Math 170S Homework #7

Damien Ha

2023-03-04

Contents

Problem 1

Part A

```r
sample1 <- c(20, 25, 18, 27, 32, 39, 21, 26, 19, 28, 33, 38, 22, 27, 23, 35, 31)
sample2 <- c(39, 21, 23, 24, 29, 28, 27, 29, 31, 33, 36, 37)

t.test(sample1, sample2, var.equal = FALSE, conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  sample1 and sample2
## t = -1.0765, df = 25.685, p-value = 0.2917
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -7.148029  2.236264
## sample estimates:
## mean of x mean of y
##  27.29412  29.75000
```

At a level of $\alpha = 0.05$, the p-value is 0.2917 which is greater than $\alpha = 0.05$, so we fail to reject the null hypothesis. There is not enough evidence to conclude that the two samples are not equal.

Part B

```r
sample1 <- c(25, 31, 22.5, 33.5, 40, 48.7, 26.25, 32.5, 23.5, 35, 41.25, 47.5,
             27.5, 33.7, 28.75, 43.7, 39, 48, 54, 56)
sample2 <- c(48, 26, 28.7, 30, 36.2, 35, 33.7, 36.2, 38.7, 41.5, 45, 46.5, 25, 26, 27)

t.test(sample1, sample2, var.equal = FALSE, conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  sample1 and sample2
## t = 0.64556, df = 32.978, p-value = 0.523
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.233373  8.168373
## sample estimates:
## mean of x mean of y
##   36.8675   34.9000
```

At a level of $\alpha = 0.05$, the p-value is 0.523 which is greater than $\alpha = 0.05$, so we fail to reject the null hypothesis. There is not enough evidence to conclude that the two samples are not equal.

Part C

```r
sample1 <- c(209, 277, 221, 215, 247.7, 253.7, 218, 212, 244.7, 250.7, 215.0, 209)
sample2 <- c(265, 201, 228, 289, 277, 289, 288, 278, 277, 265, 264, 266, 298, 291, 270)

t.test(sample1, sample2, var.equal = FALSE, conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  sample1 and sample2
## t = -4.1809, df = 24.69, p-value = 0.0003175
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -57.72604 -19.60730
## sample estimates:
## mean of x mean of y
##  231.0667  269.7333
```

Our p-value is 0.0003175. At a level of $\alpha = 0.05$, this is significant. In addition, our confidence interval does not include 0, suggesting that the true difference between the samples is not 0. Because of this, we reject the null hypothesis and conclude that the samples are not equal.

Problem 2

Part A

```r
sample1 <- c(55, 56, 59, 51, 82, 88, 72, 84, 71, 80, 66, 67, 52, 81, 57, 72, 69)
sample2 <- c(44, 43, 48, 56, 54, 53, 52, 41, 40, 43, 43, 44, 45, 45, 52, 51, 50)

t.test(sample1, sample2, paired = TRUE, alternative = "greater", conf.level = 0.99)
```

```
##
##  Paired t-test
##
## data:  sample1 and sample2
## t = 6.6656, df = 16, p-value = 2.713e-06
## alternative hypothesis: true difference in means is greater than 0
## 99 percent confidence interval:
##  12.89672      Inf
## sample estimates:
## mean of the differences
##                21.05882
```

The p-value is 0.000002713 which is less than $\alpha = 0.01$, so we reject the null hypothesis. The mean of sample 1 is greater than that of sample 2 at the 99% confidence level.

Part B

```r
sample1 <- c(140, 142, 150, 130.1, 209, 224, 183, 214, 181, 204, 168, 170, 133,
             206, 145, 183, 179)
sample2 <- c(105, 102, 115, 135.8, 130, 128, 125, 97, 95, 102, 102, 105, 108,
             107, 125, 123, 120)


t.test(sample1, sample2, paired = TRUE, alternative = "greater", conf.level = 0.99)
```

```
##
##  Paired t-test
##
## data:  sample1 and sample2
## t = 7.5548, df = 16, p-value = 5.78e-07
## alternative hypothesis: true difference in means is greater than 0
## 99 percent confidence interval:
##  40.11309      Inf
## sample estimates:
## mean of the differences
##                60.95882
```

The p-value is 0.000000578 which is less than $\alpha = 0.01$, so we reject the null hypothesis. The mean of sample 1 is greater than that of sample 2 at the 99% confidence level.

Part C

```r
sample1 <- c(53.4, 52.88, 58.13, 62.27, 70.9, 72.5, 65.7, 60.7, 55, 60, 56, 57, 53, 62)
sample2 <- c(61, 63, 71, 55, 53, 59, 63, 61, 55, 58, 63, 58, 57, 56)

t.test(sample1, sample2, paired = TRUE, alternative = "greater", conf.level = 0.99)
```

```
##
##  Paired t-test
```

```
##
## data:  sample1 and sample2
## t = 0.19829, df = 13, p-value = 0.4229
## alternative hypothesis: true difference in means is greater than 0
## 99 percent confidence interval:
##  -5.723743       Inf
## sample estimates:
## mean of the differences
##               0.4628571
```

The p-value is 0.4229 which is greater than $\alpha = 0.01$, so we fail to reject the null hypothesis. There is not enough evidence to suggest te mean of sample 1 is greater than sample 2 at a 99% considence level.

Problem 3

Part A

```r
sample1 <- c(55, 56, 59, 51, 82, 88, 72, 84, 71, 80, 66, 67, 52, 81, 57, 72, 69)
sample2 <- c(44, 43, 48, 56, 54, 53, 52, 41, 40, 43, 43, 44, 45, 45, 52, 51, 50)

wilcox.test(sample1, sample2, alternative = "greater", conf.level = 0.99)
```

```
## Warning in wilcox.test.default(sample1, sample2, alternative = "greater", :
## cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  sample1 and sample2
## W = 278, p-value = 2.264e-06
## alternative hypothesis: true location shift is greater than 0
```

The p-value is 0.000002264, which is less than $\alpha = 0.01$, so we reject the null hypothesis. The mean of sample 1 is greater than that of sample 2 at the 99% confidence level.

Part B

```r
sample1 <- c(140, 142, 150, 130.1, 209, 224, 183, 214, 181, 204, 168, 170, 133,
             206, 145, 183, 179)
sample2 <- c(105, 102, 115, 135.8, 130, 128, 125, 97, 95, 102, 102, 105, 108,
             107, 125, 123, 120)

wilcox.test(sample1, sample2, alternative = "greater", conf.level = 0.99)
```

```
## Warning in wilcox.test.default(sample1, sample2, alternative = "greater", :
## cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  sample1 and sample2
## W = 287, p-value = 4.95e-07
## alternative hypothesis: true location shift is greater than 0
```

The p-value is 0.000000495, which is less than $\alpha = 0.01$, so we reject the null hypothesis. The mean of sample 1 is greater than that of sample 2 at the 99% confidence level.

Part C

```r
sample1 <- c(53.4, 52.88, 58.13, 62.27, 70.9, 72.5, 65.7, 60.7, 55, 60, 56, 57, 53, 62)
sample2 <- c(61, 63, 71, 55, 53, 59, 63, 61, 55, 58, 63, 58, 57, 56)

wilcox.test(sample1, sample2, alternative = "greater", conf.level = 0.99)
```

```
## Warning in wilcox.test.default(sample1, sample2, alternative = "greater", :
## cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  sample1 and sample2
## W = 94.5, p-value = 0.573
## alternative hypothesis: true location shift is greater than 0
```

The p-value is 0.573 which is greater than $\alpha = 0.01$, so we fail to reject the null hypothesis. There is not enough evidence to suggest te mean of sample 1 is greater than sample 2 at a 99% considence level.

Problem 4

Part A

```r
table <- matrix(c(34, 47, 55, 49, 39, 41, 52, 42, 32, 44, 48, 45),
                nrow = 3, ncol = 4, byrow = TRUE)
rownames(table) <- c("M1", "M2", "M3")
colnames(table) <- c("A", "B", "C", "D")

chisq.test(table)
```

```
##
##  Pearson's Chi-squared test
##
## data:  table
## X-squared = 1.4185, df = 6, p-value = 0.9647
```

The p-value is 0.9647 which is greater than $\alpha = 0.05$, so we fail to reject the null hypothesis. There isn't enough evidence to suggest that there is association between the variables.

Part B

```r
table <- matrix(c(123, 156, 228, 109, 144, 201, 111, 148, 212),
                nrow = 3, ncol = 3, byrow = TRUE)
rownames(table) <- c("M1", "M2", "M3")
colnames(table) <- c("A", "B", "C")

chisq.test(table)
```

```
##
##  Pearson's Chi-squared test
##
## data:  table
## X-squared = 0.15874, df = 4, p-value = 0.997
```

The p-value is 0.997 which is greater than $\alpha = 0.05$, so we fail to reject the null hypothesis. There isn't enough evidence to suggest that there is association between the variables.

Part C

```r
table <- matrix(c(405, 656, 556, 655, 310, 608, 534, 387, 409, 626, 534, 456,
                  400, 672, 443, 542),
                nrow = 4, ncol = 4, byrow = TRUE)
rownames(table) <- c("M1", "M2", "M3", "M4")
colnames(table) <- c("A", "B", "C", "D")

chisq.test(table)
```

```
##
##  Pearson's Chi-squared test
##
## data:  table
## X-squared = 69.405, df = 9, p-value = 1.991e-11
```

The p-value is 0.00000000001991 which is less than $\alpha = 0.05$, so we reject the null hypothesis. At a 95% confidence interval, there is evidence to suggest that the variables are associated.