

Module 4: Regression Methods: Concepts and Applications

Lab 3: One-Way and Two-Way ANOVA

The goal of this lab is to answer the following scientific questions using the cholesterol dataset:

- Is rs4775401 associated with cholesterol levels?
- Are rs174548 and APOE associated with cholesterol levels?
- Does the effect of APOE on cholesterol levels depend on rs174548?

The cholesterol data set is available for download from the module Github repository and contains the following variables:

ID: Subject ID

sex: Sex: 0 = male, 1 = female

age: Age in years

chol: Serum total cholesterol, mg/dl

BMI: Body-mass index, kg/m²

TG: Serum triglycerides, mg/dl

APOE: Apolipoprotein E genotype, with six genotypes coded 1-6: 1 = e2/e2, 2 = e2/e3, 3 = e2/e4, 4 = e3/e3, 5 = e3/e4, 6 = e4/e4

rs174548: Candidate SNP 1 genotype, chromosome 11, physical position 61,327,924. Coded as the number of minor alleles: 0 = C/C, 1 = C/G, 2 = G/G.

rs4775401: Candidate SNP 2 genotype, chromosome 15, physical position 59,476,915. Coded as the number of minor alleles: 0 = C/C, 1 = C/T, 2 = T/T.

HTN: diagnosed hypertension: 0 = no, 1 = yes

chd: diagnosis of coronary heart disease: 0 = no, 1 = yes

You can download the data file and read it into R as follows:

```
cholesterol = read.csv("https://raw.githubusercontent.com/rhubb/SISG2018/master/data/SISG-D
ata-cholesterol.csv", header=T)
```

Install R packages

- For this lab you will need the *gee* and *multcomp* packages.
- If you have not already, install the packages first. You will then need to load the libraries each time you execute your R script.

```
install.packages("gee")
install.packages("multcomp")
library(gee)
library(multcomp)
```

Exercises

1. Perform a descriptive analysis to investigate the scientific questions of interest using numeric and graphical methods.
2. Conduct an analysis of differences in mean cholesterol levels across genotype groups defined by rs4775401. Is there evidence that mean cholesterol levels differ across genotypes? If so, perform all pairwise multiple comparisons using Bonferroni's adjustment. Try out different adjustment methods.
3. Compare results obtained using classical ANOVA to those based on ANOVA allowing for unequal variances, using robust standard errors, and using a nonparametric test.
4. Perform a descriptive analysis to investigate the relationships between cholesterol, APOE and rs174548. Conduct an analysis to investigate the association between cholesterol, APOE and rs174548, with and without an interaction between APOE and rs174548. Is there evidence of an interaction between APOE and rs174548?

Once your group has completed the lab exercises, please submit your R script file to the class Github repository:

<https://github.com/rhubb/SISG2018/tree/master/submit> (<https://github.com/rhubb/SISG2018/tree/master/submit>)

Sign in using the class username and password. Then click upload files to save your R script file to the repository.