

7.4: Reporting Results

Reporting results for inferential statistics are much more complex than for descriptive statistics. Where descriptive statistics are usually reported in one sentence using one to two symbols at a time, inferential statistics include both descriptive statistics as a foundation and the results of inferential tests which require additional values and their related symbols. American Psychological Association (APA) format provides useful guidance to ensure that these are reported uniformly across studies by different researchers and statisticians. This helps to both make it easy for other researchers and statisticians to quickly find the information they are looking for, but also facilitates comparisons of results across studies and reduces the chances of misinterpretation caused by formatting differences.

Structure for APA-Formatted Results

In your APA write up for a one samples t -test you should state:

1. Which test was used and the hypothesis which warranted its use.
2. Whether the aforementioned hypothesis was supported or not. To do so properly, three components must be reported:
 - a. The mean and standard deviation for the sample
 - b. The test results as $t(df) = \text{obtained value}$
 - c. The significance as $p < .05$ if significant or $p > .05$ if not significant

I will refer to component 2a as the descriptive statistics and components 2b and 2c together as the evidence string with the determination of significance in the following sections. Let's take a detailed look at both parts of an APA write-up and the components of part 2 of the write-up in detail.

Statement of Test Used and Hypothesis Tested

Your job in creating a results paragraph is to tell your reader the story of those results in a clear and succinct manner. The hypothesis is the catalyst for all the work which was done; thus, the start of the story is the hypothesis and the specific inferential test that went with it. It can be tempting to jump right to the results, but that is like starting a movie at the climax and expecting someone to know what is going on and why it is happening. Therefore, the first sentence of a results section should state these two things together: the hypothesis and statistical test used. Though it is possible and permissible to use two sentences to report this information, there are two drawbacks to doing so: 1. You miss the opportunity to emphasize their connection and 2. You are using more words and punctuation than needed to get the job done well. Therefore, let's report the hypothesis and test used for Data Set 7.1 together in a single sentence. Here is an example of how that can be done succinctly:

A one sample t -test was used to test the hypothesis that the mean test score for the school sample would be different from the population mean of 58.00.

Evidence String for t -Tests: Results and Significance

When reporting the results of an inferential test, more than just the obtained value is needed. We must also provide two additional summaries relevant to the test: the degrees of freedom and the significance using something called a p -value. These details are used together in a prescribed format to create what I will refer to as the evidence string. The evidence string goes at the end of the sentence stating the main results and includes, in this order: the symbol for the inferential test used, the degrees of freedom for the test inside parentheses, the obtained value, and the significance reported using a p -value.

Anatomy of an Evidence String

Symbol for the test	Degrees of Freedom	Obtained Value	p -Value
t	(24)	= 3.43,	$p < .05$.

Symbol for the test

The start of the evidence string provides the symbol for formula used. When a t -test is used, a lowercase letter "t" is written in italics to start the evidence string.

Degrees of Freedom

The degrees of freedom (df) are reported inside parentheses. The df summarize how many pieces of evidence (or data points) are able to either help support or refute a hypothesis. The df is calculated as the amount of data (which is equal to the sample size) minus the number of constraints placed on those data when the one sample t -test formula is used. Thus, the df can be thought of as an adjusted version of the sample size. For the one sample t -test there is one constraint so the df is equal to $n - 1$. Thus, the degrees of freedom also allow the reader to deduce the sample size used for the test (because df is an adjusted version of the sample size). For a more detailed review of df and t -tests, see the section titled "Deeper Dive: What are Degrees of Freedom?" later in this chapter.

Obtained value

The main result is known as the test statistic or obtained value. It is what is yielded by using a formula to analyze the data. The obtained value is the result of step 4 of hypothesis testing. It should be rounded and shown to the hundredths place.

p-values

p-values summarize the chance of a Type I Error when concluding that a research hypothesis is supported based on both the evidence (i.e. the data and corresponding obtained value) and the alpha level chosen. The p -value tells two things: which alpha level was used and whether the result was significant when using that alpha level. Think of the letter p standing for “probability of a Type I Error.” A Type I Error refers to when a hypothesis is supported based on sample evidence but is actually incorrect (see Chapter 6 for a detailed review of Type I Error). When a result is significant the following symbols are used: $p < .05$. Let’s translate each symbol and value then put it together: p stands for “the probability of a Type I Error”; $<$ translates to “less than”; $.05$ is the 5% alpha level written in decimal format. Taken together, $p < .05$ can be read as “The probability of a Type I Error is less than 5%.” This is just another way of saying that a result is significant. This set of symbols gets attached to the end of any sentence of a results section which declares that a hypothesis was significantly supported. Conversely, when a result is not significant, the symbols are changed and reported as $p > .05$; this version translates to “The probability of a Type I Error is greater than 5%.”

These symbols are used anytime a test for significance (i.e. an inferential statistic) is performed. The fact that these symbols are always used and that they are reported at the end of sentences makes it very easy for researchers and statisticians to scan through a results section and identify which results were or were not significant and, thus, which hypotheses were or were not supported. This is especially helpful when a study includes many hypotheses and tests, which can sometimes span a full page or more. Luckily for us, this is an introductory book, so we get to focus on just one hypothesis and its results at a time. This allows us to practice the use of symbols and where they would be located in a results sentence or paragraph one at a time until they are familiar and comfortable. After familiarity and comfort comes utility and ease and, for the focused statistician, potentially a rewarding future career in research with a nice paycheck.

Reporting p -Values

When a result is significant, the following symbols are used: $p < .05$.

When a result is not significant, the following symbols are used: $p > .05$.

The symbols are stating whether, based on the strength of the evidence, the probability of a Type I Error is less than (significant) or greater than (not significant) the generally accepted threshold of 5%.

Deeper Dive: What are Degrees of Freedom?

Degrees of Freedom (df) tell how much information you have that is free to vary. We are measuring things that can vary (i.e. variables) and are looking for specific patterns in how they vary. We use the information from a sample to estimate what is true in the population. Therefore, each case or person provides another piece of information we can use to estimate a parameter in the population.

When we estimate based on these variables, however, the variables are no longer fully independent. Each time we estimate a parameter using our data, the data are constrained. Each test that we use requires us to make at least one estimate about a population parameter using a sample statistic; each time we estimate a parameter we lose one degree of freedom.

When the constraints are groups (which is the case for t -tests), the number of independent groups is summarized with the symbol k . Degrees of freedom for each t -test are calculated as the total sample size minus the number of groups. This can be summarized as $n - k$. However, because each t -test has a different but stable number of groups, the df formula is often written with the specifics for k indicated as follows:

In the one sample t -test, there is only one sample so $k = 1$. Thus, our df for the one sample t -tests is calculated as $n - 1$.

In the dependent samples t -test, there is only one sample being tested two times so $k = 1$. Thus, our df for the dependent samples t -test is calculated as $n - 1$.

In the independent samples t -test we estimate two parameters: the mean for group 1 and the mean for group 2. Thus, our df is calculated as $n - 2$. You may also see this specified by group so that df is calculated as $n_1 - 1 + n_2 - 1$. Each of these will yield the same number.

Reading Review 7.2

1. Which two things should be stated first in a results summary paragraph?
2. What are the four parts of the evidence string and what does each one report?
3. Which set of symbols should be used at the end of a sentence to indicate that a result was not significant?

4. What does reporting Cohen's d add to a results summary?

This page titled [7.4: Reporting Results](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by .