



Association
des statisticiennes
et statisticiens
du Québec

Convergence

Volume XVI Numéro 3

Février 2012

Ce document est une reconstitution du journal *Convergence* initialement publié sous la forme d'un regroupement d'articles sur le site web de l'association.

Le mot de la présidente

Par Nathalie Madore

Je m'en voudrais de ne pas commencer ce mot en vous souhaitant à toutes et à tous une très bonne année 2012, remplie de bonheur et de défis. Pour l'ASSQ, c'est un vent de renouveau qui souffle en ce début d'année : un nouveau format pour le *Convergence*, un nouveau rédacteur en chef, trois nouveaux membres au Conseil d'administration et l'ajout d'un membre étudiant observateur. Je tiens donc à profiter de ces quelques lignes pour vous présenter celles et ceux qui ont généreusement accepté de s'impliquer activement dans notre association.

La rédaction en chef du *Convergence* est aujourd'hui assumée par M. Denis Talbot, étudiant au doctorat en statistique à l'UQAM. Après son baccalauréat en statistique à l'Université Laval, il y a fait une maîtrise qu'il a terminée en 2010. Il a occupé un poste au service de consultation statistique pendant un an et donne le cours de statistique descriptive, toujours à Laval.

Les élections tenues cet automne ont permis de nommer trois nouveaux membres au CA. D'abord, une personne qui fait un retour au Conseil, M. Éric Lacroix, qui assumera les fonctions de registraire. Éric est vice-président Solutions Web et associé chez SOM, où il travaille depuis 2007. Il a auparavant œuvré au sein du CEFRIQ et de l'Institut de la statistique du Québec. Depuis plusieurs années, SOM prend en charge le registrariat de l'ASSQ et offre un service inestimable pour le support administratif et logistique nécessaire à notre fonctionnement. Je saisis donc l'occasion pour leur renouveler notre reconnaissance pour cette grande contribution.

Le secrétariat de l'association sera assumé dorénavant par Mme Lise Charrette, statisticienne au Ministère des Ressources naturelles et de la Faune. Après l'obtention de son diplôme en statistique à l'Université Laval en 1988, Lise a toujours travaillé au sein de l'appareil gouvernemental québécois.

Le dernier poste à combler au CA sera occupé par Andréanne Moisan-Roy, fraîchement issue elle aussi du baccalauréat en statistique de l'Université Laval où elle a étudié de 2008 à 2011. Elle est depuis peu à l'emploi de Revenu Québec. Andréanne s'est impliquée dans le Comité pour l'avancement de la statistique de l'Université Laval (CASUL) à titre de vice-présidente et responsable du site Web, puis de présidente lors de sa dernière année d'études.

Lors de la dernière assemblée générale tenue en juin, l'idée d'ajouter un représentant étudiant au CA avait été accueillie favorablement par les participants. Un tel changement nécessitant une révision de nos statuts, celui-ci ne peut être officialisé avant la prochaine assemblée générale. Il a toutefois été décidé par le conseil de nommer pour l'année 2011-2012, un membre étudiant observateur qui peut ainsi dès maintenant s'impliquer dans l'administration de l'ASSQ et jouer le rôle de relayeur d'information auprès des étudiants. La personne ainsi

nommée est M. David Émond, étudiant à la maîtrise en statistique à l'Université Laval, où il a terminé son baccalauréat en 2011. Il est actuellement président du CASUL. En compagnie d'Andréanne et d'une troisième personne, il était membre de l'équipe qui a remporté le printemps dernier la bourse SFdS-ISQ attribuée au meilleur projet de fin d'études.

Je termine ici en soulignant l'excellente collaboration des membres sortants du CA : Michel Fluet, Anne-Sophie Julien et Yann Latulippe. Votre implication a contribué à faire de notre groupe une association plus vivante que jamais et toujours pleine de projets pour les années à venir. Bonne chance pour vos projets futurs.

La présidente de l'ASSQ,
Nathalie Madore, Stat. ASSQ

Le mot du rédacteur en chef

Par Denis Talbot

Nouveau départ

Voici la première édition du *Convergence* où j'occupe le poste de rédacteur en chef. Je tiens tout d'abord à remercier Jean-François Quessy pour son travail à la barre du *Convergence* au cours des dernières années ainsi que pour ses conseils.

C'est également la première édition du *Convergence* sous sa nouvelle forme, celle d'un bulletin électronique. J'espère que cette nouvelle forme pour votre journal vous plaira. Il s'agit certainement d'un format de publication intéressant qui vous permettra facilement de lire les grandes lignes de chacun des textes et de déterminer si vous voulez lire le texte au complet en un simple clic. L'édition en bulletin électronique offrira également davantage de souplesse. En effet, dans sa forme précédente en format pdf, on aurait difficilement pu imaginer des textes de plusieurs pages ou des articles comportant de nombreuses photos.

Je veux en profiter pour remercier le travail de nos nombreux collaborateurs, d'abord les auteurs des textes que vous pourrez lire dans cette édition : Abdesselam Bouhia, Anne-Sophie Julien, Yann Latulippe, Pierre Lavallée, Nathalie Madore, Joseph Nader, Claude Ouimet et Sylvain Tremblay. Ensuite, je veux remercier les correcteurs qui révisent tous les textes pour s'assurer d'éliminer les erreurs qui pourraient s'y être glissées : Steve Méthot et Myrto Mondor. Finalement, un merci tout spécial à nos lecteurs qui sont la raison d'être du *Convergence*!

Denis Talbot,
Rédacteur en chef

Les échos du C.A.

Par Yann Latulippe

Projets futurs de l'ASSQ

Chers collègues,

Voici ma dernière chronique « Les échos du C.A. » en tant que secrétaire de l'association. Ce fut un réel plaisir de faire partie du conseil d'administration ces deux dernières années. Les réunions mensuelles sont truffées de discussions intéressantes et les idées des gens sont toujours bien reçues. Comme les projets sont stimulants, c'est facile de s'impliquer et d'apporter sa contribution. Voici donc les sujets qui ont été discutés au cours des quatre dernières réunions : les jeudis de l'ASSQ, l'évolution du site Web et du *Convergence*, la composition du C.A. et les élections pour l'année 2012.

Au sujet des jeudis de l'ASSQ, le 5 à 7 de la Rentrée a eu lieu le jeudi 15 septembre dernier au « Café le Temps perdu » à Sainte-Foy. Il s'agissait d'une rencontre sociale au cours de laquelle nous avons offert deux consommations et des amuse-gueules à nos invités. Un total de 15 personnes étaient présentes. La faible participation des gens préoccupe le conseil. Nous avons discuté d'une solution pour les prochains jeudis de l'ASSQ. Ainsi, on aimerait utiliser un nouveau média : la vidéo-conférence. Pour y arriver, on aurait besoin d'un caméraman dans la salle pour filmer la conférence. Le film pourrait être diffusé en temps réel ou en différé sur notre site Web. L'avantage serait d'élargir la portée de l'activité qui pourrait avoir lieu, par exemple, à Montréal et être retransmise à Québec et à Trois-Rivières. Il s'agit d'un beau défi à réaliser. De plus, l'association prépare un document de soutien à l'organisation d'événements. Vous savez que l'ASSQ a toujours pris les devants pour organiser des activités. En parallèle, on donnerait la possibilité aux membres de suggérer et planifier une activité dans le cadre d'un jeudi de l'ASSQ. Le rôle de l'association serait de contribuer au financement et à la logistique de l'activité. Cette offre pourrait même s'élargir à des activités reliées à la statistique organisées par un organisme. Votre implication sera toujours bien appréciée. Vous serez bientôt informés de ces nouveautés.

Pour ce qui est de l'évolution du site Web, les outils suivants sont déjà en ligne : le système de paiement de comptes Desjardins et la diffusion d'un calendrier d'activités pertinentes. Le comité de travail pour développer les nouvelles sections du site Web s'est réuni cet été pour une première rencontre. Pour votre information, nous remercions les personnes suivantes faisant partie du comité : Nathalie Madore, Jean-François Plante, Denis Talbot, Eric Lacroix, Myrto Mondor et Simon Olivier Fournier. Plusieurs éléments ont été discutés et seront développés en 2012. En voici un bref résumé.

On débutera par l'ajout d'une section « Carrières en statistiques ». Le contenu de cette dernière section touchera principalement la formation et les emplois visés dans le domaine de la statistique. De plus, le formulaire d'adhésion électronique sera révisé. Il sera plus simple de le compléter. Enfin, un nouveau visuel sera testé sur notre site Web pour faciliter la diffusion d'articles ou de documents. Bien entendu, on retrouvera des articles de notre journal *Convergence*.

En effet, tel qu'annoncé dans la dernière édition, le journal *Convergence* a été converti en un bulletin électronique. Son contenu est sous la responsabilité d'un nouvel édimestre, Denis Talbot, qui prend la relève du chevronné Jean-François Quessy. Avec le bulletin électronique, il sera possible de diffuser plusieurs articles sans se soucier de l'espace disponible (élimination du format papier). De plus, les membres seront ainsi amenés plus souvent à visiter le site. Nous souhaitons que le site devienne le lieu privilégié d'échanges d'informations entre nous.

Comme troisième sujet, parlons de la composition du C.A. Depuis le mois d'octobre, le conseil permet à un étudiant d'assister aux réunions mensuelles en tant qu'observateur. Il s'agit de David Émond. Comme il est étudiant, il est facile de discuter avec ses collègues des projets en cours. De plus, David peut même inciter ses collègues à devenir membres. Pour votre information, un statut « membre étudiant » est mentionné dans les

règlements. Une proposition sera faite lors de la prochaine assemblée générale annuelle pour ajouter officiellement ce poste au sein du C.A. On prend donc les bons moyens pour assurer la relève à long terme !

Le dernier sujet concerne les élections pour l'année 2012. Comme à chaque automne, des élections se sont tenues pour combler trois postes qui devenaient libres le 31 décembre 2011. Le résultat des élections est positif : trois nouvelles personnes rejoindront prochainement le conseil d'administration. Les élus pour les deux prochaines années sont : Lise Charrette, Andréanne Moisan-Roy et Eric Lacroix qui occuperont respectivement les postes de secrétaire, trésorier et registraire. En mon nom personnel, je leur souhaite beaucoup de plaisir au sein du C.A.

Alors, voilà les sujets discutés récemment lors des dernières réunions du conseil. Il s'agit de mon dernier article; je passe le flambeau à Lise Charrette. Amuse-toi bien ! Je vous souhaite à tous une bonne année 2012. Au plaisir de se revoir !

Yann Latulippe,
Secrétaire

La minute financière

Par Anne-Sophie Julien

Service de perception de compte

Le service de perception de compte universel a été implanté au cours du mois de janvier 2011. Celui-ci permet aux membres d'effectuer leur paiement directement auprès d'une institution bancaire, soit en personne, soit par téléphone, soit par internet. Les institutions bancaires participantes sont les Caisses Desjardins, la Banque Nationale du Canada, la Banque Scotia et la Banque Royale.

Au cours des six premiers mois suivant son implantation, 31 transactions ont été effectuées par 25 membres. 25 transactions ont été effectuées auprès des Caisses Desjardins, alors que les autres membres ont utilisé le service offert par la Banque Nationale du Canada. Parmi ces transactions, 21 ont été effectuées dans le but de renouveler l'adhésion, tandis que les transactions restantes ont servi pour l'inscription au colloque.

Introduction à une méthode graphique d'inférence causale

Par Denis Talbot

Pour cette première édition, je souhaite partager avec vous un sujet souvent méconnu des statisticiens : l'inférence causale. D'abord, quel est l'intérêt d'effectuer de l'inférence causale? L'intérêt est de pouvoir tirer des conclusions du genre « Si je modifie la valeur de X, qu'advient-il à Y? », alors que l'inférence classique ou « prédictive » ne permet que des conclusions du genre « Si X change naturellement de valeur, qu'advient-il en général à Y? ». Plusieurs techniques d'inférence causale existent, mais j'aborderai ici l'approche graphique suggérée par Judea Pearl.

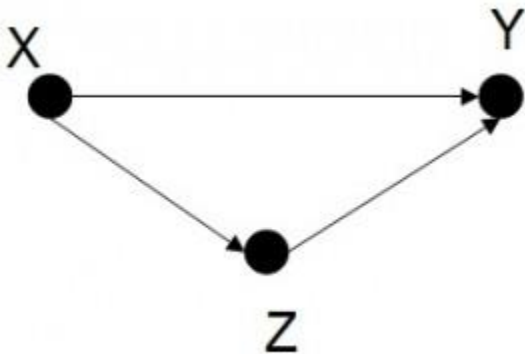
La façon la plus simple, et la plus sûre, d'effectuer de l'inférence causale est par le biais d'expériences randomisées. On divise alors les unités expérimentales aléatoirement en groupes et toutes les unités d'un groupe donné reçoivent un des traitements qu'on désire étudier. Si le hasard fait bien les choses, les groupes devraient avoir des caractéristiques similaires et toute différence dans la réponse observée entre les groupes pourra être attribuée à l'effet causal du traitement.

Qu'arrive-t-il lorsque l'étude randomisée est impossible? Doit-on abandonner complètement l'idée d'effectuer de l'inférence causale? La réponse donnée par Pearl est un «non» catégorique! Évidemment, l'inférence causale ne pourra se faire sans payer un certain prix : des hypothèses non vérifiables à partir des données devront être posées. Par contre, toute personne prête à accepter ces hypothèses devra également accepter l'interprétation

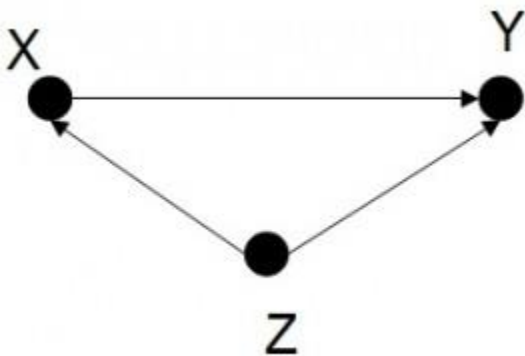
causale qui découle des analyses, puisqu'il s'agira d'une conséquence directe de ces hypothèses et des informations disponibles dans les données.

L'approche proposée par Pearl consiste à tracer un graphique acyclique orienté (« Directed Acyclic Graph » ou « DAG »). Il s'agit d'un graphique où les variables observées seront représentées par des cercles pleins, les variables non observées par des cercles vides et où les relations entre les variables seront représentées par des flèches à une pointe. Par ailleurs, si on suit le chemin dirigé (dans le sens des flèches) à partir d'une variable, on ne retournera jamais à la variable d'origine, d'où le qualificatif de graphique acyclique. Les flèches du DAG représentent des relations de cause à effet. Voici quelques exemples de DAG :

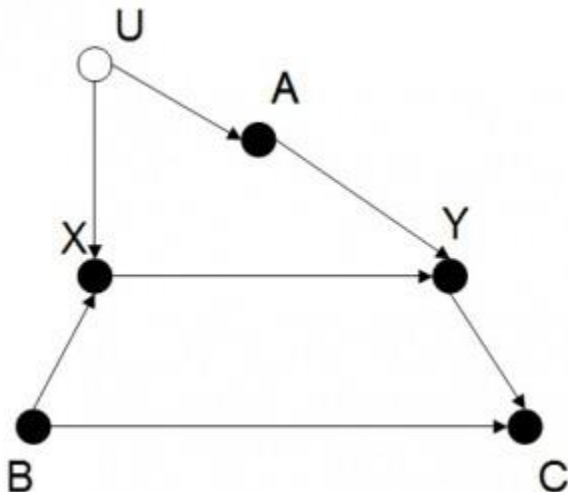
Exemple 1 :



Exemple 2 :



Exemple 3 :



Introduisons maintenant le concept de chemin porte-arrière (« backdoor path »). Si on s'intéresse à l'effet causal de X sur Y, on voudra s'assurer qu'il n'y a pas de chemin porte-arrière ouvert entre Y et X. Un chemin porte-arrière est un chemin, autre que le chemin direct entre X et Y, qui permet de relier la variable d'arrivée (Y) à la variable d'origine (X) et dont le dernier segment pointe vers la variable d'origine (X). Dans le premier exemple, les trajets de Y à X sont : $Y \leftarrow X$ et $Y \leftarrow Z \leftarrow X$. Ni l'un ni l'autre de ces trajets n'est un trajet porte-arrière, puisque le dernier segment ne pointe pas vers X. Dans l'exemple 2, les trajets sont : $Y \leftarrow X$ et $Y \leftarrow Z \rightarrow X$. Le trajet $Y \leftarrow Z \rightarrow X$ est un trajet porte-arrière. Finalement, dans le troisième exemple, nous avons : $Y \leftarrow X$, $Y \rightarrow C \leftarrow B \rightarrow X$, $Y \leftarrow A \leftarrow U \rightarrow X$, les deux derniers trajets étant des trajets porte-arrière.

À l'intérieur de certains de ces chemins, on retrouve ce qu'on appelle un collisionneur (« collider »). Un collisionneur est une variable vers laquelle deux variables pointent dans un trajet. Dans le troisième exemple, la variable C joue le rôle de collisionneur dans le trajet $Y \rightarrow C \leftarrow B \rightarrow X$.

Pour que l'effet causal de X sur Y soit correctement estimé, il est nécessaire de conditionner sur un ensemble de variables qui permettra de fermer tous les chemins porte-arrière et de ne conditionner sur aucun descendant de X. Un trajet porte-arrière sera fermé si et seulement s'il comporte un non-collisionneur sur lequel on conditionne ou s'il comporte un collisionneur sur lequel on ne conditionne pas et pour lequel on ne conditionne sur aucun de ses descendants. Par descendants, on entend les variables vers lesquelles le collisionneur pointe directement ou indirectement (dans l'exemple 3, C est un descendant de A en passant par Y). On comprend donc que le fait de conditionner sur un collisionneur ouvrira le chemin porte-arrière alors que de ne pas conditionner sur un collisionneur maintient le chemin fermé. Attention, le fait de contrôler ou de ne pas contrôler pour certaines variables afin de fermer un trajet porte-arrière peut conduire à l'ouverture d'un autre chemin porte-arrière. Il est donc essentiel de vérifier que tous les chemins ont été fermés.

Dans le premier exemple, il n'y a aucun chemin porte-arrière, il n'est donc pas nécessaire de conditionner sur la variable Z. En fait, si on conditionne sur Z, on aura conditionné sur un descendant de X et l'effet causal de X sur Y ne pourra être estimé correctement. Dans l'exemple 2, le trajet porte-arrière $Y \leftarrow Z \rightarrow X$ n'est pas bloqué. Il faut donc contrôler sur Z afin d'obtenir l'effet causal de X sur Y. Dans le troisième exemple, $Y \rightarrow C \leftarrow B \rightarrow X$ est un chemin porte-arrière naturellement bloqué par le collisionneur C. Il n'est donc pas nécessaire de faire quoi que ce soit pour bloquer ce trajet. Par contre, le trajet $Y \leftarrow A \leftarrow U \rightarrow X$ est ouvert et devra être bloqué. On pourrait penser contrôler pour A ou U, par exemple. Par contre, la variable U n'étant pas observée, il faudra contrôler pour A. Le fait de contrôler pour A n'affectera en rien le trajet $Y \rightarrow C \leftarrow B \rightarrow X$, puisque A ne se retrouve pas dans ce trajet et n'est pas un descendant de C.

Comme on pourra le constater, il est nécessaire d'incorporer plusieurs nouvelles notions pour utiliser l'approche graphique suggérée par Pearl. Ces notions sont néanmoins assez simples et, une fois bien assimilées, sont faciles à utiliser. Un autre point qu'on pourra remarquer est que l'inférence causale selon cette approche nécessite une connaissance suffisamment approfondie du domaine d'application pour être en mesure de tracer un DAG raisonnable. Tel que mentionné initialement, toute personne acceptant le DAG tracé devra également accepter les conclusions tirées à partir de ce DAG. Pour plus d'information, je vous invite à consulter *Causality: Models, Reasoning, and Inference*, Cambridge University Press, 2009 de Judea Pearl.

Denis Talbot

Chronique historique

Par Pierre Lavallée

Hypathie

Hypatie d'Alexandrie (vers 370 – vers 415)

Hypatie d'Alexandrie connut une mort violente en étant lacérée par la foule, en partie parce qu'elle n'adhérait pas aux principes de la religion chrétienne. Elle se considérait elle-même comme néoplatonicienne, païenne et disciple des idées de Pythagore. Hypatie est la première femme mathématicienne de l'histoire de l'humanité à propos de laquelle nous possédons des informations fiables et détaillées. Elle avait la réputation d'être séduisante et fermement décidée à rester célibataire. Quand on lui demandait pourquoi elle était obsédée par les mathématiques et ne prenait pas d'époux, elle répondait qu'elle était déjà mariée à la vérité.

Parmi les travaux d'Hypatie figurent ses commentaires sur l'Arithmétique de Diophante. Dans l'un des problèmes proposés à ses étudiants elle leur demandait la solution du système d'équations suivant :

$$x - y = a, x^2 - y^2 = (x - y) + b,$$

où a et b sont connus. Existe-t-il des valeurs entières de x, y qui vérifient les deux formules ?

Les chrétiens étaient ses principaux rivaux philosophiques et ont officiellement condamné ses assertions platoniciennes sur la nature de Dieu et de l'au-delà. Par un jour de mars 415, elle fut attrapée par une foule de chrétiens fanatiques, déshabillée et eut la peau lacérée jusqu'aux os par des coquilles d'huître. Son corps fut ensuite coupé en morceaux et brûlé. Comme certaines victimes du terrorisme religieux aujourd'hui, il est probable que son sort tragique soit lié au fait qu'elle ne partageait pas les mêmes croyances. Il faudra attendre le XVIIIe siècle pour qu'une autre femme, Maria Agnesi, devienne une mathématicienne célèbre.

La mort d'Hypatie entraîna le départ de nombreux savants d'Alexandrie et, à bien des égards, marqua la fin de siècles de progrès pour les mathématiques grecques. Pendant l'âge des ténèbres de l'Europe, les Arabes et les Indiens furent les seuls à jouer un rôle majeur en encourageant l'évolution des mathématiques.

[tiré de Pickover, C.A. (2010), Le beau livre des maths – De Pythagore à la 57ème dimension, Dunod, 528 pages.]

Activité de golf

Par Claude Ouimet

Septembre 2011

Le matin du 10 septembre, aux quatre coins de la province, 18 golfeurs se préparent fébrilement, chacun selon son rituel, à vivre une autre classique annuelle ASSQ qu'ils espèrent mémorable. Tous quittent bien tranquillement leur domicile, le cœur léger, sachant fort bien qu'ils n'auront pas à se soucier de la météo, les prévisions ne pouvant être plus favorables que ce qu'elles sont ce jour-là. Arrivent donc une à une au prestigieux golf de Plessisville les rutilantes voitures des participants inscrits. « Mais où sont donc les valets parking? » se disent-ils tous, les uns après les autres. « Ne sommes-nous pas attendus ? » Faisant fi de cet affront du club hôte (on ne peut pas tout avoir bien sûr pour 25\$, golf et voiturette électrique incluse), les compétiteurs ne se laissent pas déstabiliser, préférant garder leur concentration pour le duel à venir...et quel duel ce sera.



Selon une formule tellement tordue que même Einstein y perdrait son latin, les compétiteurs sont lâchés dans la fosse ~~aux lions~~ de sable et offerts en pâture aux 18 trous, tous plus féroces les uns que les autres. Et avec comme seule arme, tout de même, quatorze bâtons pour catapulter les ~~boulets~~ balles au fond des trous. Et pour ceux qui manqueraient de « visou », cinquante pieds de ficelle enroulés maladroitement autour d'un rouleau de papier de toilette. « Mais à quoi joue-t-on? » se disent, incrédules, les participants médusés.



Mais au golf pardi! C'est bien simple : les pieds de ficelle viennent au secours des coups roulés manqués. Un coup manqué de dix pieds devenant une réussite en sacrifiant dix des cinquante pieds de ficelle distribués aux quatuors. Pour ajouter un certain défi, chaque groupe doit donc prévoir son score final le plus précisément possible et ramener le moins long de corde possible à la fin de la partie. Pardon? C'est quoi ce nouveau règlement? On venait ici pour se détendre et oublier les soucis de semaine et voilà qu'il faut se faire aller les méninges au lieu de jouer une petite partie bien tranquille.



Les joueurs ont accepté de jouer le jeu et cette nouvelle formule est même devenue rapidement le sujet de conversation de la journée. Une chance, puisque la journée mémorable souhaitée le fut réellement, mais pas pour les raisons attendues. Malheureusement, les joueurs ayant été jumelés à un autre groupe en tournoi, il aura fallu environ quatre heures pour négocier les onze premiers trous. Mais, la bonne compagnie et la saine camaraderie aidant, il fallait prendre le tout en riant et c'est ce que tous ont fait. Les sept derniers trous se sont heureusement joués à un rythme plus normal. Le tout s'est terminé dans un resto sympathique, mais moins gastronomique qu'anticipé, malheureusement. L'ambiance feutrée du salon privé qui nous était réservé a favorisé des échanges bien nourris (pour ne pas dire mieux nourris que les participants), ce qui clôturait la journée en beauté.



Chronique SAS

Par Sylvain Tremblay, Groupe de formation – Institut SAS (Canada) inc.

Chiffrement des données et encodage des mots de passe

“Tire la chevillette, la bobinette cherra.”

Aujourd’hui, tout comme à l’époque de Perrault, nous cherchons à nous protéger des gens trop curieux et des voleurs, et ce en utilisant plusieurs moyens, par exemple une porte bien barrée. Dans nos milieux de travail, il en va de même pour les données.

Vous avez des fichiers de données SAS et vous cherchez à les sécuriser? Vous désirez contrôler qui peut les lire ou les modifier? Soyez rassuré, il existe plusieurs moyens de le faire. Certains sont simples et disponibles à même le module SAS/Base. D’autres sont plus sophistiqués et nécessitent le module SAS/Secure. Dans cet article, j’aimerais aborder deux techniques de base : le chiffrement (“encryption”) des fichiers de données et l’encodage des mots de passe.

Chiffrement des fichiers de données

Le chiffrement d’un fichier de données SAS offre un certain niveau de sécurité. Une fois chiffré avec un mot de passe, seulement ceux qui ont ce dernier peuvent lire, écrire ou modifier le fichier, dépendamment de la protection. Ceci est vrai à l’intérieur ou à l’extérieur de SAS (par exemple si le fichier est copié sur un CD ou envoyé par courriel). On doit noter que l’on ne peut pas chiffrer une vue SAS car cette dernière ne contient pas de données. Si un fichier de données SAS est chiffré, ses index seront aussi chiffrés. Le chiffrement demande approximativement les mêmes ressources CPU que la compression d’un fichier de données.

Afin de chiffrer un fichier de données, on doit utiliser l’option ENCRYPT=YES dans une étape DATA en spécifiant au minimum un mot de passe avec l’option PW=. Il y a toutefois des limitations pour ce dernier. Il doit être un nom

SAS valide et avoir au maximum 8 caractères (sous Windows). Les autres options possibles qui peuvent accompagner le ENCRYPT=YES sont :

READ= (sert à spécifier un mot de passe pour la lecture),

WRITE= (sert à spécifier un mot de passe pour l'écriture),

ALTER= (sert à spécifier un mot de passe pour modifier).

Si l'on ne spécifie que les options ENCRYPT=YES et PW=, alors le fichier sera automatiquement chiffré pour la lecture, l'écriture et la modification. Le programme suivant illustre cette approche (voir encadré 1).

Encadré 1

```
1 /* Chiffrement du fichier de données */
2
3 data work.salaire(encrypt=yes pw=mdepasse);
4     input nom $ salaire bonuspct;
5     datalines;
6 Muriel      34567    3.2
7 Bjorn       74644    2.5
8 Freda      38755    4.1
9 Benny      29855    3.5
10 Agnetha    70998    4.1
11 ;
12 run;
13
14 proc contents
15     data=work.salaire;
16 run;
```

Si l'on consulte ensuite le fichier journal après l'exécution de l'étape DATA, on constate que SAS a masqué le mot de passe (voir encadré 2).

Encadré 2

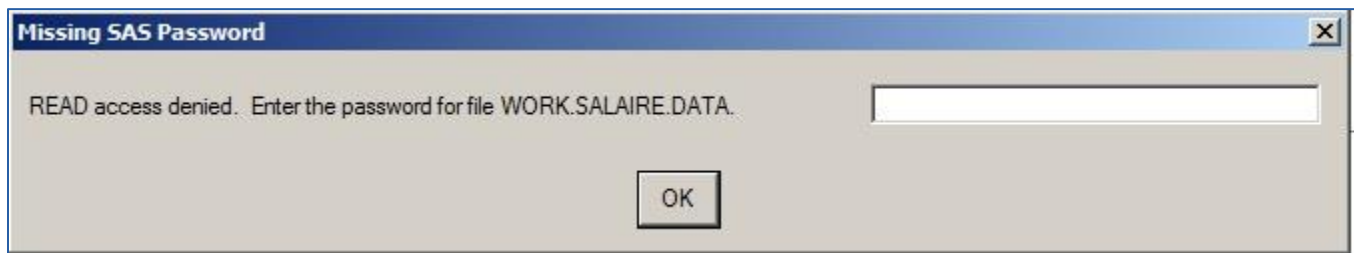
```
96 data work.salaire(encrypt=yes pw=XXXXXXXX);
97     input nom $ salaire bonuspct;
98     datalines;

NOTE: The data set WORK.SALAIRE has 5 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time           0.04 seconds
      cpu time            0.00 seconds

104 ;
105 run;
```

Lors de l'exécution de la procédure CONTENTS, SAS va nous demander le mot de passe car le fichier de données a été chiffré pour la lecture (voir encadré 3).

Encadré 3



Une fois le mot de passe entré (de façon interactive), SAS va déchiffrer les données, les lire, exécuter la procédure et afficher les résultats dans la fenêtre de sortie. Lors de l'examen de cette sortie, on remarque alors la présence de deux nouveaux attributs :

Protection (READ/WRITE/ALTER)

Encrypted (YES)

Ces attributs nous indiquent que le fichier de données a été chiffré et que l'utilisation du mot de passe est nécessaire pour sa lecture, son écriture ou sa modification (voir encadré 4).

Encadré 4

The CONTENTS Procedure			
Data Set Name	WORK.SALAIRE	Observations	5
Member Type	DATA	Variables	3
Engine	V9	Indexes	0
Created	Wed, Dec 07, 2011 03:37:06 PM	Observation Length	24
Last Modified	Wed, Dec 07, 2011 03:37:06 PM	Deleted Observations	0
Protection	READ/WRITE/ALTER	Compressed	NO
Data Set Type		Sorted	NO
Encrypted	YES		
Label			
Data Representation	WINDOWS_64		
Encoding	wlatin1 Western (Windows)		

Le fichier de données SAS est maintenant chiffré et il sera difficile pour les personnes non autorisées de l'utiliser. Mise en garde : **ne perdez pas le mot de passe!** Il vous sera en effet impossible d'accéder à votre fichier sans l'aide du personnel de SAS. Ce processus prend du temps et des ressources significatives et ne doit être utilisé qu'en dernier recours.

Mais que faire si vous devez travailler avec un fichier de données chiffré mais que vous n'êtes pas dans une session interactive avec SAS, lorsque le programme est exécuté par un ordonnanceur (scheduler) par exemple? En effet, SAS va ouvrir une fenêtre et va vous inviter à entrer le mot de passe, mais vous ne serez pas devant votre écran. La solution est simple : vous pouvez utiliser l'option de fichier de données PW= dans votre code pour spécifier directement le mot de passe. Dans l'exemple suivant, cette option est utilisée avec la procédure PRINT (voir encadré 5).

Encadré 5

```
17  
18 proc print data=work.salaire (pw=mdepasse) ;  
19 run;
```

Le tour est joué... mais d'un point de vue sécurité, cela n'est pas l'idéal car votre mot de passe est dans le programme au vu et au su de tous! Que faire? Il s'agit simplement d'encoder votre mot de passe.

Encodage des mots de passe

La procédure PWENCODE du module SAS/Base permet l'encodage des mots de passe. La sortie, sous la forme d'une chaîne de caractères, inclut un « tag » qui identifie la chaîne de caractères (le mot de passe) qui a été encodée. Un exemple de tag est {sas002}. Le tag indique la méthode d'encodage. Il existe trois méthodes d'encodage (sas001 à sas003) qui se servent de différentes clés d'encodage (32-bit, 256-bit, ..) et différents algorithmes (SASProprietary, AES, ...).

La méthode d'encodage sas002 utilise une clé de 32-bit et l'algorithme SASProprietary qui vient avec le module SAS/Base.

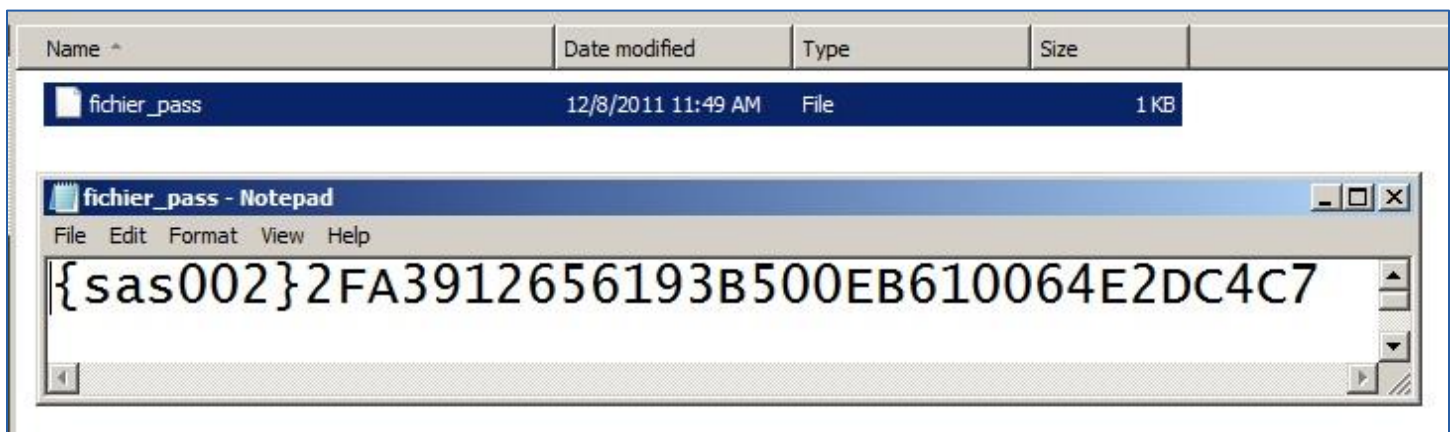
Le programme suivant illustre comment encoder le mot de passe et le sauvegarder dans un fichier externe (voir encadré 6). Cela doit être fait dans un programme à part, séparé de votre programme principal afin d'isoler votre mot de passe.

Encadré 6

```
22 /* Encodage du mot de passe dans un fichier externe */
23
24 filename fichierpw 'c:\mon_repertoire\fichier_pass';
25
26 proc pwencode
27     in='mdepasse'
28     out=fichierpw
29     method=SAS002;
30 run;
```

Notre mot de passe est maintenant à l'abri des yeux trop curieux. Le fichier ainsi créé est un fichier texte et si on l'ouvre avec un éditeur, tout ce que l'on voit est le tag et une chaîne de caractères (voir encadré 7).

Encadré 7



Ensuite, dans le programme principal qui doit traiter les données, notre but est d'utiliser le mot de passe sans que ce dernier n'apparaisse. Pour ce faire, on peut utiliser une étape DATA nulle afin de lire et déposer le mot de passe encodé dans une variable macro grâce à la routine SYMPUT et ensuite on n'a qu'à appeler cette dernière dans nos traitements (voir encadré 8). N'oubliez pas d'utiliser les doubles guillemets lorsque vous faites appel au mot de passe car vous appelez une variable macro.

Encadré 8

```
32 filename fichierpw 'c:\mon_repertoire\fichier_pass';
33
34 data _null_;
35     infile fichierpw obs=1 length=1;
36     input @;
37     input @1 line $varying1024. 1;
38     call symput('passw', substr(line, 1, 1));
39 run;
40
41 proc print data=work.salaire (pw="&passw");
42 run;
```

Conclusion

Peu importe le domaine où vous œuvrez, la sécurité est l'affaire de tous! SAS offre une panoplie de moyens pour vous permettre de sécuriser vos données. Le module SAS/Secure assure la sûreté de vos données en transit par l'entremise de protocoles tels le SSL ("Secure Sockets Layer" qui s'imbrique avec les protocoles HTTP, FTP et Telnet), le SSH ("Secure Shell") et divers algorithmes de chiffrement (RC4, DES, AES, ...).

Pour des besoins plus élémentaires, le module SAS/Base vous permet de chiffrer vos fichiers de données avec l'option ENCRYPT=YES et d'encoder vos mots de passe avec la procédure PWENCODE. Vous pouvez maintenant dormir l'esprit tranquille et tenir à l'écart le grand méchant loup!

Référence

"Encryption in SAS 9.2"

<http://support.sas.com/documentation/cdl/en/secref/62092/PDF/default/secref.pdf>

Événements à venir

Par Joseph Nader

Voici un choix d'activités futures, annoncées par différentes organisations statistiques. Si vous avez des conférences à annoncer, n'hésitez pas à nous contacter !

Quoi ?	Où ?	Organisateurs	Quand ?
2012			
Conference on Data Analysis	Santa Fe, Nouveau Mexique (É-U)	Center for nonlinear studies (subventions possibles aux étudiants)	29 février au 2 mars
Traitement des données en SHS	Paris (France)	Société française de statistique	6 avril 2012
SAS Global Forum 2012	Orlando, Floride (É-U)	SAS	22 au 25 avril 2012
44e Journées de Statistique	Bruxelles (Belgique)	Université libre de Bruxelles	21 au 25 mai 2012
Congrès annuel de la SSC	Guelph (Ontario)	Société Statistique du Canada	3 au 6 juin 2012
ICES-IV: Fourth International Conference on Establishment Surveys	Montréal (Québec)	American Statistical Association	11 au 14 juin 2012
International Symposium in Statistics (ISS) 2012 on Longitudinal Data	St John's (Terre-Neuve)	Memorial University	16 au 18 juillet 2012
Joint Statistical Meeting	San Diego, Californie (É-U)	American Statistical Association	28 juillet au 2 août 2012
2013			
45e Journées de Statistique	Toulouse (France)	Société française de statistique	27 au 31 mai 2013

Pour ceux que cela intéresse, le site de l'ISI (Institut international de statistique, basé à La Haie, aux Pays-Bas), <http://isi.cbs.nl/calendar.htm>, dresse une liste assez exhaustive des événements internationaux importants à venir.

Le site [Autres Manifestations](#) de la SFdS dresse aussi une liste d'activités complémentaires.

La Société statistique de Montréal, SSM, a repris ses activités sur une base plus régulière; plusieurs activités sont prévues en 2012; pour plus de détails, visiter son site

L'horaire ci-dessus peut être sujet à changement. Consultez le site web des organisateurs pour de plus amples informations.

Suivre son cours

Par Joseph Nader

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

Cette grille présente les cours offerts par différents organismes. La liste des cours universitaires est disponible sur demande auprès de l'ASSQ. L'information fournie dans cette grille est sujette à changements. Le lecteur est invité à communiquer avec les responsables des cours pour corroborer et compléter l'information présentée (frais d'inscription, dates, etc.).

Organisme, Lieu	Clientèle (basée sur la matière du cours)	Types de cours	Contacts
Institut de formation de Statistique Canada	Intermédiaire et avancé	Différents cours portant sur les méthodes d'analyse de données (séries chronologiques, contrôle statistique de la qualité, analyse de données de survie, ...), méthodologie d'enquête.	Céline Charrette (613) 951-1044 www.statcan.ca charcel@statcan.ca Note : Certains cours peuvent se donner à l'extérieur.
Centre interuniversitaire québécois de statistiques sociales	Avancé	Ateliers pratiques de formation intensive portant sur des problèmes méthodologiques particuliers à l'analyse de données complexes, des méthodes spécifiques et/ou des données d'enquêtes particulières.	Heather Juby, Coordonnatrice, Transfert du savoir (514) 343-2090 poste 8 www.ciqss.umontreal.ca heather.juby@umontreal.ca
École de Technologie Supérieure, Montréal	Débutant	Planification d'expérience dans le domaine industriel, contrôle statistique de la qualité.	Suzanne LeBel, Service de perfectionnement (514) 396-8830 www.perf.etsmtl.ca perf@etsmtl.ca

Institut SAS Montréal, Québec, Ottawa	Grand public, entreprises, débutant ou expert	Data mining, langage macro, programmation, rédaction de rapport, statistique.	Nathaly Renaud (514) 395-8922 poste 4071 www.sas.com/formation sastrain@can.sas.com
Creascience, Montréal	Débutant, intermédiaire, chercheurs, techniciens et autres professionnels de la R&D	Planification d'expériences, analyse en composantes principales, contrôle statistique de la qualité.	Natalie Rodrigue (514) 840-9220 poste 27 www.creascience.com info@creascience.com
Services conseils Hardy, Québec, Montréal	Débutant, intermédiaire et avancé	Différentes sessions de formation portant sur les logiciels SAS, SPSS, MINITAB, Statistica, Crystal Reports.	Monique Trempe (418) 626-1666 www.schardy.qc.ca schardy@schardy.qc.ca
Université Concordia, cours en ligne	Débutant, intermédiaire	Cours de base.	Patrick Devey, Directeur (514) 848-2424 poste 5259 devey@alcor.concordia.ca

Séminaires

Des séminaires ont lieu de façon régulière aux endroits suivants. N'hésitez pas à contacter les responsables pour plus de détails.

UQAM Pascale Rousseau Tél. (514) 987-3000 #3224 Fax (514) 987-8935 www.uqam.ca rousseau.pascale@uqam.ca	Université de Montréal Martin Bilodeau Tél. (514) 343-2410 Fax (514) 343-5700 www.umontreal.ca bilodeau@dms.umontreal.ca	Université de Sherbrooke Bernard Colin Tél. (819) 821-8000 # 62012 Fax (819) 821-7189 www.usherbrooke.ca Bernard.Colin@USherbrooke.ca	Université Laval Thierry Duchesne Tél. (418) 656-5077 Fax (418) 656-2817 www.ulaval.ca Thierry.Duchesne@mat.ulaval.ca
--	--	--	---

Nouvelles publications

Par Joseph Nader

Voici un choix de monographies, de livres de cours et autres ouvrages reliés à la statistique récemment publiés, principalement en français. Si d'autres ouvrages vous semblent intéressants, n'hésitez pas à nous contacter, nous nous ferons un plaisir de les ajouter à notre répertoire !

Titre	Auteur Date de parution Éditeur	Notes ¹
R, L'essentiel	Joseph Adler Novembre 2011 Pearson Education	Gratuit, programmable et extrêmement riche, R s'impose comme la référence open-source en matière de développements et d'analyses statistiques. Cet ouvrage est à la fois une introduction et un guide de référence. Il présente de manière la plus complète possible : le langage et ses subtilités, la programmation en R, les packages les plus importants, la liste des méthodes et des fonctions associées. 620p.
Manuel de statistique	Aubert Henry Novembre 2011 Ellipses	Cet ouvrage couvre l'ensemble des méthodes statistiques, de façon rigoureuse, mais pratique, c'est-à-dire sans tomber dans le développement de la théorie mathématique. Les «formules» de «calculs manuels» y sont également considérées comme secondaires, grâce au recours systématique au tableur Excel. L'utilisation d'Excel est complétée par la fourniture de fonctions personnalisées écrites par l'auteur quand Excel ne fournit pas la fonction en standard. Ces fonctions sont dans des classeurs Excel, un par chapitre, sur le site de l'auteur, et leur code non protégé : les lecteurs peuvent donc aussi s'initier à Visual Basic for Applications. L'accent est mis, dans un chapitre à part entière, sur l'importance de la qualité de l'échantillonnage dans les méthodes statistiques. Un chapitre est également consacré aux questionnaires d'enquêtes, un aspect de la qualité des données et par conséquent de celle des résultats qui ne doit pas être négligé. 408p.
De l'intégration aux probabilités	Olivier Garet Aline Kurtzmann Novembre 2011 Ellipses	Cette initiation à la théorie de l'intégration et au calcul des probabilités contient un cours illustré par des exercices avec les corrigés. 488p.
Comprendre et réaliser les tests statistiques à l'aide de R	Gaël Millot Octobre 2011 De Boeck	Illustré par de nombreuses figures et accompagné d'exercices avec correction, l'ouvrage aborde la statistique de la manière la plus simple qui soit, sans démonstration mathématique, mais en insistant sur les détails, afin de bien maîtriser toutes les subtilités des tests. Il traite en profondeur de notions essentielles comme la check-list à effectuer avant de réaliser un test, la gestion des individus extrêmes, l'origine de la p value, la puissance ou la conclusion d'un test. Il explique comment choisir un test à partir de ses propres données. Il décrit 35 tests statistiques sous forme de fiches, dont 24 non paramétriques, ce qui couvre la plupart des tests à une ou deux variables observées. Il traite de toutes les subtilités des tests, comme les corrections de continuité, les corrections de Welch pour le test t et l'anova, ou les corrections de p value lors des comparaisons multiples. Il propose un exemple d'application de chaque test à l'aide de R, en incluant toutes les étapes du test, et notamment l'analyse graphique des données. Ne nécessitant pas de connaissances particulières en mathématiques, ni de pré-requis en informatique, cet ouvrage devrait contenter à la fois ceux qui recherchent un manuel de statistique/biostatistique expliquant le fonctionnement des tests et ceux qui recherchent un manuel d'utilisation de R. 767 p.

Et à titre de divertissement :

Titre	Auteur Date de parution Éditeur	Notes ¹
Pourquoi faisons-nous des choses stupides ou irrationnelles ?	Sylvain Delouvé 2011 Dunod	Il vous est déjà certainement arrivé de vous trouver dans une situation où vous ne savez pas comment agir et vous comporter. Sans repère autour de vous ou sans connaissance préalable de la situation à laquelle vous êtes confrontée, qu'allez-vous faire ? Vous allez regarder à votre droite, puis à votre gauche et vous allez tout simplement prendre exemple sur les autres en ayant une idée simple en tête : si je ne sais pas quoi faire, les autres, eux, doivent savoir... Seulement voilà, les autres vont faire la même chose que vous ! Vous les regardez : ils vous observent. Vous vous trouvez alors dans une situation d'ignorance collective où chacun observe autrui pour savoir comment agir ou se comporter. 224 p.
Petites expériences scientifiques déjantées et décoiffantes!	Sean Connolly 2011 Dunod	65 + 65 = 130 expériences renversantes, à base d'explosions soudaines, de geysers bouillonnants, de fusées sifflantes, de ballons survoltés et bien d'autres curiosités bruyantes ou salissantes, voire les deux à la fois ! Pour chaque expérience, toutes facilement réalisables, vous trouverez les ingrédients nécessaires, les précautions d'usage, et bien sûr la caution scientifique. Coffret de 2 ouvrages.

¹ Les « Notes » sont extraites des notices descriptives préparées par les éditeurs.

Une visite au site <http://www.maths-cuicui.fr/accueil.php> vous permettra de participer à des chasses aux trésors mathématiques. Bon amusement !

Vous trouverez sur le site de [Dunod](#) de nombreux autres titres divertissants de loisirs scientifiques.

Mesure de l'impact de la croissance économique sur le taux de chômage au Québec

Par Abdesselam Bouhia

[Lien vers le document](#)

Lien intéressant

Suggéré par Myrto Mondor

[Estimating popularity based on Google searches: Why it's a bad idea](#)