Assignment_3

Jaimie Chin

2023-04-04

Contents

```
# Load packages
library(tidyverse)
```

1 Question 1 (10 points)

A researcher wants to know if people on vacation, engage in an "inner dialogue" less than when working. The researcher starts by obtaining a sample of 10 individuals who are about to go on a week's vacation and agree to note (on an app) each time they "hear" themselves mentally talking. Each person in the sample is asked to keep a log for the week. The daily average instances (based on the week) appears below.

```
# Instantiate dataset
sample1 = c(50, 40, 46, 49, 40, 58, 45, 47, 46, 43)
```

1.1 1. Complete a one-sample t-test where the population mean is 50. Make sure to select correctly between a one-sided and two-sided t-test given the hypothesis described in the prompt. (In R, you can specify `alternative = 'greater'` or `alternative = 'less'`to run a one-sided t-test. It runs a two-sided t-test by default). (2 pts)

```
# Running a one-sample
t.test(sample1, mu = 50, alternative='less')
```

```
##
##  One Sample t-test
##
## data:  sample1
## t = -2.1583, df = 9, p-value = 0.02962
## alternative hypothesis: true mean is less than 50
## 95 percent confidence interval:
##      -Inf 49.45763
## sample estimates:
## mean of x
##      46.4
```

1.2 2. What is the t-value? (1 pt)

The t-value is -2.1583.

1.3   3. What is the p-value? (1 pt)

The p-value is 0.02962

1.4   4. What is your interpretation of the test? (2 pts)

With a p-value of 0.02962, we can reject the null hypothesis that the true population mean is 50 at the 5% significance level. This means that we have enough evidence to conclude that people on vacation engage in an "inner dialogue" less than when working.

Furthermore, the t-value is a statistic that measures the difference between the sample mean and the null hypothesis mean (in this case, 50) in units of standard error. In this case, the t-value of -2.1583 tells us that the sample mean is 2.1583 standard errors below the null hypothesis mean of 50.

1.5   5. In Part 1, you determined that this was a one-sided or two-sided t-test. How could this experiment be adjusted so that you would run the other type of t-test? Specifically, what would the hypothesis be? (2 pts)

In Part 1, a one-sided one-sample t-test was used to compare the group mean (10 individuals on vacation) to a standard value (50). This is because the researcher wants to know if people on vacation, engage in an "inner dialogue" *less* than when working. A one-sided test is appropriate in this case because a one-tailed test looks for an "increase" or "decrease" in the parameter whereas a two-tailed test looks for a "change" (could be increase or decrease) in the parameter.

The experiment can be adjusted to run a two-sided t-test by adjusting the hypothesis to consider whether there is any difference between the daily average of "inner dialogue" instances of participants on vacation than when working. The null hypothesis would be that the mean number of instances of "inner dialogue" is the same for those on vacation and when working.

1.6   6. Run the other type of t-test. How do the test values and interpretation of the test change? (2pts)

```
# Running a two-sample t-test
t.test(sample1, mu = 50, alternative = "two.sided")
```

```
##
##  One Sample t-test
##
## data:  sample1
## t = -2.1583, df = 9, p-value = 0.05923
## alternative hypothesis: true mean is not equal to 50
## 95 percent confidence interval:
##  42.62672 50.17328
## sample estimates:
## mean of x
##      46.4
```

The t-value is the same as before (-2.1583), but the p-value is now 0.05923. With a p-value greater than 0.05, we fail to reject the null hypothesis that there is a difference in mean between those on vacation and when working. This means that we do not have enough evidence to conclude that people on vacation engage in "inner dialogue" differently than when working.

2 Question 2 (15 points)

From the experiment in (2), the researcher also obtains data from a second sample (of the same size) from individuals during a regular week of work. The daily average instances (based on a week of data) of inner dialogue appear below. The researcher wants to know if people on vacation engage in an "inner dialogue" less than when working.

```
# Instantiate dataset
sample2 <- c(53, 40, 51, 50, 43, 62, 49, 47, 51, 39)
```

2.1 1. Is the assumption of homogeneity met? Show how you determined this. (2 pts)

```
# Create dataframe with both groups
data = data.frame(inner_dialogue = c(sample1, sample2),
                  group = c("vacation", "vacation", "vacation", "vacation", "vacation", "vacation",
                            "vacation", "vacation", "vacation", "vacation", "work", "work",
                            "work", "work", "work", "work", "work", "work", "work", "work"))

# Compare variances of both groups
data %>%
  group_by(group) %>%
  summarize(variance = var(inner_dialogue))
```

```
## # A tibble: 2 x 2
##   group    variance
##   <chr>       <dbl>
## 1 vacation     27.8
## 2 work         45.8
```

Homogeneity of variance means that the two groups have the same amount of variability within them. That is, the data in one group has roughly the same amount of spread as the data in the other group, regardless of what the group means are. We can check if our data meets the assumption o homogeneity if the ratio of variances is under 3. If the ratio is under 3, we should run a student's t-test (compared to a Welch's t-test).

$$\frac{45.83333}{27.82222} = 1.647364229$$

Because the ratio of variances is below 3, our data meets the assumption of homogeneity and we can conduct a student's t-test.

2.2 2. Complete a two-sample independent t-test of your first dataset against the second. Make sure to run the appropriate type of test depending on your answer in Part 1. Also make sure to select correctly between a one-sided and two-sided t-test given the hypothesis described in the prompt. (In R, you can specify `alternative = 'greater'` or `alternative = 'less'`to run a one-sided t-test. It runs a two-sided t-test by default). (3 pts)

```
# Running a two-sample t-test
t.test(inner_dialogue ~ group, data=data, alternative = "less", var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  inner_dialogue by group
## t = -0.77378, df = 18, p-value = 0.2246
## alternative hypothesis: true difference in means between group vacation and group work is less than 0
## 95 percent confidence interval:
##      -Inf 2.606172
## sample estimates:
## mean in group vacation     mean in group work
##                   46.4                   48.5
```

2.3   3. What is the t-value? (1 pt)

The t-value is -0.77378.

2.4   4. What is the p-value? (1 pt)

The p-value is 0.2246.

2.5   5. What is your interpretation of the test? (2 pts)

Since the p-value of 0.2246 is larger than the common significance level of 0.05, we do not have sufficient evidence to reject the null hypothesis. This means we cannot conclude that there is a statistically significant difference in the mean number of instances of "inner dialogue" between vacation and work. Therefore, based on this analysis, we do not have evidence to support the hypothesis that people engage in *less* "inner dialogue" during vacation compared to when working.

2.6   6. In Part 1, you determined that this was a one-sided or two-sided t-test. How could this experiment be adjusted so that you would run the other type of t-test? Specifically, what would the hypothesis be? (2 pts)

In Part 1, a one-sided two-sample t-test was used to compare the group mean of 10 individuals on vacation to the group mean of 10 individuals during work. This is because the researcher wants to know if people on vacation, engage in an "inner dialogue" *less* than people who were working. A one-tailed t-test is appropriate in this case because a one-tailed test looks for an "increase" or "decrease" in the parameter whereas a two-tailed test looks for a "change" (could be increase or decrease) in the parameter.

The experiment can be adjusted to run a two-sided t-test by adjusting the hypothesis to consider whether there is any difference between the daily average of "inner dialogue" instances of participants on vacation compared to those who were working. The null hypothesis would be that the mean number of instances of "inner dialogue" is the same for those on vacation and those who were working.

2.7 7. Run the other type of t-test described in Part 6. How do the test values and interpretation of the test change? (2 pts)

```
# Running a two-sample t-test
t.test(inner_dialogue ~ group, data=data, alternative = "two.sided", var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  inner_dialogue by group
## t = -0.77378, df = 18, p-value = 0.4491
## alternative hypothesis: true difference in means between group vacation and group work is not equal
## 95 percent confidence interval:
##  -7.801809  3.601809
## sample estimates:
## mean in group vacation     mean in group work
##                   46.4                   48.5
```

The t-value is the same as before (-0.77378), but the p-value is now 0.4491, which is even higher than the p-value from one-sided two sample t-test. With a p-value greater than 0.05, we still fail to reject the null hypothesis that there is a difference in mean between those on vacation and when working. This means that we do not have enough evidence to conclude that people on vacation engage in "inner dialogue" differently compared to people enaging in "inner dialogue" when working.

2.8 8. In psychological experiments, we typically run two-sided t-tests even when we have a one sided hypothesis. E.g., I expect that an intervention would improve performance. Why is it standard that we run two-sided tests? That is, what do we gain by running a two-sided test over a one-sided one? (2 pts)

Running a two-sided t-test provides a more comprehensive analysis of the data as it tests for the possibility of both positive and negative effects of the intervention or treatment being studied. It is possible that an intervention could have a negative effect on performance, even if we expect it to improve performance. Therefore, by running a two-sided test, we are checking for both the possibility of an improvement and the possibility of a decrement in performance.

Running a one-sided test, on the other hand, assumes that the effect can only be in one direction and ignores the possibility of an opposite effect. While a one-sided test may be appropriate in certain cases where the direction of the effect is well-established or there is a theoretical reason to believe that the effect can only be in one direction, a two-sided test is generally considered more robust and less biased.

Moreover, it is worth noting that when running a one-sided test, the p-value is halved compared to a two-sided test. This is because in a one-sided test, all of the statistical significance is attributed to one direction. Therefore, running a two-sided test provides a more conservative measure of statistical significance.

3 Question 3. (10. points)

Redo the t-test above, but instead of a independent-samples t-test, compute the t-test as a paired samples t-test. Assume that the same people are measured during vacation and then again at a later time during a workweek.

3.1　1. Run the t-test. Run it as a two-sided t-test. (2 pts)

```
# Run paired t-test
t.test(inner_dialogue ~ group, data=data, paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  inner_dialogue by group
## t = -2.3333, df = 9, p-value = 0.0445
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -4.13594145 -0.06405855
## sample estimates:
## mean difference
##            -2.1
```

3.2　2. What is the t-value? (1 pt)

The t-value is -2.3333

3.3　3. What is the p-value? (1 pt)

The p-value is 0.0445

3.4　4. What is your interpretation of the test? (2 pts)

The p-value of 0.0445 indicates that there is a significant difference between the mean "inner dialogue" count during vacation and the mean count during a workweek. We can reject the null hypothesis and conclude that there is evidence to support the alternative hypothesis, which states that there is a significant difference in the means of inner dialogue of participants on vacation compared to when they are at work. The negative t-value of -2.3333 indicates that the mean count during vacation is lower than the mean count during a workweek.

3.5　5. Compared to the two-sided independent-samples t-test from Question 2, what changed? (2 pts)

Both the test statistic and p-value have changed from the independent-samples t-test in Question 2. This is because in the paired-samples t-test, we are comparing the same individuals under different conditions (i.e., during vacation and during a workweek), whereas in the independent-samples t-test, we are comparing two different groups of individuals (i.e., those on vacation and those during a workweek).

3.6　6. What explains the change you noted in Part 5? (2 pts)

The change in test statistic and p-value can be explained by the fact that the paired-samples t-test has more statistical power than the independent-samples t-test, since it eliminates individual differences and reduces error variance (reduces inter-subject variability). This makes it more sensitive to detecting a significant difference between the means of the two conditions.