Yu Shyr et al., J Biom Biostat 2014, 5:4
http://dx.doi.org/10.4172/2155-6180.S1.022

OMICS Group Conferences
Accelerating Scientific Discovery

3rd International Conference & Exhibition on
Biometrics & Biostatistics
October 20-21, 2014 DoubleTree by Hilton Baltimore - BWI Airport, USA

Sample size calculation based on exact test for differential expression analysis of RNA-seq data

Yu Shyr and Harold L Moses
Vanderbilt Center for Quantitative Sciences, USA

Sample size calculation is an important issue in the experimental design of biomedical research. For RNA-seq experiments, the sample size calculation method based on the Poisson model has been proposed; however, when there are biological replicates, RNA-seq data could exhibit variation significantly greater than the mean (i.e., over-dispersion). The Poisson model cannot appropriately model the over-dispersion, and in such cases, the negative binomial model has been used as a natural extension of the Poisson model. Because the field currently lacks a sample size calculation method based on the negative binomial model for assessing differential expression analysis of RNA-seq data, a method to calculate the sample size was proposed based on the exact test for assessing differential expression analysis of RNA-seq data.The proposed sample size calculation method is straightforward and not computationally intensive. Simulation studies to evaluate the performance of the proposed sample size method are presented; the results indicate the method works well, with achievement of desired power.

Biography

Yu Shyr is the Director for the Vanderbilt Center for Quantitative Sciences, received his PhD in Biostatistics from the University of Michigan (Ann Arbor) in 1994 and subsequently joined the faculty at Vanderbilt University School of Medicine. He is a Fellow of the American Statistical Association and an FDA advisory committee voting member. He has presented over 190 abstracts and has published more than 300 peer-reviewed papers. His research interests lie in developing and analyzing predictive models of the statistical relationships between multiple-variable protein and next-generation sequencing data and clinical endpoints using both supervised and unsupervised classification and pattern recognition approaches.

yu.shyr@Vanderbilt.Edu