

MA 542: REGRESSION ANALYSIS

QUIZ - 6

Name Key

The following is the R-outputs of the ANOVA table for COMMERCIAL PROPERTIES example. (Y-rental rates, X1-age, X2-expences and taxes, X3-vacancy rates, X4-squre footage).

Analysis of Variance Table

Response: Y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
X1	1	14.819	14.819	11.4649	0.001125 **
X2	1	72.802	72.802	56.3262	9.699e-11 ***
X3	1	8.381	8.381	6.4846	0.012904 *
X4	1	42.325	42.325	32.7464	1.976e-07 ***
Residuals	76	98.231	1.293		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- a) Perform a four step hypotheses test to test whether both X3 and X4 can be dropped from the regression model given that X1 and X2 are retained; use $\alpha = 0.05$. (You may use the textbook for the distribution tables).

Hypotheses:

$$H_0: \beta_3 = \beta_4 = 0 \quad \text{vs} \quad H_1: \text{not } H_0$$

$$\text{Full model: } Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \epsilon_i$$

$$\text{Reduced model: } Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i$$

$$F^* = \frac{[SSR(X_3/X_1, X_2) + SSR(X_4/X_1, X_2, X_3)]/2}{MSE(X_1, X_2, X_3, X_4)}$$

$$= \frac{(8.381 + 42.325)/2}{1.293} = 19.60$$

But $F(0.05, 2, 76) \approx F(0.05, 2, 60) = 3.15$ (approximated value)

Since $F^* > F(0.05, 2, 60)$, H_0 is rejected.

So both X_3 and X_4 can not be dropped from the model with the predictors X_1 and X_2 .

b) Refer to the previous problem and calculate $R^2_{4|123}$. Interpret your result.

$$\begin{aligned} R^2_{4|123} &= \frac{SSR(X_4 | X_1, X_2, X_3)}{SSE(X_1, X_2, X_3)} \\ &= \frac{SSR(X_4 | X_1, X_2, X_3)}{SSR(X_4 | X_1, X_2, X_3) + SSE(X_1, X_2, X_3, X_4)} \\ &= \frac{42.325}{42.325 + 98.231} = 0.3011 \end{aligned}$$

So 30.11% of left over variation of rental rates (Y) after regressing on X_1, X_2 and X_3 is explained by X_4 .