Assignment 3 - DANA 4810

Alejandro Cervantes Nassar

100388874

November 6th, 2022

Consider question 4.88 form the textbook and the data set: NAVALBASE.

The naval base would like to model the projected percentage increase in fleet effectiveness by the end of the decade as a function of the cost of modifying the fleet.

1. [1] Write a quadratic model.

E(y) = β0 + β1x + β2x2 + β3x2

1. [2] Fit the quadratic model to the data.

Text

Description automatically generated

1. [2] Interpret the value of R2 adjusted on the printout.

With an R2 adjusted of 0.8577, we can say that about 86% of the sampled percentage increase in fleet effectiveness (y) can be explained by the quadratic model.

1. [2] Find the value of *s* and interpret it.

s = 4.893 means that 95% of the percentage increase in fleet effectiveness values for the sampled fleets fall within 2s (9.098) of the model predicted values.

1. [2] Perform a test of overall model adequacy. Use α = 0.05

H0: β1 = β2 = β3 = 0

Ha: at least one β is different than 0.

With the F-statistic of 19.08 on 3 and 6 degrees of freedom, and a p-value of 0.0018 (less than α = 0.05), we reject H0 and conclude that the model is useful for predicting the percentage increase in fleet effectiveness with 95% confidence.

1. [2] Is there sufficient evidence to conclude that the percentage improvement increases more quickly for more costly fleet modifications than for less costly fleet modifications? Use *α* = 0*.*05

Because the coefficient of the quadratic term determines the speed of growth, and a positive coefficient increases the dependent variable as the independent variable increases, what we are trying to do is to determine if β2 is greater than 0.

H0: β2 <= 0

Ha: β2 > 0

The reported p-value in the printout is 0.0989 but this is for a two-tailed test. Since we are doing an upper tail test, the result must be divided by 2: 0.0989/2 = 0.04945. This p-value is barely lower than the alpha value of 0.05, but enough to reject H0. Conclusion: we have enough evidence to say that B2 is positive; the percentage-increase-in-fleet-effectiveness increases more quickly as the cost to do it increases.

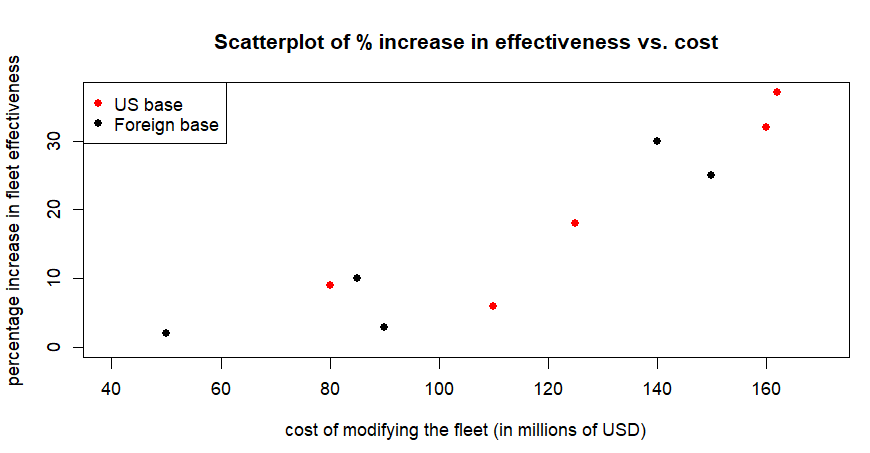
1. [2] Determine whether the calculated *s* in part (d) a reasonable potential error of prediction. What would you recommend?

Coefficient of Variability (C.V.) = 

We would like to have a C.V. of 10% or less. Such a number would indicate that our model can make precise predictions.

With our C.V. of 28.4%, we are not sure if all the coefficients in our model are statistically significant predictors.

1. [2] Create a scatterplot to display the relationship between projected percentage improvement at the end of the decade and the cost of modifying the fleet by base. Your graph should have a legend to indicate the categories of the categorical variable, (i.e. Foreign Base, U.S. Base).



1. [1] Does the same overall pattern depicted in part (h) hold for both types of the base? Explain.

Yes, the same quadratic-form pattern can be observed in the behavior of the response variable (percentage increase in fleet effectiveness) in relationship to the explanatory variable (cost), regardless of the fact if the base were foreign or in the US.

1. [4] Compare two fitted curves for both foreign Base and U.S. Base by plotting the two prediction equations on your plot.

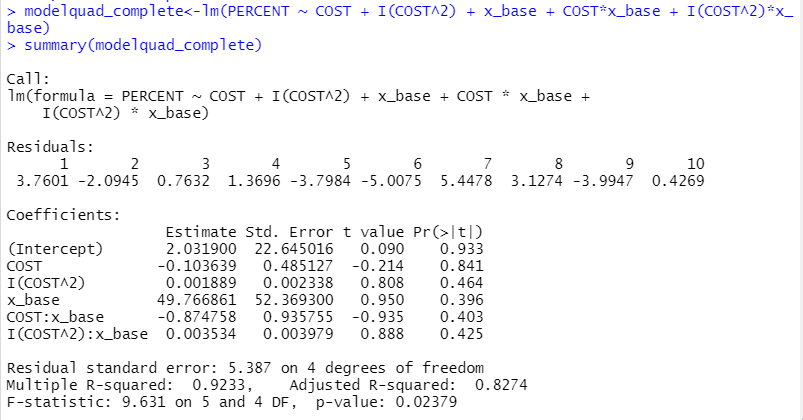
Chart, scatter chart

Description automatically generated

1. [2] Based on your conclusion in part (g), (j), propose a model. (Hint: complete second-order model)

E(y) = β0 + β1x1 + β2x12 + β3x2 + β4x1x2 + β5x12x2

1. [2] Fit the complete model of the proposed model in part (k) to the data.



1. [5] Is there sufficient evidence to indicate that type of base (U.S. or foreign) is a useful predictor of percentage improvement? Using *α* = *.*05 and the critical value approach.

H0: β3 = 0

H0: β3 != 0

test statistic = (Beta / StdError\_Beta) = 0.001889/0.002338 = 0.8079555

Critical value = |t| > tα/2 with [(n=10) – [(k=5) + 1] degrees of freedom =

qt(p=0.05/2, df=(10-(5+1)), lower.tail=FALSE) = 2.776445

Because the test statistic is lower than the critical value, we fail to reject H0; there is not sufficient evidence to indicate that the type of base is a useful predictor for the percentage-of-improvement.

1. [4] What model would you recommend as your best model? Explain your reason(s).

We will do a partial F-test to compare the nested models:

H0: β4 = β5 = 0

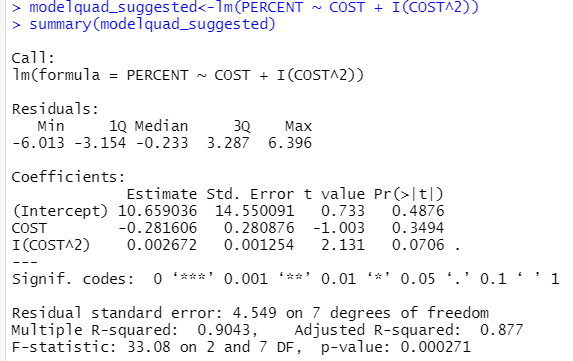
Ha: at least one of β4 or β5 is different than 0.

Text

Description automatically generated

Because the p-value is 0.6532, larger than the p-value of 0.05, we fail to reject H0. Therefore, we have sufficient evidence to say that the reduced model is better to predict the result.

Additionally, since we rejected β3 in part M, let’s see how the model with only β1 and β2 would look like:



Out of the three models, this one is the one with the highest R2adjusted and lowest *s*. I would not like to do another individual t-test and reject β1 to avoid the inflation of the Type II error. In conclusion, I would recommend this one.