



Mot du rédacteur en chef



Bonjour,

Au moment où j'écris ces lignes, il n'y a presque pas de neige au sol, on annonce une tempête de neige, mais possiblement suivie de pluie verglaçante. Bref, il y a de quoi être déprimé... Mais comme le dit mon grand copain Michel : « Quand on est déprimé, le meilleur remède est de faire des formules (mathématiques). » Bizarre, direz-vous? Peut-être pas tant que ça. En effet, les mathématiques constituent un monde cartésien où tout s'imbrique parfaitement, un monde en soi et où l'imperfection n'a pas sa place. C'est donc, d'une certaine façon, rassurant de se « promener »

dans ce monde à travers les formules mathématiques. Bon, d'accord, il n'est pas donné à tous d'être des virtuoses des équations mathématiques ou de trouver un quelconque réconfort à tracer des formules sur du papier. Heureusement cependant, pour les autres (soit la majorité des mathématiciens et statisticiens), il y a *Convergence*.

Lire *Convergence* est l'outil par excellence contre la déprime. Si, si... C'est comme faire des formules, mais les formules sont déjà faites. On n'a qu'à les lire à travers les articles tout aussi diversifiés qu'intéressants. Et c'est ce que je vous invite à faire dans le présent numéro de *Convergence*.

Dans ce numéro, vous trouverez la deuxième partie de l'excellent article d'Anne-Sophie Julien sur le « Portrait de l'emploi en statistique et en

Mot du rédacteur en chef	1
Mot du président	2
Portrait de l'emploi en statistique et en analyse de données au Québec	4
Comparaison des emplois en statistique selon leur titre	8
In memoriam : Robert Côté	12
Nouvelles de l'Institut de la statistique du Québec	14
Nouvelles de Statistique Canada	16
Pour en finir avec la signification statistique	19
Le nombre d'or	25
Les Québécois veulent accélérer sur les autoroutes	31
Recensement de la peinture	33
Aide recherchée	34
Le démon du jeu	35
À propos de l'ASSQ	36
Membres institutionnels	37

analyse de données au Québec. » ainsi qu'un autre article sur la « Comparaison des emplois en statistique selon leur titre. » Martin Provost de Statistique Canada nous parlera des dernières études de couverture du recensement de 2021. Jessica Bélisle de l'Institut de la statistique du Québec nous entretiendra sur certaines études menées à l'ISQ concernant des indicateurs pré et post-pandémie Christian Genest rendra hommage à Robert Côté, professeur retraité de l'Université Laval qui nous a malheureusement quitté en 2023. Il sera intéressant de lire l'article intitulé « Pour en finir avec la signification statistique. » Pour la chronique historique, je vous invite aussi à lire l'extrait du livre de Daniel Defays « Là où les sens se rencontrent – Harmonies » où l'auteur nous présente le nombre d'or. Vous trouverez sûrement un certain intérêt à savoir que « Les Québécois veulent accélérer sur les routes. » Et, bien entendu, ce numéro contiendra aussi des citations et dessins humoristiques. Finalement, je vous invite à porter attention à ma demande d'aide pour *Convergence*, un petit appel pour me donner un coup de main...

Bonne lecture!

Pierre Lavallée

Mot du président

Bonjour chers membres,

Dans le dernier *Convergence*, je vous offrais un résumé du Colloque 2023 de l'ASSQ. Le CA est en train d'organiser la prochaine mouture qui se tiendra un vendredi après-midi en mai prochain. La date exacte n'est pas confirmée à 100 %, mais devrait l'être sous peu. Nous planifions une rencontre sur le campus de l'Université Laval qui offrira pour cette année le double avantage d'être facilement accessible, et moins dispendieux. Une option responsable compte tenu de l'investissement important que nous avons dû consentir pour la reconstruction du site Web, et du congé de cotisation pour 2024 que nous avons précédemment voté en assemblée générale. Surveillez vos courriels au cours des prochaines semaines pour l'annonce officielle de l'événement.



Ceux d'entre vous qui ont tenté d'accéder au site Web ont sans doute eu un avant-goût du style qu'arborera notre portail lorsqu'il sera en ligne. Dans l'ombre, Simon Laflamme travaille sans relâche pour s'assurer que tout soit au point. Nous avons bien hâte que vous puissiez de nouveau avoir accès aux services électroniques de l'association. Comme je l'ai mentionné précédemment, la fin du site précédent a été précipitée par des enjeux technologiques de mise à jour. Renouveler le site Web était sur le radar, mais sa mise hors

service brutale était imprévue. Nous nous excusons de cette interruption de service, mais d'ici quelques semaines, nous pourrions tourner la page et apprivoiser notre nouvelle plateforme en ligne.

Le 8 novembre dernier, l'ASSQ a organisé un jeudi de l'ASSQ en présentiel en collaboration avec le CASUL. L'événement s'est tenu au pavillon Adrien-Pouliot de l'Université Laval et a permis à une quarantaine de participants de renouer en personne, en partageant vins et fromages. Pour être fidèle à la tradition des jeudis, une présentation de nature scientifique était à l'affiche. Anne-Sophie Julien nous a ainsi présenté la deuxième partie des résultats de l'enquête sur les emplois en statistique. Vous trouverez la version écrite de ces résultats dans cette édition de *Convergence*. Après la présentation, le CASUL a organisé une séance de « speed dating » professionnel entre les étudiants et les membres de l'ASSQ ayant une expérience de travail. Des conversations un à un étaient lancées, et à chaque 2 minutes, nous devions changer d'interlocuteur. Ce fut un excellent moyen de briser la glace et de permettre à plusieurs personnes d'échanger rapidement. Les discussions se sont ensuite poursuivies, un verre de vin (ou de jus de pomme) à la main. Les gens ont beaucoup apprécié l'événement, et un consensus a émergé voulant que la formule devrait être reprise annuellement. On se redonne sans doute rendez-vous l'automne prochain!

Dans un autre ordre d'idées, certains d'entre vous connaissent peut-être la Société statistique de Montréal (SSM). Par le passé, l'ASSQ a organisé conjointement des activités avec la SSM, et nous avons même décerné des prix communs. La structure de la SSM est beaucoup moins formelle que celle de l'ASSQ. Elle est techniquement à la fois une association régionale de la Société statistique du Canada et un chapitre de l'*American Statistical Association*. Historiquement, la SSM a connu des hauts et des bas, mais il semble acquis qu'elle s'apprête à connaître des changements majeurs au cours des prochains mois. Il est possible que ce renouveau conduise à des liens renforcés avec l'ASSQ, mais pas nécessairement non plus. Dans tous les cas, nous vous tiendrons au courant.

Statistiquement vôtre,

Jean-François Plante

« *Les seules choses certaines en ce monde, ce sont les coïncidences.* »

Leonardo Sciascia

Portrait de l'emploi en statistique et en analyse de données au Québec

PARTIE 2

L'enquête de l'ASSQ sur l'emploi en statistique et en analyse de données au Québec s'est déroulée du 7 mars au 14 avril 2023. Une première série de résultats provenant des données de 261 participants a été publiée dans l'édition précédente de *Convergence*, incluant la méthodologie de l'enquête. Dans le présent article, les résultats sur la satisfaction, les tâches effectuées dans le cadre du travail et les logiciels utilisés sont présentés.

Satisfaction

Le tableau 1 présente la distribution de la satisfaction envers trois caractéristiques de l'emploi, mesurée à l'aide d'une échelle de Likert de 1=très insatisfait à 10=très satisfait. La satisfaction est très bonne, mais un peu moins au niveau de la rémunération. La complexité des tâches (1=très simples, à 10=très complexes) et la charge de travail (1=pas du tout, à 10=trop élevée) sont considérées comme étant élevées, mais l'emploi demeure très stimulant (1=pas du tout, à 10=très stimulant).

Tableau 1 : Satisfaction

Variable	N	Moyenne	Écart-type	Minimum	25 ^e percentile	Médiane	75 ^e percentile	Maximum
Satisfaction – Tâches liées à votre emploi	252	8,07	1,56	1,00	7,00	8,00	9,00	10,00
Satisfaction – Rémunération globale	251	7,56	1,91	1,00	7,00	8,00	9,00	10,00
Satisfaction - conditions de travail	252	8,45	1,46	1,00	8,00	9,00	10,00	10,00
Complexité des tâches	250	7,26	1,56	1,00	6,00	8,00	8,00	10,00
Travail stimulant	251	7,82	1,74	1,00	7,00	8,00	9,00	10,00
Charge de travail	250	7,31	1,85	1,00	6,00	8,00	9,00	10,00

En croisant ces scores entre les différents groupes de l'échantillon, on remarque que la complexité des tâches, la charge de travail, mais aussi la stimulation au travail, augmentent avec l'expérience. Les tâches sont plus complexes pour les personnes avec un diplôme plus élevé. La satisfaction envers la rémunération et les autres conditions de travail est meilleure au privé qu'au gouvernement. Dans le milieu académique, les tâches sont plus complexes et la charge de travail est plus élevée.

Tâches

La médiane du nombre de tâches effectuées par les répondants dans le cadre de leur travail est de 8 (minimum=0, Q1=5, Q3=10, maximum=17). Sans surprise, le traitement de données, les analyses descriptives, la visualisation de données et les analyses statistiques sont les tâches les plus fréquentes, tel que le rapporte le tableau 2.

Tableau 2 : Tâches effectuées dans le cadre du travail (n = 251)

Tâches	Fréquence (%)
Création de banques de données	81 (32,3)
Gestion de banques de données	87 (34,7)
Nettoyage de banques de données	134 (53,4)
Traitement de données	198 (78,9)
Requêtes et extractions de données	152 (60,6)
Tableaux de bord	87 (34,7)
Analyses descriptives	199 (79,3)
Visualisation de données	177 (70,5)
Automatisation de processus	93 (37,1)
Simulation	73 (29,1)
Planification d'enquêtes, d'études, de sondages, d'expériences	105 (41,8)
Analyses statistiques	211 (84,1)
Rédaction de rapports ou d'articles scientifiques	139 (55,4)
Préparation/animation d'ateliers de formation	74 (29,5)
Consultation avec clients	119 (47,4)
Gestion d'employés	59 (23,5)
Gestion financière	32 (12,7)

Analyses statistiques

La médiane du nombre d'analyses statistiques distinctes effectuées est de 5 (minimum=0, Q1=3, Q3=7, maximum=11). Les modèles linéaires et non linéaires, ainsi que les tableaux de fréquences, font partie des analyses les plus fréquentes, telles que présentées au tableau 3. Les équations structurelles ainsi que les séries chronologiques sont les analyses les moins utilisées par les répondants.

Tableau 3 : Analyses statistiques effectuées (n = 240)

Analyses	Fréquence (%)
Calcul de taille d'échantillon	125 (52,1)
Analyse de données d'enquête et de sondage	117 (48,8)
Développement et validation de questionnaires et d'instruments	73 (30,4)
Équations structurelles, analyses factorielles	54 (22,5)
Modèles linéaires et non linéaires	187 (77,9)
Analyses multivariées	147 (61,3)
Tableaux de fréquences	184 (76,7)
Classification	111 (46,3)
Réduction de dimensions	81 (33,8)
Séries chronologiques	55 (22,9)
Apprentissages machine	81 (33,8)

Logiciels statistiques

La médiane du nombre de logiciels utilisés par les répondants est de 3 (minimum=0, Q1=2, Q3=5, maximum=11). L'utilisation des logiciels est présentée dans le tableau 4. Dans l'ordre, on retrouve Excel, SAS, R, SQL et Python comme les logiciels les plus fréquemment utilisés. Parmi les autres logiciels rapportés, on note entre autres Julia, Minitab, Sudaan, NQquery et Tableau qui reviennent quelques fois.

Tableau 4: Utilisation des logiciels statistiques

Logiciels (n = 257)	Fréquence (%)
Excel	191 (74,3)
SAS	189 (73,5)
R	151 (58,8)
SQL	89 (34,6)
Python	84 (32,7)
PowerBi	54 (21,0)
SPSS	43 (16,7)
Autres	29 (11,3)
G Power	21 (8,2)
STATA	20 (7,8)
MPLUS	15 (5,8)
PASS	14 (5,4)
Matlab	5 (1,9)
GraphPad Prism	3 (1,2)
SYSTAT	2 (0,8)
JMP	2 (0,8)

La fréquence d'utilisation des quatre logiciels les plus utilisés est présentée dans le tableau 5. Excel est utilisé de façon très fréquente. Parmi les logiciels de traitement statistique, les utilisateurs de SAS ont une utilisation plus régulière que les utilisateurs de R ou de Python.

Tableau 5 : Fréquence d'utilisation des logiciels statistiques les plus courants

Logiciel (n manquant)	Tous les jours	Toutes les semaines	Tous les mois	Quelques fois par années	Encore moins souvent
Excel (2)	104 (55,0)	68 (36,0)	14 (7,4)	2 (1,1)	1 (0,5)
SAS (4)	93 (50,3)	39 (21,1)	23 (12,4)	20 (10,8)	10 (5,4)
R (1)	39 (26,0)	44 (29,3)	23 (15,3)	39 (26,0)	5 (3,3)
Python (1)	20 (24,1)	16 (19,3)	21 (25,3)	23 (27,7)	3 (3,6)

L'utilisation des logiciels SAS, R et Python a été croisée avec certaines caractéristiques des employés et de l'emploi. SAS est surtout utilisé au gouvernement, dans le milieu de la santé et de la consultation, par des gens avec plus d'expérience. R est utilisé par presque tous les employés d'un établissement d'enseignement, et par près de la moitié des employés au privé et au gouvernement. R est davantage utilisé chez les gens avec moins d'expérience. Quant à Python, il est utilisé surtout dans le privé, dans le domaine des assurances, par ceux ayant moins d'expérience dans le milieu.

Limites

Les mêmes limites que celles rapportées dans la première partie du rapport publié à l'automne dernier s'appliquent aussi ici.

De plus, le portrait des logiciels utilisés sur le marché du travail n'indique pas nécessairement ce qui devrait être enseigné. On voit que les plus récents diplômés (qui ont moins d'expérience) utilisent davantage R et Python, et que SAS est bien implanté au gouvernement. Avec la hausse annoncée des coûts des licences, il est possible que cette forte présence de SAS diminue dans le futur.

Conclusion

En résumé, cette étude basée sur 261 travailleurs dans le domaine de la statistique et de l'analyse de données a permis de connaître la satisfaction des répondants, la nature du travail réalisé et avec quels logiciels ces derniers travaillent.

[Anne-Sophie Julien](#)



Comparaison des emplois en statistique selon leur titre

Lors d'un jeudi de l'ASSQ en 2022, Antoine Girard a présenté son « Enquête sur les profils data » réalisée en France. Les résultats sont disponibles en ligne¹. Cette enquête a été accomplie via son réseau LinkedIn et a inclus 150 répondants, qui étaient majoritairement des « data analysts » (incluant les chargés d'étude) ou des « data scientists » (incluant les statisticiens). Les logiciels utilisés, les types de tâches et d'analyses statistiques réalisées et préférées ont été comparés entre les deux groupes. Leur profil démographique était similaire. Python était le logiciel le plus utilisé, surtout chez les scientifiques de données. Les outils de visualisation de données ou de « business intelligence » étaient plus utilisés par les analystes des données. Ces derniers faisaient davantage de tableaux de bord, de visualisation et de requêtes, alors que les scientifiques faisaient davantage d'analyses statistiques. Quant aux tâches préférées par les employés, les scientifiques aimaient surtout prédire et modéliser, alors que les analystes aimaient mieux visualiser les données.

Cette étude a été un des motifs menant à l'enquête sur les conditions d'emploi en statistique et en analyse de données au Québec réalisée à l'hiver 2023 par l'ASSQ. Les premiers résultats descriptifs de cette enquête québécoise, ainsi que sa méthodologie, ont été présentés dans l'édition de septembre 2023 de *Convergence*. La dernière partie des résultats est présentée dans un autre article de la présente édition de *Convergence*. Ensuite, nous nous sommes intéressés à effectuer une comparaison similaire à celle réalisée en France, afin de mieux comprendre la distinction et l'utilisation des principaux titres d'emplois dans le domaine.

Pour faire cette comparaison, le titre d'emploi tel que défini par l'employeur est utilisé (par opposition au titre d'emploi utilisé par le participant pour décrire son emploi). Les quatre principaux titres présents dans l'échantillon et retenus pour les analyses actuelles, sont les statisticiens ($n = 51$), les biostatisticiens ($n = 33$), les analystes de données ($n = 25$) et les scientifiques des données ($n = 37$). Les autres titres sont exclus, comme les consultants, les méthodologistes, les professeurs, les cadres, etc.

Les caractéristiques des employés et de leur emploi sont présentées dans le tableau 1 selon le titre d'emploi. La parité des genres est atteinte chez les statisticiens, mais chez des scientifiques et les analystes des données, les trois quarts sont des hommes. Les analystes, et scientifiques dans une moindre mesure, sont plus jeunes que les autres groupes. La majorité des biostatisticiens a une maîtrise en statistique, contrairement aux autres groupes, comme les analystes qui ont seulement un baccalauréat. La majorité des statisticiens travaille au gouvernement, les scientifiques et les analystes des données sont plutôt au privé et surtout en

1

<https://public.tableau.com/app/profile/antoine.girard/viz/Rsultatsenquatedata2021/RsultatsdenqueteData>

assurances/finances/bancaire, alors que les biostatisticiens se retrouvent surtout dans les centres de recherche et autres organismes en santé.

Tableau 1 : Caractéristiques des employés et de leur emploi selon le titre d'emploi

Caractéristiques		Statisticiens	Biostatisticiens	Analystes de données	Scientifiques des données
Genre masculin		25 (49)	20 (61)	19 (76)	27 (75)
Âge	30 ans et -	7 (14)	NP	15 (60)	13 (35)
	31-40 ans	14 (27)	10 (31)	7 (28)	17 (46)
	41 ans et +	30 (59)	18 (56)	NP	7 (19)
Degré du plus haut diplôme en statistique	Aucun	NP	NP	NP	5 (14)
	Baccalauréat/Cert.	15 (29)	NP	10 (43)	10 (27)
	Maitrise	31 (61)	28 (85)	11 (48)	20 (54)
	Doctorat	NP	NP	NP	NP
Employeur	Privé	12 (24)	8 (24)	12 (48)	27 (73)
	Gouvernement	32 (63)	8 (24)	9 (36)	NP
	Autres	7 (14)	17 (51)	NP	NP
Domaine de l'employeur	Assurances / finance / bancaire	10 (20)	NP	14 (56)	23 (62)
	Santé	15 (31)	32 (97)	5 (20)	NP
	Autres	24 (49)	NP	6 (24)	12 (32)

Les valeurs sont présentées sous la forme N (% par titre d'emploi).

NP : non présenté lorsque $N < 5$.

Qu'en est-il du travail effectué par chacun de ces groupes? Les logiciels employés et les tâches effectuées dans le cadre de l'emploi ont aussi été comparés. On remarque au tableau 2 que les statisticiens et les biostatisticiens effectuent davantage de planification d'enquêtes et d'expériences, et de rédaction de rapports ou d'articles scientifiques que les deux autres groupes. Les analystes de données font davantage de tableaux de bord et de visualisation de données, mais moins d'analyses statistiques. Les scientifiques des données utilisent moins SAS et davantage Python que les trois autres groupes. Finalement, les analystes et les scientifiques font davantage de requêtes et extractions de données ou d'automatisation de processus.

Quant aux types d'analyses statistiques, les statisticiens et les biostatisticiens effectuent davantage de calculs de taille d'échantillon, des analyses de données d'enquête, du développement et validation de questionnaires, de modèles d'équations structurelles, et des analyses multivariées que les deux autres groupes. Les analystes font moins de modèles linéaires et non linéaires que les trois autres groupes. Les scientifiques de données font moins de tableaux de fréquences, mais plus de classification, de réduction de dimensions, de séries chronologiques et d'apprentissage machine que les trois autres groupes.

Si on regarde les caractéristiques quantitatives de l'emploi au tableau 3, on remarque que les analystes et les scientifiques des données ont moins d'expérience; ce sont de nouveaux titres d'emploi qui semblent être utilisés pour recruter de jeunes travailleurs.

Tableau 2 : Logiciels utilisés, tâches et analyses effectuées selon le titre d'emploi

Caractéristiques	Statisticiens	Biostatisticiens	Analystes de données	Scientifiques des données
Logiciels				
R	29 (59)	24 (75)	12 (48)	25 (69)
SAS	46 (94)	25 (78)	14 (56)	13 (36)
Python	12 (24)	6 (19)	10 (40)	29 (81)
Tâches				
Création de banques de données	15 (32)	10 (31)	8 (33)	10 (29)
Gestion de banques de données	17 (36)	19 (59)	11 (46)	10 (29)
Nettoyage de banques de données	27 (57)	21 (66)	14 (58)	17 (49)
Traitement de données	38 (81)	25 (78)	21 (88)	30 (86)
Requêtes et extractions de données	27 (57)	19 (59)	18 (75)	27 (77)
Tableaux de bord	11 (23)	8 (25)	20 (83)	13 (37)
Analyses descriptives	40 (85)	31 (97)	21 (88)	26 (74)
Visualisation de données	34 (72)	21 (66)	24 (100)	30 (86)
Automatisation de processus	16 (34)	5 (16)	12 (50)	24 (69)
Simulation	17 (36)	10 (31)	4 (17)	9 (26)
Planification d'enquêtes, d'études, de sondages, d'expériences	29 (62)	19 (59)	4 (17)	4 (11)
Analyses statistiques	43 (91)	32 (100)	17 (71)	29 (83)
Rédaction de rapports ou d'articles scientifiques	37 (79)	22 (69)	7 (29)	11 (31)
Préparation/animation d'ateliers de formation	12 (26)	14 (44)	4 (17)	6 (17)
Consultation avec clients	26 (55)	18 (56)	7 (29)	14 (40)
Gestion d'employés	7 (15)	5 (16)	2 (8)	3 (9)
Analyses				
Calcul de taille d'échantillon	33 (72)	28 (88)	4 (19)	5 (14)
Analyse de données d'enquête et de sondage	29 (63)	15 (47)	7 (33)	7 (20)
Développement et validation de questionnaires et d'instruments	21 (46)	11 (34)	1 (5)	0 (0)
Équations structurelles, analyses factorielles	13 (28)	12 (38)	2 (10)	3 (9)
Modèles linéaires et non linéaires	33 (72)	31 (97)	9 (43)	29 (83)
Analyses multivariées	33 (72)	27 (84)	7 (33)	18 (51)
Tableaux de fréquences	39 (85)	28 (88)	18 (86)	13 (37)
Classification	20 (43)	16 (50)	5 (24)	26 (74)
Réduction de dimensions	15 (33)	11 (34)	2 (10)	21 (60)
Séries chronologiques	6 (13)	5 (16)	3 (14)	15 (43)
Apprentissage machine	12 (26)	10 (31)	8 (38)	28 (80)

Les valeurs sont présentées sous la forme N (% par titre d'emploi).

Le nombre de logiciels utilisés et de tâches effectuées est similaire entre les quatre titres d'emploi. Cependant, le nombre de méthodes statistiques utilisées par les analystes est inférieur à celui des autres titres d'emploi. Aucune différence notable n'est observée quant à la perception de la complexité des tâches, de la stimulation apportée par l'emploi et de la charge de travail.

Tableau 3 : Caractéristiques quantitatives de l'emploi selon le titre d'emploi

Caractéristiques	Titre d'emploi	Minimum	25 ^e percentile	Médiane	75 ^e percentile	Maximum
Nombre d'années d'expérience	Statisticien	0	8	13	25	38
	Biostatisticien	1	4,5	14	21	30
	Analyste	0,5	3	5	7	17
	Scientifique	0,5	3	5	9	20
Nombre de logiciels	Statisticien	0	2	4	5	11
	Biostatisticien	1	2	3,5	5	8
	Analyste	1	3	4	5	7
	Scientifique	1	3	3	4	9
Nombre de tâches	Statisticien	0	6	9	10	15
	Biostatisticien	0	7	9	10	17
	Analyste	0	6	8	10	13
	Scientifique	0	6	7	9	13
Nombre de méthodes statistiques	Statisticien	0	4	5,5	8	10
	Biostatisticien	2	4	6	7,5	11
	Analyste	1	2	2	4	7
	Scientifique	0	3	5	6	9
Complexité des tâches*	Statisticien	1	6	7	8	10
	Biostatisticien	4	7	8	8	10
	Analyste	4	5	7	8	9
	Scientifique	2	6	7	8	10
Travail stimulant*	Statisticien	1	7	8	9	10
	Biostatisticien	3	8	9	9	10
	Analyste	5	6,5	8	9	10
	Scientifique	1	7,5	8	9	10
Charge de travail*	Statisticien	2	6	7	8	10
	Biostatisticien	4	6	7	8	10
	Analyste	5	6,5	8	8	10
	Scientifique	1	5	7	8	10

*Ces trois caractéristiques sont évaluées à l'aide d'une échelle de Likert de 1 à 10. La valeur 10 correspond à des tâches plus complexes, un travail plus stimulant, et une charge de travail plus importante.

En comparant les résultats actuels à ceux de l'étude réalisée en France par Antoine Girard, on constate que les titres d'analystes et de scientifiques de données sont plus fréquents en France qu'au Québec. L'utilisation des logiciels est différente, alors que Python a pris de l'avance en France. Quant à la comparaison des tâches et analyses effectuées, les résultats sont comparables dans les deux enquêtes, bien que les titres comparés ne soient pas exactement les mêmes. Il pourrait être intéressant de faire la même enquête et reprendre ces comparaisons dans quelques années afin de voir l'évolution des titres d'emploi dans ce domaine en constante évolution.

Anne-Sophie Julien

In memoriam :

Robert Côté (1938-2023)

Le 26 novembre 2023, la communauté statistique québécoise a perdu un pionnier et un bâtisseur en la personne de Robert Côté, professeur de statistique retraité de l'Université Laval, où il œuvra de 1971 à 2002.

Fils de Blanche Murray et Eugène Côté, agriculteurs, Robert naquit à Matane le 19 février 1938. Il était le 10^e enfant d'une famille qui en comptait 13. Élève brillant et appliqué, il exprima très tôt le souhait de poursuivre ses études à Rimouski en vue d'adhérer à un ordre religieux. Il fut ainsi amené à quitter le foyer familial à 12 ans, mais à l'aube de la Révolution tranquille, il résolut de revenir à la vie laïque et d'étudier les mathématiques à l'Université Laval, dont il fut diplômé en 1966. Il décida ensuite d'aller aux États-Unis pour se spécialiser en statistique, choix hardi pour l'époque au Canada français. Il fut encadré par Robert J. Hader à la North Carolina State University, à Raleigh. Sa thèse, complétée en 1971, concernait une approximation à biais minimal pour un modèle de régression général; voir par exemple [1].



De retour à Québec, Robert joignit les rangs du Département de mathématiques de l'Université Laval, où les charges de cours étaient lourdes et l'enseignement de la statistique n'en était qu'à ses balbutiements. Avec ses collègues, au premier chef Philippe Capéraà, Robert entreprit de développer un programme de baccalauréat spécialisé en statistique (BSpStat), une première au Québec. En 1980, le lancement de ce programme de conception avant-gardiste s'accompagna de crédits qui permirent à la discipline de prendre son essor. L'embauche de Louis-Paul Rivest, dès 1981, mena entre autres à la création du Service de consultation statistique, toujours très actif.

Robert fut le premier directeur du programme du BSpStat; il occupa le poste à trois reprises (1980-82, 1985-86, 1996-97). De 1986 à 1991, il assumait en outre la direction du département, qui regroupait alors 37 professeurs. Il joua un rôle de conciliateur et de facilitateur lorsque les actuaux du département émirent le vœu de se doter d'une unité distincte. Il continua de diriger le Département de mathématiques et de statistique après la fondation de l'École d'actuariat. C'est aussi pendant son mandat qu'un programme de maîtrise en statistique vit le jour. Cette formation, qui gagna rapidement en popularité, motiva l'embauche de deux nouveaux statisticiens, portant le total à 10, un sommet inégalé depuis.

Auteur de nombreux photocopiés de notes de cours, Robert s'investit beaucoup en recherche appliquée et encadra quelques étudiant.e.s de 2^e cycle en statistique. Il entretenait entre autres une

longue et fructueuse collaboration avec Michel Maziade, psychiatre et chercheur en génétique psychiatrique à la Faculté de médecine de l'Université Laval; voir par exemple [2-4], mais aussi [5]. Doté d'une force tranquille et d'une patience à toute épreuve, sourire aux lèvres, il avait un sens de l'écoute remarquable et beaucoup d'empathie. Administrateur prudent et efficace, il était un modèle de dévouement respecté et apprécié de tous. Homme de principe, il n'hésitait jamais à défendre ses collègues et sa discipline.

Mariés en 1972, Robert et son épouse, Micheline Olivier, eurent deux enfants : Marie-Andrée, née en 1973, et Martin, né en 1974. Ce dernier se révéla atteint de paralysie cérébrale et il présentait d'importants troubles moteurs. Robert et Micheline décidèrent alors de se consacrer entièrement à leur famille et d'offrir à Martin tout l'accompagnement nécessaire à la poursuite d'études normales, sans pour autant négliger leur fille aînée. À une époque où les ressources spécialisées étaient rares, ils adaptèrent leur domicile à grands frais et virent eux-mêmes à l'éducation de leur fils. Ils furent également actifs au sein de l'Association de paralysie cérébrale du Québec.

La retraite de Robert, prise à l'aube de ses 65 ans, fut assombrie par le décès de Martin dans des circonstances tragiques, en août 2002, alors que celui-ci s'apprêtait à terminer ses études de 2^e cycle en théologie. Pour se rapprocher de sa fille, audioprothésiste, et de ses deux petits-enfants, Robert quitta alors Québec pour Terrebonne avec son épouse, qui le précéda dans la mort en août 2020. Il rendit l'âme à l'hôpital Pierre-Legardeur, à l'âge de 85 ans, des suites d'une pneumonie.

Tous ceux qui ont connu Robert en garderont un souvenir impérissable. C'est lui qui m'avait embauché en 1987 et j'ai largement bénéficié de ses précieux conseils. Je le considérais comme un ami et lui serai éternellement reconnaissant.

Christian Genest, Université McGill

Références

- [1] R. Côté, A.R. Manson & R.J. Hader (1973). Minimum bias approximation of a general regression model. *Journal of the American Statistical Association*, vol. 68, no 343, pp. 633–638.
- [2] M. Maziade, C. Caron, R. Côté, C. Mérette, H. Bernier, B. Laplante, P. Boutin & J. Thivierge (1990). Psychiatric status of adolescents who had extreme temperaments at age 7. *The American Journal of Psychiatry*, vol. 147, no 11, pp. 1531–1536.
- [3] M. Maziade, R. Côté, M. Boudreault, J. Thivierge & P. Capéraà (1984). The New York Longitudinal Studies Model of Temperament: Gender differences and demographic correlates in a French-speaking population. *Journal of the American Academy of Child Psychiatry*, vol. 23, no 5, 582–587.
- [4] M. Maziade, C. Mérette, M. Cayer, M.-A. Roy, P. Szatmari, R. Côté & J. Thivierge (2000). Prolongation of brainstem auditory-evoked responses in autistic probands and their unaffected relatives. *Archives of General Psychiatry*, vol. 57, no 11, pp. 1077–1083.
- [5] J. Thivierge & R. Côté (1990). Brain-stem auditory evoked response: Normative values in children. *Electroencephalography and Clinical Neurophysiology*, vol. 77, no 4, pp. 309–313.

Nouvelles de l'Institut de la statistique du Québec

Dans un premier temps, j'aimerais me présenter en tant que nouvelle collaboratrice pour les prochaines éditions du *Convergence*. Mon nom est Jessica Bélisle, nouvelle statisticienne à l'Institut de la statistique du Québec (ISQ). Je suis heureuse de prendre le relais de mon collègue Éric Gagnon pour vous informer des différents projets de notre organisation.

Pour cette édition du *Convergence*, j'aimerais tout d'abord décrire certaines études menées à l'ISQ concernant des indicateurs pré et post-pandémie. La pandémie a eu des répercussions importantes sur la vie de tout un chacun et ses effets se font encore ressentir, notamment en matière de fréquentation muséale. Dans le bulletin *La fréquentation des institutions muséales en 2022*, on constate que malgré plusieurs améliorations encourageantes, la fréquentation totale est encore en rattrapage. Par exemple, après un creux historique en 2020, la fréquentation des institutions muséales du Québec a atteint près de 10,6 millions d'entrées en 2022. Ce résultat représente une hausse d'environ 57 % par rapport à 2021 (6,8 M), mais un recul de 30 % par rapport à la moyenne de 2015 à 2019 (15,2 M). L'un des retours majeurs a été celui des groupes scolaires et des groupes organisés. En effet, la fréquentation scolaire a atteint près de 0,6 million d'entrées en 2022, soit une forte hausse d'environ 240 % par rapport au résultat de 2021 (0,2 M). De même, le nombre de visiteurs en groupes organisés est passé de 0,1 million en 2021 à près de 0,5 million en 2022. Bien que cela constitue une importante progression, le retard par rapport aux les entrées pré-pandémiques est encore considérable, potentiellement en lien avec certaines nouvelles pratiques telles que les conférences virtuelles. Toutefois, de nouvelles habitudes telles que le mode de travail hybride pourraient bénéficier aux institutions en milieux éloignés. Les effets de cette « nouvelle normalité » pourront être mesurés dans les années à venir. Les détails de l'enquête sont disponibles via le lien suivant : [Optique culture - Numéro 91. Octobre 2023 - La fréquentation des institutions muséales au Québec en 2022](#)

Force est de constater que la COVID-19 nous a fait revoir nos pratiques. Il en va de même pour le secteur des arts de la scène au Québec qui s'est adapté de façon remarquable dans les dernières années grâce à plusieurs innovations. L'un des éléments qui ressort de *l'Enquête sur les effets de la pandémie de COVID-19 sur les entreprises et organismes du secteur des arts de la scène au Québec* est la diversité du mode de présentation des différents spectacles, le présentiel étant parfois difficile à offrir. Pour notre plus grand bonheur, la majorité des organisations prévoient que le secteur devrait avoir retrouvé un niveau d'activité pré-pandémique d'ici le printemps 2024 en ce qui concerne les cycles de création et les partenariats d'affaires. Or cela pourrait s'avérer un peu plus long pour les taux d'occupation, les taux d'abonnement et les revenus autonomes qui devraient attendre plus de deux ans pour un retour à la normale. Les informations se trouvent ici : <https://statistique.quebec.ca/fr/communiquer/innovant-secteur-arts-scene-quebec->

[strategies-adaptation-contexte-pandemique?utm_source=dialoginsight&utm_medium=email](https://statistique.quebec.ca/fr/document/strategies-adaptation-contexte-pandemique?utm_source=dialoginsight&utm_medium=email)

Toujours en lien avec la pandémie de COVID-19, nous constatons que la quarantaine a eu un impact sur la vie de plusieurs jeunes Québécois. En effet, selon l'*Enquête québécoise sur le parcours préscolaire des enfants de maternelle* (EQPPEM), plus de trois enfants de maternelle 5 ans sur quatre (77 %) ont vécu dans un ménage où au moins une personne a dû se placer en quarantaine en raison de la COVID-19 durant la période allant de mars 2020 à la réalisation de l'enquête (d'avril à août 2022). Vous pourrez avoir davantage de détails ici : <https://statistique.quebec.ca/fr/document/enfants-maternelle-5-ans-2021-2022-pandemie-covid/publication/enfants-maternelle-2021-2022-quarantaines-situation-emploi-parents-pandemie-covid-19>

Sur une autre note, le bulletin *La consommation d'enregistrements musicaux au Québec en 2022* portant sur la musique en continu au Québec nous informe qu'il y a eu près de 24 milliards d'écoutes en 2022, mais peu d'écoutes d'interprètes d'ici. Par exemple, Les Cowboys Fringants se retrouvent au 15^e rang. Il sera intéressant notamment de mesurer l'évolution des écoutes dans la prochaine année pour ce groupe lors de la parution du prochain bulletin, étant donné le décès récent de leur chanteur. Tous les détails se trouvent via le lien suivant : https://statistique.quebec.ca/fr/communiqu%C3%A9/musique-continu-24-milliards-pistes-ecoutees-quebec-2022?utm_source=dialoginsight&utm_medium=email

En terminant, je souhaite également mettre en lumière la mise à jour des Indicateurs de progrès du Québec pour la mesure du bien-être et de la richesse nationale réalisée dernièrement par l'ISQ :

- Indice de gravité de la criminalité (15 septembre 2023)
- Dépassements du seuil pour la concentration de particules fines dans l'air (17 juillet 2023)
- Taux d'accroissement de la population (17 juillet 2023)
- Espérance de vie (17 juillet 2023)
- Santé mentale excellente ou très bonne (17 juillet 2023)

Ces indicateurs visent à mesurer ce qui est le plus important pour les personnes, à partir d'un nombre restreint d'indicateurs. Vous trouverez les détails propres à chaque indicateur ici : <https://statistique.quebec.ca/fr/document/indicateurs-progres-quebec>

Bonne année 2024!

Jessica Bélisle, ISQ

**Institut
de la statistique**
Québec 

Nouvelles de Statistique Canada

Chers lecteurs de *Convergence*, dans cette édition, je partage avec vous un article préparé par mon collègue Martin Provost de la Direction des méthodes statistiques modernes et de la science des données. Son article offre un bref survol des études de couverture effectuées dans le cadre du Recensement de la population de Statistique Canada. Merci Martin!

François Brisebois, Statistique Canada

Les études de couverture du Recensement de la population de 2021

Contexte

Le Recensement de 2021 a dénombré près de 37 millions de personnes. En dépit de normes de qualité rigoureuses régissant les opérations du recensement, il est impossible de mener une telle opération en éliminant toute source d'erreur. Dans un recensement, les erreurs de couverture figurent parmi les plus importantes. Les erreurs de couverture commises lors du recensement sont mesurées à l'aide de trois études. La première, l'Enquête sur la classification des logements (ECL), porte sur les erreurs de couverture attribuables aux logements occupés classés par erreur comme inoccupés [lors du Recensement de la population]. Ainsi, les données du recensement sont rajustées en fonction de ce type d'erreur de couverture. La deuxième, l'Enquête sur le sous-dénombrement du recensement (ESoR) permet d'évaluer le nombre de personnes omises lors du recensement. La troisième, l'Étude sur le surdénombrement du recensement (ESuR), permet quant à elle d'évaluer le dénombrement de personnes à plus d'une reprise. Le surdénombrement désigne également l'inclusion de personnes qui n'auraient pas dû être dénombrées, mais ce type de surdénombrement n'est pas mesuré par l'ESuR, étant jugé négligeable en comparaison des dénombremments multiples. Les données du recensement ne sont pas rajustées en fonction des erreurs de couverture décelées par l'ESoR ou l'ESuR. Statistique Canada utilise plutôt une estimation du sous-dénombrement net (le sous-dénombrement moins le surdénombrement) pour établir des estimations démographiques.

Enquête sur la classification des logements

Depuis 1981, l'Enquête sur la classification des logements (ECL) est effectuée dans le cadre du Recensement du Canada afin d'évaluer la proportion de logements ayant été classifiés erronément comme étant inoccupés le jour du recensement. Depuis 2001, l'ECL a été élargie en vue d'estimer également le nombre de logements privés occupés parmi ceux qui n'avaient pas retourné de questionnaire de recensement et qui étaient classés comme des logements non répondants. L'ECL estime également la distribution de la taille des ménages habitant dans ces deux types de logements.

Pour ce faire, l'ECL sélectionne un échantillon de logements inoccupés et non répondants à travers le pays. À l'aide d'entrevues sur le terrain indépendantes des opérations de collecte originale du recensement, des intervieweurs doivent déterminer le statut de ces logements au jour du recensement et, pour ceux qui étaient vraisemblablement occupés, leur nombre de résidents habituels. À partir de ces informations, l'ECL estime le nombre de logements occupés de chaque taille, pour chaque type de logement (inoccupé et non répondant) et pour différentes unités géographiques. Les résultats de l'ECL sont ensuite utilisés dans le cadre de la procédure d'imputation de ménages entiers (IME), pour corriger la base de données du recensement.

Enquête sur le sous-dénombrement du recensement

L'ESoR, qui s'appelait jusqu'au Recensement de 2016 la Contre-vérification des dossiers (CVD), a débuté en 1961. Pour cette étude, un échantillon est sélectionné de huit bases de sondage différentes, cinq pour les provinces et une pour chacun des trois territoires. Dans les provinces, celles-ci sont la base du recensement précédent, celle des personnes omises par le recensement précédent, celle des immigrants intercensitaires, celle des naissances intercensitaires et celle des résidents non permanents, qui couvre les détenteurs d'un permis de travail ou d'étude valide au moment du recensement, de même que les demandeurs d'asile. Dans chacun des territoires, le fichier de santé du territoire sert de base de sondage. Un échantillon aléatoire de personnes (appelées PC pour personnes choisies) est sélectionné de chaque base. Comme il n'existe pas de liste de toutes les personnes omises au recensement précédent, les personnes classées omises par l'étude de sous-dénombrement précédente servent de base virtuelle. Toutes ces personnes se retrouvent dans l'échantillon du cycle suivant, et leur poids d'estimation final du cycle précédent sert de poids de départ pour le cycle suivant.

Des appariements sont faits aux fichiers de décès de l'état civil, de même qu'à différents fichiers fiscaux afin d'identifier les PC décédées, et d'obtenir les informations de contact les plus à jour possibles pour chaque PC. Toute l'information disponible pour chaque PC est utilisée afin d'apparier l'échantillon de l'ESoR à la base de données de réponse (BDR) du recensement. Les PC trouvées, à l'aide d'appariements automatisés ou de recherches manuelles, sont classées dénombrées. La liste des autres PC qui n'ont été ni trouvées ni classées décédées est envoyée à la collecte. L'objectif de la collecte est d'interviewer chaque PC ou un répondant par procuration acceptable afin de déterminer si la PC faisait partie de la population cible du recensement de 2021 et, le cas échéant, obtenir son adresse le jour du recensement et l'information sur les personnes qui habitaient à la même adresse, de même que toute autre adresse où la PC aurait pu être dénombrée. Cette information est ensuite utilisée pour tenter à nouveau de trouver la PC sur la BDR. À la suite de ces opérations, chaque PC se trouve classée comme étant soit dénombrée, soit omise par le recensement, soit hors de la population cible (décédée, émigrée ou à l'étranger), soit non répondante à la collecte. Après la redistribution des poids des non-répondants et d'autres ajustements de poids, l'échantillon de personnes omises est utilisé pour estimer le sous-dénombrement du recensement et ses caractéristiques.

Étude sur le surdénombrement du recensement

Depuis le Recensement de 2006, l'ESuR est l'unique véhicule permettant d'estimer le surdénombrement du recensement de la population. Auparavant, celui-ci était estimé à l'aide d'une combinaison d'études. Afin d'estimer le nombre de personnes dénombrées plus d'une fois au recensement, différents couplages sont effectués afin de créer une liste de doublons potentiels. Les variables utilisées varient d'un couplage à l'autre, mais incluent des combinaisons de noms, de dates de naissance et d'information géographique. Tout d'abord, un couplage probabiliste de la BDR à elle-même permet d'identifier des personnes aux caractéristiques semblables. Puis, un couplage de la BDR à différents fichiers administratifs représentatifs de la population canadienne permet d'identifier des enregistrements de la BDR qui s'appartiennent aux mêmes enregistrements administratifs, indiquant que ces enregistrements pourraient représenter la même personne. Ces couplages sont effectués au niveau des personnes. Le but de ces couplages n'est pas d'identifier des cas de surdénombrement avec certitude, mais bien de créer une liste de personnes potentiellement dénombrées plus d'une fois. En parallèle à ces couplages probabilistes, un couplage déterministe au niveau des ménages est effectué. Un appariement de la BDR à elle-même est effectué en comparant le niveau de similitude des noms

et des dates de naissance des membres de chaque paire de ménages. Le but de cette étape est de bénéficier de la force d'un couplage de ménages ayant plusieurs membres en commun afin d'identifier des enregistrements véritablement en double sur la BDR (doublons définitifs).

Le chevauchement entre les doublons potentiels et définitifs issus des différents couplages est éliminé et une liste de paires uniques est ainsi créée. Lorsqu'une paire d'enregistrements de la BDR a été identifiée à la fois par un des couplages probabilistes et par le couplage déterministe, ce dernier a priorité. Tous les enregistrements de la BDR reliés par du surdénombrement potentiel sont réunis dans un groupe d'enregistrements (par exemple, si les paires de personnes A-B, B-C et A-C ont été identifiées, alors le groupe A-B-C sera créé). Ces groupes d'enregistrements forment la base de sondage de l'ESuR. Les groupes constitués d'enregistrements provenant uniquement de paires identifiées par le couplage déterministe sont mis dans une strate à tirage complet et sont considérés comme étant du surdénombrement, sans vérification. Un échantillon aléatoire stratifié est sélectionné parmi les autres groupes. Toutes les paires constituant ces groupes sont vérifiées manuellement par une équipe de codeurs, qui déterminent lesquelles représentent un surdénombrement véritable. L'estimation du surdénombrement du recensement est établie à partir de ces résultats.

Résultats des études de couverture

Comme expliqué précédemment, les données du recensement sont rajustées pour tenir compte des résultats de l'ECL. Les résultats de population et de logements publiés le 9 février 2022, ainsi que tous les résultats publiés par la suite en tiennent donc compte. Pour ce qui est des résultats de l'ESoR et de l'ESuR, des estimations provisoires ont été rendues disponibles en avril 2023. Une période de validation des résultats s'ensuit, en collaboration avec les représentants de chaque province et territoire. Les [estimations définitives des erreurs de couverture du Recensement de 2021](#) ont été diffusées le 27 septembre 2023. Un rapport technique du recensement sur la couverture sera publié en octobre 2024. Ce rapport décrit en détail la méthodologie et les résultats des études de couverture du recensement. En attendant, le rapport technique sur la couverture du Recensement de 2016 peut être consulté ici : [Rapport technique sur la couverture, Recensement de la population, 2016](#).

Conclusion

Le Recensement de la population est un outil indispensable pour fournir un portrait statistique complet et détaillé du Canada. Il est donc essentiel d'en évaluer la qualité. Les études de couverture du recensement permettent aux utilisateurs d'évaluer les répercussions des erreurs de couverture et de mieux comprendre la façon dont ces erreurs surviennent. Puisque les données du recensement servent également à mettre à jour les estimations démographiques officielles, la mesure des erreurs de couverture du recensement permet d'améliorer l'exactitude des estimations démographiques. Il s'agit d'un travail complexe et une équipe dédiée de méthodologistes et d'analystes y travaille pendant le cycle complet du recensement. Les études de couverture du recensement canadien sont reconnues internationalement pour leurs méthodes robustes et rigoureuses.

Martin Provost,
Statistique Canada



Statistique
Canada

Statistics
Canada

apparition de fibrillation auriculaire (la perturbation la plus courante du rythme cardiaque) et que ces résultats contrastaient avec ceux d'une étude antérieure qui présentait un résultat statistiquement significatif.

Qu'en est-il des résultats eux-mêmes? Les chercheurs pour qui les résultats statistiques étaient non significatifs ont signalé un risque relatif de 1,2 (soit un risque 20 % plus élevé chez les patients exposés par rapport à ceux non exposés). Ils ont également présenté un intervalle de confiance à 95 % qui allait d'une légère diminution de 3 % du risque jusqu'à une augmentation considérable de 48 % du risque ($P = 0,091$, selon notre calcul). Les chercheurs de la première étude, statistiquement significative, ont trouvé exactement le même risque relatif de 1,2, mais leur étude était simplement plus précise, avec un intervalle couvrant une augmentation du risque de 9 % à 33 % ($P = 0,0003$, selon notre calcul).

Il serait aberrant de conclure qu'un résultat statistiquement non significatif ne montre pas d'association lorsque l'intervalle estimé inclut une aussi importante augmentation du risque. De plus, avec une estimation ponctuelle observée identique, il serait tout aussi absurde de prétendre que le résultat de la nouvelle étude est différent du résultat précédent. Pourtant, cette pratique courante montre comment le strict recours aux seuils de signification statistique peut induire en erreur.

Ce genre d'aberration, et d'autres similaires, est très répandu. Des études portant sur des centaines d'articles ont d'ailleurs révélé qu'un résultat statistiquement non significatif est interprété comme n'indiquant « aucune différence » ou « aucun effet » dans environ la moitié des cas.

En 2016, l'*American Statistical Association* a publié une mise en garde dans *The American Statistician* contre une utilisation abusive de la signification statistique et de la valeur P . Le numéro comprenait également de nombreux commentaires sur ce sujet. Ce mois-ci [mars 2019], un numéro spécial dans le même journal tente de pousser plus loin cette remise en cause. Il présente plus de 40 articles sur « L'inférence statistique au 21^e siècle : un monde au-delà du $P < 0,05$ ». Les éditeurs présentent la série de publications en prévenant les auteurs d'être vigilants de ne pas utiliser le terme « statistiquement significatif » [3]. Un autre article [4] entériné par des dizaines de signataires demande également aux auteurs et aux éditeurs de journaux de désavouer ce terme.

Nous sommes totalement d'accord avec ces affirmations et nous réclamons l'abandon complet du concept de signification statistique.

Et nous sommes loin d'être les seuls! Lorsque nous avons invité d'autres personnes à lire une version préliminaire du présent article, à le commenter puis à le signer, nous leur avons aussi demandé s'ils étaient d'accord avec notre message. Résultat : 250 ont entériné le tout dans les 24 premières heures. Une semaine plus tard, ils étaient plus de 800 signataires – des contributeurs ayant tous une affiliation académique ou une expérience de travail dans un domaine qui traite de la modélisation statistique. La liste comprend des statisticiens, des cliniciens et des médecins chercheurs, des biologistes et des psychologues de plus de 50 pays et de tous les continents à l'exception de l'Antarctique. Un cosignataire a qualifié notre article de frappe chirurgicale contre

l'utilisation irréfléchie des tests de signification statistique et d'une possibilité de faire entendre sa voix en faveur de meilleures pratiques scientifiques.

Nous ne réclamons pas une interdiction de l'utilisation des valeurs P ou de ne pas les utiliser comme critère de décision dans certaines applications spécialisées (telles que déterminer si un processus de fabrication répond à certaines normes de contrôle de qualité). Nous ne préconisons pas non plus que tout soit permis, comme dans le cas où de faibles différences rendent soudainement un effet plausible. À l'instar de bien d'autres personnes au fil des décennies, nous appelons plutôt à l'arrêt de l'utilisation des valeurs P de manière conventionnelle et dichotomique comme seul outil de décision pour déterminer si un résultat réfute ou soutient une hypothèse scientifique [5].

UNE CATÉGORISATION À ABOLIR

Le problème est plus humain et cognitif que statistique : regrouper les résultats en « statistiquement significatifs » et « statistiquement non significatifs » fait penser à la majorité des gens que les éléments attribués de cette manière sont catégoriquement différents [6-8]. Les mêmes problèmes sont susceptibles de se poser dans toute alternative statistique proposée impliquant une dichotomisation, qu'elle soit fréquentiste, bayésienne ou autre.

Malheureusement, la fausse croyance selon laquelle franchir le seuil de signification statistique suffit à montrer qu'un résultat est « réel » a conduit les scientifiques et les éditeurs de journaux à privilégier de tels résultats, contaminant ainsi la littérature. Les estimations statistiquement significatives sont privilégiées, et ce, potentiellement dans une large mesure, tandis que les estimations statistiquement non significatives sont beaucoup moins présentes. Par conséquent, toute discussion axée sur des estimations choisies selon leur degré d'importance sera biaisée. Par ailleurs, l'accent strictement mis sur la signification statistique encourage les chercheurs à choisir des données et des méthodes qui produisent une signification statistique pour un résultat souhaité (ou simplement publiable), ou qui donnent une signification statistique non significative pour un résultat non souhaité, comme les effets secondaires potentiels de médicaments, invalidant par le fait même les conclusions.

L'enregistrement préalable des études et l'engagement de publier tous les résultats de toutes les analyses peuvent grandement contribuer à atténuer ces problèmes. Cependant, même les résultats d'études préenregistrées peuvent être biaisés par des décisions laissées invariablement ouvertes dans le plan d'analyse [9]. Cela se produit même avec les meilleures intentions du monde.

Encore une fois, nous ne préconisons pas l'interdiction des valeurs P , des intervalles de confiance ou d'autres mesures statistiques, mais simplement qu'on ne devrait pas les traiter de manière catégorique, que ce soit la dichotomisation selon qu'un résultat est statistiquement significatif ou non, ou la catégorisation basée sur d'autres mesures statistiques telles que les facteurs de Bayes.

L'une des raisons pour éviter une telle « dichotomanie » est que toutes les statistiques, y compris les valeurs P et les intervalles de confiance, varient naturellement d'une étude à l'autre, et ce, souvent à un degré surprenant. En fait, la variation aléatoire à elle seule peut facilement conduire à de grandes disparités dans les valeurs P , bien au-delà de simplement tomber d'un côté ou de

l'autre du seuil de 0,05. Par exemple, même si des chercheurs pouvaient mener deux études parfaitement reproductibles, chacune avec une puissance de 80 % d'atteindre $P < 0,05$, il ne serait pas surprenant que l'une donne un résultat de $P < 0,01$ tandis que l'autre obtiendrait $P > 0,30$. Qu'une valeur P soit petite ou grande, la prudence est de mise dans son interprétation.

Nous devons apprendre à composer avec l'incertitude. Un moyen pratique d'y parvenir consiste à renommer les intervalles de confiance « intervalles de compatibilité » et à les interpréter de manière à éviter un excès de confiance. Plus précisément, nous recommandons aux auteurs de décrire les implications pratiques de chacune des valeurs à l'intérieur de l'intervalle, en particulier l'effet observé (l'estimation ponctuelle) et les bornes de l'intervalle. Ce faisant, ils doivent se rappeler que toutes les valeurs comprises à l'intérieur des limites de l'intervalle sont raisonnablement compatibles avec les données, compte tenu des hypothèses statistiques utilisées pour calculer l'intervalle [7,10]. Par conséquent, individualiser une valeur particulière (telle que la valeur nulle) dans l'intervalle n'a aucun sens.

Nous en avons franchement assez de voir des attestations d'absence d'effet et des affirmations de non-association dans les présentations, les articles scientifiques, les revues de littérature et le matériel pédagogique. Un intervalle contenant la valeur nulle contiendra souvent également des valeurs non nulles d'une grande importance pratique. Ceci étant dit, si on considère que toutes les valeurs à l'intérieur de l'intervalle sont pratiquement sans importance, on pourrait alors dire quelque chose comme « les résultats sont compatibles avec l'absence d'effet important ».

Lorsqu'on parle d'intervalles de compatibilité, il faut garder quatre choses à l'esprit. Premièrement, même si l'intervalle donne les valeurs les plus compatibles compte tenu des données et des hypothèses, cela ne veut pas dire que les valeurs en dehors de cet intervalle ne sont pas compatibles; elles sont juste moins compatibles. En effet, les valeurs situées juste à l'extérieur de l'intervalle ne diffèrent pas substantiellement de celles situées juste à l'intérieur de l'intervalle. Il est donc faux de prétendre qu'un intervalle montre toutes les valeurs possibles.

Deuxièmement, les valeurs contenues dans l'intervalle ne sont pas toutes compatibles de façon équivalente. L'estimation ponctuelle est la plus compatible, et les valeurs proches de celle-ci sont plus compatibles que celles proches des bornes de l'intervalle. C'est pourquoi nous invitons les auteurs à discuter de l'estimation ponctuelle, même lorsqu'ils obtiennent une valeur P élevée ou un intervalle large, et à discuter des bornes de cet intervalle. Par exemple, les auteurs auraient pu écrire : « Comme montré dans une étude précédente, nos résultats suggèrent une augmentation de 20 % du risque d'apparition d'une fibrillation auriculaire chez les patients recevant des anti-inflammatoires. Néanmoins, une différence de risque allant d'une diminution de 3 %, une légère association négative, à une augmentation de 48 %, une association positive substantielle, est également raisonnablement compatible avec nos données, compte tenu de nos hypothèses. » Interpréter l'estimation ponctuelle, tout en reconnaissant son incertitude, empêchera de faire de fausses déclarations selon lesquelles il n'y a « aucune différence » et de faire des affirmations trop confiantes.

Troisièmement, comme le seuil de 0,05 dont il est issu, la valeur par défaut de 95 % utilisée pour calculer les intervalles est elle-même une convention arbitraire. Elle repose sur l'idée trompeuse qu'il y a 95 % de chances que l'intervalle calculé contienne lui-même la vraie valeur, associée au vague sentiment qu'il s'agit là d'une base pour une décision sûre. Un niveau différent peut être

justifié, selon le domaine d'application. Dans l'exemple des anti-inflammatoires, les intervalles estimés peuvent perpétuer les problèmes de signification statistique lorsque la dichotomie qu'ils imposent est traitée comme la norme scientifique.

Enfin, et surtout, restez humble : les évaluations de compatibilité dépendent de l'exactitude des hypothèses statistiques utilisées pour calculer l'intervalle. En pratique, ces hypothèses sont, au mieux, soumises à une incertitude considérable [7,8,10]. Rendez ces hypothèses aussi claires que possible et testez celles que vous pouvez, par exemple, en mettant en graphiques vos données et en ajustant des modèles alternatifs, puis en rapportant tous les résultats.

Peu importe ce que la statistique semble dire, le résultat devrait être étoffé par une gamme d'explications potentielles, et non pas se limiter seulement aux explications favorisées. Les déductions doivent être scientifiques, et cela va bien au-delà du simple aspect statistique. Des facteurs tels que les renseignements préexistants, la conception des études, la qualité des données et la compréhension des mécanismes sous-jacents sont souvent plus importants que les mesures statistiques comme les valeurs P ou les intervalles.

L'objection la plus souvent rencontrée contre l'abandon de la signification statistique est qu'elle est nécessaire pour prendre une décision franche (qui se solde par oui ou non). Mais pour les décisions qui doivent être prises dans les environnements réglementaires, politiques ou commerciaux, les choix fondés sur les coûts, les avantages et les potentiels risques de conséquences l'emportent toujours sur les décisions fondées uniquement sur la signification statistique. De plus, pour décider de poursuivre ou non un axe de recherche, il n'existe pas de lien simple entre une valeur P et les résultats probables des études ultérieures.

À quoi peut-on s'attendre si l'on va de l'avant avec l'abandon de la signification statistique? Nous espérons que les sections sur les méthodes et la tabulation des données seront plus détaillées et plus nuancées. Les auteurs devraient mettre l'accent sur les estimations et leur incertitude, par exemple en discutant explicitement des limites inférieure et supérieure des intervalles estimés. Les conclusions ne devraient pas s'appuyer sur des tests de signification. Si des valeurs P sont rapportées, elles devraient être inscrites avec une précision raisonnable (par exemple, $P = 0,021$ ou $P = 0,13$), sans avoir recours à des étoiles ou à des lettres pour indiquer une signification statistique et sans utiliser la forme binaire $P < 0,05$ ou $P > 0,05$. Les décisions d'interpréter ou de publier les résultats ne devraient pas être basées sur des seuils statistiques. Il faudrait passer moins de temps à utiliser les logiciels statistiques et plus de temps à réfléchir.

Notre invitation à abandonner la signification statistique et à considérer les intervalles estimés comme des intervalles de compatibilité n'est pas une panacée. Même si de nombreuses mauvaises pratiques seront éliminées, il pourrait bien y en avoir de nouvelles qui seront introduites. Ainsi, la surveillance de la littérature pour détecter les abus statistiques devrait être une priorité constante pour la communauté scientifique. L'éradication de la catégorisation contribuera à mettre un terme aux affirmations trop confiantes, aux déclarations injustifiées d'absence de différence et aux allégations absurdes sur l'échec de la reproductibilité lorsque les résultats des études originales et de celles des reprises sont hautement compatibles. L'utilisation abusive de la signification statistique a causé beaucoup de tort à la communauté scientifique et à ceux qui s'appuient sur les recommandations scientifiques. Les valeurs P , les intervalles et les

autres mesures statistiques ont toutes leur place, mais il est temps de laisser tomber la signification statistique.

Valentin Amrhein, Université de Bâle, Suisse,
Sander Greenland, University of California, Los Angeles,
Blake McShane, University in Evanston, Illinois.

[Traduction libre de l'article « *Retire Statistical Significance* » paru dans *Nature*, Vol. 567, mars 2019, pp. 306-307.]

RÉFÉRENCES

1. Fisher, R. A. *Nature* 136, 474 (1935).
2. Schmidt, M. & Rothman, K. J. *Int. J. Cardiol.* 177, 1089–1090 (2014).
3. Wasserstein, R. L., Schirm, A. & Lazar, N. A. *Am. Stat.* <https://doi.org/10.1080/00031305.2019.1583913> (2019).
4. Hurlbert, S. H., Levine, R. A. & Utts, J. *Am. Stat.* <https://doi.org/10.1080/00031305.2018.1543616> (2019).
5. Lehmann, E. L. *Testing Statistical Hypotheses* 2nd edn 70–71 (Springer, 1986).
6. Gigerenzer, G. *Adv. Meth. Pract. Psychol. Sci.* 1, 198–218 (2018).
7. Greenland, S. *Am. J. Epidemiol.* 186, 639–645 (2017).
8. McShane, B. B., Gal, D., Gelman, A., Robert, C. & Tackett, J. L. *Am. Stat.* <https://doi.org/10.1080/00031305.2018.1527253> (2019).
9. Gelman, A. & Loken, E. *Am. Sci.* 102, 460–465 (2014).
10. Amrhein, V., Trafimow, D. & Greenland, S. *Am. Stat.* <https://doi.org/10.1080/00031305.2018.1543137> (2019).

Méditation scientifique

Paris, 1759...

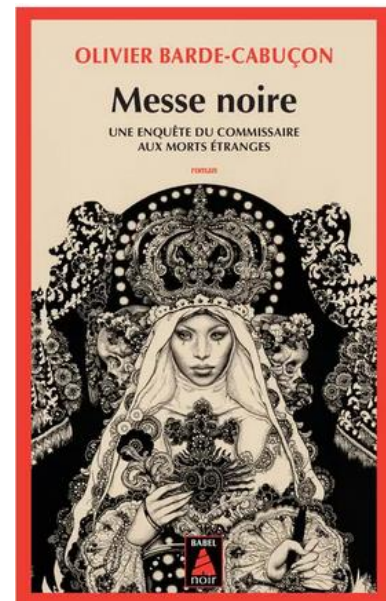
Inclinant la tête, elle le récompensa d'un sourire éclatant.

— J'admire votre science nouvelle, fit-elle.

Le moine se rengorgea.

— Ce n'est rien, j'ai toujours été en avance sur mon époque mais je n'ai guère de mérite à cela car elle est en retard sur tout!

[Extrait de « *Messe noire, une enquête du commissaire aux morts étranges* » d'Olivier Barde-Cabuçon (Éditions Babel Noir).]



Chronique historique :

Le nombre d'or

Le nombre d'or a longtemps été considéré comme un canon en matière d'esthétique

Si dans un triangle égyptien, la forme du triangle correspond à une relation mathématique entre les longueurs de ses côtés, des travaux ultérieurs ont permis d'aller beaucoup plus loin dans l'étude des rapports entre beauté, harmonie d'un côté et mathématique de l'autre. Le nombre d'or, présenté par certains comme un canon de l'esthétique, est au centre de ces travaux. Je vais m'arrêter un instant sur sa définition, son origine, les autres nombres et figures géométriques qui gravitent autour de lui, son utilisation et sa présence dans l'art et la nature.

Le nombre d'or, aussi appelé la divine proportion, le rapport doré, la division sacrée, correspond grosso modo au rapport des longueurs entre les première et deuxième phalanges d'un doigt ou entre les deuxième et troisième. Il vaut 1.618... et est souvent noté ϕ (phi) en hommage au sculpteur Phidias (480- 430 av. J.-C.) qui décora le Parthénon, et peut-être aussi à Fibonacci (Léonard de Pise, 1170-1250) qui a contribué à sa compréhension.

Euclide, le célèbre mathématicien grec, environ 300 ans avant notre ère, serait le premier à avoir formulé mathématiquement ϕ (voir plus loin). Ce rapport apparaît lorsque deux éléments forment un tout et que le rapport du tout au plus grand est exactement identique au rapport du plus grand au plus petit. Remarquons que même s'il n'était pas encore explicitement formulé, ϕ apparaît avant Euclide et se retrouve en architecture et en peinture dans les rapports entre la largeur et la hauteur de rectangles, entre les côtés de triangles, entre des segments définis par des polygones, et aussi sous différentes formes dans la nature.

Le nombre d'or est utilisé essentiellement en architecture et en peinture

La littérature sur le nombre d'or est abondante et les exemples de son utilisation sont légion. Une petite consultation sur Internet vous prouvera facilement sa popularité — peut-être usurpée comme nous le verrons plus tard. Son succès est entre autres lié à sa présence abondante en architecture, dans les œuvres artistiques et dans la nature. Le Parthénon bâti du temps de Périclès (V^e siècle av. J.-C.) sur l'Acropole est souvent donné comme exemple. Sa façade s'inscrit en effet dans un rectangle doré.

Dieu dans l'Exode a ordonné à Moïse de construire une arche (de l'Alliance) de deux coudées et demie de large et une coudée et demie de haut (le rapport de la largeur sur la hauteur qui vaut 1.66 est proche de ϕ). Cette proportion a aussi été utilisée par les Égyptiens dans leurs pyramides et plus tard dans la construction des cathédrales. Les bâtisseurs utilisaient même des compas particuliers pour pouvoir diviser facilement des longueurs données par le nombre d'or. La cathédrale de Paris respecte les canons du nombre d'or à différents endroits de sa façade occidentale (43 m de large sur 69 m de haut).

Plus récemment, l'architecte Le Corbusier (1887-1965) l'a utilisé dans certaines de ses constructions. Il affirmait : « La nature est mathématique, les chefs-d'œuvre de l'art sont en consonance avec la nature. Ils expriment les lois de la nature et ils s'en servent. ». En peinture, ϕ

a présidé à l'organisation de nombreux dessins et tableaux. Léonard de Vinci s'en est servi abondamment.

Ce nombre se retrouve aussi sous différentes formes dans la nature. Il a été montré par exemple que dans la double spirale de l'ADN, le rapport entre la longueur d'une portion de la chaîne hélicoïdale qui fait un tour complet et le diamètre de cette spirale vaut 1,619. ϕ se retrouve dans les formes de virus, dans la disposition de pétales de fleurs, de feuilles sur une tige... Il se retrouve dans des figures régulières comme les pentagones, les spirales logarithmiques que l'on trouve aussi abondamment dans la nature. Vos cartes de crédit et votre carte d'identité sont des rectangles dorés.

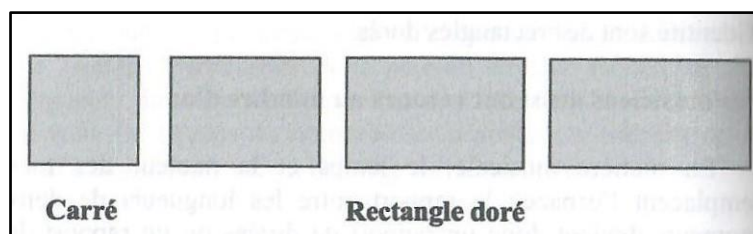
Les musiciens aussi ont recours au nombre d'or

En matière musicale, le temps et la hauteur des tons remplacent l'espace; le rapport entre les longueurs de deux segments devient donc un rapport de durées ou un rapport de hauteurs. Ainsi, ϕ (et les nombres d'une suite qui permet de définir le nombre d'or appelée suite de Fibonacci dont je parlerai un peu plus loin) est utilisé dans plusieurs compositions (Haydn, Debussy, Webern, Bartók...); les mouvements sont construits en différentes phases dont les rapports de durée correspondent à ϕ ou dont le nombre de mesures sont des nombres d'une suite de Fibonacci. Ces proportions harmonieuses peuvent être le fruit d'une construction *a priori* ou plus naturellement le produit d'une simple intuition, d'une recherche spontanée d'harmonie. André Warusfel (2012) a retrouvé des sections dorées (c'est-à-dire dont les rapports de durée sont égaux au nombre d'or) dans « The last Time » par le « Hot Five » de Louis Armstrong.

Si on exprime les intervalles entre les hauteurs des notes en demi-tons, chaque note correspond à un nombre entier. Certains ont joué avec ces nombres et ont ordonné les notes comme des éléments de nouveau d'une suite de Fibonacci (comme Bartók pour régler les entrées d'un canon dans le premier mouvement du quatuor à cordes n° 5).

Le rectangle doré est-il plus plaisant que les autres rectangles ?

Le rôle esthétique du nombre d'or a été plus explicitement mis en évidence par Gustav Fechner (1801-1887), célèbre psychologue allemand qui a entre autres étudié les rapports existant entre excitation et sensation perçue (voir plus loin). Il a publié en 1876 le résultat d'un certain nombre d'expériences sur les préférences marquées par des sujets pour des rectangles de différentes dimensions (voir la figure ci-dessous). Il en avait présenté 10 à 347 sujets et leur avait demandé d'indiquer celui qu'ils préféraient.



Bien entendu, les choix faits par les sujets ne furent pas tous identiques mais le rectangle choisi par le plus de personnes était le rectangle doré. De plus, 41 % de personnes supplémentaires avaient choisi un rectangle dont les rapports des dimensions n'étaient pas très éloignés de ϕ (1.50

ou 1.77). Ce résultat assez surprenant a alimenté depuis de nombreuses recherches sur les similarités entre les individus en matière de goût et de jugements esthétiques. Il paraissait fonder des théories qui donnaient à nos sens du beau une base commune. En fait, à la lumière d'expériences plus récentes (McManus et coll., 2010), il semblerait que la réputation flatteuse du nombre d'or dans le monde de l'esthétique est usurpée. Fechner (1871) dans son étude s'est concentré sur les similitudes des choix opérés mais ce qui frappe aussi dans ses résultats est leur diversité. Celle-ci a depuis été analysée plus en profondeur au moyen d'un dispositif expérimental un peu différent. En effet, il était difficile d'étudier la variabilité entre les individus en ne posant essentiellement qu'une seule question à chaque sujet comme Fechner l'avait fait. (En vérité il avait aussi demandé à certains sujets d'indiquer le rectangle qu'ils aimaient le moins.) Des travaux récents ont montré que la préférence pour le rectangle d'or cache en fait une grande diversité de choix individuels, voire que cette préférence collective n'existe pas. Des expériences ont été réalisées de nouveau sur des rectangles de différentes dimensions, mais présentés cette fois par paire aux sujets. Face à deux rectangles différents les sujets devaient indiquer leur préférence. Ceci a permis de collecter beaucoup de données, d'étudier la cohérence et la distribution des choix opérés par chaque individu mais aussi de contraster ces distributions : certains sujets préféraient systématiquement des formes dont les rapports sont proches de 1, c'est-à-dire proches du carré, d'autres des formes très allongées dont le rapport des dimensions est nettement supérieur à ϕ ; les derniers enfin choisissaient de préférence les rectangles dont les dimensions ont un rapport qui n'est pas très différent du nombre d'or mais ils n'étaient plus majoritaires comme dans l'expérience de Fechner. Et ces choix paraissent assez stables dans le temps car les expériences ont été répétées à différents intervalles temporels sans mettre en évidence des changements d'opinion des individus. Les résultats confirment ceux d'autres études similaires réalisées dans des conditions expérimentales proches. Le rôle prédominant de la section dorée paraîtrait donc être une légende plus qu'une réalité expérimentale et les préférences manifestées par les sujets ne sont pas homogènes. Des résultats similaires ont été observés en matière de couleur et de structures géométriques.

Les goûts seraient-ils donc purement subjectifs, et non sujets à discussion ? Les travaux dont je viens de rendre compte ne disent pas exactement cela. Les analyses réalisées à partir des données pour mieux comprendre la diversité des préférences manifestées révèlent des groupes de profils similaires, par exemple les individus qui préfèrent les formes plus carrées, ou ceux qui aiment les rectangles dont le rapport des dimensions (plus grand côté sur plus petit) est légèrement supérieur à ϕ (1.86). Il y a donc bien des préférences partagées, des groupes de sujets qui ont des préférences identiques. Et si la question n'est plus pourquoi le nombre d'or, elle devient pourquoi des rectangles très allongés, des rectangles dorés ou des carrés? De plus, l'étude commentée montre qu'il est possible d'approcher de manière quantitative les préférences, et donc de les discuter. Enfin, la diversité mise en évidence n'exclut pas des modes de fonctionnement similaires, l'existence d'invariants mais métissés par les réalités culturelles, les apprentissages.

Le nombre d'or admet plusieurs formulations mathématiques

Il est temps maintenant d'aborder les bases mathématiques du nombre d'or. J'ai donné sa valeur approchée — 1.618 — sans expliquer d'où elle venait. Il existe différentes manières de le faire et je vais en mentionner trois : une algébrique, une géométrique et une basée sur la convergence d'une suite très particulière de valeurs.

ϕ peut être présenté algébriquement comme une des solutions de l'équation du second degré :

$x^2 = x + 1$. Cette équation admet en effet deux solutions : la valeur exacte de $\phi = (1 + \sqrt{5}) / 2$ qui vaut 1,618... - et $(1 - \sqrt{5}) / 2$ (qui est au signe près l'inverse de ϕ). Il est piquant de noter que si $\phi^2 = 1 + \phi$, alors $\phi = \sqrt{1 + \phi}$. Il suffit de prendre la racine carrée des deux membres. Pourquoi ne pas aller plus loin et remplacer le deuxième ϕ de cette équation par $\sqrt{1 + \phi}$? Nous obtenons une équation un peu plus compliquée : $\phi = \sqrt{1 + \sqrt{1 + \phi}}$. En continuant, chaque fois que l'on rencontre ϕ on le remplace par $\sqrt{1 + \phi}$, on obtient la suite des égalités suivantes :

$$\phi = \sqrt{1 + \sqrt{1 + \sqrt{1 + \phi}}} = \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \phi}}}} = \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \phi}}}}} \dots$$

et ainsi à l'infini. Cette forme d'écriture de ϕ présente une belle régularité et ne contient que des 1. Déjà dans son écriture, ce nombre surprend.

Mais pourquoi cette équation? Une autre formulation permet de mieux comprendre comment cette équation est à la base de rapports harmonieux. En divisant les deux termes de l'équation par x , on obtient en effet: $x / 1 = (x + 1) / x$.

Si x représente une longueur de segment comme indiqué dans le graphique ci-dessous, l'équation exprime simplement que le rapport entre la grande partie du segment (x) et la petite partie (1) est le même qu'entre le tout ($x + 1$) et la grande partie (x). Cette subdivision est souvent appelée division en moyenne et extrême raison.

Les suites de Fibonacci convergent vers le nombre d'or

Le nombre d'or peut aussi être présenté comme la limite du rapport entre deux termes consécutifs d'une suite de Fibonacci. De quoi s'agit-il ? Une suite de Fibonacci se construit facilement. Choisissez arbitrairement les deux premiers termes de la suite. Généralement, on prend deux fois 1. Puis calculez chaque nouveau terme en faisant la somme des deux termes qui précèdent. Le troisième terme vaudra, dans notre exemple, 1+1, c'est-à-dire 2, le quatrième vaudra 1+2, donc 3, etc. Plus mathématiquement formulé, dans une suite de Fibonacci, chaque terme de rang n , noté u_n , vaudra la somme des deux termes précédents $u_{n-1} + u_{n-2}$. Ainsi, en ayant commencé par deux 1, on obtient la suite :

1 1 2 3 5 8 13 21 34 55 89 144 233 377 610 987 1597

Cette suite, qui commence par deux 1 et qui n'est en réalité qu'un cas particulier, est traditionnellement référée comme LA série de Fibonacci. Quel est le rapport avec le nombre d'or? ϕ apparaît en fait progressivement comme le rapport de deux valeurs consécutives de la suite :

1 (1/1) 2 (2/1) 1.5 (3/2) 1.66 (5/3) 1.6 (8/5) 1.625 (13/8) 1.615 (13/8) 1.615 (21/13) 1.619 (34/21) 1.617 (55/34) 1.618 (89/55) 1.618 (144/89) 1.618 (233/144)

Le lien avec la définition algébrique est facile à établir, une fois admis que la suite des rapports converge vers un nombre x . Comme $u_n = u_{n-1} + u_{n-2}$, on peut écrire

$$u_n / u_{n-1} = 1 + u_{n-2} / u_{n-1}$$

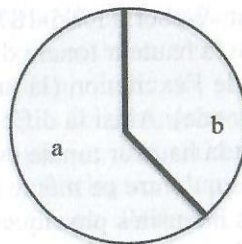
et en passant à la limite, $x = 1 + 1/x$ ou $x^2 = x + 1$ en notant que la limite x de (u_n / u_{n-1}) est la même que celle de (u_{n-1} / u_{n-2}) . Un des mérites de cette suite est d'avoir permis d'introduire des nombres entiers dans la définition du nombre d'or. Les artistes et quelquefois la nature préfèrent ces nombres aux irrationnels.

Les suites de Fibonacci se retrouvent dans la nature et dans les œuvres artistiques

Les premiers nombres de la série de Fibonacci se rencontrent sous diverses formes dans la nature. Les marguerites ont généralement 34, 55 ou 89 pétales, les écailles de pommes de pin ont 5 spirales dans un sens, et 8 dans l'autre; on retrouve des nombres consécutifs de la suite lorsqu'on compte les protubérances des ananas, les spirales constituées par les boutons de marguerite (ces nombres de spirales sont généralement de 21 dans un sens et 34 dans l'autre), ou les spirales situées dans le cœur des tournesols. La disposition des feuilles sur certaines tiges est aussi liée à cette suite qui apparaît aussi dans l'arbre généalogique des abeilles (le nombre de nœuds par niveau de l'ascendance d'une abeille mâle (un faux-bourdon) est toujours un terme de la suite de Fibonacci).

Ce petit jeu de recherche du nombre d'or ou de termes de la suite de Fibonacci peut ne jamais se terminer. Et certains ne se privent pas d'y jouer. Il suffit d'un peu d'esprit d'observation et d'un brin d'imagination. Restons raisonnables, il y a tellement de manières de regarder la nature qu'on peut toujours y retrouver ce qu'on cherche. Il importe de résister à la tendance de mettre au crédit du nombre d'or tous les phénomènes qui font intervenir soit un ou deux nombres de la suite traditionnelle de Fibonacci, soit un rapport de longueurs qui s'approchent plus ou moins de ϕ .

Les nombres dont je viens de parler ne constituent qu'une partie de l'arsenal des amis de ϕ . Pourquoi pas ne pas définir un angle d'or en partageant la circonférence d'un cercle en moyenne et extrême raison (voir la figure ci-dessous)? Cela a été fait bien sûr et l'angle ainsi défini vaut $137,5^\circ$, c'est-à-dire $360^\circ / \phi^2$ comme un calcul simple peut le montrer.



Définition de l'angle d'or $(a + b) / a = a / b$

D'autres figures mathématiques intéressantes sont liées au nombre d'or comme la spirale dorée et le pentagone.

Daniel Defays

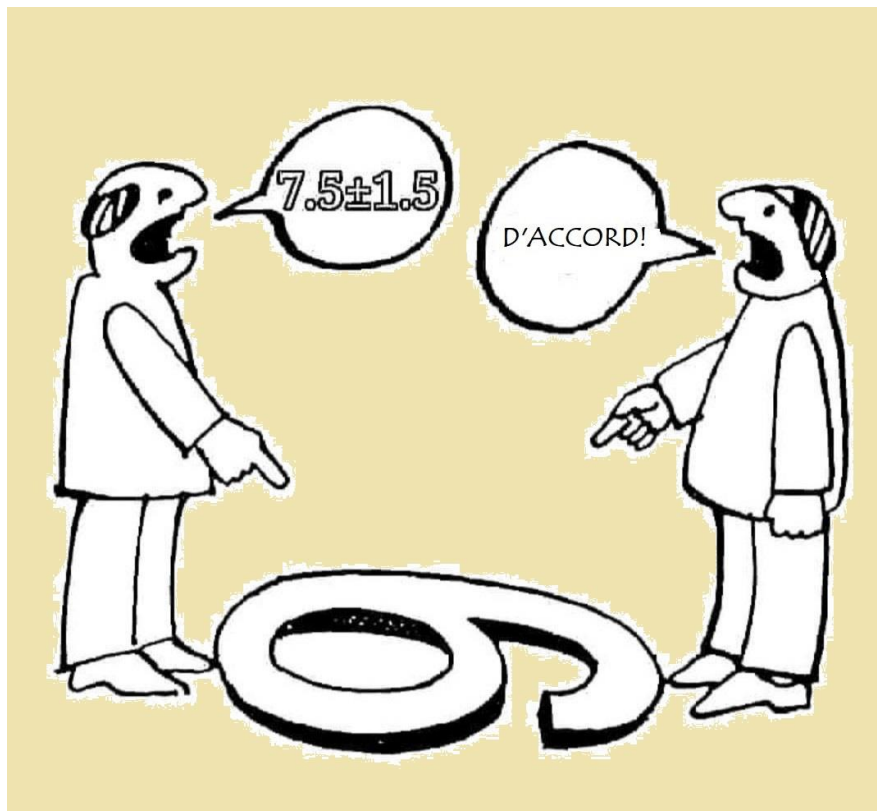
Références

Fechner, G. (1871). *Zür experimentellen Aesthetik*, Hirzel.

McManus, I., Cook, Ch.R., Hunt, A. (2010). Beyond the Golden Section and Normative Aesthetics : Why Do Individuals Differ so Much in Their Aesthetic Preferences for Rectangles? *Psychology of Aesthetics, Creativity, and the Arts*, Vol. 4, No. 2, pp. 113-126.

Warsufel, A. (2012). *Les nombres et leurs mystères*, Éditions du Seuil, Paris, France.

[Article extrait de « Là où les sens se rencontrent – HarmonieS 1 » de Daniel Dafays (Éditions l'Harmattan, 2020).]



Les Québécois veulent accélérer sur les autoroutes

Les Québécois sont en faveur d'une hausse de la limite de vitesse sur leurs autoroutes, selon un récent sondage. (Stéphane Lessard, Le Nouvelliste)

Les répondants se sont dit « plutôt » ou « très » favorables à cette idée dans une proportion de 68 %.

Le gouvernement québécois n'envisage pas d'adopter une telle loi, comme l'a fait l'Ontario avec son projet-pilote, en 2019, puis une loi permanente permettant de rouler à 110 km/h sur certains tronçons, en 2022.

Au Québec, aucune demande gouvernementale officielle n'a été faite par le ministère des Transports (MTQ).

« Le Québec voit bien ce qui se passe en Ontario, mais ce n'est pas dans les plans à court terme, dit le porte-parole du MTQ, Gilles Payer. Il n'y a pas de volonté officielle. On peut dire que la porte n'est pas fermée, mais qu'il n'y a pas de demande. »

Sondage

Selon le sondage SOM-Les Coops de l'information, réalisé au début du mois de novembre auprès de 1032 personnes, le groupe d'âge manifestant le plus d'appui à une telle mesure est celui des 55-64 ans, dans une proportion de 77 %.

Toutes générations confondues, 76 % des hommes voudraient que la limite soit de 110 km/h, tandis que les femmes appuient cette idée à 60 %.

Pour chaque groupe représenté dans le sondage en ligne, de 1 à 3 % des Québécois n'avaient pas d'opinion ou ont préféré ne pas répondre.

Parmi les autres tendances révélées à travers le sondage, c'est à l'extérieur de Montréal que les automobilistes sont le plus en faveur de la hausse de la limite permise. La métropole est d'accord à 65 %, alors que le reste de la province l'est à 70 %.

L'exemple ontarien

Dans l'ouest du Québec, l'autoroute 40 se métamorphose en autoroute 417, du côté ontarien.

Après Pointe-Fortune, la signalisation de la province voisine indique que la limite de vitesse permise est de 110 km/h.

La limite de vitesse sur l'autoroute 417, entre la frontière québécoise et Ottawa, est de 110 km/h. (Le Droit, Martin Roy)

En 2022, le gouvernement ontarien a haussé la limite de vitesse de façon permanente sur la 417, de la frontière québécoise jusqu'à Ottawa. Il s'agit du plus long tronçon de la province où il est officiellement permis de rouler à 110 km/h.

« Personne ne m’a parlé de problèmes de vitesse liés à cette nouvelle limite », dit le député provincial de Glengarry-Prescott-Russell, Stéphane Sarrazin. « Ça roule 120 sur la 417 et ça roule encore 120 en arrivant sur la 40! »

Le député franco-ontarien fait partie du gouvernement Ford, qui a amorcé le projet il y a trois ans. « Si ce n’était que de moi, j’instaurerais la limite de vitesse à 110 km/h partout en Ontario, dit-il, convaincu. La moyenne, de toute façon, c’est 110. Tout le monde le sait bien... »

Le député provincial dans Glengarry-Prescott-Russell, Stéphane Sarrazin, dresse un bilan positif du projet de son gouvernement de hausser la limite de vitesse à 110 km/h sur certains tronçons d'autoroute de l'Ontario. (Le Droit, Simon Séguin-Bertrand/Le Droit, Simon Séguin-Bertrand)

Le ministère des Transports de l’Ontario a brièvement indiqué par courriel que la modification de la limite permise n’avait pas aggravé le bilan routier, tant au chapitre des excès de vitesse qu’à celui des collisions. M. Sarrazin est aussi d’avis que la 417 n’est pas devenue le théâtre de grands excès de vitesse ou d’accidents graves.

« Il ne faut pas penser qu’on peut rouler 130 km/h dans les zones de 110, dit le député. C’est comme partout, disons. Ça tolère 120, mais pas plus... » — Stéphane Sarrazin.

Un autre sondage, effectué par le gouvernement de l’Ontario, en 2019, indiquait qu’un peu plus de 80 % des répondants étaient favorables à l’augmentation des limites de vitesse sur les autoroutes de la série 400.

« Les vitesses d’utilisation et les tendances en matière de collisions dans ces sections sont restées comparables à celles d’autres sections d’autoroutes similaires où la limite de vitesse est restée inchangée à 100 km/h », écrivait le ministère des Transports, dans son bilan du projet, en mars 2022.

D’autres provinces permettent déjà de rouler à 110 km/h sur certaines autoroutes. Il s’agit de l’Alberta, du Manitoba, du Nouveau-Brunswick, de la Nouvelle-Écosse, de la Saskatchewan et de la Colombie-Britannique.

Selon Stéphane Sarrazin, le gouvernement ne planifie pas, pour le moment, d’étendre à toutes les autoroutes du territoire la limite de 110 km/h.

En Ontario, la limite de vitesse de 110 km/h est affichée sur d’autres tronçons, plus courts, sur les 402, 417, 401, 404, 400 et 11.

Le CAA opposé

Le club automobile CAA-Québec est contre l’idée de voir la vitesse maximale établie à 110 km/h de l’autre côté de la rivière des Outaouais.

Son directeur des affaires publiques, Nicolas Ryan, s’est dit « surpris » par les résultats du sondage SOM-Les Coops de l’information.

« Qu’est-ce que les gens gagnent avec une limite de 110 au lieu de 100? » se pose-t-il comme question. Selon le CAA, les routes du Québec « ne sont pas faites » pour des vitesses supérieures.

« La configuration de nos autoroutes, leur état, les nids-de-poule, le risque d'augmentation des primes d'assurances. Nous rappelons constamment que la vitesse est directement liée à la force de l'impact en cas d'accident. Que gagne-t-on avec une limite à 110 km/h? Deux, trois minutes? »

La science

M. Payer précise que la science a aussi son mot à dire dans le débat. « Est-ce que la route est toute en courbes, combien de voies y a-t-il, quelle est la visibilité? Au Québec, le réseau a été conçu dans les années 70. »

« Le paysage est changeant, c'est boisé, les conditions sont changeantes. Ce sont des autoroutes qui existent depuis des dizaines d'années, sur la norme des 100 km/h »

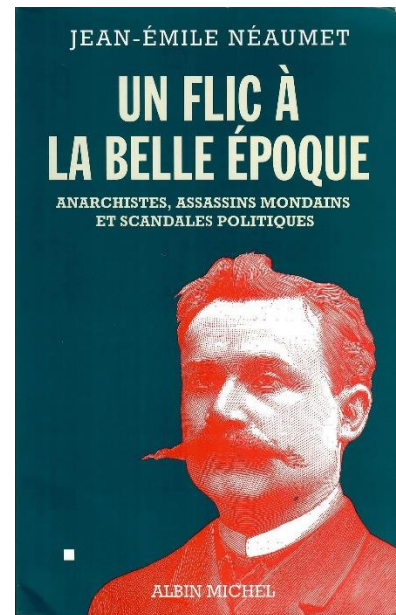
Louis-Denis Ebacher, *Les Coops de l'information*, 13 novembre 2023

Recensement de la peinture

Pour faire l'inventaire des œuvres d'art, on procède généralement simplement par dénombrement. Pour ce qui est de la peinture, on aurait pu procéder de la même méthode. Cependant, voici ce qui a été fait à Paris en 1892 :

La peinture, elle aussi, connaissait une santé sans précédent.

L'illustration notait à ce propos, se livrant au passage à cette pittoresque comptabilité: « Deux salons suffisent à peine à satisfaire la curiosité et, d'autre part, la statistique nous fournit des détails qui sont de nature à rassurer tous ceux que préoccupe l'avenir de l'art dans notre pays. Par exemple, en calculant la superficie des paysages soumis chaque année à la haute appréciation du public parisien, on trouve qu'ils représentent en moyenne une étendue de douze hectares, trois ares et cinq centiares par an. En comptant les expositions privées. Sans être aussi importantes, les scènes de guerre et d'histoire arrivent en seconde ligne avec cinq hectares et douze ares. Les natures mortes ne leur sont guère inférieures et le chiffre de quatre hectares et neuf ares prouve que l'art de représenter sur la toile des légumes, des fleurs, des fruits et de la viande de boucherie n'est pas près de tomber en désuétude. Si l'on ajoute deux hectares de portraits et un hectare et demi environ d'aquarelles, de pastels et de tableaux divers, on trouve le total imposant de vingt-cinq hectares, huit ares et cinq centiares de toiles artistiques par année. Ce qui est le maximum de peinture qu'ait jamais produit une nation civilisée. » Pas un centiare de plus ou de moins. Il est exact qu'alors la peinture française rayonnait de tous ses feux.



[Extrait de « Un flic à la Belle Époque » de Jean-Émile Néaumet, Éditions Albin Michel, 1998.]

Aide recherchée

Comme vous pouvez le constater dans chaque numéro, *Convergence* contient toutes sortes d'articles dont plusieurs ont été grappillés ici et là. La sélection de ces articles se fait au fil des jours et au hasard des lectures : un article, un passage dans un livre (roman ou autre...), une citation, un dessin humoristique, une caricature... Jusqu'ici, j'ai eu la chance de toujours trouver du contenu pour *Convergence*, mais si cela relevait quelques fois d'un blitz de dernière minute. Est-ce que cela va durer? Je l'espère, mais j'ai aussi peur d'un essoufflement de ma part. Alors pourquoi ne pas avoir un(e) assistant(e) pour m'aider au contenu de *Convergence*? Une tête, c'est bien, mais deux, ce serait bien mieux! Je cherche donc une personne curieuse et débrouillarde pour m'aider à dénicher des articles afin de meubler le contenu de *Convergence*. Ce n'est pas un boulot difficile. En fait, il s'agit seulement d'ouvrir l'œil lors des lectures de tout genre (journaux, magazines, romans, réseaux sociaux, etc.) et d'extraire le contenu pertinent et intéressant qui pourrait servir aux membres de l'association qui lisent le *Convergence*. Nous travaillerions alors ensemble pour décider de ce qui paraîtra dans le prochain numéro, ou de ce qui sera gardé pour les numéros subséquents. C'est tout, rien de plus. Es-tu intéressé(e) à me donner ce coup de main pour *Convergence*? Si oui, envoie-moi un courriel à plavall1962@gmail.com.

Merci d'avance!

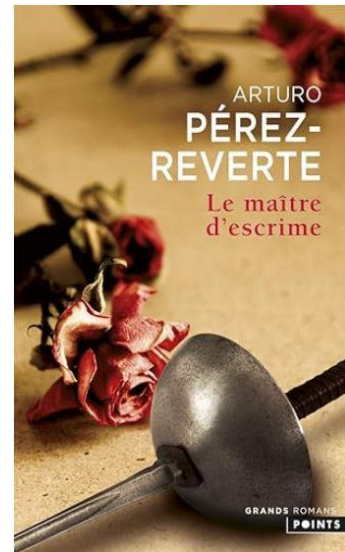
Ton rédacteur en chef

L'escrime : un sport ou une science exacte?

Mon père dit que l'escrime est bonne parce qu'elle est hygiénique, répondit courtoisement l'ainé des Carzola. C'est ce que les Anglais appellent sport.

Don Jaime regarda son élève comme s'il venait d'entendre un blasphème.

— Je ne doute pas que monsieur votre père ait des raisons pour affirmer une telle chose. Je n'en doute absolument pas. Mais moi je vous certifie que l'escrime est bien plus que cela. Elle constitue une science exacte, mathématique, dont la somme des facteurs conduit invariablement au même résultat : le triomphe ou l'échec, la vie ou la mort... Je ne suis pas ici avec vous pour que vous fassiez du sport, mais que vous appreniez une technique extrêmement pure qui, un jour, à l'appel de la patrie ou de l'honneur, pourra vous être très utile.



[Extrait de « Le maître d'escrime » d'Arturo Pérez-Reverte, Éditions du Seuil, 1994.]



LE DÉMON DU JEU

par GODI & JANSSENS

PROFESSEUR... COMMENT SE FAIT-IL QUE VOUS EN SACHEZ AUTANT SUR LE DÉMON DU JEU ?

OH, C'EST TRÈS SIMPLE, MON CHER RAOUL !

... MON PAUVRE PÈRE LUI-MÊME FUT SA VICTIME !

JE ME SOUVIENS D'AVOIR ASSISTÉ, IMPUISSANT, À LA DÉGRADATION FOLLE DE CET HOMME QUE JE VÉNÉRAIS POUR SON BON JENS ET SA FONDÉRATION

GN!

DIS-MOI, FERNAND... SI ON JOUAIT À DES JEUX ?

COMMENT ? MAIS ENFIN, PÈRE, VOUS N'Y SONGEZ PAS !

QUE DIRAIS-TU D'ESSAYER DE RETROUVER L'OMBRE DE CE CHAPEAU ? TU VAS VOIR, C'EST RIQUOIS COMME TOUT !

ET LÀ... SAURAI-TU REMETTRE DE L'ORDRE DANS CETTE VITRINE ? MARRANT, NON ?

1 BÉRET 2 HAIT-DE-FERME 3 MELON

4 PANAMA 5 CANOTIER 6 CABQUETTE

ALLONS, APPRENDS À T'AMUSER MON GARÇON...

JE REFUSE !

MAIS PUISQUE JE TE DIS QUE TA MÈRE N'EN SAURA RIEN !

3 COMBIEN DE PAIRES DE GANTS DANS CETTE BOÎTE ? DIS UN CHIFFRE, POUR VOIR ?

QUEL CHOC ÇA A DÙ ÊTRE !

HÉLAS ! VOUS COMPRENEZ MAINTENANT POURQUOI J'AI JURÉ D'ABATTRE CE FÉAU AVANT QU'IL NE DÉTRUISE COMPLÈTEMENT NOTRE SOCIÉTÉ.

HUM ! DONNEZ-MOI 12 GRILLES DE LOÏO, 5 * TAC-O-TOC, ET 3 * MILLIARDAIRES, S.V.P.

[Bande dessinée parue dans le 245^{ème} Album Spirou, Éditions Dupuis, 1998.]

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 : Audrey-Anne Vallée (Université Laval)

Trésorier : Jasmin Ricard (Centre de recherche CERVO)

Registraire : Johanne Morin (Beneva)

Directeur des communications : Simon Laflamme (Beneva)

Représentant étudiant : Élodie Bertin Comeau (Université Laval)

Membres institutionnels



Statistique
Canada

Statistics
Canada



CONVERGENCE

Convergence Vol XXIX No 1 Février 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 :

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>