STAT 260 - R Assignment 1

Parker DeBruyne - V00837207

18/01/2023

Part 1

In a study of factors thought to be responsible for the adverse effects of smoking on human reproduction, cadmium level determinations (nanograms per gram) were made on placenta tissue of a random sample of 22 mothers who were smokers and an independent random sample of 18 non-smoking mothers. The results were as follows:

Non-smokers: 17.4 14.1 19.7 20.2 15.5 28.2 18.7 18.0 21.4 14.6 14.3 13.9 19.6 18.7 17.4 18.8 20.1 16.0

Smokers: 25.5 24.3 20.4 18.5 29.0 31.1 24.6 19.9 18.3 24.8 30.0 13.6 25.2 27.1 20.9 22.6 26.8 20.2 25.3 22.9 20.2 32.4

(a) Store the data for non-smokers in a vector called nonsmokers. We want to test the research hypothesis that the mean cadmium level for non-smoking mothers is less than 19 nanograms per gram. Copy and paste the appropriate R command and output you would use to do this test.

```
nonsmokers = c(17.4, 14.1, 19.7, 20.2, 15.5, 28.2, 18.7, 18.0, 21.4,
               14.6, 14.3, 13.9, 19.6, 18.7, 17.4, 18.8, 20.1, 16.0)

t.test(nonsmokers, mu=19, alternative="less")
```

```
##
##  One Sample t-test
##
## data:  nonsmokers
## t = -1.0588, df = 17, p-value = 0.1522
## alternative hypothesis: true mean is less than 19
## 95 percent confidence interval:
##      -Inf 19.55008
## sample estimates:
## mean of x
##  18.14444
```

(b) Clearly state the p-value you found from part (a). Using the significance level $= 0.10$, should we reject the null hypothesis?

(c) Write a sentence stating your conclusion of the hypothesis test you just performed. (i.e. Write your conclusion in plain language so that a non-statistician could understand the conclusion of the test.)

(d) Store the data for smokers in a vector called smokers. Calculate a 95% confidence interval (i.e. a two-sided confidence interval) for the mean cadmium level of mothers who smoke. Copy and paste the appropriate R command and output you would use to create your confidence interval. In an additional line, clearly state the values you get for the confidence interval from your output. In an additional line, clearly state the values you get for the confidence interval from your output. In an additional line, clearly state the values you get for the confidence interval from your output. In an additional line, clearly state the values you get for the confidence interval from your output. In an additional line, clearly state the values you get for the confidence interval from your output. In an additional line, clearly state the values you get for the confidence interval from your output.

```r
smokers = c(25.5, 24.3, 20.4, 18.5, 29.0, 31.1, 24.6, 19.9,
            18.3, 24.8, 30.0, 13.6, 25.2, 27.1, 20.9, 22.6,
            26.8, 20.2, 25.3, 22.9, 20.2, 32.4)

t.test(smokers, mu=25, alternative="two.sided")
```

```
##
##  One Sample t-test
##
## data:  smokers
## t = -1.2195, df = 21, p-value = 0.2362
## alternative hypothesis: true mean is not equal to 25
## 95 percent confidence interval:
##  21.75365 25.84635
## sample estimates:
## mean of x
##      23.8
```

(e) Is it reasonable to conclude that the mean cadmium level of mothers who smoke is 25 nanograms per gram? Explain, using your confidence interval.

Part 2 From a random sample of 721 items made by a particular manufacturing process, it is found that 56 are defective.

(a) Suppose we want to test the research hypothesis that to proportion of defective items is greater than 6%. Copy and paste the appropriate R command and output you would use to do this test.

```r
binom.test(x=37, n=721, p=0.06, alternative="greater")
```

```
##
##  Exact binomial test
##
## data:  37 and 721
## number of successes = 37, number of trials = 721, p-value = 0.8563
## alternative hypothesis: true probability of success is greater than 0.06
## 95 percent confidence interval:
##  0.03849977 1.00000000
## sample estimates:
## probability of success
##             0.05131761
```

(b) Clearly state the p-value you found from part (a). Using the significance level $= 0.05$, should we reject the null hypothesis?

(c) Write a sentence stating your conclusion of the hypothesis test you just performed. (i.e. Write your conclusion in plain language so that a non-statistician could understand the conclusion of the test.)

(d) Suppose that in another manufacturing center in 529 items made by the same manufacturing process as above, it is found that 27 items are defective. Calculate a 95% confidence interval (i.e. a two-sided confidence interval) for the true proportion of defective items manufactured at this plant. Copy and paste the appropriate R command and output you would use to do this question. In an additional line, clearly state the values you get for the confidence interval from your output.

```r
binom.test(x=27, n=529, p=0.08, alternative="two.sided")
```

```
##
##  Exact binomial test
##
## data:  27 and 529
## number of successes = 27, number of trials = 529, p-value = 0.01269
## alternative hypothesis: true probability of success is not equal to 0.08
## 95 percent confidence interval:
##  0.03390134 0.07339402
## sample estimates:
## probability of success
##              0.0510397
```

(e) According to your confidence interval in part (d), is it reasonable to conclude that 8% of the items created at this plant are defective? Explain, using your confidence interval.