

1.7: T-Test

Introduction to T-Test

In the previous chapter, I demonstrated that the chi-squared test is employed when examining whether a significant association exists between two categorical variables. For instance, it can be used to test if there is a correlation between the use of credit cards (yes = 1 or no = 0) for online purchases and gender (male = 1 or female = 0). However, what if one of the variables we want to examine is continuous data, not categorical? For instance, suppose we want to investigate whether there is a difference in the number of items purchased online in the past 12 months between men and women. A chi-squared test will not be the most effective statistical tool here. Instead, the t-test is a statistical significance test frequently utilized to evaluate the difference between two group means, such as the average score on antisocial attitudes among inmates who participated in cognitive behavioral therapy (CBT) and those who did not. For example, even if inmates who completed CBT have lower antisocial attitudes than those who did not, it is a different question if there are two main types of t-tests:

- Independent-samples t-test: This test is employed when comparing the mean scores of two distinct groups of individuals or conditions.
- Paired-samples t-test: This test is used when assessing the mean scores within the same group of individuals on two separate occasions, or when dealing with matched pairs of data.

In this chapter, we will use two fictitious datasets to conduct these two types of t-tests.

Cognitive Behavioral Therapy

Those who study crime and criminals should pay great attention to the issue of rehabilitation. A significant portion of inmates eventually reintegrate into society, where they play crucial roles as members of our communities. A variety of rehabilitation programs have been developed to alter the belief systems and behaviors of offenders. Among the most prevalent programs implemented in correctional settings is CBT. CBT programs operate under the premise that holding antisocial attitudes correlates with increased engagement in antisocial behavior. Specifically, these programs posit that exposure to high-risk situations triggers antisocial thoughts and emotions, thereby heightening the probability of engaging in antisocial behavior (Vaske et al., 2011). CBT programs that target criminal activity have been used as popular interventions in correctional settings because many evaluation studies have shown that CBT reduces recidivism significantly (Landenberger & Lipsey, 2005; Lipsey et al., 2007; Zara, 2019).

Let's consider a scenario where high-risk inmates were selected at random to participate in a CBT program. This program was conducted in a group setting within the prison, with trained facilitators leading individual sessions aimed at tackling the diverse challenges offenders encounter upon their return to the community. These challenges, spanning from addiction and employment to family obligations and victimization, frequently impede successful reintegration. The program consists of 10 sessions spread over a period of 10 weeks, totaling approximately 20 hours of engagement.

How do we assess the effectiveness of this CBT program? Initially, we must define the criteria by which we will evaluate its success. Those responsible for implementing the program must first clarify its objectives. For instance, the program might aim to modify the attitudes, behaviors, or both, of inmates. For instance, if our focus is on behavioral change, we could monitor participants' recidivism rates, such as rearrest or reconviction. Alternatively, we could concentrate on shifts in participants' attitudes. Another critical consideration is the research design. We could compare participants in the program to those who did not participate, or within-group changes could be analyzed by comparing participants' attitudes before and after the program.

Independent-Samples T-Test

First, we will examine a scenario where high-risk inmates were randomly chosen to participate in a CBT program. In this context, we will compare the antisocial attitudes of those who were randomly selected for the program with those who were not. For our analysis, let's assume that antisocial attitudes were assessed using a scale developed by Farrington and McGee (2017). This scale comprises a 24-item self-reported instrument with a 4-point response format, including statements like "If someone does the dirty on me, I always try to get my own back" or "I enjoy watching people getting beaten up on TV." The higher antisocial attitudes score reflects higher levels of antisocial attitudes (e.g., aggressiveness).

We will use the fictitious dataset I constructed to contrast inmates who engaged in a CBT program with those who did not. In this hypothetical study, there were 100 high-risk inmates in the prison. Fifty inmates were randomly allocated to participate in a 10-week CBT program, while the other 50 were randomly assigned not to participate. If the program yielded an impact, we anticipate

observing reduced levels of antisocial attitudes among those who participated compared to their counterparts. Let's first download the data from [the shared Google Drive folder containing the CBT_dataset_independent.csv data](#). Then, load the data for this study using the syntax below. Each row in this data represents each inmate. The group variable is a categorical variable where the category "Participants" refers to those who participated in a CBT program, whereas "Non-Participants" refers to those who did not. The "AntiSocial" variable represents inmates' antisocial attitudes.

```
library(readr)

CBT_dataset_independent <- read_csv("C:/Users/75JCHOI/OneDrive -
West Chester University of PA/WCU Research/R/data/
CBT_dataset_independent.csv")

View(CBT_dataset_independent)
```

EExecuting `library(readr)` allows me to load the `readr` package into my R session. The `readr` package, part of the tidyverse collection, is specifically crafted to simplify importing flat-file data, such as CSVs and text files, into R. You may already be familiar with the next step, but it is worth noting that "`read_csv`" is a handy function from the `readr` package tailored for reading CSV files.

NHST Steps for Independent-Samples T-Test

Following the NHST Steps we covered in the previous chapter, we will conduct the independent-samples t-test because we are comparing the mean scores of two distinct groups of individuals (i.e., CBT participants and non-participants).

Step 1: Formulate the null and alternative hypotheses.

- H0: There is no difference in mean antisocial attitudes between CBT participants and non-participants.
- HA: There is a difference in mean antisocial attitudes between CBT participants and non-participants.

Step 2: Calculate the test statistic.

The following code facilitates the execution of an independent-samples t-test. In R, a formula typically comprises a single variable on the left, denoted by a \sim (tilde), followed by one or more predictors on the right, which help predict the variable on the left. In statistical tests, the variable on the left of the formula represents the dependent variable (e.g., antisocial attitudes), while those on the right represent the independent variables (e.g., CBT program participation).

```
twosampt <- t.test(formula = CBT_dataset_independent$AntiSocial ~
  CBT_dataset_independent$Group)

twosampt
```

The output should generate results from Welch's t-test. Welch's t-test differs slightly from the traditional t-test formula, primarily used when the data deviates from the assumption of equal variances. The `t.test()` output shows a t-statistic of 4.859. This t-value is well above a t-value of 1.96, which is the historical cut-off point for significance at the 95% confidence level.

Step 3: Determine the probability (P-Value) of Obtaining a Test Statistic at Least as Extreme as the Observed Value, Assuming no Relationship Exists.

The p-value in this output was shown in scientific notation as 4.485e-06. You may be confused because you do not know how to convert this number. Here is the way to convert p-value to a regular numeric format.

```
# Your p-value in scientific notation

p_value <- 4.485e-06

# Convert p-value to regular numeric format

formatted_p_value <- format(p_value, scientific = FALSE)

# Print the p-value

formatted_p_value
```

Now you know that the probability of obtaining a t-statistic of 4.859 is 0.000004485 if the null hypothesis is true, which is substantially smaller than .05, the conventionally used standard. R produces a regular numeric p-value, but when the p-value is very low, it may display scientific notation.

Steps 4 & 5: If the p-value is Very Small, Typically Less Than

5%, Reject the Null Hypothesis, but if the P-Value is Not Small, Typically 5% or Greater, Retain the Null Hypothesis. The t-statistic fell within the rejection region. The probability of this sample coming from a population where the mean antisocial attitudes for both CBT participants and non-participants are equal is exceedingly low. Therefore, it is probable that the sample is from a population where CBT participants and non-participants exhibit statistically significant different mean antisocial attitudes.

Reporting the Results of an Independent-Samples T-Test

The results from our independent-samples t-test can be presented as follows: An independent-sample t-test was conducted to compare the antisocial attitudes scores for CBT participants and non-participants. There was a significant difference in scores for participants ($M = 46.46$) and non-participants ($M = 55.39$; $t(97.924) = 4.86$, $p < .05$, two-tailed). If you want to see how t-test results are reported in an academic peer-reviewed journal, please see Choi et al. (2020). My colleagues and I compared levels of fear of crime between women and men using independent-samples t-tests.

Density Plot

You may wonder if there is a way to visualize the difference in antisocial attitudes between CBT participants and non-participants. Density plot is a useful tool that graphically represents a distribution of scores or values that take the form of a smooth curve. You can create a density plot using ggplot2 and tidyverse to visualize the distribution of antisocial attitudes between two groups (presumably CBT participants and non-participants) from the dataset "CBT_dataset_independent".

```
library(ggplot2)
library(tidyverse)

dens_cbt <- CBT_dataset_independent %>%
  ggplot(aes(x = AntiSocial,
    fill = Group)) +
  geom_density(alpha = .7) +
  theme_minimal() +
  labs(x = "Antisocial Attitudes", y = "Probability Density") +
  scale_fill_manual(values = c('gray', 'black'),
    name = "Group")
dens_cbt
```

- `library(ggplot2)` loads the ggplot2 package, which is used for creating data visualizations in R.
- `library(tidyverse)` loads the tidyverse package, a collection of R packages including ggplot2 for data manipulation and visualization.
- `dens_cbt <- CBT_dataset_independent %>%` creates a ggplot object called dens_cbt. It uses the pipe operator `%>%` to pass the CBT_dataset_independent data frame into the subsequent ggplot code.
- `ggplot(aes(x = AntiSocial, fill = Group))` begins the ggplot object and specifies the aesthetics (aes) mapping. It sets the x-axis to the "AntiSocial" variable and the fill color to the "Group" variable.
- `geom_density(alpha = .7) +` adds a density layer to the plot, displaying the distribution of antisocial attitudes for each group. The alpha parameter controls the transparency of the density curves.
- `theme_minimal() +` sets the plot theme to minimal, which removes gridlines and background elements for a cleaner appearance.
- `labs(x = "Antisocial Attitudes", y = "Probability Density") +` sets the x-axis and y-axis labels.
- `scale_fill_manual(values = c('gray', 'black'), name = "Group")` manually sets the fill colors for the groups (presumably CBT participants and non-participants) and provides a legend title.
- `dens_cbt` displays the density plot.

You can also conduct an independent-samples t-test with directionality in R. Specifically, if we were curious about whether CBT participants have lower levels of antisocial attitudes compared to non-participants, you can perform this in R as well. When you specify the alternative hypothesis's direction, it is called a one-tailed test. This is beyond the scope of this book though, so we will move on to the next item in this chapter: paired-samples t-test.

Paired-Samples T-Test

I mentioned that there is another way to evaluate whether a CBT program impacts participants: to compare antisocial attitudes before and after completing a CBT program. However, it should be noted that randomized experiments such as our example above are considered better than pre- and post-test experimental designs. This is because random assignment can ensure that potential confounding variables are evenly distributed between the treatment and control groups. On the other hand, pre- and post-test experimental designs using the same group of people may not control for all these possible confounding variables, such as time-related effects. Additionally, the participants may be systematically different at baseline when designing pre- and post-test experiments.

Let's first download the data from [the shared Google Drive folder containing the CBT_dataset_paired.csv data](#). In this fictitious data, 100 high-risk inmates completed a CBT program, and their survey responses to antisocial attitudes before and after the program were recorded in this dataset. Pre_CBT_Antisocial represents antisocial attitudes before a CBT program, whereas Post_CBT_Antisocial reflects antisocial attitudes after a CBT program. Let's load the data as we did in the earlier section.

```
library(readr)

CBT_dataset_paired <- read_csv("C:/Users/75JCH0I/OneDrive - West
Chester University of PA/WCU Research/R/data/
CBT_dataset_paired.csv")

View(CBT_dataset_paired)
```

NHST Steps for Paired-Samples T-Test

Step 1: Formulate the Null and Alternative Hypotheses.

- H0: There is no difference in antisocial attitudes between the pre-CBT and post-CBT assessments.
- HA: There is a difference in antisocial attitudes between the pre-CBT and post-CBT assessments.

Step 2: Calculate the Test Statistic.

We will use the `t.test()` function but we will use the `paired = TRUE` argument to conduct a paired t-test.

```
pairedsampt = t.test(x = CBT_dataset_paired$Pre_CBT_Antisocial,
  y = CBT_dataset_paired$Post_CBT_Antisocial, paired = TRUE)
pairedsampt
```

The `t.test()` output shows a t-statistic of 5.4437.

Step 3: Determine the Probability (P-Value) of Obtaining a Test Statistic at Least as Extreme as the Observed Value, Assuming no Relationship Exists.

The p-value was initially displayed in scientific notation as 3.796e-07. To convert it to a standard numeric format, we will apply the same method used above. The resulting p-value is 0.0000003796, indicating a very low probability of detecting a mean difference of 5.245796 in antisocial attitudes between the pre-CBT and post-CBT assessments.

Steps 4 & 5: If the P-Value is Very Small, Typically Less Than 5%, Reject the Null Hypothesis, but if the P-Value is Not Small, Typically 5% or Greater, Retain the Null Hypothesis.

The t-statistic fell within the rejection region. The probability of this sample coming from a population where the mean antisocial attitudes for pre-CBT and post-CBT assessments are equal is exceedingly low. Therefore, it is probable that the sample is from a population where CBT and post-CBT assessments exhibit different mean antisocial attitudes.

Reporting the Results of a Paired-Samples T-Test

Since the results do not show mean scores of antisocial attitudes for pre-CBT and post-CBT assessments, you may want to produce these statistics before reporting the results of a paired-samples t-test.

```
CBT_dataset_paired %>%  
  summarize(m_Pre_CBT_Antisocial = mean(x = Pre_CBT_Antisocial),  
            m_Post_CBT_Antisocial = mean(x = Post_CBT_Antisocial),  
            sd_Pre_Antisocial = sd(x = Pre_CBT_Antisocial),  
            sd_Post_Antisocial = sd(x = Post_CBT_Antisocial))
```

Now you have mean scores and standard deviation values of antisocial attitudes for pre-CBT and post-CBT assessments.

The results from our paired-samples t-test can be presented as follows:

A paired-sample t-test was conducted to compare the antisocial attitudes scores for pre-CBT and post-CBT assessments. There was a significant difference in antisocial attitudes for pre-CBT ($M = 49.50$) and post-CBT assessments ($M = 44.25$; $t(99) = 5.44$, $p < .05$, two-tailed).

If you want to see how t-test results are reported in an academic peer-reviewed journal, please see Choi (2020). In this paper, I used a paired-samples t-test to see changes in perceptions of the police among participants before and after watching videos related to the police.

Conclusion

We have covered two types of t-tests: the independent-samples t-test for comparing the means of two unrelated groups and the paired-samples t-test for comparing the means of two related groups. In the next chapter, we will delve into analysis of variance (ANOVA), which becomes handy when comparing mean scores across more than two groups. Additionally, you will learn about post hoc tests in ANOVA, which help identify statistically significant differences among multiple means.

References

- Choi, J. (2020). Asymmetry in media effects on perceptions of police: An analysis using a within-subjects design experiment. *Police Practice and Research*, 1-17. <https://doi.org/10.1080/15614263.2020.1749624>
- Choi, J., Yim, H., & Lee, D. R. (2020). An examination of the shadow of sexual assault hypothesis among men and women in South Korea. *International Criminal Justice Review*, 30(4), 386-405.
- Farrington, D. P., & McGee, T. R. (2017). The integrated cognitive antisocial potential (ICAP) theory: Empirical testing. In A. A. J. Blokland & V. R. Van Der Geest (Eds.), *Routledge international handbook of life- course criminology* (pp. 11–28). Routledge.
- Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1(4), 451-476.
- Lipsey, M. W., Landenberger, N. A., & Wilson, S. J. (2007). Effects of cognitive-behavioral programs for criminal offenders. *Campbell Systematic Reviews*, 6(1), 1-27.
- Vaske, J., Galyean, K., & Cullen, F. T. (2011). Toward a biosocial theory of offender rehabilitation: Why does cognitive-behavioral therapy work? *Journal of Criminal Justice*, 39(1), 90-102.
- Zara, G. (2019). Cognitive-behavioral treatment to prevent offending and to rehabilitate offenders. In D. P. Farrington, L. Kazemian, & A. R. Piquero (Eds.), *The Oxford handbook of developmental and life-course criminology* (pp. 694–725). Oxford University Press.

This page titled [1.7: T-Test](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Jaeyong Choi \(The Pennsylvania Alliance for Design of Open Textbooks \(PA-ADOPT\)\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.