Problem set 2

Luna Goldstein

15 October 2021

Question 1 (40 points): Political Science

Part 1:

The chi square test statistic is 3.791168

```
Fo_1 <- 14
Fo_2 <- 6
Fo_3 <- 7
Fo_4 <- 7
Fo_5 <- 7
Fo_6 <- 1
Fe_1 <- ((27/42) * 21)
Fe_2 <- ((27/42) * 13)
Fe_3 <- ((27/42) * 8)
Fe_4 <- ((15/42) * 21)
Fe_5 <- ((15/42) * 13)
Fe_6 <- ((15/42) * 8)
chisq <- (
  ((Fo_1 - Fe_1)^2/Fe_1) + ((Fo_2 - Fe_2)^2/Fe_2) +
    ((Fo_3 - Fe_3)^2/Fe_3) +  ((Fo_4 - Fe_4)^2/Fe_4) +
    ((Fo_5 - Fe_5)^2/Fe_5) + ((Fo_6 - Fe_6)^2/Fe_6)
)
chisq
```

Part 2:

As the p-value from the test statistic is larger than .1, we fail to reject the null hypothesis.
Code:

```
pchisq <- pchisq(3.791168, df = 2, lower.tail = FALSE)
pchisq
```

The p-value is 0.15

Part 3:

Standardized residuals for each cell :

class:	Not stopped	Bride request	Stopped/given a warning
Upper Class	0,322	-1,64	1,52
Lower Class	-0,322	1,64	-1,52

Part 4:

How might the standardized residuals help you interpret the results? The standardized residuals are the strength of the difference between the observed and the expected values. From the following we suggest that the Upper class' frequency is less than the expected frequency when 'bride request', in other circumstances, 'Not stopped' and 'Stopped/given a warning' the observed frequency is greater than the expected frequency. While, the opposite is observed for the Lower class.

Question 2 (20 points): Economics

Part 1

State a null and alternative (two-tailed) hypothesis:
The null hypothesis : No effect of the reservation policy on the number of new or repaired drinking water facilities in the villages.
The alternative hypothesis : The reservation policy leads to an increased/decreased number of new or repaired drinking facilities in the villages.

Part 2

The Bivariate regression:

```
view(Economics)
lm1 <- lm(Economics$irrigation~Economics$reserved)
summary(lm1)
```

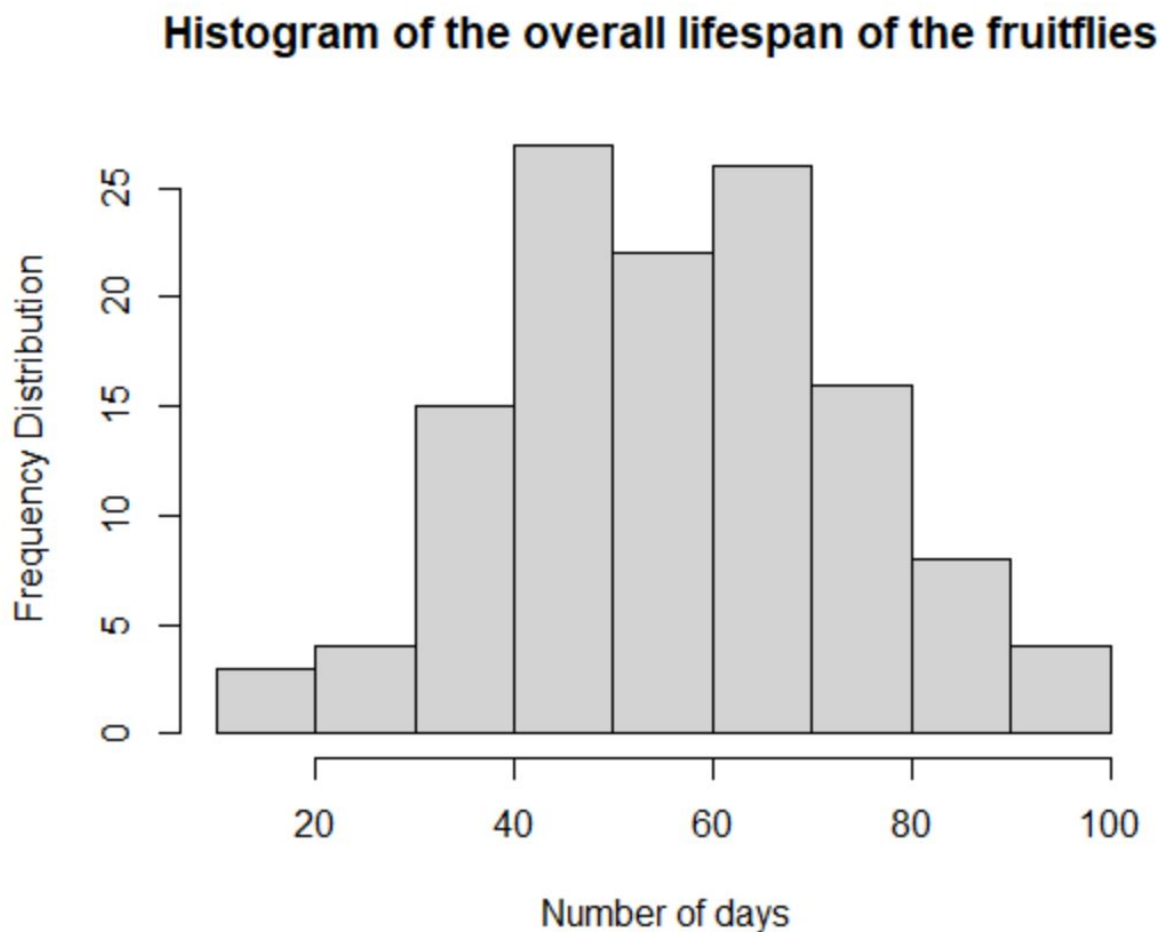
Because the p-value is higher than .05, we fail to reject the null hypothesis.

Part 3

Interpret the coefficient estimate for reservation policy: The coefficient estimate is -0.36. Because this value is negative we propose that the reservation policy has a negative relationship with the irrigation repair systems.

Question 3 (40 points): Biology

Part 1



From this graph we observe that the average of fruit flies appears to live between 40 and 75 days.

Part 2

There appears to be a linear relationship between these two variables. The correlation coefficient is 0.63, which suggests a strong positive correlation.

Part 3

The value of the slope is 144.33, this value indicates that for every 0.1 increase in thorax, longevity increases 14,43 days.

Part 4

The estimates: 144.33
The std. Error: 15.77
The test statistic: 9.15
The p-value: 1.5 e-15

From the results of the test we reject the null hypothesis (= there is no linear relationship between lifespan and thorax).

Part 5

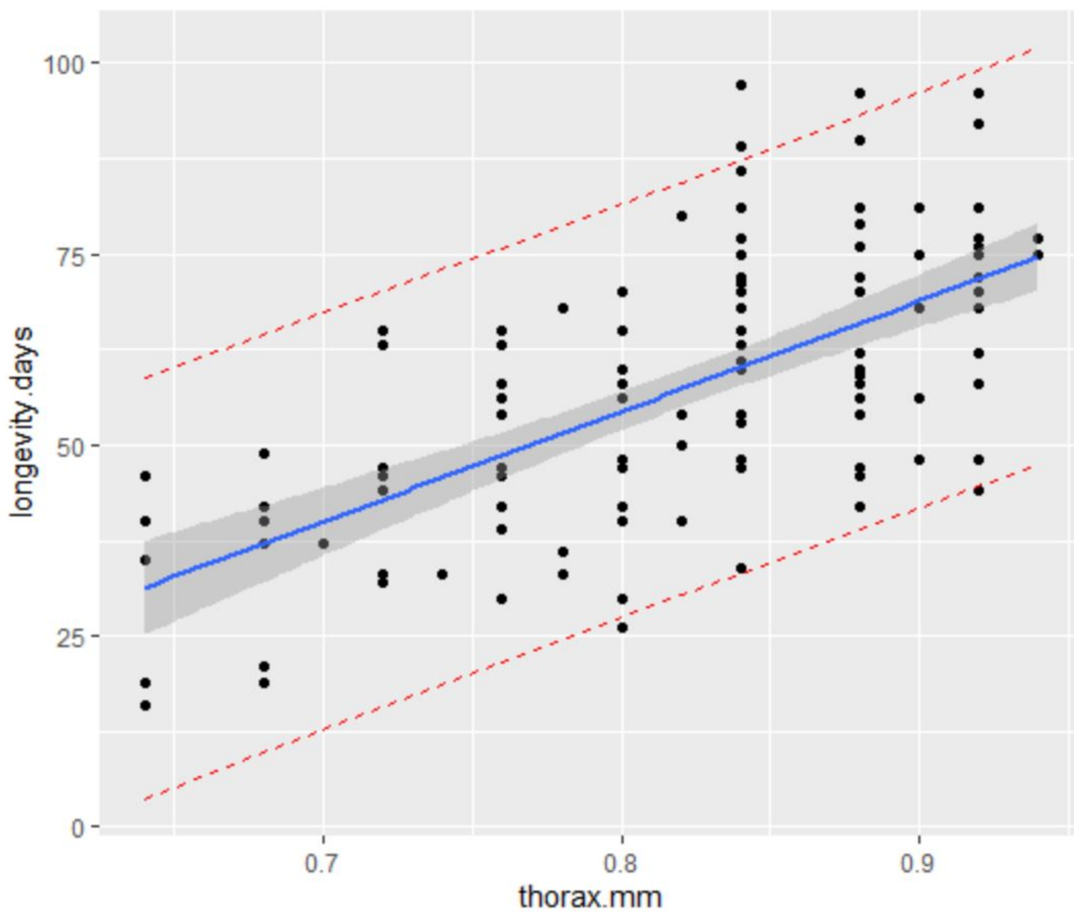
The the 90% confidence interval for the slope of the fitted model: By hand:
Lower Confidence Interval : (144.33 - ((15.77*1.645))) = 118.38
Upper Confidence Interval : (144.33 + ((15.77*1.645))) = 170.27

Part 6

The expected values of the lifespan are 54.414. This suggest that the fly would live approximately 54 days. Lower confidence Interval = 51.92 Upper Confidence Interval = 56.91

Part 7

Plot with the fitted values of the lifespan with the prediction intervals and confidence intervals:


100
75
50
25
0
longevity.days
0.7
0.8
0.9
thorax.mm