BSG-MDS practical 5 Statistical Genetics

Name Surname Name Surname

05/12/2023, submission deadline 12/12/2023

Resolve the following exercise in groups of two students. Write the R scripts, perform the computations and make the graphics that are asked for in the practical below. Take care to give each graph a title, and clearly label x and y axes, and to answer all questions asked. You can write your solution in a Word or Latex document and generate a pdf file with your solution, or generate a solution pdf file with R Markdown. Take care to number your answers exactly as in this exercise. Upload your solution in pdf format to the web page of the course at raco.fib.upc.edu no later than the submission deadline.

You can make use of the R-package genetics (and other packages) to compute your answers, as you please. The datasets can be downloaded from the web page of the course at raco.fib.upc.edu.

Genetic Association Analysis (10p)

A particular SNP is supposed to be involved in Alzheimer's disease. A case control study has been performed, obtaining the following results:

	AA	Aa	aa
Cases	112	278	150
Controls	206	348	150

1. (2p) Perform the alleles test for this data set. Provide the p-value and the odds ratio and comment on the results.
2. (2p) Test for association using a codominant, a dominant and a recessive model. Provide the p-values for all the tests and comment on the results.
3. (2p) Plot the risk of disease as a function of the number of m alleles. Comment on the results. Which model seems most appropriate?
4. (2p) Perform Armitage trend test for this data set. Does the null hypothesis $\beta_1 = 0$ hold? Comment on your response.
5. (2p) Is there evidence for association of this marker with the disease? Argument your response.