Name: Shiva Kumar Peddapuram

UnivID: 811235874  
FINAL EXAM – Math 40015/50015  
Fall 2022

SHOW ALL YOUR WORK and write complete and coherent answers. No partial credit will be given if no work is shown. Please write as clearly and neatly as possible. If I cannot read your answers, I cannot give you any credit. Feel free to ask for more paper if you need more space. GOOD LUCK !!!

**In the data set “landrent” in package alr4, the variables are average rent per acre Y planted to alfalfa,average rent paid X1 for all tillable land, density of dairy cows X2 (number per square mile), and proportion X3 of farmland used as pasture. You need to answer the following questions based on your own code (not the lm function).**

**1. For the data, the full model is E(Y ) = β0 + β1X1 + β2X2 + β3X3 + β4X2 ∗X3.**

**And the reduced model is E(Y ) = β0 + β1X1 + β2X2 + β3X3.  
Use the F-test to test which model is more appropriate for the data. Compute the F-statistic in detail. Report the p-value and summarize your conclusion.**

Solution:

Text

Description automatically generated

**Execution part:**

A picture containing text

Description automatically generated

Text, letter

Description automatically generated

***Explanation:***

***From the above we can conclude that the p value of f is 0.2076398 and F.stat value is 1.561768***

***Pvalue1= 0.04784446***

***Pvalue2= 0***

***Pvalue3= 1.097297e-05***

***Pvalue4= 0.1764994***

***Pvalue5= 0.1764994***

***Sigma hat value is 8.986852***

2.) **Suppose that the full model is chosen. Now you are asked to estimate all the parameters of the full model including the variance σ2**

**Solution:**

**Coding part:**

Graphical user interface, text, application

Description automatically generated

***Explanation part:***

All the parameter values are as below

Execution part for the above code

Graphical user interface, text, application

Description automatically generated

**So from the above code we gave initiated sigma2.hat values and beta hat value**

**So from the above the sigma2 havt value is 86.69043**

**Beta.hat values are**

-2.8282148

0.8832666

0.4317553

-11.3804544

-1.0117308

**3.) Again for the full model, you need to construct a 99% confidence interval for each of the slopes  
β1,β2,β3 and β4. Is 0 included by each confidence interval?**

**Coding part:**

**Graphical user interface, text

Description automatically generated with medium confidenceText

Description automatically generated**

**Explanation:**

**Text

Description automatically generatedGraphical user interface, text, application

Description automatically generated**

**From the above code the**

**Slope beta1 lower value is 0.7001679 and the upper value is 1.066365**

**Slope beta2 lower value is 0.145166and the upper value is** **0.7183446**

**Slope beta3 lower value is**  **-42.93181and the upper value is 20.1709**

**Slope beta4 lower value is -8.547967and the upper value is 6.524506**

4. Without actually conducting hypothesis tests, is it possible to tell whether the null hypothesis of H0 : βi = 0 vs H1 : βi 6= 0 for i = 1,2,3,4 is rejected or failed to be rejected based on the results from question 3 above? If yes, what should be the chosen significance level for each hypothesis test?

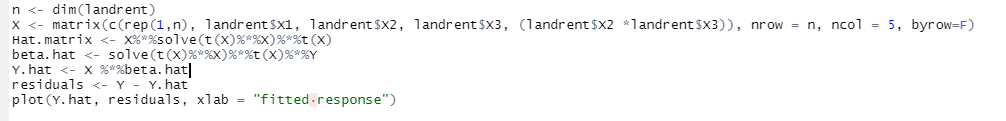
**Solution:**

Yes, it is fair to conclude that by using above confidence interval we can reject the null hypothesis with 99 percent confidence.

The null hypothesis (H0) states that there is no difference between the groups being tested. The alternative hypothesis (H1) postulates that there is a difference between the groups being tested.

5. For the full model, obtain the residuals and make the residuals vs fitted response plot. Based on the  
residual plot, can we say the linearity and constant variance assumptions hold for the model?

Code:



Plot:

Chart, scatter chart

Description automatically generated

6. For the full model, compute the Cook’s distances. Are there any influential outliers?

Text

Description automatically generated

Cookes values are:

[1,] 0.0096767714

[2,] 0.0018195043

[3,] 0.0011164589

[4,] 0.0018918651

[5,] 0.0148028993

[6,] 0.0000000000

[7,] 0.0040994600

[8,] 0.0017958516

[9,] 0.1524981177

[10,] 0.0031335811

[11,] 0.0003579510

[12,] 0.0000000000

[13,] 0.0005165651

[14,] 0.0004618965

[15,] 0.0008748523

[16,] 0.0024739824

[17,] 0.0238670892

[18,] 0.0004078618

[19,] 0.0044626645

[20,] 0.0063360857

[21,] 2.3032168019

[22,] 0.0041461266

[23,] 0.0033917323

[24,] 0.0016862318

[25,] 0.0004451163

[26,] 0.0022295390

[27,] 0.0073755391

[28,] 0.0108104446

[29,] 0.0072242908

[30,] 0.0026610166

[31,] 0.0032014278

[32,] 0.0256406766

[33,] 0.1781418197

[34,] 0.0000000000

[35,] 0.0035344830

[36,] 0.0000000000

[37,] 0.0020024529

[38,] 0.0017217584

[39,] 0.0019575243

[40,] 0.0020871259

[41,] 0.0000000000

[42,] 0.8290480686

[43,] 0.0008620414

[44,] 0.0000000000

[45,] 0.0012575940

[46,] 0.0005721686

[47,] 0.0004889101

[48,] 0.0024873404

[49,] 0.0011713613

[50,] 0.0013437320

[51,] 0.0000000000

[52,] 0.0047446153

[53,] 0.0012186882

[54,] 0.0251603604

[55,] 0.0001844792

[56,] 0.0324588260

[57,] 0.0044145153

[58,] 0.0017752066

[59,] 0.0015641335

[60,] 0.0092061620

[61,] 0.0005836189

[62,] 0.0028163517

[63,] 0.0257507351

[64,] 0.0052395773

[65,] 0.0000000000

[66,] 0.0000000000

[67,] 0.0019049670

Influential values are:

[1]  9 21 33 42

Chart, scatter chart

Description automatically generated

**The data set “Challeng” records performance of O-rings for the 23 U.S. space shuttle missions prior to the Challenger disaster of January 20, 1986. For each of the previous missions, the temperature at takeoff and the pressure of a prelaunch test were recorded, along with the number of O-rings that failed out of 6. You need to answer the following questions based on the glm function of R.**

1. **Consider “temp” and “pres” as two predictors, “fail” as the number of “successes”, and “n” as the total number of trials. Fit the binomial regression model y ∼ temp + pres + temp : pres.**

**Text

Description automatically generated with low confidence**

**Execution part:**

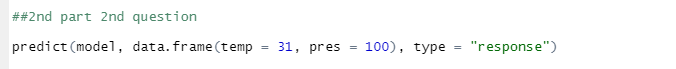
**Text

Description automatically generatedGraphical user interface

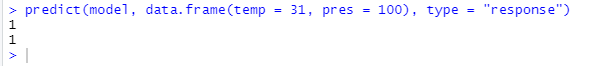
Description automatically generated with low confidence**

1. **Use your fitted model to estimate the probability of failure of an O-ring when the temperature was 31, and the pressure is 100**

**Code part:**

****

**Execution part:**

****

**Probability is 1**

1. **Consider another reduced binomial regression model y ∼ temp + pres. Test which model (full model vs reduced model) is more appropriate? To answer it, you need to compute the test statistic ∆G2 in detail, report the p-value and summarize your conclusion**

**Solution:**

****

**Execution part:**

**Text

Description automatically generated**