Lucas Faijdherbe(2594812) & Ruben van der Ham(2592271) – CS40

**Assignment 4**

**Exercise 4.1**

We are testing to see if an r = 0.876 based on n = 40 pairs are significantly different from 0, so, if the r value indicates that the Weight Watcher program is effective in reducing weight.

**Hypothesis:**

*Null hypothesis:*H0: ρ = 0

*Alternative hypothesis:*

Ha: ρ ≠ 0

**Significance level:**

α = 0.01

**Data:**

Data yields r = 0.876, n = 40.

**Test statistic:**

Tρ has a t-distribution with n – 2 = 40 – 2 = 38 degrees of freedom.

Observed value tρ = r = 11.1961483659

**Critical values:**

We have a two-tailed test with α = 0.01 and n = 40, so we get the critical values

-t38,0.01 and t38,0.01, which gives: -2.712 and 2.712.

Since tρ = 11.196 > 2.712, we reject H0.

**Conclusion:**

There is enough evidence to reject the claim that there is no linear correlation between the before weight and the after weight. So, one could argue that the value of r indicates that the Weight Watcher program is effective in reducing weight.

**Exercise 4.2**

In his book Outliers, author Malcolm Gladwell argues that more baseball players have birthdates in the months immediately following July 31, because that was the cutoff date for non-school baseball leagues. Here is a sample of frequency counts of months of birthdates of American-born major league baseball players starting with January: 387,329,366,344,336,313,313,503,421,434,398,371. Using a 0.05 significance level, is there sufficient evidence to warrant the rejection of the claim that American-born major league baseball players are born in different months with the same frequency?

**Exercise 4.3**

a)

b)

c)

d)

**Exercise 4.4**

a)

**Appendix**