DM STAT-1 CONSULTING   BRUCE RATNER, PhD

574 Flanders Drive   North Woodmere, NY 11581   br@dmstat1.com
516.791.3544 fax 516.791.5075 1 800 DM STAT-1   www.dmstat1.com

What If There Were No Significance Testing?
Bruce Ratner, Ph.D.

What If There Were No Significance Testing? This was the question faced by early researchers at the turn of the 20th century, and it has revisited us several times over the years. As early as 1901, Karl Pearson laid the groundwork for assessing a scientific hypothesis with sample data. Later, Ronald Fisher (1925) formally proposed a set of methods that, with formulation from Neyman and Pearson (1928) on power, Type I and Type II errors, evolved into the practice of null hypothesis significance testing (NHST). ""[1]

The practice of NHST, however, has not been without controversy. Beginning with Berkson (1938), scientists have periodically engaged in anything from suggestions on how to improve NHST, to advocating a Bayesian approach, to revolutionizing the way we think about scientific inference, to *doing away with NHST all together.*

The main target of these interchanges is the practice of focusing (almost exclusively) on the probability values (p values) from NHST and incorrectly using them to make dichotomous decisions as to the truth of a null hypothesis that is almost certainly false, or an alternative hypothesis, the plausibility of which has not actually been evaluated.

If the statistics community would decide to do away with NHST all together, without providing a replacement for significance testing, how would you continue your statistical craft?

If you would like to share your comments, please email me.
Thanks.

PS: I have veered to the right for my statistical modeling using a modification of the "significance testing and all its apparatus," as found throughout my book *Statistical Modeling and Analysis for Database Marketing: Effective Techniques for Mining Big Data*. As for powerful predictive models, and unique transparent data mining, I have invented the GenIQ Model©, a nonstatistical, machine-learning model that solves the workaday statistical ordinary least squares and logistic regression models. The major advantage of the GenIQ Model over the current standard statistical regression models is that the inflexible statistical issues (e.g., pre-specify the unknowable model as the true model, data preparation and EDA, missing data and imputation, multicollinearity, outliers, etc.) are nonissues for a nonstatistical method.

Reference
1. Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (Eds.). *What If There Were No Significance Testing?* Lawrence Erlbaum Associates, Publishers, New Jersey, 1997.