

LA FORMATION À LA STATISTIQUE DES ENSEIGNANTS-CHERCHEURS

L'EXPÉRIENCE DES JOURNÉES D'ÉTUDE EN STATISTIQUE (JES)

Jean-Jacques DROESBEKE¹

TITLE

Training of teachers-cum-researchers in statistics. The experience of the Journées d'Étude en Statistique (JES)

RÉSUMÉ

L'objectif de cet article est à la fois d'ordre historique et pédagogique. Il s'agit de montrer comment les *Journées d'Étude en Statistique* (JES) ont été fondées par l'*Association des Statisticiens Universitaires* (ASU), ancêtre de la *Société Française de Statistique*, avec quel objectif et selon quels critères. L'évolution de cette organisation jusqu'à nos jours, son influence sur des projets européens (*European Courses in Advanced Statistics* – ECAS et *Training of European Statisticians* – TES) et sud-américains (*Programme de Recherche et d'Enseignement en Statistique Appliquée* – PRESTA) seront aussi évoquées, en concentrant notre attention sur le contenu et les méthodes d'enseignement utilisées.

Mots-clés : enseignement, statistique, enseignants-chercheurs, praticiens.

ABSTRACT

The purpose of this paper is historical and educational. We want to point out how the *Journées d'Étude en Statistique* (JES) were created by the *Association des Statisticiens Universitaires* (ASU), ancestor of the *Société Française de Statistique*, with what objectives and criteria. The evolution of this project, its influence on European programmes (*European Courses in Advanced Statistics* – ECAS and *Training of European Statisticians* – TES) and South-American ones (*Programme de Recherche et d'Enseignement en Statistique Appliquée* – PRESTA) will also be considered, focusing on the contents and teaching methods.

Keywords : teaching, statistics, teachers-cum-researchers, practicing.

1 Préambule

Afin de rendre hommage à tous ceux qui ont contribué à l'essor des *Journées d'Étude en Statistique* et des formations en statistique qui ont suivi dans le temps, nous avons délibérément choisi d'écrire cet article² en prenant un petit air de témoignage, usant du « nous de modestie » comme il se doit. Nous avons aimé travailler avec eux !

¹ Université libre de Bruxelles, LMTD, CP 139, 50, Avenue F.D. Roosevelt, 1050 Bruxelles, Belgique, jjdroesb@ulb.ac.be

² Nous remercions Anne Gégout-Petit, Présidente du comité scientifique du CFIES 2017, de nous avoir suggéré de développer ce thème au cours de ce colloque qui s'est tenu à Grenoble, en septembre 2017.

2 La création des Journées d'Étude en Statistique

Les *Journées d'Étude en Statistique* (JES) sont créées au début des années 1980 par le bureau de l'*Association des Statisticiens Universitaires* (ASU). Cette société savante trouve son origine dans une réunion d'une trentaine de statisticiens, essentiellement des professeurs d'université, orientés surtout vers la statistique appliquée, qui souhaitent établir entre eux des relations plus structurées. Il s'agit aussi de réfléchir au contenu des enseignements de statistique – dans les cursus universitaires mais aussi dans le secondaire – et d'étudier la manière de participer aux réformes des programmes. Cette réunion se déroule en 1969, à Toulouse, à l'initiative du professeur Roger Huron.

Le projet de constituer une association fait l'objet d'une nouvelle réunion organisée à nouveau à Toulouse, le 13 juin 1970, au cours de laquelle les quarante-trois participants élaborent des statuts et discutent programmes et méthodes d'enseignement, liaisons entre universités et secteurs de production, relations avec l'étranger et carrières dans la discipline. Tout est en place pour constituer l'*Association des Statisticiens Universitaires* – que nous désignerons dans la suite uniquement par son acronyme ASU. Son premier colloque se déroule à Lyon, les 6 et 7 novembre 1970.

Les *Journées de Statistique* (JdS) sont organisées depuis cette date tous les ans – sauf en 1973 où elles ont lieu à deux reprises. Elles s'ouvrent à de jeunes chercheurs en les invitant à présenter leurs travaux, ce qui va accroître le nombre de participants... et d'exposés. Les Journées organisées à Nancy (1974) et Montpellier (1975) permettent d'initialiser cette ouverture. Depuis lors, les *Journées de Statistique* se déroulent chaque année fin mai³, début juin. Organisées dans un premier temps uniquement en France (voir Droesbeke, 2005), elles quittent pour la première fois l'hexagone en 1982, avec les JdS de Bruxelles.

C'est en 1983, sous la présidence de Michel Depaix, professeur à Nancy, que les *Journées d'Étude en Statistique* sont conçues. Lors d'une réunion du Bureau de l'ASU à Jussieu, les organisateurs des JdS présentent la liste des conférenciers invités. Parmi ceux-ci, Frank Hampel – professeur émérite à l'Institut fédéral suisse de Technologie ETH de Zurich – propose, dans sa lettre d'acceptation à faire un exposé, de dispenser aussi un cours (*tutorial*) durant le congrès. Après une longue discussion au sein du Bureau de l'association, il apparaît que cette proposition n'est pas acceptable dans la mesure où la semaine annuelle de rencontre est entièrement consacrée aux exposés et aux réunions. Il est tout à fait impossible d'y inclure un cours se déroulant pendant plusieurs jours. C'est tout autant impossible de l'organiser dans la semaine qui précède ou celle qui suit : l'ensemble serait trop long pour la grande majorité des participants.

Mais l'idée d'organiser une semaine de cours pour les enseignants-chercheurs des universités et des centres de recherche privés et publics séduit de nombreux membres du Bureau. Une proposition fuse : organiser en octobre 1984 une semaine de cours à leur intention, coupée du congrès annuel, les JdS. Quel nom peut-on donner à cette initiative ? Quel thème va-t-on choisir ? En quel lieu l'organiser ? Qui va se charger de mettre ce projet en œuvre ? Il est ainsi décidé de créer les *Journées d'Étude en Statistique* (JES). Un thème est proposé par Georges Le Calvé, professeur à Rennes : « *Choisissons les séries chronologiques, thème souvent méconnu des enseignants, dont moi-même* ». Le lieu est suggéré par Philippe Picard, professeur à Lyon, qui signale que le *Centre International de Rencontres*

³ Ce fut déjà le cas en 1971, à Orléans, mais dans leur première version, moins scientifique.

J.-J. Droesbeke

Mathématiques (CIRM), nouvellement créé à Luminy, près de Marseille, est à la recherche de candidatures pour organiser des rencontres sur une semaine. Quant au comité d'organisation, c'est parmi ses membres que le Bureau décide de le constituer. Vu le thème proposé, il s'adresse à nous, qui en avons fait notre domaine de recherche. Il se tourne ensuite vers Bernard Fichet car son lieu de travail est proche du CIRM. Il estime enfin qu'un comité chargé de prendre en compte tous les aspects logistiques et scientifiques du projet, se doit de comporter trois membres vu l'importance des tâches à réaliser. Marie-Jeanne Laurent-Duhamel, professeur à Pau et ancienne présidente de l'ASU, convainc Philippe Tassi, qui dirige à l'époque l'ENSAE, de se joindre aux deux autres en raison de son intérêt pour le sujet proposé⁴. Bruxelles, Marseille, Paris : la couverture géographique apportée par ce trio d'organisateur est large et présentera de multiples avantages.



FIGURE 1 – *Le CIRM à Luminy de nos jours* (CIRM|CNRS Le Journal)

3 L'organisation des JES

Pour les trois organisateurs, l'objectif est clair. Il s'agit d'offrir aux membres de l'ASU une formation de haut niveau, s'étendant sur une semaine et portant sur un thème intéressant de la méthodologie statistique. Pour ce premier projet, le titre est donc choisi par le Bureau de l'ASU : *Séries chronologiques. Théorie et pratique des modèles ARIMA*. Le public visé est, rappelons-le, le monde universitaire, et les centres de recherche des secteurs public et privé.

Rassembler durant une semaine, en automne 1984, dans un lieu adapté à ce type d'organisation – puisque le CIRM a marqué son accord pour accueillir les Journées – une bonne quarantaine de statisticiens venant d'horizons divers et désireux de se familiariser au thème proposé ou de l'approfondir, voilà donc l'objectif premier des trois organisateurs. Ceux-ci proposent que les deux premières journées de la semaine permettent de rentrer dans le sujet : objectifs des méthodes proposées, notations couramment utilisées, évolution historique de ces méthodes, concepts et résultats fondamentaux... Le mardi soir, les disparités entre les connaissances des participants sur le thème ont fortement diminué. Les deux journées suivantes doivent aborder des développements récents du sujet traité. Il faut garder le cap pour ne pas être perdu ! Quant à la dernière journée, elle doit principalement concerner des applications significatives des méthodes présentées durant la semaine.

⁴ Merci à Bernard Fichet de nous avoir rappelé toutes ces précisions !

Il faut à présent choisir un horaire de travail et... des conférenciers. Les trois organisateurs se mettent d'accord sur un horaire qui ne changera pratiquement plus dans la suite : deux séances de quatre-vingt-dix minutes séparées par une petite pause pour chaque matinée. Après le repas de midi, un temps libre de deux heures. On discute, on refait des démonstrations, on relit le polycopié rédigé par les conférenciers et remis aux participants en début de semaine... pour les uns ; on se détend dans les calanques, on fait une partie de ping-pong... pour les autres. Ce qui est certain, c'est qu'on reprend les deux heures de cours qui restent à suivre en fin de journée avec l'esprit bien disponible. Le mercredi après-midi doit permettre de s'adonner à des activités de détente plus importantes car aucun cours n'est prévu en ce milieu de semaine. Bernard Fichet organise une petite croisière dans les calanques : la pluie, absente les autres jours, empêche ce projet convivial de se réaliser ! La pause du vendredi en début d'après-midi n'a plus de raison d'être ; la semaine se termine à 16 heures. Cela permet aux participants de reprendre leur train ou leur avion pour rentrer chez eux le vendredi soir !

Le choix des conférenciers repose sur plusieurs critères : inviter des enseignants-chercheurs reconnus dans le thème choisi, bons pédagogues, capables de s'insérer dans un groupe de quatre à six orateurs et disponibles le plus possible pendant les séances de cours mais aussi en dehors de celles-ci. L'équipe est vite constituée. La manière dont les organisateurs ont été choisis par le Bureau de l'ASU facilite les choses : nous n'avons aucun mal à convaincre Marc Hallin et Guy Mélard, deux collègues qui travaillent dans le même service que nous à l'Université libre de Bruxelles ; il en est de même pour Philippe Tassi qui, par un cheminement parallèle, obtient l'accord de Christian Gouriéroux, en poste à l'université de Lille à ce moment-là, et Alain Monfort de l'INSEE. Un dernier « larron », Bernard Coutrot, de la COREF, vient compléter l'équipe des enseignants⁵. Il est relativement facile de leur demander de s'engager à pérenniser leur intervention en rédigeant un ou plusieurs chapitres d'un ouvrage collectif destiné à la communauté des statisticiennes et statisticiens et qui, dans un premier temps, est diffusé durant les Journées sous la forme d'un manuel de cours.

La formule semble plaire. Le bureau de l'ASU⁶ sollicite les trois organisateurs pour renouveler l'expérience avec un nouveau thème. Ce sera le tour des *Sondages* en octobre 1986. Depuis lors, tous les deux ans, les années paires, les *Journées d'Étude en Statistique* sont organisées dans le sud de la France. Les thèmes successifs proposés témoignent de la volonté des organisateurs de ces journées de varier les thèmes proposés : *Séries chronologiques. Théorie et pratique des modèles ARIMA – Les sondages – Analyse statistique des durées de vie. Modélisation des données censurées – Modèles pour l'analyse des données multidimensionnelles – Modélisation ARCH. Théorie statistique et applications dans le domaine de la finance – Inférence non paramétrique. Les statistiques de rang – Plans d'expérience. Applications à l'entreprise – Méthodes bayésiennes en statistique – Modèles statistiques pour données qualitatives – Analyse statistique des données spatiales – Analyse statistique des données longitudinales – Approches non paramétriques en régression – Modèles à variables latentes et modèles de mélange – Analyse statistique du risque – Méthodes robustes en statistique – Choix et agrégation de modèles – Apprentissage statistique et données massives*. À ce jour – ceci est écrit en 2017 – quatre-vingt-huit

⁵ Il publie cette année-là, en collaboration avec nous, un *Que-sais-je ?* consacré aux *méthodes de prévision*.

⁶ Qui devient l'*Association pour la Statistique et ses Utilisations* en 1987, nom proposé par Claude Langrand, professeur à Lille.

J.-J. Droesbeke

conférenciers ont prêté leur compétence à ce projet. Leur nom et le titre de leur contribution aux ouvrages issus de ces Journées sont présentés en annexe.



FIGURE 2 – Les seize premiers ouvrages de la collection des JES

Les quatre premiers ouvrages issus des JES ont été édités par Economica. Les deux suivants par les Editions de l'Université de Bruxelles, conjointement avec Ellipses. Depuis les

La formation à la statistique des enseignants-chercheurs

septièmes *Journées d'Étude en Statistique*, les ouvrages de la collection sont édités par Technip (voir figure 2).

Les trois organisateurs choisis au départ poursuivent leur activité de collaboration jusqu'au début des années quatre-vingt-dix. Mais les événements survenus dans leur parcours personnel ne leur permettent pas de poursuivre l'aventure commune. De ce groupe initial, nous restons seul en fonction. Il faut recomposer le comité d'organisation. Nous nous associons d'abord avec Jeanne Fine, professeur à Toulouse, pour les JES de 1994 à 1998. Pour la première fois, le comité d'organisation ne comporte que deux membres. En 1995, Gilbert Saporta, professeur au CNAM, à Paris, nous rejoint. En 1999, les activités de Jeanne Fine ne lui permettent plus de rester « gentille organisatrice ». Elle est remplacée en 1999 par Michel Lejeune, professeur à Grenoble, qui renonce à son tour, quelques années plus tard, à cette lourde tâche. Gilbert Saporta et nous, organisons seuls les Journées de 2006. Nous sommes rejoints par Christine Thomas-Agnan, professeur à Toulouse, pour les Journées de 2008 et les suivantes. Myriam Maumy-Bertrand, professeur à Strasbourg, intègre le comité d'organisation en 2015 ; nous le quittons la même année.

Comme on peut l'imaginer, le travail de publication est loin d'être facile. De nombreux échanges de courrier sont nécessaires entre les conférenciers et les organisateurs qui relisent l'entièreté des textes pour assurer cohérence et une certaine homogénéité. Comme un livre ne peut pas être publié avant que tous les auteurs n'aient rentré leur manuscrit, certains ouvrages ont mis du temps à naître. C'est ainsi que le premier livre paru est celui consacré aux sondages, en 1987, deux ans avant celui portant sur les séries chronologiques.

Seize ouvrages sont parus dans la collection des *Journées d'Étude en Statistique*, édités d'abord sous l'égide de l'ASU puis de la Société Française de Statistique (SFdS). Un dix-septième est prévu pour 2018. Après s'être déroulées au CIRM, à Luminy, près de Marseille, jusqu'en 2010, ces formations à la statistique sont, depuis lors, organisées au Centre de vacances du CNRS à Fréjus.



FIGURE 3 – *Le centre de vacances du CNRS à Fréjus*
(<https://www.caes.cnrs.fr/sejours/la-villa-clythia/>)

Le principe de ces formations est donc toujours le même. On choisit un thème d'actualité. On construit le contenu avec une équipe de conférenciers de renom, bons pédagogues dans la mesure du possible, acceptant des directives destinées à harmoniser les exposés. Partant des fondements du thème et de son histoire, on présente des résultats récents en milieu de

J.-J. Droesbeke

semaine, avant de finir par des exemples d'applications. Un manuel de cours, première mouture de l'ouvrage à paraître ultérieurement, sert de document de travail.

Plusieurs centaines d'enseignements-chercheurs ont participé à ces *Journées d'Étude en Statistique*. Les ouvrages en ont informé d'autres. Mais l'impact des JES est en réalité indirectement bien plus large comme on va pouvoir s'en rendre compte dans la suite.

4 Une première conséquence des JES : l'ECAS

Poussés par l'enthousiasme qui les imprègne à la suite des premières *Journées d'Étude en Statistique* de 1984, les trois organisateurs obtiennent du Bureau de l'ASU, présidé par Yves Escoufier, un mandat très clair en 1985 : prendre les contacts nécessaires avec les sociétés de statistique d'Allemagne, d'Italie et des Pays-Bas pour tenter d'organiser à un niveau européen une formation à la statistique en anglais, dont le fonctionnement soit analogue à celui des JES. Le cas de la Belgique est plus compliqué dans la mesure où la Société Belge de Statistique n'a plus d'activité depuis de nombreuses années. Nous contactons la *Société pour la Gestion scientifique des Entreprises* (SOGESCI) dont le siège est à Bruxelles et qui, à côté de ses activités principalement consacrées à la recherche opérationnelle, soutient un certain nombre d'initiatives d'ordre statistique. Nous obtenons aussi de la Commission européenne un soutien financier pour mener à bien ce projet.

Les discussions bilatérales portent leurs fruits : l'*European Courses in Advanced Statistics* (ECAS) est créé en tant qu'association de fait, son siège se situe à Bruxelles et son premier Conseil réunit les trois organisateurs des JES, auxquels se joignent Adriaan Abrahamse (Pays-Bas), Siegfried Heiler et Norbert Victor (Allemagne), Carlo Lauro et Alfredo Rizzi (Italie). La première session de l'ECAS est mise sur pied en 1987 par Carlo Lauro, professeur à Naples. Le thème retenu est *Methods for Multidimensional Data Analysis*.

Depuis lors, tous les deux ans, les années impaires, une session de l'ECAS est organisée dans un des pays soutenant le projet. La France, l'Allemagne, l'Italie, les Pays-Bas ne restent pas longtemps seuls. Ces pays sont rejoints ultérieurement par la Belgique⁷, l'Espagne, la Finlande, le Royaume-Uni...

Nous avons l'honneur de présider le Conseil de l'ECAS de 1987 à 1993 ; son secrétariat est assuré par Philippe Tassi. Les présidents qui nous succèdent sont Siegfried Heiler (Allemagne) de 1994 à 1997, Daniel Peña (Espagne) de 1998 à 2001, Bernard Fichet (France) de 2002 à 2005 (année au cours de laquelle de nouveaux statuts sont adoptés), Hans Nyquist (Suède) de 2006 à 2009, Ricardo Cao (Espagne) de 2009 à 2013, Anne Ruiz-Gazen (France) de 2013 à 2015 et Jean-Michel Poggi (France) depuis 2016. Notons que depuis 2014, les membres de la *Fédération des Sociétés nationales européennes de Statistique* (FENStatS) peuvent participer au projet ECAS (<http://ecas.fenstats.eu/>). En 2017, vingt-deux pays sont concernés par cette possibilité. Fort de cet élan, la présidence de Jean-Michel Poggi a impulsé une nouvelle dynamique en augmentant le rythme et en ouvrant le spectre des écoles ECAS,

⁷ Désireux de représenter la Belgique via la *Société Belge de Statistique*, alors en léthargie, nous suscitons une réunion, le 21 mai 1990, qui rassemble Jan Beirlant (Katholiek Universiteit Leuven), Marc Hallin (Université libre de Bruxelles), Henri Pastijn (SOGESCI), Pierre Dagnelie (ancien responsable de la Société Belge de Statistique) et nous. Elle est à l'origine du renouveau de la société savante belge qui rejoint l'ECAS peu après.

La formation à la statistique des enseignants-chercheurs

définissant en plus de la voie traditionnelle, à vocation universitaire, une voie orientée vers la statistique publique et une autre vers la statistique industrielle. Ainsi en 2017, trois écoles ont eu lieu : la première à Fréjus, avec pour thème *High Dimensional Statistics*, (<http://ecas2017.sfds.asso.fr/>), organisée conjointement avec la SFdS, et deux autres sur la statistique industrielle, organisées conjointement avec, d'une part, l'*European Network for Business and Industrial Statistics* – ENBIS (<http://www.unior.it/ateneo/15340/1/ecas-enbis-course-2017.html>), et d'autre part avec l'IRSDI – *Initiative de recherche en sciences des données pour l'industrie* (<https://eric.univ-lyon2.fr/iec/>)⁸. L'ECAS a ainsi récemment accru sa visibilité pour devenir ainsi un véritable acteur au service des sociétés nationales de statistique pour internationaliser leurs activités de formation doctorale et professionnelle.

L'ECAS est une formation à la statistique directement issue des *Journées d'Étude en Statistique*. Il n'est pas étonnant d'y avoir retrouvé les principes fondamentaux que nous avons décrits ci-dessus, même si les deux structures de formation ont suivi des voies distinctes ces derniers temps. Mais contrairement à ce qui s'est passé pour les JES, il n'a pas été possible de publier des ouvrages spécifiques résultant des sessions de formation.

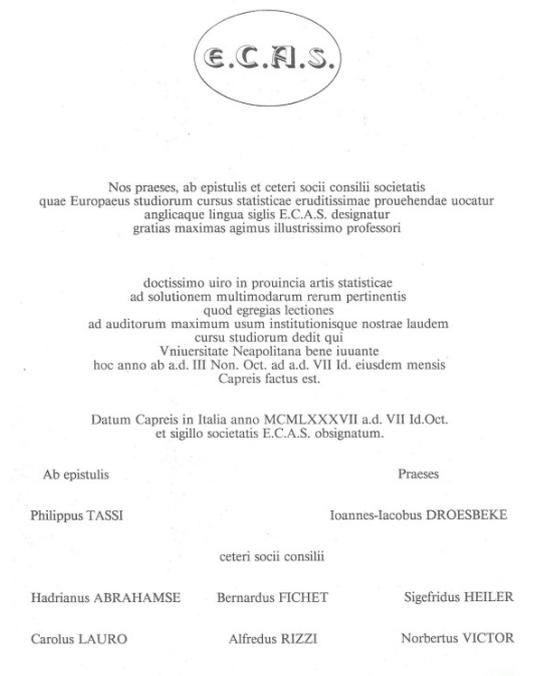


FIGURE 4 – *Diplôme de la première session de l'ECAS*

D'autres formations internationales consacrées à la statistique sont nées à la fin du vingtième siècle. Il faut en épingler deux d'entre elles qui sont aussi imprégnées par certains aspects des principes situés au cœur des JES.

⁸ En 2018, deux écoles sont d'ores et déjà prévues : en février à l'ENSAI de Rennes en statistique publique (avec EMOS et la SFdS, <http://ecas.ensai.fr/>) et en mai 2018 à Catania (Italie) sur *Clustering and Visualization of Complex Data* (avec l'IASC, www.clucla-summerschool.org).

5 Le programme PRESTA

Au moment où les premières JES voient le jour, les pays d'Amérique du Sud connaissent une mutation importante de leurs régimes politiques. Dans un contexte de récession économique sans précédent, aggravée par le poids de la dette extérieure héritée des régimes militaires qui étaient au pouvoir précédemment, les nouvelles institutions démocratiques attribuent un rôle important mais difficile aux universités et centres de recherche, en particulier dans le domaine des sciences humaines. À la fin des années quatre-vingts, l'équipement des centres universitaires de recherche s'améliore mais la destruction du potentiel scientifique que ces pays ont connue nécessite une aide extérieure pour se réapproprier une maîtrise suffisante en pratique de recherche et de gestion de données socio-économiques. La Commission Européenne décide de contribuer à cet objectif en soutenant financièrement divers projets. Parmi ceux-ci se trouve une formation à la statistique destinée aux enseignants-chercheurs et aux responsables des centres de recherche des instituts nationaux de statistique. Elle est intitulée *Programme de Recherche et d'Enseignement en Statistique Appliquée* (PRESTA).

La gestion de ce programme est confiée au *Laboratoire de Méthodologie du Traitement des Données* que nous dirigeons, à l'Université libre de Bruxelles. Après une phase d'identification destinée à évaluer les besoins en formation des institutions sud-américaines et à trouver les lieux les plus propices pour toucher le plus grand nombre de centres de recherche possible (voir Crivisqui et Droesbeke, 1991), un premier programme est mis en place à la fin des années quatre-vingts. Il est suivi d'un second programme, plus ambitieux, qui se déroule de 1994 à 1999.

Il n'est pas question de détailler ici les différents aspects des activités de formation qui se sont déroulées dans le cadre de ces programmes, et tout particulièrement le second (voir Crivisqui et Droesbeke, 1999). Notons cependant que les objectifs du programme PRESTA sont plus larges que ceux des *Journées d'Étude en Statistique*. Il y a bien sûr des cours de méthodologie statistique, comme dans ces dernières, mais on y trouve aussi des séminaires portant sur la méthodologie de la recherche quantitative, l'encadrement informatique des chercheurs, l'élaboration de documents de travail pédagogiques, la structuration d'équipes de recherche... Les participants sont invités à proposer leurs données personnelles lors des ateliers de traitement organisés dans le cadre du programme. Le fait que le responsable scientifique du programme PRESTA soit un des organisateurs des JES et que des enseignants de l'équipe de formation de ce programme aient été également enseignants ou organisateurs des Journées à Marseille explique facilement pourquoi les *Journées d'Étude en Statistique* ont eu une influence certaine sur ce programme européen.

Comme pour les JES, des manuels de cours sont publiés dans le programme PRESTA. Un autre support « papier » est engendré par les séminaires organisés sur le terrain non plus sous la forme d'ouvrages, mais d'une revue scientifique dénommée *Methodologica*, éditée jusqu'à la fin du programme, sous la responsabilité éditoriale d'Eduardo Crivisqui et d'un comité de rédaction composé de douze membres : Enrica Amaturò, Ramon Ardanuy Albajar, Monique Bécue, Atilio Boren, Pierre Bourdieu, Fernando Calderón, Philippe Cibois, Fernando Cortès, Frank Critchley, Yves Escoufier, Jeanne Fine, Eugène Horber, Jorge Raúl Jorrat, Dominique Ladiray, Claude Langrand, Carlo Natale Lauro, Jorge Padua, Hebe Roiter, Rose-Mary Salazar, Juan Ravier Sánchez Carrión, Horracio Torres, Manuel Vilares, Hugo Zemelman et nous.

La formation à la statistique des enseignants-chercheurs

Nous avons évoqué l'impact des JES. Pour se rendre compte de celui du programme PRESTA, il faut savoir que ses actions de formation ont été menées au sein de quatre zones géographiques (voir figure 5). La première, dénommée *Zona Central-Este*, regroupe le Brésil, le Paraguay, l'Uruguay et la partie nord-est de l'Argentine. La deuxième, appelée *Zona Austral*, couvre l'Argentine (sauf le nord-est) et le Chili. La troisième, dénommée *Zona Central-Oeste*, regroupe la Bolivie et le sud du Pérou. Enfin la dernière, appelée *Zona Norte*, couvre le nord du Pérou, l'Équateur, la Colombie et le Venezuela. Cette répartition était justifiée par des considérations géographiques – il fallait limiter les dépenses liées aux déplacements et faciliter les interactions ultérieures – et le fait que la demande exprimée par les communautés scientifiques ont été fort différentes d'une région à l'autre.



FIGURE 5 – Répartition des activités PRESTA en quatre zones

Le nombre de séminaires de formation qui ont été organisés n'est donc pas identique sur l'ensemble de ces zones, tout comme le nombre de participants et la répartition de leur origine géographique. Nous étions évidemment dépendants des facilités de disponibilité accordées par les institutions gouvernementales et académiques, des moyens financiers accordés localement pour soutenir le projet en complément de ceux offerts par la Communauté Européenne et du dynamisme développé par les différents centres de recherche.

TABLEAU 1 – Participation par zone aux activités de PRESTA

Zones	Nombres de séminaires	Nombres de participants
Central-Este	35	1300
Austral	24	700
Central-Oeste	13	400
Norte	34	740
Total	106	3 140

Le tableau 1, repris de l'article de Crivisqui et Droesbeke (1999), présente la participation aux cent six séminaires organisés entre 1994 et 1999, auxquels ont participé plus de trois

J.-J. Droesbeke

mille enseignants-chercheurs provenant de trois cent soixante-douze universités ou centres de recherche d'Amérique du Sud⁹.

6 Le programme TES

Signalons pour terminer la relation indirecte qu'ont eue les *Journées d'Étude en Statistique*, mais aussi l'ECAS et PRESTA, sur un autre programme européen intitulé *Training of European Statisticians* (TES). Celui-ci est mis en place par Eurostat, en 1990, pour former aux méthodes statistiques récentes les membres des instituts nationaux de statistique des pays de l'Union Européenne – appelée à l'époque Communauté Européenne – ainsi que ceux d'autres pays voisins. Ce programme débute en 1990, après consultation des pays membres. Il repose sur les travaux d'un comité de cinq experts présidé par Alberto de Michelis (Eurostat¹⁰), en charge de ce projet. Deux d'entre eux sont aussi membres du premier comité d'organisation des JES : Philippe Tassi et nous.

La demande en séminaires de formation divers s'accroît d'année en année. Pour y répondre efficacement, les offices statistiques des pays membres de l'Union Européenne et l'*European Free Trade Association* (EFTA) créent, six ans plus tard, le *TES Institute* pour gérer ce qui est devenu un impressionnant programme de formation à la statistique (voir Teekens, 2003). Cet institut, présidé par Rudolf Teekens, organise annuellement ses activités dans les pays de l'Union Européenne, mais il s'ouvre aussi à ceux de l'Europe centrale (non encore membres de l'Union Européenne), à la Russie et aux pays d'Afrique du Nord. Pour gérer ce programme de formations, il s'appuie sur les travaux d'un comité scientifique composé d'une petite dizaine de membres que nous présidons pendant quelques années, jusqu'à ce que le mode de fonctionnement de ce programme soit modifié par la Commission Européenne.

Pour des raisons analogues à ce qui s'est passé pour le programme PRESTA, les caractéristiques fondamentales des *Journées d'Étude en Statistique* – une équipe cohérente de formateurs, des documents de travail rédigés au préalable par chacun d'eux, une répartition horaire de travail de même type... – se sont trouvées utilisées dans ce projet européen dont l'impact s'est aussi révélé très important.

7 Conclusion

Le succès des *Journées d'Étude en Statistique* justifie amplement leur création. Leur extension au niveau européen et leur influence sur d'autres programmes internationaux permettent de penser que la politique de formation mise en place par l'ASU et reprise par la SFdS était pertinente.

⁹ Nous tenons à souligner l'importance du travail réalisé par Eduardo Crivisqui qui a coordonné les activités du programme PRESTA.

¹⁰ Eurostat est l'Office de statistique de l'Union européenne. Son siège se trouve à Luxembourg.

Références

- [1] Coutrot, B. et J.-J. Dreesbeke (1984), *Les méthodes de prévision*, Collection *Que sais-je ?*, **2157**, Presses Universitaires de France, Paris.
- [2] Crivisqui, E. et J.-J. Dreesbeke (1991), De la méthodologie du traitement des données socio-économiques en Amérique latine, *Bulletin des Séances de l'Académie Royale des Sciences d'Outre-Mer*, **37**, 773-794.
- [3] Crivisqui, E. et J.-J. Dreesbeke (1999), An Experience of International Cooperation: the PRESTA Programme, dans *Academic and Official Statistics Cooperation*, Eurostat, Luxembourg, 135-141.
- [4] Crivisqui, E. et J.-J. Dreesbeke (1999), Bilan d'une coopération quinquennale entre l'Europe et l'Amérique latine dans le domaine de la statistique appliquée, *Bulletin des Séances de l'Académie Royale des Sciences d'Outre-Mer*, **45**, 403-417.
- [5] Dreesbeke, J.-J. (1990), Postgraduate Programmes in Europe, Invited lecture, *Proceedings of the Finish Statistical Society*, 10 pages.
- [6] Dreesbeke, J.-J. (2005), Les racines de la Société Française de Statistique, *Journal de la Société Française de Statistique*, **146**, 5-22.
- [7] Teekens, R. (2003), Capacity Building for Social Statistics: the Role of professional Training, New-York, *United Nations*, Statistics Division, ESA/STAT/AC.88/15, 11 pages.

Annexe : Contributions des conférenciers des JES

1984 : Séries chronologiques. Théorie et pratique des modèles ARIMA

1. La place des séries chronologiques en statistique (J.-J. Dreesbeke, B. Fichet et Ph. Tassi)
2. Généralités sur les processus du second ordre (Chr. Gouriéroux)
3. Processus univariés AR, MA, ARMA. Notions de base et extensions (A. Monfort)
4. Estimation des paramètres de modèles ARMA (G. Mélard)
5. Identification et vérification (Chr. Gouriéroux)
6. Processus multidimensionnels (Chr. Gouriéroux)
7. Modèles non stationnaires, séries univariées et multivariées (M. Hallin)
8. Quatre cas pratiques (B. Coutrot, G. Mélard et Ph. Tassi)

1986 : Les sondages

1. La place des sondages en statistique (J.-J. Dreesbeke, B. Fichet et Ph. Tassi)
2. Généralités sur la théorie des sondages (Chr. Gouriéroux)
3. Sondages sans biais (Chr. Gouriéroux)
4. Modèles de surpopulation (A.-M. Dussaix)
5. Les sondages complexes (J.-Cl. Deville)
6. Effets d'un sondage : cas du χ^2 et de la régression (Chr. Gouriéroux)
7. Répliques d'échantillons : demi-échantillon, Jackknife et Bootstrap (J.-Cl. Deville)
8. Les réponses manquantes (J.-M. Grosbras)
9. Enquêtes répétées dans le temps (A.-M. Dussaix)
10. Algorithmes de tirage (J.-Cl. Deville et J.-M. Grosbras)

J.-J. Droesbeke

11. Echantillonnage séquentiel (J.-M. Grosbras)
12. L'application des sondages : trois exemples (J. Aglietta-Grivois, B. Riandey et A. Spriet)

1988 : Analyse statistique des durées de vie. Modélisation des données censurées

1. En guise d'introduction (J.-J. Droesbeke, B. Fichet et Ph. Tassi)
2. Les outils probabilistes (M. Carbon)
3. Les différents types de censure (C. Huber)
4. Estimation fonctionnelle dans les modèles de durée (C. Huber et J.-P. Lecoutre)
5. Test non paramétrique de comparaison d'échantillons (C. Huber)
6. Estimation et test dans les modèles paramétriques de durée (Chr. Gouriéroux, J.-P. Lecoutre et Ph. Tassi)
7. Hétérogénéité (Chr. Gouriéroux)
8. L'approche semi-paramétrique (Chr. Gouriéroux, J.-P. Lecoutre et Ph. Tassi)
9. Un modèle de recherche optimale d'emploi (Chr. Gouriéroux)

1990 : Modèles pour l'analyse des données multidimensionnelles

1. Introduction (J.-J. Droesbeke, B. Fichet et Ph. Tassi)
2. Modèles fonctionnels et structurels (J. Fine)
3. Modèles fonctionnels : quelques développements et applications (H. Caussinus)
4. Etudes asymptotiques (A. Pousse)
5. Extensions des analyses factorielles (Ph. Besse et A. Pousse)
6. Modèles probabilistes en classification (G. Celeux)
7. Résultats asymptotiques et validation en classification (G. Celeux)
8. Projections révélatrices (H. Caussinus)
9. Modèles graphiques d'association (J. Fine)

1992 : Modélisation ARCH. Théorie statistique et applications dans le domaine de la finance

1. Les modèles ARCH et leur histoire (J.-J. Droesbeke, B. Fichet et Ph. Tassi)
2. Modèles linéaires et non linéaires (G. Mélard)
3. Les modèles hétéroscédastiques univariés (Chr. Gouriéroux)
4. Estimation, prévision et tests : modèles univariés (Chr. Gouriéroux)
5. Généralisations et alternatives de la modélisation ARCH (A. Monfort)
6. Inférence fondée sur des simulations (A. Monfort)
7. Modèles à facteurs en finance (S. Pastorello et E. Renault)
8. Les processus ARCH comme approximations de processus en temps continu (L. Elie)
9. De l'utilité des modèles ARCH (J.-Fr. Boulier)

1994 : Inférence non paramétrique. Les statistiques de rang

1. Une approche historique des statistiques de rang (J.-J. Droesbeke et J. Fine)
2. Tests de rangs : définitions et exemples simples (B. Van Cutsem)
3. Tests sans biais, tests de permutation, tests invariants, tests de rangs (M. Hallin)
4. Éléments de la théorie asymptotique des expériences statistiques (M. Hallin)
5. Statistiques de rangs linéaires : normalité asymptotique et théorèmes de projection de Hajek (Ph. Barbe et M. Hallin)
6. Tests de rangs et tests de rangs signés pour le modèle linéaire général et les modèles autorégressifs (M. Hallin)

7. R-estimateurs dans le modèle linéaire général (B. Van Cutsem)
8. Méthodes non paramétriques : applications en marketing (M. Lejeune)

1996 : Plans d'expérience. Applications à l'entreprise

1. Le cheminement historique des plans d'expérience (J.-J. Dreesbeke, J. Fine et G. Saporta)
2. La planification des expériences et l'analyse de la variance : une introduction (P. Dagnelie)
3. Les plans factoriels (A. Kobilinsky)
4. Approche méthodologique des surfaces de réponse (D. Mathieu et R. Phan-Tan-Luu)
5. Approche méthodologique des mélanges (D. Mathieu et R. Phan-Tan-Luu)
6. L'approche Taguchi (D. Mathieu et R. Phan-Tan-Luu)
7. Plans d'expérience optimaux pour modèles linéaires (J.-P. Gauchi)
8. Plans d'expérience optimaux pour modèles de régression non linéaire (J.-P. Gauchi)
9. Trois applications à l'entreprise (J.-P. Gauchi, G. Saporta, D. Mathieu et R. Phan-Tan-Luu)

1998 : Méthodes bayésiennes en statistique

1. Thomas Bayes et son héritage (J.-J. Dreesbeke, J. Fine et G. Saporta)
2. Le paradigme bayésien (L. Simar)
3. L'inférence bayésienne : principes généraux (M. Mouchart)
4. Les modèles de base de l'analyse bayésienne (L. Simar)
5. La spécification de la distribution *a priori* (J.-P. Florens)
6. Méthodes de calcul en analyse bayésienne (Chr. P. Robert)
7. Bases décisionnelles de l'analyse bayésienne (Chr. P. Robert)
8. Tests d'hypothèses et choix de modèles (M. Mouchart)
9. Propriétés asymptotiques des estimateurs bayésiens (J.-P. Florens)
10. Estimation bayésienne de modèles de mélanges (Chr. P. Robert)
11. Quelques modèles de séries temporelles (Chr. P. Robert)
12. Inférence bayésienne non paramétrique et *bootstrap* (J.-P. Florens)
13. Modèles de durée (J.-P. Florens)
14. Estimation bayésienne d'un modèle multi-états markovien (J. A. Dupuis)

2000 : Modèles statistiques pour données qualitatives

1. Quelques aspects historiques du traitement des données qualitatives (J.-J. Dreesbeke, M. Lejeune et G. Saporta)
2. Les coefficients d'association et les tests d'indépendance pour des variables qualitatives (Chr. Croux)
3. Le modèle log-linéaire (Chr. Croux)
4. Les classes latentes (G. Saporta)
5. Modèles linéaires généralisés (P.-L. Gonzalez)
6. Modèles à réponse dichotomique (P.-L. Gonzalez)
7. Comparaison de méthodes de discrimination (P.-L. Gonzalez)
8. Régression logistique robuste (Chr. Croux et G. Haesbroeck)
9. Modèles de comptage et applications en assurance (Chr. Gouriéroux)
10. Mesures répétées (Chr. Gouriéroux)
11. Le modèle à réponse multinomiale. Application à la lutte contre le goitre endémique (M. Tenenhaus, Y. Le Roux, C. Guimart, P.-L. Gonzalez et D. J. M. Malvy)
12. La régression logistique PLS (M. Tenenhaus)

J.-J. Droesbeke

2002 : Analyse statistique des données spatiales

1. La corrélation et ses dérivés : le rôle de Galton dans leur histoire (J.-J. Droesbeke, M. Lejeune et G. Saporta)
2. Dépendance spatiale et autocorrélation (G. d'Aubigny)
3. Les modèles d'autocorrélation spatiale (G. d'Aubigny)
4. Modèles autorégressifs : tests de spécification (G. d'Aubigny)
5. Modèles autorégressifs et problèmes d'estimation (G. d'Aubigny)
6. Modèles du second ordre (X. Guyon)
7. Champ de Gibbs-Markov sur un réseau (X. Guyon)
8. Processus ponctuels spatiaux (C. Gaetan et X. Guyon)
9. Simulation des modèles spatiaux (C. Gaetan et X. Guyon)
10. Estimation des modèles spatiaux (C. Gaetan et X. Guyon)
11. Données régionales et agrégation spatiale (G. d'Aubigny)
12. Le cadre général de la géostatistique (J. Rivoirard)
13. La géostatistique linéaire, prédiction linéaire par krigeage (J. Rivoirard)
14. Krigeage disjonctif (Chr. Lajaunie)
15. Prédiction non linéaire par simulation conditionnelle (Chr. Lantuéjoul)
16. Discontinuités statistiques et discontinuités spatiales : l'exemple des inégalités de richesse par habitant en Italie 1951-1991 (Cl. Grasland)
17. Validation d'un modèle géostatistique pour l'interpolation : application à un événement pluvieux (D. Allard)
18. Poisson non stationnaire et échantillonnage semi-raréfié : comment les cèdres du Lubéron envahissent-ils l'espace ? (J. Chadœuf, M. Goulard, F. Lefèvre, P. Monestiez et C. Pichot)
19. Quelles sont les échelles spatiales importantes dans un écosystème ? (P. Legendre et D. Borcard)

2004 : Analyse statistique des données longitudinales

1. Les modèles et leur histoire (J.-J. Droesbeke et G. Saporta)
2. Modèles marginaux et modèles mixtes (G. Verbeke et G. Molenberghs)
3. Données manquantes (G. Molenberghs, C. Beunckens, I. Jansen, H. Thijs, G. Verbeke et M. G. Kenward)
4. Analyse longitudinale exploratoire (P. M. Kroonenberg)
5. Modèles dynamiques en économétrie des panels (A. Trognon)
6. Modèles non linéaires en économétrie des panels (A. Trognon)
7. L'approche biographique des trajectoires individuelles (E. Lelièvre)

2006 : Approches non paramétriques en régression

1. Les premiers pas de la régression (J.-J. Droesbeke et G. Saporta)
2. Les estimateurs à noyaux (M. Delecroix)
3. Fonctions orthogonales (M. Carbon)
4. Noyaux auto-reproduisants à base d'ondelettes (A. Antoniadis)
5. Fonctions splines (Chr. Thomas-Agnan)
6. Le fléau de la dimension et ses parades (M. Delecroix)
7. Les modèles de régression à directions révélatrices (M. Delecroix)
8. Données censurées (I. Van Keilegom)
9. Prédiction non paramétrique (M. Carbon)
10. Données spatiales (L. Cucala et Chr. Thomas-Agnan)

11. Données fonctionnelles (D. Bosq)
12. Quantiles de régression : applications à la construction de courbes de référence (J. Saracco)
13. Modèles à direction révélatrice unique : application en économie (M. Simioni)
14. La modélisation des courbes de croissance (M. Lejeune)

2008 : Modèles à variables latentes et modèles de mélange

1. De Poisson à Lazarsfeld, en passant chez Quetelet, Bertillon, Galton et Pearson (J.-J. Dreesbeke et G. Saporta)
2. Modèles de mélange et estimation (B. Garel)
3. Modèles de mélange : le nombre de composants (B. Garel)
4. Modèles de mélange et classification (G. Govaert)
5. Modèles à variables latentes : introduction et cadre général d'étude (J. K. Vermunt)
6. Analyse factorielle, réponse de type item et modèles à classes latentes (J. K. Vermunt)
7. Modèles à effets aléatoires (J. K. Vermunt)
8. Modèles d'équations structurelles, approches basées sur les composantes (V. E. Vinzi et L. Trinchera)
9. Modèles locaux issus d'un modèle d'équations structurelles (V. E. Vinzi et L. Trinchera)
10. Modèles dynamiques à variables latentes (A. Monfort)
11. Modèles espace-état linéaires et discrets (A. Monfort)
12. Inférence indirecte (A. Monfort)
13. Modèles de mélange en capture-recapture (L. Rouan, S. Cubaynes, Chr. Duchamp, Chr. Miquel, A.-M. Reboulet, O. Gimenez, J.-D. Lebreton, R. Choquet et R. Pradel)
14. Variables latentes et analyse de la satisfaction (E. Jakowicz)

2010 : Analyse statistique du risque

1. Risque et probabilité : les premières rencontres (J.-J. Dreesbeke et G. Saporta)
2. Théorie de la décision en univers risqué et incertain : une introduction (J.-M. Talon)
3. Mesures de risque (A. Charpentier)
4. Modèles univariés de valeurs extrêmes (A.-L. Fougères)
5. Théorie de la ruine (P. Bertail et St. Loisel)
6. Copules et risques multiples (A. Charpentier)
7. Valeurs extrêmes multivariées (A.-L. Fougères)
8. Théorie de la ruine multivariée (R. Biard et St. Loisel)
9. Solvabilité (St. Loisel)
10. Risque et assurance (A. Charpentier)
11. Considérations décisionnelles pour la construction d'un ouvrage de protection contre les crues (E. Parent, A. Pasanisi, N. Bousquet, M. Keller et J. Bernier)
12. Méthodes d'évaluation des risques alimentaires (P. Bertail, St. Clémenton et J. Tressou)

2012 : Méthodes robustes en statistique

1. La médiane et ses petits frères : une croissance difficile malgré leur robustesse (J.-J. Dreesbeke, G. Saporta et Chr. Thomas-Agnan)
2. Mesures de robustesse (Chr. Croux, C. Dehon et A. Ruiz-Gazen)
3. Analyse multivariée (Chr. Croux et G. Haesbroeck)

J.-J. Droesbeke

4. Analyse en composantes principales robuste. Application à la problématique des classements d'universités (C. Dehon)
5. Classification robuste (Chr. Ruwet et G. Haesbroeck)
6. Régression linéaire robuste (E. Cantoni et C. Dehon)
7. Le modèle linéaire généralisé robuste (W. Aeberhard et E. Cantoni)
8. Régularisation robuste de matrices de variances-covariances (G. Haesbroeck)
9. Mesures d'influence et robustesse en sondages (J.-Fr. Beaumont, D. Haziza et A. Ruiz-Gazen)
10. Détection graphique de valeurs atypiques pour données dépendantes (A. Ruiz-Gazen)

2014 : Model Choice and Model aggregation

1. A model selection tale (J.-J. Droesbeke, G. Saporta et Chr. Thomas-Agnan)
2. Introduction (P. Massart)
3. Non linear Gaussian model selection (P. Massart)
4. Bayesian model choice (J.-M. Marin et Chr. P. Robert)
5. Some computational aspects of Bayesian model choice (J.-M. Marin et Chr. P. Robert)
6. Randomization and aggregation for predictive modeling with classification data (N. Vayatis)
7. Mixture models (Chr. Biernacki)
8. Calibrations of penalties (P. Massart)
9. High dimensional clustering (Chr. Biernacki et C. Maugis-Rabusseau)
10. Clustering of co-expressed genes (M.-L. Martin-Magnette, C. Maugis-Rabusseau et A. Rau)
11. Forecasting the French national electricity consumption: from sparse models to aggregated forecasts (M. Mougeot)

2016 : Apprentissage statistique et données massives

1. Historique du sujet (M. Maumy-Bertrand, Chr. Thomas-Agnan et G. Saporta)
2. Fondamentaux de l'apprentissage statistique (S. Arlot)
3. Principaux modèles, erreur de prévision et risque (Ph. Besse)
4. Séparateurs à vaste marge (St. Canu)
5. Perceptron, apprentissage connexionniste (N. Villa-Vialanex)
6. Analyse des réseaux sociaux (E. Viennet)
7. Arbres, bagging boosting (J.-M. Poggi)
8. Forêts aléatoires, mesures d'importance des prédicteurs (J.-M. Poggi)
9. Validation croisée (S. Arlot)
10. Machines à noyau (St. Canu)
11. Apprentissage pour données massives (N. Villa-Vialanex)
12. Systèmes de recommandation, filtrage collaboratif (E. Viennet)
13. Passage à l'échelle « volume » des méthodes d'apprentissage (Ph. Besse)