19

One Sample Parametric Tests

The One-Sample z-Test
The One-Sample t-Test
The Confidence Interval

Dr. Roberta Damon developed a marital profile of Southern Baptist missionaries in eastern South America in 1985. Part of her study involved measuring whether SBC missionary husband-wife pairs differed from American society-at-large on the variable "Couple Agreement on Religious Orientation," as measured by the Religious Orientation Test. She set $\alpha = 0.01$ and decided to use a two-tail test. The American mean (μ) at the time was **56**. The mean score ($\overline{X}$) and estimated standard deviation (**s**) of her sample of 330 missionaries was 86.3 and 20.847 respectively. Applying the sampling distribution z-formula we introduced on page 17-9, she computed **z** as follows:

$$z = \frac{\overline{X}-\mu}{s_{\overline{X}}} = \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} = \frac{86.3 - 56}{\frac{20.847}{\sqrt{330}}} = \frac{30.3}{1.1476} = 26.403$$

Remember that a two-tail test (α=0.01) requires only **z=2.58** for declaring a difference significant. Here we see **z=26.403**. Southern Baptist missionaries serving in eastern South America in 1984, as a group, scored over ***26 standard errors above*** the national mean in "Couple Agreement"[1] on the Religious Orientation Scale!

In chapters 17 and 18 we explored the theoretical basis for hypothesis testing with sample means. Here in Chapter 19 we will use these principles to establish our first practical use for hypothesis testing. **One-sample tests** compare a population mean (μ) with a single sample mean($\overline{X}$). We have already seen the **one-sample z-test** in action. We will also discuss the **one-sample t-test**.

The One-Sample z-Test

The procedure we used in chapter 17 to compute the significance of a church's attitude toward building renovation, as well as in the analysis shown above, used the ***one-sample z-test***. The formula has two forms. When the population standard deviation (σ) is *known*, we use the equation on the left below. When sigma (σ) is *not known*

[1]"Couple Agreement"measures the degree to which religion serves as a relationship strength. A high score means that there is high agreement on religious issues between husbands and wives, and that subjects consider religion an important part of their marriage relationship. Roberta McBride Damon, "A Marital Profile of Southern Baptist Missionaries in Eastern South America," (Fort Worth, Texas: Southwestern Baptist Theological Seminary, 1985), 37

(which is usually the case), we use **s** to estimate σ, and so use the equation on the right (the more popular of the two).

$$z = \frac{\overline{X}-\mu}{\sigma_{\overline{X}}} \qquad z = \frac{\overline{X}-\mu}{s_{\overline{X}}} \qquad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \quad s_{\overline{X}} = \frac{s}{\sqrt{n}}$$

The limitation to using **s** to estimate **σ** is whether the sample is large enough to approximate a normal curve. "Large enough" means at least ***n=30 subjects***. The normal curve table (z-) requires a normal distribution of scores in order to give accurate proportions under the curve. ***When a sample contains less than 30 scores*, the requirement of normality is not met, and we cannot use the normal curve table or the *z-test*. In this case, use the one sample t-test.**

The One-Sample t-Test

The **one-sample t-test** must be used to test the difference between a population mean (**μ**) with a single sample mean($\overline{X}$) when **both** (1) the population standard deviation (**σ**) is unknown, and (2) the sample size (**n**) is less than 30 (That is, the z-test cannot be used under these conditions).

However, note that the one-sample t-test **can be used for samples of any size**, and is often used instead of the z-test, making it a one of the most popular tests of difference.

The value of the one-sample t is obtained exactly like the one-sample z. The formula looks just like the z-test formula given earlier:

$$\mathbf{t} = \frac{\overline{X}-\mu}{s_{\overline{X}}}$$

The logic behind the one-sample t is the same as we've used for the z-test, with one exception. A smaller n means lower power (see 18-5). Since the t-test is designed to be used with smaller n samples, the t-distribution critical value table assigns a **slightly larger critical value to reflect the loss of power.**

The t-Distribution Table

Look at the t-distribution table in appendix A3-2 at the back of the text. Notice that the column headings reflect levels of significance ranging from 0.10 to 0.005. These are one-tailed probabilities.

The heading ***df*** on the left bold-faced column is an abbreviation for "degrees of freedom." Degrees of freedom is directly related to the size of a sample, but it is a concept that is much easier to illustrate than define.

Let's say we have three variables that add to 10, such that **X + Y + Z = 10**. What number can we substitute for "X"? The answer is ***any number***. The variable X is **free to vary**. We have one degree of freedom. So let's arbitrarily assign the value "2" to X.

We now have **2 + Y + Z = 10.** What number can we substitute for "Y"? The answer is ***any number***. Variable Y is **free to vary**. We now have two degrees of freedom. Let's assign the value "5" to Y.

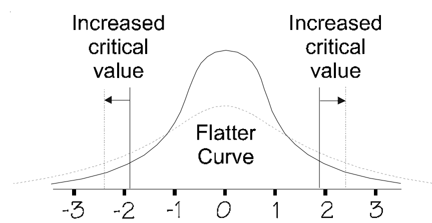
We now have **2 + 5 + Z = 10.** What number can we substitute for "Z"? There is only one number in the universe: ***Z must equal "3"*** for the equation to be true. Z is **NOT free to vary**. So for X, Y, and Z, three numbers (n=3), we have two (n-1) degrees of freedom. This relationship holds true for any size n: ***df = n-1.*** Or in English, if you have **26** subjects in a study and use the one-sample t-test, then the degrees of freedom for the analysis is (n-1 = 26-1 =) **25**.

Recall that in the Normal Curve table we can find proportion values for any z-score. We focused on the four primary critical values of 1.65, 1.96, 2.33, and 2.58, but could calculate proportions for any z from 0 to 3.

Look at the t-table under the column labelled "0.05" and move down to the bottom of the column, next to df=∞. The value of the t-test critical value (**1.645**) is exactly the same as the value in the Normal Curve table (0.05, 1-tail test). To the right of "1.645" is the .025 column with the value "**1.96**." This is exactly the same as the 2-tail 0.05 Normal Curve value of z. The heading ".025" in the t-table refers to the $\alpha/2$ area of the two-tail test: 0.05/2 = 0.025.

The next value to the right is "**2.326**," the one-tail 0.01 value of z. And the next, "**2.576**." *These are the same four essential critical values we studied for the normal curve: 1.65, 1.96, 2.33, and 2.58.*

The t-Distribution table differs from the Normal Curve table in that it contains nothing but critical values for a given level of significance and df. ***Each "df" row provides the critical value for a unique t-distribution.*** As we have just seen, for n=∞, the t-distribution is the same as the Normal Curve distribution. *As n decreases, the t-distribution becomes increasingly platykurtic.* That is, the smaller the sample size, the flatter and wider the distribution. This pushes the critical values out farther on the tails, making "significance" harder to establish, as you can see in the diagram at right.



Now choose the column in the t-table ending with "1.645." Move up the column and watch how the critical values increase. ***As df decreases, critical values increase.*** This is just another way of saying "as n decreases, power decreases." Thus, the t-table accounts for lower power which derives from smaller n.

Computing t

Let's revisit the congregation study of attitude toward building renovation we did with the z-test in Chapter 17. Suppose our random sample is a small one, made up of only 25 members rather than 100. Let's use the same hypothesized population value of 4.0 and the same mean and standard deviation of 3.8 and 0.7 respectively. Computing t we have

$$\mathbf{t} = \frac{\overline{X}-\mu}{s_{\overline{X}}} = \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} = \frac{3.8-4.0}{\frac{0.7}{\sqrt{25}}} = \frac{-0.20}{0.14} = \mathbf{-1.43}$$

The critical value for α=.05, 1-tail, and df = 24 is **±1.711**. Notice that this critical value for t, symbolized as t_{cv}, is a larger value than the comparable z-test value of 1.65. But since our value of -1.43 is smaller (not as far out on the left tail) than -1.711 , *we retain the null hypothesis.*

While the difference between sample mean and hypothesized population mean is the same as before (-0.2), the standard error of the mean ($\mathbf{s}_x$) was twice as large -- 0.14 in the t-test as opposed to 0.07 in the z-test. This is due to the smaller number of subjects (n=25) in this sample as compared to the former one (n=100).

Additionally, the critical value "bar" for the t-test (**1.711**) is higher than the z-test (**1.645**). So the smaller sample size yields less power. The same hypothesis (H_0:μ=4.0) and the same difference (-0.2) **yielded two different results**. *The z-test yielded a significant difference. The t-test did not. Why? Because in the second case we lacked sufficient power to declare the difference significant*. The t-test does not correct for lack of power. It simply allows us to test samples too small for the z-test.

Up to now we have used the z- and t- formulas to test single hypotheses. Another, less common, use for these formulas is in the creation of a **confidence interval.**

Confidence Intervals

The ***confidence interval*** offers another approach to making statistical decisions. In this approach, we set an interval around the population mean, bordered by confidence limits. We can then state, with a given degree of confidence, that the null hypothesis is true if any sample mean which falls within this interval , and false if it falls outside this interval. The benefit of using a confidence interval is that any number of sample means can be tested with only one computation.

A z-Score Confidence Interval

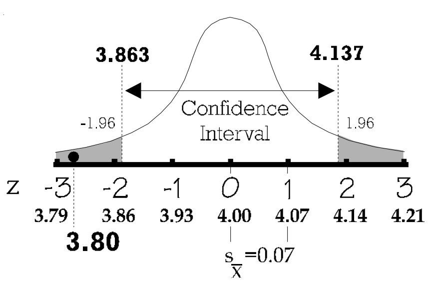
The z-score confidence interval equation looks like this:

$$CI_{95} = \mu \pm z s_{\bar{X}}$$

where $\mathbf{CI_{95}}$ is the 95% confidence interval consisting of *two endpoint values* (x.x, x.x), $\boldsymbol{\mu}$ is the population mean, **z** is the z-score for the given level of significance (in this case, α= 0.05, so z = 1.96), and $s_{\bar{X}}$ is the standard error of the mean.

Now let's revisit the church attitude study again using n=100 subjects and a confidence interval based on the z-score. Using the formula above we have the following computation:

$$\begin{aligned} CI_{95} = \mu \pm z s_{\bar{X}} &= 4.0 \pm (1.96)(0.07) \\ &= 4.0 \pm (0.1372) \\ &= 4.0 - .1372,\ 4.0 + .1372 \\ &= 3.863,\ 4.137 \end{aligned}$$



In the diagram at left you can see the church's mean score (n=100) of **3.8**, and the confidence interval values of 3.863 and 4.137. The mean (3.8) falls outside the confidence interval. We therefore reject the null hypothesis (just as we did with the hypothesis test in Chapter 17). The church has a negative attitude toward renovation. But let's assume we tested twenty churches in our association.[2] All we need do is calculate the mean score of each church and compare that to the interval "3.863 - 4.137." Any church mean falling below this interval

---

[2]I've oversimplified this to focus your attention on the meaning of confidence interval. But for this to actually work as I've described it, we have to assume that all twenty churches produced the same standard deviation (s) on their attitude scores -- and this is an unreasonable assumption. If we were to actually do this study, we would compute the overall standard deviation from all twenty churches, and use that value to construct the confidence interval. Then, any church falling outside the interval would be significantly different from 4.00, the hypothesis neutral point. But even so, we make one computation, 20 comparisons.

reflects a significant negative attitude. Any church mean falling above the interval reflects a significant positive attitude. Any mean falling within the interval reflects no attitude (neutral).

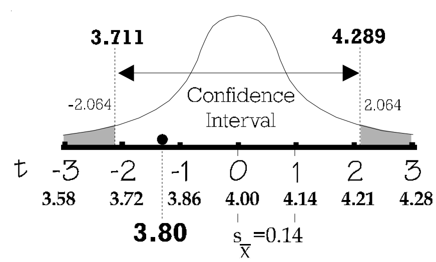
Confidence intervals are *always based on two-tailed tests.* The two values, 3.863 and 4.137, are called the ***confidence limits***. The range of scores between the limits, shown at right by the two-headed arrow, is the ***confidence interval***.

A t-Score Confidence Interval

We can also develop a confidence interval with t-scores. We simply substitute the appropriate t_{cv} value for z_{cv}. In the above example, the appropriate t-table critical value (0.05, 2-tail, df=24) is **2.064**. The standard error of the mean is different because of the change in sample size (25 vs. 100). It is **0.14** rather than 0.07, which we used with the z-test (note the equation in the middle of page 19-3). Therefore, the t-value confidence interval equation is

$$
\begin{aligned}
CI_{95} = \mu \pm ts_{\bar{X}} &= 4.0 \pm (2.064)(0.14) \\
&= 4.0 \pm (0.289) \\
&= 4.0 - 0.289, 4.0 + 0.289 \\
&= 3.711, 4.289
\end{aligned}
$$

The flatter curve of the t-distribution forces the **interval to be wider** than the one we computed with z. Because of the larger standard error of the mean ($s_{\bar{X}}$), the t-score cut-off values are farther apart (***0.14*** vs. 0.07). Notice also that the sample mean of 3.8, which fell ***outside the z-score interval***, falls ***inside the t-score confidence interval***. This agrees with the hypothesis test on page 19-3. Since the mean of 3.8 falls within the confidence interval, it is declared not significantly different. The different result (z- vs. t-) is **due directly to different n's** (100 subjects vs. 25 subjects). The loss of power with n=25 changed our significant finding to a non-significant finding.



Summary

This chapter is built on Chapters 17 and 18. The principal extensions we made in this chapter include the use of the t-distribution table, the concept of "degrees of freedom" and "confidence interval." The next chapter extends these concepts still further: to ***testing mean differences between <u>two</u> samples.***

Vocabulary

Confidence interval	Distance between the 2-tail critical values centered on mean. "Region of acceptance (of null)"
Degrees of freedom	the number of values free to vary given a fixed sum (*n-1* for one group)
One sample z-test	tests difference between μ and X-bar - if σ is known, or if σ is unknown and n>30
One sample t-test	tests difference between μ and X-bar - if σ unknown and estimated by s (must use if n<30)

Remember:
The one-sample t-test can be used with any sample size, which makes it one of the most popular statistics of difference.

Study Questions

1. A sample mean on an attitude scale equals 3.3 with a standard deviation of 0.5. There were 16 people tested in the group. Test the hypothesis: "The group will score significantly higher than 3.0 on attitude X." (Use =0.01)
   A. State the statistical hypothesis.
   B. Compute the proper statistic to answer the question. *(con't)*
   C. Test the statistic with the appropriate table.
   D. State your conclusion.
   E. Establish a 99% confidence interval about the sample mean. (Careful here...)
   F. Draw the sampling distribution and the confidence interval.

2. Repeat #1 above, but with a sample size of 49.

3. A study in 1980 revealed that the average salary of Southern Baptist ministers of education in Fort Worth was $29,000 (fictitious data). You randomly sample 28 ministers of education (1995) and find their average salary is $37,000 with a standard deviation of $3,000. Have salaries improved significantly?

   A. Draw and label an appropriate sampling distribution.
   B. State the research hypothesis.
   C. State the statistical hypothesis.
   D. Select the proper test and compute the statistic.
   E. Test the statistic with the appropriate table.
   F. Establish a 95% confidence interval about the sample mean.
   G. Does this confidence interval agree with your hypothesis test? Explain how.

Sample Test Questions

1. When must you use the one-sample t-test to test differences between two groups of scores?
   A. When you do not know σ and $n<30$.
   B. When you do not know μ.
   C. When you know σ and $n>30$.
   D. When you do not know s and $n<30$.

2. The denominator of the one-sample t-test equals
   A. $n-1$
   B. $s/\sqrt{n}$
   C. σ
   D. s

3. What would be your best estimate of the critical value for t (.05, df=19, 1-tail)?
   A. 1.65
   B. 1.73
   C. 1.96
   D. 2.01

4. T F A confidence interval produced by a z-score is narrower than one produced by a t-score — particularly for samples of 30 or fewer subjects.

5. T F If a confidence interval, centered at the hypothesized population mean, includes a sample mean, then the sample mean is considered "significantly different."

