DETERMINING THE DIFFERENTIALLY EXPRESSED GENES BY BAYESIAN DECISION RULE

by

Tao Cui, B.S., M.S.

A Dissertation submitted to the faculty of the Graduate School,
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

Milwaukee, Wisconsin

August 2024

ABSTRACT
DETERMINING THE DIFFERENTIALLY EXPRESSED GENES BY BAYESIAN DECISION RULE

Tao Cui, B.S., M.S.

Marquette University, 2024

The identification of differentially expressed genes (DEGs) is crucial for understanding the molecular mechanisms underlying various biological conditions and diseases. Traditionally, this process utilizes statistical methods which manage high-dimensional data and adjust for multiple testing to ensure statistical and biological relevance. These methods assume the independence of genes, which is not usually the case, many of genes are correlated when there is a stimila. In this dissertation, we will work on the methods that could consider the correlation in finding the DEGS to improve the power of tests.

We have developed a novel approach to multiple hypothesis testing using Bayesian decision rules that account for correlation effects among targeted genes, transitioning from traditional 0-1 loss functions to a more flexible 0-K loss to control the False Discovery Rate (FDR). Our methodology includes innovative parameter estimation techniques tailored for the complexities of gene expression data. We rigorously validated our approach through extensive simulations that demonstrate the efficiency and robustness of our tests. Specifically, we applied our methods to COVID-19 data to highlight potential improvements in identifying differentially expressed genes. Additionally, we expanded our framework to include analyses based solely on p-values, allowing our methods be be used for many published studies where only p-values are available.. This comprehensive approach not only enhances the statistical power but also provides a more accurate tool for biological discovery in genomics research.