

Agathe Guilloux

LaMME - Université d'Evry - CNRS - Université Paris
Saclay
23 boulevard de France
91037 Evry Cedex

Née à Paris le 28 août 1978
Mariée, deux enfants
Nationalité française

email : agathe.guilloux@math.cnrs.fr

WWW : <http://www.math-evry.cnrs.fr/members/aguilloux/welcome>

Tel : 01 44 27 85 57

Situation actuelle

Professeur. Université d'Evry Val d'Essonne

Directrice. Laboratoire de mathématiques et modélisation d'Evry : LaMME - UMR CNRS 8171 - Université d'Evry val-d'Essonne - Université Paris Saclay

Responsable. Master 2 Data Science : santé, assurance, finance de Paris Saclay

Cursus universitaire et professionnel

Maître de Conférences. Université Pierre et Marie Curie (UPMC) (2005-2016)

Habilitation à diriger les recherches. Soutenue le **9/12/2010** à l'UPMC.

Contributions to nonparametric statistics for survival analysis

Jury : N. Keiding, V. Patilea, K. Tribouley, Z. Ying (Rapporteurs), L. Birgé (Président), G. Biau, F. Comte, J.Y. Dauxois (Examineurs).

ATER. (2004-2005) à l'IUT de Sceaux, Université Paris-Sud

Thèse de doctorat. Soutenue le **11/10/2004** à l'Université de Rennes 1

DEA de Mathématiques. (2000-2001), spécialité Statistique à l'Université de Rennes 1.

ENSAI. (1998-2001), spécialité Biostatistiques et Epidémiologie.

Projets, distinctions

Consortium "Variabilité génomique". Co-principal investigateur avec Anne-Sophie Jannot (Hôpital européen Georges Pompidou - Assistance Publique - Hôpitaux de Paris). Novembre 2018-

Master 2+2. Lauréat de l'appel à projet de l'Initiative BigData de Sorbonne Université, co-principal investigateur avec Anne-Sophie Jannot (Hôpital européen Georges Pompidou - Assistance Publique - Hôpitaux de Paris). Janvier 2016-Août 2016. Budget 8keuros.

Délégation au Centre de Mathématiques Appliquées. Ecole Polytechnique, Sept 15-Août 16.

INCA-DGOS. . Lauréat de l'appel à projet du programme de recherche translationnelle en cancérologie INCA-DGOS, co-principal investigateur avec Alex Duval (CdR Saint Antoine), Jean-Francois Fléjou et Tabassome Simon (APHP). Janvier 2015-Décembre 2017. Budget 636keuros.

Porteur de l'ANR Prognostic. ANR-09-JCJC-010 (Octobre 2009 - Septembre 2012). Point processes : non-parametric inference, learning and applications. Budget : 90keuros.

Partenariats, contrats

CEMKA/UEVE-Paris-Saclay/Paris Descartes. (CIFRE) : Prédiction dynamique de survie pour l'optimisation de planning de surveillance individualisé des patients atteints de cancer.

CEMKA/UEVE-Paris-Saclay/Paris Descartes. (CIFRE) Machine learning pour l'optimisation des poids des modèles structuraux marginaux dans le cadre de données longitudinales.

Ubisoft/UPMC/CNAM. (CIFRE) Approches big data pour la modélisation de la rétention et de la monétisation des joueurs.

MutRé/UPMC. (Contrat de recherche)

PwC/UPMC. (CIFRE) "Statistique non-paramétrique pour les calculateurs assurantiers de solvabilité".

Responsabilités administratives

- Directrice du Laboratoire de mathématiques et modélisation d'Evry : LaMME - UMR CNRS 8171 - Université d'Evry val-d'Essonne - Université Paris Saclay (2020-)
- Responsable du Master 2 Data Science : santé, assurance, finance de Paris Saclay (2017-)
- Membre élu du Conseil d'Administration de l'Université d'Evry-val-d'Essonne (2018-)
- Membre du Conseil scientifique de l'école doctorale SDSV (Paris Saclay)
- Membre élu du Conseil National des Universités de la 26-ième section (2015-2016)
- Membre élu du Conseil d'Administration de l'Institut Henri Poincaré (2015-2016)
- Membre du Comité de Pilotage de l'initiative "Big data" de Sorbonne Universités (2015-2016)
- Membre de la commission des thèses et habilitations de l'Ecole Doctorale des Sciences Mathématiques de Paris Centre (2012-2016)
- Membre élu du Conseil d'UFR (2009-2016)
- Membre des Comités de Sélection Evry (2018, 2019), Versailles (2019), Descartes (2019), UPMC (2006, 2007, 2008, 2009, 2010, 2014, 2015, 2016, 2018), Lille 1 (2017), INSA Toulouse (2017), INSA Rennes (2017), Lyon-1 et École Centrale de Lyon (2012), Paris-Descartes (2010), Paris-Ouest (2009 et 2010)
- Membre élu du Conseil de Direction de l'ISUP (2006-2017)
- Membre du Jury d'admission sur titre de l'ISUP (2009-2017)

Animation scientifique et collective

Présidente du comité HCERES. Unité MIAT (2019)

Editrice associée. International Journal of Biostatistics (2019-)

Génomique numérique. Comité scientifique du colloque (2019)

FRANCE INNOVATION. Groupe de travail « Impacts de l'IA sur les industries » (hiver 2018)

ICS Summer School. Comité scientifique de la semaine "Data Science" avec Patrick Gallinari et Sylvie Thiria (été 2016)

Institut des actuaires. Coordinateur du groupe de travail "Big data" (2015-2016)

Société Française de Statistique. Membre élu du bureau du Groupe "Statistique Mathématique" de la SFDS (jusqu'à 2016)

European research consortium for Informatics and Mathematics. Membre du Groupe "Statistical Analysis of Event Time" de l'ERCIM

Workshop SAM. Membre du comité d'organisation de SAM ("The statistical analysis of multi-outcome data") organisé à l'UPMC les 9 et 10 juillet 2012.

Rapport sur articles. Annals of Statistics, Journal of Multivariate Analysis, International Journal of Biostatistics, Bernoulli, Electronic Journal of Statistics, Journal of Statistical Planning and Inference, Scandinavian Journal of Statistics, NonParametric Statistics, ALEA, Life Time Data Analysis, ...

Encadrement scientifique

Thèses.

- Co-encadrement de la thèse CIFRE d'Aziliz Cottin avec Sandrine Katshaian (Descartes) depuis décembre 2018, *Prédiction dynamique de survie pour l'optimisation de planning de surveillance individualisé des patients atteints de cancer.*
- Co-encadrement de la thèse CIFRE de Camille Nevoret avec Sandrine Katshaian (Descartes) depuis décembre 2018, *Machine learning pour l'optimisation des poids des modèles structuraux marginaux dans le cadre de données longitudinales.*
- Co-encadrement de la thèse de Alain Virouleau avec Stéphane Gaïffas (CMAP - X) depuis septembre 2015, *Outliers with high dimensional data*
- Co-encadrement (50%) de la thèse CIFRE de Quang Dien Duong avec Olivier Lopez (UPMC), depuis décembre 2015, *Statistique non-paramétrique pour les calculateurs assurentiels de solvabilité.*
- Co-encadrement de la thèse de Simon Bussy avec Anne-Sophie Jannot (AP-HP) depuis septembre 2015, *Données temporelles massives en apprentissage supervisé et application à la cohorte ARTEMIS*
- Co-encadrement de la thèse de Thibault Allart avec Stéphane Natkin, Guillaume Levieux (CNAM) et Michel Pierrefite (Ubisoft) depuis décembre 2014, *Approches big data pour la modélisation de la rétention et de la monétisation des joueurs*
- Co-encadrement de la thèse de Mokhar Alaya avec Stéphane Gaïffas (CMAP, X) depuis septembre 2012, *Segmentation pour l'intensité de processus ponctuels*
- Co-encadrement de la thèse de Sarah Lemler avec Marie-Luce Taupin (UEVE) - depuis septembre 2011, *Réduction de dimension en analyse de survie. Application en statistique génétique et génomique.*
- Co-encadrement de la thèse de Ségolen Geffray (dirigée par Paul Deheuvels, LSTA -UPMC), soutenue le 3 novembre 2006, *Estimation non-paramétrique de données censurées dans un cadre multi-états.*

Ingénieurs d'étude.

- Encadrement (2017-) des travaux de Adeline Denis dans le cadre d'une collaboration avec l'Institut Institut Imagine - INSERM
- Encadrement (2016) des travaux de Malek Boumelil dans le cadre du contrat MutRé/UPMC

Mémoires et stages.

- Encadrement (2019) du stage de M1 de Katherine Morales *Screening with a partially observed Gold standard*
- Encadrement (2019) du stage de M1 de Thomas Marcoux *Machine learning pour l'Analyse et prédiction du phénomène d'exon skipping dans le cadre de cancers MSI*
- Encadrement (2018) du stage de 2A (ENSIIE) de Mayssa Labidi *Données MIMIC 3 : deep learning*
- Encadrement (2018) du stage de L3 d'Ayoub Abraich *Données MIMIC 3 : benchmarking*
- Encadrement (2017) du mémoire de M2 de Kossi Doulabé *Comparaisons des algorithmes d'estimation pour le modèle LDA*
- Co-encadrement (2017) avec Emmanuelle Six (Institut Imagine - INSERM) du stage de Romain Villette *Analyse de données pour la mise en évidence de l'hétérogénéité de l'hématopoïèse humaine*

- Co-encadrement (2016) avec Emmanuelle Six (Institut Imagine - INSERM) du stage de 3^{ème} année de l'ENSAE/Master BigData d'Arnaud Lecoules *Modélisation statistique pour l'hématopoïèse*
- Co-encadrement (2016) avec Anne-Sophie Jannot (AP-HP) du stage de 3^{ème} année de l'X de Emeline Fay *Score de propension, quelles limites pour leur utilisation ?*
- Co-encadrement (2016) avec Anne-Sophie Jannot (AP-HP) des stages de M2/AgroParisTech/ENSAI (dans le cadre du projet Master 2+2 de l'Initiative Big Data de Sorbonne Université) de Clément Morel (3A ENSAI) et Christophe Botella *Apprentissage statistique sur données d'entrepôt*
- Encadrement (2014-2015) du stage de M2 d'Alain Virouleau *Détection de valeurs aberrantes dans des modèles linéaires généralisés, application à des données génétiques.*
- Encadrement (2014-2015) du stage de fin d'étude de l'ISUP de Sami Lagnaoui *Apports du machine learning et de l'open-data pour la modélisation du risque en assurance santé.*
- Encadrement (2014-2015) du stage de S3 de l'Ecole des Mines de Paris de Soufiane Mourragui *Statistique pour la diffusion dans les réseaux sociaux*

Jurys de thèse et d'HDR.

- Membre du jury de la thèse de Babacar Diallo - Université Paris Saclay (2019)
- Rapporteur de la thèse de Pascal Lu - Sorbonne Université (2019)
- Rapporteur de la thèse de Camille Sabathé - Université de Bordeaux ISPED (2019)
- Membre du jury de l'HDR d'Olivier Bouaziz - Université Paris Descartes (2018)
- Rapporteur de la thèse de Paulien Jeunesse - Université Paris Dauphine (2018)
- Rapporteur de l'HDR de Eve Leconte - Toulouse School on Economics (2017)
- Rapporteur de l'HDR de Julien Stirnemann - Univ. Paris Descartes (2017)
- Membre du jury de la thèse de Thibault Allart - Conservatoire National des Arts et Métiers (2017)
- Membre du jury de la thèse de Guillaume Biessy - Univ. d'Evry (2016)
- Membre du jury de la thèse de Quentin Grimonprez - Univ. Lille 1 (2016)
- Membre du jury de la thèse de Gwenaëlle Mabon - Univ. Paris Descartes (2016)
- Membre du jury de la thèse de Mokhtar Alaya - Univ. Pierre et Marie Curie (2016)
- Membre du jury de la thèse de Sarah Lemler - Univ. Evry (2014)
- Membre du jury de la thèse de Cécile Chauvel - Univ. Pierre et Marie Curie (2014)
- Membre du jury de la thèse de Svetlana Gribkova - Univ. Pierre et Marie Curie (2014)
- Rapporteur de la thèse de Paul Blanche - Univ. Bordeaux (2013)
- Membre du jury de la thèse de Gaele Chagny - Univ. Paris-Descartes (2013)
- Membre du jury de la thèse de Margarita Moreno - Univ. Paris-Sud (2013)
- Membre du jury de la thèse d'Abdullah Oueslati - Univ. Pierre et Marie Curie (2013)
- Rapporteur de la thèse de Damien Bousquet - Univ. Montpellier 2 (2012)
- Membre du jury de la thèse de Juliette Penichoux - Univ. Paris-Sud (2012)
- Membre du jury de la thèse d'Alexis Flesch - Univ. de Franche Comté (2012)
- Membre du jury de la thèse de Segolen Geffray - Univ. Pierre et Marie Curie (2006)

Publication

Articles publiés

- “Clonal tracking in gene therapy patients reveals a diversity of human hematopoietic differentiation programs” (2020). In : *Blood*.
- Alaya, M. Z., S. Bussy, S. Gaïffas et A. Guilloux (2019). “Binarsity : a penalization for one-hot encoded features in linear supervised learning”. In : *Journal of Machine Learning Research* 20.118, p. 1-34.
- Bussy, S., R. Veil, V. Looten, A. Burgun, S. Gaïffas, A. Guilloux, B. Ranque et A.-S. Jannot (2019). “Comparison of methods for early-readmission prediction in a high-dimensional heterogeneous covariates and time-to-event outcome framework”. In : *BMC medical research methodology* 19.1, p. 50.
- Morel, M., E. Bacry, S. Gaïffas, A. Guilloux et F. Leroy (2019). “ConvSCCS : convolutional self-controlled case-seris model for lagged adverser event detection.” In : *Biostatistics*, kxz003.

- Cohen, R., E. Hain, O. Buhard, A. Guilloux, A. Bardier, R. Kaci, P. Bertheau, F. Renaud, F. Bibeau, J.-F. Fléjou, T. Andre, M. Svrcek et A. Duval (2018). “Association of Primary Resistance to Immune Checkpoint Inhibitors of Metastatic Colorectal Cancer With Misdiagnosis of Microsatellite Instability or Mismatch Repair Deficiency Status”. In : *JAMA Oncology*.
- Jonchere, V., L. Marisa, M. Greene, A. Virouleau, O. Buhard, R. Bertrand, M. Svrcek, P. Cervera, A. Goloudina, E. Guillermin et al. (2018). “Identification of Positively and Negatively Selected Driver Gene Mutations Associated with Colorectal Cancer with Microsatellite Instability”. In : *Cellular and Molecular Gastroenterology and Hepatology*.
- Bussy, S., S. Gaïffas, A. Guilloux et A. Jannot (2017). “C-mix : a high dimensional mixture model for censored durations, with applications to genetic data”. In : *Statistical Methods for Medical Research*.
- Buhard, O., A. Lagrange, A. Guilloux, C. Colas, M. Chouchène, K. Wanherdrick, F. Coulet, E. Guillermin, C. Dorard, L. Marisa et al. (2016). “HSP110 T17 simplifies and improves the microsatellite instability testing in patients with colorectal cancer”. In : *Journal of Medical Genetics*.
- Guilloux, A., S. Lemler et M.-L. Taupin (2016a). “Adaptive estimation of the baseline function in a high-dimensional Cox model via model selection”. In : *Journal of Statistical Planning and Inference*.
- (2016b). “Adaptive kernel estimation of the baseline function in the Cox model, with high-dimensional covariates”. In : *Journal of Multivariate Analysis*.
- Alaya, M., S. Gaïffas et A. Guilloux (sept. 2015). “Learning the Intensity of Time Events With Change-Points”. In : *IEEE Transactions on Information Theory* 61.9, p. 5148-5171.
- Bodo, S., C. Colas, O. Buhard, A. Collura, J. Tinat, N. Lavoine, A. Guilloux, A. Chalastanis, P. Lafitte, F. Coulet et al. (2015). “Diagnosis of Constitutional Mismatch Repair-Deficiency Syndrome Based on Microsatellite Instability and Lymphocyte Tolerance to Methylating Agents”. In : *Gastroenterology* 149.4, p. 1017-1029.
- Bouaziz, O. et A. Guilloux (2014). “A penalized algorithm for event-specific rate models for recurrent events”. In : *Biostatistics*, kxu046.
- Collura, A., A. Lagrange, M. Svrcek, L. Marisa, O. Buhard, A. Guilloux, K. Wanherdrick, C. Dorard, A. Taieb, A. Saget et al. (2014). “Patients With Colorectal Tumors With Microsatellite Instability and Large Deletions in HSP110 T 17 Have Improved Response to 5-Fluorouracil-Based Chemotherapy”. In : *Gastroenterology* 146.2, p. 401-411.
- Dauxois, J.-Y., A. Guilloux et S. N. Kirmani (2014). “Estimation in a competing risks proportional hazards model under length-biased sampling with censoring”. In : *Lifetime data analysis* 20.2, p. 276-302.
- Bouaziz, O., F. Comte et A. Guilloux (2013). “Nonparametric estimation of the intensity function of a recurrent event process”. In : *Statistica Sinica* 23.2, p. 635-665.
- Bouaziz, O. et A. Guilloux (2013). “A lasso estimator for event-specific rate models for recurrent events”. In : *DYNSTOCH 2013*.
- Brunel, E., F. Comte et A. Guilloux (2013). “Nonparametric estimation for survival data with censoring indicators missing at random”. In : *Journal of Statistical Planning and Inference* 143.10, p. 1653-1671.
- Gaïffas, S. et A. Guilloux (2012). “High-dimensional additive hazards models and the lasso”. In : *Electronic Journal of Statistics* 6, p. 522-546.
- Comte, F., S. Gaïffas et A. Guilloux (2011). “Adaptive estimation of the conditional intensity of marker-dependent counting processes”. In : *Annales de l'institut Henri Poincaré (B)* 47.4, p. 1171-1196.
- Geffray, S. et A. Guilloux (2011). “Maximum likelihood estimator for cumulative incidence functions under proportionality constraint”. In : *Sankhya A* 73.2, p. 303-328.
- Zaanan, A., P. Cuilliere-Dartigues, A. Guilloux, Y. Parc, C. Louvet, A. DeGramont, E. Tiret, S. Dumont, B. Gayet et P. Validire (2010). “Impact of p53 expression and microsatellite instability on stage III colon cancer disease-free survival in patients treated by 5-fluorouracil and leucovorin with or without oxaliplatin”. In : *Annals of oncology* 21.4, p. 772-780.
- Brunel, E., F. Comte et A. Guilloux (2009). “Nonparametric density estimation in presence of bias and censoring”. In : *Test* 18.1, p. 166-194.

- El-Bchiri, J., A. Guilloux, P. Dartigues, E. Loire, D. Mercier, O. Buhard, I. Sobhani, P. de la Grange, D. Auboeuf, F. Praz et al. (2008). “Nonsense-mediated mRNA decay impacts MSI-driven carcinogenesis and anti-tumor immunity in colorectal cancers”. In : *PLoS One* 3.7, e2583.
- Brunel, E., F. Comte et A. Guilloux (2008). “Estimation Strategies for Censored Lifetimes with a Lexis-Diagram Type Model”. In : *Scandinavian Journal of Statistics* 35.3, p. 557-576.
- Dauxois, J.-Y. et A. Guilloux (2008). “Nonparametric inference under competing risks and selection-biased sampling”. In : *Journal of Multivariate Analysis* 99.4, p. 589-605.
- Guilloux, A. et P. Saint-Pierre (2008). “Estimateur de la fonction de répartition bivariée avec censures à droite et à gauche”. In : 52.1-2, p. 157-164.
- El Bichri, J., A. Guilloux, P. CuillereDartigues, O. Buhard, S. Legouvello, J. Flejou et A. Duval (2007). “Functional consequences due to differential frameshift mutation-derived mRNAs decay in colorectal cancers”. In : *Molecular Cancer Therapeutics* 6.11 Supplement, p. C247-C247.
- Guilloux, A. (2007). “Nonparametric estimation for censored lifetimes suffering from unknown selection bias”. In : *Mathematical Methods of Statistics* 16.3, p. 202-216.
- Guilloux, A. (2006). “Estimation sous biais de sélection et avec fonction de poids inconnue”. In : *Comptes Rendus Mathématique* 342.4, p. 275-278.
- Dauxois, J.-Y., A. Guilloux et S. N. Kirmani (2005). “Nonparametric estimation from proportional hazards competing risks data under selection bias”. In : *Nonparametric Statistics* 17.6, p. 717-731.
- Geffray, S. et A. Guilloux (2005). “Estimation dans un modèle de Koziol–Green généralisé”. In : *Comptes Rendus Mathématique* 341.1, p. 49-52.
- Dauxois, J.-Y., A. Guilloux et S. N. Kirmani (2003). “Estimation de la fonction de survie sous biais de longueur pour des risques concurrents et proportionnels”. In : *Comptes Rendus Mathématique* 337.11, p. 745-748.
- Bray, F., A. Guilloux, R. Sankila et D. M. Parkin (2002). “Practical implications of imposing a new world standard population”. In : *Cancer causes & control* 13.2, p. 175-182.
- Lambert, R., A. Guilloux, A. Oshima, V. Pompe-Kirn, F. Bray, M. Parkin, W. Ajiki et H. Tsukuma (2002). “Incidence and mortality from stomach cancer in Japan, Slovenia and the USA”. In : *International Journal of Cancer* 97.6, p. 811-818.

Prépublications sur ArXiv

- Bussy, S., M. Z. Alaya, A.-S. Jannot et A. Guilloux (2020). *Binacox : automatic cut-point detection in high-dimensional Cox model with applications in genetics*. arXiv : arXiv:1807.09813.
- Virouleau, A., A. Guilloux, S. Gaïffas et M. Bogdan (2017). *High-dimensional robust regression and outliers detection with SLOPE*. arXiv : arXiv:1712.02640.
- Achab, M., A. Guilloux, S. Gaïffas et E. Bacry (2015). *SGD with Variance Reduction beyond Empirical Risk Minimization*. arXiv : arXiv:1510.04822.
- Gaïffas, S. et A. Guilloux (2009). *Learning and adaptive estimation for marker-dependent counting processes*. arXiv : arXiv:0909.5342.
- Comte, F., S. Gaïffas et A. Guilloux (2008). *Adaptive estimation of the conditional intensity of marker-dependent counting processes*. arXiv : arXiv:0810.4263.

Prépublications

- Alaya, M., S. Bussy et A. Guilloux (2018). *Features binarisation for Cox models*.

Chapitres dans des ouvrages collectifs

- Brunel, E., F. Comte et A. Guilloux (2014). “Estimation/Imputation strategies for missing data in survival analysis”. In : *Statistical Models and Methods for Reliability and Survival Analysis*. John Wiley & Sons, Inc., Hoboken, USA, p. 229-252.
- Guilloux, A. (2008). “Sampling in Survival Analysis and Estimation with Unknown Selection Bias and Censoring”. In : *Statistical Models and Methods for Biomedical and Technical Systems*. Springer, p. 213-224.

Articles dans des actes de conférences

- Allart, T., G. Levieux, M. Pierfitte, A. Guilloux et S. Natkin (2017). “Difficulty influence on motivation over time in video games using survival analysis”. In : *Proceedings of the 12th International Conference on the Foundations of Digital Games*. ACM, p. 2.
- (2016). “Design influence on player retention : A method based on time varying survival analysis”. In : *Computational Intelligence and Games (CIG), 2016 IEEE Conference on*. IEEE, p. 1-8.
- Comte, F., S. Gaïffas et A. Guilloux (2009). “Estimation adaptative pour des processus de comptage dépendant de covariables”. In : *41èmes Journées de Statistique, SFdS, Bordeaux*.
- Gaïffas, S., A. Guilloux et al. (2009). “Apprentissage pour des processus de comptage avec covariables”. In : *41èmes Journées de Statistique, SFdS, Bordeaux*.
- Saint-Pierre, P. et A. Guilloux (2009). “Fonction de survie bivariable de variables censurées à droite et à gauche”. In : *41èmes Journées de Statistique, SFdS, Bordeaux*.