Midterm Make-up

Tuesday, October 27, 2020 11:40 AM

15) → a) → simple regression equation y = murder/mil
x = % unemployed

c) → R^2 value

16) Cand d?

d → total variation left unexplained by the model
c → $R^2 = SSR/SST$

14)
1. Homoscedasticity
2. SSE is the total variation left unexplained. It's the difference between the data points and the line of best fit.
3. [0, 1]
4. R^2
5. [-1, 1]
6. The assumptions are that the data is normal, the extraneous values have a mean value of zero, and that there is no correlation between x and the extraneous values dependent given independent is normal

15) a) $Y = -28.52671 + 7.079554X$ → x = % unemployed

b) As unemployed goes up 1, murders go up by 7.079
At 0 unemployment, there are -28.53 murders

c) $R^2 = .7480$

16) a) $Y/X \rightarrow (y_i - \bar{y})(x_i - \bar{x}) / (x_i - \bar{x})^2 \rightarrow SPXY/SS_X = Y/X$

slope = 296446.059 / 147911.938 ≈ 2.004 = slope

107.43 = (128.88)(2.004) + y
107.43 = 258.304 + y
y intercept = -150.87 This doesn't make sense because time has to be positive but I don't have enough time to fix it

b) For every unit increase in distance, time increases by 2.004

When distance = 0, time = -150.87

c) $R^2 = SSR/SST = 1 - SSE/SST$
$SSE = 592892.117$ $SST = 1185784.23$
$R^2 = 1 - .500... \rightarrow R^2 = .4999$

e) $Y = (103)(2.004) - 150.87$
$Y = 55.542$

d) $SSE = 592892.117$
SSE is the total variation left unexplained by the model