

Taxonomic Enrichment Analysis with Isometric Log-Ratios

Quang P. Nguyen^{1,2}, Anne G. Hoen^{1,2}, and H. Robert Frost¹

¹*Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth College, Hanover, NH 03755, USA*

²*Department of Epidemiology, Geisel School of Medicine at Dartmouth College, Hanover, NH 03755, USA*

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Pellentesque arcu libero, suscipit sed enim nec, posuere eleifend nunc. Phasellus vitae augue orci. Sed vestibulum nisi id augue porta, a sagittis magna accumsan. Sed augue mi, venenatis sed fringilla nec, ultrices id ipsum. Suspendisse ac sapien eu mi laoreet fringilla. Nulla facilisi. Sed eget feugiat erat, et efficitur risus. Duis sit amet nulla at leo dignissim porta. Morbi nec ligula non sapien fringilla congue. Proin consequat volutpat nulla, eu convallis leo tempor in. Mauris elit sem, dignissim sit amet sapien sed, varius laoreet felis. Etiam elementum vulputate justo non malesuada. Suspendisse a libero id massa pellentesque convallis at in nibh. Ut nec consequat ante, vitae convallis dui. Ut eu pharetra nisi.

Background

Methods

Simulation Analysis

We chose to simulate data according to [1]

Real data analysis

We benchmark type I error rate on 16S and WGS data from the Human Microbiome Project (HMP) obtained from the packages *HMP16SData* (ver. 1.9.3) and *curatedMetagenomicData* packages in R. For each data set, we filtered out samples with less than 1000 total number of reads, as well as taxa with a proportion of zeroes of 0.9 or more. We then randomly assigned each sample into one of two arbitrary groups.

References

- [1] Michael B. Sohn, Ruofei Du, and Lingling An. A robust approach for identifying differentially abundant features in metagenomic samples. *Bioinformatics*, 31(14):2269–2275, July 2015.