

Convergence

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



Volume III • Numéro 2

Août 1998

Table des matières

Mot du rédacteur	P. 3
Le Billet : Interprétation trompeuse, dit Statistique Canada (Mike Sheridan)	P. 3
À propos de l'ASSQ (Nadia Ghazzali, Nathalie Hamel, Daniel Proteau)	P. 4
Les fleurs et les pots (Marc Duchesne)	P. 6
Chronique SAS : La fonction PUT, la clé des formats SAS (Jacques Pagé, Jean Hardy)	P. 7
Chronique Internet : les outils de recherche (Sylvain Végiard)	P. 8
Marketing et Institutions Financières : le rôle essentiel du statisticien (Nicolas De Kufrin)	P. 9
Erreur de probabilité	P. 10
«Comment expliquez-vous ... le coefficient de variation?» (Sylvie Gauthier, Denis Malo)	P. 11
Suivre son cours	P. 12

Le mot du président

Avons-nous les moyens de nos ambitions ? I. La loi des grands nombres

Si vous lisez ceci, c'est que vous faites probablement (un statisticien n'est jamais certain, ça c'est sûr !) partie de l'ASSQ. Avant d'adhérer à l'association ou de renouveler votre carte de membre, vous vous êtes possiblement interrogé sur les raisons qui vous ont incité à poser ce geste qui vous honore. Tout comme ceux et celles qui ont concrétisé il y a plus de trois ans ce projet de mettre sur pied l'association que l'on connaît aujourd'hui, vous nourrissez peut-être vous aussi le rêve de voir un jour la statistique occuper sa juste place dans le « concert des professions ». Mais...

Car il y a un mais... Pour que cet objectif à long terme se réalise dans un avenir prévisible, toute une série d'actions devra être posée, et l'ASSQ devrait normalement être au cœur de celles-ci, à la mesure de ses moyens. Quel pourrait être le rôle d'une association comme la nôtre dans l'atteinte de cet objectif ? Évidemment, si l'on s'assoit tous ensemble, je suis certain (sic !) que l'on établirait rapidement une longue liste d'actions toutes plus pertinentes les unes que les autres. Mais serait-ce là un plan réaliste ?

Dresser une liste d'actions à entreprendre est une chose, identifier des personnes prêtes à les réaliser en est une autre. Et que dire de veiller à ce que ces mêmes personnes effectuent dans les délais prescrits les actions qui leur ont été assignées ? Sans oublier que tout ceci se déroule dans un contexte de bénévolat et dans une société où vie professionnelle et familiale ne laissent plus beaucoup de place pour d'autres activités, ce qui implique une disponibilité limitée à quelques heures tout au plus par semaine.

Ce qui m'amène à poser cruellement la question : avons-nous les moyens de nos ambitions ?

(suite page 8)



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 et est distribué gratuitement aux membres de l'ASSQ.

Rédacteur en chef

Marc Duchesne, Viasystems

Rédacteurs adjoints

Sylvie Gauthier, Statistique Canada

Daniel Hurtubise, Statistique Canada

Pierre Lavallée, Statistique Canada

Denis Malo, Statistique Canada

Steve Méthot, Agriculture et Agroalimentaire Canada

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, soumission et résolution 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 courrier 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.: therrien@som-inc.com

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

Dépôt légal - Bibliothèque nationale du Québec, 1996

MEMBRES INSTITUTIONNELS:

bsq Bureau de
la statistique
du Québec

 Statistique
Canada

 UNIVERSITÉ
LAVAL

Département de mathématiques et de 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 souhaitant 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 500 \$ par année pour les institutions du secteur privé et de 300 \$ par année pour les institutions du secteur public et ainsi bénéficier de plusieurs privilèges dont l'adhésion gratuite comme membres statisticiens ou affiliés pour trois de leurs employés.

Conseil d'administration

Président:	Sylvain Végiard, <i>Ministère des ressources naturelles</i>
Vice-présidente:	Nadia Ghazzali, <i>Université Laval</i>
Secrétaire:	Gilles Therrien, <i>SOM Inc.</i>
Trésorier:	Daniel Proteau <i>C.M.P. Ltée.</i>
Directrice de l'admissibilité des membres:	Nathalie Hamel <i>Statistique Canada</i>
Directeur des Communications:	Pierre Lavallée <i>Statistique Canada</i>

Mot du rédacteur

Bonjour tout le monde,

Après les vacances, le retour de *Convergence*... (à chanter sur un air connu !). Plus sérieusement, dans cette édition, l'équipe poursuit ses efforts afin de vous offrir un contenu varié et enrichissant (on l'espère !). En regardant le contenu de ce deuxième numéro, je n'ai pas pu m'empêcher de prendre du recul et de constater les liens avec les objectifs de notre association.

Par exemple, nos chroniques Internet, SAS et 'Comment expliquez-vous' illustrent de belle façon notre objectif de développer un réseau d'aide, d'information et de conseils en statistique. Le 'Billet', de son côté, rejoint celui de sensibiliser au bon usage de la statistique. Quant à la page 'Suivre son cours', elle sert de rappel afin de nous inciter à la formation continue. Et, *Convergence* dans son ensemble constitue un bon moyen de favoriser les échanges entre les membres. J'espère que ce petit exercice vous fait réaliser, tout comme moi, que *Convergence* essaie de remplir sa mission (sans s'empêcher d'être divertissant !).

Vous avez d'autres idées d'articles ou de chroniques ? N'hésitez pas à me rejoindre par téléphone, courrier ou via le courrier électronique – marc.duchesne@viasystems.com. Pour notre édition prévue en décembre prochain, il n'est pas trop tard, loin de là : on devrait finaliser le contenu vers la mi-octobre. L'invitation vous est lancée !

Bonne lecture,

Marc Duchesne, rédacteur en chef



Le <<Billet >>

Interprétation trompeuse, dit Statistique Canada

Mike Sheridan, Statistique Canada

Mike Sheridan est directeur général de la Direction des enquêtes des ménages et du travail à Statistique Canada

Statistique Canada attire votre attention sur la conclusion inexacte d'une étude publiée par la Banque Canadienne Impériale de Commerce (CIBC). Le 20 mars, une dépêche de la *Presse Canadienne*, intitulée <<Le taux de chômage ne tient pas compte des jeunes>> mentionne l'étude en question dans laquelle la Banque soutient que les statistiques de chômage chez les 15-24 ans masquent 200 000 sans-emploi. Toutefois, cette interprétation est à la fois fautive et trompeuse. Regardons la situation de plus près.

Il est vrai qu'il existe environ 200 000 jeunes qui ne sont ni des employés, ni des étudiants, ni à la recherche d'un emploi. Il n'est pas vrai, ni raisonnable, par contre, de présumer que tous ou même la plupart sont, en effet, des sans-emploi.

Si ces jeunes gens, qui ne sont pas des étudiants, font partie en fait des <<chômeurs latents>>, pourquoi ne sont-ils pas à la recherche d'un emploi ? Vraisemblablement parce que, selon eux, il n'y en a pas. Cependant, dans ses enquêtes, Statistique Canada ne manque pas de fournir des estimations sur ces jeunes qu'il classe dans la catégorie des <<travailleurs découragés>>. Ces jeunes

sont au nombre de 11 000 environ et représentent 5,5% des 200 000 jeunes en question.

Près de la moitié du nombre qui reste, soit 72 000, sont des jeunes femmes qui ont au moins un enfant. Pour une raison ou pour une autre, elles ont choisi de ne pas faire partie de la population active et ont privilégié leur famille. Les classer dans la catégorie des sans-emploi serait nettement incorrect.

Les deux catégories susmentionnées représentent presque la moitié des 200 000 jeunes en question. Un nombre subséquent est en transition dans les écoles secondaires et post-secondaires.

Tout comme les jeunes mères, le choix de ne pas travailler ou de ne pas chercher un emploi leur appartient. Cependant, il n'y a aucun moyen de savoir leur nombre exact, mais ils ne représentent surtout pas tous les <<chômeurs latents>>.

Tirer des conclusions trop hâtives à propos des jeunes en transition est certes dangereux, mais il est vrai qu'il y a très peu d'éléments probants pour déterminer ce qui pousse les jeunes à faire de tels choix. Statistique Canada entreprendra bientôt plusieurs enquêtes pour faire la lumière sur la question.

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.

Congrès annuel 1998

Par Nadia Ghazzali

L'ASSQ a tenu deux sessions dans le cadre du XXVI^{ème} congrès annuel de la Société Statistique du Canada (SSC) qui a eu lieu à l'Université de Sherbrooke du 31 mai au 3 juin 1998.

La première session organisée et présidée par François Pageau (Les technologies industrielles SNC) portait sur la statistique en milieu industriel. Ainsi, Marc Duchesne (ViaSystems) a développé une expérience factorielle visant à améliorer un processus de placage du cuivre. Il a présenté une analyse des résultats ainsi que l'impact de ses conclusions sur les actions entreprises en fabrication. La conférence de Daniel Proteau (C.M.P. Metal Products) mettait l'accent sur l'importance à accorder à la prise en compte de l'erreur d'échantillonnage lorsqu'on estime l'aptitude d'un processus avec les indices d'aptitude usuels. Différents intervalles de confiance pour ces indices ont été discutés.

Pour clore la session, François Pageau a présenté le développement d'une carte de contrôle adaptée à des données provenant d'une distribution de Rayleigh. Cette carte est utilisée pour prévenir l'usure prématurée des outils de formage du métal. Mis à part l'aspect technique, ces trois conférenciers ont aussi souligné les contraintes pratiques propres au secteur manufacturier qui ont un impact sur l'intervention du statisticien.

La deuxième session organisée et présidée par Sylvain Végiard (Ministère des ressources naturelles du Québec) concernait la consultation statistique. Cet atelier a réuni trois panelistes qui proviennent de trois domaines différents : public, privé et universitaire. Il s'agit respectivement de Diane Leroux (Ministère des transports du Québec), de Michel Fluet (SOM recherches et sondages) et d'Hélène Crépeau (Service de consultation statistique, Université Laval).

Le but de cette session était de partager diverses expériences et méthodes de travail, et de confronter les visions de chacun des panelistes. Plus spécifiquement, la session a abordé les cinq thèmes suivants : compréhension du besoin du client, gestion de la relation avec le client, résolution de problèmes, présentation de la solution et satisfactions et frustrations.

Les organisateurs, ainsi que leurs conférenciers et panelistes, sont vivement remerciés pour leur contribution à la réalisation de ces deux sessions qui ont été fort

intéressantes. Sans oublier la participation de l'auditoire dont les échanges avec les conférenciers et les panelistes ont permis d'enrichir le contenu des thèmes abordés.

Ces deux sessions ont été suivies par l'assemblée annuelle de l'ASSQ. Trente membres sur un total de 117 membres en règle étaient présents (quorum atteint). La réunion a été assez brève et technique laissant peu de place aux discussions de fond sur les orientations de l'Association compte tenu des contraintes de temps pour pouvoir assister au banquet de la SSC.

La première partie de la réunion a été consacrée à la présentation de différents rapports annuels par l'exécutif tel que le rapport annuel des activités par le président et le rapport financier par le trésorier. Différentes propositions, nécessitant l'approbation des membres lors de l'assemblée annuelle, ont été adoptées durant la deuxième partie de la réunion. Ces propositions concernaient en grande partie des modifications aux statuts. Ainsi ont été adoptées les propositions suivantes :

- 1) la dispense de nomination d'un vérificateur externe pour l'année 1998 ;
- 2) le trésorier est désormais autorisé à signer seul les chèques de moins de 1000\$;
- 3) la durée des mandats des administrateurs est portée à deux ans plutôt qu'un an comme à l'heure actuelle ;
- 4) un vote postal peut être tenu pour modifier les statuts si le quorum n'est pas atteint lors d'une assemblée générale ;
- 5) le titre de directeur de l'admissibilité des membres est dorénavant remplacé par celui de registraire ;
- 6) l'adoption d'une nouvelle grille tarifaire pour les membres corporatifs et institutionnels, devant entrer en vigueur pour 1999.

Une autre proposition de modification aux statuts portait sur la création d'un nouveau type de membre, les <<associations affiliées>>, mais cette proposition a fait l'objet d'une motion de renvoi à l'an prochain et n'a donc pas été adoptée. Toutefois, les membres ont fortement encouragé l'exécutif à poursuivre ses démarches pour se rapprocher des autres grandes associations statistiques comme la SSC.

Finalement, Stéphane Tremblay (Statistique Canada) a proposé de former un comité sur la place des femmes au sein de l'ASSQ et a invité les personnes intéressées à entrer en contact avec lui pour mettre sur pied ce comité et à en définir la mission et les activités.

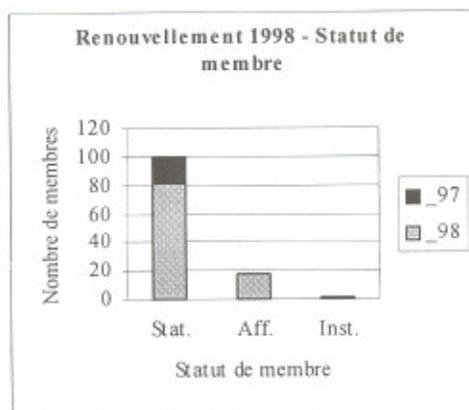


Campagne de renouvellement 1998

Par Nathalie Hamel

La campagne de renouvellement 1998 est pratiquement terminée. Rappelons que cette campagne visait toutes les personnes et institutions devenues membres de l'ASSQ jusqu'au 31 août 1997. On compte 100 membres ayant répondu favorablement à la campagne sur un potentiel de 120, soit un taux de renouvellement de 83%. De plus, 7 personnes ont réintégré les rangs de l'ASSQ après une absence d'au moins un an. Nous sommes très heureux de vous revoir !

Le graphique qui suit présente la distribution des membres de l'association de 1997 visés par la campagne de renouvellement 1998, selon leur statut.



La partie qui correspond à « _98 » représente les membres de 1997 qui ont renouvelé en 1998. La partie qui correspond à « _97 » représente les membres de 1997 n'ayant pas encore renouvelé pour 1998. Ils ont malheureusement perdu leur statut de membre en règle puisque celui-ci était effectif jusqu'au 30 avril 1998, selon les statuts de l'association. À noter qu'il n'est jamais trop tard pour renouveler son adhésion à l'ASSQ !

Depuis septembre 1997, l'association compte 3 nouveaux membres affiliés, 12 nouveaux membres statisticiens et un nouveau membre institutionnel, soit le Département de mathématiques et de statistique de l'Université Laval. Bienvenue à tous ces nouveaux membres!

À noter que les membres ayant adhéré à l'ASSQ entre septembre et décembre 1997 n'ont pas été touchés par la campagne de renouvellement de 1998. Ils ont conservé leur statut de membre pour l'année en cours.

Un gros merci au nom de l'ASSQ.

Parlons Finances !

Par Daniel Proteau

Bonjour tout le monde ! À la fin de cette colonne, vous pourrez voir un tableau reflétant les dépenses et les revenus de notre association pour la période du 1^{er} janvier 1998 au 30 juin 1998. Vous remarquerez que le poste des communications forme notre plus grosse dépense. La définition de ce poste comptable a été faite dans le *Convergence* d'avril 1998. L'impression du bottin des membres devient la deuxième dépense d'importance. Afin d'améliorer notre journal, nous avons acheté une imprimante, les logiciels d'Office 97 et Windows 95. Ces trois items se chiffrent à 912,91\$. Ensuite viennent l'impression de *Convergence* (décembre 97 et avril 98) et l'achat d'un MODEM. La création de notre nouveau site Web nous a coûté 120,78\$. Les frais bancaires terminent la liste pour un total de dépenses de \$4259,09.

Du côté des revenus, 95 membres réguliers et 15 étudiants ont renouvelé leur cotisation. On vous remercie vraiment beaucoup pour votre encouragement. Les membres institutionnels (Université Laval, Bureau de la Statistique du Québec et Statistique Canada) constituent la deuxième source de revenus. Lors du congrès de la SSC à Sherbrooke, nous avons vendu pour 133,00 \$ de casquettes et 't-shirts'. Un total de revenu pour cette période de 6083,00\$. Nous avons donc un montant de 3995,49\$ en banque le 30 juin 1998.

Tableau 1

Encaisse le 1er janvier 1998	\$ 2171,58	
	1 janvier 1998 au 30 juin 1998	
	Dépenses	Revenus
Communications	\$1165,21	
Bottin des membres	\$1035,23	
Imprimante, Office 97, Windows 95	\$ 912,91	
Convergence	\$ 792,14	
MODEM	\$ 142,63	
Site Web	\$ 120,78	
Frais Bancaires	\$ 90,19	
Total (Dépenses)	\$ 4259,09	
Renouvellement (1998) (15 étudiants et 95 réguliers)		\$ 5050,00
Membres institutionnels		\$ 900,00
t-shirts et casquettes		\$ 133,00
Total (Revenus)		\$ 6083,00
Encaisse le 30 juin 1998	\$ 3995,49	

Les fleurs ...



Par Marc Duchesne

Vous avez sûrement entendu ou lu les résultats d'un "sondage" téléphonique effectué par les médias sur une très courte période de temps. Finalement, le journal La Presse reconnaît que ces consultations n'ont rien d'une enquête bien structurée en ajoutant dorénavant une note au bas des résultats de leur 'Question du jour'. Celle-ci se lit comme suit : "Ce sondage maison sert uniquement à prendre le pouls de nos lecteurs et n'a aucune prétention scientifique." Bien que cette remarque modérée puisse être plus explicite, elle a au moins le mérite de distinguer ce genre de consultation d'un 'vrai' sondage et d'en reconnaître le caractère scientifique. Pour cet effort, décernons-leur une fleur, par le fait même, un pot pour tous les autres médias qui ignorent toujours cette différence.

et les pots

AVIS DE RECHERCHE

Par la présente, un avis de recherche est officiellement lancé pour retracer : articles de journaux, revues ou magazines présentant des déficiences (ou des opportunités d'amélioration !) au point de vue statistique.

S.V.P. veuillez envoyer vos trouvailles par courrier ou télécopieur au numéro de l'ASSQ ou à un membre du c.a. près de chez vous

P.S.: Aucune rançon n'est promise mais votre satisfaction personnelle de contribuer à la mission de l'ASSQ pourrait s'accroître...

NOVAXA

Optimisation des
processus d'affaires

Conseiller(ère) en statistique industrielle

NOVAXA Conseil offre des services de consultation en **productivité** et **qualité** auprès d'entreprises manufacturières d'envergure. La croissance de ses affaires l'amène à rechercher des candidats pour combler un poste de **conseiller(ère) – statisticien(ne)** ayant pour intérêt l'application des statistiques dans le milieu manufacturier (contrôle de la qualité, analyse des données, Six Sigma, etc.).

Qualifications requises :

- Détenir un baccalauréat en mathématiques ou statistique
- Détenir une maîtrise en statistique serait un atout
- Connaître les logiciels d'environnement Windows
- Être familier(e) avec des logiciels statistiques (Statistica, Minitab, SAS, etc.)
- Posséder d'excellentes aptitudes en communication verbale et écrite, en français et en anglais
- Avoir un intérêt marqué pour la consultation et une expérience pertinente en milieu manufacturier
- Faire preuve de polyvalence, de créativité et de flexibilité
- Être de nature organisée, structurée et rigoureuse
- Avoir un intérêt pour enseigner auprès de nos clients
- Être prêt à voyager

Nous offrons un environnement de travail des plus stimulants ainsi que des conditions d'emploi avantageuses.

Les personnes intéressées sont priées de faire parvenir leur curriculum vitae en spécifiant le poste et à l'attention de Cécile Berthier :

NOVAXA CONSEIL
4, Place du Commerce, Bur. 320
Île des Sœurs, Verdun
Québec, H3E 1J4

Téléphone : (514) 767-0070
Télécopieur : (514) 767-0050
E-mail : novaxa@novaxa.ca

Date limite pour postuler : 11 septembre 1998

CHRONIQUE SAS:

La fonction PUT, la clé des formats SAS

Jacques Pagé et Jean Hardy
Services Conseils Hardy Inc.

Utilisée avec un format numérique standard, la fonction **PUT** constitue la seule façon acceptable de convertir une variable numérique en variable caractère. Exemple : convertir la variable *ANNEE* en une variable caractère de 4 octets :

```
AN = PUT(ANNEE,4.)
```

La variable résultante *AN* peut ensuite être traitée correctement par une fonction caractère telle *SUBSTR* sans encombrer le journal de messages inutiles et sans ralentir l'exécution de l'étape *DATA*. À noter que la fonction **PUT** produit toujours un résultat de type caractère.

Lorsque conjointement employée avec un format personnalisé, la même fonction convertit encore une variable numérique en variable caractère, mais en appliquant les règles de transformation inhérentes au format utilisé. Cet usage transcende le format personnalisé SAS de simple format d'impression en outil de conversion ou de codification sophistiqué. Voyons un exemple : un format personnalisé, nommé *FMTAGE*, assigne un code de groupe d'âge à une valeur numérique entre 20 et 50 et un code X à toute autre valeur.

```
proc format;
  value fmtage 20-25 = "A"
              26-30 = "B"
              31-35 = "C"
              36-40 = "D"
              41-45 = "E"
              46-50 = "F"
              other = "X";
data out1;
  input ndossier $ age;
  grage = put(age,fmtage.);
  lines;
1234 22
2345 34
3456 43
4567 18
run;
proc print data=out1;
run;
```

OBS	NDOSSIER	AGE	GRAGE
1	1234	22	A
2	2345	34	C
3	3456	43	E
4	4567	18	X

Associée à un format SAS créé dynamiquement, la même fonction **PUT** remplacera avantageusement (exécution plus rapide) une fusion commandée par l'énoncé *MERGE* traditionnel, à condition de ne vouloir fusionner qu'une seule variable. Exemple : on doit associer le nom d'un professeur correspondant au numéro de professeur présent dans la table *SAS BIB1.RESULTAT*. Le numéro et le nom du professeur font partie d'une seconde table *SAS* nommée *BIB1.PROFCLA*.

```
data prof (keep=start label fmtname type);
  set bib1.prof_cla (keep=profess nom_prof);
  start = profess;
  label = nom_prof;
  fmtname = "FMTPROF";
  type = "N";
run;
proc format cntlin=prof;
run;
data resultat(keep=eleve profess nomprof);
  set bib1.resultat;
  nomprof = put(profess,fmtprof.);
run;
proc print data=resultat;
run;
```

OBS	ELEVE	PROFESS	NOMPROF
1	4	2	MAURICE
2	7	1	JEAN-MARC
3	8	2	MAURICE
4	9	8	JEAN-PIERRE
5	11	2	MAURICE
6	17	2	MAURICE
7	26	2	MAURICE
8	28	3	FRANCINE
9	29	3	FRANCINE
10	30	3	FRANCINE
11	41	9	ANNE-MARIE

Qui ne connaît pas l'utilité d'un format SAS personnalisé dans la définition de strates servant à contrôler le processus d'agrégation réalisé par certaines procédures telles que *TABULATE*, *FREQ* ou *SUMMARY* ? Le format démontré précédemment pourrait fort bien servir à une telle agrégation.

L'énoncé **FORMAT** permet donc d'associer un format SAS au processus d'agrégation des procédures sommaires de SAS sauf une ! Il s'agit de la procédure **SQL** qui utilise le paramètre **FORMAT=** plutôt que l'énoncé. Mais cette seule invocation d'un format ne produit pas les résultats escomptés. Voyons ce que produit *SQL* en présupant que notre table *SAS OUT1* contiendrait mille individus âgés entre 20 et 50 ans :

```
proc sql;
  select age format=fmtage.,
         count(*) as nombre
  from out1
  group by age;
```

AGE	NOMBRE
A	27
A	32
A	39
A	29
A	34
...	

Le format agit sur l'impression mais n'agit pas sur l'agrégation. Encore une fois, la fonction **PUT** doit être utilisée. Voici le programme corrigé :

```
proc sql;
  select put(age,fmtage.) as gr_age,
         count(*) as nombre
  from out1
  group by gr_age;
```

Le mot du président (suite)

Si on se compare avec d'autres associations professionnelles, force est d'admettre que nous sommes une petite association, certes, mais dont la proportion de ses membres «actifs», c'est-à-dire impliqués d'une manière ou d'une autre dans une activité quelconque, est remarquablement élevée (je la situe quelque part autour de 15%). Avec un tel taux d'implication, nous constituerions une véritable armée si notre association comptait 10 000 membres, ou encore une brigade entière si nous étions 1 000, mais admettons-le, une bien petite patrouille quand on ne regroupe qu'un peu plus d'une centaine de personnes. Avez-vous pensé un instant au budget et à la force de frappe que peut avoir par exemple l'Ordre des ingénieurs du Québec, avec ses quelque 40 000 membres et une cotisation annuelle de près de 200 dollars ? L'embauche de personnel permanent devient alors possible, et les activités de représentation (et de lobby, appelons les choses par leur nom) permettent une visibilité et un rayonnement qui ne nous est pas - et ne nous sera jamais - accessible.

Vue sous cet angle, la petite taille de notre association nous condamne presque comme membres à devoir s'impliquer activement dans celle-ci, sous peine de ne pas atteindre les grands objectifs qui nous sont si chers.

Vous trouvez mon propos trop défaitiste ? Possible. Mais rassurez-vous. Dans le prochain numéro de *Convergence*, vous pourrez lire la seconde partie de ce texte, qui proposera une vision plus positive sur les défis qui attendent notre association.

D'ici là, vous avez amplement le temps de vous interroger sur votre implication!

Sylvain Végiard, stat. ASSQ
président



Chronique Internet: les outils de recherche

par Sylvain Végiard (svegiard@mrn.gouv.qc.ca)

À peine quelques heures de navigation sur le Web suffisent pour prendre la mesure du volume colossal d'informations qui s'y trouve. Compte tenu de la nature non structurée de ces données (elles sont disséminées sur une multitude de serveurs aux quatre coins de la planète !), des outils de recherche ont été développés pour faciliter la localisation de renseignements portant sur un sujet particulier.

Voici les adresses de quelques outils ou moteurs de recherche (« search engines ») parmi les plus populaires :

- Alta Vista (<http://altavista.digital.com/>)
Probablement le plus gros répertoire général.
- Excite Netsearch (<http://www.excite.com/>)
Un répertoire général avec classification des sujets.
- Lycos (<http://www.lycos.com/>)
Le "petit frère" de Alta Vista.
- Yahoo (<http://www.yahoo.com/>)
Un des répertoires les plus fréquentés en raison de son système de classification par sujets.
- La toile du Québec (<http://www.toile.qc.ca/>)
Un excellent répertoire général des pages publiées au Québec.

Ces moteurs de recherche sont à peu près tous conçus de la même façon pour ce qui est de leur utilisation :

- une boîte de saisie permet de coder la requête que l'on veut effectuer ;
- il y a le plus souvent deux modes de recherche, un mode simplifié et un autre dit avancé, ce dernier étant utile lors de requêtes impliquant plusieurs critères ;
- il y a presque toujours de l'aide disponible (« search tips » ou « more tips ») pour chaque moteur qui décrit en détail la syntaxe et les opérateurs logiques à employer lors de requêtes plus complexes.

Conseils pratiques

Évitez de mettre dans votre requête des caractères accentués ou des majuscules. En effet, les accents ou majuscules sont restrictifs dans une requête. Donc, si vous voulez obtenir des informations sur le Québec, vous mettez toutes les chances de votre côté en codant "quebec", ce qui vous permettra d'intercepter la mention "Québec", mais aussi "Quebec", "québec" et "quebec".

Enfin, si vous voulez chercher une chaîne de caractères comprenant plusieurs mots, vous n'avez qu'à coder celle-ci entre guillemets, comme par exemple : "echantillonnage systematique".

Bonnes recherches !



Marketing & Institutions Financières: le rôle essentiel du statisticien

par Nicolas De Kufrin, Reader's Digest Canada

Pour la plupart des gens, le mot statistique signifie principalement sondages, moyennes, tableaux de données, etc. Cependant, plusieurs domaines d'application des diverses méthodes statistiques demeurent inconnus. À cet effet, il m'a semblé important d'écrire cet article afin de faire connaître un autre champ d'application des statistiques plutôt méconnu du public et des entreprises canadiennes et qui peut avoir un impact financier important pour ces dernières.

Que peut bien faire un statisticien au sein d'un département de marketing ou encore dans une institution bancaire offrant des cartes de crédit? Étant donné la compétitivité grandissante dans le secteur privé, les compagnies se doivent de cibler leurs marchés :

- Qui sont leurs clients ?
- Quel est leur profil ?
- Pourquoi ces derniers achètent-ils un certain produit plutôt qu'un produit similaire offert par un compétiteur?
- Comment acquérir de nouveaux clients ou encore augmenter sa part du marché ?
- Quel est le meilleur prix pour vendre un produit afin de maximiser le profit ?
- Quel est l'impact à long terme d'une nouvelle politique de prix ?
- Etc.

Les réponses à ces questions sont vitales pour toutes les compagnies. Malgré leur complexité apparente, un nombre appréciable de méthodes statistiques a été développé et nous permettent de répondre à ces questions. Voici quelques-unes de ces méthodes statistiques.

Modèles prédictifs (Predictive modeling)

L'utilisation de modèles prédictifs, telle la régression logistique, est essentielle dans plusieurs milieux, tels le marketing direct ainsi que les institutions financières offrant des cartes de crédit. Un modèle prédictif établi sur un échantillon de la population cible nous permet par la suite d'identifier à qui attribuer une carte de crédit, quelle limite de crédit donner ou encore à qui envoyer par la poste une offre d'achat pour un nouveau produit.

Segmentation du marché

Pour la plupart des entreprises, une segmentation de sa clientèle est primordiale afin de connaître les différents besoins de chaque client pour ainsi leur offrir le bon produit et maximiser les profits. À cet effet, les techniques telles l'analyse de facteur (*Factor Analysis*), l'analyse par

grappe (*Cluster Analysis*) de même que les arbres de régression (*Regression trees*) sont parmi les méthodes les plus populaires en ce qui a trait à la segmentation du marché.

L'Analyse conjointe (Conjoint Analysis)

Afin de comprendre les fondements de cette technique, considérons la situation suivante. On désire déterminer l'importance des trois attributs suivants : prix, couleur, offre promotionnelle, lors de l'achat d'une nouvelle voiture pour ainsi déterminer la combinaison préférée des consommateurs.

Prix de la voiture	Couleur	Offre promotionnelle
1- 10 000 \$	1- rouge	1- air climatisé gratuit
2- 20 000 \$	2- bleue	2- radiocassette inclus
3- 25 000 \$	3- verte	
4- 50 000 \$		

A l'aide d'un panel de répondants, on demande à ces derniers d'indiquer leurs préférences pour divers scénarios (ex.: voiture bleue de \$ 10 000 avec air climatisé gratuit). Grâce à ces informations, il nous est possible d'identifier les attributs les plus importants pour l'achat d'une voiture de même que la combinaison préférée. De façon générale, l'analyse conjointe nous permet d'estimer :

- 1- l'importance de chacun des attributs ainsi que leurs interactions,
- 2- quelle est la combinaison préférée des répondants,
- 3- quelle part du marché occuperait un produit spécifique.

Les méthodes présentes sont nées au début du siècle et sont utilisées depuis le début des années 60 dans la plupart des compagnies américaines. Ayant moi-même travaillé aux États-Unis, j'ai constaté que nos voisins du Sud ont quelques décennies d'avance sur le Canada en terme d'utilisation de la statistique dans le secteur privé. Malgré le fait que nous possédons des statisticiens compétents, il semble que le secteur privé soit réticent à engager ces derniers. Plus souvent qu'autrement, on engage des individus avec diverses formations ayant certaines connaissances limitées en analyse quantitative. Et malheureusement, dans beaucoup de cas, ces personnes ne sont pas préparées pour faire face aux problèmes spécifiques des entreprises. N'importe qui peut utiliser un logiciel d'analyse statistique et obtenir des résultats. Cependant, seul un statisticien peut conseiller sur la méthode optimale à utiliser dans un contexte particulier et ainsi assurer la validité des résultats. ■

Erreur de probabilité

Jusqu'au début de 1654, Antoine Gombaud, chevalier de Méré, gagnait très souvent aux tables de jeux. Puis la chance tourna... et, grâce à lui, la science naissante des calculs de probabilités trouva sa première application pratique.

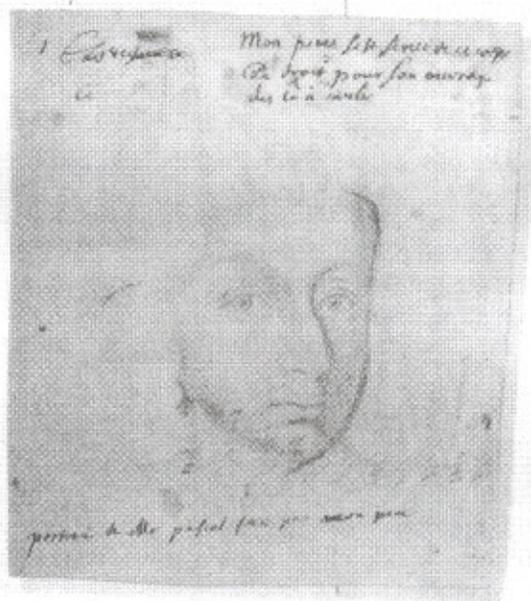
L'un des jeux préférés de Méré consistait à lancer un dé à quatre reprises ; pour gagner, il fallait faire sortir le six au moins une fois. Le chevalier réussissait si souvent que plus personne n'accepta bientôt de le défier. Alors Méré modifia le jeu : il fallut jouer avec deux dés et réussir à faire un double-six en vingt-quatre coups. Méré avait calculé qu'à ce nouveau jeu, il gagnerait deux fois sur trois.



Fermat

Ce jeu remporta beaucoup de succès, mais le chevalier perdit plus souvent que ses calculs ne l'avaient prévu. Il était convaincu que ces échecs répétés provenaient d'une erreur de raisonnement; mais laquelle ? Perplexe, Méré écrivit à son ami Blaise Pascal, le célèbre mathématicien français : «Combien de lancers, lui demanda Méré, faut-il faire pour être à coup sûr plus souvent gagnant que perdant ?» Pascal ne fréquentait pas les tables de jeux, mais le problème philosophique et pratique que pose tout pari - comment se comporter en se confrontant au hasard ? - l'intéressait au plus haut point. Il soumit le

problème au philosophe Pierre de Fermat et, pendant quatre mois, de juillet à octobre 1654, les deux grands mathématiciens étudièrent la question. Après d'enrichissantes discussions avec Fermat, Pascal rédigea un traité du triangle arithmétique, qui fit date dans l'histoire du calcul des probabilités. Pascal et Fermat découvrirent que Méré, avec son jeu de dés, était très loin de pouvoir gagner deux fois sur trois, comme il le pensait : les calculs de Pascal révélèrent que le chevalier avait seulement 49% de chances de gagner.



Pascal

Mais Pascal démontra que, pour arriver à inverser la tendance, il suffisait de lancer les dés une fois de plus. En vingt-cinq coups, Méré avait cinquante et une chances sur cent de faire un double-six.

L'histoire ne dit pas si Méré mit en pratique le conseil de Pascal, mais il ne trouva sans doute aucun jeu de dés susceptible de le faire gagner deux fois sur trois car, pour atteindre cet objectif, il faudrait lancer les dés... trente-neuf fois !

[Article tiré de *La chance et le hasard*, Collection *L'univers de l'étrange*, Éditions Time-Life, Amsterdam, 1992.]

«Comment expliquez-vous ... le coefficient de variation?»

par Sylvie Gauthier et Denis Malo

Étroitement relié à l'intervalle de confiance, thème abordé par Denis Malo dans *Convergence*, Vol.III, numéro 1, le sujet du présent article porte sur le coefficient de variation.

La mise en contexte pour expliquer le coefficient de variation devrait insister sur le fait que toute enquête par échantillon comporte des mesures de l'erreur d'échantillonnage, telles que la variance, le coefficient de variation, etc. qui déterminent la fiabilité des résultats et qui sont utilisées pour l'analyse des données.

Il serait ensuite primordial de mentionner que la variance en soi, représentée parfois par des nombres gigantesques, n'indique pas facilement le niveau de fiabilité d'un estimateur. À titre d'exemple, vous pourriez mentionner qu'une variance de 100 par rapport à une estimation de 50 et une variance de 100 par rapport à une estimation de 400 ne représente pas le même niveau de précision. Puis, vous pourriez enchaîner avec la définition du coefficient de variation (CV) à proprement dite, c'est-à-dire en mentionnant que le CV est le quotient de l'écart-type de l'estimation sur l'estimation ponctuelle. Votre client pourrait alors constater de lui-même que le CV permet d'évaluer la fiabilité d'une estimation plus aisément, et de surcroît, lui offre la possibilité de comparer la précision de plusieurs estimations.

Vous pourriez ensuite le surprendre en lui révélant le lien entre le CV et l'intervalle de confiance. Votre client, muni d'une estimation et du CV associé, sera heureux de constater qu'il est maintenant en mesure de construire un intervalle de confiance et de faire de l'inférence. Rappelons que si l'on suppose une distribution normale, le calcul de l'intervalle de confiance (I.C.) à 95% pour Y est donné par l'expression suivante :

$$I.C.(Y) = ESTIMATION \pm 1.96 * (ESTIMATION) * (C.V.(Y))$$

En dernier lieu, pour l'intérêt des lecteurs de *Convergence*, le tableau ci-dessous propose des lignes directrices pour les CVs qui sont utilisées dans certaines enquêtes de Statistique Canada.

Indicateur alphabétique	Intervalle de C.V.	Lignes directrices
A	0.0% - 0.5%	Sans réserve
B	0.6% - 1.0%	
C	1.1% - 2.5%	
D	2.6% - 5.0%	
E	5.1% - 10.0%	
F	10.1%-16.5%	
G	16.6%-25.0%	Sous réserve
H	25.1% - 33.3%	Confidentielle
J	33.4% - +	Diffusion interdite

Cet été, soyez

significativement différent!

Elles sont là! Prêtes pour l'été!

Les *CASQUETTES DE L'ASSQ* sont disponibles!

En coton bleu royal, courroie de cuir, logo de l'ASSQ en blanc avec le slogan "significativement différent!".



12 \$ ch. (sans taxe)

Il ne reste que quelques *T-SHIRTS DE L'ASSQ*!

À 50% de rabais, pourquoi s'en passer?!



5 \$ ch. (sans taxe)

Pour commander, contactez Pierre Lavallée au (613)951-2892.

(fax: (613)951-1462, plavall@statcan.ca)

Merci de votre appui!

Suivre son cours ...

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

Titre du cours	Lieu	Clientèle (3)	Durée totale	Date du cours / Inscription	Contact	But du cours
Statistique Canada (4)						
Composantes des séries chronologiques	STC, Ottawa	D	3 jours	21 au 23 septembre 1998/ dès maintenant	(1)	Examiner en profondeur les diverses composantes des chroniques socio-économiques ainsi que leur interprétation.
Désaisonnalisation	STC, Ottawa	I et A	4 jours	13 au 16 octobre 1998/ dès maintenant	(1)	Appliquer la méthode de désaisonnalisation X-11-ARIMA/88 sur des chroniques, en choisir les options les plus appropriées, évaluer les résultats obtenus avec cette méthode.
Modélisation et prévision ARIMA	STC, Ottawa	I et A	5 jours	11 au 15 janvier 1999 / 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..
Enquêtes longitudinales	STC, Ottawa	I et A	4 jours	26 au 29 octobre 1998/ dès maintenant	(1)	Familiariser les participants avec les méthodes propres aux enquêtes longitudinales, i.e. enquêtes pour lesquelles on effectue un suivi des répondants dans le temps.
Introduction à l'échantillonnage statistique	STC, Ottawa	D et I	3 jours	Hiver 1999 / dès maintenant	(1)	Dispenser des connaissances de base dans la pratique de l'échantillonnage. L'accent est mis sur les diverses méthodes, leurs avantages et leurs inconvénients, avec un minimum de formules.
Conception de questionnaires	STC, Ottawa	D et I	3 jours	13 au 15 janvier 1999 / 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 module 1	STC, Ottawa	D	3 jours	28 au 30 septembre 1998/ dès maintenant	(1)	Initier les participants aux concepts probabilistes de base.
Méthodes statistiques module 2	STC, Ottawa	D et I	3 jours	28 au 30 octobre 1998/ dès maintenant	(1)	Initier les participants aux concepts d'inférence statistique de base et de niveau intermédiaire.
Méthodes statistiques module 3	STC, Ottawa	I et A	5 jours	23 au 27 novembre 1998/ dès maintenant	(1)	Approfondir les concepts d'inférence statistique de niveaux intermédiaires et avancés des points de vue théorique et pratique.
Méthodes statistiques module 4	STC, Ottawa	I et A	6 jours	18 au 25 janvier 1999 / dès maintenant	(1)	Initier les participants aux méthodes statistiques nécessaires à l'analyse de données obtenus à l'aide de plans d'échantillonnage complexes..
École de technologie supérieure						
Maîtriser ses procédés: les méthodes Taguchi et traditionnelles	ETS, Montréal	D	3 jours	Automne 98 / dès maintenant	(2)	Planification d'expérience dans le domaine industriel.
Le contrôle statistique de la qualité	ETS, Montréal	D	2 jours	Automne 98 / dès maintenant	(2)	Contrôle statistique de la qualité (plans d'échantillonnage, CSP).

(1) Hew Gough (613) 951-3067

(2) Line Beauchamp (514) 396-8830 (Service de perfectionnement)

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

(4) 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.

SÉMINAIRES

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

UQAM

Pascale Rousseau
Tél. (514) 987-3000, #3224
Fax (514) 987-8935

Université Laval

Louis-Paul Rivest
Tél. (418) 656-7353
Fax (418) 656-2817

Bureau de la statistique du Québec

Paul Berthiaume
Tél. (418) 691-2410
Fax (418) 643-4129

Université de Sherbrooke

Jean Vaillancourt
Tél. (819) 821-8000, #2007
Fax (819) 821-8200

Université de Montréal

Roch Roy
Tél. (514) 343-7977
Fax (514) 343-5700
royr@dms.umontreal.ca