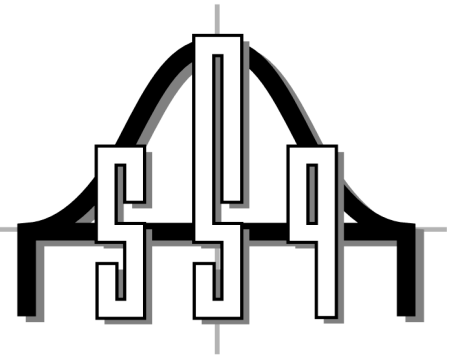


# Convergence

Le journal de l'Association des statisticiennes et statisticiens du Québec



Volume V • Numéro 1

Avril 2000

## Le mot du président

**B**onjour,  
Dans ce premier mot de l'année, permettez-moi de souhaiter la bienvenue à nos nouveaux membres et de réitérer mes remerciements à ceux qui nous appuient dans nos démarches en renouvelant leur cotisation. Vous avez d'ailleurs pu constater le fruit de nos efforts lors du premier trimestre. En effet, deux activités ont eu lieu avec un vif succès qui réanime notre désir de poursuivre notre mission de promotion et de perfectionnement tout en entretenant les liens entre nous. Et ça va continuer: plusieurs sessions sont prévues lors du congrès annuel à l'Université d'Ottawa et nous évaluons la possibilité de présenter des cours sur divers sujets. Évidemment, si vous avez des idées ou si vous voulez collaborer, nous sommes toujours ouverts à recevoir vos commentaires.

Dans la même ligne de pensée, on ne peut que se réjouir de la participation au sondage sur Convergence. Au nom du C.A., je vous remercie de votre participation qui illustre à nouveau un intérêt certain quant aux affaires de votre association. En ce qui a trait aux quelques commentaires faits sur le fonctionnement de l'ASSQ, j'ai regroupé nos réactions dans ce qui suit. Premièrement, nous allons continuer à favoriser la communication par courrier électronique, étant donné que la vaste majorité de nos membres y ont accès facilement. De plus, d'ici le mois prochain, Sylvain Végiard consacra plus de temps à notre site Internet à titre de webmestre. Nous espérons ainsi rendre notre site plus vivant et plus au fait des activités courantes de l'ASSQ. Finalement, l'heure de notre assemblée annuelle a été déplacée en après-midi lors du congrès annuel ce qui augmentera, on l'espère, la participation de nos membres à cette rencontre. Je profite de l'occasion pour vous rappeler que vous pouvez proposer des amendements à nos statuts. Vous devez nous aviser d'ici le 1er mai afin que nous puissions produire l'ordre du jour dans les délais requis.

Je termine en espérant vous revoir lors de notre rencontre annuelle du mardi 6 juin prochain à Ottawa ou lors d'une de nos prochaines activités.

À la prochaine!

Marc Duchesne, stat. ASSQ

### *Dans ce numéro:*

<b>Mot du rédacteur</b>	3
<b>À propos de l'ASSQ</b> (Diane Leroux, Nathalie Hamel et Daniel Proteau)	4
<b>Propos sur l'art, la créativité et la statistique</b> (Isme Alam)	6
<b>Le traitement des données manquantes</b> (Jean Hardy et Jacques Pagé)	7
<b>Un vote de confiance pour Convergence</b> (Denis Malo et Isabelle Marchand)	8
<b>Techniques d'enquête célèbre ses 25 ans d'existence</b> (Denis Lemire)	9
<b>L'erreur quadratique moyenne : l'erreur qu'a drastiquement fait la moyenne des gens</b> (Pierre Lavallée)	10
<b>La cryptographie Internet</b> (Thierry Petitjean-Roget)	12
<b>Info-média : La statistique pour comprendre le génie de Mozart</b>	13
<b>Les pots... et les fleurs!</b> (Denis Malo et Daniel Hurtubise)	14
<b>«Des lettres et des... lettres» Solution du jeu no. 3</b> (Paul Bergeron)	15
<b>Suivre son cours</b>	16

# CONVERGENCE

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

## Rédacteur en chef

Daniel Hurtubise, Statistique Canada

## Rédacteurs adjoints

Sylvie Gauthier, Statistique Canada

Pierre Lavallée, Statistique Canada

Denis Malo, Statistique Canada

Isabelle Marchand, Statistique Canada

Steve Méthot, Agriculture et Agroalimentaire Canada

Thierry Petitjean-Roget, IRSST

## 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 soumis sous forme de fichiers de traitement de texte Microsoft Word. Les auteurs doivent faire parvenir, à l'adresse de l'ASSQ, leur fichier sur disquette 3,5 pouces (ou un fichier convertible et lisible sur courriel électronique) ainsi qu'une copie papier de leur texte avant la date de tombée du prochain numéro. 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, à l'adresse de l'ASSQ, leur document prêt pour l'impression avant la date de tombée du prochain numéro. Les tarifs pour la parution dans un numéro de *Convergence* sont les suivants:

Tarifs	page intérieure	endos(publicité seul.)
Carte d'affaires	15 \$	20 \$
1/4 page	40 \$	50 \$
1/2 page	80 \$	100 \$
page entière	150 \$	200 \$

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

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

## Toute correspondance doit être adressée à:

### *Convergence*

Association des statisticiennes et statisticiens du Québec

Boîte postale 94

Loretteville (Québec), G2B 3W6, Canada

Fax/répondeur: (418) 871-1945

Adresse élect.: dleroux@mtq.gouv.qc.ca

Page internet: <http://www.assoc-stat.qc.ca>

Dépôt légal - Bibliothèque nationale du Québec, 1996

## MEMBRES INSTITUTIONNELS:



Statistique  
Canada



Département de mathématiques et de statistique



Statistical consultants

STATEX

Experts-conseils en statistique

## Mission

L'ASSQ a pour mission de regrouper les statisticiennes et les statisticiens de tous les domaines en vue de promouvoir la statistique et d'en favoriser la bonne utilisation.

## Membres

L'ASSQ offre deux catégories de membres aux personnes intéressées par ses activités:

**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 probabilité). Les personnes ne répondant 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.

Frais d'adhésion annuels pour chacune des deux catégories:  
50 \$ (régulier) 20 \$ (étudiant)

Les organismes peuvent devenir membres institutionnels de l'ASSQ au coût de 300 \$ par année et ainsi bénéficier de plusieurs privilèges dont l'adhésion gratuite comme membres statisticiens ou affiliés pour trois de leurs employés.

## Conseil d'administration

Président: Marc Duchesne (*Viasystems*)

Vice-président: Gilles Therrien (*SOM Inc.*)

Secrétaire: Diane Leroux (*Transports Québec*)

Trésorier: Daniel Proteau (*C.M.P. Ltée*)

Registraire: Nathalie Hamel (*Statistique Canada*)

Directeur des Communications:  
Nicolas de Kuffrin (*Reader's Digest Global*)

# Mot du rédacteur

Bonjour,

Volume 5! Nous voici à la 5<sup>e</sup> année de parution du périodique *Convergence*! 5 ans de *Convergence* signifie aussi que ça fait 5 ans que votre association existe (enfin un peu plus parce qu'il y a eu au préalable 2 ans avec un conseil d'administration provisoire). En abordant cette 5<sup>e</sup> année, nous avons voulu sonder le pouls de notre lectorat avec un sondage afin de réajuster le contenu et le format de notre périodique au désir de ses lecteurs. Isabelle Marchand et Denis Malo nous dressent un portrait des résultats de ce sondage. Nous vous remercions d'avoir pris le temps de nous répondre. Nous ferons tout notre possible pour répondre à vos attentes. Ces pages sont à vous, que vous désiriez nous faire part de vos questions, de vos commentaires ou nous soumettre un article. Votre participation est la bienvenue! Nous désirons souhaiter la bienvenue à deux nouveaux collaborateurs de *Convergence*, Isabelle Marchand et Thierry Petitjean-Roget, qui ont profité du sondage pour s'embarquer dans cette belle équipe. Merci!

Ce numéro en est un de sondage. En effet, outre les résultats de notre sondage, nous publions un article qui décrit le sondage réalisé par la société *National Geographic* sur son site web. Tenez-vous bien, il semble que la qualité dans les résultats soit assurée par la quantité des répondants plutôt que la qualité de la méthode de collecte. Dans le cadre de la chronique Internet, Thierry Petitjean-Roget nous parle de la sécurité des transactions sur Internet. D'un autre côté, ce numéro en est aussi un artistique. La chronique info-média nous parle d'un logiciel qui permet d'analyser statistiquement des pièces

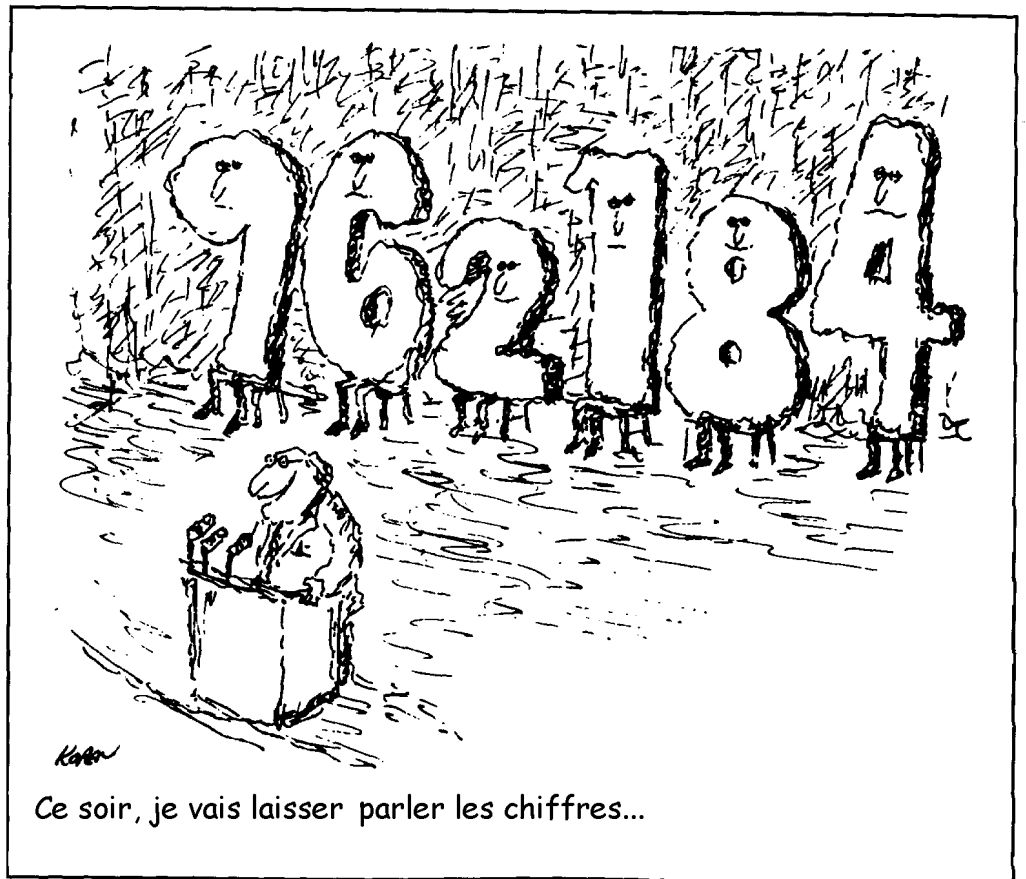
musicales, dont des œuvres de Mozart ou Bach. De même, Isme Alam de Statistique Canada nous

fait part de sa vision de l'apport de l'art lors de l'analyse des données. Laissons aller notre imagination pour analyser des données.

Avec le printemps arrive également les différents congrès. C'est une bonne façon de se retrouver et d'échanger sur nos différents travaux. L'ASSQ est également impliquée dans le congrès de la SSC, qui aura lieu à Ottawa au mois de juin. S'y tiendra aussi l'assemblée annuelle de votre association. Vous avez des questions, des propositions à faire? C'est le bon endroit pour en faire part à votre conseil d'administration.

Je vous souhaite une bonne lecture! ■

Daniel Hurtubise, [hurtdan@statcan.ca](mailto:hurtdan@statcan.ca)



Ce soir, je vais laisser parler les chiffres...

Tiré de Moore, D.S. (1998). *Statistics - Concepts and Controversies*. 4<sup>e</sup>ème édition, W.H. Freeman and Company, New York, 526 pages.

# À propos de l'ASSQ

Cette rubrique est préparée par les membres du Conseil d'administration et a pour but de fournir de l'information continue aux membres de l'ASSQ.

## Les échos du C.A.

**Diane Leroux, Secrétaire**

Depuis la dernière parution de *Convergence*, le conseil d'administration a tenu deux réunions, l'une le 3 décembre 1999 et la dernière, le 11 février 2000. Ces réunions ont surtout permis d'organiser les activités de visionnement de la vidéocassette et du déjeuner-conférence sur le « Data mining ». Puisque vous trouverez dans ce numéro de *Convergence* des détails sur ces activités, j'attirerai plutôt votre attention sur quelques autres points qui étaient aussi à l'ordre du jour de nos rencontres.

- Implication de l'ASSQ auprès des étudiants :
  - Grâce à la participation de la firme de sondage SOM, un étudiant en statistique de l'Université Laval aura la chance d'être initié au déroulement d'un sondage. Celui-ci nous fera part de son expérience par la rédaction d'un article dans un prochain numéro de *Convergence*;
  - L'ASSQ sera présente lors de la prochaine journée de la statistique organisée, le 31 mars 2000, par le Comité pour l'avancement de la statistique à l'Université Laval (CASUL);
  - La compagnie Logicarrières, qui produit des logiciels d'orientation scolaire et professionnelle, fournira maintenant les coordonnées de l'ASSQ comme source supplémentaire de renseignements sur la profession de statisticien.
- Bottin des membres

Puisque la majorité des membres de l'ASSQ ont accès au Web, nous avons convenu de ne pas rééditer sur papier le bottin des membres comme nous l'avions fait en 1998. Rappelons que la liste des membres est régulièrement mise à jour sur notre site Internet. Toutefois, ceux et celles qui voudraient en avoir une copie sur papier pourront en faire la demande auprès de notre registraire, Nathalie Hamel.

La prochaine réunion du C.A. est prévue pour le 28 avril 2000. ■

## Activités de l'ASSQ

**Diane Leroux, Secrétaire**

Notre première activité multi-régionale, le visionnement d'un vidéo sur la méta-analyse, a été un succès avec près de 60 participants. Bien que la participation n'a pas atteint les objectifs escomptés à Montréal et à Hull, le processus de tenir des activités a été démarré et nous a donné le goût de répéter l'expérience. D'ailleurs, les personnes présentes ont bien apprécié le contenu de la cassette qui dressait un historique de la méthode tout en soulignant les problèmes actuels. Et l'occasion était belle de se revoir tout en apprenant un peu plus sur ce sujet qui touche plusieurs domaines de recherche.

Notre deuxième activité, un déjeuner-conférence, s'est déroulé à l'Hotel Inter-Continental de Montréal, le 21 mars dernier. Notre conférencier-invité, Nicolas De Kufrin (qui est aussi le directeur des communications de l'ASSQ), a brossé un survol clair et précis du 'data mining'. L'assistance de 38 personnes (dont plusieurs non-membres) a bien apprécié l'effort de vulgarisation des méthodes d'analyse regroupées sous le chapeau de « data mining ». Nicolas a aussi profité de cette tribune pour souligner l'importance du rôle du statisticien même si les logiciels en vente offrent des programmes qui donnent des solutions 'automatiquement'. Les exemples pratiques d'utilisation de ces méthodes et les commentaires pertinents en ont satisfait plus d'un, tout en donnant le goût d'en savoir plus. La consultation effectuée après le déjeuner-conférence nous a permis de confirmer cette impression avec un taux global de satisfaction de 4,1 sur 5. De plus, les deux tiers des répondants souhaitent avoir un cours sur le sujet. Finalement, tous les commentaires reçus sont très positifs et nous encouragent à poursuivre dans cette voie. ■

## Campagne de renouvellement 2000

**Nathalie Hamel, Registraire**

Depuis janvier 2000, la campagne de renouvellement s'est amorcée. Cette campagne vise les personnes devenues membres de l'ASSQ avant le 31 août 1999 et pour lesquelles la carte de membre expirait le 31 décembre 1999. Les personnes qui ont adhéré à partir du 1<sup>er</sup> septembre 1999 ne sont pas considérées par cette campagne puisque leur cotisation est valide jusqu'au 31 décembre 2000.

En vertu de l'article 6 des statuts de l'association, tout membre ayant payé ses frais de cotisation pour l'année 1999 sera considéré en règle jusqu'au 30 avril 2000. Il importe donc d'effectuer son renouvellement avant cette

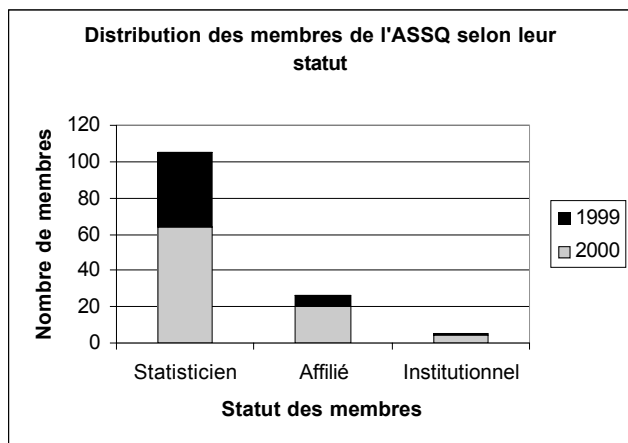
date pour s'assurer de conserver les avantages et de bénéficier des services offerts par l'association. Évidemment, toute cotisation reçue après cette date permettra la réintégration à l'association avec les mêmes privilèges.

Et maintenant, quelques statistiques... Au 31 décembre 1999, l'association comptait 136 membres. Sur ce nombre, l'association a reçu à ce jour (27 mars 2000) 88 renouvellements, soit 65% des membres potentiels, comparativement à 43% il y a un an. C'est une bonne amélioration. Merci! Pour les autres, **N'OUBLIEZ PAS DE RENOUELER LE PLUS RAPIDEMENT POSSIBLE!** Vous réduirez ainsi nos procédures de suivi auprès des membres de 1999 non renouvelés pour l'an 2000. Si vous avez égaré votre formulaire de renouvellement, n'hésitez pas à me contacter pour en obtenir une nouvelle copie.

Le tableau qui suit présente la distribution des membres de l'ASSQ en 1999 selon leur statut de renouvellement. La partie qui correspond à «2000» représente les membres de 1999 ayant renouvelé depuis le début de la campagne de renouvellement 2000 ou encore ceux qui avaient adhéré à l'ASSQ après le 1<sup>er</sup> septembre 1999. La partie qui correspond à «1999» représente les personnes qui n'ont pas encore renouvelé et dont on espère recevoir le formulaire de renouvellement sous peu!

Depuis janvier 2000, 14 nouveaux membres se sont joints à l'association, soient 2 membres affiliés et 12 membres statisticiens. Bienvenu à vous tous!

Un petit rappel aux membres affiliés... Si vous désirez obtenir le statut de membre «statisticien» au sein de l'association, vous devez présenter votre diplôme et/ou relevé de notes. Une deuxième façon d'obtenir ce statut est d'effectuer une demande d'accréditation. Pour de plus



amples renseignements, n'hésitez pas à me contacter par courriel à «[hamenat@statcan.ca](mailto:hamenat@statcan.ca)» ou par téléphone au (613)951-2071. ■

## Parlons Finances!

**Daniel Proteau, Trésorier**

Bonjour tout le monde! En bas de cette colonne, vous pourrez trouver le tableau qui résume les dépenses et les

revenus de notre association pour la période du 1<sup>er</sup> janvier 1999 au 31 décembre 1999.

Il faut remarquer que cette année, le budget est un peu plus élaboré. J'ai inclus un peu plus de détail pour un peu plus de clarté et de transparence.

Vous constaterez au tableau 1 que nous avons dépensé 5308.07\$. Notre plus grosse dépense (1529.90\$) a été faite pour l'impression de 4 parutions de Convergence, dont celle de décembre 1998. J'en profite pour féliciter l'équipe de Convergence qui fait un excellent travail qui trop souvent passe dans l'ombre. Le recueil des présentations au congrès de Sherbrooke de 1998 vient en deuxième place avec 862.69\$

Du côté des revenus, cette année n'a pas été différente des années passées. Comme à l'habitude, la très grande majorité des revenus proviennent des renouvellements individuels. En fait, cette catégorie comprend plus de 77% de nos recettes. Je remercie ceux qui ont déjà renouvelé pour l'an 2000. Au deuxième rang, nous retrouvons nos 5 membres institutionnels: Institut de la statistique du Québec, SOM Inc., STATEX, Statistique Canada et l'Université Laval (département de mathématiques et de statistique). Au total, les revenus pour 1999 se chiffrent à 6894\$. ■

Tableau 1

Encaisse le 1 <sup>er</sup> janvier 99	\$ 2983.21	
	1 <sup>er</sup> janvier 99 au 31 décembre 99	
	Dépenses	Revenus
Convergence (12/98, 04/99, 08/99, 12/99)	\$ 1529.90	
Recueil des présentations Sherbrooke 1998	\$ 862.69	
Réunions du CA	\$ 584.09	
Papeterie (enveloppes, étiquettes, timbres, etc.)	\$ 568.25	
Envois postaux de Convergence (12/98, 04/99, 08/99, 12/99)	\$ 551.21	
Recueil des présentations ACFAS 1999	\$ 308.75	
Télécopieur ASSQ	\$ 274.83	
Cassettes vidéos ASA (R&D)	\$ 215.20	
Boîte postale	\$ 82.82	
Frais bancaires	\$ 77.80	
Golf	\$ 68.00	
Envoi postal des renouvellements 99	\$ 55.78	
Écran de veille ASSQ	\$ 50.00	
Association de recherche en communication	\$ 46.75	
IGIF	\$ 32.00	
<b>Total (Dépenses)</b>	<b>\$ 5308.07</b>	
Renouvellement (1999)		\$ 5360.00
Membres institutionnels		\$ 1500.00
Objets promotionnels		\$ 34.00
	<b>Total (Revenus)</b>	<b>\$ 6894.00</b>
<b>Encaisse le 31 décembre 99</b>	<b>\$ 4569.14</b>	

## Arts et statistique :

# Propos sur l'art, la créativité et la statistique

## Isme Alam, Statistique Canada

*(N.D.L.R. : Isme a exposé ses créations artistiques pour stimuler la réflexion sur l'utilité et la pertinence de l'art en statistique. Lors de l'exposition d'art et d'artisanat de Statistique Canada à l'automne 1999, elle rédigeait un rapport analytique pour la publication intitulée Le Canada, sa culture, son patrimoine et son identité : Perspective statistique, à Culture, tourisme et Centre des statistiques sur l'éducation. Actuellement, elle travaille à la Division des statistiques des entreprises.)*

L'art est une expression de la créativité. Il peut nous ramener à notre enfance, une période où rien que de regarder les couleurs de l'arc-en-ciel dans un paquet de crayons suscitait une grande joie en nous.

Mais une fois l'enfance finie, beaucoup d'entre nous ont remplacé leurs crayons, pinceaux et pastels par des stylos et des ordinateurs. Si nous sortions nos crayons maintenant, ce ne serait très souvent qu'à la demande de nos enfants. Pourquoi ça? Peut-être parce que nous hésitons à nous exprimer à travers l'art, croyant que la créativité n'a pas droit de cité dans le monde du travail, qu'elle ne fait pas bon ménage avec les statistiques qui, elle, font partie de notre vie quotidienne au bureau.

En fait, la nature de notre cerveau nous prédispose à être à la fois créatifs et analytiques.

### L'hémisphère droit dans l'art...

Chaque côté de notre cerveau est lié à la main du côté opposé. Certains disent que le travail accompli par la main gauche permet une plus grande expression des émotions, voire de l'inconscient, tandis que le travail fait par la main droite est le résultat d'une réflexion plus concrète. (Bien sûr, il faut des recherches plus poussées pour vérifier cette hypothèse.)

Bien que je sois naturellement droitier, mon travail artistique représente les expériences que je fais avec les deux mains — et avec les deux hémisphères du cerveau. Pouvez-vous deviner lesquelles ont été faites avec ma main droite et lesquelles ont été faites avec ma main gauche?

### ...et en statistique

Une bonne partie de notre travail consiste à planifier la collecte des données, phase pendant laquelle nous utilisons souvent le côté gauche du cerveau, c.-à-d. le côté où se trouve le centre des fonctions verbales, analytiques, rationnelles et logiques.

Toutefois, après la collecte, nous devons cerner les tendances qui se dégagent des données. Pour le faire — c.-à-d. pour trouver le récit dans les données —, nous faisons souvent usage de

l'hémisphère droit du cerveau, où se trouve le centre des fonctions non verbales, intuitives, holistiques et non rationnelles.

### Donner vie aux chiffres

Être créatif, c'est pouvoir interpréter les données de façon intéressante et révélatrice.

Les gens créatifs reconnaissent toujours les différences entre les deux processus de collecte et d'interprétation créative des données. Apprendre à connaître les deux côtés de votre cerveau est un pas important dans la libération de votre potentiel créatif.

Une fois ce potentiel exploité, nous pouvons en imprégner nos statistiques et leur donner vie! ■

Quand les statistiques ne sont pas strictement basées sur des calculs précis, elles induisent en erreur au lieu de guider. La pensée se laisse facilement mener par les fausses apparences d'exactitude que les statistiques possèdent, et adopte en confiance les erreurs présentées sous forme de vérité mathématique.

Alexis de Tocqueville,  
*Democracy in America*

## Chronique SAS :

# Le traitement des données manquantes

*Jean Hardy et Jacques Pagé, Services Conseils Hardy*

Un des atouts importants d'un logiciel tel que SAS est sa capacité à gérer les données manquantes. Il suffit d'avoir cherché à analyser des données d'enquêtes avec Excel ou avec un système de gestion de bases de données pour comprendre à quel point cette caractéristique est importante. Rappelons que la gestion des données manquantes préoccupe l'analyste d'affaires ou le spécialiste du marketing autant que le sondeur.

Par exemple, une analyse du volume d'appels interurbains non facturés aux clients doit distinguer notamment les appels non complétés (mais qui ont tout de même entraînés des "coûts cachés"), les appels crédités en raison d'une communication médiocre, les appels faits à partir d'installations ou de cartes volées, les appels "gratuits" en raison d'un bogue du logiciel de facturation, etc. Se contenter d'assigner un code de données manquantes, ou pire la simple valeur 0, à tous ces appels sans distinction nous prive d'une information capitale.

Dans les faits, SAS est en mesure de manipuler plusieurs codes de données manquantes. Supposons les données de la Figure 1, représentant les revenus (non pas les profits) en milliers de dollars pour plusieurs sociétés.

Figure 1 - Contenu de la table SAS MAST

OBS_ID	REVENUS
10100	22,000
10103	-998
10106	51,000
10108	.
10234	37,500
10665	-997
11054	0
11067	-998
12093	-999

Le • représente un donnée non disponible tandis que le zéro représente des revenus nuls. Les valeurs négatives représentent ici divers types de données manquantes:

-997 → Société privée  
-998 → Refus de répondre  
-999 → Société inopérante

La première étape consiste à recodifier les données négatives en divers codes de données manquantes: SAS permet en effet de représenter les manquantes par le symbole • et par les symboles .A à .Z, soit un total de 27 codes distincts:

```
DATA EXTR;  
SET MAST;  
IF REVENUS=-997 THEN REVENUS=.P;  
ELSE IF REVENUS=-998 THEN REVENUS=.R;  
ELSE IF REVENUS=-999 THEN REVENUS=.I;  
RUN;
```

Ensuite, il s'agit d'assigner un format qui identifie ces valeurs (et qui regroupe les autres en catégories au besoin, au lieu d'employer un énoncé IF!):

```
PROC FORMAT;  
VALUE REVF  
LOW-<35000="Moins de 35 M$"   
35000-HIGH="35M$ et plus"   
  .="Non-disponible"   
  .P="Société privée"   
  .R="Refus de répondre"   
  .I="Société inopérante";  
RUN;
```

Si une liste de fréquences est commandée, les résultats ressembleront à la Figure 2:

```
PROC FREQ DATA=EXTR;  
TABLES REVENUS / NOCUM;  
FORMAT REVENUS REVF.;  
RUN;
```

Figure 2 - Fréquences sans données manquantes

REVENUS	Frequency	Percent
Moins de 35M\$	2	50.0
35M\$ et plus	2	50.0

Frequency Missing = 5

Comme d'habitude, l'ajout du paramètre **MISSING** dans **TABLES** force l'inclusion des valeurs manquantes comme s'il s'agissait de valeurs valides (Figure 3).

Figure 3 - Fréquences avec données manquantes

REVENUS	Frequency	Percent
Non disponible	1	11.1
Refus de répondre	2	22.2
Société inopérante	1	11.1
Société privée	1	11.1
Moins de 35M\$	2	22.2
35M\$ et plus	2	22.2

Les autres procédures SAS (**MEANS**, **SUMMARY**, **TABULATE**, **REG**, etc.) traitent les données manquantes de la même façon. Il faut toutefois se rappeler que lors de l'emploi des énoncés **IF** et **WHERE**, certaines formulations ne toucheront que les valeurs contenant un • (deux premiers énoncés ci-contre), tandis que d'autres affecteront toutes les données manquantes, quelle qu'en soit la définition (trois derniers énoncés ci-contre):

```
IF REVENU=.;  
WHERE REVENU=.;  
IF REVENU IN (.,.A,.B,.C);  
WHERE REVENU IS NULL;  
WHERE REVENU IS MISSING; ■
```

# Un vote de confiance pour **Convergence**

**Denis Malo et Isabelle Marchand, pour l'équipe de rédaction**

**V**oici les résultats du petit sondage qui était joint au dernier envoi du Convergence et qui voulait mesurer votre degré de satisfaction envers le journal de l'ASSQ. D'abord, le nombre de questionnaires reçus s'élève à 78, soit un taux de réponse de 60,5%. Parmi les répondants, les étudiants et les enseignants sont sous-représentés par rapport à leur proportion dans les rangs de l'ASSQ et qu'en contrepartie, les professionnels sont sur-représentés. Nous vous remercions pour votre collaboration et vos nombreux commentaires, que nous essaierons de prendre en considération dans la mesure du possible. Il va s'en dire que votre aide est précieuse pour réaliser les suggestions que vous nous avez faites. D'ailleurs, un résumé des commentaires et suggestions est incorporé à cette analyse.

La première partie du questionnaire portait sur l'utilisation du journal. Une forte majorité de répondants disent lire le Convergence dans sa presque totalité. En effet, 29,9% le lisent intégralement et 51,9% lisent la majorité des articles. Quant aux autres lecteurs (18,2%), ils disent lire seulement quelques articles reliés à leur domaine. Après une première lecture, 67,5% des membres conservent le journal pour consultation en cas de besoin, 15,6% en disposent à leur guise, et 13,0% le conservent et le prêtent à des collègues. Finalement, 24,7% des lecteurs ont déjà consulté le Convergence pour tenter de solutionner un problème.

La seconde partie traitait plus spécifiquement du contenu de la publication. Parmi les rubriques qui sont traitées régulièrement, voici les cinq plus populaires en ordre de préférence: les chroniques SAS, la série d'articles traitant de vulgarisation (Comment expliquez-vous...?), les annonces au sujet de l'ASSQ, les chroniques historiques et les profils de carrière. De la même façon, voici une liste de cinq chroniques que vous voudriez voir plus souvent dans le Convergence : des outils pouvant aider le travail du statisticien, le courrier des membres, une critique de logiciel, l'enseignement des statistiques et des nouveautés informatiques. Dans cette optique, vous nous avez clairement indiqué qu'une plus grande diversité d'auteurs était souhaitable. En effet, 70,1% des répondants ont indiqué une réponse en ce sens. Votre participation est donc la bienvenue, nous vous attendons.

Nous avons reçu peu de suggestions concernant le contenu de Convergence. Les commentaires obtenus proposaient une augmentation du contenu utilitaire comme par exemple des techniques et trucs pour solutionner des problèmes concrets. Bien que quelques chroniques déjà en place répondent à ce besoin (chroniques SAS, comment expliquez-vous ?, Internet – sites d'organisations statistiques) le contenu utilitaire pourrait en effet être augmenté quelque peu. Naturellement, ceci serait facilité par un plus grand

nombre de collaborateurs. Vos articles sont donc les bienvenus, nous les attendons.

La mise en page du journal plaît grandement à 94,8% des répondants. Certains trouvent la mise en page trop dense. Malheureusement, une aération du contenu du journal signifie une augmentation du nombre de pages, d'où une augmentation des coûts de publication.

Deux questions portaient sur la diffusion du journal. D'abord, les dates actuelles de publication satisfont 92,2% des lecteurs. Plusieurs souhaiteraient voir la fréquence de diffusion augmentée à quatre fois par année plutôt que trois. En effet, 35,1% des répondants ont indiqué une réponse en ce sens. Évidemment le choix de passer de trois à quatre numéros par année a des conséquences directes sur les coûts de publications, demande une participation de plus de membres pour répartir le travail et aurait un impact sur le contenu. L'équipe du Convergence entend poursuivre avec trois publications pour l'année en cours mais considère la possibilité d'augmenter cette fréquence, tout en évitant de diluer le contenu pour avoir un numéro de plus et d'augmenter la charge de travail des bénévoles de l'équipe. Donc, pour avoir un numéro de plus par année, une plus grande participation des membres au journal est requise. Votre participation est donc la bienvenue, nous vous attendons.

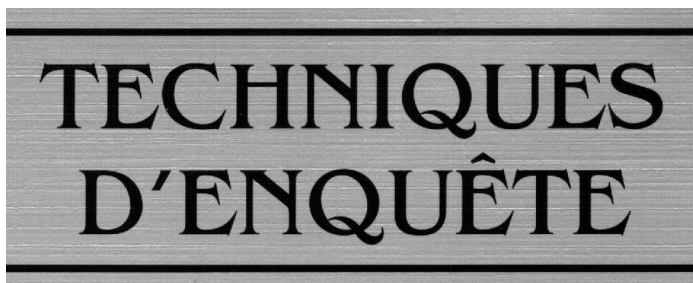


Nous avons également abordé quelques sujets reliés à l'Association elle-même. D'abord, on y a appris que 92,2% des répondants ont accès au site Internet de l'ASSQ, 76,6% l'ont déjà consulté et que 18,2% avaient des commentaires à son sujet. Les commentaires reçus portaient principalement sur une amélioration souhaitée au niveau de la mise à jour du site que l'on veut plus fréquente. Enfin, 94,8% des répondants sont satisfaits du contenu de l'article 'À propos de l'ASSQ' qui, entre autres, résume les activités du conseil d'administration. Il semble donc que l'article soit un excellent moyen de transmettre ce genre d'information aux membres.

Beaucoup d'encouragements et de messages positifs. En particulier, la chronique de SAS semble bien appréciée. La plupart des lecteurs aiment et veulent plus de contenu statistique. Bien que le but premier de *Convergence* ne soit pas de devenir un recueil d'articles statistiques mais plutôt de favoriser la communication et de circuler certaines informations parmi ceux qui sont touchés de près ou de loin par le monde de la statistique, une attention particulière pourrait être apportée au contenu utilitaire des rubriques présentes ou lorsque de nouvelles rubriques verront le jour.

Donc, en résumé, le journal semble plaire à la majorité des membres et plusieurs souhaite-

raient augmenter la fréquence et la qualité du contenu à un niveau statistique supérieur. Comme il a été mentionné plus haut, le but de ce journal est la communication entre les membres sans devenir une publication scientifique. Le journal est un moyen d'aider l'Association à atteindre ses objectifs de regroupement des statisticiennes et statisticiens, de même que de favoriser une meilleure utilisation de la statistique. Le journal est ouvert aux membres pour leur permettre d'échanger sur des sujets statistiques en dehors des revues spécialisées scientifiques. Donc, les pages de ce journal vous sont grandes ouvertes. ■



célèbre ses 25 années d'existence

**Denis Lemire, Statistique Canada.**

Avec son numéro de décembre 1999, *Techniques d'enquête* referme un cycle de 25 années de publication. Afin de souligner l'occasion, nous avons décidé de produire un numéro spécial. On y trouve des articles sollicités auprès d'éminents statisticiens d'enquête sur un certain nombre de thèmes actuels en méthodologie susceptibles d'intéresser leurs collègues de partout dans le monde. Devant la réaction quasi unanime de nos collaborateurs, notre prochain numéro de juin 2000 poursuivra sur le même thème.

*Techniques d'enquête* publie des articles abordant divers aspects du développement de la statistique pertinents pour un bureau de statistique. En plus d'articles d'intérêt général, *Techniques d'enquête* présente des sections spéciales qui permettent un traitement plus approfondi des techniques et expériences nouvelles sur des sujets choisis.

La revue est semestrielle. Tous les articles sont soumis à un comité de lecture anonyme et publiés dans les deux langues officielles. Les auteurs sont invités à présenter leurs articles dans la langue de leur choix. Les textes doivent être soumis en quatre exemplaires à M.P. Singh, rédacteur en chef, *Techniques d'enquête*, Direction de la méthodologie, Statistique Canada, Ottawa, Ontario, Canada K1A 0T6 (courriel : [singhmp@statcan.ca](mailto:singhmp@statcan.ca)). Les articles qui traitent de divers aspects du développement de la statistique et de leurs applications sont toujours les bienvenus. Veuillez communiquer avec le rédacteur en chef pour obtenir plus de renseignements.

Les membres de l'Association des statisticiennes et des statisticiens du Québec sont invités à s'abonner à *Techniques d'enquête* en même temps qu'ils renouvellent leur cotisation annuelle. Le coût de l'abonnement est de 30\$. Les membres de l'ASSQ bénéficient donc d'un rabais de 17\$.

## Courrier des lecteurs :

# L'erreur quadratique moyenne : l'erreur qu'a drastiquement fait la moyenne des gens

**Pierre Lavallée, Statistique Canada**

Il y a quelques mois, je me rendais à une conférence où l'on discutait, entre autres, de l'estimateur de James-Stein. Cet estimateur est suggéré dans le cas suivant. Supposons que l'on désire estimer un certain total  $Y_d$  pour un certain domaine d'intérêt  $d$ . Par exemple, les domaines d'intérêt pourraient être les régions touristiques du Québec. Pour ce faire, on tire un échantillon d'où on obtient une estimation  $\hat{Y}_d$  sans biais de  $Y_d$ . Malheureusement, il s'avère que la taille de l'échantillon du domaine  $d$  est trop faible et ainsi la variance de  $\hat{Y}_d$  est trop grande. On désire alors améliorer cette estimation. James et Stein (1961) nous proposent alors de choisir une valeur  $Y_d^0$  en guise de valeur supposée de  $Y_d$ . L'estimateur résultant de James-Stein est

$$\hat{Y}_{JS,d} = \alpha \hat{Y}_d + (1 - \alpha) Y_d^0$$

où  $\alpha \in [0,1]$ . On peut déterminer une valeur optimale  $\alpha_{opt}$  pour laquelle l'erreur quadratique moyenne  $EQM(\hat{Y}_{JS})$  est minimale. Avec la valeur optimale  $\alpha_{opt}$ , et si le nombre de domaines est égal ou supérieur à trois, on peut démontrer que :

$$EQM(\hat{Y}_{JS,d}) \leq EQM(\hat{Y}_d).$$

Ce résultat est pour le moins puissant. Il veut dire qu'en utilisant n'importe quelle valeur supposée  $Y_d^0$  de  $Y_d$ , on peut obtenir une estimation avec une erreur quadratique moyenne (EQM) inférieure ou égale à l'estimation de départ  $\hat{Y}_d$ . C'est puissant, mais pourtant il y a quelque chose qui cloche. Ce quelque chose vient du fait qu'on ne regarde que l'EQM et non pas les composantes de cette dernière.

L'EQM est constituée de deux composantes, soit le biais  $B$  et la variance  $V$ . On a en fait

$$EQM = B^2 + V.$$

L'EQM d'un estimateur est donc composée de deux sources dont seule la seconde s'estime généralement sans biais. En effet, si on possède un estimateur  $\hat{B}$  sans biais pour  $B$ , on peut toujours constituer un nouvel estimateur sans biais en

soustrayant le biais de l'estimateur initial. Par exemple, pour  $\hat{Y}_{JS,d}$ , on pourrait construire l'estimateur  $\hat{Y}_{JS,d}^*$  sans biais :

$$\hat{Y}_{JS,d}^* = \hat{Y}_{JS,d} - \hat{B}_{JS,d}.$$

Malheureusement, en général, on ne possède pas d'estimateur sans biais du biais  $B$  et l'estimateur  $\hat{Y}_{JS,d}^*$  est donc rarement possible.

On ne dispose habituellement que d'une estimation sans biais (ou approximativement sans biais) de la variance  $V$ . Comme on ne peut estimer sans biais la valeur de  $B$  pour avoir une idée de l'EQM associée à l'estimateur, on ne peut que se fier à l'estimation  $\hat{V}$  de  $V$ . L'estimation  $\hat{V}$  est donc directement utilisée pour l'estimation de l'EQM. Si le ratio  $B^2/V$  (ou  $B/\sqrt{V}$ ) est élevé, ceci produira une sous-estimation de l'EQM et provoquera une distorsion de l'intervalle de confiance.

Pour illustrer le sérieux du problème, prenons un cas extrême. On demande à un statisticien d'estimer le nombre de fumeurs dans la région de Montréal au Québec. Sans aucune mesure ou sondage, ce dernier répond : « 500,000 ». Cette "estimation" possède une variance nulle. En effet, pour n'importe quelle mesure, le statisticien aurait répondu la même réponse. L'estimation donnée possède cependant un biais non nul qui peut être élevé. Son EQM correspond ici simplement au carré du biais puisque la variance est nulle. Pour mesurer la précision de cet estimateur, comme on ne peut estimer que la variance, on aura alors une EQM estimée de zéro. L'estimation donnée du nombre de fumeurs sera donc faussement considérée comme très précise (et même exacte!), même si elle peut être tout à fait à côté de la valeur réelle. On voit donc le problème de considérer l'EQM globalement sans tenir compte de l'importance relative de ses deux composantes, soit le biais et la variance. Pour ne pas avoir trop de problème avec la sous-estimation de l'EQM, Cochran (1977) suggère que le ratio  $B/\sqrt{V}$  ne dépasse pas 10%. Sinon, on obtient une distorsion trop grande de l'intervalle de confiance.

Revenons maintenant à l'estimateur de James-Stein. En supposant la valeur  $Y_d^0$ , on ne fait que donner une valeur qui n'a pas de variance, mais qui peut avoir un énorme biais. Bien qu'on peut démontrer que  $EQM(\hat{Y}_{JS,d}) \leq EQM(\hat{Y}_d)$ , notre situation n'est pas vraiment mieux si le biais de l'estimation  $\hat{Y}_{JS,d}$  est trop grand. Il peut en effet être préférable de choisir  $\hat{Y}_d$  au lieu de  $\hat{Y}_{JS,d}$  parce qu'avec l'estimateur  $\hat{Y}_d$ , on est au moins assuré d'avoir un faible biais par rapport à la variance et d'avoir ainsi une EQM estimée plus proche de la réalité.

En conclusion, retenons qu'une petite EQM n'est pas nécessairement une bonne chose si le biais de

l'estimateur s'avère relativement grand par rapport à la variance. Malheureusement, choisir la plus faible EQM est le choix que fait drastiquement la moyenne des gens...

### Références

Cochran, W,G, (1977), « *Sampling Techniques* ». John Wiley and Sons, New York, 1977.

James, W., Stein, C. (1961), « Estimation With Quadratic Loss », *Proceedings of the Fourth Berkeley Symposium of Mathematical Statistics and Probability*, Vol. 1, University of California Press, Berkeley, pp. 361-379. ■

**Vous cherchez  
de la formation SAS®**

LES SERVICES CONSEILS  
**HARDY**

- ✓ offerte par des professionnels,
- ✓ en séminaires publics ou dans votre organisation,
- ✓ adaptée à vos besoins,
- ✓ neutre et appuyée par 30 ans d'expérience,
- ✓ articulée autour de 25 cours distincts?



**CONTACTEZ-NOUS!**

Tél.: (514) 866-0871  
(418) 626-1666  
Fax: (418) 626-2097  
[www.schardy.qc.ca](http://www.schardy.qc.ca)

4715, des Replats, bureau 260  
Québec Qc G2J 1B8

® SAS est une marque de commerce déposée de SAS Institute Inc., Cary, N.C.

# La cryptographie Internet

*Thierry Petitjean-Roget, IRSST*

**J**e me suis souvent fait demander : « *les transactions dites «sécuritaires» sur Internet le sont-elles vraiment ?* » la réponse habituelle « ça dépend » ne suffisant généralement pas, j'ai dû me résoudre à fournir les explications ci-dessous.

La cryptographie, ou chiffrement, est l'action de transformer une information claire, compréhensible de tout le monde, en une information chiffrée, incompréhensible pour les non-initiés. Le chiffrement est toujours associé au déchiffrement, l'action inverse.

Dès le 5<sup>ème</sup> siècle avant Jésus Christ, on retrouve des procédés de cryptographie, l'utilisation de codes secrets, chez les Grecs. Le premier procédé de chiffrement par substitution est proposé par Polybe, écrivain grec ; un exemple de codage par cette méthode est un texte de César à Cicéron.

Cependant, la cryptographie moderne, telle qu'utilisée à travers les réseaux informatiques, découle de la puissance de calcul des processeurs modernes et de la représentation numérique de l'alphabet (codes ASCII, EBCDIC, ...). Elle est de deux types :

L'algorithme D.E.S., pour « Data Encryption Standard », est un algorithme très répandu créé à l'origine par IBM en 1977 ; D.E.S. utilise une clé secrète de 64 bits, qu'il transforme en 16 "sous-clés" de 48 bits chacune, le cryptage se déroulant sur 19 étapes. Le nombre de combinaisons possibles est de l'ordre de 72 quadrillions de combinaisons.

Le hic ? La clé secrète est dite symétrique : elle sert autant au codage qu'au déchiffrement par algorithme inverse ; elle doit être connue des deux correspondants.

Sans entrer dans les détails, cette méthode d'encryptage, adoptée par le ministère de la défense des Etats-Unis, simple à programmer, est très utilisée en réseaux privés, le procédé étant directement implanté dans les machines inter reliées ( ex : quichet automatique).

La deuxième méthode est actuellement utilisée par la plupart des fureteurs. L'algorithme R.S.A., du nom de ses inventeurs Rivest-Shamir-Adleman (1978), est en partie basé sur le calcul modulaire et sur le petit théorème de Fermat (1601-1665) : « Soient  $p$  un nombre premier et  $a$  dans  $N^*$ . Alors  $p$  divise  $a^p - a$ . » En fait, on choisit un très grand nombre  $n$ , produit de 2 nombres premiers  $n=pq$  ; on détermine ensuite 2 nombres  $d$  et  $e$ , tel que :

$$de = 1 \text{ mod } (p-1)(q-1).$$

La clé de codage est alors le couple  $(n, e)$  et le codage d'un entier  $x < n$  est le nombre  $y = x^e \text{ ( mod } n)$  ; le décodage se fait à l'aide de la clé  $(n, d)$  en calculant  $x = y^d \text{ ( mod } n)$ . La clé de codage de l'individu X est appelée clé publique et est mise à la disposition de tous ceux qui veulent communiquer avec lui ; elle ne peut qu'encoder l'information. Seul X peut lire le message car il est le seul à connaître la clé de décryptage, dite clé privée. Le secret dans ce codage, dit asymétrique, vient du fait que le couple  $(n, e)$  ne permet pas de trouver dans un laps de temps raisonnable la clé privée  $d$  de décodage avec les moyens actuels de calcul, si  $n$  est très grand ; on parle actuellement de codage à 128 bits ( à titre indicatif  $2^{32} = 4\ 294\ 967\ 296$  ), et rien n'empêche d'utiliser des nombres encore plus grands, à 512 et même à 1024 bits. À titre indicatif, il a fallu pour factoriser un nombre de 129 digits ( 426 bits) en mars 1994 les ressources de 1600 ordinateurs pendant plus de 8 mois, soit l'équivalent de 4 à 6000 années/MIPS (Million d'opérations par seconde).

Pour parfaire la sécurité de la communication sur Internet, le protocole SSL ( Secure Socket Layer) s'exécute de la façon suivante :

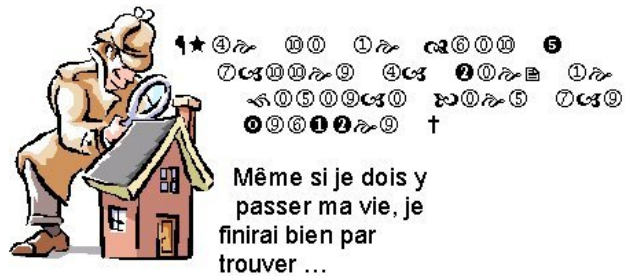
- Une fois la sécurité activée, le fureteur client génère une paire de clés publique/privée.
- Il demande au serveur sa clé publique.
- Il encrypte sa clé publique avec la clé publique du serveur avant de la lui transmettre.

L'échange de clés achevé, le transfert sécurisé de l'information peut commencer.

Les avantages : la clé publique du client n'est connue que du serveur et une nouvelle paire de clés

publique/privée est générée à chaque session, ce qui minimise encore plus le risque de décryptage frauduleux ; pas besoin non plus de retenir un « mot de passe ».

En guise de conclusion, que l'on se rassure: les données transmises en mode sécurisé à travers l'Internet sont difficilement accessibles aux intrus ! Mais quel que soit le processus de sécurité utilisé, la fraude est toujours possible au-delà de la transaction, que ce soit de la part du marchand ou du client !



Même si je dois y  
passer ma vie, je  
finirai bien par  
trouver ...

Pour ceux qui voudront en savoir plus sur le sujet que ce trop bref résumé, nous conseillons les lectures suivantes :

Livre : Singh, Simon, *Histoire des codes secrets*, JC Lattès, Paris, 1999.

Sites Internet :

- « Le paiement électronique sur Internet » <http://rambit.qc.ca/plamondon/crypto.htm>
- « FIPS PUB 46-2 - Announcing the DATA ENCRYPTION STANDARD (DES) » <http://security.isu.edu/isl/fips46-2.html>
- « RSA Laboratories » <http://www.rsa.com/rsalabs/>
- « PGP pour les francophones » <http://www.cl.cam.ac.uk/~fapp2/pgp/>
- « Pierre de FERMAT » [http://www.maths.tcd.ie/pub/HistMath/People/Fermat/RouseBall/RB\\_Fermat.html](http://www.maths.tcd.ie/pub/HistMath/People/Fermat/RouseBall/RB_Fermat.html) ■

# Info-média : La statistique pour comprendre le génie de Mozart

**La Presse, samedi le 4 mars 2000, page A26.**

« Faut-il disposer de puissants ordinateurs et maîtriser la statistique avancée pour comprendre le génie de Mozart, de Bach ou de Beethoven? » C'est en ces termes que le journaliste André Pratte, de La Presse, commence son article sur les travaux d'un professeur à l'Université de Zurich, Guerino Mazzola. M. Mazzola et son collègue, Jan Beran, ont mis au point une méthode qui permet d'analyser les partitions musicales. Un logiciel a même été écrit qui permet d'encoder numériquement des partitions. Ce logiciel permet d'attribuer un poids numérique à chaque note indiquant l'importance qu'elle a dans la structure rythmique, harmonique et mélodique de la pièce. Chaque note est analysée par rapport à ses voisines sur la portée et aussi par rapport à l'ensemble du morceau. Le journaliste semble peu convaincu de l'approche scientifique de l'analyse de la musique, disant de la statistique qu'elle est froide et bêtement rationnelle, la musique étant quant à elle chaleureuse et suprêmement émotive.

Un article scientifique a été écrit par MM. Mazzola et Beran sur ce sujet et a été publié dans *Statistical Science*.

Référence : J. Beran & G. Mazzola, *Analyzing musical structure and performance – A statistical approach*, *Statistical Science*, vol. 14, no. 1, 1999. ■

# Les pots...

**Denis Malo, Statistique Canada**

« Nous avons reçu plus de 50 000 réponses, soit deux fois plus que le nombre minimal requis pour atteindre la validité scientifique... »

Tels sont les propos tirés du site Internet de la populaire revue *National Geographic* (<http://survey2000.nationalgeographic.com>).

Cette société a mené une importante enquête sur une période de dix semaines à l'automne de 1998 (Survey 2000). Dans son édition de décembre 1999, ils y présentent les résultats et y vont de conclusions plutôt douteuses étant donné les méthodes d'enquêtes employées.

Il faut préciser que cette consultation, dont la seule méthode de collecte consistait à l'utilisation du site Internet de l'organisation, avait comme objectif de mesurer la mobilité des gens, de recueillir des informations au sujet de leur identité culturelle, de leur utilisation d'Internet et de divers comportements de la vie contemporaine. La sélection de l'échantillon se déroulait de façon volontaire et n'avait donc aucune composante aléatoire. Les écoles, centres communautaires et les bibliothèques offraient un accès au site tandis que les journaux et magazines publicisaient cette enquête.

Étant donné ce qui précède, d'importantes lacunes se sont présentées dans l'analyse des résultats car aucune mention des limitations évidentes ne fut ajoutée. Pire, on annexe aux résultats une comparaison du profil des répondants de l'enquête au profil de la population adultes des États-Unis provenant de l'enquête sociale générale. Comme on pouvait s'y attendre, les répondants à l'enquête sont beaucoup plus éduqués que la population générale étant donné la corrélation entre cette variable et l'accès au réseau Internet. Autre exemple, la population noire n'est presque pas couverte par le sondage tandis qu'elle représente près de 10% de la popula-

tion adulte. N'y a-t-il donc pas de biais de couverture? En plus du problème précédent, rien n'indique si des mesures ont été mises en place pour éviter qu'un répondant ne puisse avoir accès au questionnaire plus d'une fois...

Nous vous invitons à consulter le site Internet du *National Geographic* pour plus de détails et à mettre la main sur l'édition de décembre 1999 du magazine. Vous comprendrez qu'il nous a été facile de décerner ce pot, surtout avec le genre d'affirmation qui débute le présent article et celle qui le conclut : «*Les répondants au sondage Survey 2000 ont participé volontairement et nécessitaient un accès à internet mais les informations détaillées qu'ils ont fournies ainsi que l'énorme taille d'échantillon offrent aux scientifiques des données d'une très grande valeur qui*

# et les fleurs!

*ne peuvent être trouvées nul part ailleurs* ». ■

**Daniel Hurtubise, Statistique Canada**

En contrepartie, une fleur est attribuée à l'émission Enjeux, de la Société Radio-Canada, diffusion du 7 mars dernier. Il est fait mention lors de l'émission qu'un sondage visant les adolescents a été réalisé sur internet. Le journaliste mentionne aussi que ce sondage n'a aucune valeur scientifique, qu'il ne sert qu'à obtenir un portrait des jeunes internautes. Bravo pour cette mention! ■

**Qui se gausserait d'une telle célébri-**



### Paul Bergeron, Transports Québec

- |               |                 |                 |                |            |              |
|---------------|-----------------|-----------------|----------------|------------|--------------|
| 1. PURISME    | 8. ORDINAL      | 15. PIFFER      | 22. COMMANDE   | 29. AN     | 36. DOUÉ     |
| 2. PROBABLE   | 9. PONDÉRER     | 16. FUTURES     | 23. PROGRAMME  | 30. EFFETS | 37. S.P.S.S. |
| 3. ANALYSE    | 10. INSTABILITÉ | 17. VALEUR      | 24. RÉSIDENT   | 31. QUEL   | 38. JOUAL    |
| 4. DÉDUCTIONS | 11. INÉGALITÉ   | 18. DOSSIERS    | 25. PASCAL     | 32. DUAL   | 39. NUIT     |
| 5. DISTANCE   | 12. AMASSER     | 19. FICHIERS    | 26. MILLÉNAIRE | 33. MODALE | 40. 1577     |
| 6. LOGARITHME | 13. INDÉCIS     | 20. RÉPERTOIRES | 27. SIÈCLE     | 34. NULLES | 41. 1578     |
| 7. NATURELS   | 14. SONDEURS    | 21. ALGORITHMES | 28. DÉCADE     | 35. EXAMEN | 42. KSI      |

#### Références:

L'histoire des mathématiques d'après 3 grands scientifiques, <http://www.netrover.com/~razvan/Gauss-histoire.html>

Statistical Methods For Business and Economics, édition révisée, Pfaffenberger, Roger C., Patterson, James H., Richard D. Irwin inc., Homewood, Illinois, Etats-Unis, 1981, 828 p.

**CARL FRIEDRICH GAUSS (1777 - 1855) ST**  
 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

**ATISTICIEN\*, ASTRONOME ET PHYSICIEN**  
 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58

**ALLEMAND. FAMEUSEMENT CONNU DANS LE**  
 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88

**DOMAINE PAR LA LOI NORMALE (LAPLACE**  
 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117

**- GAUSS), UNE DISTRIBUTION QU'IL A D**  
 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143

**ÉVELOPPÉE PAR LA MESURE DES ERREURS**  
 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173

**DE PRÉDICTIONS DES ORBITES PLANÉTAI**  
 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204

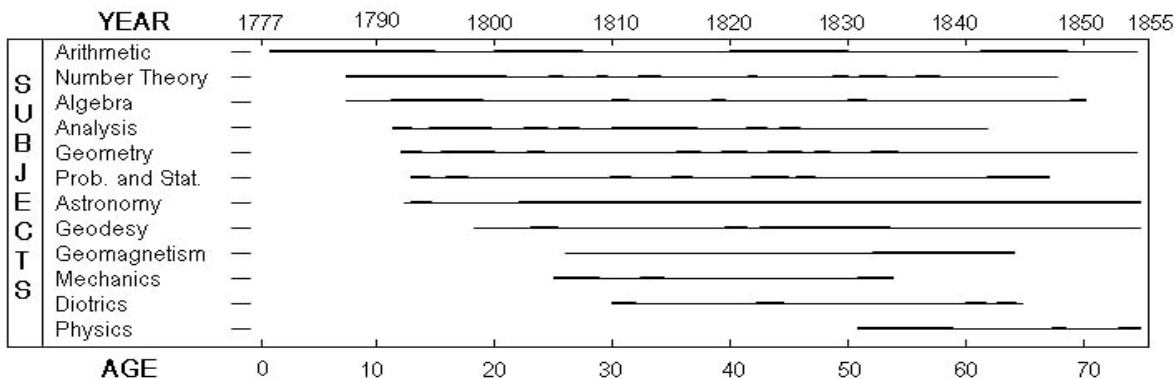
**RÉS. JEUNE, IL FORMULA LA MÉTHODE D**  
 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231

**ES MOINDRES CARRÉS. SON EFFIGIE FIG**  
 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260

**URE SUR LES BILLETS DE DIX MARKS.**  
 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286

\*: À la manière des moines copistes, nous avons pris la liberté de modifier les faits, en réalité il était plutôt mathématicien!

#### Graphique des réalisations de Carl Friedrich Gauss tout au long de sa vie





# Suivre son cours ...

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

Titre du cours	Lieu	Clientèle (4)	Durée totale	Date du cours / Inscription	Contact	But du cours
<b>Statistique Canada (5)</b>						
Modélisation et prévision ARIMA	STC, Ottawa	I et A	5 jours	27,29 novembre 2000+ 15-17 janvier 2001 / dès maintenant	(1)	Étudier les modèles ARIMA avec intervention pour décrire et prévoir des chroniques, calculer l'impact d'événements, tels que des données aberrantes et des changements de réglementation.
Les composantes des séries chronologiques	STC, Ottawa	I et A	3 jours	5-7 septembre / dès maintenant	(1)	Au terme de ce cours, le participant reconnaîtra, comprendra et interprétera les variations survenant dans les séries chronologiques. Il se sera aussi familiarisé avec la représentation graphique des données.
Conception de questionnaires	STC, Ottawa	D et I	3 jours	18-20 septembre 2000/ dès maintenant	(1)	Comprendre les concepts et les pratiques de base liés à la conception de questionnaires; concevoir des questionnaires qui servent à recueillir des données utiles et importantes.
Méthodes statistiques – Niveau avancé	STC, Ottawa	I et A	4 jours	4-8 décembre 2000 / dès maintenant	(1)	Approfondir les concepts d'inférence statistique de niveaux intermédiaire et avancé des points de vue théorique et pratique à l'aide de SAS.
Méthodes statistiques – Niveau approfondi	STC, Ottawa	I et A	4 jours	Dates non déterminées	(1)	Initier les participants aux méthodes statistiques nécessaires à l'analyse de données obtenues à l'aide de plans d'échantillonnage complexes.
<b>Services Conseils Hardy</b>						
SAS Version 8 (Nashville)	Montréal, Québec	D et I	1 jour	Printemps 2000 (consulter le site web pour les dates)/ dès maintenant	(3)	différents aspects des modifications apportées par la version 8 (disponible depuis janvier 2000) et couvrant notamment la méthodologie à suivre pour effectuer une migration sécuritaire de vos fichiers SAS.
Programmation SAS/IML	Montréal, Québec	D et I	1 jour	Printemps 2000 (consulter le site web pour les dates)/ dès maintenant	(3)	Cette session couvre en détail tous les énoncés et les fonctions de SAS/IML, le langage matriciel du système SAS.
<b>École de technologie supérieure</b>						
Maîtriser ses procédés : les méthodes Taguchi et traditionnelles	ETS, Montréal	D	3 jours	17-19 mai 2000 / dès maintenant	(2)	Planification d'expérience dans le domaine industriel.
Le contrôle statistique de la qualité	ETS, Montréal	D	2 jours	25-26 mai / dès maintenant	(2)	Contrôle statistique de la qualité (plans d'échantillonnage, CSP).

(1) Hew Gough (613) 951-3067 ou Céline Charrette (613) 951-1044

(2) Line Beauchamp (514) 396-8830 (Service de perfectionnement), site web, [www.perf.etsmtl.ca](http://www.perf.etsmtl.ca), courriel : [perf@etsmtl.ca](mailto:perf@etsmtl.ca)

(3) Monique Trempe (418) 626-1666, site web : <http://www.schardy.qc.ca>, courriel : [schardy@schardy.qc.ca](mailto:schardy@schardy.qc.ca)

(4) D : Débutant, I : Intermédiaire, A : Avancé (basé sur la matière du cours)

(5) Certains cours de Statistique Canada peuvent se donner à l'extérieur. De plus, le contenu peut varier selon les besoins exprimés. Certains cours sont donnés à raison d'une demi-journée par semaine pendant une certaine période pour permettre l'assimilation de la matière et l'essai des techniques étudiées. Les cours sont offerts en anglais et en français. Les coûts sont de 200 \$ par jour et par personne.

\* La grille précédente présente les cours offerts par différents organismes. La liste des cours universitaires est disponible sur demande à l'ASSQ. Notez que l'information fournie dans la grille des cours est sujette à changement. Le lecteur est invité à entrer en communication avec le ou la responsable des cours pour corroborer et compléter l'information présentée (p. ex.: frais d'inscription). Le lecteur pourra par le fait même faire ses propres démarches pour suivre le ou les cours de son choix.

## SEMINAIRES

Des séminaires ont lieu de façon régulière aux endroits suivants. N'hésitez pas à contacter le ou la responsable pour plus de détails.

<b>UQAM</b>	<b>Université Laval</b>	<b>Institut de la statistique du Québec</b>	<b>Université de Sherbrooke</b>	<b>Université de Montréal</b>
Pascale Rousseau Tél. (514) 987-3000, #3224 Fax (514) 987-8935 <a href="mailto:rousseau.pascale@uqam.ca">rousseau.pascale@uqam.ca</a>	Christian Genest Tél. (418) 656-5280 Fax (418) 656-2817 <a href="mailto:genest@mat.ulaval.ca">genest@mat.ulaval.ca</a>	Françoise Tarte Tél. (418) 691-2410 Fax (418) 643-4129 <a href="mailto:francoise.tarte@stat.gouv.qc.ca">francoise.tarte@stat.gouv.qc.ca</a>	Bernard Collin Tél. (819) 821-8000, #2012 Fax (819) 821-8200 <a href="mailto:bernard.colin@dm.usherb.ca">bernard.colin@dm.usherb.ca</a>	Christian Léger Tél : (514) 343-7824 Fax : (514) 343-5700 <a href="mailto:leger@dms.umontreal.ca">leger@dms.umontreal.ca</a>