

## Tests for Zero Inflation in Count Data with Measurement Error in Covariates

**Authors :** Man-Yu Wong, Siyu Zhou, Zhiqiang Cao

**Abstract :** In quality of life, health service utilization is an important determinant of medical resource expenditures on Colorectal cancer (CRC) care, a better understanding of the increased utilization of health services is essential for optimizing the allocation of healthcare resources to services and thus for enhancing the service quality, especially for high expenditure on CRC care like Hong Kong region. In assessing the association between the health-related quality of life (HRQOL) and health service utilization in patients with colorectal neoplasm, count data models can be used, which account for over dispersion or extra zero counts. In our data, the HRQOL evaluation is a self-reported measure obtained from a questionnaire completed by the patients, misreports and variations in the data are inevitable. Besides, there are more zero counts from the observed number of clinical consultations (observed frequency of zero counts = 206) than those from a Poisson distribution with mean equal to 1.33 (expected frequency of zero counts = 156). This suggests that excess of zero counts may exist. Therefore, we study tests for detecting zero-inflation in models with measurement error in covariates. Method: Under classical measurement error model, the approximate likelihood function for zero-inflation Poisson regression model can be obtained, then Approximate Maximum Likelihood Estimation (AMLE) can be derived accordingly, which is consistent and asymptotically normally distributed. By calculating score function and Fisher information based on AMLE, a score test is proposed to detect zero-inflation effect in ZIP model with measurement error. The proposed test follows asymptotically standard normal distribution under  $H_0$ , and it is consistent with the test proposed for zero-inflation effect when there is no measurement error. Results: Simulation results show that empirical power of our proposed test is the highest among existing tests for zero-inflation in ZIP model with measurement error. In real data analysis, with or without considering measurement error in covariates, existing tests, and our proposed test all imply  $H_0$  should be rejected with P-value less than 0.001, i.e., zero-inflation effect is very significant, ZIP model is superior to Poisson model for analyzing this data. However, if measurement error in covariates is not considered, only one covariate is significant; if measurement error in covariates is considered, only another covariate is significant. Moreover, the direction of coefficient estimations for these two covariates is different in ZIP regression model with or without considering measurement error. Conclusion: In our study, compared to Poisson model, ZIP model should be chosen when assessing the association between condition-specific HRQOL and health service utilization in patients with colorectal neoplasm. and models taking measurement error into account will result in statistically more reliable and precise information.

**Keywords :** count data, measurement error, score test, zero inflation

**Conference Title :** ICSA 2017 : International Conference on Statistics and Analysis

**Conference Location :** Singapore, Singapore

**Conference Dates :** September 11-12, 2017