



Association
des statisticiennes
et statisticiens
du Québec

Convergence

Volume XV • Numéro 1

Avril 2010

Chers membres,

Le printemps est dorénavant des nôtres, et il amènera avec lui un grand événement à ne pas manquer à Québec : le 38^e Congrès annuel de la Société statistique du Canada, qui aura lieu du 23 au 26 mai à l'Université Laval. Vous obtiendrez plus de détails sur l'événement à l'adresse http://www.ssc.ca/2010/index_f.html.

Cette rencontre scientifique sera une occasion unique d'avoir accès, ici-même au Québec, aux travaux et à l'expertise de plusieurs diplômés et professionnels de la statistique œuvrant partout en Amérique du Nord. Sans compter les nombreux représentants des autres continents. En parcourant le programme du Congrès, vous découvrirez que les sujets discutés permettront de rejoindre un vaste auditoire de statisticiens. Il est certain que tous y trouveront leur compte, peu importe leur champ d'expertise. Vous y apprendrez certainement des applications nouvelles et vous aurez, en prime, une occasion unique d'étoffer votre réseau d'experts de qualité parmi les centaines de statisticiens canadiens et étrangers qui seront sur place.

Je vous invite particulièrement à ne pas manquer la session « 05D » du lundi 24 mai à 16h30. Cette session sera parrainée par l'ASSQ et aura pour titre « Quel est le rôle d'une association de statisticiens pour assurer la rigueur dans la publication médiatique de résultats statistiques? » Ce panel de discussion permettra à trois invités de débattre la question. Deux invités sont confirmés jusqu'ici, à savoir M. Michel Saulnier, président du chapitre du Québec de l'Association de la recherche et de l'intelligence marketing, et Mme Lise Millette, de la Presse canadienne et membre du conseil d'administration de la Fédération professionnelle des journalistes du Québec. Un gros merci à notre vice-présidente, Nathalie Madore, de la Régie des rentes du Québec, pour l'organisation de l'événement.

L'ASSQ profitera également de ce congrès pour y tenir sa traditionnelle assemblée générale annuelle. Vous recevrez

plus de détails à ce sujet dans les prochaines semaines. Vous pouvez néanmoins réserver la plage horaire du lundi 24 mai, entre 17h30 à 18h30, pour participer à cette activité vitale pour le bon fonctionnement de notre association. Oui, j'en conviens, ce lundi est un jour férié, mais une fois ne sera pas coutume, c'est que nous devons cette année suivre l'agenda du Congrès de la SSC. Vous pourrez ensuite demeurer des nôtres en début de soirée pour un souper de l'ASSQ en compagnie de tous les membres présents à l'assemblée. Prendre note qu'il sera possible de participer uniquement à notre assemblée générale annuelle sans devoir s'inscrire au Congrès en tant que tel.

Au plaisir de vous rencontrer en grand nombre le lundi 24 mai!

Le président de l'ASSQ,

[Martin Rioux](#), B.Sc., M.B.A.

Le mot du président	1
Le mot du rédacteur en chef	3
Stat et vie : Pauvres fumeurs, ou fumeurs pauvres?	3
Les échos du CA	4
Le mot du registraire	5
Compte-rendu de la 23^e Journée de la statistique	6
Chronique historique : le baron de Férussac	7
La chronique SAS : Proc DATASETS	10
Caricature : l'inflation!	13
À inscrire à votre agenda	14
Conférences à venir	15
Nouvelles publications	16
Stat et vie : productivité statistiquement significative	17
Suivre son cours	18

CONVERGENCE

Convergence, le journal de l'Association des statisticiennes et statisticiens du Québec (ASSQ), est publié trois fois l'an, soit en avril, juillet et décembre. Il est distribué gratuitement aux membres de l'ASSQ.

Rédacteur en chef

Jean-François Quessy, Université du Québec à Trois-Rivières (Jean-Francois.Quessy@uqtr.ca)

Rédacteurs adjoints

Steve Méthot, Agriculture et agroalimentaire Canada
Myrto Mondor, C. H. affilié universitaire de Québec

AVIS AUX AUTEURS

La rédaction de *Convergence* invite les statisticiens et toutes les personnes intéressées par la statistique et ses applications à lui faire parvenir leurs articles, questions, commentaires, soumissions et résolutions de problèmes. Les textes doivent être envoyés, sous forme de fichiers Microsoft Word, à l'adresse électronique du rédacteur en chef (voir ci-haut). Les dates de tombée sont approximativement les 1^{er} mars, 1^{er} juin et 1^{er} novembre pour les numéros d'avril, de juillet et de décembre, respectivement. La rédaction ne s'engage pas à publier tous les textes reçus et se réserve le droit de n'en publier que des extraits sur approbation de l'auteur.

AVIS AUX ANNONCEURS / EMPLOYEURS

Les entreprises ou les personnes qui désirent faire paraître de la publicité ou des offres d'emploi dans *Convergence* doivent faire parvenir par courriel au rédacteur en chef leur document électronique prêt pour l'impression avant la date de tombée du prochain numéro. Les membres institutionnels ont une gratuité de publicité allant jusqu'à une page par numéro. Les tarifs pour la parution dans un numéro de *Convergence* sont les suivants (à noter que l'annonce de cours et de séminaires en dernière page est gratuite) :

Tarifs

Carte d'affaires	15 \$
1/4 page	40 \$
1/2 page	80 \$
Page entière	150 \$

Note liminaire : la forme masculine est employée dans le but d'alléger le texte et désigne les deux sexes, à moins d'une mention contraire de l'auteur.

La rédaction de *Convergence* laisse aux auteurs l'entière responsabilité de leurs opinions. La reproduction des articles est autorisée, sous réserve de mention de la source.

Toute correspondance doit être adressée à :

Convergence

Association des statisticiennes et statisticiens du Québec
2136, chemin Sainte-Foy
Suite 200
Québec (Québec) Canada G1V 1R8
Courriel : assq@association-assq.qc.ca
Page Internet : <http://www.association-assq.qc.ca>
Dépôt légal - Bibliothèque nationale du Québec, 2010

MEMBRES INSTITUTIONNELS :



Département de mathématiques et de statistique



Mission

L'ASSQ vise à regrouper les statisticiennes et les statisticiens de tous les domaines afin de promouvoir la statistique et d'en favoriser la bonne utilisation.

Membres

L'ASSQ offre deux types d'adhésion aux personnes intéressées par ses activités, à savoir membres *statisticien* et *affilié*.

Membre statisticien : Toute personne possédant au moins un baccalauréat en statistique ou l'équivalent (baccalauréat avec au moins 24 crédits de cours reconnus en statistique ou en probabilité). Les personnes qui ne répondent pas à cette condition peuvent accéder à la catégorie de membre statisticien si leur expérience professionnelle est jugée équivalente aux connaissances acquises lors de la formation académique.

Membre affilié : Toute personne qui souhaite faire partie de l'ASSQ.

Membres institutionnel : Les organismes peuvent devenir membres de l'ASSQ et ainsi bénéficier de plusieurs privilèges, dont l'adhésion gratuite comme membres statisticiens ou affiliés pour trois de leurs employés

Les frais d'adhésion annuels sont de 50 \$ pour les membres réguliers, 10 \$ pour les étudiants et 300 \$ pour les membres institutionnels

Conseil d'administration

Président : [Martin Rioux](#) (Geysier Statistique)

Vice-présidente : [Nathalie Madore](#) (Régie des rentes du Québec)

Secrétaire : [Yann Latulippe](#) (Régie des rentes du Québec)

Trésorière : [Anne-Sophie Julien](#) (INSPQ)

Registraire : [Michel Fluet](#) (SOM Inc.)

Directeur des communications : [Jean-François Plante](#) (HEC Montréal)

Le mot du rédacteur en chef

C hers membres,

L'édition printanière de *Convergence* recèle encore une fois une panoplie d'informations sur les activités chapeautées par l'ASSQ. À commencer par les traditionnels «Échos du C.A.», signés de la plume de Yann Latulippe, notre secrétaire. Yann y dresse un bilan des discussions qui ont eu lieu pendant les trois réunions de l'hiver. Notamment, il fait un retour sur les deux « Jeudi de l'ASSQ » qui ont récemment eu lieu et mentionne le rôle que l'ASSQ entend jouer lors du prochain congrès annuel de la Société statistique du Canada. Il décrit également quelques projets futurs, notamment la refonte du site web.

Dans un bref billet, le registraire Michel Fluet dresse un portrait du nombre de membres réguliers et institutionnels. Ensuite, pour ceux qui n'ont pu assister à la 23^e édition de la Journée de la statistique, tenue le 19 mars dernier, un compte-rendu vous est offert.

Avec bonheur, j'en suis sûr, vous retrouverez deux chroniques régulières : la chronique historique de Pierre Lavallée et la chronique SAS signée par Sylvain Tremblay. D'une part, vous en apprendrez beaucoup sur le baron de Férusac (1786-1836), un adepte de la pensée rigoureuse propre à la statistique. D'autre part, vous verrez comment manipuler plus efficacement de grands jeux de données à l'aide de l'instruction Proc DATASETS de SAS.

Finalement, je vous invite à prendre connaissance des habituels « À inscrire à votre agenda », « Conférences à venir », « Nouvelles publications » et « Suivre son cours ». J'en profite aussi pour vous inciter à nourrir ces sections, par exemple en nous envoyant vos propres suggestions de lecture, ou en portant à notre attention des événements qui sont susceptibles d'intéresser notre lectorat.

Bonne lecture!

Jean-François Quessy, Ph. D.

Statistique et vie :

Pauvres fumeurs, ou fumeurs pauvres?

L'Organisation mondiale de la santé dénonce le cercle vicieux qui lie le tabac et la pauvreté. Pas moins de 84 % des fumeurs vivent dans des pays en développement ou des pays en transition. Le tabac y représente d'ailleurs un poids financier important : au Vietnam, les ménages les plus démunis dépensent presque quatre fois plus pour le tabac que pour l'éducation. Au Royaume-Uni, une étude a montré que seuls 10 % des femmes et 12 % des hommes fument dans la catégorie socio-économique la plus élevée, contre 35 % des femmes et 45 % des hommes dans la catégorie socio-économique la plus basse. Pauvres fumeurs, ou fumeurs pauvres?

[Tiré de l'Almanach Vermot 2006]

Les échos du CA

Rubrique préparée par les membres du Conseil d'administration de l'ASSQ afin de fournir de l'information continue à ses membres

Chers collègues,

Voici ma deuxième chronique « Les échos du C.A. » en tant que secrétaire de l'association. Je suis toujours heureux d'en faire partie. Les discussions sont stimulantes et les projets en 2010 s'annoncent passionnants.

Trois réunions se sont tenues depuis le mois de janvier. Plusieurs projets sont en cours et certaines activités organisées par l'association en 2010 méritent d'être soulignées. Il s'agit des jeudis de l'ASSQ, du colloque de la Société statistique du Canada en mai 2010 et de la refonte du site internet de l'association.

Concernant les jeudis de l'ASSQ, deux « 5 à 7 » ont eu lieu le 26 novembre et le 11 mars dernier. Une statisticienne de l'Institut de la statistique du Québec, Marie-Ève Tremblay, a discuté des défis méthodologiques liés à une enquête de santé réalisée l'an passé par l'Institut. Le même exposé a été présenté à l'Université Laval et à l'UQAM. Globalement, la participation a été excellente avec environ 25 participants à Québec et 10 participants à Montréal.

De plus, tel que mentionné dans le mot du président, le colloque de la Société statistique du Canada aura lieu du 23 au 26 mai prochain à l'Université Laval. L'association est, sans surprise, impliquée dans la préparation de cette activité. En plus de commanditer l'événement de façon monétaire, l'ASSQ est responsable d'organiser une table ronde d'une durée de 90 minutes avec trois panelistes. Un animateur dirigera les discussions autour du sujet suivant : « Quel est le rôle d'une association de statisticiens pour assurer la rigueur dans la publication médiatique de résultats statistiques? » Il sera intéressant de voir, au terme de

cette conférence, si l'association peut intervenir pour préserver la bonne utilisation des résultats de sondage. La table ronde aura lieu le 24 mai en fin d'après-midi. Elle sera suivie de l'assemblée générale annuelle des membres de notre association (AGA) et d'un souper dans un restaurant bien connu de Ste-Foy. C'est un rendez-vous à ne pas manquer!

Enfin, un autre projet d'envergure pour l'année 2010 est la refonte du site Internet de l'ASSQ. Plusieurs fonctionnalités sont à revoir ou à ajouter. Par, on envisage l'implantation d'un système de paiement en ligne. Les membres pourraient alors payer leur inscription à un guichet automatique, à un comptoir ou par téléphone en mentionnant le numéro de compte de l'association. De plus, la diffusion d'un calendrier d'activités est prévue. Pour ce faire, un recensement des activités intéressantes liées à la statistique sera effectué bientôt auprès des membres. Une section rassemblant les formulaires pertinents est envisagée, de même qu'une section « Manchettes ». Une section « Blogue » pourrait aussi voir le jour. L'objectif est de divulguer des informations pertinentes et de recueillir des commentaires des visiteurs sur le site. Ce sujet sera abordé lors de l'assemblée générale du 24 mai prochain.

Voilà les sujets chauds discutés récemment lors des réunions mensuelles du C.A. Les événements à venir seront très intéressants. Plusieurs d'entre vous ont d'ailleurs renouvelé leur adhésion à l'association. Il me fera donc plaisir de vous rencontrer!

Yann Latulippe
Secrétaire

Le mot du registraire

Compte rendu de la campagne d'adhésion et de renouvellement préparé par le registraire de l'ASSQ, Michel Fluet

Voici quelques nouvelles concernant les adhésions à l'Association des statisticiennes et statisticiens du Québec. En date du 24 mars 2010, l'association comptait cent sept membres actifs. De ce nombre, sept sont des membres institutionnels :

- SOM ;
- L'Institut de la statistique du Québec ;
- Techno 5 ;
- Le département de mathématiques et de statistique de l'Université Laval ;
- les Services conseils Hardy ;
- StatExpert ;
- l'Université du Québec à Montréal.

À noter également que Statistique Canada a annoncé le renouvellement de son adhésion. À titre comparatif, l'Association comptait cent huit membres actifs à pareille date l'an dernier. J'invite donc tous ceux et celles qui n'ont pas encore renouvelé à le faire dans les prochains jours afin de soutenir l'Association et de profiter de ses services.

Michel Fluet
Registraire



VOTRE GÉNIE AU SERVICE D'UN CÉGEP!

LES 48 CÉGEPs DU QUÉBEC EMPLOIENT PLUS DE 35 000 PERSONNES.

POUR PLUS D'INFORMATION ET POUR CONSULTER LES POSTES DISPONIBLES, VISITEZ NOTRE SITE WEB :

emploi cegep
qc.ca

Les cégeps sont actuellement à la recherche de :

ENSEIGNANT OU ENSEIGNANTE EN GÉNIE ÉLECTRIQUE

Nature du travail : Spécialiste de l'ingénierie et pédagogue capable de transmettre ses connaissances aux étudiants.

Qualifications requises : Diplôme universitaire de premier cycle en génie électrique, génie électronique ou génie des communications.

ENSEIGNANT OU ENSEIGNANTE EN MATHÉMATIQUES

Nature du travail : Spécialiste des mathématiques et pédagogue capable de transmettre ses connaissances aux étudiants.

Qualifications requises : Diplôme universitaire de premier cycle en mathématiques ou statistiques.

À titre de membre du personnel, vous contribuerez à créer un environnement éducatif propice aux études et à la réussite. Vous jouerez un rôle essentiel dans la mission éducative des cégeps.

JOIGNEZ-VOUS À UN EMPLOYEUR DE CHOIX QUI CONTRIBUE À BÂTIR L'AVENIR

LES CÉGEPs ONT BEAUCOUP À VOUS OFFRIR :

- Un milieu de travail dynamique, stimulant et enrichissant
- Divers avantages sociaux et services liés à la conciliation travail et vie personnelle
- Le remboursement de certaines activités de formation et de perfectionnement
- Des horaires de travail avantageux

Compte-rendu de la 23^e édition de la Journée de la statistique

Denis Talbot,
Responsable des relations publiques
Comité pour l'avancement de la statistique à l'Université Laval

Le vendredi 19 mars dernier avait lieu, à l'Hôtel Universel de Québec, la 23^e édition de la Journée de la statistique. Organisée par le Comité pour l'avancement de la statistique à l'Université Laval, cette rencontre a pour but principal de permettre aux personnes intéressées par la statistique d'en apprendre davantage sur les différentes possibilités d'emplois dans le domaine. L'activité a attiré plus de trente participants.

Cette année, les participants ont d'abord pu assister à une conférence donnée par Michel Gariépy, du Mouvement des caisses Desjardins. M. Gariépy a exposé sa vision du rôle d'un statisticien dans le domaine financier. Martin Perron, de la Régie des rentes du Québec, a ensuite discuté de l'utilisation de la statistique dans l'amélioration de la performance d'une organisation. Finalement, Aurélie Dauré et Gina Maucieri, de l'entreprise Techno 5, ont entretenu l'auditoire sur la façon dont un statisticien peut améliorer ses démarches de recherche d'emploi dans les domaines financier et pharmaceutique.



Sur la photo : les membres du comité exécutif du Comité pour l'avancement de la statistique à l'Université Laval, accompagnés par les conférenciers. De gauche à droite: Frédérick Beaupré, Sandra Larrivée, Gina Maucieri, Aurélie Dauré, Michel Gariépy, Jonathan Athot, Claudia Beaudoin, Andréanne Moisan-Roy, Denis Talbot et Anne-Sophie Julien. (absent lors de la photo : Martin Perron).

Chronique historique :

Le baron de Férussac (1786-1836)

André Étienne Juste Pascal Joseph François d'Audebard, baron de Férussac, est né le 30 décembre 1786 au Chartron près de Lauzerte, dans le Quercy blanc, rattaché en 1790 au département du Lot, puis en 1808 à celui du Tarn-et-Garonne. Nous ne connaissons pas son prénom d'usage; ses mémoires sont généralement signés baron de Férussac. Nous nous conformerons à la coutume qui le prénomme André. Son père, Jean Baptiste Louis Joseph, est militaire et naturaliste. Sa mère, Marie Catherine Joséphe de Rozet, appartient à la noblesse de Lauzerte. La famille Audebard est d'ancienne noblesse militaire. Son origine géographique est controversée : certains la situent en Limousin, d'autres en Lorraine. Sa date d'installation en pays agenais n'est pas connue, mais semble se situer entre le X^e et la fin du XV^e siècle, où l'on trouve ici et là, sous diverses orthographes, des Audevars, Audevards, Daudebard, etc. Rien n'indique qu'elle ait été particulièrement riche et puissante. Le titre de baron, dont André de Férussac se recommande volontiers, est lui-même sujet à caution et son origine demeure obscure. À une date inconnue, vraisemblablement à la fin du XVI^e siècle, les Daudebard ont pris le nom du château-moulin de Ferrussac, construit aux XII^e et XIV^e siècles, et qui existe toujours. Il se trouve sur la commune de Saint-Maurin à l'est d'Agen, dans le Lot-et-Garonne. André de Férussac meurt à Paris, le 21 janvier 1836, 25 rue de l'Université, où il réside.

La pensée statistique de Férussac s'exprime principalement dans deux textes, publiés en 1821 et 1829, qui encadrent d'une certaine façon son activité éditoriale. Le premier s'intitule *De la géographie et de la statistique*, ces deux domaines étant considérés dans leurs rapports avec les sciences qui les avoisinent de plus près. Suit le *Plan sommaire du cours de Géographie et Statistique*, qu'il enseigne à l'École d'état-major. Citons la première phrase de ce texte remarquable, pour donner une idée du style contourné de son auteur et sa hauteur de vue : « Il est dans la nature de nos facultés d'élever assez haut l'édifice des sciences, avant d'employer pour le soutenir et pour assurer l'harmonie de toutes ses parties, les règles d'une

saine critique, et cette méthode rigoureuse, qui peut seule assigner sa place véritable aux matériaux divers entassés par les efforts successifs des générations. » Férussac stigmatise ensuite les savants, poussés par « le besoin impérieux de rendre raison des choses », qui bâtissent hâtivement des systèmes destinés à s'écrouler à brève échéance, parce qu'ils n'ont pas su « attendre que les faits soient observés sous leur vrai jour ». Plutôt que des guides, ces systèmes sont des obstacles aux « progrès des connaissances ». Ils mêlent ensemble des faits, des observations, sous des intitulés trop larges et mal délimités, Physique, Histoire naturelle, Géographie, et maintenant Statistique. Toutes ces « sciences » parlent de la Nature, « l'Univers, les Êtres en général, le Globe et les sociétés humaines en particulier ». Qu'ont-elles de spécifiques? C'est ce que notre savant auteur entend clarifier.



Se recommandant de Say (pour mieux le réduire ensuite), Férussac introduit le concept de « champ » d'une science, ou, de façon plus géométrique, de cercle, ou sphère d'une science, expressions synonymes qui reviennent fréquem-

ment sous sa plume. Ces champs, cercles ou sphères ne sont pas placés au hasard dans l'espace du savoir, ils ont un « lieu », une place, d'où ils voient le monde, à leur manière propre, jusqu'aux frontières « naturelles » de leurs domaines. Cette géographie des sciences permet de classer les faits, de leur donner leur « vrai jour », de préciser leurs places, leurs enchaînements, leurs dépendances. Par exemple, nous explique le baron : « Au commencement du XVIII^e siècle, le défaut d'esprit critique et d'une saine méthode firent bannir de la médecine, l'application des procédés chimiques. Après 50 ans de proscription, la révolution opérée par Lavoisier, dans la chimie, lui fit recouvrer l'honneur d'être un des plus utiles appuis de la médecine ». La chimie médicale appartient au cercle de la médecine, la chimie des corps animés ou inanimés à celui de la chimie. Le système de Férussac se complique d'ailleurs singulièrement par ce fait que certaines sciences s'intéressent aux « rapports simples des choses », d'autres, différentes, aux « rapports complexes » de ces mêmes choses. Chacune a son importance et son point de vue qui doit être caractérisé et distingué. Le baron a une vision multidimensionnelle de la géographie des sciences, et ses cercles peuvent appartenir à des plans différents, parallèles ou non, comme dans un atlas à trois dimensions, voire à davantage de dimensions encore, au moins métaphoriquement. On doit même imaginer des atlas partiels dans le grand atlas général des sciences, toute une arborescence d'atlas, qu'il convient de décrire avec ordre et méthode. Il faut enfin prendre en considération la dimension temporelle. Certaines sciences sont immobiles, toutes les géographies par exemple, d'autres sont mobiles, les histoires en général, l'histoire politique ou la géologie en particulier. Il y a donc des cercles, orthogonaux aux autres, qui incluent la flèche du temps. Vision, ou rêve, ou utopie, qui ne sont pas sans grandeur, ni beauté.

Le baron remarque ensuite qu'un problème particulier se pose « de nos jours pour la statistique et l'économie politique ». Par absence de sens critique et de véritable méthode, ces deux sciences d'une importance capitale, sont « méconnues » : elles n'ont trouvé ni leur place, ni leur cercle, ni leur façon de voir. Nous n'insisterons pas sur l'économie politique dont le but est de déterminer les « lois d'existence » des sociétés en général, de devenir la « physique des sociétés » ; ce qui la situe, la munit de son point de vue, et la différencie d'autres sciences qui lui ressemblent superficiellement, la géographie politique par exemple, mais dont elle est naturellement différente, bien qu'elle emprunte certains matériaux. Nous nous arrêterons plutôt sur la façon dont notre baron voit le « cercle scientifique de la statistique ».

La statistique est une science potentiellement très ancienne, mais qui n'a été reconnue que fort récemment. Les Allemands amalgament sous ce nom depuis assez longtemps la géographie politique, l'économie politique et diverses autres sciences. À l'inverse, les Français n'y voient qu'un auxiliaire mineur de la géographie ou de l'économie. Ce faisant, les uns comme les autres broient « une couleur habilement composée de beaucoup d'autres... avec une de celles qui servirent à la créer : on n'aurait plus ni l'une ni l'autre ; détruites par l'amalgame, la combinaison qui en résulterait resterait inutile, comme ces couleurs fausses que le peintre rejette de sa palette ». Le baron de Férussac, ce « guerrier passionné par le beau », qui depuis sa plus tendre enfance identifie et classe les mollusques non marins, et qui s'efforce de « peindre » ce qu'il a vu, sait de quoi il parle. Quelle est la couleur de la statistique? En quel lieu se situe-t-elle? Qu'a-t-elle en vue? Modestement, le baron précise que tout a déjà été dit dans le « programme publié par l'Institut, en 1818, pour le concours au prix de cette science ». La réalité est un peu différente, mais il y a incontestablement des analogies entre le texte du Programme académique et la définition du cercle statistique (ou plutôt des cercles statistiques) du baron. D'abord l'affirmation nette du Programme, selon laquelle « la statistique est une science de plus », distincte des autres, que le baron reprend et détaille (contre Say). Ensuite le but de la statistique : « reconnaître et constater les effets généraux des institutions civiles et de tous les éléments de la puissance respective et de la richesse des nations », qui devient sous la plume du baron : « connaître la situation générale et respective des sociétés, pour juger comparativement la force, les ressources de tout genre, la richesse, la gloire, la prospérité, la puissance en un mot des nations entre elles ». Toutefois, il y a entre le Programme et l'Essai, certaines différences de tonalité (et de couleur).

Pour la commission académique, la Statistique reconnaît et constate la richesse des nations. La Statistique de Férussac le fait tout autant, mais elle se voit pourvue d'un rôle en quelque sorte stratégique, au moins praxéologique. Non seulement elle reconnaît les richesses, mais elle en « compare » les avantages et les inconvénients et elle « éclaire » l'homme dans sa vie privée, le savant dans ses recherches, l'officier sur le champ de bataille, comme la nation dans son gouvernement. Citons encore le baron : « un ouvrage qui présenterait, dans ses généralités, le tableau statistique des grandes régions naturelles du globe; qui ensuite offrirait la situation respective de chacune des sociétés politiques qui s'y trouvent comprises; où la partie narrative de chacune de ces situations serait accompagnée de tableaux synoptiques où tous les résultats numériques seraient présentés dans un ordre convenable, tableaux qui sont pour la statistique ce que les cartes sont

pour la géographie; un tel ouvrage, dis-je, serait aussi neuf qu'il serait important et utile. Des topographies statistiques conçues sur le même plan deviendraient le flambeau de l'homme de guerre et le salut des peuples et des armées. » La statistique ne se contente donc pas de dresser des tableaux numériques, elle doit les « présenter dans un ordre convenable », et les résumer en « tableaux synoptiques » de façon à faciliter les prises de décision sur le terrain comme au laboratoire. C'est au fond le programme de l'École de Condorcet-Laplace, mais revu à la manière d'un officier d'état-major de la Grande Armée, naturaliste quercynois, spécialiste des tableaux de classifications conchyliologiques, à l'usage des futurs officiers d'état-major de l'Armée française.

Le baron ignore évidemment le calcul des probabilités qui donne, à l'École dont il s'agit, des formules uniformes capables de supputer et de juger le plus avantageusement du monde, mais André de Férussac est certainement un pionnier de l'analyse topographique et picturale des données de toute nature, qui permet une véritable critique, dès que l'on a affaire à des données très multidimensionnelles, et les statisticiens du XX^e siècle, certains d'entre eux au moins, ne le renieraient pas tout à fait. Cournot, lui-même, qui, plus que tout autre, entend promouvoir la statistique contrôlée par le calcul des probabilités, sait si bien les limites de la statistique mathématique, qu'il a dû forger, pour les besoins de la critique des sciences, le concept de « probabilité philosophique », non numérique, à laquelle il n'est pas absolument déraisonnable d'associer l'ordre et la méthode des couleurs de Férussac.

Férussac aborde ensuite deux objections qu'on oppose communément à la statistique. Les tableaux statistiques sont illusoire et inutile parce qu'ils fluctuent d'un dénombrement à l'autre. De plus, ils sont variables dans le temps, il y a des années d'abondance et des années de disette. La seconde objection a, selon notre savant, été levée par le

Programme de l'Académie : les chiffres de la statistique sont soumis à un principe général selon lequel « les éléments variables conservent entre eux une relation qui subsiste toujours ou du moins pendant un laps de temps considérable », ce sont les fameux rapports constants de la statistique laplacienne, le taux de masculinité ou le rapport de la population aux naissances. La première objection relève, elle, d'un principe plus général encore. Si l'on fait abstraction des variations temporelles, les chiffres de la statistique fluctuent, en effet, d'un groupe à l'autre. Mais, en moyenne, dans les grandes lignes, ils sont fixes, du moins assez fixes pour permettre des comparaisons et des « jugements ». Il existe ainsi une autre sorte de rigueur, la rigueur statistique, qui a lieu dans l'ensemble, à un ordre de fluctuation près, que l'on peut négliger pour les buts qu'on s'assigne, si bien que cette rigueur approximative possède une valeur pratique et scientifique, qui n'est pas moindre que la rigueur minutieuse des sciences de la précision. Comme l'écrit le baron : « ne serait-ce pas une folie, par exemple, de vouloir absolument connaître la quantité de pain, à une livre près, qui se consomme annuellement à Paris, pour assurer l'approvisionnement de cette place? ».

Nous l'avons dit, l'Essai de Férussac est d'abord un discours de la méthode, pour la critique et la topographie des sciences. Programme ambitieux, hors de portée, certes. Quoi qu'il en soit, il est manifeste que le baron ressent, en 1821, le besoin, la « nécessité impérieuse », de réaliser une topographie statistique de la science en général, qui éclairerait la marche des savants, en leur montrant l'ordre et la généalogie des sciences et en leur permettant d'en suivre les progrès continuels et d'autant plus accélérés qu'ils sont plus largement connus.

[Extraits de Bru, B., Martin, Th. (2005), Le baron de Férussac, la couleur de la statistique et la topologie des sciences, *Journal Électronique d'Histoire des Probabilités et de la Statistique*, <http://www.emis.de/journals/JEHPS/index-3.html>.]

Citation

Les statistiques sont vraies quant à la maladie et fausses quant au malade; elles sont vraies quant aux populations et fausses quant à l'individu.

Léon Schwartzberg

La chronique SAS :

La gestion des données avec *Proc DATASETS*

[Sylvain Tremblay](#)
Institut SAS (Canada) Inc.

Lorsque l'on travaille avec des données dans SAS, il y a souvent plusieurs façons de faire la même chose. On opte parfois pour la technique qui nous apparaît la plus évidente, sans tenir compte de son efficacité. Lorsque le nombre d'observations est relativement peu élevé, la différence est peu perceptible. Par contre, quand la taille des fichiers de données SAS est grande, un choix judicieux s'impose afin d'éviter certains problèmes de performance.

Dans cette chronique printanière, j'aimerais faire un survol de la procédure DATASETS de SAS. Méconnue de plusieurs, cette procédure permet de faire une gestion efficace des fichiers de données dans SAS.

Une question de structure

Lors de l'exercice de mes fonctions, je vois souvent des gens utiliser une étape DATA pour modifier ou ajouter des attributs à une variable, tel un format ou une étiquette. Bien que cette approche soit fonctionnelle, elle n'est pas efficace. Pourquoi? Parce qu'un fichier de données SAS possède une structure particulière. En effet, ces fichiers ont deux composantes : (i) une zone descriptive et (ii) une zone de données.

Dans la zone descriptive, on retrouve des informations techniques sur la table et les méta-données (c'est-à-dire les renseignements sur les données). C'est là que sont stockés les attributs de la table, comme le nombre d'observations, le nombre de variables et la date de création de la table, et les attributs des variables, tels leur nom, leur type et leur longueur. Pour consulter la zone descriptive, on peut utiliser la procédure CONTENTS. Dans la zone de données, on retrouve les données elles-mêmes.

L'étape DATA : lecture et écriture des données

Supposons que l'on veuille ajouter un format à une variable d'un fichier de données SAS qui existe déjà. Pour le faire de façon permanente, on pourrait choisir d'utiliser une étape DATA. L'encadré 1 illustre une telle approche.

Encadré 1

```
3 data lib.dsn;  
4     set lib.dsn;  
5     format var format. ;  
6 run;
```

Que se passe-t-il lors de l'exécution de cette étape DATA? SAS va ajouter ce format à la variable, mais va aussi, et de façon itérative, lire et écrire toutes les observations. Cet *input/output* constitue une pure perte de temps!

Il existe une approche beaucoup plus efficace qui utilise la procédure DATASETS de SAS. Celle-ci a la particularité de ne pas lire la zone de données de la table SAS. En fait, elle ne lit et ne modifie que la zone descriptive. Donc, pour une table qui contiendrait des millions d'observations, on réaliserait des gains de performance importants, simplement parce que les données ne se sont pas lues! L'encadré 2 illustre cette approche efficace. L'instruction LIBRARY indique la bibliothèque de données, alors que l'instruction MODIFY spécifie le fichier de données SAS à modifier. Ensuite, il ne reste qu'à ajouter l'instruction FORMAT et le tour est joué!

Encadré 2

```
10 proc datasets library=lib nolist;
11     modify dsn;
12     format var format.;
13 quit;
14 run;
```

Un vrai couteau suisse

La procédure DATASETS est très polyvalente. En plus de modifier les attributs des variables d'un fichier de données SAS, on peut aussi s'en servir pour d'autres aspects de la gestion de tables SAS. La liste suivante, même si elle est loin d'être exhaustive, donne un aperçu des capacités de cette procédure.

- Création d'un index pour une table ;
- Effacement d'une table dans une bibliothèque de données ;
- Écriture du contenu de la zone descriptive d'une table dans le fichier journal ;
- Ajout de contraintes d'intégrité à une table ;
- Gestion des nouvelles versions des tables (*generation datasets*)

À l'encadré 3, la procédure DATASETS est utilisée pour initialiser les *generation datasets* en SAS. Dans cet exemple, on demande à SAS de conserver les changements apportés à cette table, à partir de quatre copies de table dsn (genmax=4), et de les écrire dans un *audit trail*. On peut ainsi faire la gestion de la nouvelle version de cette table.

Encadré 3

```
29 proc datasets library=lib;
30     modify dsn (genmax=4);
31     audit dsn; initiate;
32 run;
33 quit;
```

La création d'index, qui servent à accélérer les requêtes, est un vrai jeu d'enfants avec la procédure DATASETS! Dans l'exemple de l'encadré 4, on utilise l'instruction INDEX CREATE afin de créer un index simple sur la variable **var**. À noter que la création d'index composite est aussi possible.

Encadré 4 - Création d'un index

```
35 proc datasets library=lib nolist;
36     modify dsn;
37     index create var / nomiss unique;
38 quit;
```

Dans l'exemple de l'encadré 5, des contraintes d'intégrité (*Integrity Constraints*) pour les variables **gender** et **age** sont créées à l'aide de l'instruction IC CREATE. Lorsque cette table sera ultérieurement modifiée, par exemple en lui ajoutant des observations avec la procédure APPEND, les valeurs de ces variables devront respecter ces contraintes, à défaut de quoi la modification ne sera pas acceptée.

Encadré 5 - Création de contraintes d'intégrité

```
49 proc datasets
50     library=lib nolist;
51     modify dsn;
52     ic create val_sex = check(where=(gender in ('M','F')))
53     message = "Valeurs valides pour GENDER: 'M' ou 'F'.";
54     ic create val_age = check(where=(age >= 8 and age <= 120))
55     message = "Une valeur invalide pour AGE a été donnée.";
56 quit;
```

Conclusion

Quand on travaille avec SAS, il est important de développer des techniques de travail efficaces. Elles-ci seront de véritables alliées quand on aura à se frotter à des fichiers de données SAS comportant un grand nombre d'observations. Pour faire la gestion des tables SAS, la procédure DATASETS est toute indiquée : elle est très efficace car elle ignore la zone de données et elle ne travaille que sur la zone descriptive. Si vous ne la connaissiez pas déjà, je vous invite à découvrir ses multiples usages. Bon printemps!

Référence

Pour de la documentation technique concernant le module SAS/BASE pour la version 9.2, visitez <http://support.sas.com/documentation/cdl/en/proc/61895/HTML/default/a000055590.htm>

Caricature : l'inflation !



Caricature parue le 25 février 2010 dans le journal *Le Droit*

À inscrire à votre agenda

XXVIe Symposium international sur les questions de méthodologie

Lieu : Ottawa, Ontario, Canada

Dates : du 26 au 29 octobre 2010

Le Symposium international de 2010 sur les questions de méthodologie de Statistique Canada se déroulera du 26 au 29 octobre 2010 à l'hôtel Crowne Plaza (situé en plein cœur du centre-ville d'Ottawa).

Ce Symposium sera intitulé « Statistiques sociales : interaction entre recensements, enquêtes et données administratives ». Y sont conviés tous les membres de la communauté, qu'ils proviennent d'organismes de recherche privés, gouvernementaux, universitaires ou autres, et particulièrement ceux qui s'intéressent aux enjeux méthodologiques ou statistiques qui découlent de l'utilisation de sources multiples de données (recensements, enquêtes ou données administratives).

Des ateliers auront lieu le premier jour, tandis que des séances plénières et plusieurs séances de présentations en parallèle portant sur une vaste gamme de sujets se tiendront les jours suivants. D'autres recherches et résultats seront communiqués au moyen de séances de présentations par affiches.

Consultez notre site web régulièrement afin d'obtenir plus de détails sur le programme, les ateliers, l'inscription, l'hébergement, et les activités touristiques au <http://www.statcan.gc.ca/conferences/symposium2010/index-fra.htm>.

Statistique et vie :

Vous avez dit statis ... quoi?

64,8% des personnes équipées d'un appareil auditif le portent du matin au soir, selon une étude de *l'Institut de recherche pour le confort auditif et d'audition santé et bien-être*, un regroupement d'audioprothésistes. Ce qui n'est pas normal, ont dû reconnaître les professionnels. Mais pourquoi 35% des utilisateurs laissent-ils leur prothèse dans le tiroir de la table de nuit?

Il semble que ce soit parce qu'elle est mal réglée et que l'on néglige le suivi de l'appareillage. Pour les lunettes, ce serait un peu la même chose : on attend, on ne sait quoi d'ailleurs, avant d'aller consulter lorsque nécessaire. De toute manière, il semble que « l'appareil auditif ait encore du mal à décoller ».

[Tiré de l'Almanach Vermot 2006]

Conférences à venir

Quoi ?	Où ?	Organisateurs ?	Quand ?
			2010
<u>38^e Congrès annuel de la SSC</u>	Québec (Québec)	<u>Société statistique du Canada</u>	23 au 26 mai
<u>10^e Symposium international en hydraulique stochastique & 5^e Conférence internationale en ressources en eau et recherche en environnement</u>	Québec (Québec)	<u>AIRH (Association internationale en ressources hydrauliques)</u>	5 au 7 juillet
<u>Joint Statistical Meetings</u>	Vancouver (Canada)	<u>American Statistical Association</u>	31 juillet au 5 août
COMPSTAT 2010 - The 19 th International Conference on Computations Statistics	Paris (France)	<u>Société francophone de classification</u>	22 au 27 août
The 31 st ISCB Annual conference	Montpellier (France)	<u>Société internationale en biostatistiques cliniques</u>	29 août au 2 septembre
<u>2^eme Colloque francophone international sur l'enseignement de la statistique</u>	Bruxelles (Belgique)	<u>Société française de statistique</u>	8 au 10 septembre

Le site Web de l'Institut international de statistique (<http://isi.cbs.nl/calendar.htm>), basé à La Haie (Pays-Bas), dresse une liste assez exhaustive des événements internationaux importants à venir. Le calendrier ci-dessus est sujet à changement; consultez le site Web des organisateurs pour de plus amples informations.

Nouvelles publications

Voici un choix de monographies, de livres de cours et d'ouvrages divers récemment publiés concernant la statistique. Si d'autres ouvrages vous semblent intéressants, n'hésitez pas à nous contacter. Il nous fera plaisir de les ajouter à notre répertoire.

Ouvrages pédagogiques en probabilités et en statistique ...

Titre	Auteur (s), éditeur, date de parution	Notes ⁽¹⁾
Les fractales – Apprenez à les connaître, les dessiner, les aimer	Florence Messineo, <i>Ellipses</i> (2010)	Ce livre s'adresse à tout public intéressé aussi bien par l'aspect esthétique des fractales que par leurs si intrigantes propriétés. Il offre de faire un tour d'horizon ludique des fractales; s'amuser à les connaître, les dessiner; se donner des idées pour ensuite faire ses propres essais et constructions à l'ordinateur.
Les petites tables : 17 tables financières et statistiques	Jean-Pascal Gayant, <i>Dunod</i> (2010)	Dix-sept tables statistiques et financières indispensables pour tout étudiant souhaitant réussir ses examens de statistique, mathématiques financières, économétrie et finance. L'utilisation des tables est facilitée par une explication pratique et synthétique.
Probabilités, variables aléatoires, lois classiques (Tome 1)	Virginie Delsart & Nicolas Vaneecloo, <i>Septentrion</i> (2010)	Ce premier tome de la série « Méthodes statistiques de l'économie et de la gestion » est consacré à la maîtrise du calcul des probabilités et de la notion de variable aléatoire, ainsi qu'à la connaissance des lois de probabilités classiques et de leurs applications.
Statistiques médicales et biologiques	Fethi Borsali, <i>Ellipses</i> (2010)	Le présent ouvrage est destiné aux étudiants et aux chercheurs des disciplines médicales et biologiques désirant compléter leur travail de recherche par des méthodes statistiques, ou s'initier à des méthodes statistiques plus approfondies. Chaque chapitre contient de nombreux exemples basés sur des situations rencontrées en médecine et en biologie. La plupart des démonstrations ont été évitées. La lecture de cet ouvrage ne nécessite aucune connaissance de mathématiques, sauf quelques notions de l'enseignement secondaire.

Et à titre de divertissement scientifique ...

Titre	Auteur (s), éditeur, date de parution	Notes ⁽¹⁾
Le nombre d'or – Quelques propriétés, quelques constructions	Florence Messineo, <i>Ellipses</i> (2010)	Ce livre s'adresse à un public intéressé aussi bien par l'aspect esthétique du nombre d'or que par ses intrigantes propriétés. Il offre de faire un tour d'horizon ludique du nombre d'or et de s'amuser à découvrir les innombrables figures géométriques où il est présent.
21 énergies renouvelables insolites	Denis Bonnelle & Renaud de Richter, <i>Ellipses</i> (2010)	Les limites des énergies fossiles sont aujourd'hui évidentes, qu'il s'agisse de la ressource ou de la capacité de l'atmosphère à absorber leurs émissions de CO ₂ . Certaines des énergies renouvelables les plus connues, comme les biocarburants, ont montré qu'elles pouvaient, elles aussi, être limitées. L'avenir nous réserve-t-il d'autres mauvaises surprises ?

¹Les « notes » sont extraites des notices descriptives préparées par les éditeurs.

Pour les intéressés, vous trouverez sur le site de l'AMQ, dans la rubrique « Lu pour vous » du Bulletin AMQ, des présentations d'un choix de livres de mathématiques et de statistique.

Statistique et vie :

Productivité statistiquement significative

Une entreprise britannique, *Farrelly Facilites and Engineering*, a mis au point une tactique pour améliorer sa productivité. Ce n'est d'ailleurs pas la tactique du gendarme, mais plutôt le contraire. Voyez le programme : les heures supplémentaires sont proscrites, on diffuse de la musique douce, les cinquante salariés n'ont pas le droit de travailler le jour de leur anniversaire. Des sorties aux frais de la société sont programmées, jamais de travail après 17 heures et des embrassades à l'arrivée et au départ du bureau. Beau programme, direz-vous, mais cela donne quoi?

Et bien si l'on en croit la direction de l'entreprise, cela a suffi à faire grimper les bénéfices de 30 % en trois ans, et la productivité a triplé. C'est le nouveau directeur, initié aux philosophies orientales, qui a mis au point la nouvelle formule de travail. Il a donné des idées, mais c'est surtout le personnel qui a enclenché la machine, chacun apportant les siennes. Allez, on se fait la bise et on va bosser...!

[Tiré de l'Almanach Vermot 2006]

Suivre son cours ...

La vie suit son cours, mais qu'en est-il du statisticien qui sommeille en vous ?

Cette grille présente les cours offerts par différents organismes. La liste des cours universitaires est disponible sur demande auprès de l'ASSQ. L'information fournie dans cette grille est sujette à changements. Le lecteur est invité à communiquer avec les responsables des cours pour corroborer et compléter l'information présentée (frais d'inscription, dates, etc.).

Organisme	Clientèle	Types de cours	Contacts
Creascience	Débutant, intermédiaire, chercheurs, techniciens, professionnels de la R&D	Planification d'expériences, analyse en composantes principales, contrôle de la qualité	Natalie Rodrigue (514) 840-9220, poste 27 www.creascience.com info@creascience.com
Services conseils Hardy	Débutant, intermédiaire, avancé	Formation sur les logiciels SAS, SPSS, MINITAB, Statistica, Crystal Reports	Monique Trempe (418) 626-1666 www.schardy.qc.ca schardy@schardy.qc.ca
Institut SAS	Grand public, entreprises, débutant ou expert	Data mining, langage macro, programmation, rédaction de rapports, statistique	Chantal Lessard (514) 395-4071 http://support.sas.com/training/canada/sastrain@can.sas.com
École de Technologie Supérieure	Débutant	Planification d'expérience en industrie, contrôle de la qualité	Suzanne LeBel (514) 396-8830 (Service de perfectionnement) www.perf.etsmtl.ca perf@etsmtl.ca
Institut de formation de Statistique Canada	Intermédiaire et avancé	Analyse de données (séries chronologiques, contrôle de la qualité, données de survie) et méthodologie d'enquête	Céline Charette (613) 951-1044 www.statcan.ca charcel@statcan.ca Note : Certains cours peuvent se donner à l'extérieur.
Centre interuniversitaire québécois de statistiques sociales	Avancé	Formations intensives sur des problèmes méthodologiques particuliers à l'analyse de données complexes et données d'enquêtes particulières	Heather Juby, Coordinatrice, Transfert du savoir (514) 343-2090, poste 8 www.cigss.umontreal.ca heather.juby@umontreal.ca
Université Concordia cours en ligne	Étudiants inscrits	Cours en ligne	Patrick Devey, Directeur Patrick.devey@concordia.ca

Séminaires

Des séminaires ont lieu régulièrement aux endroits suivants. N'hésitez pas à contacter les responsables pour plus de détails.

<p>UQAM Pascale Rousseau Tél. 514-987-3000, # 3224 Fax 514-987-8935 www.uqam.ca rousseau.pascale@uqam.ca</p>	<p>Université de Montréal Martin Bilodeau Tél. 514-343-2410 Fax 514-343-5700 www.umontreal.ca bilodeau@dms.umontreal.ca</p>	<p>Université de Sherbrooke Bernard Colin Tél. 819-821-8000, # 2012 Fax 819-821-8200 www.usherbrooke.ca bernard.colin@dmi.usherb.ca</p>	<p>Université Laval Thierry Duchesne Tél. 418-656-5077 Fax 418-656-2817 www.ulaval.ca Thierry.Duchesne@mat.ulaval.ca</p>
---	--	--	---