \geq 1\}$ est une suite monotone d'entiers, non nécessairement constante, et vérifiant la condition $n^{-1}k_n \to 0$ lorsque $n \to \infty$. J'ai pu résoudre *intégralement* ce problème en montrant dans [60] et [81] que, sous réserve de conditions générales de croissance et de régularité portant sur $\{k_n : n \geq 1\}$, les résultats suivants sont vérifiés. Je donne ci-dessous la version correspondant à des variables aléatoires $U_1, U_2, \ldots$ indépendantes et de loi uniforme sur (0,1). Cette restriction n'implique d'ailleurs aucune perte de généralité, puisqu'on peut écrire l'identité $X_{i,n} = Q(U_{i,n})$, pour $1 \leq i \leq n$, où la *transformation de quantiles* $Q(.)$ est définie par $Q(t) = \inf\{x : F(x) \geq t\}$ pour $0<t<1$, lorsque $F(x) = \mathbb{P}(X_1 \leq x)$.

Soient $U_{1,n} < \ldots < U_{n,n}$ la *statistique d'ordre* de $U_1, \ldots, U_n$, obtenue en rangeant ces observations par ordre croissant. Ici, $U_1, U_2, \ldots$, comme ci-dessus, désigne une suite de variables aléatoires indépendantes de même loi uniforme sur (0,1). J'ai établi que :

1) Si $(c_n - k_n)/\sqrt{k_n} \ \to \ \infty$, alors

$$\mathbb{P}\left(U_{k_n,n} \geq n^{-1}c_n \ i.s.\right) = 0 \ \text{(resp. } =1)$$

$$\Leftrightarrow \ \sum_{n=1}^{\infty} \frac{1}{n}\left(e\frac{c_n}{k_n}\right)^{k_n} \exp\left(-c_n\right) < \infty \ \text{(resp. } =\infty).$$

2) Si $(c_n - k_n)/\sqrt{k_n} \ \to \ -\infty$, alors

$$\mathbb{P}\left(U_{k_n,n} \leq n^{-1}c_n \ i.s.\right) = 0 \ \text{(resp. } =1)$$

$$\Leftrightarrow \ \sum_{n=1}^{\infty} \frac{1}{n}\left(e\frac{c_n}{k_n}\right)^{k_n} \exp\left(-c_n\right) < \infty \ \text{(resp. } =\infty).$$

Ces résultats ([60], [81]) sont optimaux au sens qu'ils ne pourraient être éventuellement améliorés qu'en affaiblissant les conditions de régularité et de monotonie supposés sur k_n. Ils comprennent comme cas particuliers les critères de Barndorff-Nielsen (1961), de Robbins et Siegmund (1972), Kiefer (1972) et de Shorack et Wellner (1978). Enfin, ils présentent la particularité de faire intervenir la *même* série pour caractériser les encadrements inférieurs et supérieurs.

Mes recherches, portant sur les lois fortes de statistiques d'ordre (ou de quantiles) ont été poursuivies dans plusieurs directions, comprenant l'étude des *statistiques de queue*, des *processus empiriques de queue*, des *statistiques locales*, des *processus empiriques locaux* et des *records*.

4.b.b. Records.

Les *records classiques*, pour une suite de variables aléatoires $X_1, X_2, \ldots$ sont définis comme suit. Soit $X_{1,n} \leq \ldots \leq X_{n,n}$ la statistique d'ordre de $X_1, \ldots, X_n$, obtenue en rangeant ces observations par ordre croissant. Posons, par convention, $X_{0,0} = -\infty$. Pour chaque valeur de l'entier $j \geq 1$, le $j^{\text{ème}}$ *temps de record* (ou *temps de* 1-*record*) $n_j^{(1)}$ est la plus petite valeur de l'entier $n \geq 1$ pour laquelle on observe, pour la $j^{\text{ème}}$ fois consécutive, l'événement $X_{m,m} > X_{m-1,m-1}$, pour *m* variant de 1 à n. La $j^{\text{ème}}$ *valeur de record* (ou *valeur de* 1-*record*) $R_j^{(1)}$ désigne alors la valeur prise par $X_{n,n}$ pour la valeur de *n* donnée par $n = n_j^{(1)}$.

On définit, pour $k \geq 1$ fixé, la suite $n_1^{(k)} = k < n_2^{(k)} < \ldots < n_j^{(k)} < \ldots$ des *temps de k-records* comme l'ensemble ordonné des indices $n \geq k$ tels que $X_{n-k+1,n} > X_{n-k,n-1}$, et la $j^{\text{ème}}$ *valeur de k-records*, $R_j^{(k)}$, comme la valeur obtenue en posant $R_j^{(k)} = X_{n-k+1,n}$ pour la valeur de *n* correspondant à $n = n_j^{(k)}$.

L'étude de la suite double des temps et valeurs de *k*-records $\left\{\left(n_j^{(k)}, R_j^{(k)}\right) : j \geq 1\right\}$, qui se ramène
à l'étude de la trajectoire du $k^{\text{ème}}$ maximum de la suite $\left\{X_{n-k+1,n} : n \geq k\right\}$, a fait l'objet de très nombreux travaux, initiés par les recherches originales de Chandler (1952) et Rényi (1970). Lorsque les observations $X_1, X_2, \ldots,$ sont indépendantes et de même loi, de fonction de répartition continue $F(x) = \mathbb{P}\left(X_1 \leq x\right)$, la loi des temps de records ne dépend pas de *F*, tandis que celle des valeurs de records ne dépend de *F* que par le biais d'un changement d'échelle. Il est alors commode de se ramener au cas de la *loi exponentielle standard,* pour laquelle $F(x) = \mathbb{P}\left(X_1 \leq x\right) = 1 - e^{-x}$ pour $x \geq 0$. Nous supposerons par la suite que cette dernière hypothèse est satisfaite.

Dans la période 1981-1986, j'ai pu établir des résultats définitifs ([29], [32], [33], [35], [39], [40], [49], [50], [73]) sur l'*approximation forte* optimale des temps de records et des valeurs de records, à l'aide de processus de Wiener. C'est ainsi, entre autres, que j'ai montré que, pour $k \geq 1$ fixé, il était possible de construire (sur un espace de probabilités convenable) un processus de Wiener (ou mouvement Brownien unidimensionnel) standard $\{W(t) : t \geq 0\}$ tel que la suite des temps de *k*-records, $\left\{n_j^{(k)} : j \geq 1\right\}$, vérifie

$$n_j^{(k)} = \exp\left(\frac{j}{k} + \frac{1}{k}W(j) + O\left(\log j\right)\right) \quad \text{p.s., lorsque } j \to \infty.$$

Ce résultat est obtenu avec une vitesse d'approximation *optimale*, au sens que le "$O(\log j)$" ne peut être remplacé par un "$o(\log j)$". Il permet de ramener l'étude des fluctuations des temps de *k*-records à celle des fluctuations du processus de Wiener.

En 1983, j'ai démontré (cf. [40]) (sous l'hypothèse d'une loi exponentielle standard, comme ci-dessus, avec $F(x)=\mathbb{P}(X_1 \le x)=1-e^{-x}$ pour $x>0$) que les processus ponctuels $\{N_k(t):t\ge 0\}$, $k=1,2,\ldots,$ ayant pour temps d'arrivée $\{R_j^{(k)}:j\ge 1\}-\{R_j^{(k-1)}:j\ge 1\}$, avec la convention que $\{R_j^{(0)}:j\ge 1\}=\varnothing$, composent, $k=1,2,\ldots,$ une suite de processus de Poisson *indépendants de même loi*. J'ai appris ensuite que ce résultat avait été annoncé indépendamment, mais sans démonstration par Ignatov (1978). J'en ai donc obtenu la première démonstration publiée. Depuis, ce théorème a été établi à nouveau par d'autres méthodes (voir par ex. Goldie et Rogers (1984), Vervaat, (1986)).

Dans [51], j'ai montré que, l'indépendance des variables de la suite $\{R_j^{(k)}-R_{j-1}^{(k)}:j\ge 2\}$, pour une valeur de $k\ge 1$ donnée, caractérisait la loi exponentielle (au sens que cette propriété équivaut à ce que l'on ait $\mathbb{P}(X_1\le x)=1-\exp(-\lambda(x-\theta))$ $\forall x\ge\theta$, pour des valeurs convenables de $\lambda>0$ et $\theta\in\mathbb{R}$). Ceci m'a permis d'unifier tout un ensemble de résultats disparates déjà établis sur le sujet dans la littérature scientifique consacrée à la caractérisation des lois de probabilité.

Plus récemment, j'ai obtenu dans [112] et [115] (en collaboration avec V. B. Nevzorov), les extensions de ces résultats d'approximation forte à un contexte considérablement plus général, et ceci, par une nouvelle technique. C'est ainsi que j'ai démontré des théorèmes d'approximation forte, d'une part, pour les $k_n^{\text{èmes}}$ *temps de record*, correspondant au remplacement de l'entier $k\ge 1$ fixé ci-dessus par une suite d'entiers $\{k_n:n\ge 1\}$, et d'autre part, dans le cadre (voir [131]) de records basés sur des observations $\{X_n:n\ge 1\}$, indépendantes, mais dont la loi $F_n(x)=\mathbb{P}(X_n\le x)$ varie avec *n*. On suppose ici que F_n est de la forme $F_n=F^{\alpha_n}$, où *F* désigne une loi de répartition fixe, et $\{\alpha_n:n\ge 1\}$, une suite d'entiers positifs. Il s'agit du "F^α - *scheme*", introduit par V. B. Nevzorov (1986).

Ce modèle est le seul pour lequel les indicatrices $1_{\{X_n>X_{n-1,n-1}\}}$, des événements que n est un temps de record, composent une suite de variables aléatoires *indépendantes*. Pour mémoire, le fait que les indicatrices des temps de record $1_{\{X_n>X_{n-1,n-1}\}}$ soient indépendantes dans le cas d'une suite $\{X_n, n\ge 1\}$ de variables aléatoires indépendantes de même fonction de répartition continue *F*, a été pour la première fois mis en évidence par A. Rényi (1962). Cette dernière situation correspond au cas particulier de "F^α - *scheme*" obtenu pour $\alpha_n=1,\ \forall n\ge 1.$ Dans [136] (avec V. B. Nevzorov) j'étudie des propriétés de rééchantillonnage par *bootstrap* de la suite de ces records.

4.b.c. Extrêmes Multivariés.

Dans une série d'articles publiés à partir de 1978, je me suis intéressé aux *lois limites pour les extrêmes multivariés*, ce qui revient à étudier, pour $p \geq 1$ fixé, la convergence en loi lorsque $n \to \infty$, du vecteur aléatoire

$$V_n = \left(\frac{Y_n(1) - b_n(1)}{a_n(1)}, \ldots, \frac{Y_n(p) - b_n(p)}{a_n(p)} \right),$$

où, pour $j = 1, \ldots, p$, $Y_n(j) = \max\{X_1(j), \ldots, X_n(j)\}$, et $\{(X_n(1), \ldots, X_n(p)) : n \geq 1\}$ désigne une suite de vecteurs aléatoires indépendants de même loi à valeurs dans $\mathbb{R}^p$. Pour chaque choix de $j = 1, \ldots, p$, $a_n(j) > 0$ et $b_n(j)$, pour $n = 1, 2, \ldots$, désignent des suites de constantes réelles choisies de telle sorte que V_n converge en loi, lorsque $n \to \infty$, vers une loi limite dans $\mathbb{R}^p$ à marges non dégénérées. Le problème est alors de déterminer la structure des lois limites ainsi obtenues.

J'ai caractérisé, dans [13], [15], [16], [23], [48], l'ensemble de toutes les lois limites possibles pouvant être générées par ce modèle. Ces résultats ont généralisé au cas $p \geq 3$ des travaux antérieurs de Geffroy (1958), Sibuya (1960) et Tiago de Oliveira (1961), obtenus dans le cas particulier de $p = 2$. Le résultat pour $p = 1$ avait été décrit bien antérieurement, dans un ensemble de travaux, dûs à Fréchet, Weibull et Gendenko, et réalisés de 1923 à 1943.

La structure particulière de ces lois limites multivariées a toutes sortes de propriétés remarquables. L'étude de celles-ci m'a permis, notamment, d'introduire, dans [38], les processus appellés depuis *max-stables.*

Ces recherches m'ont amené à étudier de manière approfondie les *copules* ou fonctions de dépendance. Il semble, en particulier, que je sois le premier à avoir démontré, dans [14], l'existence, quelle que soit la fonction de répartition multivariée $F(x_1, \ldots, x_p)$ admettant pour marges $F_j(x_j) = F(\infty, \ldots, \infty, x_j, \infty, \ldots, \infty)$, pour $j = 1, \ldots, p$, d'une *copule* $C(u_1, \ldots, u_p)$, fonction de répartition d'un vecteur aléatoire à marges uniformes sur $[0,1]$, telle que

$$C(F_1(x_1), \ldots, F_p(x_p)) = F(x_1, \ldots, x_p),$$

en tout point de continuité. Ce résultat est loin d'être trivial lorsque F et $p \geq 2$ sont quelconques. Il est naturel d'utiliser également une telle construction lorque F est la fonction de répartition empirique d'un échantillon. Ceci sera évoqué plus loin dans le §2d.

L'estimation des lois extrêmes multivariées a fait l'objet des articles [22], [48], [86] (en collaboration avec J. Tiago de Oliveira), et [97], où je décris le comportement limite d'estimateurs de la loi jointe d'extrêmes bivariés, dûs à Pickands (1981). Ces recherches sont actuellement développées afin d'appliquer de tels résultats limites pour obtenir la loi de probabilité de *tests d'ajustement* du type Cramér-Von Mises ([125], [146]) (voir le §2d).

4.b.d. Copules et tests d'indépendance.

La *fonction de dépendance empirique* (ou *copule empirique*) est une fonction de répartition C_n dans $\mathbb{R}^d$, à lois marginales uniformes sur $[0,1]$, qui vérifie l'identité

$$C_n(F_{n,1}(x_1),\dots,F_{n,d}(x_d)) = F_n(x_1,\dots,x_d),$$

où F_n désigne la fonction de répartition empirique d'un échantillon de taille *n* dans $\mathbb{R}^d$, de fonctions de répartition marginales

$$F_{n,j}(x_j) = F_n(\infty,\dots,\infty,x_j,\infty,\dots,\infty), \quad \text{pour} \quad j=1,\dots,d.$$

J'ai entrepris dans [13], [14], [19], [20], [22], [24], [27], [28] l'étude des propriétés limites de C_n. Un exemple des résultats ainsi obtenus est la détermination, dans [26], de la loi exacte, pour $d \geq 1$ quelconque, de la *statistique de Cramér-Von Mises multivariée*, définie par

$$T_n^2 = \int_{\mathbb{R}^d} \left(F_n(x_1,\dots,x_d) - F(x_1,\dots,x_d)\right)^2 dF_n(x_1,\dots,x_d),$$

où $F(x_1,\dots,x_d)$ désigne la fonction de répartition exacte de la variable générique dans $\mathbb{R}^d$, qui engendre un échantillon de taille $n \geq 1,$ lequel a pour fonction de répartition empirique $F_n(x_1,\dots,x_d)$.

La caractérisation de la loi limite de T_n^2 lorsque $n \to \infty$ n'était jusque là connue que pour *d*=1, (Von Mises (1933)), et pour *d*=2 (Blum, Kiefer et Rosenblatt (1961)). Plus récemment, des travaux entrepris avec G. Martynov, m'ont permis d'établir des tabulations précises des lois asymptotiques de statistiques type Cramér-Von Mises ([125]) ces dernières étant d'un intérêt tout particulier pour les données multivariées (cf. §2c).

Ces recherches ont permis d'obtenir des résultats assez spectaculaires dans l'étude du processus défini par

$$Z_n(u) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\min\left\{ \frac{X_i / \overline{X}_n}{u}, \frac{Y_i / \overline{Y}_n}{1-u} \right\} - 1 \right) \quad \text{pour} \quad 0 \leq u \leq 1,$$

avec

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i,$$

et où $\{(X_n, Y_n) : n \geq 1\}$ est une suite de vecteurs aléatoires indépendants, à valeurs dans $\mathbb{R}^2,$ et telles que, pour $n \geq 1,$ et des constantes convenables $\lambda > 0$ et $\mu > 0,$

$$\mathbb{P}(X_n > x, Y_n > y) = \exp(-(\lambda x + \mu y)) \quad \text{pour} \quad x, y > 0.$$

Dans [97] et [146] (pour ce dernier article, en collaboration avec G. Martynov), j'établis la convergence en loi de ce processus vers un processus gaussien centré $\{Z(u) : 0 \leq u \leq 1\}$. J'obtiens, en particulier, la décomposition de Karhunen-Loeve explicite de $\{Z(u) : 0 \leq u \leq 1\}$, en faisant intervenir les valeurs propres de l'équation de Fredholm associée au noyau de covariance. Ces dernières sont données par

$$\lambda_k = \frac{6}{k(k+1)(k+2)(k+3)} \quad \text{pour} \quad k = 1, 2, \dots$$

Les fonctions propres associées s'expriment directement et explicitement à partir de polynômes de Jacobi. Ce résultat permet, entre autres, d'établir la convergence en loi, lorsque $n \to \infty$,

$$\int_0^1 Z_n^2(t)dt \xrightarrow{\mathrm{d}} \int_0^1 Z^2(t)dt = \sum_{k=1}^{\infty} \lambda_k \omega_k^2,$$

où $\{\omega_k : k \geq 1\}$ désigne une suite de variables aléatoires normales $N(0,1)$ standard. Ceci permet de tabuler la distribution correspondante et de construire ainsi un test d'indépendance des paires exponentielles $\{(X_n, Y_n) : n \geq 1\}$. De plus, c'est un des rares exemples connus où une telle décomposition est explicite pour un processus gaussien d'intérêt statistique. Les autres cas concernent le processus de Wiener, le pont brownien, le processus d'Ornstein-Uhlenbeck, et surtout le processus d'Anderson et Darling (1952), défini par $B(u)/\sqrt{u(1-u)}$, où $\{B(u) : 0 \leq u \leq 1\}$ est un pont brownien. Dans ce dernier cas, les valeurs propres sont données par

$$\lambda_k = \frac{1}{k(k+1)} \quad \text{pour} \quad k = 1, 2, ...,$$

et les fonctions propres sont aussi des polynômes de Jacobi. Je travaille actuellement à montrer que ces résultats peuvent être unifiés et généralisés en faisant usage de la théorie de la représentation des groupes, où les polynômes de Jacobi interviennent directement.

4.b.e. Tests d'ajustement.

Dans [34] et [36], j'ai établi que certains tests d'ajustement proposés par E. Parzen en 1979 pour vérifier des hypothèses sur une loi de probabilité, indépendamment de ses paramètres de location et de dispersion, avaient une loi limite basée sur le pont Brownien, si et seulement si la loi de l'échantillon était uniforme, exponentielle, ou exponentielle après un retournement d'échelle. Cette propriété s'exprime sur le pont Brownien $\{B(u) : 0 \leq u \leq 1\}$, en constatant que

$$I(u) = B(u) - \int_0^u B(v)\phi(v)dv + u\int_0^1 B(v)\phi(v)dv$$

est un pont Brownien, *si et seulement si* $\phi(u)$ est égal à 1, $1/u$, $1/(1-u)$, ou une combinaison de ces fonctions.

Dans un tout autre ordre d'idées, j'ai développé la théorie asymptotique des méthodes de comparaison non-paramétriques de deux échantillons $X_1, ..., X_n$ et $Y_1, ..., Y_n$, de fonctions de répartition empiriques $H_n(x) = n^{-1}\#\{X_i \leq x : 1 \leq i \leq n\}$ et $K_n(x) = n^{-1}\#\{Y_i \leq x : 1 \leq i \leq n\}$, et de fonctions de quantiles $H_n^{-1}(s) = \inf\{x : H_n(x) \geq s\}$ et $K_n^{-1}(s) = \inf\{x : K_n(x) \geq s\}$. Ces comparaisons sont basées sur les *P-P* et *Q-Q* plots, ces derniers étant, par définition, les statistiques

$$K_n\left(H_n^{-1}(s)\right) \quad \text{et} \quad K_n^{-1}\left(H_n(s)\right).$$

J'ai obtenu les lois limites de tests basés sur ces statistiques, ainsi que des principes d'invariance forts dans [82] (avec D. M. Mason), [89] (avec J. Beirlant) et [109], pour le processus de Kaplan-Meier (avec J. H. J. Einmahl).

4.c. Approximation de Poisson.

Depuis les travaux fondamentaux de Poisson au XIX[ème] siècle, l'étude de l'approximation de la loi $L(S_n)$ de la somme partielle $S_n = X_1+...+X_n$ d'ordre n d'une suite de variables indépendantes $X_1,X_2,...$, de Bernoulli, telles que, pour n=1,2,...,

$$P(X_n=1) = 1 - P(X_n=0) = p_n \in [0,1],$$

par la loi $\Pi(\mu_n) = L(T_n)$ d'une variable de Poisson T_n, telle que, pour n=1,2,...,

$$P(T_n=k) = \frac{\mu_n^k}{k!}\exp(-\mu_n) \text{ pour } k=0,1,...,$$

où $\mu_n = p_1+...+p_n$, est resté une question classique du calcul des probabilité, d'une importance comparable à celle du théorème central limite (convergence vers la loi de Laplace-Gauss), dont la résolution était, cependant, restée incomplète. Il peut paraître donc surprenant que la détermination précise de la *vitesse exacte de convergence en loi de* $L(S_n)$ vers $L(T_n)$ (la *loi de Poisson*) n'ait pu être déterminée qu'au cours de la décennie 1980-90, et ce, malgré l'ancienneté du problème.

Pour mesurer la distance entre les lois $L(S_n)$ et $L(T_n)$, il est commode d'utiliser des métriques probabilistes telles que la *distance en variation*

$$d_V\left(L(S_n),L(T_n)\right) = \sup_{A\subseteq\mathbb{N}}\left|\mathbb{P}(S_n\in A)-\mathbb{P}(T_n\in A)\right|.$$

J'ai pu apporter, dans une série de travaux ([62], [64], [69], [77], [78], [79], [80], [84], [103]), en collaboration avec A. Karr, R. Serfling, et D. Pfeifer, une solution presque complète au problème de l'évaluation asymptotique et à distance finie de d_V, et ce, par une technique originale, basée sur les *semigroupes d'opérateurs dans les espaces de Banach.* Antérieurement, l'évaluation de d_V, pour le choix de $\mu_n = \mathbb{E}(T_n) = \mathbb{E}(S_n) = p_1+...+p_n$, avait fait l'objet de nombreux travaux. Par exemple, en 1960, L. Le Cam avait obtenu la borne supérieure

$$d_V\left(L(S_n),L(T_n)\right) \le \min\left\{\sum_{i=1}^n p_i^2, \max\left(p_1,...,p_n\right)\right\}.$$

En 1984, Barbour et Hall précisèrent cette borne en obtenant l'inégalité

$$d_V\left(L(S_n),L(T_n)\right) \le \left\{\sum_{i=1}^n p_i\right\}^{-1}\left\{1-\exp\left(-\sum_{i=1}^n p_i\right)\right\}\sum_{i=1}^n p_i^2.$$

J'ai obtenu une évaluation asymptotique exacte de cette distance, en montrant, en particulier, dans les travaux [62] et [80], cités plus haut, que

$$d_V\left(L(S_n),L(T_n)\right) = D_V + r_V.$$

Dans cette expression, le *terme principal* D_V est égal à

$$D_V = \frac{1}{2}\left(\sum_{i=1}^n p_i^2\right)\left(\frac{\mu^{a-1}(a-\mu)}{a!}-\frac{\mu^{b-1}(b-\mu)}{b!}\right),$$

où

$$\mu=\mu_n=\sum_{i=1}^n p_i,\ \ a=\left\lfloor \mu+\frac{1}{2}+\left(\mu+\frac{1}{4}\right)^2\right\rfloor,\ \ b=\left\lfloor \mu+\frac{1}{2}-\left(\mu+\frac{1}{4}\right)^2\right\rfloor,$$

et $\lfloor u\rfloor\le u<\lfloor u\rfloor+1$ désigne la partie entière de *u*.

Le *terme résiduel* r_V de l'expression ci-dessus est asymptotiquement négligeable relativement à D_V, et peut être évalué à l'aide de développements, ou, plus simplement, borné supérieurement. L'une de ces bornes, donnée dans [80] (en collaboration avec D. Pfeifer), montre que

$$|r_V|\le\frac{(2\theta)^{3/2}}{2\sqrt{1-2\theta}}\quad\text{lorsque}\quad \theta=\left(\sum_{i=1}^n p_i^2\right)/\left(\sum_{i=1}^n p_i\right)<\frac{1}{2}.$$

Par exemple, lorsque

$$\left(\sum_{i=1}^n p_i^2\right)/\left(\sum_{i=1}^n p_i\right)\to 0\ \text{ et }\ \mu_n=\sum_{i=1}^n p_i\ \to\infty,$$

ces résultats permettent d'obtenir l'équivalent asymptotique (cf. [62]) remarquable

$$d_V\left(L(S_n),L(T_n)\right)=\frac{1+o(1)}{\sqrt{2\pi e}}\left(\sum_{i=1}^n p_i^2\right)/\left(\sum_{i=1}^n p_i\right),$$

qui améliore du facteur multiplicatif $1/\sqrt{2\pi e}$ la borne de Barbour et Hall (1984).

J'ai également obtenu des évaluations analogues en remplaçant la distance en variation par d'autres métriques mesurant les distances entre lois de probabilité. Parmi celles-ci, il faut mentionner les distances de Kolmogorov, de Fortet-Mourier, et la catégorie générale formée par les distances de Wasserstein.

Dans le cas particulier de la distance de Kolmogorov, définie ici par

$$d_K\left(L(S_n),L(T_n)\right)=\sup_{-\infty<x<\infty}\left|\mathbb{P}(S_n\le x)-\mathbb{P}(T_n\le x)\right|,$$

j'ai montré, dans [80] et [84], (avec D. Pfeifer, M. Puri et S. Ralescu) que, si Z_n désigne une variable aléatoire normale, de paramètres

$$N\left(\sum_{i=1}^n p_i,\sum_{i=1}^n p_i(1-p_i)\right),$$

où $\sum_{i=1}^n p_i$ désigne l'espérance, et $\sum_{i=1}^n p_i(1-p_i)$ la variance, et si

$$\left(\sum_{i=1}^n p_i^2\right)/\left(\sum_{i=1}^n p_i\right)\to 0,\ \ \sum_{i=1}^n p_i\to\infty,\ \text{ et }\ \left(\sum_{i=1}^n p_i^2\right)/\left(\sum_{i=1}^n p_i\right)^{1/2}\to\alpha,$$

on a alors, pour tout *n* suffisamment grand,

$$d_K\left(L(S_n),L(T_n)\right)>d_K\left(L(S_n),L(Z_n)\right),\ \text{resp.}\ d_K\left(L(S_n),L(T_n)\right)<d_K\left(L(S_n),L(Z_n)\right),$$

lorsque α vérifie

$$\alpha>\Gamma\sqrt{2}\ ,\ \text{resp.}\ \ \alpha>\Gamma\sqrt{2}\ ,$$

où Γ=0.2784... est solution de l'équation $1+x+\log x=0$.

Ces résultats permettent de tracer une frontière précise entre le cas ou l'approximation normale est meilleure que l'approximation Poissonienne de celui où l'inverse a lieu (ceci, pour l'approximation de la loi de sommes de variables de Bernoulli vis à vis du critère donné par par la distance de Kolmogorov).

Les méthodes basées sur l'approximation de Poisson fournissent un outil particulièrement puissant *sur le plan théorique*, dans l'étude des *processus empiriques*. D'un point de vue *pratique*, ces techniques sont également extrêmement utiles pour analyser les *événements rares*, en *fiabilité*. Soit, par exemple, un équipement industriel possédant un (grand) nombre *n* de composants indépendants ayant des probabilités individuelles $p_i,\ i=1,\dots,n,$ de défaillance, très petites, dans une période donnée. Le nombre de défaillances pouvant être enregistrées est alors une variable aléatoire S_n qui se trouve être une somme d'indicatrices de Bernoulli comme ci-dessus. Le calcul de la distribution exacte de S_n est très long et difficile lorsque le nombre des composants est élevé, et les p_i non-identiques. Par contre, l'approximation de la loi de S_n par une loi de Poisson est facile, car il suffit de sommer les probabilités individuelles de défaillance pour en obtenir le paramètre.

Bien entendu, une telle approximation ne peut être utilisée que si on en maîtrise bien le terme d'erreur, ce qui motive les recherches précédentes, dont les résultats sont utilisés tant dans des modèles actuariels (par exemple, pour les calculs de primes pour les équipements pétroliers "off-shore"), que dans des modèles industriels (étude des défaillances de grands équipements).

4.d. Processus de sommes partielles et de renouvellement.

Soit $\{X_n : n\geq 1\}$ une suite de variables aléatoires centrées, de variance égale à 1, indépendantes et de même loi. Pour $t\geq 0,$ on désigne par

$$S(t)=\sum\nolimits_{i=1}^{\lfloor t\rfloor}X_i,$$

le processus de sommes partielles associé à cette suite (ici, $\lfloor t\rfloor\leq t<\lfloor t\rfloor+1$ désigne la partie entière de *t,* et $\sum_{\varnothing}(\bullet)=0$). Sous réserve d'existence, au voisinage de 0, de la *fonction génératrice des moments* $\psi(t)=\mathbb{E}\left(\exp\left(tX_1\right)\right),$ le théorème d'approximation forte de Komlós, Major et Tusnády (1975, 1976) montre qu'on peut construire, sur un espace de probabilités convenable, un processus de Wiener standard $\{W(t):t\geq 0\},$ de telle sorte que

$$|S(t)-W(t)|=O(\log t)\text{ p.s. lorsque }t\to\infty.$$

Pour une suite de constantes positives $\{k_n : n\geq 1\},$ considérons les accroissements maximaux

$$I_n(k_n)=\sup_{0\leq t\leq n-k_n}\{S(t+k_n)-S(t)\}\quad\text{et}\quad I_n^*(k_n)=\sup_{0\leq t\leq n-k_n}\{W(t+k_n)-W(t)\}.$$

La vitesse d'approximation de $S(t)$ par $W(t)$ montre que, lorsque $k_n/\log n\to\infty,$ le comportement, au premier ordre, de $I_n(k_n)$ s'apparente à celui de l'expression analogue $I_n^*(k_n),$ obtenue en remplaçant $S(.),$ par le processus de Wiener $W(.).$ Par contre, lorsque

$k_n / \log n \to c < \infty,$ cette propriété n'est plus satisfaite, puisqu'il est alors possible de montrer que $k_n^{-1} I_n(k_n)$ converge vers presque sûrement vers $\alpha = \inf \{x \geq 0 : \Psi(x) \geq 1/c\}$, où

$$\Psi(x) = \sup_{t:\psi(t)<\infty} \{tx - \log \psi(t)\}$$

est la *fonction de Chernoff* (ou *transformée de Legendre*) associée à la loi commune des $X_i,\ i = 1, 2, \ldots,$ par l'intermédiaire de sa transformée de Laplace $\psi(t) = \mathbb{E}\big(\exp(tX)\big)$.

Ce résultat, dont une version a été mise en évidence pour la première fois par Shepp (1964),est aujourd'hui connu sous le nom de théorème d'Erdős-Rényi (1970). J'ai consacré une part significative de mes travaux ([53], [58], [59], [65], [66], [74], [83], [90], [101]) (avec, notamment, L. Devroye, J. Lynch, J. Steinebach, P.Erdős et P. Révész) à l'étude de ce théorème et des questions qui lui sont apparentées, relevant de l'étude générale des fluctuations de sommes partielles de variables aléatoires indépendantes. J'ai, en particulier, obtenu en 1986-87, dans [53] et [67], la forme complète du théorème d'Erdős-Rényi (celui-ci n'avait été établi antérieurement qu'avec des restrictions sur la valeur de $c \geq 0$), ainsi que la solution du problème, lontemps resté ouvert, de la détermination de la vitesse de convergence de cette loi limite. Dans [101], j'ai obtenu la forme générale du théorème d'Erdős-Rényi fonctionnel, en établissant que l'ensemble $\left\{k_n^{-1}\big(S(t+k_n s) - S(t)\big) : 0 \leq t \leq n - k_n\right\}$, de fonctions de $s \in [0,1]$, converge, dans une topologie convenable, vers un ensemble limite de fonctions sur $[0,1]$, dont j'obtiens également la caractérisation. Dans le cas où $\Psi(u)/|u| \to \infty$ lorsque $|u| \to \infty,$ ce résultat (également obtenu dans ce cas particulier par Borovkov (1991) et Sanchis (1994)) donne un ensemble limite qui n'est autre que la boule unité d'un *espace d'Orlicz*. Celle-ci est composée (pour un choix d'échelle convenable) de l'ensemble des fonctions absolument continues de la forme

$$f(s) = \int_0^s \dot{f}(u)du, \quad \text{pour } s \in [0,1], \quad \text{avec } \int_0^1 c\Psi\left(c^{-1}\dot{f}(u)\right)du \leq 1,$$

où $\dot{f}(t) = \dfrac{d}{dt} f(t)$ désigne la dérivée de Lebesgue de *f*.

Ces recherches ont débouché sur l'étude systématique des processus empiriques locaux (voir les §5-6) et de queue, à l'aide de lois fonctionnelles et de techniques basées sur les *grandes déviations*.

4.e. Processus empiriques.

Etant donné une suite $U_1, U_2, \ldots,$ de variables aléatoires indépendantes de loi uniforme sur $[0,1]$, on définit, pour $n \geq 1$, la *fonction de répartition empirique uniforme* par

$$F_n(x) = n^{-1}\#\{U_i \leq x : 1 \leq i \leq n\} \quad \text{pour} \quad x \in \mathbb{R},$$

où $\#E$ désigne le nombre d'éléments (ou la cardinalité) de *E*. On définit la *fonction empirique de quantile uniforme* par la formule:

$$G_n(t) = \inf\{s \geq 0 : F_n(s) \geq t\} \quad \text{pour} \quad 0 \leq t \leq 1.$$

Le *processus empirique uniforme* α_n et le *processus de quantile uniforme* β_n sont alors définis respectivement par

$$\alpha_n(t) = n^{1/2}\left(F_n(t) - t\right) \quad \text{et} \quad \beta_n(t) = n^{1/2}\left(G_n(t) - t\right) \quad \text{pour} \quad 0 \le t \le 1.$$

Une partie importante de mes travaux concerne l'étude des propriétés asymptotiques de ces processus. Les principaux résultats que j'ai obtenus sont détaillés dans les paragraphes ci-dessous.

4.e.a. Représentation de Bahadur-Kiefer.

En 1990, avec D. M. Mason, j'ai résolu, dans [88], la *conjecture de Kiefer*, restée ouverte depuis 1970. Pour énoncer ce résultat, posons, pour toute fonction *g*, bornée sur [0,1], $\|g\| = \sup_{0\le t\le 1} |g(t)|$. Bahadur (1966), et Kiefer (1967, 1970), ont montré que:

(i) Pour tout $t_0 \in [0,1]$ fixé,

$$\limsup_{n\to\infty} \left\{ n^{1/4} (2\log\log n)^{-3/4} \left|\alpha_n(t_0) + \beta_n(t_0)\right| \right\} = \left(t_0\left(1-t_0\right)\right)^{1/4} 2^{1/2} 3^{-3/4} \quad \text{p.s.}$$

(ii) On a

$$\limsup_{n\to\infty} \left\{ n^{1/4} (2\log\log n)^{-3/4} \left\|\alpha_n + \beta_n\right\| \right\} = 2^{-1/4} \quad \text{p.s.}$$

(iii) On a

$$\lim_{n\to\infty} \left\{ n^{1/4} (\log n)^{-1/2} \frac{\left\|\alpha_n + \beta_n\right\|}{\left\|\alpha_n\right\|^{1/2}} \right\} = 1 \quad \text{en probabilité.}$$

Kiefer (1970) a émis la conjecture que la limite *(iii)* avait lieu presque sûrement. J'ai démontré ([88]) (avec D. Mason) que c'était effectivement le cas. Cette propriété permet alors de ramener la démonstration de *(ii)* à celle de la loi du logarithme itéré de Chung (1948), c'est à dire, au fait que

$$\limsup_{n\to\infty} \left\{ (\log\log n)^{-1/2} \left\|\alpha_n\right\| \right\} = 2^{-1/2} \quad \text{p.s.}$$

Dans une série d'articles ([82], [88], [93], [99], [100], [105], [106]), écrits pour partie avec J. Beirlant, D. Mason, J. H. J. Einmahl et J. Steinebach, j'ai pu développer ces résultats et les étendre à d'autre processus tels que le *processus de sommes partielles,* le *processus de renouvellement*, le *processus de Kaplan-Meier* et les *processus empirique* et *de quantiles* associés aux *espacements*.

Par exemple, dans un travail récent ([117]), avec D. Mason, j'ai pu élucider le mystère de la constante bizarre $2^{1/2}3^{-3/4}$ qui apparaît dans le membre de droite de *(i).* Grâce à une nouvelle démonstration de cette propriété, basée sur une *loi fonctionnelle du logarithme itéré*, il a été possible d'en expliquer que l'origine, se ramenait la formule simple

$$\sup_{0\le x\le 1} \sqrt{s\left(1-s^2\right)} = 2^{1/2}3^{-3/4}.$$

Les résultats ci-dessus sont intimement liés à l'étude des *modules de continuité* de α_n et β_n.

4.e.b. Lois limites fonctionnelles locales.

Le comportement des *fluctuations locales* du *processus empirique uniforme* α_n et du *processus empirique uniforme de quantiles* β_n a fait l'objet d'une grande partie de mes travaux récents. Dès (1984), j'avais décrit, dans [44] avec L. Devroye, le comportement presque sûr des k_n-*espacements uniformes* pour $k_n = O(\log n)$, ce qui se ramène à l'étude des incréments d'ordre $n^{-1}k_n$ du processus empirique de quantiles uniforme $\{\beta_n(t): 0 \le t \le 1\}$. Depuis 1988, j'ai entrepris l'étude plus systématique de ces quantités, à l'aide, principalement, de *lois fonctionnelles limites*.

Considérons une suite de constantes positives $\{h_n : n \ge 1\}$, vérifiant
$0 < h_n < 1,\quad h_n \downarrow 0,\quad nh_n \uparrow \infty,\quad nh_n / \log n \to c \in [0,\infty],\quad (\log(1/h_n))/\operatorname{loglog} n \to d \in [0,\infty]$.
On considère les ensembles aléatoires de fonctions définis par

$$\mathfrak{E}_n = \left\{ \frac{\alpha_n(t+h_n I) - \alpha_n(t)}{\sqrt{2h_n \log(1/h_n)}} : 0 \le t \le 1-h_n \right\},$$

où $I(t) = t$ désigne la fonction identité définie sur [0,1], et

$$\mathfrak{F}_n = \left\{ \frac{\beta_n(t+h_n I) - \beta_n(t)}{\sqrt{2h_n \log(1/h_n)}} : 0 \le t \le 1-h_n \right\}.$$

Désignons par $\mathbb{S}$, l'ensemble des fonctions *f*, définies dans $[0,1]$, telles que $f(0)=0$, et absolument continues, de dérivées de Lebesgue $\dot{f}$ telles que

$$|f|_{\mathbb{H}} = \left\{ \int_0^1 \dot{f}(t)^2 \, dt \right\}^{1/2} \le 1.$$

Cet ensemble n'est autre que la boule unité de *l'espace de Hilbert à noyau autoreproduisant* associé à la fonction de covariance du processus de Wiener, utilisée par Strassen (1964) dans la *loi fonctionnelle du logarithme itéré*. Avec la notation $\|g\| = \sup_{0 \le t \le 1} |g(t)|$, posons, pour toute partie *A*, non vide de l'ensemble $B[0,1]$ des fonctions bornées sur $[0,1]$,

$$A^{\varepsilon} = \{ f \in B[0,1] : \exists g \in A, \|f-g\| < \varepsilon \}.$$

Enfin, si $A \subseteq B[0,1]$, et $B \subseteq B[0,1]$, on définit la *distance de Hausdorff* entre *A* et *B*, par

$$\Delta(A,B) = \inf \{ \varepsilon > 0 : A \subseteq B^{\varepsilon} \ \text{ et } \ B \subseteq A^{\varepsilon} \},$$

si un tel $\varepsilon > 0$ existe, et $\Delta(A,B) = \infty$ autrement. Dans [108] (avec D. M. Mason), j'ai montré que, sous ces hypothèses, lorsque $c = d = \infty$,

$$\lim_{n\to\infty} \Delta(\mathfrak{E}_n, \mathbb{S}) = \lim_{n\to\infty} \Delta(\mathfrak{F}_n, \mathbb{S}) = 0 \quad \text{p.s.}$$

Le comportement de ces suites lorsque $c < \infty$ ou $d < \infty$ est décrit dans [107], [108], [120], [126], [167].

Lorsque $t_0 \in [0,1)$ est fixé, le *processus empirique local* en t_0 est défini par la suite

$$\frac{\alpha_n(t_0 + h_n s) - \alpha_n(t_0)}{\sqrt{2h_n \operatorname{loglog} n}} \quad \text{pour} \quad s \in [0,1].$$

On suppose que

$$h_n \downarrow 0, \quad nh_n \uparrow \infty, \quad nh_n / \log\log n \to r \in [0,\infty].$$

Dans [91] (avec D. M. Mason), j'ai décrit le comportement presque sûr de $\{f_n\}$, en montrant notamment que cette suite était presque sûrement relativement compacte dans l'ensemble $B[0,1]$ des fonctions bornées sur $[0,1]$, muni de la topologie uniforme. De plus, j'ai établi que, pour $r < \infty$, l'ensemble limite s'exprimait comme la boule unité d'un certain espace d'Orlicz.

Récemment, dans [137], j'ai montré que, pour toute fonction *f* de l'ensemble de Strassen, i.e. telle que $f(0) = 0$, avec

$$|f|_{\mathbb{H}} = \left\{ \int_0^1 \dot{f}^2(t)dt \right\}^{1/2} \leq 1,$$

où $\dot{f}(t) = \frac{d}{dt} f(t)$ est la dérivée de Lebesgue de *f*, et, pour toute suite $\{h_n : n \geq 1\}$ telle que

$$h_n \downarrow 0, \quad nh_n \uparrow \infty, \quad nh_n / \log\log n \to \infty,$$

on a, en notant $\|g\| = \sup_{0 \leq t \leq 1} |g(t)|$,

$$\liminf_{n\to\infty} (2\log\log n) \left\| \frac{\alpha_n(h_n \bullet)}{\sqrt{2h_n \log\log n}} - f \right\| = \frac{\pi}{2\sqrt{1-|f|_{\mathbb{H}}^2}} \quad \text{p.s.}$$

Dans [117], [120], [143], les résultats correspondant au processus empirique α_n sont generalisés au cas de variables aléatoires à valeurs dans $\mathbb{R}^d$, pour des processus indexés par des fonctions ou par des ensembles. Ces travaux permettent de décrire, par des corollaires simples de théorèmes généraux, le comportement presque sûr d'un très grand nombre de statistiques non paramétriques *locales* (telles que développées dans le cadre de l'estimation non paramétrique de la densité, cf. §). Nous traitons plus en détail dans ce qui suit le cas particulier de l'étude des *statistiques des queues de distribution*.

4.e.c. Approximation forte et fluctuations du processus des quantiles.

Dans l'article récent [129], j'ai montré que lecomportement asymptotique du processus empirique local de quantiles, basé sur β_n, pouvait différer très notablement de celui du processus empirique local basé sur α_n. Le résultat s'énonce comme suit. Fixons $t_0 \in]0,1[$. Considérons la suite de fonctions de $s \in [0,1]$, définie par

$$g_n(s) = \frac{\beta_n(t_0 + h_n s) - \beta_n(t_0)}{\sqrt{2h_n \left\{ \log_+ \left(\frac{1}{h_n \sqrt{n}} \right) + \log\log n \right\}}},$$

où $\log_+ u = \log(\max\{u, e\})$. Alors, sous des hypothèses générales de comportement et de régularité portant sur $\{h_n : n \geq 1\}$, supposant, en particulier, que $nh_n / \log n \to \infty$, la suite $\{g_n : n \geq 1\}$ est presque sûrement relativement compacte dans l'espace des fonctions bornées muni de la topologie uniforme, et a pour ensemble limite l'ensemble de Strassen (défini plus haut au §5b). Il est remarquable que ce résultat diffère de celui qui serait obtenu, soit pour

$t_0 = 0,$ soit en remplaçant β_n par α_n. Dans chacun de ces cas, le terme $\log_+\left(1/\left(h_n\sqrt{n}\right)\right)$ disparaît.

Une application directe de cette propriété montre que, pour tout processus de Kiefer $\{K(s,t): s \geq 0, 0 \leq t \leq 1\}$, et pour tout $\varepsilon > 0$, on a

$$\limsup_{n\to\infty} n^{1/4}(\log n)^{\varepsilon}\left\{\sup_{0\leq t\leq 1}\left|\beta_n(t) - n^{-1/2}K(n,t)\right|\right\} = \infty \text{ p.s.}$$

L'approximation des processus empiriques et de quantiles uniformes par des processus de Kiefer est d'une importance particulière en statistique asymptotique. Un processus de Kiefer $K(n,t)$ est une somme

$$K(n,t) = \sum_{i=1}^{n} B_i(t)$$

de ponts browniens $B_i(t) = W_i(t) - tW_i(1)$, indépendants (construite à partir d'une suite de processus de Wiener $\{W_i(t): t \geq 0\}$, $i = 1,2,\ldots,$ indépendants). La meilleure vitesse connue d'approximation du processus empirique par un processus de Kiefer est due à Komlós, Major et Tusnády (1975). Ils ont établi que, (sur un espace de probabilités convenable, et pour un processus de Kiefer $K'(n,t)$ approprié), lorsque $n \to \infty$,

$$\sup_{0\leq t\leq 1}\left|\alpha_n(t) - n^{-1/2}K'(n,t)\right| = O\left(n^{-1/2}\log^2 n\right) \quad \text{p.s.}$$

En ce qui concerne le processus des quantiles, le meilleur résultat connu est dû à Csörgő et Révész (1975,1976) qui ont montré que (sur un espace de probabilités convenable, et pour un processus de Kiefer $K''(n,t)$ approprié), lorsque $n \to \infty$,

$$\sup_{0\leq t\leq 1}\left|\beta_n(t) - n^{-1/2}K''(n,t)\right| = O\left(n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}\right) \quad \text{p.s.}$$

Csörgő et Révész (1975,1976) ont émis la conjecture que le meilleur ordre possible d'approximation uniforme de $\beta_n(t)$ par $n^{-1/2}K(n,t)$ pourrait être analogue à celui obtenu pour α_n, à savoir en $n^{-1/2}(\log n)^2$. Mon résultat de [129] infirme cette conjecture. Plus récemment (dans [134]), j'ai pu montrer, en fait, que la vitesse d'approximation obtenue par Csörgő et Révész (1975,1976) était *optimale*, en apportant une solution finale à ce problème, resté ouvert depuis 1975. Mon résultat montre l'existence d'une constante positive *C*, comprise entre $1/11$ et $2^{-1/4}$, et telle que, quel que soit le processus de Kiefer *K* considéré, on ait

$$\limsup_{n\to\infty}\left(n^{1/4}(\log n)^{-1/2}(\log\log n)^{-1/4}\right)\left|\beta_n(t) - n^{-1/2}K''(n,t)\right| \geq \mathrm{C} \quad \text{p.s.}$$

Plus récemment, dans [142], j'ai montré qu'on pouvait approximer le processus empirique des quantiles avec une vitesse d'approximation d'ordre $n^{\varepsilon-1/2}$, pour $\varepsilon > 0$ arbitraire, à l'aide de processus de Kiefer itérés. Dans [133] et [139], je décris le comportement local du processus des quantiles pour des accroissements d'ordre $h_n = cn^{-1}\log n$.

4.e.d. Espacements.

Soit $X = X_1,\ldots,X_n$ un échantillon de taille *n* d'une loi de probabilité réelle, dont les statistiques d'ordre sont notées

$$X_{1,n} \leq \ldots \leq X_{n,n}.$$

L'étude des *espacements* $X_{i+1,n}-X_{i,n}$, $i=1,\ldots,n-1$ associés (qui peuvent aussi être interprétés comme les accroissements d'ordre 1/n de la fonction de quantile empirique de l'échantillon), a été développée dans mes articles [31], [32], [41], [44], [45], [56], [57], [61], [75] essentiellement de 1981 à 1988, puis, à nouveau en 2003, dans [148].

Mes travaux ont porté tout d'abord sur les *espacements maximaux*

$$M_n = \max_{\{1\leq i<n\}}\{X_{i+1,n} - X_{i,n}\},$$

dans le cas de lois uniformes, ainsi que dans le cas de lois plus générales. En particulier, j'ai obtenu des encadrements asymptotiques presque sûrs pour M_n et d'autres statistiques analogues, ainsi que plusieurs théorèmes limites en loi décrivant le comportement de ces suites aléatoires. Par exemple, dans [61], j'ai établi que, si la loi de X a une densité f, continue et positive sur $[0,\infty)$ et nulle sur $]-\infty,0]$, et si, lorsque $n \to \infty$, le maximum $X_{n,n}$ de $X_1,\ldots,X_n$ est tel que $a_n^{-1}(X_{n,n} - b_n)$ converge en loi vers une loi de Gumbel}, de fonction de répartition exp($-e^{-x}$), alors

$$\lim_{n\to\infty}\mathbb{P}\left(a_n^{-1}M_n > x\right) = \prod_{k=1}^{\infty}\left(1-e^{-kx}\right) \quad \text{pour} \quad x > 0.$$

4.f. Statistiques actuarielles - Statistiques des queues de distributions.

4.f.a. Estimation de l'index de Pareto.

Une variable aléatoire Y est dite (cf. [54]) d'*index de Pareto* $\lambda > 0$, si

$$P(Y > y) = y^{-1/\lambda}\,L(y),$$

où L(.) est une fonction *à variation lente à l'infini*, c'est à dire, telle que pour tout choix de $c>0$, $L(cy)/L(y) \to 1$ lorsque $y \to \infty$.

Ce type de loi est souvent associé à des phénomènes physiques, qui se rencontrent typiquement en assurance, dans le cas de *grands sinistres*. Il est également présent dans de nombreux exemples industriels (corrosion), ou naturels (météorologie). Il est donc utile d'estimer le paramètre λ à partir d'échantillons observés, selon le schéma suivant.

Soit une suite $Y_1, Y_2,\ldots$, de variables aléatoires indépendantes de même loi d'index de Pareto $\lambda>0$. Désignons par $Y_{1,n} \leq \ldots \leq Y_{n,n}$ la statistique d'ordre de $Y_1,\ldots,Y_n$. Pour estimer λ, Hill (1975) a introduit l'estimateur

$$\lambda_n^* = \frac{1}{k_n}\sum_{i=1}^{k_n}(\log Y_{n-i+1,n} - \log Y_{n-i,n}),$$

où $\{k_n : n \geq 1\}$ désigne une suite de constantes entières telles que

$$1 \leq k_n < n,\ k_n \uparrow \infty, \text{ et } n^{-1} k_n \to 0.$$

J'ai établi dans [76], en 1988 (avec D. M. Mason et E. Haeusler), qu'une condition *nécessaire et suffisante* pour que $\lambda_n^* \to \lambda$ p.s. lorsque $n \to \infty$ est que

$$k_n / \text{loglog } n \to \infty.$$

En 1985, j'ai introduit, dans [55] (avec S. Csörgő et D. M. Mason), un estimateur, comprenant l'estimateur de Hill comme cas particulier, et défini par

$$\lambda_n^G = \left(\sum_{j=1}^{k_k} \frac{1}{k_n} K\left(\frac{j}{k_k}\right)\right)^{-1} \left(\sum_{j=1}^{k_k} \frac{j}{k_n} K\left(\frac{j}{k_k}\right) \left\{ \log Y_{n-j+1,n} - \log Y_{n-j,n} \right\} \right),$$

où le "noyau" $\{K(t) : t \geq 0\}$ est une fonction, vérifiant des propriétés de régularité convenables (comme celle d'être nulle en dehors d'un intervalle compact de $\mathbb{R}^+$, et d'être à variation bornée), ainsi que:

$$\int_0^\infty K(t)dt = 1 \text{ et } \int_0^\infty K^2(t)dt < \infty.$$

J'ai établi dans [55] diverses propriétés de cet estimateur, dont des conditions impliquant sa normalité asymptotique et sa convergence en probabilité. De plus, j'ai déterminé des choix asymptotiquement optimaux du noyau K et de la suite k_n permettant de choisir ces facteurs dans les applications pratiques. J'ai également montré que l'estimateur de Hill ne permettait généralement pas d'obtenir une vitesse de convergence optimale en moyenne quadratique realtivement à l'ensemble des estimateurs issus de cette famille élargie.

4.f.b. Sommes d'extrêmes, coefficient d'ajustement en théorie du risque

Il s'agit d'étudier le comportement limite d'expressions de la forme

$$S_n(k_n) = \sum_{i=1}^{k_n} Y_{n-i+1,n},$$

où $Y_{1,n} \leq \ldots \leq Y_{n,n}$ désigne la statistique ordonnée des n premières observations $Y_1, \ldots, Y_n$ d'une suite de variables aléatoires indépendantes de même loi. Ce type de problème a des applications naturelles pour l'étude des *sommes de grands sinistres* en théorie actuarielle. Il présente aussi un grand intérêt théorique dans le cadre de l'étude des statistiques censurées et tronquées. Enfin, nous mentionnerons plus loin une application obtenue pour l'estimation du *coefficient d'ajustement* en théorie du risque.

Mes contributions dans ce domaine ont débuté par les articles, [72] et [87], consacrés au cas où le maximum $Y_{n,n} = \max\{Y_1, \ldots, Y_n\}$ est dans le domaine d'attraction de la loi de Gumbel. On suppose, dans ce cas, qu'il existe des constantes $a_n > 0$ et b_n telles que, lorsque $n \to \infty$,

$$\mathbb{P}\left(a_n^{-1}\left(Y_{n,n} - b_n\right) \leq x\right) \to \exp\left(-e^{-x}\right) \text{ pour tout } x \in \mathbb{R}.$$

Dans ces travaux, j'ai montré (avec D. M. Mason et E. Haeusler) qu'on pouvait se ramener essentiellement au cas où la loi des Y_i est exponentielle pour $i = 1, 2, \ldots,$ c'est à dire, telle que $\mathbb{P}(Y_i > t) = e^{-t}$ pour $t > 0$. Un exemple surprenant des résultats ainsi obtenus est le suivant. Supposons que $k_n / \text{loglog } n \to c \in]0, +\infty[$. Alors, on a, presque sûrement,

$$\limsup_{n\to\infty} \left\{ \pm \left(\frac{1}{k_n} \sum_{i=1}^{k_n} Y_{n-i+1,n} - \log\left(n / k_n\right) \right) \right\} = \frac{\pm 2c}{2c + 1 \mp \sqrt{1+4c}} \pm \log\left(\frac{2c}{2c + 1 \mp \sqrt{1+4c}} \right).$$

La constante bizarre trouvée dans le membre de droite de cette expression a pu être expliquée par mon article [96] de 1989 (avec D. M. Mason). J'y ai montré que toute une série de lois limites du calcul des probabilités pouvaient être obtenues comme conséquences de *lois fonctionnelles du logarithme itéré*. Dans l'exemple présent, on raisonne comme suit. Tout d'abord, on établit ([91]) que la suite de fonctions définie par

$$g_n(s) = \frac{n}{\log\log n}\ G_n\left(\frac{k_n s}{n}\right) \ \text{pour} \ 0 \le s \le 1,$$

où $G_n(t)$ désigne la fonction empirique de quantiles uniforme (voir le5 ci-dessus), est presque sûrement relativement compacte pour la topologie faible des fonctions de répartition de mesures positives. Cette suite a comme ensemble limite l'ensemble, noté $\Gamma(c)$, de toutes les fonctions de répartition $g(s)$, de mesures positives sur [0,1], ayant une composante singulière $g_S(s)$, et une composante absolûment continue $\int_0^s g'(t)dt$, telles que

$$\Gamma(c) = \left\{ g : g(0) = 0,\ g_S(1+) + \int_0^1 c\ell\left(c^{-1} g'(u)\right) du \le 1 \right\},$$

où $\ell(u) = u - 1 - \log u$ pour $u > 0$. Ensuite, on vérifie que

$$\left\{ \frac{1}{k_n} \sum_{i=1}^{k_n} Y_{n-i+1,n} - \log\left(n / k_n\right) \right\} = \int_0^1 \{-\log g_n(s)\} ds + \log\left(\frac{k_n}{\log\log n} \right).$$

Il suffit alors de constater que

$$\sup \pm \left\{ \int_0^1 \{-\log g_n(s)\} ds + \log c \right\} = \frac{\pm 2c}{2c + 1 \mp \sqrt{1+4c}} \pm \log\left(\frac{2c}{2c + 1 \mp \sqrt{1+4c}} \right),$$

pour retrouver la conclusion de [96] citée ci-dessus. Ce genre de résultat explique naturellement la génération de constantes inhabituelles dans les théorèmes limites, à partir de la solution de certains problèmes de calculs d'extrema dans des espaces fonctionnels, rentrant dans la catégorie générale des *espaces d'Orlicz*.

Le *coefficient d'ajustement* en théorie du risque est défini comme le plus grand nombre $a > 0$ tel que la probabilité de ruine d'une compagnie d'assurance ayant un capital initial égal à *t* soit uniformément bornée supérieurement par une expression de la forme $C\exp\left(-at\right)$, où C désigne une constante convenable. Dans l'article [92] (en collaboration avec J. Steinebach), j'ai introduit (et étudié les propriétés de convergence correspondantes à) des estimateurs de ce coefficient d'ajustement basés sur les fluctuations négatives du *processus de risque*. Ces estimateurs s'apparentent à des sommes pondérées de valeurs extrêmes, et sont aujourd'hui couramment utilisés, entre autres, pour valider des modèles paramétriques sur la loi des sinistres.

4.f.c. Approximation forte du processus de risque.

Le *processus de risque actuariel* est défini par

$$S(t) = \sum_{i=1}^{N(t)} Y_i,$$

où $\{Y_i : i \ge 1\}$ désigne la suite des *coûts de sinistres*, et $N(t)$ représente le *nombre de sinistres* observés dans l'intervalle de temps $\left(0, t\right]$. On suppose ici que le $n^{\text{ème}}$ sinistre survenient à l'instant $X_1 + \ldots + X_n$. On suppose que $\{X_i : i \ge 1\}$ et $\{Y_i : i \ge 1\}$ sont deux suites indépendantes, chacune d'entre elles étant composée de variables aléatoires indépendantes de même loi, ayant une fonction génératrice des moments finie au voisinage de l'origine. On pose

$$\mathbb{E}\left(X_1\right)=m,\ \ \mathbb{E}\left(Y_1\right)=\mu,\ \ \operatorname{Var}\left(X_1\right)=s^2,\ \ \operatorname{Var}\left(Y_1\right)=\sigma^2.$$

En 1987, dans [71], j'ai prouvé (avec M.Csörgő et L.Horváth), qu'il existe un espace de probabilité supportant à la fois le processus $\{S(t): t\geq 0\}$, et un processus de Wiener standard $\{W(t): t\geq 0\}$, de telle mani\` ere que, pour $x\geq 0$ et $T>0$, on ait l'inégalité

$$\mathbb{P}\left(\sup_{0\leq t\leq T}\left|S(t)-\frac{\mu}{m}t-\left(\frac{\sigma^2m^2+s^2\mu^2}{m^3}\right)^{1/2}W(t)\right|\geq x+A\log T\right)\leq Be^{-Cx},$$

pour des $A>0,\ B>0,\ C>0$ convenables.

Des résultats analogues sont obtenus lorsque X_1 et Y_1 possèdent seulement des moments d'ordre $r>2.$ Ces travaux comprennent, comme cas particuliers, les approximations fortes de Komlós, Major et Tusnády (1975,1976) (en prenant les X_i constants), et ceux de Mason et Van Zwet (1987) (en prenant les Y_i constants). De plus, ils permettent diverses applications actuarielles, ainsi que dans la théorie des files d'attente. L'approximation du processus de risque par un processus de Wiener permet de simplifier considérablement certaines analyses du risque actuariel lié aux fluctuations de ce processus. On ne peut en effet, sauf exception, décrire ce dernier par des calculs exacts, que le cas où le processus d'arrivée des sinistres est un processus de Poisson, ce qui est, en général, très sensiblement différent de la réalité observée.

4.g. Lois fonctionnelles et théorèmes du type Strassen.

Soit $\{W(t): t\geq 0\}$ un processus de Wiener standard. La loi fonctionnelle du logarithme itéré de Strassen (1964) établit que, si, $T>3$,

$$f_T(s)=\frac{W(Ts)}{\sqrt{2T\ \operatorname{loglog} T}}\quad \text{pour}\ \ 0\leq s\leq 1,$$

alors, pour toute suite $3<T_1<\ldots<T_n<\ldots$, avec $T_n\to\infty$, la suite $\left\{f_{T_n}(.): n\geq 1\right\}$ est presque sûrement relativement compacte dans l'ensemble $C[0,1]$ des fonctions continues sur $[0,1]$, muni de la topologie de la convergence uniforme, et que l'ensemble limite, composé des limites de toutes les sous-suites convergentes de $\left\{f_{T_n}(.): n\geq 1\right\}$, est presque sûrement égal à

$$\mathbb{S}=\left\{f: f(t)=\int_0^t\phi(u)du,\ 0\leq t\leq 1,\ \int_0^1\phi^2(u)du\leq 1\right\}.$$

Dans un travail récent, en collaboration avec M. Lifshits, j'ai établi ([114], [118]) des *conditions nécessaires et suffisantes* portant sur une semi-norme $|\bullet|$ générale (pouvant prendre des valeurs infinies) sur $C[0,1]$ pour que le résultat ci-dessus reste vrai pour la topologie définie par $|\bullet|$. Ces résultats apportent une solution complète à un problème posé 1964, et abordé, entre autres, par Lerche (1992) et Baldi, Ben Arous et Kerkyacharian (1992), pour des normes particulières (telle que la norme de Hölder).

Dans [113], en collaboration avec P.Révész, j'ai décrit le comportement limite de la suite d'ensembles aléatoires

$$\left\{ \frac{W_i(Ts)}{\sqrt{2T\left\{\log \kappa_T + \operatorname{loglog} T\right\}}} : 0 \le s \le 1,\ i = 1, \dots, \kappa_T \right\},$$

où $\left\{\kappa_T : T > 0\right\}$ est une fonction monotone de *T*. Ces résultats ont des applications nombreuses, et sont liés à des travaux analogues sur les sommes partielles dans les tableaux de variables aléatoires ([122], [123]), et les incréments de processus empiriques ([108]).

4.h. Objets fractals aléatoires.

Dans plusieurs articles récents ([116], [121], [127], [128], [132]), écrits, pour partie, en collaboration avec D. M. Mason et M. A. Lifshits, j'ai entrepris l'étude des *objets fractals aléatoires* engendrés par les oscillations de divers processus, tels que le processus empirique, le processus des quantiles, le processus de Wiener, etc. Voici un exemple de ces résultats. Soit une suite de constantes positives $\left\{h_n : n \ge 1\right\}$, vérifiant

$$0 < h_n < 1,\quad h_n \downarrow 0,\quad nh_n \uparrow \infty,\quad nh_n / \log n \to \infty,\quad \left(\log\left(1/h_n\right)\right)/\operatorname{loglog} n \to \infty.$$

On considère les ensembles aléatoires définis par

$$D(\lambda) = \left\{ t \in [0,1] : \limsup_{n\to\infty} \left(\frac{\alpha_n\left(t + h_n I\right) - \alpha_n(t)}{\sqrt{2h_n \log\left(1/h_n\right)}} \right) \ge \lambda \right\}.$$

Alors, avec probabilité 1 pour tout $\lambda \in \left]0,1\right[$, l'ensemble $D(\lambda)$ est partout dense dans $[0,1]$ et de *dimension de Hausdorff* égale à $1-\lambda^2$.

De nombreuses extensions de ces résultats sont données dans [116], [121], [130], comme, par exemple, des raffinements de résultats dûs à Orey et Taylor (1974) sur le processus de Wiener. Plus récemment, j'ai obtenu dans [132] (avec D. M. Mason) une version nouvelle de la loi de Chung pour le processus de Wiener. Celle-ci est décrite en introduisant une nouvelle famille d'objets fractals aléatoires gengendrés par les points exceptionnels du processus de Wiener. Dans [127], avec M. A. Lifshits, j'ai montré que ces résultats étaient valables sous forme fonctionnelle, indépendamment de la norme utilisée, pourvu que celle-ci demeure compatible avec la loi du logarithme itéré pour le processus de Wiener.

4.i. Statistique appliquée et industrielle.

A côté des travaux de statistique théorique ayant fait l'objet de publications dans des journaux scientifiques et évoqués ci-dessus une partie importante de mes travaux concerne la *statistique appliquée*. Au-delà des articles [25], [46], [98], [102], [111], [119], [140], ont trait à quelques unes de ces recherches, la majeure partie de celles-ci a fait l'objet de rapports internes et de notes techniques au sein des organismes où ils ont été effectués, principalement des compagnies pétrolières (Compagnie Française des Pétroles, TOTAL, ELF) et phamaceutiques (ELF-SANOFI, SANOFI-SYNTHELABO, SANOFI-AVENTIS). Compte tenu de leur importance industrielle, ces travaux sont pour l'essentiel du ressort de la confidentialité, et je n'en mentionnerai que quelques exemples significatifs.

Un résultat ayant fait l'objet de publication partielle ([98], [111]) concerne l'industrie pétrolière dans le cadre de la modélisation et du contrôle des *écoulements diphasiques.* Schématiquement, si on pompe dans un pipe-line un mélange de gaz et de liquide, il peut se former des b*ouchons* liquides en alternance avec des *bulles* gazeuses. Il est alors important de pouvoir disposer de modèles fiables permettant de décrire la longueur des bouchons dans un écoulement stationnaire. Dans des travaux, en collaboration avec M. Bernicot et H. Dhulesia ([121]), puis avec M. Bernicot ([119]), j'ai pu établir que cette loi était bien représentée, suivant les circonstances, soit par des modèles *gaussiens inverses,* soit par des modèles *demi-normaux*. La loi gaussienne inverse n'est rien d'autre que la distribution du premier temps de passage d'un processus de Wiener avec dérive, de la forme $\nu t+\sigma W(t)$ (avec $\nu>0$) à un niveau donné $a>0$. Cette loi a une densité donnée par

$$f(t)=\frac{a}{\sigma\sqrt{2\pi t^3}}\ \exp\left(-\frac{\left(a-\nu t\right)^2}{2\sigma^2 t}\right)\ \text{ pour }\ t>0.$$

Une analyse statistique, comprenant l'ajustement des paramètres et la comparaison avec d'autres familles de distributions possibles, d'un grand nombre d'observations expérimentales de systèmes déjà installés (notamment en Indonésie), ainsi qu'une modélisation théorique du phénomène ont permis de mettre en évidence et d'interpréter la génération de ces lois. Ces recherches contribuent à la prédiction de la taille des bouchons apparaissant à la sortie des pipe-lines. Les caractéristiques de ces derniers doivent être convenablement évaluées pour concevoir une bonne gestion des installations.

Un autre exemple de recherche appliquée est la mise au point de protocoles expérimentaux pour l'analyse de données pharmaceutiques. Dans [146], en collaboration avec G. Derzko, j'ai répertorié une famille de plans d'expérience en blocs incomplets équilibrés *optimaux* dans le cas de liaisons temporelles. Ces plans sont utilisés pour la comparaison de nouveaux médicaments en milieu hospitalier, notamment par ELF-SANOFI. Dans le même champ de recherche, j'effectue des analyses de durées de survie en données censurées par des méthodes inspirées par les articles [89], [109] et [125].

Ces activités de statistique appliquée ont le double objet de motiver mes recherches de statistique fondamentale par l'observation in situ des problèmes théoriques posés par la pratique permanente du traitement des données, et inversement de pouvoir mettre en oeuvre en industrie des résultats issus de la recherche théorique, par l'intermédiaire de stagiaires (DEA, ISUP, etc.) issus des formations d'enseignement auxquelles je participe, et dont l'insertion en milieu professionnel est favorisée par ces relations.

Dans la cadre de ces applications, j'ai travaillé récemment, avec D. M. Mason et G. Shorack sur le *bootstrap*. Dans [110], nous avons caractérisé les cas où les bootstraps des extrêmes, de sommes d'extrêmes, de statistiques tronquées et de la moyenne, étaient convergents. Ces recherches, de nature théorique, sont associées à des applications industrielles.

Plus récemment, j'ai coordonné un contrat d'assistance technique pour la Commission Européenne (en 2002-2003). Le rapport correspondant a été rendu disponible sur internet sur les sites web :

http://europa.eu.int/comm/consumers/cons_safe/news/prod_saf_fr.pdf

et

http://europa.eu.int/comm/consumers/cons_safe/news/prod_saf_en.pdf

Mes activités d'expertise sur la sécurité des produits et sur la comparaison des protocoles de mesure ont donné lieu à de nombreux rapports d'expertise, notamment, auprès de la DGCCRF, et de plusieurs organismes internationaux (sécurité des mesures d'incendie en Grande Bretagne).

5. Séjours dans des centres universitaires étrangers (liste partielle)

Allemagne :

Technische Universität Aachen
Universität Bremen
Universität Hannover
Universität Marburg
Universität Oldenburg
Universität Gesamthochschule Siegen

Autriche:

Technische Universität, Wien

Belgique:

Katholieke Universiteit Leuven
Limburg Universitair Zentrum

Canada:

Carleton University, Ottawa
McGill University, Montréal
Université Laval, Québec

Espagne :

Universidad de Granada
Universidad de Malaga
Universidad de Santander
Universidad Complutense, Madrid

Etats Unis:

University of California, Santa Barbara
Columbia University, New York
University of Delaware, Newark, Delaware
The Johns Hopkins University, Baltimore
University of Indiana, Bloomington
University of North Carolina, Chapel Hill
University of Washington, Seattle

Hongrie:

Matematikai Kutato Intezete, Budapest

Italie:

Universita degli Studi Torino

Pays Bas:

Katholieke Universiteit Nijmegen
Universiteit Eindhoven
Universiteit Leiden
Universiteit Maastricht
Erasmus Universiteit, Rotterdam

Portugal :

Universita de Coimbra
Facultade de Ciencias de Lisboa
Universita Nova de Lisboa
Universita de Porto

Roumanie :

Centrul de Statistica Matematica, Bucarest
Universitate din Cluj
Universitate din Craiova
Universitate din Iasi
Universitate din Timisoara

Russie :

Université de Saint Petersbourg

Suisse :

ETH Zürich
Universität Zürich

6. Congrès (liste partielle)

1**979**:

16-18 avril Journées de Statistiques, ASU, Paris
28-31 mai Statistique non Paramétrique, Rouen
13-14 juin Asymptotic Theory of Statistical Tests and Estimation, Chapel Hill, USA (en l'honneur de Wassily Hoeffding)

1980:

19-22 mai Journées de Statistiques, ASU, Toulouse
8-4 juin Analytische Methoden der Wahrscheinlichkeitsrechnung, Oberwolfach, Allemagne
23-28 juin János Bolyai Symposium on Nonparametric Statistics, Budapest
5-12 juillet XII° Reunion de la Sociedad Española de Investigación, Estadistica e Informatica, Jaca, Espagne

1981:

6--10 avril Exchangeability in Probability and Statistics, Rome, en l'honneur de Bruno de Finetti
29 avril-1 mai Canadian Conference in Applied Statistics, Montréal, Canada
14--20 juin IInd Pannonian Symposium on Mathematical Statistics, Bad

Tatzmannsdorf, Autriche

9-13 novembre — II^{o} Colloquio de Estadistica, Coimbra, Portugal

30 nov.-1 déc. — 43rd Session of the International Statistical Institute, Buenos Aires

14-18 déc. — Density Estimation, Oberwolfach

1982:

21-26 juin — János Bolyai Symposium on Limit Theorems in Probability and Statistics, Veszprém, Hongrie

29 août-4 sep. — 7th Conference in Probability Theory, Brasov, Roumanie

14-19 sep. — IIIrd Pannonian Symposium on Mathematical Statistics, Visegrad, Hongrie

4 juin — Contact Group in Probability Theory, Diepenbeek, Belgique

1983:

31 août-4 sep. — Nato-ASI Conference on Statistical Extremes and Applications, Vimeiro, Portugal

25-29 juillet — Multivariate Analysis VI, Pittsburgh, USA

24--25 nov. — 4ème Rencontre Franco-Belge de Statisticiens, Louvain la Neuve, Belgique

1984:

25-31 mars — Order Statistics, Quantile Processes and Extreme Value Theory, Oberwolfach

9-14 avril — XIV^{o} Reunion de la Sociedad Española de Investigació, Estadistica e Informatica, Malaga, Espagne

1985:

4-8 août — IMSIBAC III, Bilbao

8-20 août — 45th Session of the International Statistical Institute, Amsterdam

1986:

14-21 sep. — Sixth Pannonian Symposium on Mathematical Statistics, Bad Tatzmannsdorf, Burgenland, Autriche

3-16 nov. — XVI^{o} Reunion de la Sociedad Española de Investigación, Estadistica e Informatica, Malaga, Espagne

1987:

22-26 juin — NSF Workshop on Applied Probability: Extremes of Random Processes in Applied Probability, Santa Barbara, USA

6-12 déc. — Extremwerttheorie, Oberwolfach

1988:

22-26 août — 18th European Meeting of Statisticians, Berlin

1989:

3-7 juillet — Third Hungarian Colloquium on Limit Theorems in Probability and Statistics, Pécs, Hongrie

23-24 nov. — X^{ème} Rencontre Franco-Belge de Statisticiens, Bruxelles

1990:

1-4 avril — IMS Eastern Regional Meeting, Baltimore, USA, Special Invited Lecture

28 mai-1 juin — XXII^{èmes} Journées de Statistique, ASU, Tours

29 juil.-10 août Nato Advanced Study Institute Nonparametric Functional Estimation and

Related Topics, Spetses, Grèce
16 novembre Colloque de Biostatistique, Paris, France

1991:
18-19 mars Alfred Rényi Memorial Conference, Budapest
23-26 avril Fifth International Symposium on Applied Stochastic Models and Data Analysis, Granada, Espagne
4-8 mai International Symposium on Nonparametric Statistics and Related Topics, Ottawa
27-30 mai XXIII$^{\text{èmes}}$ Journées de Statistique, ASU, Strasbourg
19--21 juin 5$^{\text{th}}$ International Conference on Multiphase Production, Cannes

1992:
18-22 mai XXIV$^{\text{èmes}}$ Journées de Statistique, U.L.B., Bruxelles, Belgique
18-20 nov. XIII$^{\text{èmes}}$ Rencontres Franco-Belges Statisticiens, Villeneuve d'Ascq, Résultats Nouveaux en Théorie des Valeurs Extrêmes et Applications

1993:
25 avril-2 mai Geometria Stocastica, Corpi Convessi, Misure Empiriche, Palerme, Italie
2-7 mai Conference on Extreme Values Theory and its Applications, Gaithersburg, USA
16-18 juin 6$^{\text{th}}$ International Conference on Multiphase Production, Cannes, France
1-5 août Distributions with Fixed Marginals, Doubly Stochastic Measures and Markov Operators, Seattle, USA
16-21 août 9$^{\text{th}}$ Conference on Probability in Banach Spaces, Sandbjerg, Danemark

1994:
24-27 mai XXVI$^{\text{èmes}}$ Journées de Statistique, ASU, Neuchâtel, Suisse
8-9 juin Scientific Session at the Mathematical Institute of the Hungarian Academy of Sciences in celebration of Pál Révész's 60$^{\text{th}}$ birthday, Budapest, Hongrie
26-28 mai Multivariate Extreme Value Estimation with Applications in Economics and Finance, Rotterdam, Pays Bas

1995:
11-13 sept. IFU-Workshop 1995, Weimar, Allemagne
18-24 sept. COPAM 95, Fourth AMU Pan-African Congress of Mathematicians, Ifrane.

1996:
26-31 août 4$^{\text{th}}$ World Congress of the Bernoulli Society, Vienne, Autriche
8-9 sept. Seconda Conferenze Internazionale, Geometrica Stocastica Corpi Convessi, Misure Empiriche, Agrigento, Sicile
21-22 nov. XXVII$^{\text{ème}}$ Rencontre Franco-Belge de Statisticiens, Marne-la-Vallée

1997
26-30 mai XXIX$^{\text{èmes}}$ Journées de Statistique, Caracassonne
8-13 juillet ICAMPS'97, International Conference on Asymptotic Methods in Probability and Statistics, Ottawa, Canada.
10-13 sep. XXI$^{\text{o}}$ Convegno Annuale A.M.A.S.E.S., Rome, Italie.
17-20 sep. IFU--Workshop 1997, Istanbul, Turquie
13--24 oct. Symposium on Nonparametric Functional Estimation. Centre de Recherches

Mathématiques, CRM, Université de Montréal, Canada

1998:

13-24 juillet Random Walks - Workshop and Summer School, Paul Erdős Research Center for Mathematical Sciences, Budapest.

20-22 août Perspectives in Modern Statistical Inference, Parametrics, Semiparametrics, Nonparametrics, Prague.

1999:

28 juin-2 juil. Fourth Hungarian Colloquium on Limit Theorems in Probability and Statistics (J. Bolyai Mathematical Society), Balatonlelle, Hongrie.

22-24 oct. Statistical Models: Probabilistic Background and Inference, Cagliari, Italie.

2000:

20-25 mars Hamburg Stochastik Tage, Hamburg, Allemagne.

15-21 mai 5th World Congress of the Bernoulli Society, Guanajuato, Mexique.

29-31 mai International Workshop on Goodness-of-fit Tests and Validity of Models, Paris.

4-7 juillet Mathematical Methods in Reliability: Methodology, Practice and Inference (MMR-2000), Bordeaux.

2001:

14-18 mai 33èmes Journées de Statistique, Nantes

19-20 mai Colloquium in Honor of George Roussas, Davis, California

24-29 sep. IVth International Conference in Stochastic, Geometry, Convex Bodies, Empirical measures and Applications, Tropea, Italie.

7-9 déc. Characterizations, Modelling and Applications, Antalya, Turquie.

2002:

15-23 mai 34èmes Journées de Statistique, Bruxelles

17-22 juin Mathematical Methods in Reliability 2002, Trondheim, Norvège

24-28 juin The 3rd International Conference on High Dimensional Probability, Sandbjerg Estate, Danemark

14-17 août Perspectives in Modern Statistical Inference II, Brno, République Tchèque

2004 :

17-19 juin Conference in Honour of Endre Csáki, Budapest.

2005 :

17-25 mai Applied Stochastic Models and Data Analysis, Brest.

20-24 juin The 4th International Conference on High Dimensional Probability, Santa Fe, New Mexico.

14-15 octobre Conference in Honour of Wolfgang Wertz, Vienne

2006:

15-20 mai IWAP International Workshop in Applied Probability (Storrs, Connecticutt)

22-28 mai 37èmes Journées de Statistique, Paris.

7. Compléments personnels

- Famille :

Marié le 4 mars 1971 (civ.) &19 mars 1971 (rel.) avec Joële Cormerais

Quatre enfants:

Fleur	née le 19 juillet 1972
Sophie	née le 11 décembre 1975
Camille	née le 16 décembre 1979
Aurore	née le 22 avril 1987

- Loisirs :

Escalade rocheuse, pyrénéisme.