**MODERN COLLEGE OF ARTS,SCI. & COMM. PUNE-05.**

**DEPARTMENT OF STATISTICS. M.Sc.( I ) Sem II**

**ST- 25 EXPT.NO. 02B**

**Title : Multiple Linear Regression ( Estimation, Testing, prediction, confidence interval, lack of fit)**

Q.1 Consider the National Football League data as given as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Team | y | x1 | x2 | x3 |
| Washington | 10 | 1985 | 59.7 | 2205 |
| Minnesota | 11 | 2855 | 55 | 2096 |
| New England | 11 | 1737 | 65.6 | 1847 |
| Oakland | 13 | 2905 | 61.4 | 1903 |
| Pittsburgh | 10 | 1666 | 66.1 | 1457 |
| Baltimore | 11 | 2927 | 61 | 1848 |
| Los Angeles | 10 | 2341 | 66.1 | 1564 |
| Dallas | 11 | 2737 | 58 | 1821 |
| Atlanta | 4 | 1414 | 57 | 2577 |
| Buffalo | 2 | 1838 | 58.9 | 2476 |
| Chicago | 7 | 1480 | 67.5 | 1984 |
| Cincinnati | 10 | 2191 | 57.2 | 1917 |
| Cleveland | 9 | 2229 | 58.8 | 1761 |
| Denver | 9 | 2204 | 58.6 | 1709 |
| Detroit | 6 | 2140 | 59.2 | 1901 |
| Green Bay | 5 | 1730 | 54.4 | 2288 |
| Houston | 5 | 2072 | 49.6 | 2072 |
| Kansas City | 5 | 2929 | 54.3 | 2861 |
| Miami | 6 | 2268 | 58.7 | 2411 |
| New Orleans | 4 | 1983 | 51.7 | 2289 |
| New York Giants | 3 | 1792 | 61.9 | 2203 |
| New York Jets | 3 | 1606 | 52.7 | 2592 |
| Philadelphia | 4 | 1492 | 57.8 | 2053 |
| St. Louis | 10 | 2835 | 59.7 | 1979 |
| San Diego | 6 | 2416 | 54.9 | 2048 |
| San Francisco | 8 | 1638 | 65.3 | 1786 |
| Seattle | 2 | 2649 | 43.8 | 2876 |
| Tampa Bay | 0 | 1503 | 53.5 | 2560 |

where

y: Games won (per 14- game season)

x1: Passing yards (season)

x2: Percent rushing (rushing plays/total plays)

x3: Opponents’ rushing yards (season)

1. Fit a multiple regression model relating the number of games won to the team’s passing yardage (x1),the percentages of rushing plays (x2), and the opponents’ yard rushing (x3)
2. Construct the analysis of variance table and test for significance of regression.
3. Calculate t statistics for testing the hypothesis H0: β1=0, H0: β2=0, H0: β3=0. What conclusions can you draw about the role of the variables x1, x2, and x3 play in the model?
4. Calculate R2 and for this model.
5. Show numerically that the square of the simple correlation coefficient between the observed values yi and the fitted values equals R2.
6. Find a 95 % confidence interval on β2.
7. Find a 95 % confidence interval on the mean number of games won by a team when x1 =2300 x2=56.0 and x3=2100.

Q.2 Fit a model to above data using only x2 and x3 as the regressors. a. Test for significance of regression.

b. Calculate R2 and . How do these quantities compare to the values computed for the model in Q.1, which included an additional regressor x1.

c. Find a 95 % confidence interval on β2. Also Find a 95 % confidence interval on the

mean number of games won by a team when x2=56.0 and x3=2100. Compare the

lengths of these confidence intervals to the length of the corresponding confidence

interval from Q.1.

d. What conclusions can you draw from this problem about the consequences of

omitting an important regressor from a model?

Q.3 The concentration of NbOCl3 in a tube-flow reactor as a function of several controllable

variables is shown as below:

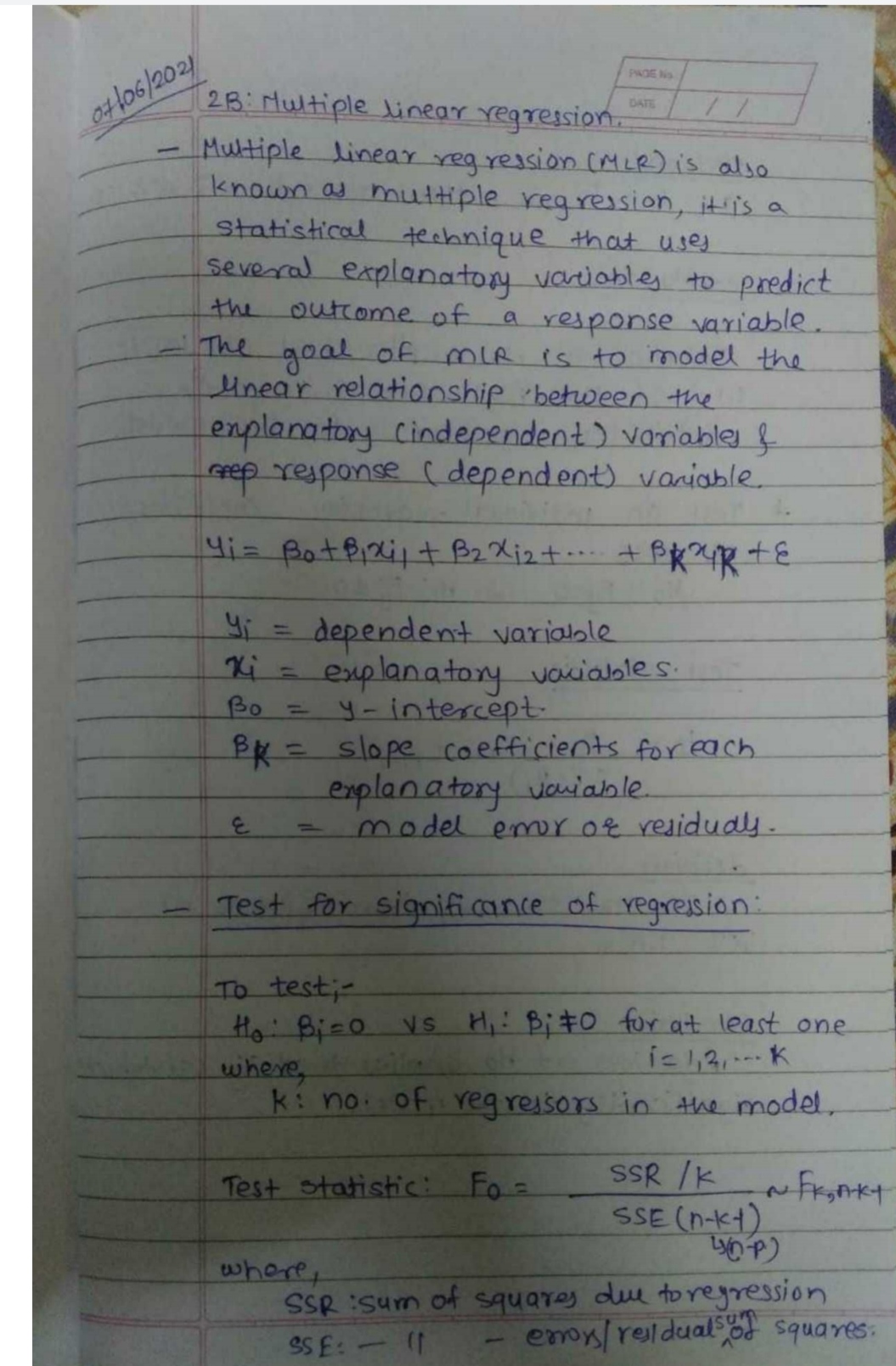
|  |  |  |
| --- | --- | --- |
| y | x1 | x2 |
| 0.000450 | 0.0105 | 0.0177 |
| 0.000450 | 0.011 | 0.0172 |
| 0.000473 | 0.0106 | 0.0157 |
| 0.000507 | 0.0116 | 0.0082 |
| 0.000457 | 0.0121 | 0.007 |
| 0.000452 | 0.0123 | 0.0065 |
| 0.000453 | 0.0122 | 0.0071 |
| 0.000426 | 0.0122 | 0.0062 |
| 0.001215 | 0.0123 | 0.0153 |
| 0.001256 | 0.0122 | 0.0129 |
| 0.001145 | 0.0094 | 0.0354 |
| 0.001085 | 0.01 | 0.0342 |
| 0.001066 | 0.0101 | 0.0323 |
| 0.001111 | 0.0099 | 0.0337 |
| 0.001364 | 0.011 | 0.0161 |
| 0.001254 | 0.0117 | 0.0149 |
| 0.001396 | 0.011 | 0.0163 |
| 0.001575 | 0.0104 | 0.0164 |
| 0.001615 | 0.0067 | 0.0379 |
| 0.001733 | 0.0066 | 0.0360 |
| 0.002753 | 0.0044 | 0.0327 |
| 0.003186 | 0.0073 | 0.0263 |
| 0.003227 | 0.0078 | 0.02 |
| 0.003469 | 0.0067 | 0.0197 |
| 0.001911 | 0.0091 | 0.0331 |
| 0.002588 | 0.0079 | 0.0674 |
| 0.002635 | 0.0068 | 0.077 |
| 0.002725 | 0.0065 | 0.078 |

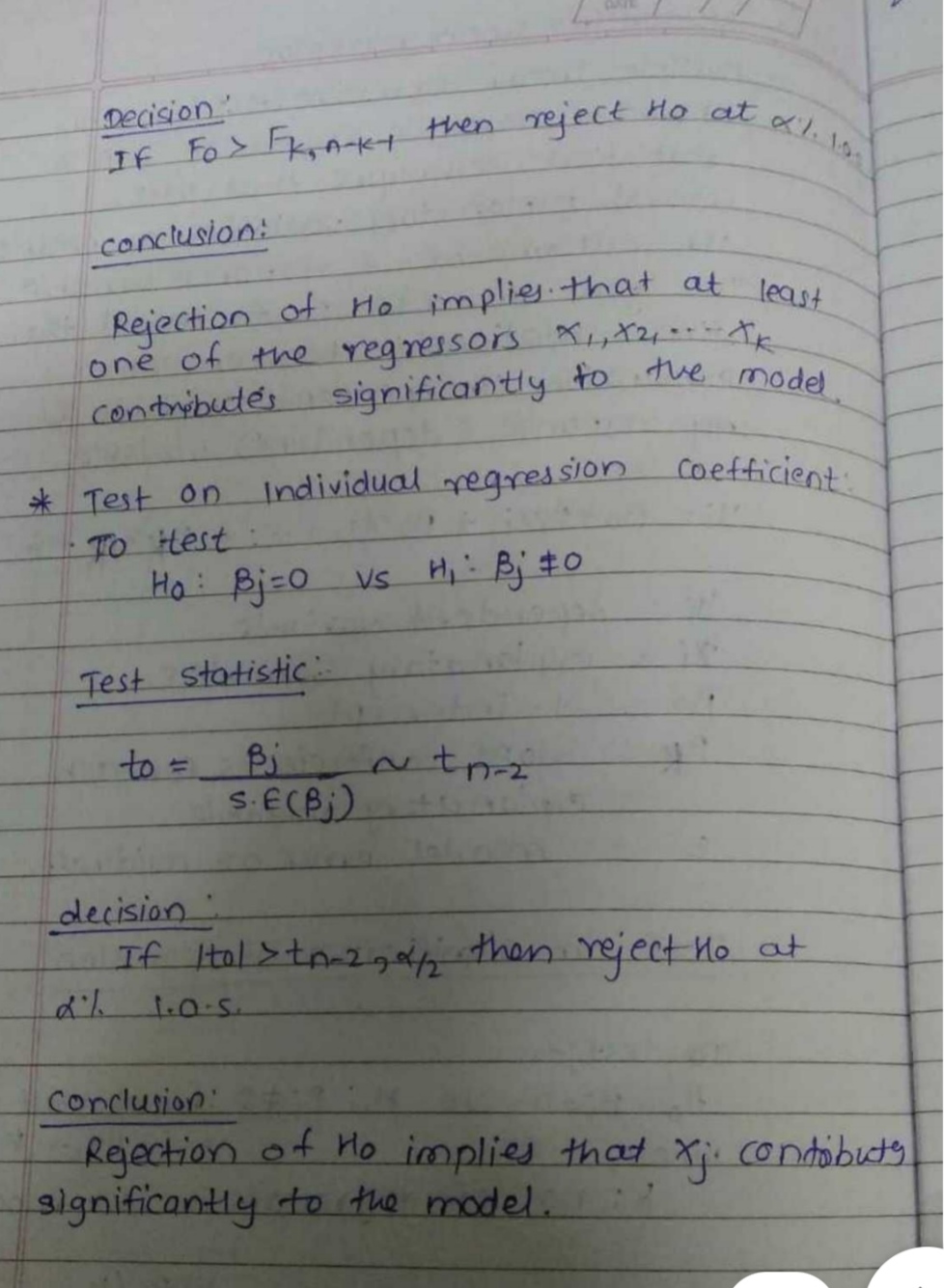
where y: NbOCl3 (g-mol/l)

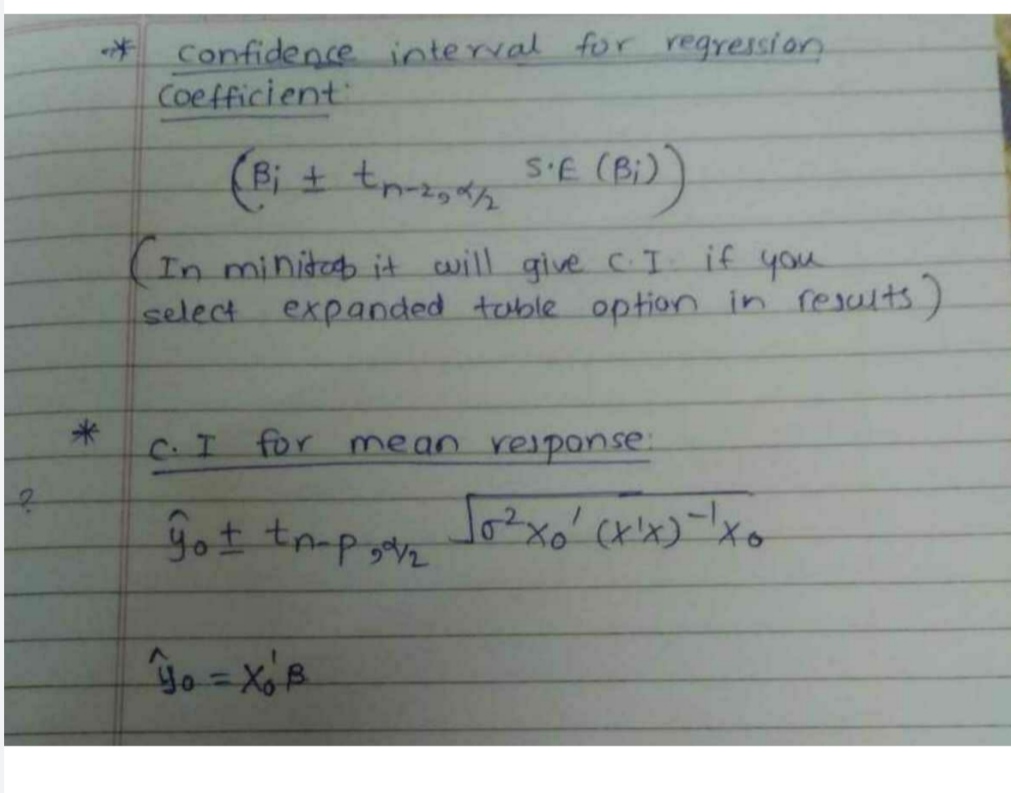
x1: COCl2 concentration (g-mol/l)

x2 : Mole fraction CO2.

1. Fit a multiple linear regression model relating concentration of NbOCl3 (y) to concentration of COCl2(x1) and mole fraction (x2).
2. Test for significance of regression.
3. Calculate R2 and for this model.
4. Using t test, determine the contribution of x1 and x2 to the model. Are both regressors x1 and x2 necessary?
5. Is multicollinearity a potential concern in this model?







**Process in Minitab**

**Q1**

**Copy the data of all response and regressor variables in Minitab.**

**Initially select Stat menu from the menu bar.**

**Under that menu bar select regression.**

**Again select regression and then finally select fit regression model.**

**Select the column of Y under response variable.**

**Select the columns of all regressors under continuous predictors.**

**Then go to storage. A dialog box will appear.**

**Check mark fits,residuals and coefficients as the solution for this question requires all these values**

**Then click OK.**

**You will get the equation of fitted line as your output.**

**ANOVA table,model summary, coefficient table and columns of fitted value and residuals will also be displayed as your output.**

**The table value for F test can be found by selecting the F dist option under probability distribution of the Stat menu. Select inverse cumulative probability and then enter the degrees of freedom,enter constant as 0.975 and click OK.**

**The table value for t test can be found by selecting the t dist option under probability distribution of the Stat menu. Select inverse cumulative probability and then enter the degrees of freedom,enter constant as 0.975 and click OK.**

**For F test,the mean sum of squares of residuals and regressors can be obtained from the ANOVA table.**

**For t test the estimated value of coefficients and the standard error can be obtained from the coefficients table.The calculated values can be cross checked with the t-cal values of the coefficients table.We can easily calculate confidence intervals also using this table.These values can be verified from the CI values of expanded table of regression.**

**R-square and R-square adjusted values can be obtained from model summary.**

**To calculate correlation select the Stat menu and then select basic statistic.Under the basic statistic option select correlation.Select the column of fitted values and observed values.Click OK.**

**Under Calc menu select calculator and take square of the obtained correlation coefficient.Thus it is proved that this value matches with the R2 value from model summary.**

**Finally to solve question (g) of the first question we need to create a matrix Xo of the given values x1, x2 and x3.Enter x1,x2 and x3 in a column. To create a matrix under Data menu select Change data type and then select the option Columns to Matrix. A dialog box will appear on the screen.Select the column of x1,x2 and x3 values. Enter a matrix name.Click OK.**

**This will create the matrix Xo.**

**Similarly create X matrix by selecting the columns of x1 ,x2 and x3 regressors containing 28 observations.Further to view the matrix go to display data under the data menu and select the matrices you want to be displayed.Repeat same process to form matrix of coefficients. Click OK.**

**Under Calc menu go to matrices**

**and select the transpose option.Then select the matrix Xo and store it in some other variable. Display that with the help display data option.Do the same for X matrix.**

**Now under Arithmetic option of Matrix in Calc menu multiply the transpose of X matrix to the matrix of coefficients from which betao has been removed. Finally add betao value to the obtained result. Thus we get value of estimated value of yhat.To find C.I for yhat we need (X'X)-1. You will find the inverse option in the matrix option itself.Thus all further calculations can be done under Calc menu only.**

**Q2**

**Copy the data of all response and regressor variables in Minitab.**

**Initially select Stat menu from the menu bar.**

**Under that menu bar select regression.**

**Again select regression and then finally select fit regression model.**

**Select the column of Y under response variable.**

**Select the columns of all regressors under continuous predictors.**

**Then click OK.**

**You will get the equation of fitted line as your output.**

**ANOVA table,model summary, coefficient table will also be displayed as your output.**

**R-square and R-square adjusted values can be obtained from model summary.**

**The table value for t test can be found by selecting the t dist option under probability distribution of the Stat menu. Select inverse cumulative probability and then enter the degrees of freedom,enter constant as 0.975 and click OK.**

**The estimated value of coefficients and the standard error can be obtained from the coefficients table.The calculated values can be cross checked with the t-cal values of the coefficients table.We can easily calculate confidence intervals also using this table.These values can be verified from the CI values of expanded table of regression.**

**To calculate the C.I for mean number of games we need to create a matrix Xo of the given values x2 and x3.Enter x2 and x3 in a column. To create a matrix under Data menu select Change data type and then select the option Columns to Matrix. A dialog box will appear on the screen.Select the column of x2 and x3 values. Enter a matrix name.Click OK.**

**This will create the matrix Xo.**

**Similarly create X matrix by selecting the columns of x2 and x3 regressors containing 28 observations.Further to view the matrix go to display data under the data menu and select the matrices you want to be displayed.Repeat same process to form matrix of coefficients. Click OK.**

**Under Calc menu go to matrices**

**and select the transpose option.Then select the matrix Xo and store it in some other variable. Display that with the help display data option.Do the same for X matrix.**

**Now under Arithmetic option of Matrix in Calc menu multiply the transpose of X matrix to the matrix of coefficients from which betao has been removed. Finally add betao value to the obtained result. Thus we get value of estimated value of yhat.To find C.I for yhat we need (X'X)-1. You will find the inverse option in the matrix option itself.Thus all further calculations can be done under Calc menu only.**

**Q3**

**Copy the data of all response and regressor variables in Minitab.**

**Initially select Stat menu from the menu bar.**

**Under that menu bar select regression.**

**Again select regression and then finally select fit regression model.**

**Select the column of Y under response variable.**

**Select the columns of all regressors under continuous predictors.**

**Then click OK.**

**You will get the equation of fitted line as your output.**

**ANOVA table,model summary, coefficient table will also be displayed as your output.**

**The table value for F test can be found by selecting the F dist option under probability distribution of the Stat menu. Select inverse cumulative probability and then enter the degrees of freedom,enter constant as 0.975 and click OK.**

**The table value for t test can be found by selecting the t dist option under probability distribution of the Stat menu. Select inverse cumulative probability and then enter the degrees of freedom,enter constant as 0.975 and click OK.**

**For F test,the mean sum of squares of residuals and regressors can be obtained from the ANOVA table.**

**For t test the estimated value of coefficients and the standard error can be obtained from the coefficients table.The calculated values can be cross checked with the t-cal values of the coefficients table.**

**R-square and R-square adjusted values can be obtained from model summary.**

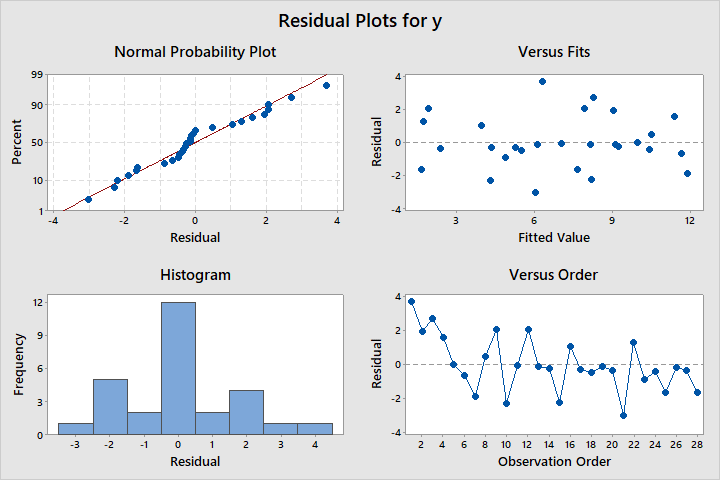
**For multicollinearity we check the VIF values from the coefficient table.If the VIF is equal to 1 there is no multicollinearity among factors, but if the VIF is greater than 1, the predictors may be moderately correlated.**

**Q1.**

**a)**

**Regression Equation:**

**y = -1.81 + 0.003598 x1 + 0.1940 x2 - 0.00482*x3***



1. *Anova Table:*

*Analysis of Variance*

*Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value*

*Regression 3 257.09 78.63% 257.09 85.698 29.44 0.000*

*x1 1 76.19 23.30% 78.03 78.028 26.80 0.000*

*x2 1 139.50 42.67% 14.07 14.068 4.83 0.038*

*x3 1 41.40 12.66% 41.40 41.400 14.22 0.001*

*Error 24 69.87 21.37% 69.87 2.911*

*Total 27 326.96 100.00%*

*ii) Test Of Significance Of Regression:*

*Here We Test The Hypothesis*

*Ho:βj= 0 Vs H1:βj≠ 0 for at least one j = 1,2,3*

*for α = 5% l.o.s we have test statistic as*

*Fo = ~ FK,n-K-1*

Thus, Fcal =29.44 ~ F3,24

*While Ftab = F3,24 = 3.00879*

*Therefore, As Fcal > Ftab Reject Ho at 5% l.o.s*

**Conclusion: At least one of the regressor contribute significantly to the Regression Model.**

c.

i) Here we want to test the hypothesis,

H0: β1=0 VS H1: β1 ≠ 0

for α=5% l.o.s we have test statistic as,

tcal = ~ t (n-2)

thus, tcal= 5.18 ~ t(26)

while, tα/2(n-2)= t0.025,(26)=2.05553

Decision Rule: If tcal > tα/2(n-2) then reject H0.

Here 5.18 > 2.05553 thus we reject H0.

i.e. β1  do not differ significantly to the regression model

**Conclusion:Thus regressor X1 i.e. Passing Yards contributes significantly to the given model of Y i.e Winning Games**

ii) Here we want to test the hypothesis,

H0: β2=0 VS H1: β2 ≠ 0

for α=5% l.o.s we have test statistic as,

tcal = ~ t (n-2)

thus, tcal= 2.20 ~ t(26)

while, tα/2(n-2)= t0.025,(26)=2.05553

Decision Rule: If |tcal |> tα/2(n-2) then reject H0.

Here 2.20 > 2.05553 thus we reject H0.

i.e. β2  do not differ significantly to the regression model .

**Conclusion:Thus regressor X2 i.e. Percent Rushing contributes significantly to the given model of Y i.e Winning games**

iii) Here we want to test the hypothesis,

H0: β3=0 VS H1: β3≠ 0

for α=5% l.o.s we have test statistic as,

tcal = ~ t (n-2)

thus, tcal= 3.77 ~ t(26)

while, tα/2(n-2)= t0.025,(26)=2.05553

Decision Rule: If tcal > tα/2(n-2) then reject H0.

Here 3.77 > 2.05553 thus we reject H0.

i.e. β3  do not differ significantly to the regression model .

**Conclusion:Thus regressor X3 i.e Opponents Rushing Yard contributes significantly to the given model of Y i.e. Winning games.**

d.

i) Coefficient Of Determination = 1 – (SSRes/SST) =78.63%

i.e. 78.63% variability in Y explained by regression model.

ii) Coefficient Of Determination Adj. = 1 – (SSRes/SST)\*(27/25) =75.96%

i.e. 75.96% variability in Y explained by regression model.

e.

the simple correlation coefficient between the observed values yi and the fitted values = ρ = = 0.887

square of the simple correlation coefficient between the observed values yi and the fitted values = = 0.7867

Conclusion: = 0.7867 is approxi. equal to Coefficient Of Determination = 0.7863.

F.

95%Confidence Interval on β2  :

C.I. for β2=( β2-tα/2\*SE(β2) , β2+tα/2\*SE(β2))

95% Confidence interval for β2 is = ( 0.0119, 0.3761).

G.

X =

1.0e+03 \*

0.0010 1.9850 0.0597 2.2050

0.0010 2.8550 0.0550 2.0960

0.0010 1.7370 0.0656 1.8470

0.0010 2.9050 0.0614 1.9030

0.0010 1.6660 0.0661 1.4570

0.0010 2.9270 0.0610 1.8480

0.0010 2.3410 0.0661 1.5640

0.0010 2.7370 0.0580 1.8210

0.0010 1.4140 0.0570 2.5770

0.0010 1.8380 0.0589 2.4760

0.0010 1.4800 0.0675 1.9840

0.0010 2.1910 0.0572 1.9170

0.0010 2.2290 0.0588 1.7610

0.0010 2.2040 0.0586 1.7090

0.0010 2.1400 0.0592 1.9010

0.0010 1.7300 0.0544 2.2880

0.0010 2.0720 0.0496 2.0720

0.0010 2.9290 0.0543 2.8610

0.0010 2.2680 0.0587 2.4110

0.0010 1.9830 0.0517 2.2890

0.0010 1.7920 0.0619 2.2030

0.0010 1.6060 0.0527 2.5920

0.0010 1.4920 0.0578 2.0530

0.0010 2.8350 0.0597 1.9790

0.0010 2.4160 0.0549 2.0480

0.0010 1.6380 0.0653 1.7860

0.0010 2.6490 0.0438 2.8760

0.0010 1.5030 0.0535 2.5600

>> X'\*X

ans =

1.0e+08 \*

0.0000 0.0006 0.0000 0.0006

0.0006 1.3343 0.0345 1.2543

0.0000 0.0345 0.0010 0.0340

0.0006 1.2543 0.0340 1.2828

>> inv(X'\*X)

ans =

21.4422 -0.0009 -0.2278 -0.0029

-0.0009 0.0000 0.0000 0.0000

-0.2278 0.0000 0.0027 0.0000

-0.0029 0.0000 0.0000 0.0000

>> X0 = [1;2300 ;56 ;2100]

X0 =

1

2300

56

2100

>> B = [-1.81000;0.00360;0.19400;-0.00482]

B =

-1.8100

0.0036

0.1940

-0.0048

>> Y0 = X0'\*B

YO =

7.2120

>> X0'\*inv(X'\*X)\*X0

ans =

0.0491

>> 2.911\*X0'\*inv(X'\*X)\*X0

ans =

0.1429

>> sqrt(2.911\*X0'\*inv(X'\*X)\*X0)

ans =

0.3780

Confidence interval for mean response is:

( Y0 – t(n-2),α/2\* , Y0 -+t(n-2),α/2\* )

(7.2120 – 2.0595\*0.3780 , 7.2120 – 2.0595\*0.3780)

Conclusion:95% Confidence interval is ( 6.42712, 7.98768)

Q2.

***Ans:***

**Regression Equation :**

***y = 17.94 + 0.048 X2 - 0.00654 X3***

**a)**

1. **Anova Table:**

**Analysis of Variance**

**Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value**

**Regression 2 179.07 54.77% 179.066 89.5331 15.13 0.000**

**X2 1 97.24 29.74% 0.974 0.9739 0.16 0.688**

**X3 1 81.83 25.03% 81.828 81.8280 13.83 0.001**

**Error 25 147.90 45.23% 147.898 5.9159**

**Total 27 326.96 100.00%**

**ii) Test Of Significance Of Regression:**

Here We Test The Hypothesis

Ho:βj= 0 Vs H1:βj≠ 0 for at least one j = 2,3

for α = 5% l.o.s we have test statistic as

Fo = ~ FK,n-K-1

Thus, Fcal =15.13 ~ F2,25

While Ftab = F2,25 = 3.38519

Therefore, As Fcal > Ftab Reject Ho at 5% l.o.s

**Conclusion: At least one of the regressor contribute significantly to the Regression Model**

**b.**

i) Coefficient Of Determination = 1 – (SSRes/SST) = 54.77%

**Conclusion:54.77% variability in Y explained by regression model.**

ii) Coefficient Of Determination Adj. = 1 – (SSRes/SST)\*(27/26) = 51.15%

**Conclusion: 51.15% variability in Y explained by regression model.**

**Higher value of indicate better fitting of model.**

**As Compare to Q1 both and Adj. of Q2 are less .**

**Conclusion:**

1. **model Y with three regressor in Q1 is better fitted than model Y with two regressor in Q2**
2. **X1 contributes more significantly to the model.**

**c.**

**i) 95% Confidence Interval on β2  :**

C.I. for β2=( β2-t (n-2),α/2\*SE(β2) , β2+t (n-2),α/2\*SE(β2))

**Conclusion: 95% Confidence interval for β2 is = (-0.197,0.294)**

Length of CI for is =0.491

For Q1) we have confidence interval as;

95% Confidence interval for β2 is = ( 0.0119, 0.3761).

**Length of CI for Q1) = 0.3642**

1. **95% Confidence Interval for mean response**

***X =***

1.0e+03 \*

0.0010 0.0597 2.2050

0.0010 0.0550 2.0960

0.0010 0.0656 1.8470

0.0010 0.0614 1.9030

0.0010 0.0661 1.4570

0.0010 0.0610 1.8480

0.0010 0.0661 1.5640

0.0010 0.0580 1.8210

0.0010 0.0570 2.5770

0.0010 0.0589 2.4760

0.0010 0.0675 1.9840

0.0010 0.0572 1.9170

0.0010 0.0588 1.7610

0.0010 0.0586 1.7090

0.0010 0.0592 1.9010

0.0010 0.0544 2.2880

0.0010 0.0496 2.0720

0.0010 0.0543 2.8610

0.0010 0.0587 2.4110

0.0010 0.0517 2.2890

0.0010 0.0619 2.2030

0.0010 0.0527 2.5920

0.0010 0.0578 2.0530

0.0010 0.0597 1.9790

0.0010 0.0549 2.0480

0.0010 0.0653 1.7860

0.0010 0.0438 2.8760

0.0010 0.0535 2.5600

>> X'\*X

ans =

1.0e+08 \*

0.0000 0.0000 0.0006

0.0000 0.0010 0.0340

0.0006 0.0340 1.2828

>> inv(X'\*X)

ans =

16.4418 -0.1909 -0.0025

-0.1909 0.0024 0.0000

-0.0025 0.0000 0.0000

>> X0 = [ 1;56;2100]

X0 =

1

56

2100

>> B=[17.94;0.048;-0.00654]

B =

17.9400

0.0480

-0.0065

>> Y0 = X0'\*B

Y0 =

6.8940

>> 5.9159\*X0'\*inv(X'\*X)\*X0

ans =

0.2840

>> sqrt(5.9159\*X0'\*inv(X'\*X)\*X0)

ans =

0.5329

***Confidence interval for mean response is:***

*( Y0 – t(n-2),α/2\* , Y0 +t(n-2),α/2\* )*

*(6.8940 -2.0595\*0.5329, 6.8940 + 2.0595\*0.5329)*

95% Confidence interval is ( 5.7964, 7.9915)

*Length of CI for is =2.120*

For Q4) we have confidence interval as;

95% Confidence interval is ( 6.42712, 7.98768)

Length of CI for Q4) = 1.56056

Conclusion : Length of CI of β2 and Mean Response for Q1) is less than for Q2) thus CI for Q1) is better than that for Q2) i.e model in Q1) is better than in Q2)

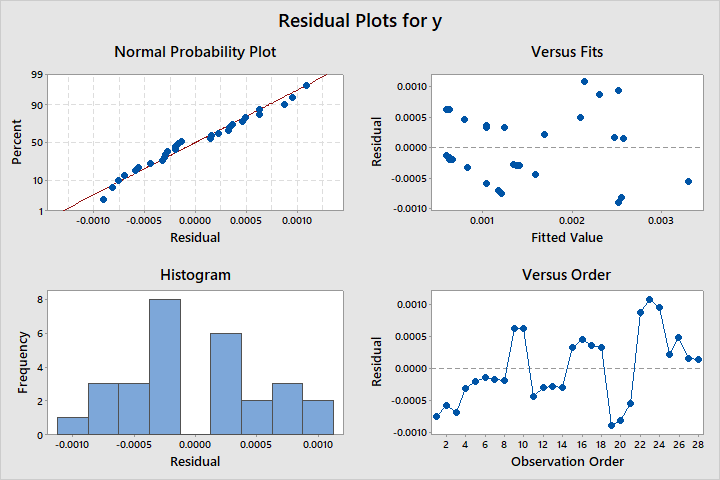
**d.**

**Conclusion: If we omit important regressor from the model then the model is affected significantly as we see there is significant decrease in value of coefficient of determination .**

**Q3.**

**Regression Equation :**

**y = 0.004833 - 0.3450 x1 - 0.00014 x2**



**b.**

***Analysis of Variance***

***Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value***

***Regression 2 0.000017 66.36% 0.000017 0.000008 24.66 0.000***

***x1 1 0.000017 66.36% 0.000009 0.000009 26.20 0.000***

***x2 1 0.000000 0.00% 0.000000 0.000000 0.00 0.986***

***Error 25 0.000008 33.64% 0.000008 0.000000***

***Total 27 0.000025 100.00%***

**Regression Equation:**

**y = 0.004833 - 0.3450 x1 - 0.00014 x2**

**C.**

**Test Of Significance Of Regression:**

Here We Test The Hypothesis

Ho:βj= 0 Vs H1:βj≠ 0 for at least one j = 1,2

for α = 5% l.o.s we have test statistic as

Fo = ~ FK,n-K-1

Thus, Fcal =24.66 ~ F2,25

While Ftab = F2,25 = 3.38519

Therefore, As Fcal > Ftab Reject Ho at 5% l.o.s

**Conclusion:At least one of the regressor contribute significantly to the Regression Model.**

***2.***

**i) *Coefficient Of Determination = 1 – (SSRes/SST) =66.36%***

i***.***e. 66.36% variability in Y explained by regression model.

**ii) Coefficient Of Determination Adj. = 1 – (SSRes/SST)\*(27/26) =63.67%**

**i**.e. 63.67% variability in Y explained by regression model.

**d.**

i) Here we want to test the hypothesis,

H0: β1=0 VS H1: β1 ≠ 0

for α=5% l.o.s we have test statistic as,

tcal = ~ t (n-2)

thus,|tcal |= 5.12 ~ t(26)

while, tα/2(n-2)= t0.025,(26)=2.05553

Decision Rule: If tcal > tα/2(n-2) then reject H0.

Here 5.12 > 2.05553 thus we reject H0.

**i.e. β1  differ significantly to the regression model**

**Coclusion: Thus regressor X2 i.e. Concentration of COCl2do not contributessignificantly to the given model of concentration of NbOCl3**

ii) Here we want to test the hypothesis,

H0: β2=0 VS H1: β2 ≠ 0

for α=5% l.o.s we have test statistic as,

tcal = ~ t (n-2)

thus, |tcal |= 0.02 ~ t(26)

while, tα/2(n-2)= t0.025,(26)=2.05553

**Decision Rule: If tcal < t(n-2),α/2 then reject H0.**

**Here 0.02 < 2.05553 thus we accept H0.**

**i.e. β2  do not differ significantly to the regression model .**

**Coclusion*:* Thus regressor X2 i.e. mole fraction do not contributes significantly to the given model of concentration of NbOCl3**

**Conclusion: Only Regressor X1 i.e. concentration of COCl2 is necessary to the model of Concentration of NbClO3**

**e.**

Pearson correlation of x1 and x2 = -0.687

Here, Multicollinearity is less than 0.7

**Conclsion: So, multicollinearity is not potential concern to the model**