

HW 2

Exercise 2.1 (Sweetness of Orange Juice). (Page 105 Question 3.13, Dataset: OJUICE)

- (a) Fit the model and find a 90% confidence interval for the true slope of the line. Interpret the result.
- (b) Fit the model and determine whether there is a positive or negative linear relationship between the amount of pectin x and the sweetness y . That is, determine if there is sufficient evidence (at $\alpha = 0.05$) to indicate that β_1 , the slope of the straight-line model, is significantly different from zero.

Exercise 2.2 (Joint Strike Fighter program). (Page 105 Question 3.12, Dataset: F35)

- (a) Fit the simple linear regression model, $E(y) = \beta_0 + \beta_1$, to the data.
- (b) List assumptions required for the regression analysis.
- (c) Find the value of SSE.
- (d) Find the estimated standard error of the regression model, s .
- (e) Give a practical interpretation of s .
- (f) Find a 95% confidence interval for the true slope of the line.
- (g) Interpret the confidence interval in (f).
- (h) Find the p -value for testing $H_0 : \beta_1 = 0$ versus $H_a : \beta_1 \neq 0$. Use this result to test the simple linear regression model is statistically useful for predicting the annual cost using the year of initial operation. (Test using $\alpha = 0.05$)
- (i) Find and interpret the coefficient of determination, r^2 .
- (j) A researcher wants to estimate of the average annual cost of all the military aircraft with the year of initial operation in 1980. Which interval is desired by the researcher, a 95% prediction interval for y or a 95% confidence interval for $E(y)$? Use R to calculate the desired interval.
- (k) Give a practical interpretation of the interval in part (j).

Exercise 2.3. Fill in the blanks in the table and answer questions:

For future planning and budgeting, the researchers want to analyze the relationship between the total area of structurally deficient bridges in a state and the number of deficient bridges. A simple linear regression model was fitted. In this analysis,

- x = number of structurally deficient bridges,
- y = the total area (thousands of square feet) of the deficient bridges.
- The MINITAB output is as follows (don't worry about MINITAB, just look at the output table:)

Regression Analysis: SDArea versus NumberSD

The regression equation is
 $SDArea = 120 + 0.346 \text{ NumberSD}$

Predictor	Coef	SE Coef	T	P
Constant	119.9	123.0	0.97	0.335
NumberSD	0.34560	0.06158	(1)	(2)

$S =$ (3) $R\text{-Sq} = 38.7\%$ $R\text{-Sq(adj)} = 37.4\%$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	12710141	(4)	31.50	0.000
Residual Error	50	20173111	(5)		
Total	51	32883252			

(a) Please fill in the blanks in the table.

- (1)
- (2)
- (3)
- (4)
- (5)

- (b) Find and interpret the coefficient of determination, r^2 .
- (c) Calculate the coefficient of correlation, r .

Exercise 2.4 (Recalling student names). (Page 106 Question 3.15, Dataset: NAMEGAME2)

- (a) Find a 99% confidence interval for the mean recall proportion for students in the fifth position during the "name game". Interpret the result.
- (b) Find a 99% prediction interval for the recall proportion of a particular student in the fifth position during the "name game". Interpret the result.
- (c) Compare the two intervals, part (a) and part (b). Which interval is wider? Will this always be the case? Explain.