Sample Size Calculation for Assessing Differential Expression Analysis in RNA-seq Data

Chung-I Li
Department of Applied Mathematics, National Chiayi University

Sample size determination is an important issue in the experimental design of biomedical research. Because of the complexity of RNA-seq experiments, the field currently lacks a sample size method widely applicable to differential expression studies utilizing RNA-seq technology. In this research, we propose sample size calculation formulas based on several test statistics, such as Wald test, score test and likelihood ratio rest, under Poisson distribution assumption for assessing differential expression analysis in RNA-seq data. For dealing with over-dispersion issue, we further propose sample size calculation method based on exact test under negative binomial distribution assumption. Simulation studies to evaluate the performance of the proposed sample size formulas are presented; the results indicate our method works well, with achievement of desired power. Finally, our sample size calculation methods are applied to two real RNA-seq data sets.

Keywords: sample size calculation, false discovery rate, RNA-seq