

Famille : Marié et deux beaux-fils de 11 et 15 ans. J'avoue que je trouve plus facile le programme de mathématique de 6^e année cette fois-ci.

Animal : « Spike », notre dragon barbu. Spike était le plus faible de sa portée et ne voit pas d'un œil. On doit le nourrir de force avec des fruits, des légumes, des comprimés de protéines et des grillons lyophilisés (Miam!). Mais c'est tout un grimpeur!

Sports/Passe-temps : Apprendre à jouer au hockey sur glace, la planche à neige, le tennis, l'entretien ménager et la menuiserie... Le camping, la randonnée, le canoë, le kayak, le bridge et bien sûr... le programme de math de 6^e année!

Ma fin de semaine idéale :

L'été :

1. Faire du canoë ou du kayak dans le Nord de l'Ontario avec mon épouse.
2. Approfondir mes compétences en menuiserie (et enrichir ma collection d'outils) en m'attaquant à un projet à la maison.

L'hiver :

1. Faire de la planche à neige dans l'abondante poudreuse fraîche.
2. Perdre la notion du temps en me plongeant dans un bon livre au coin du feu (et oui, nous avons un foyer).

Mets préférés : Au restaurant *Lahore Tikka* dans le quartier Little India de Toronto, sur Gerard St. East. Toute ma famille raffole de cet endroit!

Si je pouvais être quelqu'un d'autre (qu'un programmeur SAS), je voudrais être... menuisier je pense. Possédant une maison centenaire, j'ai découvert que j'aimais vraiment la rénovation et la restauration. Pas de construction neuve, mais je pense que j'aimerais rénover et restaurer de vieilles maisons.

Lorsque je ne fais pas de programmation SAS, j'aime ... faire mes devoirs de math, 6^e année ☺. N'importe quel des passe-temps énumérés précédemment.

Une chose que tout programmeur SAS devrait savoir ... c'est où trouver de l'aide pour programmer en SAS. Chaque programmeur SAS devrait

1. être à l'aise et se retrouver dans la documentation en ligne de SAS
2. lire la section *What's New ...* de la documentation en ligne de SAS après chaque installation de version de maintenance
3. savoir comment soumettre des questions à l'équipe de soutien technique sur le Web (<http://support.sas.com/techsup/intro.html>).
4. assister à des réunions de groupes d'utilisateurs internes ou externes.

Techno-conseil : Choix des échantillons *bootstrap* au moyen de PROC SURVEYSELECT

Les méthodes de *réplication* ou de *rééchantillonnage* comme les méthodes *bootstrap* (BS), *jackknife* (JK) et répliques répétées équilibrées (BRR) pour estimer la variance

d'une estimation ponctuelle ont connu une hausse de popularité. L'option `REP=nrep` dans `PROC SURVEYSELECT` vous permet de sélectionner facilement des échantillons répliques *nrep* de taille et de conception identiques. Ces répliques peuvent alors être utilisées à titre d'échantillons *bootstrap*.

Par exemple, le code `SURVEYSELECT` qui suit sélectionne 1000 répliques d'un échantillon aléatoire simple avec remise (échantillon aléatoire sans restriction : `METHOD=URS`). Chaque réplique est un échantillon à 5 % réparti proportionnellement entre les régions. L'échantillon original est spécifié comme la table d'entrée de l'option `DATA=`. Les 1000 répliques sont regroupées pour créer la table de sortie spécifiée par l'option `OUT=`. `OUTHITS` indique que la table de sortie doit contenir une observation distincte à chaque fois qu'une observation est sélectionnée plus d'une fois pour une réplique simple. La variable **numberhits** dans la table de sortie enregistre le nombre de fois que chaque observation est sélectionnée (pour chaque réplique) et la variable **replicate** indexe la réplique courante dans laquelle chaque sélection a été faite.

```
proc surveyselect    data=suvprb.HH_sample
                    samprate=.05
                    out=HH_BS_samples
                    method=urs
                    rep=1000
                    seed=13510
                    outhits;

strata region;
run;
```

Pour réussir à bien utiliser la méthode *bootstrap*, vous devez :

1. Imiter l'échantillon original expérimental au moment de sélectionner les répliques *bootstrap*. Cela inclut la conception de fonctions comme la stratification, le regroupement et la pondération inégale.
2. Ajuster les facteurs de pondération de l'échantillonnage d'origine pour chaque réplique pour rendre compte du plus petit échantillonnage. Voir l'article de Rust et Rao (1996).
3. Répéter tous les ajustements des facteurs de pondération après l'enquête/ expérience pour chaque réplique. Cela inclut les ajustements comme ceux qui sont requis pour le dénombrement incomplet et la non-réponse.

Autres lectures :

Efron, B. et Gong G., « A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation », *The American Statistician*, 1983, vol. 37, n° 1, pp. 36-48.

Izrael, D. et coll., « Use of the ROC Curve and the Bootstrap in Comparing Weighted Logistic Regression Models ». *SUGI 27 Proceedings* 2003, document n° 248.

<http://www2.sas.com/proceedings/sugi27/p248-27.pdf>

Rust, K. et Rao, J.N.K., « Variance Estimation for Complex Surveys Using Replication Techniques ». *Statistical Methods in Medical Research*, 1996, vol. 5, pp. 283-310.