STAT2008/STAT4038/STAT6038 - Regression Modelling - Sem 1 2017

QUIZ NAVIGATION

1 2 3 4 5

Show one page at a time

Finish review

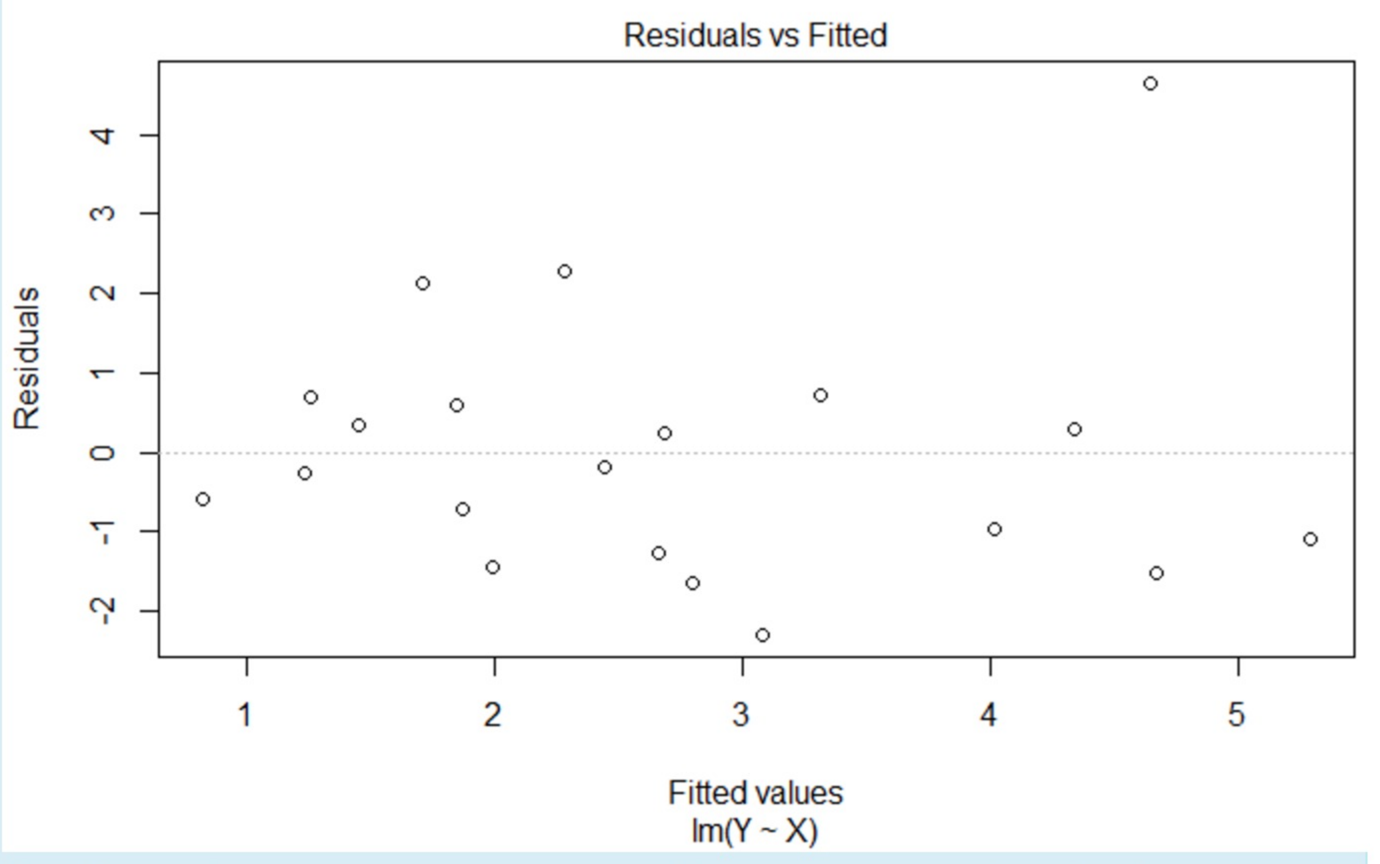
Started on	Tuesday, 28 March 2017, 5:10 PM
State	Finished
Completed on	Tuesday, 28 March 2017, 5:25 PM
Time taken	15 mins 16 secs
Grade	4.00 out of 5.00 (80%)
Feedback	Detailed feedback on the quiz questions will be available again, once the quiz has closed at 3pm on Friday 31 March 2017. I hope you are enjoying the course and getting a lot out of it. There will be a short poll available in the first topic on Wattle, during the break (from 6 April 2017 to 15 April 2017) for you to give me some anonymous feedback on the course. Any constructive comments will be welcome.

Question **1**

Correct

Mark 1.00 out of 1.00

Flag question



What (if anything) is wrong with the above residual plot? Please select the best ONE of the following options:

Select one:

- A. Residuals are not independent (obvious pattern)
- B. Residuals do not display constant variance
- C. Residuals are not normally distributed
- D. There are obvious outliers and/or influential observations ✓ This is the best option with this plot.
- E. More than one of the above problems
- F. No obvious problems

Your answer is correct.

There is an obvious vertical outlier (an observation which is outlying in the vertical or Y direction) located in upper right hand corner of the plot. This data point has probably also been influential in determining the overall fit of the regression model and may also have caused other apparent problems in the plot (a slight downwards trend in the other residuals, possible decreasing variance if you ignore the outlier).

The correct answer is: There are obvious outliers and/or influential observations

Question **2**

Correct

Mark 1.00 out of 1.00

Flag question

In a simple linear regression model, the estimated regression equation is $\hat{Y} = 5 - 2X$ and the coefficient of determination, R^2, is 0.81, then what is the value of r, the sample correlation coefficient between X and Y?

Answer: -0.9 ✓

Good, spot on!

The correlation coefficient, r, is equal to the square root of the coefficient of determination, with the same sign as the slope coefficient of the estimated regression equation.

So in this instance, the sample correlation coefficient between X and Y will be −0.9, indicating that X and Y are strongly negatively correlated (Y will decrease as X increases).

The correct answer is: -0.9

Question **3**

Correct

Mark 1.00 out of 1.00

Flag question

The sample residuals e_i are independently and identically (normally) distributed with mean 0 and constant variance σ^2.

Select one:

- True
- False ✓

Good, your answer is correct.

The errors ε are assumed to be independently and identically (normally) distributed with mean 0 and constant variance σ^2.

But the sample residuals (the estimated errors) are neither independent nor do they have constant variance.

The variance-covariance matrix of the residuals is $\sigma^2(\mathbf{I} - \mathbf{H})$, which is typically not a diagonal matrix. As a result the sample residuals are correlated and the variance of each residual is $\sigma^2(1 - h_{ii})$, which is not necessarily constant.

The correct answer is 'False'.

Question **4**

Correct

Mark 1.00 out of 1.00

Flag question

Given data $(w_1, z_1), \ldots, (w_n, z_n)$, you define two new transformed variables:

$$x_i = \frac{(w_i - \bar{w})}{s_w} \;\; \& \;\; y_i = \frac{(z_i - \bar{z})}{s_z}, \qquad i = 1, \ldots, n$$

where s_w and s_z are the sample standard deviations of w and z respectively.

You then use least squares to estimate the simple linear regression model with y as the response variable and x as the explanatory variable.

Which of the following statements about this estimated simple linear regression model is NOT true:

Select one:

- A. The fitted line passes through the origin
- B. The sample residuals from the fitted model sum to zero
- C. The fitted line has intercept equal to zero
- D. The fitted line passes through the point $(\overline{w}, \overline{z})$. ✓ This is the only option that is NOT guaranteed to be true. The estimated regression model will pass through the mean of x and y (which are the two transformed variables used to estimate the model), but will not necessarily pass through the mean of the original untransformed variables.

Your answer is correct.

The two new transformed variables x and y are standardized versions of w and z respectively. In the top or numerator of these standardization transformations, the two variables are first mean-centered, then divided by the respective standard deviations.

Variables which are mean-centered (i.e. have had the mean or expected value subtracted from each of the data values) have mean 0, so both x and y have mean 0 and will sum to a total of 0. The sample residuals from a regression model are a similar example of a mean-centered variable.

An simple linear regression model estimated using least squares will always pass through the point defined by the mean of the two variables used to fit that model, so the estimated regression model will pass through the point (0, 0) and will therefore also have an intercept equal to 0.

The correct answer is: The fitted line passes through the point $(\overline{w}, \overline{z})$.

Question **5**

Incorrect

Mark 0.00 out of 1.00

Flag question

If the sample correlation coefficient between x and y is exactly zero, then the least squares estimate (b_1) of the slope coefficient in a simple linear regression model will also be exactly zero.

Select one:

- True
- False ✗

Sorry, your answer is not correct.

The sample correlation coefficient is equal to the sample covariance of x and y divided by the product of the sample standard deviations of both x and y. If the sample correlation coefficient is zero, then the sample covariance of x and y will also be zero.

The least squares estimate (b_1) of the slope coefficient in a simple linear regression model is estimated by dividing the sample covariance of x and y by the sample variance of x. So, it will also be zero, if the sample covariance is zero.

The correct answer is 'True'.

Finish review