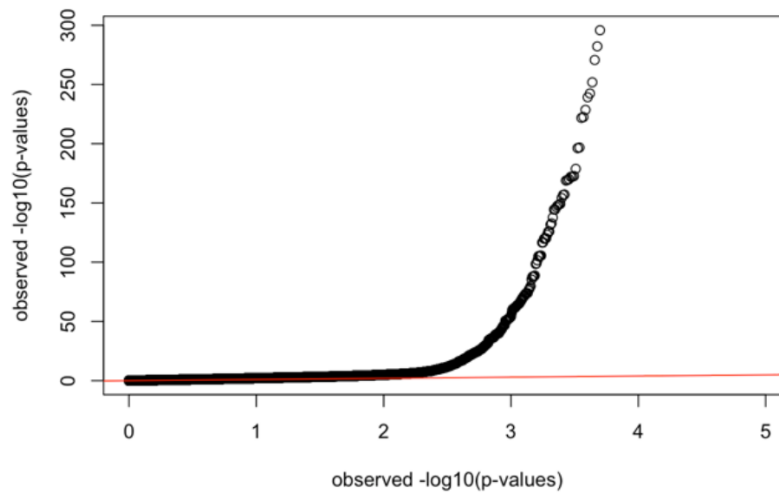
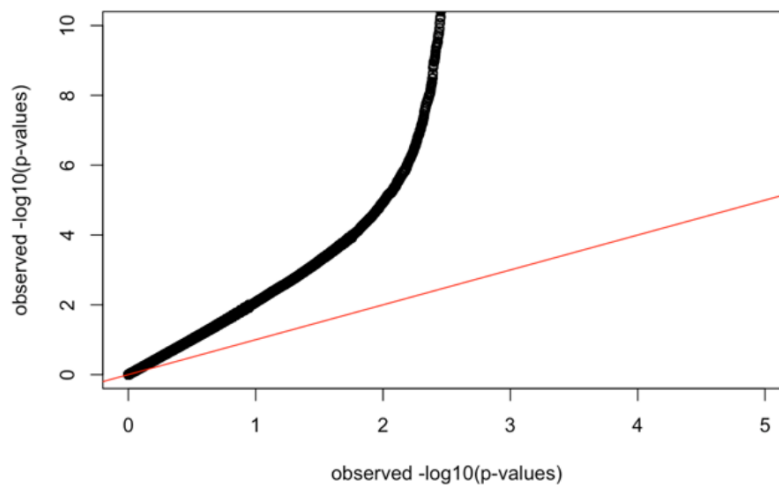


CMM 262 2021

Homework 3: Genetic Variation and GWAS (Graham McVicker)



Upon zoom in:



Instructions

Answer the following questions in your own words and upload a PDF of your answers to Gradescope. This assignment is due **3/10/21 at 9:00AM**.

1. Under the null hypothesis, how should the p-values be distributed (i.e. What distribution should they come from)? (1 point)
2. Do you notice a problem with the p-values from the zoomed-in plot? (1 point)
3. Why is this a problem?(2 points)
4. What could cause this problem with the p-values? (2 points)
5. How could you correct this problem? (2 points)

Bonus Question. An association between a SNP and a phenotype can be computed with either a linear or a logistic regression. Explain the difference between the two and for each, give a corresponding phenotype you would analyze with that type of regression. (1 point)