Statistical Methods - Assignment 3

Michel Mooiweer (1866761) Thomas Webbers (2560695) Eirik Kultorp (2544992)

December 2016

Theoretical exercises

3.1

$$n = \frac{[z_{\alpha/2}]^2 \hat{p}\hat{q}}{E^2}$$

Sample Proportion $\hat{p} = 0.15$

$1 - \hat{p} = \hat{q} = 0.85$

Confidence level $\alpha = 0.05$

$\alpha/2 = 0.025$

Critical value: $z_{\alpha/2} = 1.96$

$E = 0.03$

$$n = \frac{1.96^2 \cdot 0.15 \cdot 0.85}{0.03^2}$$
$$n = 544.226 \rightarrow 545$$

Sample size needed is 545

3.2

Claim males (12-16) correctly report their height:

- Population parameter: difference population means
- $H_0 : \mu_1 = \mu_2$
- $H_a : \mu_1 \neq \mu_2$ -Significance level: $\alpha = 0.1$
- Test statistic:

$$T_d = \frac{\bar{D} - (\mu_1 - \mu_2)}{\frac{S_d}{\sqrt{n}}}$$

  which has a t-distribution with n-1 degrees of freedom under the null hypothesis
- Observed value: $\bar{d} = -1$

with $\mu_1 - \mu_2 = 0$

with $S_d = 3.520$

with n = 12

$$t_d = \frac{-1 - 0}{\frac{3.520}{\sqrt{12}}}$$

$$t_d = -0.984$$

From Table 3-A Critical value (2-tailed $\alpha = 0.1$) $= 1.796$

-0.984 is between -1.796 and 1.796 Thus, we can't reject the null hypotheses. Therefore, there is not enough evidence to say males(12-16) do not correctly report their height.

R-Exercises

Exercise 3.3

a)

```
## [1] "Point estimate :  0.2024 (Alice works 0.2024 hours more than Bob)"
## [1] "Confidence interval (0.95) : -0.057 : 0.462"
```

b)

We do a t-test because we're testing for difference between population means. The conditions are also met:

- The sampling method for each sample is simple random sampling
- The samples are independent.
- Each population is at least 20 times larger than its respective sample. (We don't know this, but we assume it)
- The sampling distribution is approximately normal and the sample size is greater than 40, without outliers.

Hypothesis test report

- Null hypothesis: True difference in means is equal to 0

- Alternative hypothesis: NOT True difference in means is equal to 0

- Significance level: 0.05

- Test statistic: $-Teq = \frac{((\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2))}{\sqrt[2]{S_p^2/n_1 + S_p^2/n_2}}$

  with $S_p^2 = \frac{(n_1-1)*S_2^1 + (n_2-1)*S_2^2}{n_1+n_2-2}$

- Observed value: 0.202

- p-value: 0.125

We reject the null hypothesis because p<=0.05.In other words, we can reject the claim that Alice and Bob work the same amount of hours. We could also have reached the same conclusion just by looking at the confidence interval, which

c)

We do a t-test for the same reasons as above. This time we test the claim that Alice works less than Bob

Hypothesis test report

- Null hypothesis: True difference in means is less than or equal to 0

- Alternative hypothesis: NOT True difference in means is less than or equal to 0

- Significance level: 0.05

- Test statistic: $Teq = \frac{((\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2))}{\sqrt[2]{S_p^2/n_1 + S_p^2/n_2}}$

  with $S_p^2 = \frac{(n_1-1)*S_2^1 + (n_2-1)*S_2^2}{n_1+n_2-2}$

- Observed value: 0.202

- p-value: 0.062

We reject the null hypothesis because p<=0.05.In other words, we can reject the claim that Alice works more than Bob. We could also have reached the same conclusion just by looking at the confidence interval, which

d)

`0.1247/0.06235 == 2`

The p-values represent the probability of obtaining something at least as extreme as the observed value if the null hypothesis is true. We performed two tests on the same data, `u!=0` and `u>=0`, a one-tailed and two-tailed test. Because the tails are of equal area, the two-tailed test has twice as much area outside the confidence interval.

3.4

a)

`## [1] 0.16`

b)

We perform a test for population proportions.

Hypothesis test report

- Null hypothesis: True difference in proportions is less than or equal to 0

- Alternative hypothesis: NOT True difference in proportions is less than or equal to 0

- Significance level: 0.05

- Test statistic: $Z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$

- Observed value: 0.16

- p-value: 0.059

We reject the null hypothesis because p<=0.05.In other words, we can reject the claim that The proportion of evenings on which Alice worked more than 3.5 hours is larger than the proportion of evenings on which Bob worked more than 3.5 hours. We could also have reached the same conclusion just by looking at the confidence interval, which
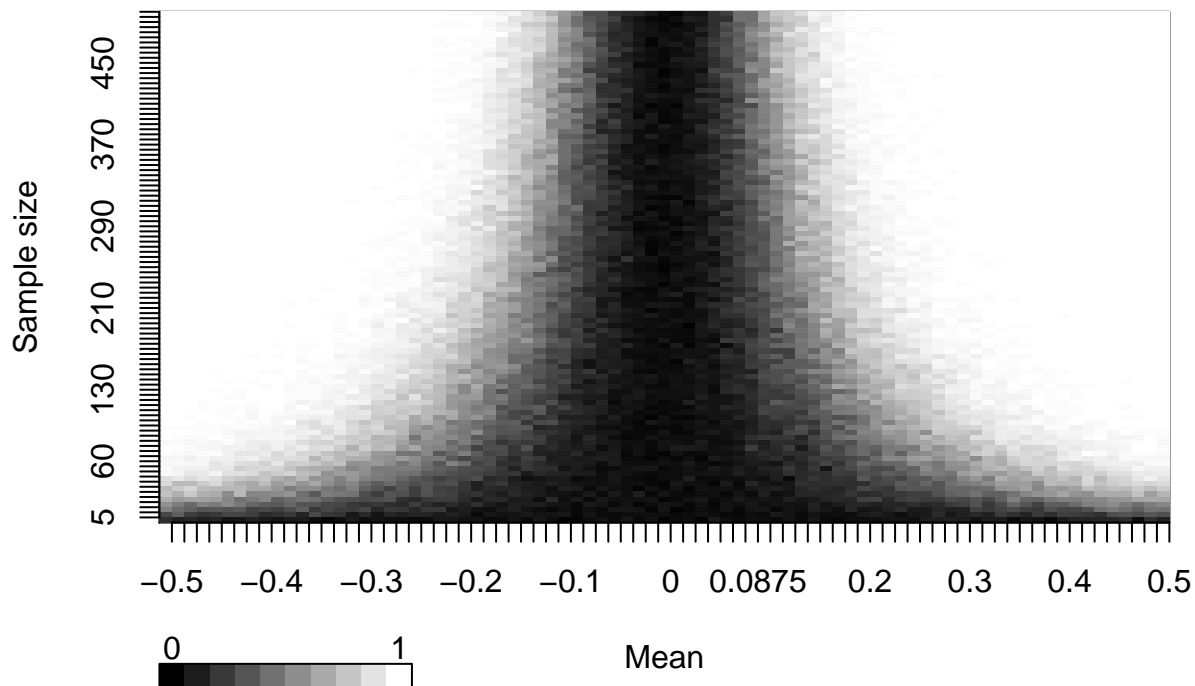
3.5

a,b,c)

```
##        [,1] [,2] [,3] [,4]
## [1,] 0.25 0.11 0.10 0.86
## [2,] 0.77 0.16 0.31 1.00
## [3,] 0.95 0.24 0.55 1.00
## [4,] 1.00 0.54 0.83 1.00
```

d,e)

We repeat for many more values and visualize it, so as to see patterns more clearly

In the above graphic, lighter cells indicate a higher proportions of rejected hypotheses. Since we know that the null hypothesis really does hold, this translates into the proportion of type II errors (false negative). It is difficult to see in the above graphic, but as the sample size increases the proportion should approach 0.95 for the samples of mean=0, because the significance level is 0.05.

Appendix

printhypotest

```
printhypotest <- function(res,test_statistic,null_hypothesis,claim){

  SIGNIFICANCE_LVL=0.05

  mycat <- function(text){return(cat(gsub(pattern = "\n", replacement = "  \n", x = text)))}

  s=paste("#### Hypothesis test report\n\n - Null hypothesis:",null_hypothesis,
        "\n - Alternative hypothesis:","NOT",null_hypothesis,
        "\n - Significance level:",SIGNIFICANCE_LVL,
        "\n - Test statistic:",test_statistic,
        "\n - Observed value:",round(res$estimate[1]-res$estimate[2],3),
        "\n - p-value:",round(res$p.value,3))

  a = "We could also have reached the same conclusion just by looking at the confidence interval, which
  fir <- if (res$p.value<0.05) paste0("\n\nWe fail to reject the null hypothesis because p>",SIGNIFICANC
  return(mycat(paste(s,fir,a)))
}
```

3.3.a

```
load("Ass3.RData")
res=t.test(Alice,Bob)
pe = res[['estimate']][1]-res[['estimate']][2]
pe = round(pe,4)
print(paste('Point estimate : ',pe,'(Alice works',pe,'hours more than Bob)'))
print(paste('Confidence interval (0.95) :',round(res[['conf.int']][1],3) , ':' , round(res[['conf.int']]
```

3.3.b

```
printhypotest(res,"-$Teq = \\frac{((\\bar{x}_1-\\bar{x}_2)-(\\mu_1-\\mu_2))}{\\sqrt[2]{S_p^2/n_1 + S_p^
```

3.4.a

```
(sum(Alice > 3.5)/length(Alice))-(sum(Bob > 3.5)/length(Bob))
```

3.4.b

```
d <- matrix(c(sum(Alice>3.5),sum(Bob>3.5),length(Alice)-sum(Alice>3.5),length(Bob)-sum(Bob>3.5)),nrow=2
colnames(d) <- c("T","F")
rownames(d) <- c("A","B")
d <- as.table(d)
res = prop.test(d,alternative="greater")
printhypotest(res,"$Z=\\dfrac{(\\hat{p}_1-\\hat{p}_2)-(p_1-p_2)}{\\sqrt{\\hat{p}(1-\\hat{p})\\left(\\df
```

tte

```r
tte <- function(means,sample_sizes,repetitions){

  m = c()

  for (mean in means){
    for (sample_size in sample_sizes){
      num_rejects = 0
      for (i in seq(1,100)){
        x = rnorm(sample_size,mean,1)
        t = t.test(x)
        if (t$p.value<=0.05) num_rejects=num_rejects+1
      }
      m[length(m)+1] = num_rejects/repetitions
    }
  }

  return(matrix(m,nrow=length(sample_sizes),ncol=length(means)))
}
```

3.5.a,b,c

```r
tte(c(-0.5, -0.2, 0.3, 1),c(10, 25, 50, 100),100)
```

3.5.d,e

```r
means = seq(-0.5,0.5,0.0125)
sample_sizes = rev(seq(5,500,5))

r = tte(means, sample_sizes,100)

library(plotrix)

color2D.matplot(
  r,
  show.legend=TRUE,
  xlab="Mean",
  ylab="Sample size",
  axes=FALSE,
  border=NA
)
axis(1,at=1:length(means),labels=means)
axis(2,at=length(sample_sizes):1,labels=sample_sizes)
```