

AMS394 Midterm2

October 13, 2021

This is an open book exam. However, no communication is allowed between students. Please provide complete codes and conclusions for full credit. Good luck!

Question 1 (35 points):

Use the “d_logret_6stocks” dataset to answer the questions. Test by using $\alpha = 0.05$.

- (1) Regress the return of Pfizer on the returns of AmerExp (with intercept). Report the estimated coefficients.
- (2) Generate an ANOVA table to conclude if regression effects are significant.
- (3) Compute the correlation of Pfizer and AmerExp, and test if their correlation is zero.
- (4) Test if the proportion of returns of AmerExp greater than Intel is 0.7.
- (5) We now ignore the time series features of all returns and consider them independent. We also treat all weekly returns of Exxon, AmerExp and Pfizer as Groups 1, 2, and 3, respectively. Perform one-way ANOVA for Groups 1-3.

Question 2 (20 points):

Use the ‘cement’ dataset in ‘MASS’ package to answer the question.

- (1) Conduct the multiple linear regression, regress y value on x1, x2, x3 and x4 (without intercept). Report the estimated coefficients.
- (2) What is the adjusted R square of your regression?
- (3) Conduct a best subset regression (with intercept) with the function ‘regsubsets’ and find the best model with two independent variables.

Question 3 (25 points):

We want to test the body weight of three different species of cicadas. The body weights are measured in grams. There are three species of cicadas: tredecula, tredecassini, tredecim. Please write up the R code to do the following.

Species	Body weights
tredecula	0.09, 0.17, 0.19, 0.16, 0.27, 0.15, 0.25, 0.16, 0.14
tredecassini	0.28, 0.15, 0.22, 0.11, 0.20, 0.17, 0.23, 0.18, 0.23
tredecim	0.28, 0.29, 0.14, 0.18, 0.32, 0.39, 0.33, 0.26, 0.12

- (1) Please check the normality of the body weight of each species of cicadas at significance level 0.1.
- (2) Compare the variance of the body weights of different species of cicadas with Bartlett’s test at significance level 0.1.
- (3) Draw a plot with ‘stripchart’ function of body weights versus species of cicadas with ‘jitter’ method.
- (4) Perform one-way ANOVA for the body weight of different species of cicadas and conclude if the means of body weights of different species of cicadas are the same at significance level 0.1.

Question 4 (20 points):

(1) Construct the following table that summarizes the number of people who smoke in a company.

Age	Smoke	
	Yes	No
20-30	4	6
30-40	10	18
40-50	14	20
50-60	5	10

(2) Use the Chi-Square test to conclude if the group (smoke and non-smoke) and the Age range effects are independent. Test by using $\alpha = 0.1$.

(3) Report expected value of each cell under the null hypothesis (smoke and age are independent).