TIM-7100

DATA FILE 7

ASSIGNMENT 8

1. Suppose you fit the multiple regression model

*y = β0 + β1x1 + β2x2 + ϵ*

 to *n* = 30 data points and obtain the following result: $\hat{y}=3.4-4.6x\_{1}+2.7x\_{2}+0.93x\_{3}$

 The estimated standard errors of $\hat{β}\_{2}$ and $\hat{β}\_{3}$ are 1.86 and .29, respectively.

1. Test the null hypothesis *H0: β2 = 0* against the alternative hypothesis *Ha: β2 ≠0.*  Use α = .05.
2. Test the null hypothesis *H0: β3 = 0* against the alternative hypothesis *Ha: β3 ≠0.*  Use α = .05.
3. The null hypothesis *H0: β2 = 0* is not rejected. In contrast, the null hypothesis *H0: β3 = 0* is rejected. Explain how this can happen even though $\hat{β}\_{2}$ > $\hat{β}\_{3}.$
4. Use SPSS to fit a second-order model to the following data:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **x** | **0** | **1** | **2** | **3** | **4** | **5** | **6** |
| **y** | **1** | **2.7** | **3.8** | **4.5** | **5.0** | **5.3** | **5.2** |

* 1. Find SSE and s2.
	2. Do the data provide sufficient evidence to indicate that the second-order term provides information for the prediction of *y?* [*Hint:* Test *H0: β2 = 0*].
	3. State the least squares prediction equation.
1. Running a manufacturing operation efficiently requires knowledge of the time it takes employees to manufacture the product, otherwise the cost of making the product cannot be determined. Estimates of production time are frequently obtained using time studies. The data in the table below came from a recent time study of a sample of 15 employees performing a particular task on an automobile assembly line.

|  |  |
| --- | --- |
| **Time to Assemble, y (minutes)** | **Months of Experience, x** |
| **10** | **24** |
| **20** | **1** |
| **15** | **10** |
| **11** | **15** |
| **11** | **17** |
| **19** | **3** |
| **11** | **20** |
| **13** | **9** |
| **17** | **3** |
| **18** | **1** |
| **16** | **7** |
| **16** | **9** |
| **17** | **7** |
| **18** | **5** |
| **10** | **20** |

* 1. Run the multiple linear regression model in SPSS. State the least squares prediction equation.
	2. Test the null hypothesis *H0: β2 = 0* against the alternative *Ha: β2 ≠0*. Use α = .01. Does the quadratic term make an important contribution to the model?
	3. Your conclusion in part b should have been to drop the quadratic term from the model. Do so and fit the “reduced model” y = *β0 + β1x + ϵ* to the data.
	4. Define β1 in the context of this exercise. Find a 90% confidence interval for β1 in the reduced model of part c.
1. Suppose you fit the model *y = β0 + β1x1 + β2x2 + β3x1x2 + β4x12+ β5x22ϵ* to n = 30 data points and get SSE = .46 and R2 = .87.
	1. Do the values of SSE and R2 suggest that the model provides a good fit to the data? Explain.
	2. Is the model of any use in predicting *y*? Test the null hypothesis that *E(y) = β0*; that is, test

H0: *β1*= *β2* = *β3 = β4 = β5 = 0*

against the alternative hypothesis

Ha: At least one of the parameters *β1, β2, … β5* is nonzero.

Use α = .05.

1. A company that services copy machines is interested in developing a regression model that will assist in personnel planning. It needs a model that describes the relationship between the time spent on a preventive maintenance service call to a customer, *y*, and two independent variables: the number of copy machines to be serviced, *x1*, and the service person’s experience in preventive maintenance, *x2.*

|  |  |  |
| --- | --- | --- |
|  Hours of maintenance | Number of copy machines | Months of experience |
| 1.0 | 1 | 12 |
| 3.1 | 3 | 8 |
| 17.0 | 10 | 5 |
| 14.0 | 8 | 2 |
| 6.0 | 5 | 10 |
| 1.8 | 1 | 1 |
| 11.5 | 10 | 10 |
| 9.3 | 5 | 2 |
| 6.0 | 4 | 6 |
| 12.2 | 10 | 18 |

* 1. Fit the model y = *β0 + β1x1 + β2x2 + ϵ* to the data.
	2. Investigate whether the model is useful. Test α = .10.
	3. Find R2 for the fitted model. Interpret your results.
	4. Fit the model y = *β0 + β1x1 + β2x2 + β3x1x2 + ϵ* to the data.
	5. Find R2 for the model in part d.
	6. Explain why you should not rely solely on a comparison of the two R2 values for drawing conclusions about which model is more useful for predicting *y.*
	7. Do the data provide sufficient evidence to indicate that the interaction term, x1x2 contributes information for the prediction of *y?* [Hint: Test *H0: β3 = 0*] Which model is more useful for predicting *y?*
	8. Can you be certain that the model you selected in part g is the best model to use in predicting maintenance time? Explain.