

Family: Married with two stepsons, 11 and 15. I am finding the grade 6 math curriculum easier this time around.

Pets: “Spike”, our bearded dragon. Spike was the runt of his litter and is blind in one eye. He must be force fed fruits, vegetables, protein pellets and freeze-dried crickets (Yum!). But what a climber!

Sports/Hobbies: Learning to play ice hockey. Snowboarding. Tennis. Home maintenance and carpentry. Camping, hiking, canoeing and kayaking. Playing bridge. And, of course, the grade 6 curriculum.

My ideal weekend:

Summertime:

1. Kayaking or canoeing in northern Ontario with my wife.
2. Expanding my carpentry skills (and tool collection) via a home project.

Wintertime:

1. Snowboarding in abundant fresh powder.
2. Getting lost in a good book in front of a fire in my living room. (Yes, we have a fireplace).

Favourite Food: *Lahore Tikka* in Toronto’s Little India, Gerard St. East. My entire family loves this place!

If I could be anything at all (besides a SAS programmer), I would be ... I think I would like to be a carpenter. Owing a 100 year old home, I have learned that I really enjoy home renovation/restoration. Not new construction, but renovating/restoring older homes would be fun, I think.

When not programming in SAS, I like to ... do my grade 6 homework ☺. Any of the hobbies listed above.

One thing every SAS programmer should know ... is where to find SAS programming help. Every SAS programmer should

1. be comfortable navigating the SAS Online Documentation
2. read the *What’s New ...* section of the Online Documentation following each maintenance release
3. know how to submit questions to tech support from the web (<http://support.sas.com/techsup/intro.html>).
4. attend internal/external User Group meetings.

Techie Tip: Selecting Bootstrap Samples Using PROC SURVEYSELECT

Replication or *Resampling* methods such as the Bootstrap (BS), Jackknife (JK), and Balanced Repeated Replication (BRR) for estimating the variance of a point estimate have grown in popularity. The `REP=nrep` option in PROC SURVEYSELECT allows you

to easily select *nrep* replicate samples of the same size and design. These replicates can then be used as bootstrap samples.

For example, the following SURVEYSELECT code selects 1000 replicates of a with-replacement simple random sample (unrestricted random sampling: METHOD=URS). Each replicate is a 5% sample proportionally allocated across regions. The original sample is specified as the input data set on the DATA= option. The 1000 replicates are appended together to create the output data set specified on the OUT= option. OUTHITS requests a separate observation in the output data set for each selection of an observation that is selected more than once for a single replicate. The variable **numberhits** in the output data set records the number of times each observation is selected (within each replicate) and the variable **replicate** indexes the current replicate into which each selection was made.

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

To be successful using the bootstrap method, you must:

1. Mimic the original sample/experimental design when selecting the bootstrap replicates. This includes design features such as stratification, clustering, and unequal weighting.
2. Adjust the original sampling weights for each replicate to account for the smaller sample. See Rust and Rao (1996).
3. Repeat any post-survey/experiment weight adjustments for each replicate. This includes adjustments such as those for undercoverage and nonresponse.

Further Reading:

Efron, B. and Gong G. (1983), "A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation", *The American Statistician*, 37:1, 36-48.

Izrael, D. et al. (2003), "Use of the ROC Curve and the Bootstrap in Comparing Weighted Logistic Regression Models," *SUGI 27 Proceedings*. Paper 248.

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

Rust, K. and Rao, J.N.K. (1996), "Variance Estimation for Complex Surveys Using Replication Techniques," *Statistical Methods in Medical Research*, 5, 283-310.