A Note from the Editor of *MLRV*

T. Mark Beasley
University of Alabama at Birmingham

I would like to thank the members of the Multiple Linear Regression – General Linear Model Special Interest Group (MLR/GLM SIG) for affording me this opportunity to be the Editor of *Multiple Linear Regression Viewpoints* (*MLRV*). It is a great honor to have my name follow the likes of Isadore Newman, John Pohlmann, Keith McNeil, and Randall Schumacker. I am please to name Robin K. Henson of the University of North Texas as my Associate Editor. I believe he will be a great help.

Many changes have come along with this editorship. Namely, since being elected Editor of *MLRV*, I have taken a faculty position in the Department of Biostatistics at the University of Alabama at Birmingham (UAB). After eight years as an Assistant/Associate Professor teaching service courses in Statistics and Research Methods at St. John's University in New York, I felt it was time for a change and a new challenge. UAB presented me great prospects for research along with the opportunity to be closer to my family in my home state of Tennessee. I believe that the opportunities at UAB will help me advance my career as a statistician, but I also feel that the resources at UAB should help me to advance *MLRV* as a publication.

Several years ago, the members of the SIG officially changed the name to the Multiple Linear Regression – General Linear Model SIG. The impetus for this change was to incorporate a wider variety of topics for (a) proposals to the SIG for the American Educational Research Association (AERA) convention and (b) submissions to *MLRV*. In the 1990's, the SIG and *MLRV* recognized the linear model basis of Hierarchical Linear Models (HLMs) and Structural Equation Models (SEMs) and started including them as acceptable subject matter. To this extent, I propose further expansion to include *Generalized Linear Models*. Outside of the fact that it does not even require a change in acronym and we can still use the same monogrammed towels at the AERA convention, *Generalized Linear Models* subsume all the topics currently covered in *MLRV*.

For those of you who are not familiar with the topic, *Generalized Linear Models* are a broad class of models that include regression models for continuous dependent variables, alternative models for continuous dependent variables that do not assume normality or homoscedasticity of the residuals, and also models for discrete dependent variables (e.g., dichotomies, counts). Therefore, the multiple regression and ANOVA models typically covered in MLRV are special cases of *Generalized Linear Models*.

There are three basic components to *Generalized Linear Models*:

1. The *systematic component* specifies the explanatory variables used as the predictors, *X*. This is no different than the more familiar General Linear Models. The systematic component may include fixed effects, random effects, or may involve a mixed model as is common in longitudinal data and nested data structures (e.g., HLMs).

2. The *random component* identifies the dependent variable, *Y*, and specifies a probability distribution for the residuals. The normal distribution is but one of many known probability distributions that can be specified.

3. The *link function* specifies a function of the expected value of *Y*, E(*Y*) or $\hat{Y}$, to be predicted by the *X* variables. This component *links* the systematic and random components by specifying the connection between the *X* predictor variables and the predicted value of *Y*. That is, it specifies the predicted value of *Y* as a function of the *X* variables through an equation of linear form:

$$f(\hat{Y}) = \beta_0 + \beta_1 X_1 + \beta_1 X_1 + \ldots + \beta_k X_k \ .$$

For the familiar Ordinary Least Squares (OLS) linear regression model with a continuous *Y*, there is an *identity link function* (i.e., the predicted values are not transformed in any way) and the random error component (i.e., residuals) is assumed to have a normal distribution. However, in the closely related logistic regression model, Y is dichotomous and therefore has a binomial distribution. Thus, logistic regression is a *Generalized Linear Model* that specifies a *logit link function* that transforms $\hat{Y}$ to be a linear function of the X variables,

$$f(\hat{Y}) = \text{logit}(\hat{Y}) = \log\left(\frac{\hat{Y}}{1-\hat{Y}}\right),$$

and identifies a binomial distribution for the random error component.

For continuous dependent variables, it is possible to specify an error distribution other than the normal. For example, when the residuals display fan-shaped heteroscedasticity, it is possible that the variance of the residuals increases as a systematic function of $\hat{Y}$. One approach to this analytic problem is to transform the data in order to remove the heteroscedasticity. This approach is problematic if the researcher is interested in prediction and the variable "loses its meaning" when transformed. In an experimental design context, it is possible that a treatment may increase average level and reduce variability. Instead of transforming this "nuisance" into submission, a *Generalized Linear Model* can parameterize and model this interesting phenomenon. However, identifying the "correct" error distribution can be arduous, and therefore, nonparametric methods are still viable approaches.

Thus, the investigating the merits and statistical properties of these *Generalized Linear Models* relative to other procedures should provide a wealth of research opportunities for statisticians who conduct simulation studies. For example, if the simulation researcher creates a situation where the random error distribution is radically skewed, (s)he could compare *Generalized Linear Models*, with a variety of specified random components, to OLS regression and nonparametric methods. For applied statisticians and data analysts, manuscripts elaborating (a) how *Generalized Linear Models* relate to more familiar data analytic techniques or (b) how *Generalized Linear Models* can be applied to answer educational research questions are welcomed.

To my recollection, there has been only one paper specifically concerning *Generalized Linear Models* presented in the MLR/GLM SIG at AERA. I hope to see more and see such submissions to *MLRV*. This does not preclude other submissions, I simply want to expand the horizon for *MLRV*. This inclusion of *Generalized Linear Models* will not change the face of *MLRV* but hopefully enlarge the number of subscriptions and submissions.

Again, I thank the MLR/GLM SIG members for this wonderful opportunity and hope that I serve the editorship well.

Sincerely,

T. Mark Beasley, Editor
Multiple Linear Regression Viewpoints
Department of Biostatistics
University of Alabama at Birmingham
343C Ryals Public Health Building
1665 University Boulevard
Birmingham AL 35294
mbeasley@uab.edu