LECTURE - APPLIED Regression - MLR Models

Extra Credit Problem

Problem: Cooling method for gas turbines. Refer to the Journal of Engineering for Gas Turbines and Power (January 2005) study of a high-pressure inlet fogging method for a gas turbine engine. A number of independent variables were used to predict the heat rate (kilojoules per kilowatt per hour) for each in a sample of 67 gas turbines augmented with high-pressure inlet fogging. The available independent variables are:

X_1 = RPM = cycle speed (revolutions per minute), X_2 = CPRATIO= cycle pressure ratio,

X_3 = INLET-TEMP = inlet temperature, X_4 = EXH-TEMP= exhaust gas temperature,

X_5 AIRFLOW = air mass flow rate (kg per second), X_6 = POWER = horsepower (Hp units).

The researcher first used all the independent variables to predict the HEATRATE (Model-1)and found the following ANOVA USING R.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
RPM	1	119598530.46	119598530.46	568.28	0.0000
CPRATIO	1	22745478.32	22745478.32	108.08	0.0000
INLETTEMP	1	9020839.00	9020839.00	42.86	0.0000
EXHTEMP	1	2915998.19	2915998.19	13.86	0.0004
AIRFLOW	1	774427.05	774427.05	3.68	0.0598
POWER	1	214461.94	214461.94	1.02	0.3168
Residuals	60	12627473.49	210457.89		
Total	66	167897207.45			

(1) Is this model (Model-1) significant at 5% level? (Must Write Hypothesis)

(2) Considering Backward selection, researcher drop POWER and rerun the model (Model-2), is AIRFLOW a significant at 5% level in this new model (Model-2) ?

Solution:

(1) Is this model (Model-1) significant at 5% level? (Must Write Hypothesis)

The Model is: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_6 X_6 + error$

$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{12627473.49}{167897207.45} = 0.9248$

Now to test the null hypothesis we set up the hypothesis as

$H_0 : \beta_1 = \beta_2 = \ldots = \beta_6 = 0$

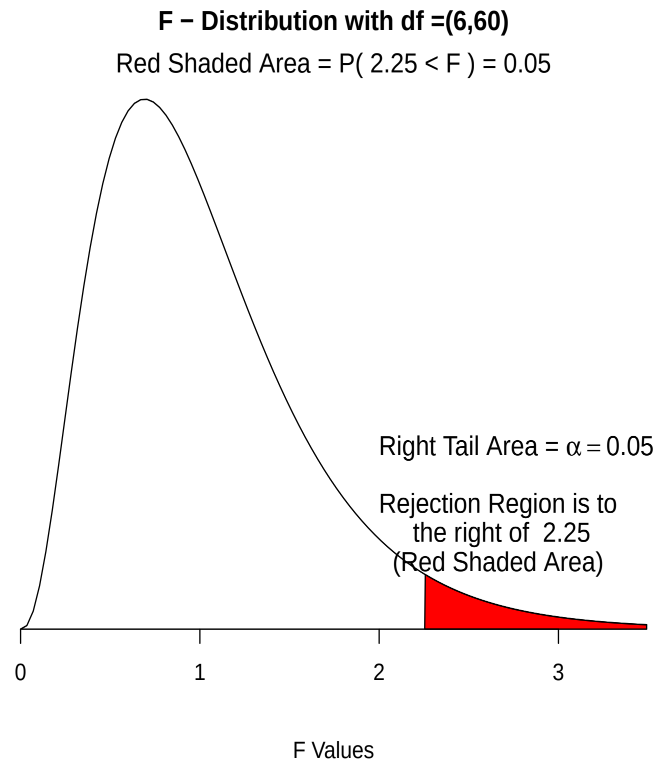
vs

H_1 : at least one of them is non-zero

$F - stat = \frac{R^2/K}{(1-R^2)/(n-k-1)} = 122.97$

As F-stat=122.97 $>$ $F_{0.05,6,60} = 2.25$ we reject the null hypothesis.

Conclusion: Model is significant at 5% level.



(2) Considering Backward selection, researcher drop POWER and rerun the model (Model-2), is AIRFLOW a significant at 5% level in this new model (Model-2) ?

To check whether X_5 = AIRFLOW is significant or not we compare the SSE of Model-2 with the model after we drop X_5 = AIRFLOW (Call it Model-3).

Model-2: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + Error$

Model-3: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + Error$

SSE of Model-2 $= SSE_2 = 214461.94 + 12627473.49 = 12841936.43$

SSE of Model-3 $= SSE_3 = 214461.94 + 12627473.49 + 774427.05 = 13616363.48$

We need to test $H_0 : \beta_5 = 0 \ vs \ \beta_5 \neq 0$

$F-stat = \frac{(SSE_3 - SSE_2)/1}{SSE_2/61} = \frac{774427.05/1}{12841936.43/61} = 3.678$

As F-Stat $= 3.678 \not> F_{0.05,1,61} = 4$ we do not reject the null hypothesis p-value $= 0.0598$

Conclusion: X_5 = AIRFLOW is NOT significant at 5% level.