Lecture 8 Exercises:

ALL of the code and scripts are inside of the file: Lecture_8_code.txt. Because the code for doing some of these tasks is quite involved, please copy and paste from that file, but do make efforts to understand what the code is generally doing.

Exercise 1. Read in the variants within "Gene1.txt" into R. Calculate the minor allele frequencies of the variants and plot a histogram. What do you notice?

Exercise 2. Read in the trait value in the file "Trait1.txt". Then test for an association between each variant in Gene1 and the Trait1. Is anything significant after adjusting for multiple testing?

Exercise 3. Test for an association between the rare variants in Gene1 and the quantitative trait. For now, let's define rare variants to be the variants with MAF < 3%. Specifically, please apply

- CAST (Binary collapsing approach)
- MZ Test/GRANVIL (Count based collapsing)
- Weighted Count Based Collapsing where weights can depend on MAF

Exercise 4. Repeat Exercise 3 using the variants inside of "Gene2.txt" and "Trait2.txt".

Exercise 5. Apply CAST and/or GRANVIL to test for an association between the variants inside of "Gene3.txt" and "Trait3.txt". Let's consider alternative definitions for rare variant: MAF < 5%, < 3%, < 1%, < 0.5%.