Identifying $Z = -1.37$ in the normal probability table or using statistical software, we can determine that the left tail area is 0.0853. Doubling this value yields the total area in the two tails: about 0.17. This is the estimated p-value for the hypothesis test. However, there's a problem: this is very different than the earlier p-value we computed: 0.2444.

The discrepancy is explained by normal model's poor representation of the null distribution in Figure 2.35. As noted earlier, the null distribution from the simulations is not very smooth, and the distribution itself is slightly skewed. That's the bad news. The good news is that we can foresee these problems using some simple checks. We'll learn about these checks in the following chapters.

In Section 2.5 we noted that the two common requirements to apply the Central Limit Theorem are (1) the observations in the sample must be independent, and (2) the sample must be sufficiently large. The guidelines for this particular situation – which we will learn in Section 3.1 – would have alerted us that the normal model was a poor approximation.

2.7.4 Conditions for applying the normal model

The success story in this section was the application of the normal model in the context of the opportunity cost data. However, the biggest lesson comes from our failed attempt to use the normal approximation in the medical consultant case study.

Statistical techniques are like a carpenter's tools. When used responsibly, they can produce amazing and precise results. However, if the tools are applied irresponsibly or under inappropriate conditions, they will produce unreliable results. For this reason, with every statistical method that we introduce in future chapters, we will carefully outline conditions when the method can reasonably be used. These conditions should be checked in each application of the technique.

2.8 Confidence intervals

A point estimate provides a single plausible value for a parameter. However, a point estimate is rarely perfect; usually there is some error in the estimate. In addition to supplying a point estimate of a parameter, a next logical step would be to provide a plausible *range of values* for the parameter.

2.8.1 Capturing the population parameter

A plausible range of values for the population parameter is called a **confidence interval**. Using only a point estimate is like fishing in a murky lake with a spear, and using a confidence interval is like fishing with a net. We can throw a spear where we saw a fish, but we will probably miss. On the other hand, if we toss a net in that area, we have a good chance of catching the fish.

If we report a point estimate, we probably will not hit the exact population parameter. On the other hand, if we report a range of plausible values – a confidence interval – we have a good shot at capturing the parameter.

⊙ **Guided Practice 2.59** If we want to be very certain we capture the population parameter, should we use a wider interval or a smaller interval?[44]

[44]If we want to be more certain we will capture the fish, we might use a wider net. Likewise, we use a wider confidence interval if we want to be more certain that we capture the parameter.

2.8.2 Constructing a 95% confidence interval

A point estimate is our best guess for the value of the parameter, so it makes sense to build the confidence interval around that value. The standard error, which is a measure of the uncertainty associated with the point estimate, provides a guide for how large we should make the confidence interval.

> **Constructing a 95% confidence interval**
> When the sampling distribution of a point estimate can reasonably be modeled as normal, the point estimate we observe will be within 1.96 standard errors of the true value of interest about 95% of the time. Thus, a **95% confidence interval** for such a point estimate can be constructed:
>
> $$\text{point estimate} \ \pm \ 1.96 \times SE \tag{2.60}$$
>
> We can be **95% confident** this interval captures the true value.

Guided Practice 2.61 Compute the area between -1.96 and 1.96 for a normal distribution with mean 0 and standard deviation 1. [45]

Example 2.62 The point estimate from the opportunity cost study was that 20% fewer students would buy a DVD if they were reminded that money not spent now could be spent later on something else. The point estimate from this study can reasonably be modeled with a normal distribution, and a proper standard error for this point estimate is $SE = 0.078$. Construct a 95% confidence interval.[46]

Since the conditions for the normal approximation have already been verified, we can move forward with the construction of the 95% confidence interval:

$$\text{point estimate} \ \pm \ 1.96 \times SE \quad \rightarrow \quad 0.20 \ \pm \ 1.96 \times 0.078 \quad \rightarrow \quad (0.047, 0.353)$$

We are 95% confident that the DVD purchase rate resulting from the treatment is between 4.7% and 35.3% lower than in the control group. Since this confidence interval does not contain 0, it is consistent with our earlier result where we rejected the notion of "no difference" using a hypothesis test.

In Section 1.1 we encountered an experiment that examined whether implanting a stent in the brain of a patient at risk for a stroke helps reduce the risk of a stroke. The results from the first 30 days of this study, which included 451 patients, are summarized in Table 2.36. These results are surprising! The point estimate suggests that patients who received stents may have a *higher* risk of stroke: $p_{trmt} - p_{ctrl} = 0.090$.

---

[45] We will leave it to you to draw a picture. The Z scores are $Z_{left} = -1.96$ and $Z_{right} = 1.96$. The area between these two Z scores is $0.9750 - 0.0250 = 0.9500$. This is where "1.96" comes from in the 95% confidence interval formula.

[46] We've used $SE = 0.078$ from the last section. However, it would more generally be appropriate to recompute the SE slightly differently for this confidence interval using the technique introduced in Section 3.2.1. Don't worry about this detail for now since the two resulting standard errors are, in this case, almost identical.

	stroke	no event	Total
treatment	33	191	224
control	13	214	227
Total	46	405	451

Table 2.36: Descriptive statistics for 30-day results for the stent study.

Example 2.63 Consider the stent study and results. The conditions necessary to ensure the point estimate $p_{trmt} - p_{ctrl} = 0.090$ is nearly normal have been verified for you, and the estimate's standard error is $SE = 0.028$. Construct a 95% confidence interval for the change in 30-day stroke rates from usage of the stent.

The conditions for applying the normal model have already been verified, so we can proceed to the construction of the confidence interval:

$$\text{point estimate} \ \pm\ 1.96 \times SE \quad \rightarrow \quad 0.090 \ \pm\ 1.96 \times 0.028 \quad \rightarrow \quad (0.035, 0.145)$$

We are 95% confident that implanting a stent in a stroke patient's brain increased the risk of stroke within 30 days by a rate of 0.035 to 0.145. This confidence interval can also be used in a way analogous to a hypothesis test: since the interval does not contain 0, it means the data provide statistically significant evidence that the stent used in the study *increases* the risk of stroke, contrary to what researchers had expected before this study was published!

As with hypothesis tests, confidence intervals are imperfect. About 1-in-20 properly constructed 95% confidence intervals will fail to capture the parameter of interest. Figure 2.37 shows 25 confidence intervals for a proportion that were constructed from simulations where the true proportion was $p = 0.3$. However, 1 of these 25 confidence intervals happened not to include the true value.

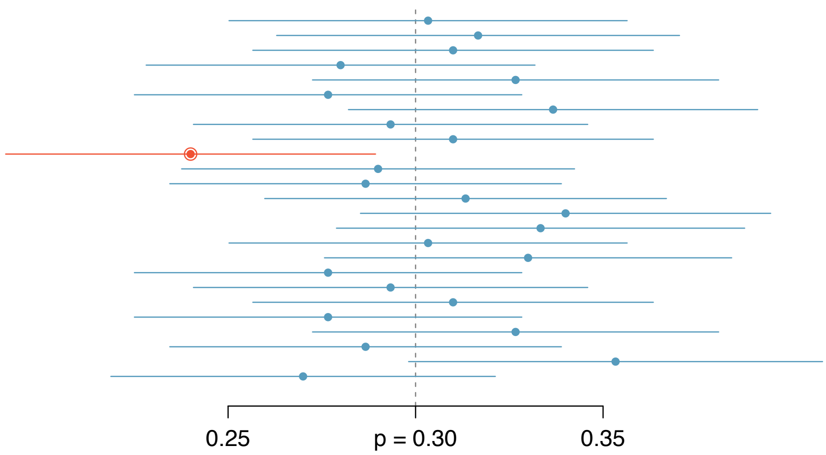

Figure 2.37: Twenty-five samples of size $n = 300$ were simulated when $p = 0.30$. For each sample, a confidence interval was created to try to capture the true proportion p. However, 1 of these 25 intervals did not capture $p = 0.30$.

⊙ **Guided Practice 2.64** In Figure 2.37, one interval does not contain the true proportion, $p = 0.3$. Does this imply that there was a problem with the simulations run?[47]

2.8.3 Changing the confidence level

Suppose we want to consider confidence intervals where the confidence level is somewhat higher than 95%: perhaps we would like a confidence level of 99%. Think back to the analogy about trying to catch a fish: if we want to be more sure that we will catch the fish, we should use a wider net. To create a 99% confidence level, we must also widen our 95% interval. On the other hand, if we want an interval with lower confidence, such as 90%, we could make our original 95% interval slightly slimmer.

The 95% confidence interval structure provides guidance in how to make intervals with new confidence levels. Below is a general 95% confidence interval for a point estimate that comes from a nearly normal distribution:

$$\text{point estimate} \ \pm \ 1.96 \times SE \tag{2.65}$$

There are three components to this interval: the point estimate, "1.96", and the standard error. The choice of $1.96 \times SE$ was based on capturing 95% of the data since the estimate is within 1.96 standard errors of the true value about 95% of the time. The choice of 1.96 corresponds to a 95% confidence level.

⊙ **Guided Practice 2.66** If X is a normally distributed random variable, how often will X be within 2.58 standard deviations of the mean?[48]

To create a 99% confidence interval, change 1.96 in the 95% confidence interval formula to be 2.58. Guided Practice 2.66 highlights that 99% of the time a normal random variable will be within 2.58 standard deviations of its mean. This approach – using the Z scores in the normal model to compute confidence levels – is appropriate when the point estimate is associated with a normal distribution and we can properly compute the standard error. Thus, the formula for a 99% confidence interval is

$$\text{point estimate} \ \pm \ 2.58 \times SE \tag{2.67}$$

The normal approximation is crucial to the precision of these confidence intervals. The next two chapters provides detailed discussions about when the normal model can safely be applied to a variety of situations. When the normal model is not a good fit, we will use alternative distributions that better characterize the sampling distribution.

---

[47] No. Just as some observations occur more than 1.96 standard deviations from the mean, some point estimates will be more than 1.96 standard errors from the parameter. A confidence interval only provides a plausible range of values for a parameter. While we might say other values are implausible based on the data, this does not mean they are impossible.

[48] This is equivalent to asking how often the Z score will be larger than -2.58 but less than 2.58. (For a picture, see Figure 2.38.) To determine this probability, look up -2.58 and 2.58 in the normal probability table (0.0049 and 0.9951). Thus, there is a $0.9951 - 0.0049 \approx 0.99$ probability that the unobserved random variable X will be within 2.58 standard deviations of the mean.

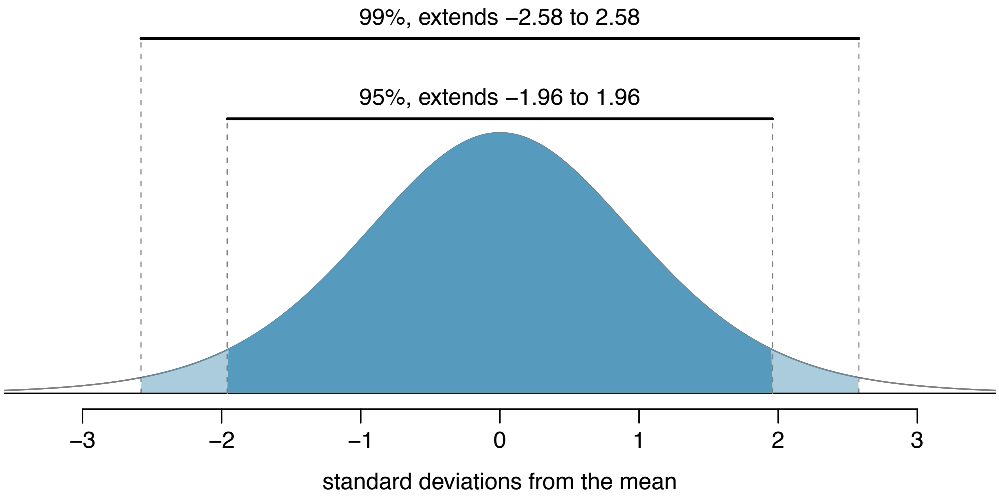

Figure 2.38: The area between -$z^\star$ and $z^\star$ increases as $|z^\star|$ becomes larger. If the confidence level is 99%, we choose $z^\star$ such that 99% of the normal curve is between -$z^\star$ and $z^\star$, which corresponds to 0.5% in the lower tail and 0.5% in the upper tail: $z^\star = 2.58$.

⊙ **Guided Practice 2.68** Create a 99% confidence interval for the impact of the stent on the risk of stroke using the data from Example 2.63. The point estimate is 0.090, and the standard error is $SE = 0.028$. It has been verified for you that the point estimate can reasonably be modeled by a normal distribution.[49]

Confidence interval for any confidence level

If the point estimate follows the normal model with standard error SE, then a confidence interval for the population parameter is

$$\text{point estimate} \ \pm \ z^\star \times SE$$

where $z^\star$ corresponds to the confidence level selected.

Figure 2.38 provides a picture of how to identify $z^\star$ based on a confidence level. We select $z^\star$ so that the area between -$z^\star$ and $z^\star$ in the normal model corresponds to the confidence level.

Margin of error

In a confidence interval, $z^\star \times SE$ is called the **margin of error**.

[49]Since the necessary conditions for applying the normal model have already been checked for us, we can go straight to the construction of the confidence interval: point estimate $\pm\ 2.58 \times SE \rightarrow (0.018, 0.162)$. We are 99% confident that implanting a stent in the brain of a patient who is at risk of stroke increases the risk of stroke within 30 days by a rate of 0.018 to 0.162 (assuming the patients are representative of the population).

Guided Practice 2.69 In Example 2.63 we found that implanting a stent in the brain of a patient at risk for a stroke *increased* the risk of a stroke. The study estimated a 9% increase in the number of patients who had a stroke, and the standard error of this estimate was about $SE = 2.8\%$. Compute a 90% confidence interval for the effect.[50]

2.8.4 Interpreting confidence intervals

A careful eye might have observed the somewhat awkward language used to describe confidence intervals. Correct interpretation:

> We are XX% confident that the population parameter is between...

Incorrect language might try to describe the confidence interval as capturing the population parameter with a certain probability. This is one of the most common errors: while it might be useful to think of it as a probability, the confidence level only quantifies how plausible it is that the parameter is in the interval.

Another especially important consideration of confidence intervals is that they *only try to capture the population parameter*. Our intervals say nothing about the confidence of capturing individual observations, a proportion of the observations, or about capturing point estimates. Confidence intervals only attempt to capture population parameters.

[50] We must find $z^\star$ such that 90% of the distribution falls between -$z^\star$ and $z^\star$ in the standard normal model, $N(\mu = 0, \sigma = 1)$. We can look up -$z^\star$ in the normal probability table by looking for a lower tail of 5% (the other 5% is in the upper tail), thus $z^\star = 1.65$. The 90% confidence interval can then be computed as point estimate $\pm\ 1.65 \times SE \to (4.4\%, 13.6\%)$. (Note: the conditions for normality had earlier been confirmed for us.) That is, we are 90% confident that implanting a stent in a stroke patient's brain increased the risk of stroke within 30 days by 4.4% to 13.6%.