



Association  
des statisticiennes  
et statisticiens  
du Québec

# Convergence

Volume XXV Numéro 1

Février 2020

## Mot du rédacteur en chef

Voici, chères lectrices et chers lecteurs, votre *Convergence*. Un numéro assez particulier parce que la grande majorité des articles proviennent de collaborateurs externes à l'ASSQ. Pour être franc, nos collaborateurs habituels ont été très occupés ces derniers temps et ils ont eu quelques problèmes à écrire un article pour le présent numéro. Eh bien, qu'à cela ne tienne! Je les laisse se reposer un peu pour que le numéro d'août prochain soit aussi touffu que celui d'août 2019. Je profite donc de l'espace libre pour remplir d'articles divers sur la statistique dont plusieurs sont tirés de l'*Almanach Vermot*. Pour ceux qui ne le connaissent pas, cet almanach se dit le « petit musée des traditions et de l'humour populaires français ». Il contient souvent des articles statistiques à saveur humoristique que je me plais à lire. Je vous en ai sélectionné quelques-uns que je partage avec vous dans ce *Convergence*.



Ce numéro contient aussi un excellent article de Jean-François Cliche du *Soleil* de Québec : « Les sondages influencent-ils le vote? » Un sujet qui passionne, notamment, nos cousins français. Je remercie grandement mon collègue et ami, Jean-Jacques Droesbeke de l'Université Libre de Bruxelles, de me permettre de reproduire son article à propos du site « Histoire des mathématiques » de Bernard Ycart, une perle pour tout amateur de l'histoire de la statistique. Un gros merci également à Yanick Beaucage, de Statistique Canada, pour son article sur la science des données. Finalement, j'ai inséré un article tiré du magazine *7 Jours* (si, si...) sur le nombre total d'humains ayant vécu sur Terre.

Somme toute, je crois que ce *Convergence* n'est pas mal du tout...

Bonne lecture!

Pierre Lavallée

<b>Mot du rédacteur en chef</b>	<b>1</b>
<b>Mot du président</b>	<b>2</b>
<b>Le cerveau des femmes</b>	<b>8</b>
<b>Nouvelles de Statistique Canada</b>	<b>3</b>
<b>Au total, combien d'êtres humains ont vécu sur terre?</b>	<b>5</b>
<b>Les sondages influencent-ils le vote ?</b>	<b>6</b>
<b>Le site « Histoire des mathématiques » de Bernard Ycart</b>	<b>8</b>
<b>Démographie des fourmis</b>	<b>12</b>
<b>Statistique et vie</b>	<b>13</b>
<b>À propos de l'ASSQ</b>	<b>15</b>
<b>Membres institutionnels</b>	<b>23</b>

## Mot du président

L'automne dernier, le CA a été remanié pour pourvoir tous les postes. Nous avons d'abord accueilli une nouvelle secrétaire, Audrey-Anne Vallée. Audrey-Anne vient d'être engagée comme professeure de statistique à l'Université Laval. Elle a fait ses études de baccalauréat et de maîtrise en statistique à l'Université de Montréal. Son directeur de maîtrise était David Haziza. Pour ses études doctorales, elle est allée à l'Université de Neuchâtel avec le professeur Yves Tillé, un spécialiste en échantillonnage. En fait, j'ai rencontré Audrey-Anne lors d'un stage de recherche à Neuchâtel; Audrey-Anne, Yves et moi avons travaillé ensemble pour concevoir un plan de sondage pour l'inventaire forestier québécois qui réponde, le mieux possible, aux objectifs de maintenir les coûts à un niveau acceptable tout en captant la diversité de la forêt québécoise. Bastien Raymond-Ferland qui, à l'époque, travaillait aux inventaires forestiers nous a fourni des données pour tester la méthodologie que nous avons développée. Ceci a donné lieu à un article scientifique :

A.-A. Vallée, B. Raymond-Ferland, L.-P. Rivest & Y. Tillé (2015) Incorporating Spatial and Operational Constraints in the Sampling Designs for Forest Inventories. *Environmetrics*, 26; 557–570.

Revenons au CA. Roxanne Brousseau, notre ancienne secrétaire, a accepté le poste de directrice des communications et c'est elle qui pilote le dossier de la mise aux normes du site web de l'association. C'est un projet en cours dont nous pourrions discuter lors de la prochaine assemblée générale annuelle. Le CA est donc complet; les autres membres sont Éric Lacroix (registraire), Laurence Desbois-Bédard (trésorière), Valérie Roy (vice-présidente) et Jasmin Ricard (représentant étudiant).

La date du prochain colloque est maintenant arrêtée; il aura lieu le vendredi 8 mai 2020. Réservez cette date à votre agenda. Cette année nous avons retenu l'Auberge Saint-Antoine dans le vieux Québec comme lieu de rencontre (voir : <https://www.saint-antoine.com/fr>).

C'est un site patrimonial au cœur du Vieux-Québec. Les inondations record de 2019 ont naturellement suggéré un thème pour notre prochain colloque : événements extrêmes et gestion des risques. Nous avons retenu des conférenciers très intéressants. La tête d'affiche sera Christian Genest de l'Université McGill qui nous entretiendra de « Méthodes statistiques d'évaluation des risques de catastrophes environnementales ». Il sera accompagné de Alexis Hannart (Ouranos) et de Bastien Raymond-Ferland (Desjardins). Nous espérons que vous serez des nôtres.

Louis-Paul Rivest



## Nouvelles de Statistique Canada : La science des données fait sa place à Statistique Canada

*Dans cette édition du Convergence, je partage avec vous un article de mon collègue Yanick Beaucage, qui décrit l'évolution de la place prise par la science des données à Statistique Canada. Yanick est bien placé pour nous en parler, lui qui est au cœur des activités de ce domaine dans son rôle de directeur adjoint de la nouvelle Division de la science des données à Statistique Canada. N'hésitez pas à le contacter pour en savoir plus. Bonne lecture!*

*François Brisebois, Statistique Canada (francois.brisebois@canada.ca)*

Dans les dernières années, en raison de la « révolution des données » et au développement fulgurant des infrastructures informatiques, la science des données a pris du galon. Tous les jours, on entend parler de nouvelles applications révolutionnaires qui font appel à l'intelligence artificielle et à l'apprentissage automatique, que ce soit pour contrôler les voitures, créer de nouvelles compositions musicales ou encore déceler des signes d'un cancer lors de la lecture de radiographies. Mais comment les organismes nationaux statistiques peuvent-ils en faire usage pour, par exemple, faciliter et accélérer leur travail, accéder à de nouvelles sources de données ou encore accélérer le traitement et le niveau de détails des publications? L'article du *UNECE Machine Learning Team* (2018) fait un bon survol des possibilités d'utilisation de ces nouveaux outils dans la production de statistiques officielles, par exemple lors de la classification ou encore de l'estimation.

À Statistique Canada, une petite équipe s'est d'abord intéressée aux données provenant de lecteurs optiques et a tenté de s'en servir pour estimer le total des ventes de marchandises d'un détaillant national. Le défi consistait à coder automatiquement des millions de produits selon le système de classification des produits de l'Amérique du Nord (SPAN) à partir de l'information contenue dans les données provenant de lecteurs optiques, puis de sommer les ventes. En conséquence du succès de ce projet, l'intérêt pour ce genre de travail s'est accru, puis, pour aborder une nouvelle vague d'initiatives, l'équipe en place s'est agrandie, devenant alors formellement connue sous le nom de l'Accélérateur de la science des données! Tout à coup, plusieurs projets qu'on croyait jusqu'alors un peu trop ambitieux devenaient possibles! L'équipe s'est rapidement développée pour pouvoir faire face à de nouveaux défis tels que l'interprétation d'images satellites pour déterminer le type de grandes cultures ensemencées dans les provinces de l'ouest, le dénombrement de serres ou encore de panneaux solaires; l'interprétation des données de sites web d'entreprises pour pouvoir lister les produits vendus; la lecture des journaux pour tenter de déceler les tendances économiques ou pour signaler les changements affectant les entreprises canadiennes. En septembre 2019, Statistique Canada reconnaissait officiellement la place importante que la science des

données prenait dans l'univers des nouveaux outils statistiques en créant la Division de la science des données au sein de la Direction de la méthodologie, nouvellement renommée Direction des méthodes statistiques modernes et de la science des données, pour justement refléter cette nouvelle réalité. Cette nouvelle division montre bien la diversité de la science des données puisqu'elle réunit à la fois des scientifiques des données, des méthodologistes et des économistes.

Parmi les projets actuels de la division, notons une collaboration avec Santé Canada pour un projet concernant les produits naturels. À partir d'une photo de l'étiquette, on détecte les ingrédients qui composent le produit afin de déterminer si le produit devrait être effectivement considéré comme un produit naturel. Une autre avenue étudiée présentement est l'utilisation des algorithmes d'apprentissage automatique pour remplacer les algorithmes traditionnels dans la recherche. Que ce soit avec des données sur la santé mentale ou sur les opioïdes, notre équipe tente de tirer profit de ces algorithmes qui souvent permettent de modéliser des relations non linéaires. Nous travaillons également sur un projet explorant la possibilité de produire des estimations en temps réel. L'idée est de confirmer les tendances des estimations produites à l'aide de l'analyse de journaux. Nous nous penchons aussi sur un grand nombre de projets de classification, que ce soit pour trier des commentaires ou encore pour coder automatiquement les marchandises vendues, importées ou transportées. D'ailleurs, la nouvelle version (version 3.2) du progiciel G-Code inclut dorénavant des algorithmes utilisant l'apprentissage automatique tel que XGBoost et fastText (développé par Facebook). Maintenant que nous avons commencé à utiliser ces nouvelles techniques, une de nos prochaines tâches sera d'uniformiser les applications de codage automatique pour l'ensemble des variables qui nécessitent une classification standardisée. Nous travaillons à l'aide de logiciels de type code source ouvert tel que R ou encore Python, et ce, dans plusieurs environnements informatiques dont l'infonuagique ou encore sur des ordinateurs munis de GPU (Graphics Processing Unit), qui permettent le traitement d'images. D'ailleurs nous travaillons également sur un projet de chiffrement homomorphique qui permettra de faire des calculs dans le nuage sans jamais y décrypter nos données.

L'équipe fait partie du comité de gestion ou dirige un certain nombre de communautés de pratiques reliées aux sciences des données à Statistique Canada. Que ce soit la communauté de pratique de l'apprentissage automatique, le traitement du langage naturel, le traitement des images satellites ou encore le groupe des utilisateurs de R/Python, les membres y sont très actifs, présentant leurs travaux et participant à l'avancement de la connaissance dans ces domaines. Nous tentons également d'améliorer le niveau de connaissances en apprentissage automatique en coordonnant des activités de formation pour permettre à nos méthodologistes d'acquérir de nouvelles connaissances. Puisque ces activités ont besoin d'encadrement et de direction, nous sommes en train de développer un cadre de travail pour un usage responsable de l'apprentissage automatique. Ce cadre mettra en lumière les efforts requis pour assurer un usage responsable de l'apprentissage automatique pour différents thèmes tels que le

respect des Canadiens, le respect des données, l'utilisation de méthodes et d'applications valides. De plus, nous sommes en train de mettre sur pied un comité de coordination qui pourra valider le suivi des lignes directrices du cadre avant de mettre les nouvelles applications en production. Comme vous pouvez le voir, Statistique Canada est en train de s'adapter aux nouveaux outils reliés à l'apprentissage automatique et est en bonne posture pour prendre le virage de la science des données!

Yanick Beaucage, Statistique Canada ([yanick.beaucage@canada.ca](mailto:yanick.beaucage@canada.ca))



Statistique  
Canada

Statistics  
Canada

### Référence:

UNECE Machine Learning Team: Wesley Yung (Canada), Jukka Karkimaa (Finland), Monica Scannapieco (Italy), Giulio Barcarolli (Italy), Diego Zardetto (Italy), José Alejandro Ruiz Sanchez (Mexico), Barteld Braaksma (Netherlands), Bart Buelens (Netherlands), Joep Burger (Netherlands), [The use of machine learning in official statistics](#), Novembre 2018.

---

## Au total, combien d'êtres humains ont vécu sur terre?

Si l'on en croit les démographes, plus de 80 milliards d'hommes et de femmes ont peuplé notre planète depuis le début de l'humanité. La plupart y ont vécu au cours des derniers siècles. Entre 600 000 et 10 000 ans avant Jésus-Christ, 10 milliards d'êtres humains ont existé; de l'an -10 000 à -5000, 10 milliards; de l'an -5000 à l'ère chrétienne, 29 milliards; jusqu'à l'an 1000, 9 milliards; de 1001 à 1800, 12 milliards; de 1801 à 1900, 4,6 milliards; de 1901 à 1997, 9,4 milliards. La Terre a été peuplée pour la première fois par un milliard d'hommes et de femmes à la fois vers 1801. Cent vingt-quatre ans plus tard, en 1925, elle comptait deux milliards d'habitants. Puis, la progression s'est accélérée : trois milliards en 1959 (34 ans plus tard); quatre milliards en 1974 (15 ans plus tard); cinq milliards en 1987 (13 ans plus tard); six milliards en 1999 (12 ans plus tard). À ce rythme, la population mondiale pourrait atteindre quelque 10 milliards d'individus à la fin du XXI<sup>e</sup> siècle.

[Article tiré du magazine *7 Jours* de juillet 2002.]

## Les sondages influencent-ils le vote ?

### L'affirmation

« Les sondages sont de plus en plus nombreux et de plus en plus pointus [par circonscription] pendant les campagnes électorales, et on entend souvent dire qu'ils influencent le vote. Mais jusqu'à quel point est-ce vrai? Y a-t-il des études là-dessus ? Des gens sont-ils vraiment tentés de voter comme la majorité ou alors contre un candidat, et ainsi priver un tiers parti des votes nécessaires à sa survie? » demande Gilles Labrecque, de Saint-Jean-Port-Joli.

### Les faits

Sans le contexte, la plupart des gens diraient que la citation suivante doit dater d'il y a à peine 10 ou 15 jours, alors que la campagne fédérale s'achevait : « Était-il bon ou néfaste que ces sondages aient été rendus publics? Ces chiffres peuvent-ils avoir influencé l'opinion publique? Et si tel est le cas, était-ce une mauvaise influence? »

En fait, il s'agit d'un extrait du livre *Are Opinion Polls Useful?* (Les sondages d'opinion sont-ils utiles?) qui fut publié en... 1946, lit-on sur le site de l'Association d'histoire américaine. [bit.ly/2q8bacY]. Alors oui, il y a eu des études sur l'influence des sondages et, depuis le temps, il y en a même eu beaucoup.

Que disent-elles? Dans l'ensemble, on ne peut pas dire que leurs résultats montrent clairement une influence des sondages sur le vote populaire. Du moins, pas une grosse, et pas au sens où on l'entend habituellement, où des individus décident de changer leurs intentions de vote pour « suivre » la majorité — j'y reviens tout de suite. *Certaines* études ont trouvé des signes d'un effet des sondages, mais d'autres non.

Par exemple, une étude danoise récente [bit.ly/2NsOrjs] a comparé les intentions de vote de trois groupes lors d'une campagne électorale. Un groupe s'est simplement fait demander pour qui ils comptaient voter : 18 % d'entre eux ont dit qu'ils comptaient appuyer le Parti social-démocrate (PSD). Un autre groupe a dû lire au préalable un (faux) article de nouvelles sur un sondage qui montrait prétendument le PSD en perte de vitesse: 16 % ont ensuite dit qu'ils allaient voter PSD. Et un troisième groupe a été exposé à un autre faux sondage qui indiquait, lui, que le PSD était en progression : 22 % de ces gens ont ensuite affirmé qu'ils appuieraient le PSD.

Cependant, les différences entre ces groupes n'étaient pas toutes « statistiquement significatives », ce qui signifie que les chances pour que ces écarts soient simplement dus au hasard sont trop fortes pour qu'on les considère tous comme « réels », et ce, même si l'échantillon était assez grand (autour de 600 personnes par groupe), ce qui suggère fortement que l'effet des sondages, en présumant qu'il existe, serait plutôt faible.

D'autres travaux ont eux aussi conclu à un effet réel, mais assez faible (par exemple : stanford.io/34jyAdR), mais il y a aussi des études qui n'ont pas trouvé de signes montrant que les sondages ont vraiment un « effet de contagion » (les gens suivent la majorité) ou de

« découragement » (des électeurs qui ne votent pas parce qu'un autre parti que le leur semble trop fort). Ainsi, des chercheurs des universités de Montréal, McGill et Waterloo ont publié l'année dernière des résultats portant précisément là-dessus dans la revue *Options politiques* Il s'agissait de sondages « en deux temps » administrés avant et après les élections fédérales et provinciales de 2007 à 2015 au Québec. En demandant aux participants s'ils lisaient les sondages, ces chercheurs ont réalisé qu'en fait, les électeurs qui prêtent le plus attention aux sondages sont ceux qui sont *les moins susceptibles* de changer de parti au cours de la campagne. De plus, leur taux de participation est plus élevé que la moyenne (donc pas d'effet de découragement) et ils ne sont pas plus nombreux que les autres à voter pour le parti qui est en tête dans les sondages.

Ce sont là des comportements qui ne sont pas du tout cohérents avec l'idée que les sondages influencent vraiment le vote, et ils ne surprennent plus la chercheuse Claire Durand, spécialiste des sondages de l'Université de Montréal et coauteure de cette étude : « ça fait cinq élections qu'on fait cet exercice-là et qu'on voit la même chose à chaque fois », m'a-t-elle dit en entrevue.

Ces résultats montrent par ailleurs que les électeurs les plus susceptibles de changer leur fusil d'épaule pendant une campagne électorale sont ceux qui ont l'intention de voter pour des « petits » partis. Par exemple, aux provinciales de 2008, les partisans de Québec solidaire, d'Option nationale et du Parti vert ont été (proportionnellement) entre 3 et 17 fois plus nombreux à changer d'idée avant le scrutin que ceux du Parti québécois.

Évidemment, on peut voir là-dedans un signe que les sondages favorisent le « vote stratégique » (voter pour un choix autre que son favori afin de bloquer un autre parti), et il est entièrement possible que les sondages aient ce genre d'effet. Mais il s'agit là d'un « effet d'information », souligne Mme Durand, pas d'un effet d'influence (au sens de « contagion ») à proprement parler, comme on l'entend si souvent.

« Tout indique, concluent Mme Durand et ses collègues dans *Options politiques*, que la décision prise par les électeurs dépend beaucoup plus des positions et messages véhiculés par les partis et leurs chefs, ou un simple attachement à un parti, que de l'information que fournissent les sondages sur la popularité des partis. Nous pensons qu'il y a une surestimation de l'effet des sondages dans la sphère et les débats publics, puisque les analyses montrent que les sondages n'ont pas une influence importante sur le choix électoral des citoyens. Ils sont plutôt une information parmi d'autres utilisée dans le processus de décision. »

### **Le verdict**

Loin d'être clair. Certaines études suggèrent que les sondages provoquent bel et bien un effet d'entraînement à la faveur du meneur ou d'un parti en progression, mais d'autres travaux arrivent à des conclusions contraires.

Dans tous les cas, si effet il y a, il semble qu'il soit de faible envergure.

[Jean-François Cliche, Le Soleil](#)

[Article paru dans le quotidien *Le Droit* du 31 octobre 2019.]

## Chronique historique : Le site « Histoire des mathématiques » de Bernard Ycart



Bernard Ycart est professeur à l'Université de Grenoble-Alpes. Parmi ses nombreux centres d'intérêt, on relève l'histoire des mathématiques. Il a développé au cours des années un site ([hist-math.fr](http://hist-math.fr)) sur lequel on trouve des animations d'une vingtaine de minutes, consacrées à des questions diverses de cette discipline.

Six domaines sont concernés par ce site. Ils présentent tous entre 20 et 30 animations chacun. Les trois premiers forment le socle de l'enseignement obligatoire des mathématiques : la géométrie, l'algèbre et l'arithmétique. Viennent ensuite deux autres sujets : les savants et l'informatique. La lectrice et le lecteur trouveront en annexe de cette note la liste des thèmes traités dans ces cinq parties. Nous préférons nous centrer sur le dernier domaine concerné par ce site. Il est intitulé : Statistique.

Il eût peut-être été utile d'intituler ce domaine « probabilités et statistique ». En parcourant la liste des 23 sujets qui y sont traités, on trouve en effet une petite moitié d'animations consacrées aux probabilités. La lectrice et le lecteur trouveront ci-dessous la liste des titres, chacun d'eux étant accompagné d'un sous-titre. Nous ne commenterons pas en profondeur l'intégralité de ces animations. Nous préférons en donner les caractéristiques essentielles, ce qui ne nous empêchera pas de faire certaines remarques spécifiques.

**1. La combinatoire des dés** – du Mahābhārata à Galilée : après quelques considérations sur les jeux, nous découvrons un manuscrit médiéval De Vetula, pour ensuite côtoyer les auteurs classiques du genre Cardano (*De ludi aleae*) et Galilée (*Considerazione sopra il giuco de dadi*). Pour les amateurs !

**2. La proportion des voyelles** – d'Alberti à Markov : un traité indien sur les arts de la scène, une lettre sur la musique française de Rousseau, le Dictionnaire philosophique de Voltaire, deux citations de Charles Quint (?) nous mènent jusqu'au peintre, architecte, mathématicien... Alberti, dont l'étude des proportions de voyelles et de consonnes dans les poèmes est relue grâce à la probabilité. Un petit détour par le Monthly Magazine de 1799 et nous nous retrouvons dans l'*Encyclopedia Americana* de Liber avant de découvrir Andrei Markov dont les chaînes doivent beaucoup aux « 20000 épreuves connectées qui sont soit une voyelle, soit une consonne ». Un parcours peu fréquenté dans les enseignements.

**3. La géométrie du hasard** – le problème des partis : partant de la divine proportion de Pacioli et remontant un peu le temps, on n'est pas surpris de trouver Pascal, sa correspondance avec Fermat pour terminer avec Huygens. Un classique évidemment.

**4. La combinatoire des hexamètres** – de Bernoulli à Knuth : « Belle Marquise, vos beaux yeux me font mourir d'amour. » Cette réplique que Molière met dans la bouche du Bourgeois Gentilhomme (acte II, scène IV) est une belle introduction au sujet de cette animation. Un Epigrammatum, des dactyles, spondées et trochées nous conduisent à l'hexamètre, cette technique de versification antique revenue à la mode à la Renaissance. Un petit détour chez Vossius et nous voilà chez Prestet et ses *Éléments de mathématiques* (1675 puis 1689). Le voyage n'est pas fini. On passe chez Wallis, puis chez Leibniz pour terminer chez Bernoulli (Jacques !) et... Knuth ! Oui, le voyage est un peu déroutant.

**5. Tables de mortalité** – parier sur la vie : débiter avec la mortifère peste du Decameron de Boccace est, il faut le dire, original. Les épidémies de peste qui se succédèrent ensuite nous amènent à Graunt, la référence en matière de tables de mortalité. Il occupe le cœur de l'animation, c'est normal, et est accompagné de son complice Petty pour présenter l'arithmétique politique qui fit fureur au XVII<sup>e</sup> siècle. Quant aux successeurs, Halley, Kerseboom, Süßmilch et autres, sans oublier Quételet, ils sont aussi présents. À recommander à ceux qui aiment la vie et la survie.

**6. L'état de l'État** – la statistique au siècle des Lumières : la statistique allemande d'Achenwald et le goût des enquêtes de Vauban sont mis à l'honneur ici, surtout le second. On y trouve aussi Uztariz et von Bliefeld pour compléter l'animation précédente et on termine par Sinclair. Une belle illustration d'une époque qui s'interroge.

**7. Un argument pour la divine providence** – la première  $p$ -valeur : les amateurs de tests statistiques ont déjà entendu parler d'Arbutnot. Ils pourront approfondir son importance et les conséquences de ses réflexions sur Nicolas Bernoulli, Gravesande, de Moivre et même Laplace.

**8. L'aiguille de Buffon** – naissance des probabilités géométriques : à conseiller aux amateurs de Buffon et du jeu du franc-carreau. Son aiguille les mènera à Laplace, mais aussi à Barbier et Bertrand, sans oublier Poincaré.

**9. Le paradoxe de Saint-Pétersbourg** – jouer en martingale : un épisode incontournable de l'histoire des probabilités initié dans une lettre de Nicolas Bernoulli à Montmort (1713). On y retrouve Buffon, mais aussi d'Alembert. Un autre Bernoulli, Daniel — quelle famille prolifique — intervient dans cette histoire qui en a troublé plus d'un. Laplace, Poisson sont sollicités. L'espérance morale a la cote. Nous voilà partis dans les salles de jeu du XVIII<sup>e</sup> siècle, la martingale devient un sujet à la mode. Borel s'en saisira au début du XX<sup>e</sup> siècle.

**10. La puissance du binôme** – le premier théorème central limite : bien sûr, de Moivre ouvre le bal. Le mémoire de Maty sur sa vie sert de fil conducteur. Son exode pour raison religieuse, ses rencontres sur le sol anglais, sa *Doctrine of chances* nous sont contés en détail jusqu'à son

argument-clé qui le conduisit à un intervalle de fluctuation. On vit la convergence de la loi binomiale vers celle qui deviendra « normale » un siècle plus tard.

**11. La probabilité des jugements** – de Cicéron à l'affaire Dreyfus : on sait que, dans l'Antiquité, la probabilité est un concept totalement qualitatif permettant de parler de ce qui se trouve entre le certainement vrai et le certainement faux. Il faut ensuite attendre le seizième siècle pour aborder les prémisses de la quantification de ce concept : les jeux de dés et lancements de pièces équilibrées permettent de citer Pascal, Huygens... mais aussi Arnaud et Nicole. Bien sûr, on débouche sur Jacques Bernoulli et son *Ars Conjectandi* qui a introduit l'approche fréquentiste de la probabilité. Pour introduire l'usage de la probabilité comme instrument de jugement, rien de tel que parler de Bayes et de son ami Price. « Le » théorème redécouvert par Laplace un peu plus tard est à l'honneur. Un petit détour chez Condorcet et l'admission des femmes au droit de cité, et nous voilà chez Poisson et ses Recherches sur la probabilité des jugements. À propos de jugement, celui de Dreyfus est bien connu. C'est l'occasion de rappeler le « J'accuse » de Zola et le rôle peu glorieux d'Alphonse Bertillon. Heureusement, Darboux, Appell et Poincaré étaient là ! Une des histoires que nous préférons !

**12. Le Mont-Blanc mathématique** – théorie analytique des probabilités : ne pas consacrer une animation à Laplace eût été difficile à justifier. Il est au centre de cette animation avec ses œuvres maîtresses en probabilité et son goût pour les approximations. On trouve aussi Fourier qui fit son éloge, et Arago qui fit l'éloge de ce dernier dont les Recherches statistiques... sont devenues célèbres. Auguste Comte est aussi convoqué ainsi que de Morgan qui a qualifié la Théorie des probabilités de Laplace de « Mont-Blanc de l'analyse mathématique ».

**13. Sympathique escroc** – la représentation graphique des données : Bernard Ycart nous conte par le détail la vie de William Playfair, cet original auteur d'un Statistical Breviary qui a vraiment donné le feu vert à une utilisation de graphiques en statistique. Oresme, Halley, Barbeau du Bourg, Fourcroy et Peuchet sont aussi appelés à témoigner.

**14. Loi de Gauss** – la probabilité des erreurs : partant des Éléments de la théorie des probabilités de Borel (1909), la loi de Gauss et ses fondements nous sont rappelés. Les écrits de Gauss sont bien sûr évoqués comme ceux de Legendre pour la méthode des moindres carrés liée à cette loi ainsi qu'à la moyenne, ainsi que ceux de Boscovich dont le Voyage astronomique et géographique rappelle le célèbre problème de la mesure d'un arc de méridien et les expéditions de La Condamine et Maupertuis.

**15. La difficulté du recueil** – statistique napoléonienne : vous rappelez-vous le Directoire, puis le Consulat, puis l'Empire ? Et les ministres de l'Intérieur (de Neufchâteau et Chaptal) qui ont tant peiné à recevoir des préfets les informations statistiques dont ils avaient besoin pour gérer la France ? Bien sûr, on ne peut éviter de parler du préfet Bottin qui a laissé son nom à un répertoire de moins en moins utilisé de nos jours. C'est aussi l'occasion d'évoquer les relations entre l'inventeur de la chaptalisation avec Mademoiselle Bourgoïn, de la Comédie française, et le plaidoyer de Fiévée qui proclame, dans le Journal des Débats, que « la statistique ne repose sur rien ».

**16. Statistique à la mode** – la carte choroplèthe : Dupin est l’auteur d’une carte thématique où les régions sont remplies d’une couleur en fonction des valeurs d’une statistique particulière. C’est le point de départ d’une promenade à l’époque de Dupin et d’Andraud, auteur d’une statistique morale de la France qui en a choqué plus d’un. C’est l’occasion de rappeler cette phrase de Balzac : « la statistique est devenue à la mode et c’est une position que de statistiquer ». Les Vivacités du Capitaine Tic de Labiche et Madame Bovary de Flaubert, sans oublier Offenbach, Courteline ou Alphonse Allais nous rejoignent à la fin de l’histoire. Il faut s’amuser de temps en temps !

**17. Des amateurs scrupuleux** – intuitions non paramétriques : on repart de Dupin pour une autre ballade au XIX<sup>e</sup> siècle, avec cette fois, Lamartine, Cournot, Bignon, d’Angeville, toujours dans le même ton.

**18. La fièvre puerpérale** – les statistiques de l’horreur : changement de décor ! Nous voilà au cœur de la médecine et du traitement des maladies « qui sentent ». C’est le moment de se rappeler qu’on a de la chance de vivre à notre époque où les statistiques médicales sont plus sérieuses.

**19. The Lady with a Lamp** – les maladies zymotiques : qui dit médecin dit infirmière. Et qui dit infirmière pense à Florence Nightingale, cette admiratrice de Quételet qui s’est illustrée pendant la guerre de Crimée. Le rôle de Farr et celui de Snow sont aussi évoqués dans cette animation consacrée à la santé publique et aux moyens d’en parler.

**20. La statistique littéraire** – croire en la loi des grands nombres : voilà un sujet intéressant qui nous permet de parler de cryptage, de fréquence d’apparition de lettres, de Georges Pécoc... On y rencontre à nouveau Arbuthnot, Bernoulli (Jacques) et la loi des grands nombres, Quételet et son penchant au crime, sans oublier ceux qui ont voulu comparer des auteurs avec des moyens statistiques.

**21. Dr. Jekyll et Mr. Hyde** – statistique et eugénisme : pour qui ne connaît pas le rôle de Galton, Bertillon, Pearson, Fisher dans cette philosophie douteuse dont les effets furent désastreux.

**22. Trop beau pour être vrai** – les expériences de Mendel : de quoi nous rappeler nos cours de botanique, de biologie, de zoologie ou de génétique.

**23. Une histoire de bière** – la statistique des petits échantillons : ce n’est pas une histoire belge, puisqu’il s’agit de Guinness et de son homme de science Gosset, célèbre pour avoir signé ses deux articles fondamentaux du nom de Student. Ses relations avec Pearson et Fisher sont aussi évoquées.

La lectrice et le lecteur l’auront compris, Bernard Ycart nous conte (c’est sa voix qu’on entend) des histoires souvent sérieuses avec un ton plaisant, voire ironique. Certaines sont plus anecdotiques que d’autres. Elles sont probablement trop longues pour être écoutées dans un cours, mais l’enseignant ou le curieux, comme l’enseignante ou la curieuse, pourront y trouver des anecdotes et des références qui sont parfois les bienvenues à un moment donné.

Voilà pour les histoires. Elles sont richement illustrées, agréablement présentées, jamais ennuyeuses. Un mode d'emploi est aussi proposé dans un forum de questions possibles. Si ce sont les personnages qui vous intéressent, le site vous propose d'en rencontrer quelques-uns (529, d'après mes calculs). Ils ne sont pas tous mathématiciens, mais on en parle dans les histoires. Dommage qu'ils ne soient pas mentionnés par ordre alphabétique ! Une autre amélioration possible serait d'étoffer les mots-clés en y intégrant quelques domaines importants (statistique publique, visualisation des données ...) et en spécifiant davantage les méthodes statistiques auxquelles les animations sont attachées. Mais le travail est déjà prodigieux ainsi. Quant aux textes présentés dans la troisième partie du site, Bernard Ycart en répertorie 129.

Pour un statisticien qui défend depuis près de cinquante ans l'importance d'accompagner l'enseignement d'un concept de quelques considérations d'ordre historique qui le font vivre et revivre, ce site est un trésor. Il faut le partager avec tous ceux qui pensent de même. Et quand c'est fait avec érudition et pédagogie, le plaisir de la consultation n'en est que plus grand. Courez-y vite et picorez ce qui vous intéresse, même si cela prend un peu de temps pour y arriver. Le jeu en vaut la chandelle !



[Jean-Jacques Droesbeke, Université libre de Bruxelles, Belgique](#)

[Article paru dans *Statistique et Société*, vol. 7, n° 1, [www.statistique-et-societe.fr](http://www.statistique-et-societe.fr), © Société française de Statistique (SFdS), juin 2019.]

---

## Démographie des fourmis

Encore des petites bêtes, mais bien visibles par rapport aux bactéries... Les fourmis nous impressionnent toujours par leur organisation, leur entêtement, et surtout par leur nombre. On en recense environ 12 000 espèces dans le monde, contre une seule espèce humaine. Combien d'individus cela fait-il? Un entomologiste britannique téméraire, C. B. Williams (1889-1981), a mené de savants calculs avant d'aboutir à un résultat qui donne le tournis : il estime que la Terre est peuplée de 10 millions de milliards de fourmis, à comparer aux 7,5 milliards d'êtres humains. Heureusement qu'elles sont minuscules! Pourtant, chacune pesant de 1 à 5 mg, la masse totale de leur population serait comparable à celle de tous les humains!

[Tiré de *l'Almanach Vermot 2019*, Éditions Hachette.]

## Statistique et vie :

### À la santé du tabac!

L'État finance les campagnes antitabac et... le train de vie des buralistes! C'est le constat dressé par la Cour des comptes dans son rapport de 2017. Chaque année, 300 millions d'euros d'argent public sont (re)distribués aux vendeurs de cigarettes pour financer leurs retraites complémentaires, la sécurisation des locaux et les embauches en contrats aidés\*. Pourtant, d'après les Sages\*\*, « les revenus des débiteurs continuent d'augmenter et le nombre d'établissements en difficulté apparaît limité ». Mieux, le chiffre d'affaires moyen des buralistes a explosé : il est passé de 447 238 euros en 2002 à 723 778 euros en 2015, soit 60 % de hausse! La ponction sur les deniers publics est d'autant plus toxique que les campagnes de sensibilisation antitabac ne sont pas toujours efficaces : un Français sur trois continue de fumer. Mais il est vrai que l'État dépense en moyenne cent fois plus d'argent pour la prévention routière que contre le tabagisme, alors que ce dernier fait chaque année « vingt-trois fois plus de morts » que les accidents de la route.

\* Un contrat aidé, ou emploi aidé, est un contrat de travail pour lequel l'employeur reçoit une aide financière qui réduit le coût du travail.

\*\* Le Conseil constitutionnel ou « Sages » est une institution française créée par la Constitution de la Cinquième République du 4 octobre 1958. Il se prononce sur la conformité à la Constitution des lois et de certains règlements dont il est saisi. Il veille à la régularité des élections nationales et référendums.

[Tiré de l'*Almanach Vermot 2019*, Éditions Hachette.]

### Maladies... imaginaires

Quel jour les arrêts maladie sont-ils le plus nombreux? Vous l'avez deviné, c'est le lundi! Une étude menée par une grande mutuelle française a démontré que non seulement leur nombre est en hausse depuis 7 ans, mais qu'en plus, 28 % des arrêts de travail survenaient le lundi. Et dans le cas des arrêts de 24 heures et moins, c'est le vendredi qui est le plus répandu. Que voulez-vous, on peut être malade et bien organiser son temps!

[Tiré de l'*Almanach Vermot 2019*, Éditions Hachette.]

### Slogan inégal

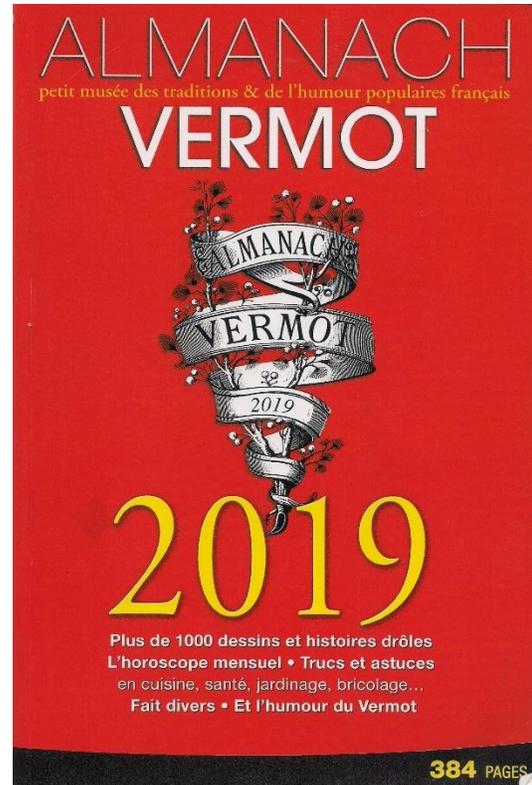
« Nous sommes les 99 % » est le slogan signature du mouvement Occupy Wall Street, né aux États-Unis en 2011. Cette formule fait référence aux inégalités économiques : 1 % de la population mondiale détiendrait autant de richesses que les 99 autres pour cent combinés. C'est donc parce qu'il montre à la fois l'importance de ces inégalités et la puissance potentielle du mouvement que ce slogan est devenu emblématique de la lutte contre la corruption et la répartition inégale des richesses.

[Tiré de l'*Almanach Vermot 2019*, Éditions Hachette.]

## Avez-vous vos 8 heures 15 quotidiennes?

Nous passons en moyenne 118 jours par an dans les bras de Morphée, ce qui représente... un tiers de notre vie! Mais plus on vieillit, moins on dort : les nouveau-nés roupillent 12 à 18 heures par jour; les 3 à 5 ans entre 11 et 13 heures; les 5 à 12 ans, 10 à 11 heures; les ados, 8,5 à 10 heures. Quant aux adultes, s'ils ont besoin grosso modo de 8 h 15 de sommeil (de 7,5 à 9 heures selon les profils), ils dorment en moyenne 7 h 45 par nuit. Il y a pire : depuis les années 1990, la durée moyenne du sommeil des adolescents s'est écourtée de 50 minutes et celle des adultes de 18 minutes. La faute à notre mode de vie stressé et ultra-connecté! Il est temps de tout débrancher et d'aller se coucher!

[Tiré de l'*Almanach Vermot 2019*, Éditions Hachette.]



## Crimes majeurs de mineurs...

Tous les ans, en France, en moyenne 57 mineurs âgés de moins de 15 ans sont victimes d'homicide, selon les chiffres de l'Observatoire national de la délinquance et des réponses pénales. Dans un tiers des cas, les enfants ont moins d'un an. Dans 72 % des cas, les victimes ont un lien familial avec leur bourreau, ce qui n'est pas le cas lors de crimes d'adultes. L'étude souligne également que « les personnes condamnées pour homicide sur mineurs de moins de 15 ans sont majoritairement des femmes ». C'est la triste réalité pour 227 des condamnations prononcées entre 1996 et 2015, soit dans environ 70 % des affaires présentées devant les tribunaux. Et ce sont souvent des mères qui tuent leur jeune enfant. Si tous les milieux sociaux sont touchés, l'isolement affectif et social est une constante de ces affaires. L'Observatoire rappelle néanmoins que le nombre d'homicides sur les enfants a tendance à baisser : une diminution qui s'expliquerait par la meilleure prise en charge psychologique et médicale des femmes pendant et après la grossesse. D'une manière générale, rappelons que, contrairement à une idée reçue, les homicides volontaires ont diminué en France ces vingt-cinq dernières années. À Paris, on est passé de 256 homicides en 1994 à 90 en 2013, soit une baisse de 65 %.

[Tiré de l'*Almanach Vermot 2019*, Éditions Hachette.]

## À propos de l'ASSQ

### **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.

*Member 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.

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

*Member 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 : Louis-Paul Rivest (Université Laval)

Vice-présidente : Valérie Roy (Institut de la statistique du Québec)

Secrétaire : Audrey-Anne Vallée (Université Laval)

Trésorière : Laurence Desbois-Bédard (Retraite Québec)

Registraire : Éric Lacroix (SOM Inc.)

Directrice des communications : Roxanne Brousseau (Retraite Québec)

Représentant étudiant : Jasmin Ricard (Université Laval)

## Membres institutionnels



### CONVERGENCE

*Convergence* Vol XXIII No 2 Septembre 2018

*Convergence*, le journal de l'Association des statisticiens et statisticiens du Québec (ASSQ), est publié deux ou trois fois l'an. Il est distribué gratuitement aux membres de l'ASSQ.

Rédacteur en chef :

Pierre Lavallée (plavall1962@gmail.com)

Rédacteurs adjoints :

Steve Méthot, Agriculture et agroalimentaire Canada

Myrto Mondor, Unité de recherche en santé des populations, CHAUQ

**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). 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. 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 est gratuite) :

<b>Tarifs</b>	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 :

**Pour nous contacter :**

Association des statisticiens et statisticiens du Québec

3340, rue de La Pérade, 3e étage

Québec (Québec), G1X 2L7

Courriel : [assq@association-assq.qc.ca](mailto:assq@association-assq.qc.ca)

Page Internet : <http://www.association-assq.qc.ca>