

CSE 5307 Bioinformatics Homework 1

Name: Venkat Ankit Gundala

Student ID: 1002069069

1.2 Fisher's Exact Test

Null hypothesis:

The null hypothesis for the above study that neither of the two classifications, T alleles nor C alleles, belong to a complex trait.

Alternative Hypothesis:

As stated in the question, if the odd ratio for allele C is greater than one, then allele C is one of the causes of the complex genetic trait. I used alternative value as greater because the odd ratio is greater than one to find the odd ratio and p value.

The result.csv file has the p values and SNPs that are Significant as first and second column respectively. The Significant column consist of True and False values. The SNP which has true value is significant.

The number of significant SNP can be written as number of true in significant SNPs column.

The significant count is 164

1.3 Corrected P-Values

Given effective p-value is 5×10^{-8} Bonferroni corrected p value is calculated by dividing effective p value and number of tests performed. bonferroni corrected $p = p \text{ effective} / n$.

Where n is the number of test performed for this problem are 1000 Then, Bonferroni correct p value is $4.999999999999995e-11$.

The number of SNPs that are significant under the corrected p value

is 114 Including the SNPs that are significant under the significant as the third columns in the result.csv file.

1.4 Manhattan Plot

Taking the x axis as the SNPs and the y axis as the $-\log_{10}(p \text{ value})$. Considering two values T1 and T2. Where the T1 indicates the p effective on the plot and T2 is drawn considering the p value as Bonferroni corrected p value. Differentiating two lines using different colors red as p effective and bonferroni corrected p is indicated with green color.

The graphs help understand the difference between p effective and bonferroni corrected p value.

The SNPs above Bonferroni corrected p value states that the null hypothesis is wrong. The SNPs above p value reject the null hypothesis they might be few SNPs may accept at certain conditions like false positive.

Difficulty Adjustment

It took 20 hours to complete this assignment.
I was struck while doing the Fisher's exact Test.