
Group:**Name:****Problem 1 (22 pts)**

The takings of five different vendors at a street market were monitored over the seven different days of the same week. The total revenue each vendor made over the week measured in pounds was 1017 for A, 965 for B, 477 for C and 408 for D. The value for vendor E is not given. The following information is provided as well.

Source	DF	Sum of Squares	Mean Square	F-score	F_{crit} (5%)
Day					
Vendor				4.17	
Error			3876		
Total		303678			

- (a) **(8 pts)** Complete the table using the information provided above.
- (b) **(2 pts)** Is there a significant difference between the daily takings of different vendors? What about the takings of all vendors on different days?
- (c) **(4 pts)** Construct a 90% confidence interval for the difference between the daily takings of vendors A and D. Would you say there is evidence of a difference?
- (d) **(8 pts)** Construct 95% simultaneous confidence intervals for the difference of daily takings between vendors A and D and the difference between daily takings of vendors B and C.

Group:**Name:****Problem 2 (22 pts)**

A study looked into the views of workers towards a government policy of school closures in order to reduce coronavirus transmission. 300 participants from three sectors (hospitality, banking and construction) were interviewed and their responses were classified into three categories, namely positive, neutral and negative. The data are shown below.

	View Towards School Closure			Total
	Positive	Neutral	Negative	
Hospitality	35	25	40	100
Banking	28	47	25	100
Construction	38	22	40	100
Total	101	94	105	300

- (a) **(3 pts)** Based on the data in the table, and without conducting a significance test, describe the differences of views across the different sectors.
- (b) **(8 pts)** Calculate the χ^2 statistic and use it to test for independence, using a 5% significance level. What do you conclude?
- (c) **(3 pts)** Based on a sample of 25 observations, the population regression model

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

was estimated. The least squares estimates obtained were

$$\hat{\beta}_0 = 15.6, \quad \hat{\beta}_1 = 1.3.$$

The total and error sums of squares were $SST = 268$ and $SSE = 204$.

Find and interpret the coefficient of determination.

- (d) **(8 pts)** Find a 95% confidence interval for β_1 .

Group:**Name:****Problem 3 (22 pts)**

The car rental company is interested in the relationship between the mileage of cars X and the cost of monthly maintenance Y . To clarify the nature of this connection, 8 cars were selected.

Car	1	2	3	4	5	6	7	8
Mileage	6	7	8	10	13	17	18	20
Maintenance cost	13	16	15	19	24	35	34	39

- (a) **(2 pts)** Graph the source data and comment the type of the dependence.
- (b) **(5 pts)** Calculate the sample correlation coefficient, test its significance (on non-zero value) at $\alpha = 0.05$.
- (c) **(5 pts)** Using the OLS method estimate a linear regression coefficients and graph the regression line.
- (d) **(10 pts)** The data below shows lengths (inches) of fish caught in one day in a river. Construct a histogram for these data (Do not use more than seven categories along the x axis). Label it carefully.

15.2 15.6 14.6 10.5 13.7 10.6 13.9 14.7 13.9 10.6
 10.3 10.6 12.1 12.8 14.9 15.2 13.8 14.5 15.0 11.7
 10.9 18.4 14.1 14.9 10.2 13.9 11.6 13.9 12.5 11.1
 15.1 14.9 15.0 12.6 10.9 13.7 12.9 12.1 14.0 13.5

Group:**Name:****Problem 4 (10 pts)**

An agency offers students preparation courses for an achievement test. As a part of an experiment to evaluate the merits of the course, twelve students were chosen and divided into six pairs, in such a way that the two members of any pair had similar academic record. Before taking the test, one member of each pair was assigned at random to take the preparation course, while the other member took no course. The achievement test scores are shown in the accompanying table. Assuming that the differences in scores follow a normal distribution, test at 5% level the null hypothesis that the two population means are equal, against the alternative that the true mean is higher for students taking the preparation course.

Student pair	1	2	3	4	5	6
Course	82	73	59	48	69	93
No course	75	71	52	46	70	83