



## Stats Happening - December 2012

1. **Multilevel Modeling Workshop in January**
2. **What is a p-value?**
3. **New Bayesian Procedure in SAS**
4. **Multilevel Complex Survey Analysis in Stata**
5. **A New Environment for R Users: RStudio**
6. **Structural Equation Models in JMP**

### 1. **Multilevel Modeling Workshop: Jan 17<sup>th</sup> & 18<sup>th</sup> 2013**

Start the spring semester early by attending the CSCU Multilevel Modeling workshop right before classes begin: Thursday January 17<sup>th</sup> and Friday January 18<sup>th</sup> 2013. Multilevel models (also referred to as mixed models or hierarchical models) are used when observations are not independent. Data clustering occurs with many experimental designs (e.g., split plot designs), with social science data collected simultaneously on different units of analysis (e.g., households and their individual members), with measurements taken on the same units over several time periods (longitudinal studies) and with spatial data. The purpose of this workshop is to introduce the concepts that form the basis of these models, the underlying statistical model and the estimation techniques that are used. Many examples will be presented during the workshop to enable participants to recognize such models when encountered in their own research, to analyze them and interpret the results. The workshop is intended for participants who have the equivalent of two semesters of statistics and some previous experience with ANOVA and linear regression. The workshop will be taught through a combination of lectures and hands-on computer exercises. The lecture will be on Thursday January 17<sup>th</sup> from 9:00am until 12:00am in 100 Savage Hall. For the hands-on session in Mann library you have a choice to sign up for Thursday 17<sup>th</sup> from 2:00-3:30pm or on Friday from 10:00-11:30am.

For more information and registration: <http://www.cscu.cornell.edu/workshops/multilevel.php>

### 2. **What is a p-value?**

Most research papers involving statistical testing use the p-value to determine significance. However, there are common misinterpretations of what the p-value is. We recommend two easy-to-read sources on the subject. Andrew Vickers' book "What is a p-value anyway? 34 Stories to Help You Actually Understand Statistics", explains many ideas in statistics through simple examples. In a 1988 paper titled "Statistical Illusion and the Illusion of Objectivity", published in American Scientist (Vol. 76), Berger and Berry focus on the p-value, how it is sometimes misinterpreted, and how it compares with the Bayesian approach.

### **3. Bayesian Procedure in SAS**

Have you been interested by Bayesian statistics but overwhelmed by the not too user friendly Bayesian software packages? Although Bayesian methods had previously been incorporated in several of the SAS procedures such as PROC GENMOD and PROC PHREG, SAS has released a new procedure that will allow you to solve many problems using a Bayesian approach. The MCMC procedure (PROC MCMC) is a general purpose Markov Chain Monte Carlo (MCMC) simulation procedure that estimates Bayesian models. Specifying a likelihood function for the data and a prior distribution for the parameters, Bayesian methods will estimate the posterior distribution of the parameters and present results in terms of summary and diagnostic statistics of these posterior distributions.

### **4. Multilevel Complex Survey Analysis in Stata**

Working with data obtained with a complex survey requires the use of specialized software to obtain correct point estimates and standard errors. The statistical analysis of complex survey data that has a multilevel structure such as repeated measures on the same person or clustering of the observations within schools present an additional complication. Until recently the software used by researchers to analyze complex surveys did not provide commands to estimate mixed models. Stata version 12 is attempting to fill this gap by allowing the possibility of including sampling weights and robust variance estimators with their xtmixed command designed to run mixed models.

### **5. A New Environment for R Users: RStudio**

RStudio is a free and open source integrated development environment that has the potential to make R coding much easier. You can run it on all major platforms (Windows, Mac, or Linux), and can even run it on a server accessed over the web. RStudio brings all of the R components into one place and keeps things nicely organized. RStudio includes a variety of powerful coding tools designed to increase your productivity, and enables quick navigation to files and functions. RStudio projects make it easy to manage multiple work contexts. Also, it supports interactive graphics using the “manipulate” package. To download RStudio and for more detailed information, please go to: <http://www.rstudio.com/>

### **6. Structural Equation Models in JMP**

If you are looking for a GUI to the SAS structural Equation Modeling (SEM) Procedure PROC CALIS, the new JMP version 10 might be exactly what you are looking for. The new menu-driven SEM graphic user interface for the latest version of SAS is very user-friendly and comes with good documentation that includes a detailed example enabling you to get started in a very short time. Unfortunately all this comes with a major drawback, namely that you need to have both software packages installed on your computer as the JMP SEM menu simply relies on SAS for the actual estimation. If you are an avid SAS user who needs to use the SEM framework, it might be well worth installing JMP version 10 on your computer to benefit from this very intuitive interface. PROC CALIS offers many different estimation methods but does not allow bootstrapping. If you have JMP and SAS on your computer and want to download the SAS add-in SEM package 1.3 for JMP, go to: [http://support.sas.com/demosdownloads/sysdep\\_t1.jsp?packageID=000818](http://support.sas.com/demosdownloads/sysdep_t1.jsp?packageID=000818)