

Supplemental Information: Examples of sLDA Permutation Distributions

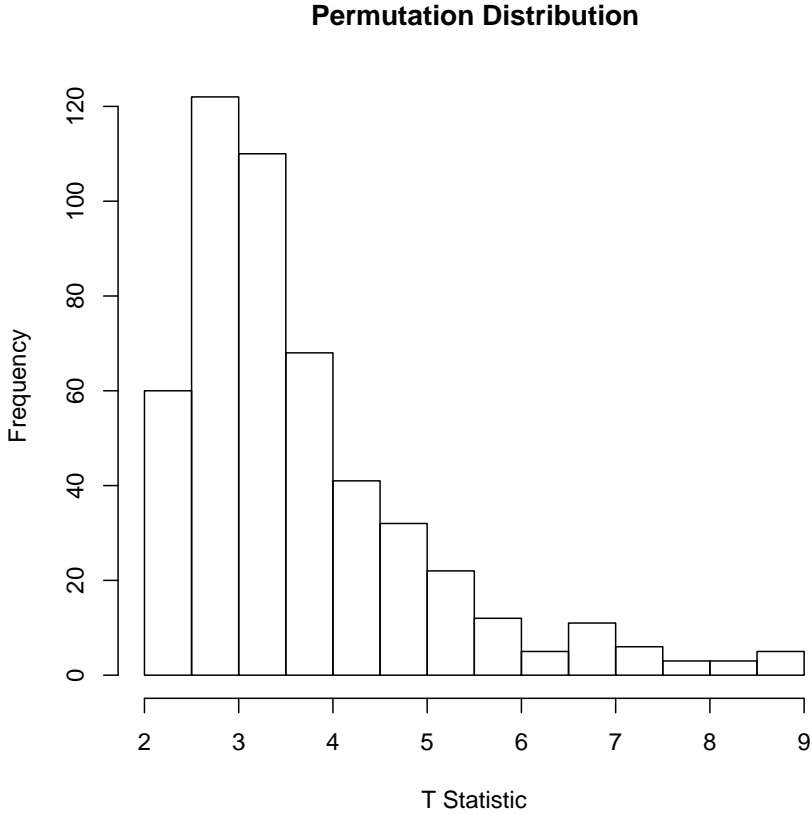


Figure 1: The distribution of the $|T|$ statistics found under each permutation for a single data set generated under simulation Setting 2. Because the sLDA approach is a supervised method, the T does not follow a usual t -distribution and permutation is necessary in order to generate the null distribution.

Number of Genes Selected under Permutation

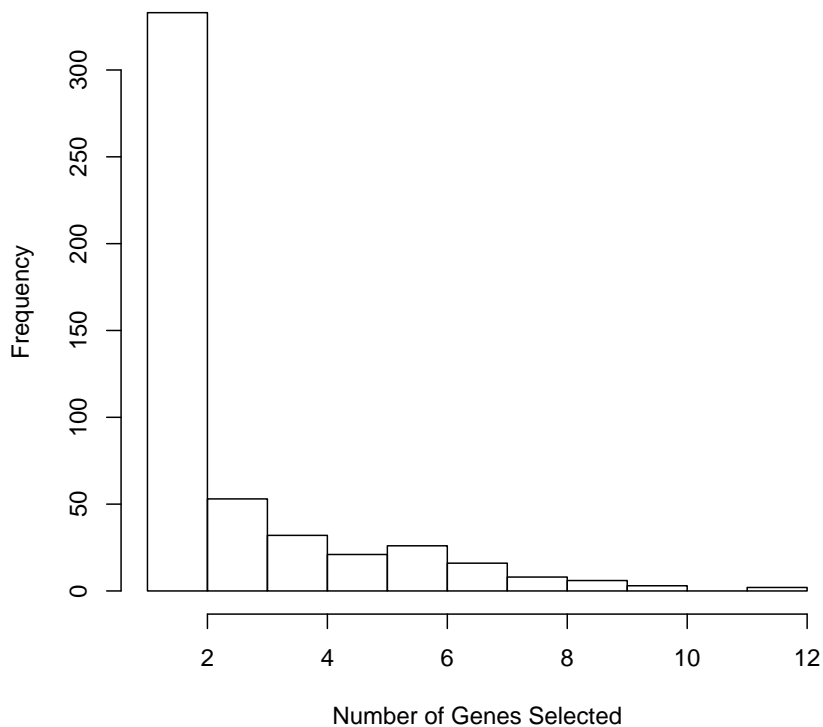


Figure 2: The distribution of the number of gene selected by sLDA under each permutation for a single data set generated under simulation Setting 2. Under permutation, we do not expect an association between the gene set and the outcome so we expect relatively few genes to be selected. We see that in most cases (65%), only a single gene was selected. Our approach restricts the number of genes selected to be at least 1 since the problem is not well defined if no genes are selected.