On Monday May 1st, Dr. Christopher Gotwalt from the SAS Institute will be visiting UNH and giving two talks. Please see below for the titles, abstracts, times, and locations of each of the talks. A bio for Dr. Gotwalt is also provided below.

Presentation 1:
**When:** May 1st from 1pm to 3:00pm.
**Where:** M.S. Analytics building, 10 Pettee Brook Lane (next to Papa John’s and across from the Health Center).

**Text Exploration and Sentiment Analysis**

The new Text Explorer platform in JMP Pro 13 statistical software opens the door to new insights, providing an easy to use platform for discovery via text data. Our mission for JMP Pro 13 was to make Text Explorer a powerful tool that is accessible to as wide an audience as possible. The result is a unique “JMP-like” approach to text analysis that can be learned fairly quickly by anyone already familiar with JMP. We break the process of analyzing text analysis into three phases: feature creation, exploration, and modeling. The first phase, feature creation, is the process of using Text Explorer’s convenient interactive tools to clean up text data, correcting misspellings and identifying the key phrases to be used as terms in the analysis. The second phase, text exploration, utilizes the high performance analytical techniques such as the Singular Value Decomposition and Latent Class Analysis built into the Text Explorer platform for finding patterns in the data to identify themes in the data and clusters of documents that are similar to one another. In the third phase you use JMP Pro’s Text Scoring capability to save prediction formulas from the text analysis back to the data table. At that point you can unleash all of the rest of JMP Pro’s analysis and visualization capabilities on your text data. **We will demonstrate, using penalized regression, how to build predictive models and perform custom sentiment analyses** that use the data to determine which words and phrases are predictive of outcomes like purchasing behavior, review ratings, and future problems with products. Overall, most of text analysis uses concepts that are familiar to those with a modeling background, although the names are often different. As much as possible, we will translate text analysis concepts into terms familiar with traditional analysis techniques. Starting with simple, easy to understand, examples we will tackle progressively larger problems to clearly illustrate the fundamental ideas of text analysis and demonstrate Text Explorer’s capabilities. JMP Pro has made text analysis easy and broadly accessible, and this talk will be both informative and entertaining.

Presentation 2:
**When:** Monday May 1st from 3:30pm to 5:00pm
**Where:** The Mathematics and Statistics Department Conference Room, W390

**What is REML? Why Does it Work? And How Do We Extend it to the Generalized Linear Mixed Model?**

Restricted Maximum Likelihood (REML) estimates of the linear mixed model parameters is one of the most commonly used methods in statistics because REML estimates of variance components have lower bias than Maximum Likelihood estimates. Outside of balanced cases, the reason why it works so well has been in many ways a mystery. We present a new derivation for the REML method for the estimating the parameters of the linear mixed model. We demonstrate that in the case of variance models the REML estimator is an instance of a Firth adjusted estimator, a technique that reduces the bias of maximum likelihood estimators. The new derivation is advantageous because it suggests a new method for reducing the small sample bias of variance estimators in the generalized linear mixed model
(GLMM). It is well known that in small samples the standard methods of estimating variance parameters in GLMM’s are biased downward, leading to unacceptably high Type I error rates in tests of the fixed effects in these models. We apply the Firth technique to the one and two treatment logistic regression model with simple random effects. We demonstrate via simulation that, relative to estimators using pseudo likelihood and quadrature based maximum likelihood, the Firth estimates are both less biased and the Type I error rate of tests on the fixed effects is superior.

Biographical Information:

Christopher Gotwalt, Ph.D., Director of Statistical Research and Development, JMP Division of the SAS Institute.

Dr. Gotwalt has a Ph.D. in Statistics from North Carolina State University. He was the recipient of the inaugural Computational Statistics Fellowship awarded by the SAS Institute. Currently, Chris leads the statistical software development and testing teams. Since joining SAS as a Senior Research Statistician in 2003, Dr. Gotwalt has made numerous contributions to JMP, including developing algorithms for optimal design of experiments, neural networks, time-to-event modeling, measurement systems analysis, linear mixed models, multivariate analysis, optimization algorithms, and text analytics. Since 2007 he has led the analytical development of JMP, and as a Director has led successful development initiatives in manufacturing quality methodology, reliability modeling, consumer research and data mining. Dr. Gotwalt is also an Adjunct Professor in the statistics departments of two universities; the University of Nebraska and North Carolina State University, where he oversees the dissertation research of PhD students in the area of generalized linear mixed models, the analysis of large-scale computer simulation experiments, and big data approaches to imputing missing values.