STATISTICS
IN EVERYDAY LIFE

In social science research, the 0.05 alpha level has become the standard indicator of a significant difference. This alpha level means that we will incorrectly reject the null hypothesis only five times out of every 100 tests. These might seem like excellent odds, but research that involves potentially harmful drugs would call for even lower alpha levels (0.001, 0.0001, or even lower) to minimize the possibility of making incorrect decisions and endangering health.

of a significant result. However, the widespread use of the 0.05 level is simply a convention, and there is no reason that alpha cannot be set at virtually any sensible level (such as 0.04, 0.027, 0.083). The researcher has the responsibility of selecting the alpha level that seems most reasonable in terms of the goals of the research project.

The Student's *t* Distribution

To this point, we have considered only situations involving single-sample means in which the value of the population standard deviation (σ) was known. Needless to say, in most research situations, the value of σ will be unknown. However, a value for σ is required in order to compute the standard error of the mean (σ/N), convert our sample outcome into a Z score, and place the Z(obtained) on the sampling distribution (step 4). How can we reasonably obtain a value for the population standard deviation?

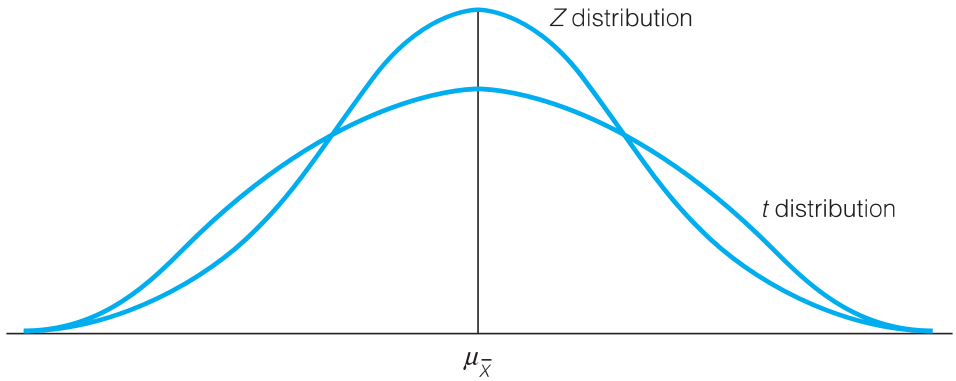
It might seem sensible to estimate σ with s, the sample standard deviation. As we noted in Chapter 6, s is a biased estimator of σ, but the degree of bias decreases as the sample size increases. For large samples (that is, samples with 100 or more cases), the sample standard deviation yields an adequate estimate of σ. Thus, for large samples, we simply substitute s for σ in the formula for Z(obtained) in step 4 and continue to use the standard normal curve to find areas under the sampling distribution.[1]

However, for smaller samples, when σ is unknown, an alternative distribution called the **Student's *t* distribution** must be used to find areas under the sampling distribution and establish the critical region. The shape of the t distribution varies as a function of the sample size. The relative shapes of the t and Z distributions are depicted in Figure 7.7. For small samples, the t distribution is much flatter than the Z distribution, but as the sample size increases, the t distribution comes to resemble the Z distribution more and more until the two are essentially identical for sample sizes greater than 120. As N increases, the sample standard deviation (s) becomes a more and more adequate estimator of the population standard deviation (σ) and the t distribution becomes more and more like the Z distribution.

---

[1]Even though its effect will be minor and will decrease with the sample size, we will always correct for the bias in s by using the term $N - 1$ rather than N in the computation for the standard deviation of the sampling distribution when s is unknown.

FIGURE 7.7 The *t* Distribution and the *Z* Distribution



Applying Statistics 7.1 Testing Sample Means for Significance: The One-Sample Case

For a random sample of 152 felony cases tried in a local court, the average prison sentence was 27.3 months. Is this significantly different from the average prison term for felons nationally (28.7 months)? We will use the five-step model to organize the decision-making process.

Step 1. Making Assumptions and Meeting Test Requirements.

Model: Random sampling
Level of measurement is interval-ratio
Sampling distribution is normal

Because this is a large sample ($N > 100$) and the length of a sentence is an interval-ratio variable, we can conclude that the model assumptions are satisfied.

Step 2. Stating the Null Hypothesis (H_0). The null hypothesis would say that the average sentence locally (for *all* felony cases) is equal to the national average. In symbols:

$$H_0: \mu = 28.7$$

The research question does not specify a direction; it only asks if the local sentences are "different from" (not higher or lower than) national averages. This suggests a two-tailed test:

$$(H_1: \mu \neq 28.7)$$

Step 3. Selecting the Sampling Distribution and Establishing the Critical Region. Because this is a large sample, we can use Appendix A to establish the critical region and state the critical scores as *Z* scores (as opposed to *t* scores).

$$\text{Sampling distribution} = Z \text{ distribution}$$

$$\alpha = 0.05$$

$$Z(\text{critical}) = \pm 1.96$$

Step 4. Computing the Test Statistic. The necessary information for conducting a test of the null hypothesis is:

$$\bar{X} = 27.3 \qquad \mu = 28.7$$

$$s = 3.7$$

$$N = 152$$

The test statistic—*Z*(obtained)—would be:

$$Z(\text{obtained}) = \frac{\bar{X} - \mu}{s/\sqrt{N-1}}$$

$$Z(\text{obtained}) = \frac{27.3 - 28.7}{3.7/\sqrt{152-1}}$$

$$Z(\text{obtained}) = \frac{-1.40}{3.7/\sqrt{151}}$$

$$Z(\text{obtained}) = \frac{-1.40}{0.30}$$

$$Z(\text{obtained}) = -4.67$$

Step 5. Making a Decision and Interpreting the Test Results. With alpha set at 0.05, the critical region begins at *Z*(critical) $= \pm 1.96$. With an obtained *Z* score of -4.67, the null would be rejected. The difference between the prison sentences of felons convicted in the local court and felons convicted nationally is statistically significant. The difference is so large that we may conclude that it did not occur by random chance. The decision to reject the null hypothesis has a 0.05 probability of being wrong.

The Distribution of *t*: Using Appendix B. The *t* distribution is summarized in Appendix B. The *t* table differs from the *Z* table in several ways. First, there is a column at the left of the table labeled *df* for "degrees of freedom."[2] Because the exact shape of the *t* distribution varies by sample size, the exact location of the critical region also varies. Degrees of freedom, which are equal to $N - 1$ in the case of a single-sample mean, must first be computed before the critical region for any alpha can be located. Second, alpha levels are arrayed across the top of Appendix B in two rows—one row for the one-tailed tests and one for two-tailed tests. To use the table, begin by locating the selected alpha level in the appropriate row.

The third difference is that the entries in the table are the actual scores—called *t*(critical)—that mark the beginnings of the critical regions and not areas under the sampling distribution. To illustrate the use of this table with single-sample means, find the critical region for alpha equal to 0.05, two-tailed test, for $N = 30$. The degrees of freedom will be $N - 1$, or 29; reading down the proper column, you should find a value of 2.045. Thus, the critical region for this test will begin at *t*(critical) $= \pm 2.045$.

Notice that this *t*(critical) is larger in value than the comparable *Z*(critical), which for a two-tailed test at an alpha of 0.05 would be ± 1.96. This is because the *t* distribution is flatter than the *Z* distribution (see Figure 7.7). On the *t* distribution, the critical regions will begin farther away from the mean of the sampling distribution; therefore, the null hypothesis will be harder to reject. Furthermore, the smaller the sample size (the lower the degrees of freedom), the larger the value of *t*(obtained) necessary to reject the H_0.

Also note that the values of *t*(critical) decrease as degrees of freedom increase. For one degree of freedom, the *t*(critical) is 12.706 for an alpha of 0.05 with a two-tailed test, but this score grows smaller for larger samples. For degrees of freedom greater than 120, the value of *t*(critical) is the same as the comparable value of *Z*(critical), or ± 1.96. As the sample size increases, the *t* distribution resembles the *Z* distribution more and more until, with sample sizes greater than 120, the two distributions are essentially identical.[3]

Using *t* in a Test. To demonstrate the uses of the *t* distribution in more detail, we will work through an example problem. Note that in terms of the five-step model, the changes occur mostly in steps 3 and 4. In step 3, the sampling distribution will be the *t* distribution and degrees of freedom (*df*) must be computed before locating

[2]Degrees of freedom refers to the number of values in a distribution that are free to vary. For a sample mean, a distribution has $N - 1$ degrees of freedom. This means that for a specific value of the mean and of N, $N - 1$ scores are free to vary. For example, if the mean is 3 and $N = 5$, the distribution of five scores would have $5 - 1 = 4$ degrees of freedom. When the values of four of the scores are known, the value of the fifth is fixed. If four scores are 1, 2, 3, and 4, the fifth must be 5 and no other value.

[3]Appendix B abbreviates the *t* distribution by presenting a limited number of critical *t* scores for degrees of freedom between 31 and 120. If the degrees of freedom for a specific problem equal 77 and alpha equals 0.05, two-tailed, we have a choice between a *t*(critical) of 2.000 ($df = 60$) and a *t*(critical) of 1.980 ($df = 120$). In situations such as these, take the larger table value as *t*(critical). This will make rejection of H_0 less likely and is therefore the more conservative course of action.

the critical region or the t(critical) score. In step 4, a slightly different formula for computing the test statistic—t(obtained)—will be used. As compared with the formula for Z(obtained), s will replace σ and $N - 1$ will replace N.

Specifically:

FORMULA 7.2

$$t\,(\text{obtained}) = \frac{\bar{X} - \mu}{s/\sqrt{N-1}}$$

A researcher wonders if commuter students are different from the general student body in terms of academic achievement. She has gathered a random sample of 30 commuter students and has learned from the registrar that the mean grade point average for all students is 2.50 ($\mu = 2.50$), but the standard deviation of the population (σ) has never been computed. Sample data are reported here. Is the sample from a population that has a mean of 2.50?

Student Body	Commuter Students
$\mu = 2.50 (= \mu_{\bar{X}})$	$\bar{X} = 2.78$
$\sigma = ?$	$s = 1.23$
	$N = 30$

Step 1. Making Assumptions and Meeting Test Requirements.

Model: Random sampling
Level of measurement is interval-ratio
Sampling distribution is normal

Step 2. Stating the Null Hypothesis.

$$H_0: \mu = 2.50$$

$$(H_1: \mu \neq 2.50)$$

You can see from the research hypothesis that the researcher has not predicted a direction for the difference. This will be a two-tailed test.

Step 3. Selecting the Sampling Distribution and Establishing the Critical Regions. Since σ is unknown and the sample size is small, the t distribution will be used to find the critical region. Alpha will be set at 0.01:

$$\text{Sampling distribution} = t \text{ distribution}$$

$$\alpha = 0.01, \text{two-tailed test}$$

$$df = (N - 1) = 29$$

$$t\,(\text{critical}) = \pm 2.756$$

ONE STEP AT A TIME **Completing Step 4 of the Five-Step Model: Computing *t*(obtained)**

Follow these procedures when using Student's t distribution.

To Compute the Test Statistic by Using Formula 7.2

Step	***Operation***
1.	Find the square root of $N - 1$.
2.	Divide the quantity you found in step 1 into the sample standard deviation (s).
3.	Subtract the population mean (μ) from the sample mean ($\bar{X}$).
4.	Divide the quantity you found in step 3 by the quantity you found in step 2. This value is *t*(obtained).

Step 4. Computing the Test Statistic.

$$t(\text{obtained}) = \frac{\bar{X} - \mu}{s/\sqrt{N - 1}}$$

$$t(\text{obtained}) = \frac{2.78 - 2.50}{1.23/\sqrt{29}}$$

$$t(\text{obtained}) = \frac{.28}{.23}$$

$$t(\text{obtained}) = +1.22$$

Step 5. Making a Decision and Interpreting Test Results. The test statistic does not fall into the critical region. Therefore, the researcher fails to reject the H_0. The difference between the sample mean (2.78) and the population mean (2.50) is not statistically significant. The difference is no greater than what would be expected if only random chance were operating. The test statistic and critical regions are displayed in Figure 7.8.

To summarize, when testing single-sample means, we must make a choice regarding the theoretical distribution we will use to establish the critical region. The choice is straightforward. If the population standard deviation (σ) is known or the sample size is large, the Z distribution (summarized in Appendix A) will be used. If σ is unknown and the sample is small, the t distribution (summarized in Appendix B)

ONE STEP AT A TIME **Completing Step 5 of the Five-Step Model: Making a Decision and Interpreting Results**

Step	***Operation***
1.	Compare the *t*(obtained) to the *t*(critical). If *t*(obtained) is *in* the critical region, *reject* the null hypothesis. If *t*(obtained) is *not in* the critical region, *fail to reject* the null hypothesis.
2.	Interpret your decision in terms of the original question. For example, our conclusion for the example problem used in this section was "There is no significant difference between the average GPAs of commuter students and the general student body."

FIGURE 7.8 Sampling Distribution Showing *t*(Obtained) versus *t*(Critical) ($\alpha = 0.05$, two-tailed test, $df = 29$)

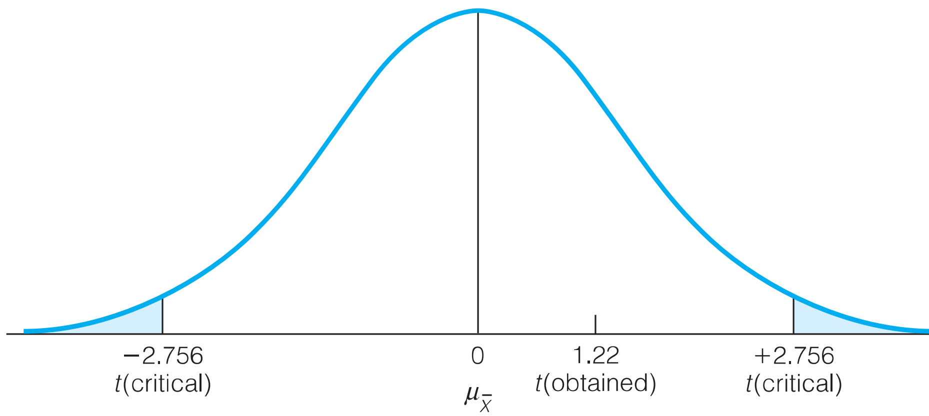


TABLE 7.6 Choosing a Sampling Distribution When Testing Single-Sample Means for Significance

If Population Standard Deviation (σ) Is	Sampling Distribution
Known	*Z* distribution
Unknown and the sample size (*N*) is large	*Z* distribution
Unknown and the sample size (*N*) is small	*t* distribution

will be used. These decisions are summarized in Table 7.6. *(For practice in using the t distribution in a test of hypothesis, see problems 7.8, 7.10 to 7.12, and 7.19 .)*

Tests of Hypotheses for Single-Sample Proportions (Large Samples)

In many cases, we work with variables that are not interval-ratio in level of measurement. One alternative in this situation would be to use a sample proportion (P_s) rather than a sample mean as the test statistic. As we shall see, the overall procedures for testing single-sample proportions are the same as those for testing means. The central question is still "Does the population from which the sample was drawn have a certain characteristic?" We still conduct the test based on the assumption that the null hypothesis is true, and we still evaluate the probability of the obtained sample outcome against a sampling distribution of all possible sample outcomes. Our decision at the end of the test is also the same. If the obtained test statistic falls into the critical region (is unlikely, given the assumption that the H_0 is true), we reject the H_0.

Of course, there are also some important differences in significance tests for sample proportions. These differences are best related in terms of the five-step model. In step 1, we assume that the variable is measured at the nominal level. In step 2, the symbols used to state the null hypothesis are different even though it is still a statement of "no difference."