

# Convergence

Volume XXX Numéro 1 Février 2025

# Mot du rédacteur en chef



Pour passer les longs mois d'hiver, il y a le temps de s'asseoir devant un bon feu, de s'abriller avec une douce couverture, de boire un bon vin ou un chocolat chaud, selon vos goûts, et de savourer le dernier *Convergence*. Fidèle à mes habitudes, et maintenant aidé de Josiane Guedem, je vous ai concocté un numéro de *Convergence* diversifié et instructif.

Bien sûr, il y a les incontournables « Mot du rédacteur en chef » (que, en passant, vous êtes en train de lire...) et « Mot du président ». Jessica Bélisle de l'Institut de la statistique du Québec (ISQ) nous a préparé un texte soulignant les 25 ans de l'ISQ : un article agrémenté de statistiques intéressantes. De son côté, Statistique Canada nous

parle d'une vaste initiative de transition vers les sources libres, plus particulièrement en ce qui

concerne la transition des systèmes généralisés de Statistique Canada. Ce numéro de Convergence contient aussi un article intitulé « Le pari perdant sur sa propre mort » qui traite de la pension du Régime de rentes du Québec (bon, vous qui êtes jeunes, ça ne vous touche peut-être pas, mais moi si, bizarrement...). Dans ce numéro, vous en saurez un peu plus sur les

Sumériens et la mesure du temps, ainsi que sur un certain Thomas Fuller qui fut un génie du calcul mental. Je me suis aussi permis de mettre un long texte sur les échecs, la musique et les paradoxes, tiré du



Mot du rédacteur en chef	1
Mot du président	2
Le pari perdant sur sa propre mort	3
Les échecs, la musique et les paradoxes	
	6
Arobase	8
Nouvelles de l'ISQ	9
Nouvelles de Statistique Canada	11
Thomas Fuller	15
Les Sumériens et la mesure du temps	15
Des corrélations complètement loufoques	
:	<b>17</b>
14 citations de Galilée	20
Statistiques diverses	21
À propos de l'ASSQ	22
Membres institutionnels	23

livre « Le tableau du Maître flamand » de Arturo Perez-Reverte. Il y aura aussi un article à la fois intéressant et divertissant nous présentant des corrélations complètement loufoques. Pour qui aiment Galilée, vous vous délecterez avec 14 citations de ce dernier. Finalement, comme toujours, ce numéro est agrémenté de dessins humoristiques, de pensées et de citations, notamment de Galilée.

Bonne lecture!

Pierre Lavallée

# Mot du président

Bonjour chers membres,

La campagne de renouvellement est bien enclenchée. Si vous n'avez pas déjà renouvelé votre adhésion, ne tardez pas à le faire! Comme je l'ai mentionné dans mes messages précédents, et comme il a été discuté en AGA l'an dernier, les mésaventures liées au site web ont épuisé les réserves que nous avions depuis quelques années, d'où la décision de rétablir une cotisation. Tout de même, après quelques années de congé de cotisation, le renouvellement représente encore une somme modique.

Comme vous le savez tous, nous sommes passés bien près d'avoir des élections l'automne dernier. Les bulletins de vote électroniques étaient même prêts à être envoyés!



Audrey-Anne Vallée nous a alors annoncé qu'elle s'apprêtait à prendre un congé de maternité en 2025 et qu'elle préférait retirer sa candidature pour le poste de secrétaire plutôt que de s'absenter pour une partie substantielle de son mandat. Félicitations à Audrey-Anne pour cette excellente nouvelle, et merci pour tes années de service au CA! Josiane Guedem, une statisticienne établie à Montréal depuis plus d'une quinzaine d'années et qui a surtout œuvré dans le domaine bancaire et financier, a donc été élue par acclamation. Elle est entrée en poste le 1<sup>er</sup> janvier et elle était donc bien aux commandes lors de notre rencontre du CA de janvier. Bienvenue au CA, Josiane!

Parlant du CA, le poste de représentant étudiant est toujours vacant. Je profite donc de ce message pour inviter les membres étudiants qui lisent *Convergence* à se manifester s'ils désirent contribuer à l'ASSQ. L'an dernier, notre représentante étudiante avait notamment facilité l'organisation d'événements conjoints avec le Comité d'Avancement de la Statistique à l'Université Laval (CASUL). Espérons que nous pourrons poursuivre cette collaboration dans les prochains mois.

Pour ce qui est des activités, l'automne a finalement été plus calme que prévu, mais nous espérons toujours organiser un ou deux événements du type « jeudi de l'ASSQ » au printemps. Heureusement, l'organisation du prochain colloque va bien. Le 15 mai prochain, vous êtes conviés au COPL de l'Université Laval où se tiendra notre Colloque 2025. Il s'agit du même lieu où s'est tenu le Colloque 2024. La même formule que l'an dernier prévaudra également, à savoir trois présentations scientifiques en présentiel, l'AGA en mode hybride, puis un vins et fromages pour poursuivre la discussion. L'expérience avait été fort agréable l'an dernier. Par ailleurs, la partie scientifique sera également publicisée comme une activité conjointe avec le CIMMUL. Le CIMMUL est un groupe de recherche en mathématiques comptant quelques membres dans l'ASSQ. Faire de la partie scientifique un événement conjoint permet de donner de la visibilité à la statistique dans le contexte d'un centre où la statistique n'est qu'une discipline des mathématiques parmi d'autres.

Au plaisir de vous voir sous peu!

Jean-François Plante

# Le pari perdant sur sa propre mort

L'année où on choisit de réclamer sa rente du Régime de rentes du Québec ressemble à une gageure.

Année après année, des dizaines de milliers de Québécois parient sur l'âge de leur décès, et une majorité d'entre eux misent sur un scénario improbable dont les gains potentiels n'ont rien d'extraordinaire. Des perdants.

C'est la conclusion à laquelle arrivent trois enseignants au département des sciences comptables de l'Université du Québec à Trois-Rivières (UQTR). Le trio a publié ses résultats dans un cahier de recherche au titre fort inspiré, *Le pari du perdant ravi*.

De quoi est-il question, au juste? Eh oui! De l'âge auquel les Québécois demandent leur prestation du Régime de rentes du Québec (RRQ). Deux professeurs (Marc Bachand et Nicolas Lemelin) et un chargé de cours de l'UQTR (Nicolas Monette) proposent un angle d'approche auquel je n'avais pas songé, le coup de poker!

L'année où on choisit de réclamer sa rente du RRQ ressemble en effet à une gageure. Quand je la prends dès que possible, c'est miser sur son décès précoce. À l'inverse, quand j'attends le plus tard, c'est croire en mes chances de vivre au-delà d'un certain âge.

Pour ceux qui auraient raté mes chroniques précédentes sur le sujet, il faut rappeler qu'un participant au Régime de rentes du Québec est admissible à sa pleine allocation de retraite à 65 ans. Il peut la demander dès 60 ans, mais les mensualités sont alors amputées de 36 % (0,6 % par mois d'anticipation). Il peut aussi la différer jusqu'à 72 ans, ce qui lui assure une prestation améliorée de 58,8 % jusqu'au décès (0,7 % par mois de report). Lorsqu'on la repousse, on démarre avec un retard, mais la bonification permet de rattraper ce déficit avec les années, à moins de rencontrer la faucheuse entre-temps.

Le parallèle utilisé par les universitaires trifluviens est assez éloquent, mais le plus frappant, c'est lorsqu'on prend conscience des probabilités de gagner et des gains maximaux potentiels. Ici, on présente deux options: lorsqu'on choisit de toucher sa rente à 60 ans plutôt qu'à 65 ans; et à 65 ans plutôt qu'à 72 ans.

Voici tout de suite la conclusion: avec les mêmes chances et les mêmes gains espérés, personne ne parierait dans une machine à sous ou à la roulette, comme le font pourtant des milliers de Québécois avec leurs rentes de retraite.

### Moins d'une chance sur 10 de gagner

Dans le cas d'un homme qui réclame sa rente à 60 ans plutôt qu'à 65 ans, les probabilités qu'il sorte gagnant s'élèvent à 9 % seulement. Pour femme, ce choix serait plus audacieux encore, avec 7 % de chances de gagner. (Les hypothèses et les formules retenues sont simples, des calculs plus élaborés ne produiraient pas des



résultats radicalement différents.)

Au casino, moins les possibilités sont de notre côté, plus l'enjeu est important. C'est moins évident ici.

Le gain le plus élevé qu'un rentier du RRQ peut espérer en demandant sa rente à 60 ans au lieu de 65 ans monte à 52 000 \$ (de 2024, dans le cas de participant ayant cotisé le maximum). Pour remporter ce « jack pot », le bénéficiaire doit toutefois mourir dans une fenêtre plutôt restreinte, juste avant son 65e anniversaire. Le gain décroît ensuite de mois en mois, jusqu'à ce que le retraité arrive à 73,9 ans, bien vigoureux. Il bascule alors du côté des perdants, un scénario dont la probabilité s'élève à 91 % pour un homme et à 93 % pour une femme.

À chaque mois supplémentaire que le retraité poursuit sa vie, la perte découlant de son choix se creuse. S'il atteint l'âge vénérable de 100 ans, le déficit totalisera 160 000 \$ (de 2024).

Statistique amusante avancée par les auteurs : les probabilités de vivre jusqu'à 100 ans sont plus élevées que de mourir la veille de son 65e anniversaire. Autrement dit, une personne en bonne santé qui choisit de prendre sa prestation du RRQ à 60 ans court plus de risques de laisser 160 000 \$ sur la table qu'il a de chances de s'éteindre à 64 ans et 11 mois, plus riche de 52 000 \$.

Et pourtant, une majorité de Québécois penchent encore pour la rente à 60 ans.

Les auteurs de l'étude tirent les mêmes conclusions à l'égard de l'autre option, soit la prise de la rente à 65 ans plutôt qu'à 72 ans. Une petite minorité de participants au RRQ demandent leur rente après 65 ans, un nombre infinitésimal patiente jusqu'à la limite. Pourtant, un homme en bonne santé a presque trois chances sur quatre de sortir gagnant en encaissant sa pension à partir de 72 ans, et pour une femme, c'est mieux encore: 80 %. Le point de bascule se situe alors à 83,9 ans, c'est l'âge où celui qui a attendu à 72 ans rattrape l'autre qui touche sa prestation depuis son 65e anniversaire.

#### La faute de biais cognitifs

Deux situations justifient le choix d'une rente précoce: une piètre santé et l'absence d'autres sources de revenus pour compenser l'argent auquel on renonce durant la période de report. Alors, faut-il conclure que les autres sont stupides?

Non, affirment les auteurs de l'étude. Comme c'est souvent le cas lors de la prise d'une décision financière, un ensemble de biais cognitifs viennent brouiller les cartes.

Je retiens du rapport quelques explications parmi d'autres. D'abord, les participants n'ont pas conscience qu'il s'agit d'un pari et ils ignorent les probabilités qui jouent contre eux.

Nous sommes tous plus ou moins affectés par l'aversion à la perte, un biais cognitif largement documenté.

La douleur ressentie par une perte est toujours plus forte que le plaisir engendré par un gain. La crainte de laisser de l'argent sur la table pèse beaucoup plus sur nos décisions que la perspective de réaliser un profit.

Le monde accorde également plus d'importance aux événements négatifs. Quand une célébrité décédée prématurément fait les manchettes, ça nous frappe davantage que les 100 ans de Jeanette Bertrand.

Les trois profs recommandent à Retraite Québec d'indiquer sur les avis envoyés aux participants les probabilités de se sortir gagnant pour chacune des options.

Gagnants? Ça dépend pour qui, car tous les « perdants ravis » font diminuer le coût du régime pour la minorité de rentiers qui patientent et qui gagnent.

#### Daniel Germain, Les Coops de l'information

[Article paru dans le quotidien Le Droit du 5 décembre 2024.]

# Les échecs, la musique et les paradoxes

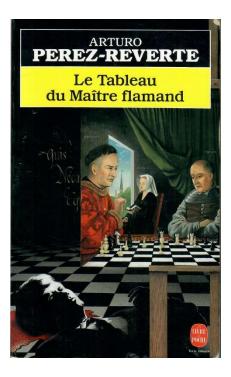
- Vous jouez? demanda Muñoz d'un air détaché.
- —Autrefois. Pratiquement plus maintenant... Mais à vrai dire, je n'ai jamais pensé que cette partie pouvait se jouer à l'envers... pensif, il tapotait ses genoux. Jouer à l'envers... C'est amusant! Vous savez que Bach adorait les inversions musicales? Dans certains canons, il renverse le thème, si bien que la mélodie descend chaque fois que le thème original monte... L'effet est peut-être un peu bizarre, mais quand on s'habitue, on finit par le trouver tout à fait naturel. Il y a même un canon dans l'Offrande qui s'exécute à l'envers de ce qui est écrit il regarda Julia. Je crois vous avoir déjà dit que Jean-Sébastien était un fieffé farceur. Son œuvre est remplie de pièges. Comme si, de temps en temps, une note, une modulation ou un silence vous disait: « Je renferme un message; découvrez-le. »
- —Comme le tableau, dit Muñoz.
- Oui. À cette différence près que la musique ne se résume pas à des images, à des dispositions de pièces ou, dans le cas particulier, à des vibrations dans l'air, mais qu'elle réside dans les émotions que ces vibrations produisent dans le cerveau de chacun... Vous rencontreriez de sérieuses difficultés si vous tentiez d'appliquer à la musique les méthodes que vous avez employées pour résoudre l'énigme de la partie d'échecs du tableau... Il vous faudrait découvrir la note qui contient les effets affectifs en question. Ou plutôt, la combinaison de notes... Vous ne croyez pas que c'est infiniment plus difficile que de jouer aux échecs?

Munoz réfléchit longuement.

— Je crois que non, dit-il enfin. Les règles générales de la logique sont les mêmes dans tous les domaines. La musique, comme les échecs, obéit à des règles. Il s'agit simplement de se mettre au travail jusqu'à isoler un symbole, une clé — un léger rictus tordit la moitié de sa bouche. Comme la pierre de Rosette des égyptologues. Une fois que vous l'avez, ce n'est plus qu'une question de travail, de méthode. Et de temps.

Belmonte battait des paupières, moqueur.

- Vous croyez? ... Vous pensez vraiment que tous les messages secrets peuvent se déchiffrer? ... Qu'il est toujours possible de trouver une solution exacte en appliquant un système?
- J'en suis sûr. Parce qu'il existe un système universel, des lois générales qui permettent de démontrer ce qui peut l'être et d'écarter ce qui doit l'être.



Le vieillard paraissait sceptique.

— Je ne suis pas du tout de votre avis, pardonnez-moi. Je pense au contraire que toutes les divisions, classifications, distributions et règles que nous attribuons à l'univers sont fictives, arbitraires... Il n'y en a pas une seule qui ne renferme sa propre contradiction. C'est un vieil homme qui vous le dit, un vieil homme qui a beaucoup vécu.

Muñoz changea de position dans son fauteuil et regarda autour de lui. Il ne semblait pas très heureux du tour que prenait la conversation, mais Julia eut l'impression qu'il ne souhaitait pas non plus changer de sujet. Elle savait que cet homme ne parlait pas pour ne rien dire et qu'il devait donc avoir une idée derrière la tête. Peut-être Belmonte devinait-il quelque chose dans les pièces que le joueur d'échecs étudiait pour résoudre le mystère.

- —C'est discutable, dit finalement Muñoz. L'univers est rempli d'infinis démontrables. Les nombres premiers, les combinaisons aux échecs...
- Vous le croyez vraiment?... Que tout est démontrable? Permettez au musicien que j'ai été le vieillard montra ses jambes invalides avec une expression de mépris tranquille —, ou que je continue d'être malgré tout, permettez-moi de vous dire que tout système est incomplet. Que la démontrabilité est un concept de bien moindre valeur que la vérité.
- La vérité est comme le meilleur coup aux échecs: elle existe, mais il faut la chercher. Avec le temps, elle est toujours démontrable.

Belmonte sourit d'un air malicieux.

—Je dirais plutôt que ce coup parfait, que vous l'appeliez ainsi ou que vous l'appeliez tout simplement la vérité, existe peut-être. Mais qu'on ne peut pas toujours le démontrer. Et que tout système qui tente de le faire est limité et relatif. Envoyez mon [tableau de] Van Huys sur Mars ou sur la planète X, pour voir si quelqu'un là-bas est capable de résoudre le problème. Je dirais même plus : envoyez-lui ce disque que vous écoutez en ce moment. Tant qu'à faire, envoyez le disque cassé. Que signifie alors ce qu'il contient?... Et puisque vous paraissez amateur de règles exactes, je vous rappelle que la somme des angles d'un triangle est de cent quatre-vingts degrés en géométrie euclidienne, mais plus en géométrie elliptique et moins en géométrie hyperbolique... C'est qu'il n'y a pas de système unique, pas d'axiomes. Les systèmes sont dissemblables, même à l'intérieur du système ... Vous aimez résoudre les paradoxes? Il n'y a pas que la musique, la peinture et, j'imagine, les échecs, qui en soient remplis. Voyez plutôt — il tendit la main vers la table, prit un papier et un crayon, puis écrivit quelques lignes qu'il montra ensuite à Muñoz. Regardez-moi ceci, je vous prie.

Le joueur d'échecs lut à haute voix:

- La phrase que j'écris en ce moment est celle que vous lisez en ce moment... il regarda Belmonte, surpris. Oui, et puis?
- C'est tout. J'ai écrit cette phrase il y a une minute et demie et vous venez de la lire, il n'y a que quarante secondes. En d'autres termes, mon écriture et votre lecture correspondent à des moments différents. Mais sur le papier, ce moment et ce moment

sont indubitablement le même moment... Donc, la phrase est à la fois vraie et fausse... Ou est-ce le concept de temps que nous laissons de côté?... N'est-ce pas un bon exemple de paradoxe?... Je vois que vous n'avez rien à répondre, et il en va de même lorsqu'il s'agit du vrai fond des énigmes que peut poser mon Van Huys ou autre chose... Qui vous dit que votre solution du problème est correcte? Votre intuition et votre système? Bien. Et quel est le système supérieur qui vous permet de démontrer que votre intuition et votre système sont valides? Quel autre système confirme ces deux systèmes?... Vous jouez aux échecs. Ces vers vous intéresseront, je crois.

Et Belmonte se mit à réciter, avec de longs silences:

Prisonnier, le joueur l'est aussi

— la sentence est d'Omar — d'un autre échiquier de noires nuits et de blanches journées.

Dieu déplace le joueur, et celui-ci la pièce.

Quel Dieu derrière Dieu commence donc la trame de poussière et de temps, de rêve et d'agonies...?

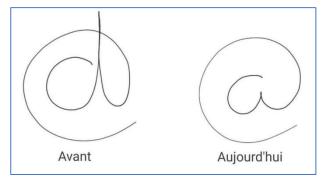
—Le monde est un immense paradoxe, conclut le vieil homme. Et je vous mets au défi de démontrer le contraire.

[Extrait de « Le tableau du Maître flamand » de Arturo Perez-Reverte, *Le livre de poche*, 1993.]

#### **Arobase**

L'origine de l'arobase provient du Moyen Âge, les moines copistes utilisaient le signe « @ » comme une abréviation du son « ad » en combinant les lettres « a » et « d ». Cela a évolué au fil du temps pour donner le symbole actuel « @ ». Plus tard, lors de l'invention des imprimeries, les caractères d'impression étaient disposés différemment. L'arobase

était placée dans la rangée inférieure, où les majuscules et caractères spéciaux étaient rangés. Les imprimeurs l'ont nommé « a-rond-bas » car elle était située en bas et avait une forme ronde autour du « a ». Cette appellation s'est progressivement transformée pour devenir « arobase ».



[Tiré de la page Facebook d'Olivier Morattel.]

# Nouvelles de l'Institut de la statistique du Québec

Pour cette édition, j'aimerais débuter en soulignant le 25<sup>e</sup> anniversaire de la création de l'Institut de la Statistique du Québec. L'Institut de la statistique du Québec (ISQ) est né le 1<sup>er</sup> avril 1999 de la fusion du Bureau de la statistique du Québec (BSQ), de l'Institut de recherche et d'information sur la rémunération, de Santé Québec et de la direction du ministère du Travail affectée à l'Enquête sur la rémunération globale. L'ancêtre de l'ISQ, le BSQ, a pour sa part vu le jour il y a plus de 100 ans! Il a été fondé en 1913 et était la toute première agence statistique au pays. (Source : 25 ans de l'ISQ)

Il va sans dire que la réalité a beaucoup changé depuis un quart de siècle, tant au niveau démographique qu'au niveau culturel. Il est intéressant de mentionner certains faits marquants :

- Salaire minimum : 6,90 \$ en 1999 vs 15,75 \$ le 1<sup>er</sup> mai 2024
- Nombre de personnes de 75 ans et plus : 374 500 en 1998 vs 809 848 en 2023 (données provenant des éditions du Québec chiffres en main de 1999 et de 2024)
- Films ayant généré le plus d'assistance en salle : *Titanic* (1998, 2,68 M\$) vs *Super Mario Bros, le film* (2023, 0,64 M\$)
- Films québécois ayant généré le plus d'assistance en salle : Les Boys (1998, 1,01 M) vs Le temps d'un été (2023, 0,22 M)
- Nombre d'albums vendus : 12,30 M en 2002 vs 1,36 M en 2023
- Album le plus vendu : A new day has come de Céline Dion (2002) vs Neo-Romance d'Alexandra Stréliski (2023)

Autre fait intéressant : le taux de chômage a beaucoup diminué en 25 ans. Il est passé de 10,3 % en 1998 à 4,5 % en 2023. En lien avec cette donnée, mentionnons que le poids démographique a subi des changements notoires depuis 1998. Effectivement, la part des 65 ans et plus a presque doublé, passant de 12,5 % à 20,8 %, alors que le poids démographique des moins de 20 ans est passé de 25,3 % à 20,8 %. Outre le vieillissement de la population, ce qui retient également l'attention est son accroissement, la population québécoise étant passée de 7 310 238 en 1999 à 8 984 918¹ en 2024. Ces deux éléments pourraient être à l'origine de certains défis vécus par l'industrie des soins de santé et de l'assistance sociale. Dans ce secteur, entre 2019 et 2023, il s'est ajouté environ 25 000 postes vacants, pour atteindre près de 43 000 en 2023. Il faut savoir qu'en parallèle, on note également une croissance d'environ 36 000 emplois salariés dans cette industrie, qui ont été pourvus. Vous pourrez retrouver ces données dans cette publication.

Pour poursuivre sur la thématique de la santé, l'espérance de vie de la population il y a 25 ans était de 77,9 ans alors qu'en 2023, elle était de 82,5 ans<sup>2</sup>, soit l'une des plus élevées au Canada et dans le monde. De plus, la part de jeunes du secondaire fumant la

\_

<sup>&</sup>lt;sup>1</sup> Donnée provisoire

<sup>&</sup>lt;sup>2</sup> Donnée provisoire

cigarette a chuté entre 1998 (33 %) et 2019 (9 %). Or, le vapotage gagne du terrain dans cette partie de la population, ce qui sera à surveiller dans les prochaines années. La consommation d'alcool et de drogue a aussi diminué chez les jeunes entre 1998 et 2019. D'autres statistiques sur la même thématique sont disponibles ici : Portrait du Québec : taux de chômage élevé, faible espérance de vie et importante proportion de jeunes qui fument la cigarette (quebec.ca)

Finalement, j'aimerais terminer par de nouveaux indicateurs pré et postpandémiques concernant les activités culturelles. Dans mon dernier article, j'avais abordé la fréquentation muséale. Cette fois-ci, j'ai plutôt été interpellée par le nombre de spectateurs dans les cinémas. L'achalandage est en hausse pour une troisième année consécutive, mais il n'est pas de retour au niveau prépandémique et bien loin de celui observé en 1998! En effet, en 1998, 26 millions de spectateurs avaient franchi les portes des cinémas au Québec, alors qu'en 2023, on a compté 14,6 millions de spectateurs. La hausse de 29 % par rapport à 2022 n'a pas permis de rejoindre le résultat observé en 2019, étant encore 22 % moins élevé que celui-ci. Le moins grand nombre de projections, la baisse du nombre de cinémas actifs entre 2019 et 2023 (96 à 93) et les changements dans les habitudes de consommation des films pourraient expliquer cette diminution. Toutefois, les films québécois ont su se démarquer. En effet, la part de marché des films québécois dans les entrées en salle s'établit à 12 % au Québec en 2023 (9 % en 2022 et 8 % en 2019). À noter que le nombre de nouveautés québécoises projetées en salle a augmenté de façon importante depuis 25 ans, passant de 19 en 1998 à 66 en 2023. Vous pourrez retrouver ces statistiques dans la publication La fréquentation des cinémas en 2023.

Pour avoir accès aux faits marquants cités ainsi qu'à d'autres informations intéressantes, vous pouvez vous rendre sur cette <u>page</u> créée pour les 25 ans de l'ISQ.

Jessica Bélisle



« Ce que l'on conçoit bien s'énonce clairement, et les mots pour le dire arrivent aisément. »

Nicolas Boileau (1636 – 1711)

# **Nouvelles de Statistique Canada**

Chers lecteurs de *Convergence*, dans cette édition, je partage avec vous un article préparé par mes collègues Darren Gray et Fritz Pierre de la Direction des méthodes statistiques modernes et de la science des données. Leur article traite d'une vaste initiative de transition vers les sources libres, plus particulièrement en ce qui concerne la transition des systèmes généralisés de Statistique Canada. Merci Darren et Fritz!

# François Brisebois, Statistique Canada

# De SAS vers les sources libres : conversion des systèmes généralisés de Statistique Canada

# Adopter les outils de source libre à Statistique Canada

Quiconque travaille aujourd'hui pour un organisme statistique doit être bien conscient de la prévalence croissante des langages de programmation de source libre, comme R et Python. Nous voyons leur application dans les statistiques, les méthodes d'enquête, la science des données et l'intelligence artificielle, pour ne citer que quelques domaines, avec de nouveaux outils qui apparaissent presque tous les jours. Ces langages sont de plus en plus utilisés dans le monde académique, et bon nombre de nos nouveaux employés les maîtrisent au moins à un certain degré. Comment un organisme comme Statistique Canada peut-il tirer parti des possibilités qu'offrent ces nouveaux outils?

Bien que R et Python soient disponibles à Statistique Canada, nos processus de production à grande échelle, c'est-à-dire ceux utilisés pour la plupart de nos statistiques officielles, reposent encore largement sur une infrastructure SAS. Statistique Canada a toutefois entrepris une vaste initiative de transition vers les sources libres d'ici 2028, qui sera notamment simplifiée par le fait que les processus de production utilisent des systèmes généralisés.

#### Quels sont les systèmes généralisés ?

Au fil des ans, Statistique Canada a consacré d'importantes ressources à l'élaboration d'un ensemble d'outils statistiques communs appelés systèmes généralisés remplissant des fonctions statistiques clés. Ils sont développés et maintenus par l'agence et ils sont entièrement pris en charge par des équipes de méthodologistes et d'informaticiens. Un certain nombre de ces systèmes, dont certains remontent aux années 1980, ont connu plusieurs itérations dans différents langages de programmation. Les systèmes suivants sont actuellement programmés en SAS :

• *G-Coup* : Couplage d'enregistrements probabiliste

• *G-Éch* : Échantillonnage

• *Banff* : Vérification et imputation

• *G-Est* : Estimation

*G-Séries*: Séries chronologiques *G-Confid*: Contrôle de la divulgation

À Statistique Canada, l'utilisation de systèmes généralisés permet de gagner du temps, de réduire les risques d'erreurs de traitement, de faciliter le transfert des connaissances, d'assurer l'uniformité des méthodes statistiques dans l'ensemble des projets et de favoriser l'utilisation de méthodes approuvées, bien établies scientifiquement et modernes. Les systèmes sont également offerts gratuitement aux utilisateurs externes, tels que les organismes statistiques canadiens et internationaux, dont l'Institut de la statistique du Québec (ISQ). La conversion rapide des systèmes généralisés vers R et Python aidera les programmes et utilisateurs dans leur transition d'ici 2028.

## Où en sommes-nous aujourd'hui?

Le travail a progressé rapidement au cours des dernières années, en partie parce que certaines des équipes responsables avaient déjà exploré ou même débuté le processus de conversion vers R ou Python. *Banff* est offert sous forme de logiciel libre depuis le mois de janvier 2025, *G-Séries* le sera en mars et les autres devraient l'être au cours des deux prochaines années.

Voici quelques leçons apprises que nous pouvons transmettre à d'autres organismes qui font face à des défis similaires.

1. Évaluation des progiciels existants : Compte tenu du grand nombre de progiciels statistiques existants, y compris ceux que l'on trouve sur la liste impressionnante des logiciels statistiques (une liste d'outils statistiques de source libre élaborés et utilisés par des organismes statistiques nationaux), il doit sûrement y avoir des outils existants qui pourraient remplacer convenablement les systèmes généralisés ? Eh bien, oui et non. Nous avons examiné un certain nombre de ces progiciels et constaté que, même dans les cas où les progiciels appliquaient les mêmes méthodes statistiques, les différences dans leur mise en œuvre signifiaient souvent qu'ils produisaient des résultats différents ou qu'ils ne pouvaient pas régler efficacement les problèmes les plus complexes et à plus grande échelle que nous rencontrons dans la production des statistiques officielles canadiennes; même lorsqu'ils donnent les mêmes résultats, ils n'ont pas toujours la robustesse attendue à des fins de production. De plus, la dépendance à l'égard de logiciels libres développés et maintenus à l'extérieur de l'agence comporte ses propres risques. Compte tenu de ces risques et de la quantité de travail qu'il faudrait pour modifier ces progiciels afin d'atteindre nos objectifs de production, la décision a été prise de développer nous-mêmes des versions en source libre de nos systèmes

- généralisés. Cela nous permet de contrôler pleinement les systèmes, mais aussi de garantir une assistance et une maintenance à long terme pour nos utilisateurs.
- 2. Éviter une « conversion directe » : L'idée d'une « conversion directe » est simple : vous prenez du code existant et le traduisez, ligne par ligne ou fonction par fonction, d'un langage à l'autre. C'est logique, non ? Eh bien, la réalité est que R et Python sont très différents de SAS, et que les approches qui fonctionnent bien dans un langage fonctionnent rarement aussi bien ou parfois pas du tout dans un autre. Pour chaque système, nous avons plutôt cherché des solutions qui tirent pleinement parti de leurs forces relatives, et nous avons essayé de coder nos fonctions statistiques d'une manière qui semble « naturelle » pour chaque langage. Cela prend plus de temps, mais permet d'obtenir un outil plus robuste, plus efficace et plus facile à maintenir et à améliorer dans le futur.
- 3. **Collaboration**: C'est l'un de ces concepts qu'il est facile de prêcher, mais qui demande des efforts pour être mis en pratique. Dans le cas des systèmes généralisés, la collaboration a été multidimensionnelle. La conversion de *Banff* et *G-Séries* de SAS vers Python n'aurait pas été possible sans une bonne collaboration entre les méthodologistes et les programmeurs du projet, facilitée par des outils collaboratifs comme GitLab. La communication entre les équipes a également été exemplaire. Par exemple, les équipes développant *G-Éch, G-Est* et *G-Séries* ont travaillé ensemble pour évaluer divers optimisateurs dans R, ce qui a permis d'économiser d'innombrables heures d'évaluation et de tests. Le groupe d'utilisateurs de R et de Python (GURP) de Statistique Canada a été une ressource essentielle, tout comme les consultations externes avec des organismes comme le Bureau central de la statistique (Pays-Bas) et l'Institut national de la statistique et des études économiques (France) qui investissent déjà massivement dans l'utilisation d'outils de source libre dans leurs programmes.
- 4. Intrépidité! C'est peut-être la leçon la plus importante. Lorsque vous entreprenez une tâche aussi immense, vous devez être prêt à essayer de nouvelles choses, presque quotidiennement, jusqu'à ce que vous trouviez quelque chose qui fonctionne. La plupart d'entre nous qui travaillons sur les systèmes généralisés avons consacré des années, voire des décennies, à perfectionner nos compétences en programmation dans SAS et à apporter des améliorations progressives à nos systèmes. Réaliser la même chose dans de nouveaux langages de programmation, en seulement quelques années, est sans aucun doute intimidant, et nécessite un esprit d'aventure.

Nous n'avons pas encore tout compris. Avec les deux nouveaux systèmes généralisés en sources libres (*Banff* et *G-Séries*) actuellement offerts, nous verrons sous peu comment les utilisateurs réagiront. Le support continu présente également de nouveaux défis. Bien que nous ayons délibérément choisi de minimiser le nombre de programmes utilitaires

insérés dans les systèmes généralisés, certains sont inévitables et doivent être mis à jour très fréquemment.

#### Réflexions finales

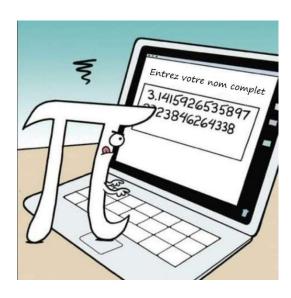
Dans l'ensemble, les systèmes généralisés sont sur la bonne voie d'être offerts en tant que logiciels libres d'ici la fin de 2026, ouvrant la voie à d'autres systèmes de production. Au-delà du simple passage d'un langage à l'autre, le processus de transition nous a obligés à repenser nos approches de programmation et à adopter des outils modernes. En cours de route, nous avons tiré de nombreuses leçons que nous pourrons transmettre à nos collègues de Statistique Canada et aux autres organismes. La conversion réussie des systèmes généralisés est également une étape clé de l'initiative de modernisation de l'agence, car nous adoptons non seulement de nouveaux langages de programmation (pour nous), mais nous embrassons aussi la communauté caractéristique et la culture de collaboration et de transparence de R et Python. C'est aussi une excellente occasion de mettre à profit les compétences de la prochaine génération de statisticiens, de méthodologistes, de scientifiques des données, d'analystes et de programmeurs qui se joindront à Statistique Canada.

Si vous utilisez actuellement l'un des systèmes généralisés de Statistique Canada, demeurez à l'affût des nouvelles concernant le lancement des prochains systèmes généralisés en sources libres et communiquez directement avec nous pour obtenir de plus amples renseignements!

Darren Gray (<u>darren.gray@statcan.gc.ca</u>) et Fritz Pierre (<u>fritz.pierre@statcan.gc.ca</u>)



Statistique Canada Statistics Canada



# **Chronique historique:**

# **Thomas Fuller**

Le 29 décembre 1790 marque la mort du célèbre mathématicien Thomas Fuller, connu sous le nom de « Mental Calculator ».

Mort le 29 décembre 1790, feu Thomas Fuller était un esclave africain connu pour ses compétences en mathématiques. Il a été capturé en Afrique par des esclaves blancs et expédié aux États-Unis en 1724 alors qu'il avait seulement 14 ans.

Il était si bon en mathématiques qu'il pouvait faire des calculs inimaginables. Un jour, quand ils lui ont demandé combien de secondes il y avait en un an et demi, il a répondu en environ deux minutes, 47 304 000. Les proabolitionnistes et philanthropes blancs ont utilisé son talent comme preuve que les esclaves noirs étaient égaux aux blancs en intelligence.



Thomas Fuller était un grand mathématicien, mais malheureusement oublié par l'Histoire.

[Tiré de la page Facebook « Black Power – all power to the people".]

# Les Sumériens et la mesure du temps

Les Sumériens ont regardé vers le ciel lorsqu'ils ont inventé le système du temps... et nous l'utilisons encore aujourd'hui.

On pourrait trouver curieux que l'on divise les heures en 60 minutes et les jours en 24 heures — pourquoi pas un multiple de 10 ou 12?

En termes simples, la réponse est que les inventeurs du temps n'ont pas fonctionné sur un système décimal (base 10) ou duodécimal (base 12), mais sur un système sexagésimal (base 60).



Pour les anciens innovateurs sumériens qui furent les premiers à diviser les mouvements en intervalles dénombrables, 60 était le nombre parfait. Le nombre 60 peut être divisé par 1, 2, 3, 4, 5, 6, 10, 12, 15, 20 et 30 parties égales.

De plus, les anciens astronomes croyaient qu'il y avait 360 jours dans une année, un nombre dans lequel 60 correspond parfaitement à six fois.

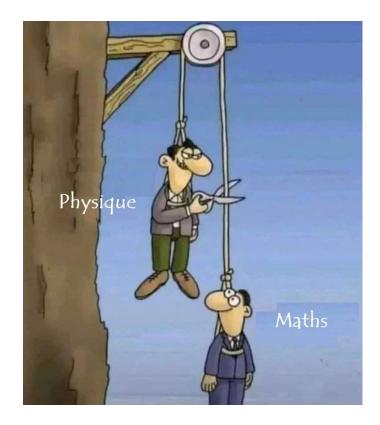
L'empire sumérien n'a pas duré. Cependant, depuis plus de 5000 ans, le monde est rattaché à cette délimitation du temps.

[Tiré de la page Facebook de Scriven Bisigo.]

« Tenez-vous loin des gens négatifs, ils ont un problème pour chaque solution »

Albert Einstein

Sans les mathématiques...

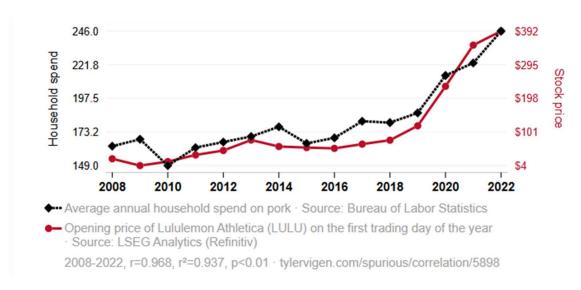


# Des corrélations complètement loufoques

Des ressemblances entre deux courbes statistiques peuvent parfois mener à de drôles de conclusions... C'est en tout cas ce que nous rappelle le site *Spurious Correlations*, qui met en évidence des corrélations absurdes et qui démontre que « courbes semblables » ne veut pas toujours dire « lien de cause à effet » entre deux phénomènes. En voici quelques exemples bien cocasses.

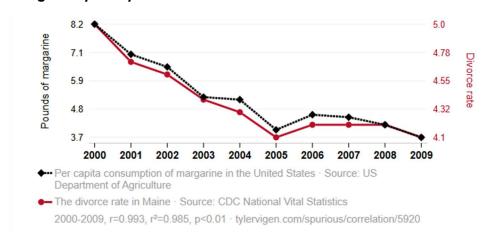
1. Quand la consommation de porc augmente, l'entreprise LuluLemon fait plus de profits et, par conséquent, son cours en bourse augmente.

Figure 1 : Corrélation entre la consommation de porc des ménages et les cours boursiers de l'entreprise LuluLemon



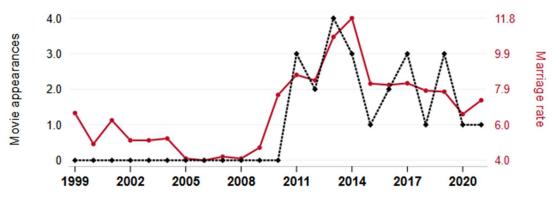
2. Quand les divorces diminuent; il se mange moins de margarine.

Figure 2 : Corrélation entre le taux de divorce au Maine et la consommation de margarine per capita.



3. Lorsque la filmographie (nombre d'apparitions dans les films) de l'actrice Elisabeth Olsen augmente, les mariages aussi.

Figure 3 : Corrélation entre la filmographie d'Elisabeth Olsen et le taux de mariage dans le District de Columbia (É.-U.)

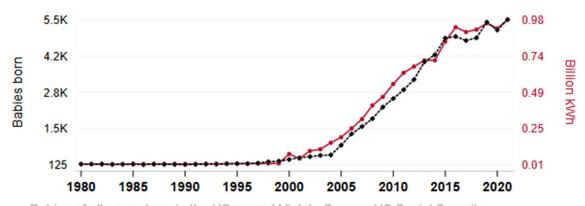


- •-- The number of movies Elizabeth Olsen appeared in · Source: The Movie DB
- The marriage rate in District of Columbia · Source: CDC National Vital Statistics

1999-2021, r=0.847, r<sup>2</sup>=0.718, p<0.01 · tylervigen.com/spurious/correlation/5873

4. Plus le prénom Violette devient populaire, plus l'utilisation des combustibles fossiles en augmente.

Figure 4 : Corrélation entre le nombre d'enfants nommées Violette à la naissance et l'utilisation des combustibles fossiles en Guinée équatoriale.

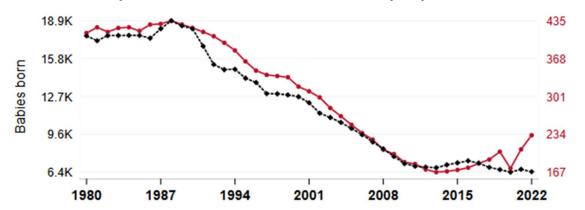


- Babies of all sexes born in the US named Violet · Source: US Social Security Administration
- Total fossil fuel use in Equatorial Guinea in billion kWh · Source: Energy Information Administration

1980-2021, r=0.995, r<sup>2</sup>=0.990, p<0.01 · tylervigen.com/spurious/correlation/3022

5. Plus le prénom Thomas devient populaire, plus la consommation de l'essence à la pompe augmente.

Figure 5: Corrélation entre le nombre d'enfants nommés Thomas à la naissance et la quantité d'essence consommée à la pompe en France



- Babies of all sexes born in the US named Thomas · Source: US Social Security Administration
- Volume of gasoline pumped consumed in France in millions of barrels per day
   Source: Energy Information Administration

1980-2022, r=0.986, r<sup>2</sup>=0.972, p<0.01 · tylervigen.com/spurious/correlation/5024

D'ailleurs, le fait de confondre corrélation et causalité porte un nom, il s'agit de l'effet cigogne, un phénomène peut-être plus répandu que vous ne pourriez le croire... Sur ce, n'oubliez pas : la corrélation n'implique pas la causalité! Et pour plus de corrélations loufoques, visitez le site « Spurious Correlations ».

# Économie 101

La ville médiévale de Visby et ses commerces se préparaient à l'approche des fêtes de fin d'année. Les taux directeurs stagnaient à près de zéro virgule zéro, encourageant les consommateurs à emprunter l'argent qu'ils n'avaient pas afin que, cette année encore, les achats de Noël battent les records précédents. En conséquence, les gens conservèrent leur emploi et, ainsi, les moyens de rembourser les prêts que les banques venaient de leur accorder. L'économie est une science subtile.

[Tiré de « L'assassin qui rêvait d'une place au paradis » de Jonas Jonasson, Presse de la Cité.]



# 14 Citations de Galilée

« La mathématique est une science dangereuse : elle dévoile les supercheries et les erreurs de calcul. »

« Et pourtant, elle tourne! »

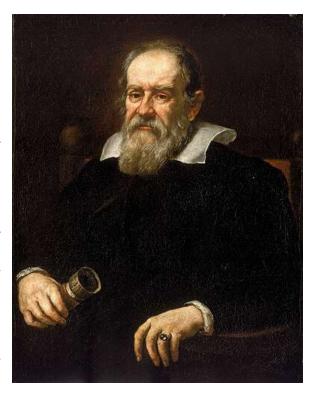
« On ne peut rien apprendre aux gens. On peut seulement les aider à découvrir qu'ils possèdent déjà en eux tout ce qui est à apprendre. »

« Le soleil, avec toutes ces planètes qui gravitent sous sa gouverne, prend encore le temps de mûrir une grappe de raisin, comme s'il n'y avait rien de plus important. »

« L'autorité d'un seul homme compétent, qui donne de bonnes raisons et des preuves certaines, vaut mieux que le consentement unanime de ceux qui n'y comprennent rien."

« Le doute est père de la création. »

« Je n'ai jamais rencontré d'homme si ignorant qu'il n'eut quelque chose à m'apprendre. »



- « Il est certainement nocif pour les âmes de transformer en hérésie le fait de croire ce qui est prouvé. »
- « Tu ne peux pas tout enseigner à un homme ; tu peux seulement l'aider à le trouver en lui. »
- « Le vin est de l'eau emplie de soleil. »
- « Le vin, c'est la lumière du soleil captive dans l'eau. »
- « On ne peut rien enseigner à autrui. On ne peut que l'aider à le découvrir lui-même. »
- « La philosophie est écrite dans ce grand livre qui s'étend chaque jour devant nos yeux : l'univers. Mais on ne peut le comprendre si nous n'apprenons d'abord son langage et si nous ne comprenons les symboles avec lesquels il est écrit. »
- « Les noms et les attributs doivent être donnés selon l'essence des choses, et non l'essence des noms, puisque les choses viennent en premier et les noms ensuite. »

# **Statistiques diverses**

#### Le nombre

109 est le nombre de jours consécutifs au cours desquels la température a dépassé 38 °C [100 °F] à Phoenix, en Arizona. Un nouveau record. Du 28 mai au 12 septembre 2024, la capitale de l'Arizona a rôti. Le mercure a même dépassé 45 °C [113 °F] durant 24 jours. En plein soleil, la température des rues et des trottoirs pouvait atteindre 82 °C [180 °F].

Sources: accuweather.com et The Atlantic

### La comparaison

En moyenne, combien font-ils de gestes par minute lorsqu'ils parlent?

— Un adulte italien: 40;

— Un adulte suédois: **20**.

Source: Panorama, traduit par Courrier international

#### En hausse

Si la tendance se maintient, on dénombrera environ 1400 hippopotames à travers l'Amazonie, en 2040. Les animaux sont les descendants de quatre hippopotames qui faisaient partie du Jardin zoologique personnel de Pablo Escobar, le chef du cartel de Medellín, en Colombie. En 1993, après la mort du narcotrafiquant, le domaine fut saccagé. Mais personne n'osa toucher les « hippos », jugés trop dangereux. Plusieurs ont ensuite pris la clé des champs. En 2008, on en dénombrait 24. L'an dernier, leur nombre était passé à 169. Sachant qu'une femelle hippopotame peut engendrer 25 petits « hippos » durant 50 années d'existence, nous vous laissons faire le calcul...

Source: The Guardian

#### Le sondage

51 % est la proportion des Noirs aux États-Unis qui estiment que le gouvernement américain fait la promotion de la contraception et de l'avortement afin de limiter la croissance de la population noire.

Source: Pew Research Center

Jean-Simon Gagné, Le Soleil

[Article tiré du quotidien *Le Soleil* du 16 septembre 2024.]

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.

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.

*Membre 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 : Jean-François Plante (HEC Montréal)

Président sortant : Louis-Paul Rivest (Université Laval)

Vice-président : Yohann Courtemanche (CHU de Québec / Université Laval)

Secrétaire : Josiane Guedem (Banque Laurentienne du Canada)

Trésorier : Jasmin Ricard (Centre de recherche CERVO)

Registraire: Johanne Morin (Beneva)

Directeur des communications : Simon Laflamme (Beneva)

Représentant étudiant : (poste vacant)

# **Membres institutionnels**





Département de mathématiques et de statistique











Statistique Canada

Statistics Canada



#### **CONVERGENCE**

Convergence Vol XXIX No 2, septembre 2024

Convergence, le journal de l'Association des statisticiennes 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 :

Josiane Guedem, Banque Laurentienne du Canada

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) :

 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 :

#### Pour nous contacter :

Association des statisticiennes et statisticiens du Québec Département de mathématiques et de statistique Université Laval, 1045 avenue de la Médecine

Québec (Québec) G1V 0A6

Courriel: assq@association-assq.qc.ca

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