Stat-440/640 Regression and Time Series Analysis
Fall 2018

Exam 1A

Name _____

show work to ensure full credit

1. Accu-Copiers Inc., sells and services Accu-500 copying machines. As a part of its standard service contact, the company agrees to perform routine service on this copier. To obtain information about the time it takes to perform routine service, Accu-Copiers has collected data for 11 service calls. The data are given in the following table:

Service Call	Number of Copiers Serviced (x)	Number of Minutes Required (y)
1	4	109
2	2	58
3	5	138
4	7	189
5	1	37
6	3	82
7	4	103
8	5	134
9	2	68
10	4	112
11	6	154

The R output of a simple liner regression analysis of the service time data

Coefficients	Estimate	Std. Error	t value	P
(Intercept)	11.4641	3.4390	3.334	0.00875
x	24.6022	0.8045		

Analysis of variance

Source	DF	Sum of Squares	Mean Squares	F
Regression			19918.84	
Error		191.70		
Total	10	20110.55		

(a) (7 pts) Some entries in the above tables have been erased. Complete the tables by putting the correct numbers in the boxes.

(b) (4 pts) Find the least-squares regression line fitted to the service time data.

(c) (4 pts) Use the least-squares line to obtain a point estimate of the mean time to service four copiers on a single call.

(d) (4 pts) Use the least-squares line to obtain a point prediction of the time to service four copiers on a single call.

(e) (5 pts) Identify s_{b_1} and the t- statistic for testing the significance for the slope. Show how t has been calculated using b_1 and s_{b_1}.

(f) (9 pts) Using the t-statistics and the appropriate rejection region, test $H_0 : \beta_1 = 0$ versus $H_A : \beta_1 \neq 0$ by setting $\alpha = 0.05$. What do you conclude about the relationship between y and x?

(g) (5 pts) How much of the variation in time for service is explained by the proposed model?

(h) (5 pts) Calculate the value of adjusted R^2.

(i) (10 pts) Calculate the simple correlation r between x and y. Use this to test $H_0 : \rho = 0$ versus $H_A : \rho \neq 0$ by setting $\alpha = 0.05$. What can you say about the relationship between y and x?

2. (10 pts) In multiple linear regression with k regressors we use $D = x_0^T (X^T X)^{-1} x_0$ to find confidence interval for the mean response of the dependent variable, where

$$x_0 = \begin{bmatrix} 1 \\ x_{01} \\ x_{02} \\ \vdots \\ x_{0k} \end{bmatrix} \text{ and } X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

Show that for the simple linear regression the expression $x_0^T (X^T X)^{-1} x_0$ is same as $\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{s_{xx}} \right)$, where $s_{xx} = \sum x_i^2 - (\sum x_i)^2 / n$

3. (12 pts) Circle the best possible answer of the following:

(a) The MSE is an estimator of:

$$(a)\ \epsilon \qquad (b)\ \sigma^2 \qquad (c)\ 0 \qquad (d)\ Y$$

(b) In multiple regression with p predictor variables, when constructing a confidence interval for any β_j, the degrees of freedom for the tabulated value of t should be:

$$(a)\ n-1 \quad (b)\ p-1 \quad (c)\ n-p \quad (d)\ n-p-1$$

(c) In a regression study, a 95% confidence interval for β_3 was given as:$(-5.65, 2.61)$. What would a test for $H_0 : \beta_3 = 0$ versus $H_A : \beta_3 \neq 0$ conclude?

i. reject the null hypothesis at $\alpha = 0.05$ and all smaller α
ii. fail to reject the null hypothesis at $\alpha = 0.05$ and all smaller α
iii. reject the null hypothesis at $\alpha = 0.05$ and all larger α
iv. fail to reject the null hypothesis at $\alpha = 0.05$ and all larger α

(d) In general, the Least Squares Regression approach finds the equation:

i. that includes the best set of predictor variables
ii. of the best fitting straight line through a set of points
iii. with the highest R^2, after comparing all possible models
iv. that has the smallest sum of squared errors