hw3

313707025 jebuhdah

2025-03-15

'##
qn1: a. State the null and alternative hypotheses in terms of the model parameters. Null Hypothesis (H): There is no relationship between the number of medals won and GDP. In terms of the model parameters, this means ß2 = 0. Alternative Hypothesis (H): There is a positive relationship between the number of medals won and GDP. In terms of the model parameters, this means ß2 > 0.

b. What is the test statistic for part (a) and what is its distribution if the null hypothesis is true?

Test statistic for testing hypothesis 2 is given by: t= (b2-ß2)/(se(b2))

For the null hypothesis (ß2=0), the test statistic would be: t= b2/(se(ß2))

c. What happens to the distribution of the test statistic for part (a) if the alternative hypothesis is true?

the centre of the distribution will be shifted to the right relative to the usual t-distribution due to the expected value of b2 would be positive, leading to a larger t-statistic.

d. For a test at the 1% level of significance, for what values of the t-statistic will we reject the null hypothesis in part (a)? For what values will we fail to reject the null hypothesis?

```
## [1] 2.388011
```

so when the t-statistics is larger than 2.39, we will reject the null hypothesis, and we will fail to reject when its smaller than 2.39.

e. Carry out the t-test for the null hypothesis in part (a) at the 1% level of significance. What is your economic conclusion? What does 1% level of significance mean in this example?

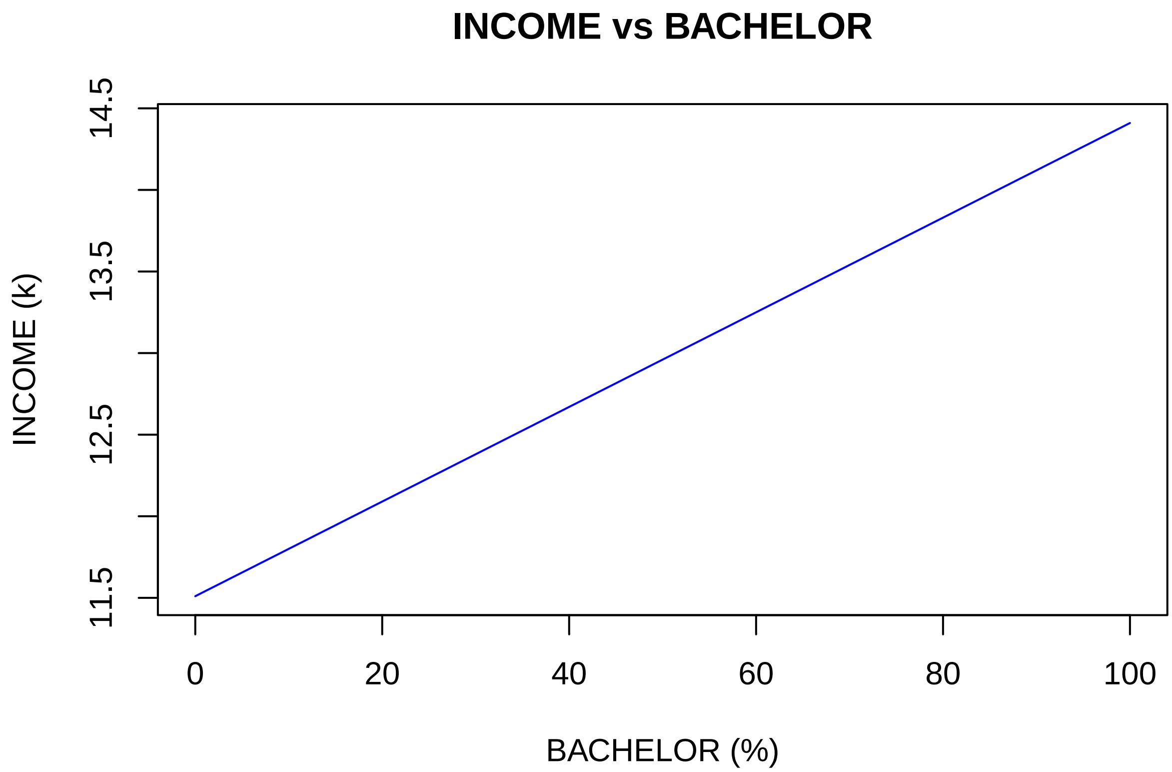
from above 2.39 is the critical t value.

```
## [1] 6.088372
```

since 6.09 > 2.39, so we reject the null hypothesis. so there is a positive relationship between gdp and medals.
'##
qn.7 a. Using the information provided calculate the estimated intercept.

a/2.672 = 4.31, a = 11.52

b. Sketch the estimated relationship. Is it increasing or decreasing? Is it a positive or inverse relationship? Is it increasing or decreasing at a constant rate or is it increasing or decreasing at an increasing rate?



it would be increasing since the slope coefficient 0.029 is positive.

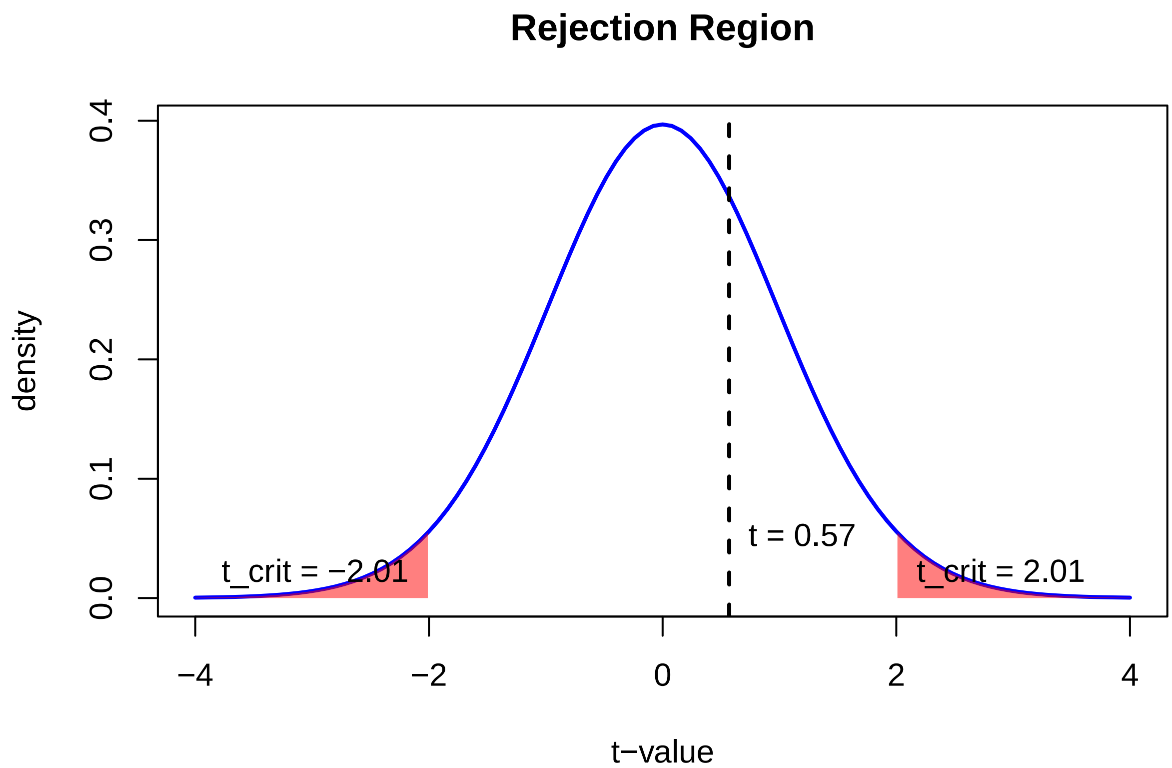
c. Using the information provided calculate the standard error of the slope coefficient. Show your work.

t = b/(se(b)), se(b) = b/t se(b) = 0.029/10.75 = 0.0027

d. What is the value of the t-statistic for the null hypothesis that the intercept parameter equals 10?

t = (11.52 - 10) / 2.7 = 0.56

e. The p-value for a two-tail test that the intercept parameter equals 10, from part (d), is 0.572. Show the p-value in a sketch. On the sketch, show the rejection region if a = 0.05



f. Construct a 99% interval estimate of the slope. Interpret the interval estimate.

```
## [1] 0.02176413
```

```
## [1] 0.03623587
```

g. Test the null hypothesis that the slope coefficient is one against the alternative that it is not one at the 5% level of significance. State the economic result of the test, in the context of this problem.

```
## [1] -359.6296
```

```
## [1] 2.009575
```

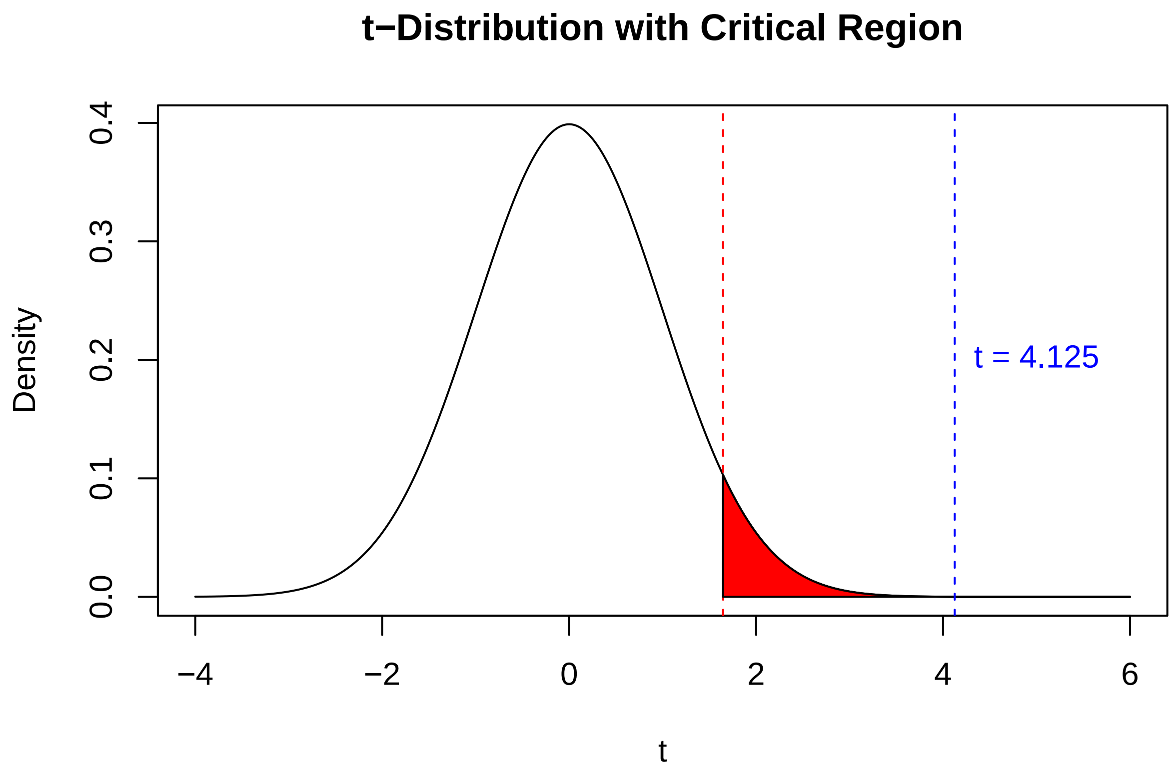
t statistics > critical t value, so we reject.

'##
qn17, a. Using the urban regression, test the null hypothesis that the regression slope equals 1.80 against the alternative that it is greater than 1.80. Use the a = 0.05 level of significance. Show all steps, including a graph of the critical region and state your conclusion.

t = (t2-1.8)/se(b2) = (2.46 - 1.8)/0.16 = 4.125

```
## [1] 1.646404
```

since the t statistics is greater than critical, null hypothesis is rejected.



b. Using the rural regression, compute a 95% interval estimate for expected WAGE if EDUC=16. The required standard error is 0.833. Show how it is calculated using the fact that the estimated covariance between the intercept and slope coefficients is -0.761.

```
## [1] 23.92
```

```
## [1] 1.971217
```

```
## [1] 22.27798
```

```
## [1] 25.56202
```

c. Using the urban regression, compute a 95% interval estimate for expected WAGE if EDUC=16. The estimated covariance between the intercept and slope coefficients is -0.345. Is the interval estimate for the urban regression wider or narrower than that for the rural regression in (b). Do you find this plausible? Explain.

```
## [1] 28.6
```

```
## [1] 0.8163945
```

```
## [1] 1.962378
```

```
## [1] 26.99793
```

```
## [1] 30.20207
```

it is narrowier.

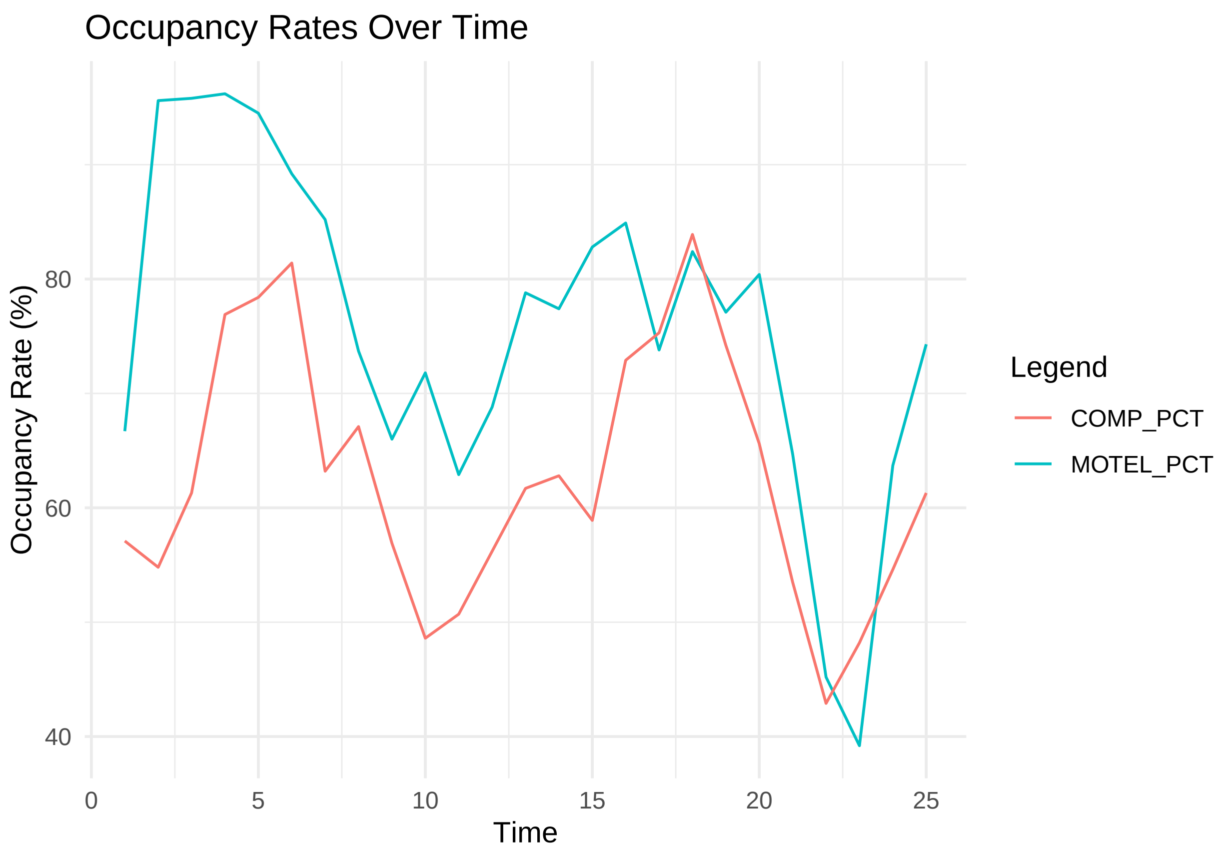
d. Using the rural regression, test the hypothesis that the intercept parameter ß1 equals four, or more, against the alternative that it is less than four, at the 1% level of significance.

t=(-4.88-4)/3.29= -2.7

```
## [1] -2.344066
```

the t statistics is smaller than the critical value, thus we reject the null

'##
qn19, a. Plot MOTEL_PCT and COMP_PCT versus TIME on the same graph. What can you say about the occupancy rates over time? Do they tend to move together? Which seems to have the higher occupancy rates? Estimate the regression model……




```
##
## Call:
## lm(formula = motel_pct ~ comp_pct, data = motel_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -23.876  -4.909  -1.193   5.312  26.818
```

```
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  21.4000    12.9069   1.658 0.110889
## comp_pct      0.8646     0.2027   4.265 0.000291 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.02 on 23 degrees of freedom
## Multiple R-squared:  0.4417, Adjusted R-squared:  0.4174
## F-statistic: 18.19 on 1 and 23 DF,  p-value: 0.0002906
```

```
##                 2.5 %    97.5 %
## (Intercept) -5.2998960 48.099873
## comp_pct     0.4452978  1.283981
```

PCT of Motel is larger than Comp and they move in unison.

b. Construct a 90% interval estimate of the expected occupancy rate of the motel in question, MOTEL_PCT, given that COMP_PCT = 70.

```
##        fit      lwr      upr
## 1 81.92474 77.38223 86.46725
```

c. In the linear regression model MOTEL_PCT = ß1 + ß2_COMP_PCT_ + e, test the null hypothesis H0:ß2<=0 against the alternative hypothesis H0:ß2>0 at the a = 0.01 level of significance. Discuss your conclusion. Clearly define the test statistic used and the rejection region.

```
## comp_pct
##  4.26536
```

```
## [1] 2.499867
```

since t_stat is larger than critical, we reject the null hypothesis.

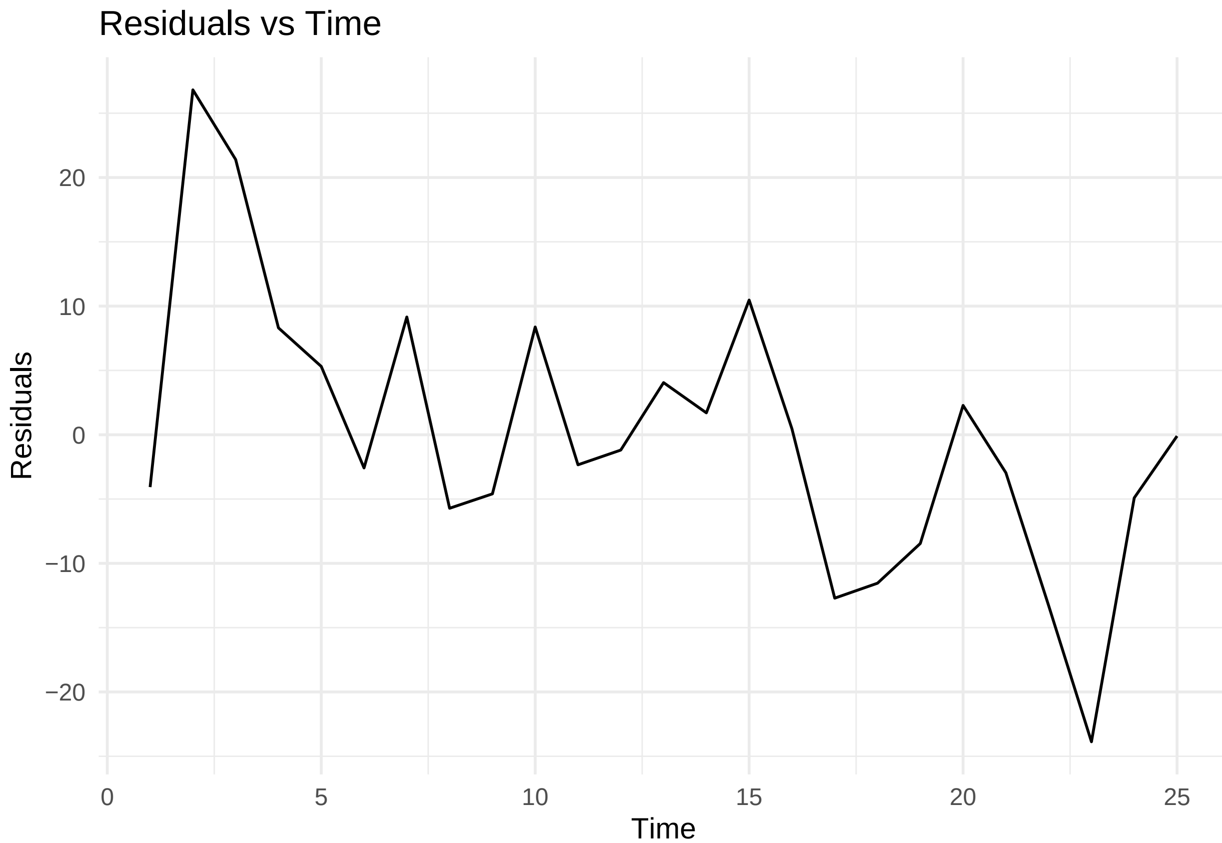
d. In the linear regression model MOTEL_PCT = ß1 + ß2_COMP_PCT_ + e, test the null hypothesis H0:ß2=1 against the alternative hypothesis H0:ß2=/=1 at the a = 0.01 level of significance. If the null hypothesis were true, what would that imply about the motel's occupancy rate versus their competitor's occupancy rate? Discuss your conclusion. Clearly define the test statistic used and the rejection region.

```
##   comp_pct
## -0.6677491
```

```
## [1] 2.807336
```

since the t stat is smaller than the critical value, we failed to reject the null hypothesis.

e. Calculate the least squares residuals from the regression of MOTEL_PCT on COMP_PCT and plot them against TIME. Are there any unusual features to the plot? What is the predominant sign of the residuals during time periods 17–23




```
## [1] "Negative"
```

it seems like the motel's occupancy rate was consistently lower than expected based on the competitor's occupancy rate.