2012 Hawaii University International Conferences
Education, Math & Engineering Technology
July 31st to August 2nd
Waikiki Beach Marriot Resort & Spa
Honolulu, Hawaii

Comparison Study of Handling Measurement Error Models: Regression Calibration, SIMEX and Quasi-Likelihood Variance Function Based on Empirical Likelihood Approach

Naima Shifa*, **Rong Wang**** and **Mamunur Rashid*****
Department of Mathematics
DePauw University
Department of Mathematics
DePauw University
Department of Mathematical Sciences
Indiana University-Purdue University Indianapolis
Naima Shifa
Department of Mathematics
DePauw University
Greencastle, IN 46135 USA
naimashifa@depauw.edu, rongwang_2012@depauw.edu

Comparison Study of Handling Measurement Error Models: Regression Calibration, SIMEX and Quasi-Likelihood Variance Function Based on Empirical Likelihood Approach

Statistics

Naima Shifa[*], Rong Wang[**] and Mamunur Rashid[***]

Department of Mathematics
DePauw University

Department of Mathematics
DePauw University

Department of Mathematical Sciences
Indiana University-Purdue University Indianapolis

Naima Shifa
Department of Mathematics
DePauw University
Greencastle, IN 46135 USA

naimashifa@depauw.edu, rongwang_2012@depauw.edu

Abstract

Statistical model analysis is helpful in assessing specific form of the relationship between variables and the ultimate goal is to predict or estimate the value of one variable corresponding to a given value of another variable. In analyzing data in all disciplines, it is frequently desirable to learn something about the relationship between two numeric variables. For example, we may be interested in studying the relationship between yield of corn and available nitrogen in the soil, the consumption level of some nutrient and weight gain, reproduction of a particular strain of bacterium and temperature over a specified period of time, blood pressure and age, the intensity of a stimulus and reaction time, or total family income and medical care expenditures. By statistical model analysis, another word, by regression analysis we examine the nature and the strength of the relationships between variables. In fact, in many situations we cannot observe the exact value of the independent variables; like, we cannot detect the exact amount of soil nitrogen in a particular land; to estimate the available soil nitrogen, we need to sample the soil of the plot of land and do some laboratory analysis for the samples. As a result of sampling and laboratory analysis we can obtain an estimate of the amount of nitrogen but still the true amount is unknown. Besides, the content of nitrogen is quite different in different part of the field because of the application of varying rates of fertilizer and the growing of the different fraction of legumes. Consequently, we cannot observe the exact quantity of nitrogen in a particular field; we only observe the estimated amount, therefore, the variable "Amount of Soil Nitrogen" is subjected to error. In epidemiological study, if we examine the relationship between blood pressure and pulse rate, a single measurement of blood pressure, considered as a measure of an individual's underlying average blood pressure. A single measurement of blood pressure is not good enough to explain the true health condition of the person since it depends on the temperature of the place, time of the day and the work load of the individual at that moment. In a physics laboratory, weighing an object by using a spring balance might show some variation in measurement due to fluctuations in temperature, conditions of loading and unloading, etc. A measuring instrument with a higher precision means there will be lesser fluctuations in its measurement. In fact, there are lots of situations where the measurements of the variables of interest are noisy error-prone version of the quantity of primary interest. Such errors are called measurement errors and can arise due to lack of knowledge of the respondent, technical error induced by imperfect measurement instruments and short-term fluctuations over time. The

consequences of measurement error in a variable depend on the variable's role in the substantive model of interest; measurement error in the independent variables of statistical models does cause bias, in general. Measurement error in a variable of interest may distort estimates of the variable's effect on the outcome of interest, and may not disappear even if the sample size increases to infinity.

There are some techniques currently available: regression calibration, simulation extrapolation (SIMEX), and the score function method for correcting effect estimates in the presence of measurement error. In this research work, a comparison study will be done among the existing techniques and a newly proposed simulation-based approach. This simulation-based method is developed by considering quasi likelihood variance function (QVF) and empirical likelihood approach. In this article, we will describe all the available techniques to manage measurement errors in the models and finally we will compare the length of the confidence intervals of the error prone regression parameters obtained from these techniques.