COVER PAGE

STAT 608 Homework 07 Summer 2017

Please TYPE your name and email address below, then convert to PDF and attach as the first page of your homework upload.

NAME: RAJAN KAPOOR

EMAIL: r.kapoor@tamu.edu

35/35

Solution 1)

Assuming unrelated regression lines model
$Y =$ TOTALCWG , $x =$ STAGE

$$M1 \begin{cases} Y = \beta_0 + \beta_1 x + e & TRT = 0 \\ Y = \beta_0 + \beta_2 + (\beta_1 + \beta_3)x + e & TRT = 1 \end{cases}$$

We call model M1 as full model.

Now to test null hypothesis

$H_0 : \beta_3 = 0$ (identical slopes)

$H_A : \beta_3 \neq 0$

The following is the reduced model

$$M2 \begin{cases} Y = \beta_0 + \beta_1 x + e & (TRT = 0) \\ Y = \beta_0 + \beta_2 + \beta_2 x + e & (TRT = 1) \end{cases}$$

The null hypothesis can be tested using partial-F test

$$F = \frac{\dfrac{RSS(\text{reduced}) - RSS(\text{full})}{df(\text{red}) - df(\text{full})}}{RSS(\text{full})/df_{\text{full}}}$$

For model M1

$$Y = 5.4760 + 2.4934x \qquad TRT=0$$

$$Y = 7.7143 + 0.4785x \qquad TRT=1$$

For reduced model

$$Y = 7.2298 + 1.5113x \qquad TRT=0$$

$$Y = 5.91248 + 1.5113x \qquad TRT=1$$

$$F = \frac{\frac{187.59 - 155.22}{1}}{\frac{155.22}{19}} = 3.9623$$

p value $= 0.001786 < \alpha (= 0.05)$

$\Rightarrow$ Reject Null Hypothesis

Solution 2

2)

$$\log \frac{p(x)}{1-p(x)} = \hat{\beta}_0 + \hat{\beta}_1 x$$

$$\Rightarrow \quad p(x) = \exp(\hat{\beta}_0 + \hat{\beta}_1 x) + \exp(\hat{\beta}_0 + \hat{\beta}_1 x)\, p(x)$$

$$\Rightarrow \quad p(x) = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x)} = \frac{1}{\exp(-\hat{\beta}_0 - \hat{\beta}_1 x) + 1}$$

at $x = 3$

$$p(x) = \frac{1}{\exp[-(-2.643 + 0.674 \times 3)] + 1}$$

$\Rightarrow P(x) = 0.3495$

$\therefore$ probability that the insect will survive $= 1 - p(x)$
$= 0.6504$

Expected no. of surviving insects $= 200 \times (1 - p(x))$
$= 130.$

2.2 Odds on dying $= \dfrac{p(x)}{1-p(x)} = \exp(\hat{\beta}_0 + \hat{\beta}_1 x)$

if x increases by 1 unit, odds on dying increases by $\phi = \exp(0.674)$
$= 1.96207$

2.3 90% CI for ϕ $\exp\left(\hat{\beta}_1 \pm z_{0.05} * s.e(\hat{\beta}_1)\right)$ $\alpha = 0.1$, $\alpha/2 = 0.05$

$\exp(0.674 \pm 1.6448 \times 0.039)$

$(1.840, 2.092)$

Solution 3

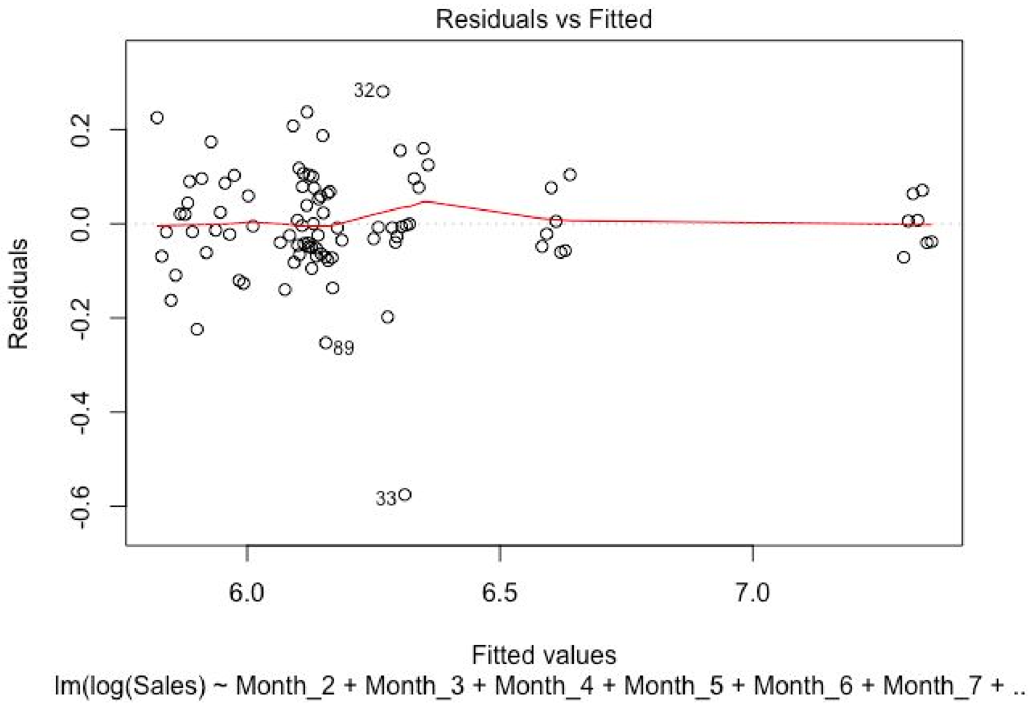
$\log(Sales) = \beta_0 + Time + Month\_i \quad i = 2, \ldots, 12$

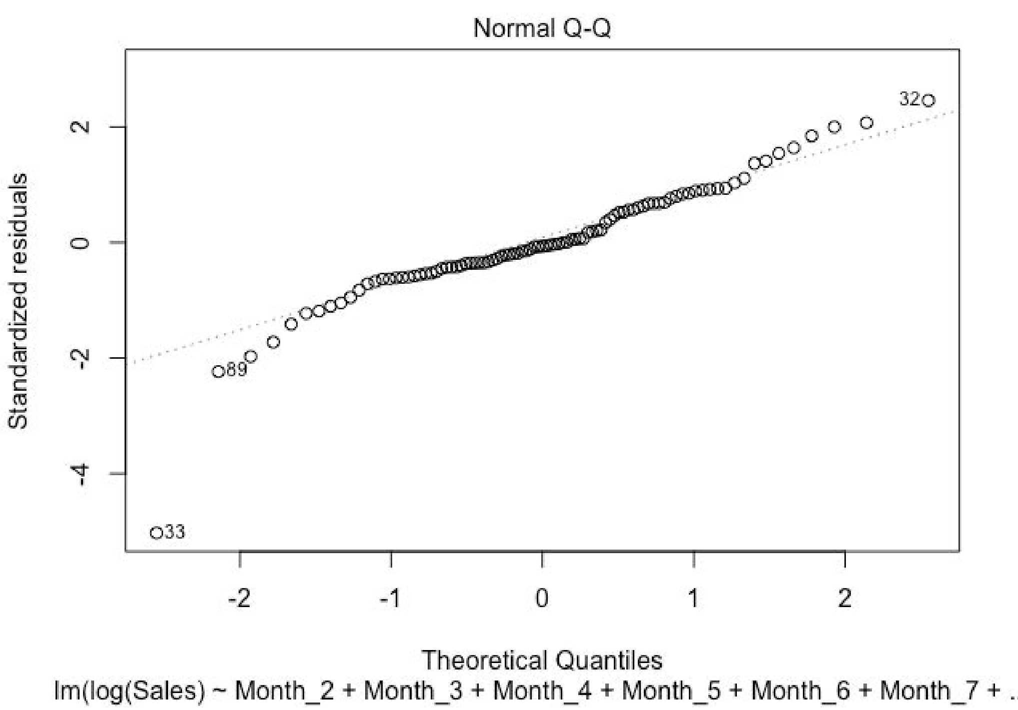
The diagnostic plots and summary from R is attached.

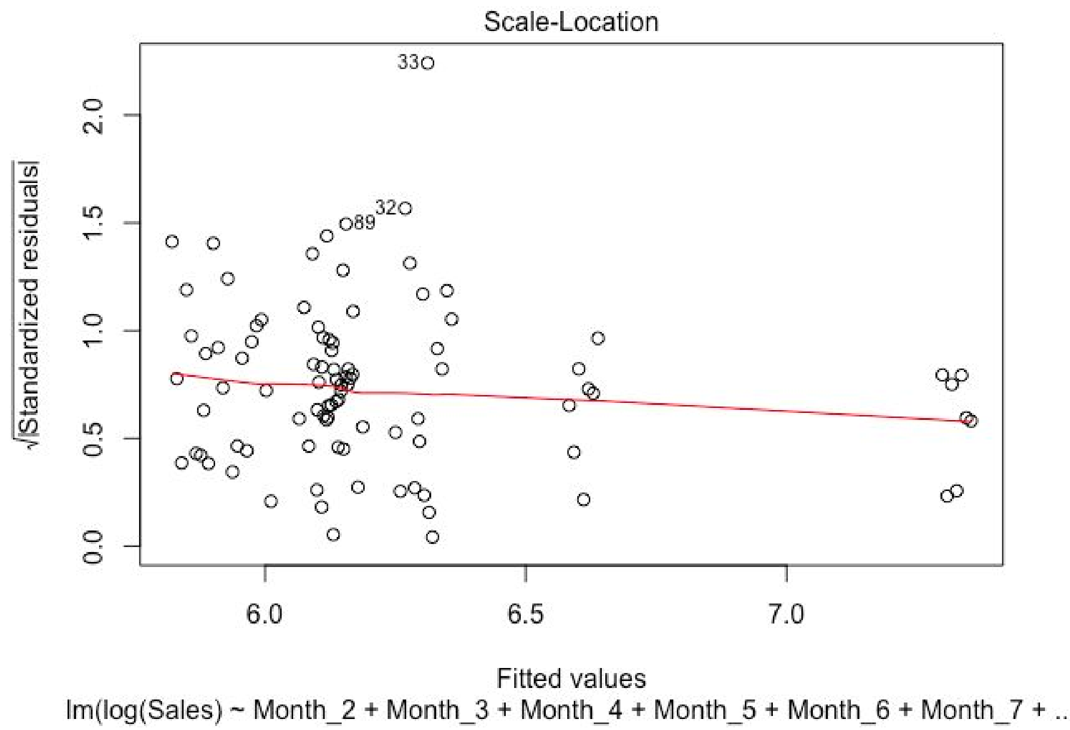
Weakness - Variables Time, Month_2, Month_7 are not statistically significant
- Case 32, 33, 41, 89 have standard resid > 2.

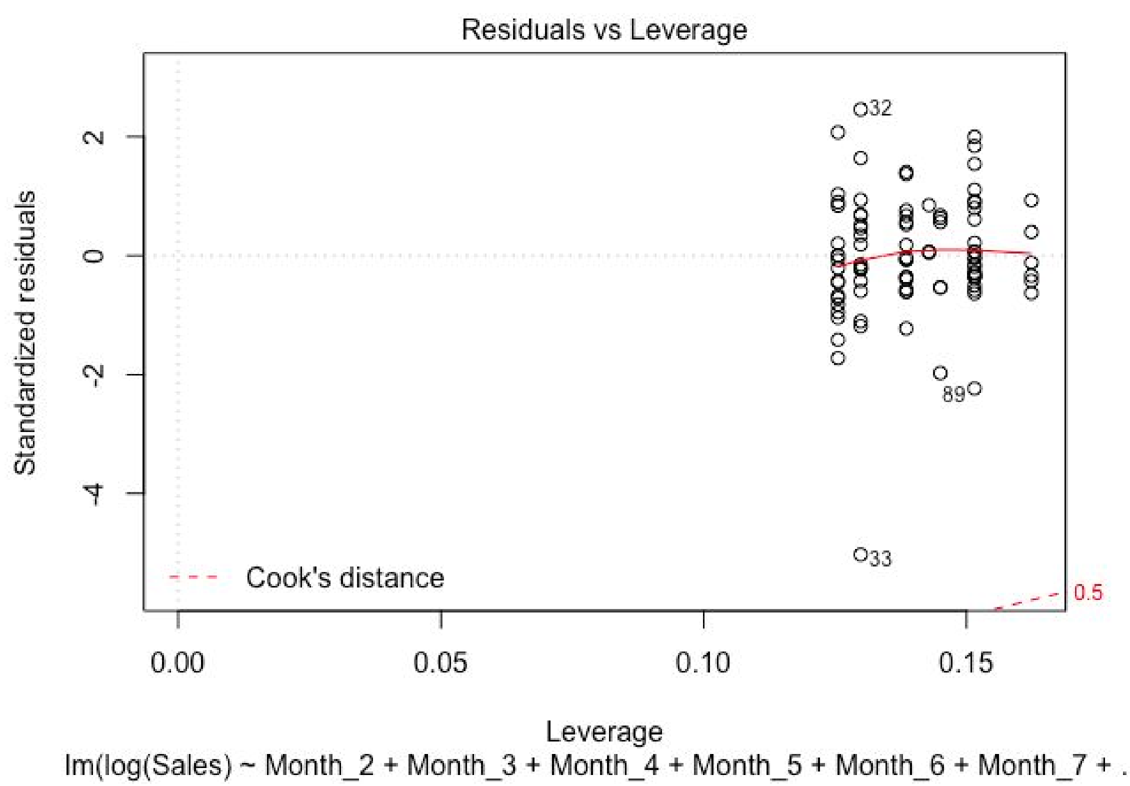
```
Call:
lm(formula = log(Sales) ~ Month_2 + Month_3 + Month_4 + Month_5 +
    Month_6 + Month_7 + Month_8 + Month_9 + Month_10 + Month_11 +
    Month_12 + Time)

Residuals:
     Min       1Q   Median       3Q      Max
-0.57520 -0.05166 -0.00739  0.07133  0.28116

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.8723775  0.0518913 113.167  < 2e-16 ***
Month_2     -0.0522214  0.0635314  -0.822 0.413533
Month_3      0.1907035  0.0635154   3.002 0.003572 **
Month_4      0.2474630  0.0635029   3.897 0.000201 ***
Month_5      0.2270750  0.0634940   3.576 0.000595 ***
Month_6      0.2138238  0.0634887   3.368 0.001167 **
Month_7      0.0690719  0.0634869   1.088 0.279874
Month_8      0.2208295  0.0634887   3.478 0.000819 ***
Month_9      0.3708060  0.0634940   5.840 1.07e-07 ***
Month_10     0.4133160  0.0635029   6.509 6.13e-09 ***
Month_11     0.7023337  0.0655759  10.710  < 2e-16 ***
Month_12     1.4169818  0.0655707  21.610  < 2e-16 ***
Time         0.0007670  0.0004756   1.613 0.110755
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1227 on 80 degrees of freedom
Multiple R-squared:  0.9111,    Adjusted R-squared:  0.8978
F-statistic: 68.33 on 12 and 80 DF,  p-value: < 2.2e-16
```

Diagnostic Plots:


Residuals vs Fitted
0.2
0.0
-0.2
-0.4
-0.6
Residuals
32
89
33
6.0
6.5
7.0
Fitted values
lm(log(Sales) ~ Month_2 + Month_3 + Month_4 + Month_5 + Month_6 + Month_7 + ..


Normal Q-Q
2
0
-2
-4
Standardized residuals
32
89
33
-2
-1
0
1
2
Theoretical Quantiles
lm(log(Sales) ~ Month_2 + Month_3 + Month_4 + Month_5 + Month_6 + Month_7 + ..


Scale-Location
√|Standardized residuals|
Fitted values
lm(log(Sales) ~ Month_2 + Month_3 + Month_4 + Month_5 + Month_6 + Month_7 + ..
0.0
0.5
1.0
1.5
2.0
6.0
6.5
7.0
33
32
89


Residuals vs Leverage
Standardized residuals
Leverage
lm(log(Sales) ~ Month_2 + Month_3 + Month_4 + Month_5 + Month_6 + Month_7 + ..
Cook's distance
0.5
-4
-2
0
2
0.00
0.05
0.10
0.15
32
89
33

Solution 4

(4.1)
$$\underbrace{\begin{bmatrix} e_2 \\ e_3 \\ e_4 \\ \vdots \\ e_n \end{bmatrix}}_{Y} = \rho \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_{n-1} \end{bmatrix}}_{X} + \begin{bmatrix} \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad \text{— (M1)}$$

Clearly a linear regression of Y on X through the origin will estimate ρ. Since all 4 assumptions of Regression Model –

- Y is related to X by simple reg. model ($\because e_t = \rho e_{t-1} + \epsilon_t$)
- ϵ_t are independent of each other
- ϵ_t have common variace
- ϵ_t are normally distributed with mean zero & common variance (=1 $\because$ standard normal) distribution given for ϵ_t (least sq estimate)

$$\hat{\rho} = \text{slope} = 0.412$$

(4.2) YES. the valid s.e. for $\hat{\rho}$ should be 0.133, since Model M1 is valid regression model (linear).