
MODULE 14

CONFIDENCE REGIONS

Objectives:

1. Describe the concept underlying confidence intervals.
2. Construct confidence intervals for parameters.
3. Use the confidence interval formula to estimate desired sample sizes.

Contents

14.1 Confidence Concept	179
14.2 Constructing Confidence Regions	182
14.3 Inference Type Relationship	185
14.4 Precision and Sample Size	186

THE FINAL RESULT FROM A HYPOTHESIS TEST (Module 13) can feel uneventful – i.e., either conclude that the parameter may be equal to or different from the hypothesized value.¹ If the parameter is thought to be different from the hypothesized value we might then say that our best guess at the parameter is the observed statistic. However, as shown in Module 12, a statistic is an imperfect estimate of the unknown parameter because of sampling variability. This imperfectness can be recognized by computing a range of values that is likely to contain the parameter. For example, we may make a statement such as this – “Our best guess for the true population mean length of fish in Square Lake is the sample mean of 98.5 mm; however, we are 95% confident that the mean of ALL fish in the lake is between 95.9 and 101.1 mm.” The range in the last phrase that acknowledges sampling variability is called a confidence region. In this module, the concept, calculation, and interpretation of confidence regions is explored.

¹Depending on the H_A it may be known if the parameter is more or less than the hypothesized value.

14.1 Confidence Concept

An understanding of what it means to be “95% confident” requires examination of multiple samples from a population as was done when considering sampling variability in Module 12. In this initial discussion, only 95% confidence intervals (CI) where a range (i.e., bounded on both ends) is computed are considered. These simplifying restrictions and unrealistically knowing population values are used here only so that the **concept** of confidence intervals can be more easily explored. General methods for constructing other types of confidence regions with other levels of confidence are in Section 14.2.

In the Square Lake example (introduced in Module 1), it was known that $\mu=98.06$ and $\sigma=31.49$ (Table 2.1) and $\bar{x} = 100.04$ from the first sample of $n=50$ (Table 2.2). A 95% CI for μ is defined as $\bar{x} \pm 2SE_{\bar{x}}$. Thus, the 95% CI for the mean total length for the Square Lake population is $100.04 \pm 2\frac{31.49}{\sqrt{50}}$, 100.04 ± 8.91 , or $(91.13, 108.95)$. In this exploratory example, this interval does contain μ . In other words, this particular CI accomplished what it was intended to do, i.e., provide a range that contains μ .

Despite the success observed in this first sample, not all confidence intervals will contain μ . For example, four of 100 95% confidence intervals shown in Figure 14.1 did not contain μ . Thus, the researcher would have concluded that μ was in an incorrect interval four times in these 100 samples. The concept of “confidence” in confidence regions is related to determining how often the intervals correctly contain the parameter.

From the Central Limit Theorem, the sampling distribution of \bar{x} for samples of $n=50$ is $N(98.06, \frac{31.49}{\sqrt{50}})$ or $N(98.06, 4.45)$ for this known population. According to the 68-95-99.7% Rule, it is known that 95% of the sample means in this sampling distribution will be between $\mu \pm 2SE$ or, in this specific case, between $98.06 \pm 2(4.45)$. The sampling distribution and this range of expected sample means is shown at the top of Figure 14.1. In addition, the range of expected sample means is extended down through all of the CI lines in Figure 14.1. Note that any sample that produced a mean (solid dot on the CI line) inside the expected range of sample means also produced a 95% CI that contained μ (i.e., blue CI line). Because 95% of the sample means will be within the expected range of sample means, 95% of the 95% CIs will contain μ . So, “95% confident” means that 95% of all 95% CIs will contain the parameter and 5% will not. In other words, the mistake identified above will be made with 5% of all 95% confidence intervals.

The specifics for constructing confidence regions with different levels of confidence will be described below. However, at this point, it should be noted that the number of CIs expected to contain the parameter of interest is set by the level of confidence used to construct the CI. For example, 80% of 80% CIs and 90% of 90% CIs will contain the parameter of interest. In either case, a particular CI either does or does not contain the interval and, in real-life, we will never know whether it does or does not (i.e., we won’t know the value of the parameter). However, we do know that the technique (i.e., the construction of the CI) will “work” (i.e., contain the parameter) a set percentage of the time. To reiterate this point, examine the 100 90% CIs (Figure 14.2-Left) and 100 80% CIs (Figure 14.2-Right) for the Square Lake fish length data.

◊ **The number of confidence intervals expected to contain the parameter of interest is set by the level of confidence used to construct the confidence interval.**

One should consider the following subtleties when considering the concept of a confidence region,

- A CI is a random variable like any other statistic. That is, each sample results in a different 95% CI (see CI lines in Figure 14.1) just like it results in a different \bar{x} (see dots on CI lines in Figure 14.1).
- Any CI either contains the parameter (e.g., μ) or not. However, on average, 95% of 95% CIs will contain the parameter and 5% will not. That is, 95% of all possible 95% CIs will contain the parameter.
- A 95% CI is a technique that “works correctly” 95% of the time. In other words, 95% of all 95% CI “capture” the unknown parameter.

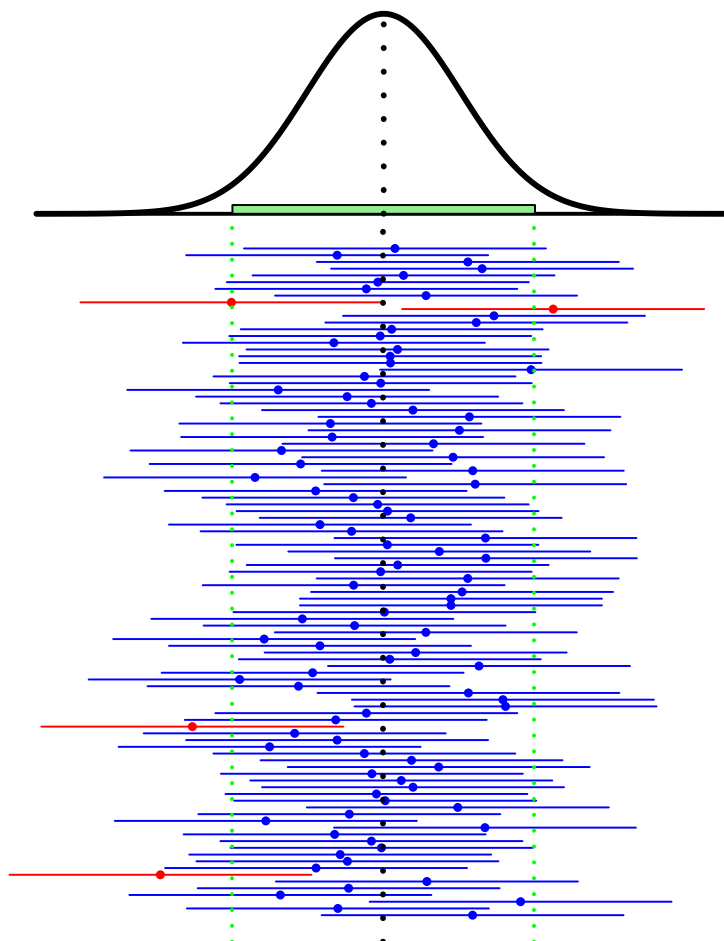


Figure 14.1. Sampling distribution of the sample mean (top) and 100 random 95% confidence intervals (horizontal lines) from samples of $n=50$ from the Square Lake population. Confidence intervals that do NOT contain $\mu=98.06$ are shown in red.

Because of these subtleties, confidence regions are often misinterpreted. Common misinterpretations are listed below with an explanation for the misinterpretation in parentheses. These misinterpretations should be studied, compared to the interpretations discussed above, and avoided.

1. “There is a 95% probability that the population mean is contained in the confidence interval.” [*This is incorrect because the population mean is constant (not random), it either is or is not in a particular CI, and it will never change whether it is or is not in that CI. The CI, not the parameter, is random.*]
2. “95% of all 95% confidence intervals are contained in the confidence interval.” [*First, this is physically impossible at this point (i.e., using Z^*) because each CI is the same width (if n and the level of confidence stay constant). Second, it is not important how many CI are contained in a CI; interest is in whether the parameter is in the interval or not.*]
3. “There is a 95% probability that the sample mean is contained in the confidence interval.” [*This is incorrect for the simple fact that CI are not used to estimate sample means (or, generally, statistics); they are used to estimate population means (or parameters). Furthermore, the sample mean has to be exactly in the middle of the CI (see next section).*]

◇ Confidence intervals are constructed for parameters, not statistics.

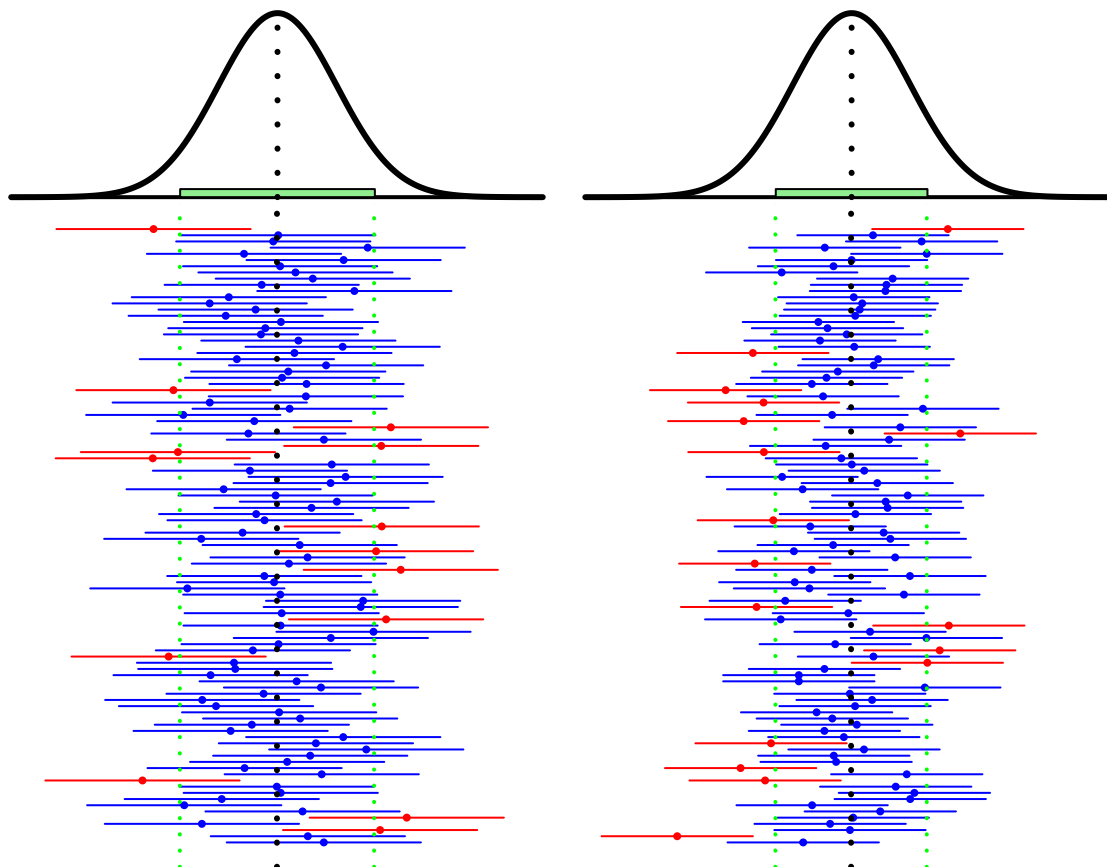


Figure 14.2. Sampling distribution of the sample mean (**tops**) and 100 random 90% (**Left**) and 80% (**Right**) confidence intervals (horizontal lines) from samples of $n=50$ from the Square Lake population. Confidence intervals that do NOT contain μ are shown in red.

◇ Care and specificity must be used when interpreting and describing confidence intervals.

Review Exercises

- 14.1** True or False – A 95% confidence region can be constructed for \bar{x} ? [Answer](#)
- 14.2** True or False – A 95% confidence region can be constructed for the population median? [Answer](#)
- 14.3** True or False – A 95% confidence region can be constructed for σ ? [Answer](#)
- 14.4** Yes, No, Can't tell – Is the estimated parameter in the CI: (111.12, 123.32). [Answer](#)
- 14.5** Replace “XXX” to make this statement correct – “I am 99% confident that the XXX of interest is in my confidence interval?” [Answer](#)

14.2 Constructing Confidence Regions

Not all confidence regions are designed to contain the parameter 95% “of the time,” are intervals, or are computed to contain μ . Confidence regions can be constructed for any level of confidence, as intervals or bounds, and for nearly all **parameters**.

The level of confidence (C) used will be determined by the α chosen for the hypothesis test. Specifically, the level of confidence will be $100(1 - \alpha)\%$. For example, if α is set at 0.05, then the level of confidence will be 95% or if α is set at 0.01, then a 99% level of confidence will be used. From this, one can see that if α is decreased such that fewer Type I errors are made, then the confidence level will increase and more of the confidence regions will contain the parameter of interest (i.e., fewer errors). In this manner the proportion of Type I errors in hypothesis testing is linked to the proportion of errors made with confidence regions.

◇ The level of confidence (C) is determined from α ; i.e., $C = 100(1 - \alpha)\%$.

The type of confidence region depends on the type of alternative hypothesis. If the alternative hypothesis is two-tailed (i.e., \neq), then the confidence region will be an interval (i.e., a range will be computed, as in Section 14.1). However, if the alternative hypothesis is one-tailed, then a confidence bound is used. For example, if the alternative hypothesis is a “less than”, then interest lies in determining what is the “largest possible value” for the parameter (rather than a range of possible values). In other words, if the alternative hypothesis is a “less than”, then an upper confidence bound for the parameter is constructed. In contrast, if the alternative hypothesis is a “greater than”, then a lower confidence bound is constructed to estimate the “smallest possible value” for the parameter.

◇ A confidence interval should be constructed when a two-tailed H_A is used.

◇ A confidence bound should be constructed when H_A is one-tailed. If H_A is a “greater than”, then the smallest possible value of the parameter is sought and a lower bound is constructed. If H_A is a “less than”, then the largest possible value of the parameter is sought and an upper bound is constructed.

Fortunately, most confidence regions follow the same basic form of

$$\text{“Statistic”} + \text{“scaling factor”} * SE_{\text{statistic}}$$

where “Statistic” represents the statistic used to estimate the parameter, $SE_{\text{statistic}}$ is the standard error of that statistic, and “scaling factor” * $SE_{\text{statistic}}$ is called the margin-of-error. The scaling factor is computed from a known distribution. When σ is known, the scaling factor is computed from a $N(0, 1)$ and is called Z^* . Thus, in the case when a confidence interval is being constructed for μ and σ is known, the specific formula for the confidence region is

$$\bar{x} + z^* \frac{\sigma}{\sqrt{n}}$$

The “scaling factor” serves to control the width and type of confidence region. The magnitude of the scaling factor controls the relative width of the region such that the parameter is contained in the region at a rate according to the level of confidence. For example, the scaling factor for a 99% confidence region will be set such that 99% of the confidence regions will contain the parameter.

The sign of the scaling factor controls whether an interval, upper bound, or lower bound is computed. For example, if the alternative hypothesis is two-tailed, then Z^* is the two values such that an area equal to

the level of confidence is contained between them (Figure 14.3-Left). The two values that delineate these boundaries will be the same value but with different signs because the $N(0,1)$ is symmetric about zero. Thus, a confidence interval is computed with a scaling factor of $\pm Z^*$.

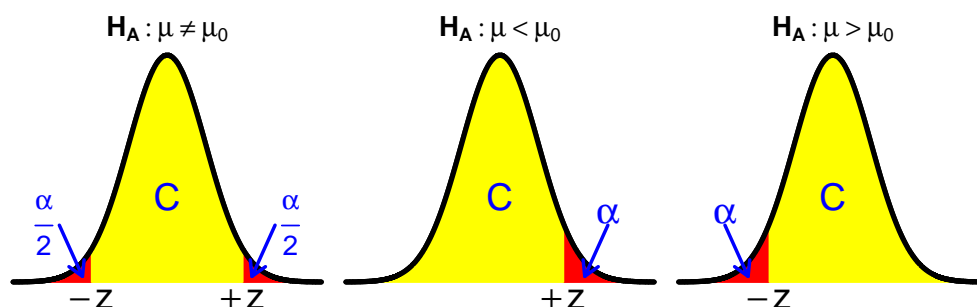


Figure 14.3. Areas (yellow) that define Z^* for confidence regions of a parameter in a hypothesis test.

In contrast, if the alternative hypothesis is a “less than”, then an upper confidence bound is desired and Z^* has an area equal to the level of confidence LESS THAN it (Figure 14.3-Middle). As the level of confidence will always be greater than 50%, this definition will produce a positive Z^* so that the scaling factor will be $+Z^*$. Similarly, if the alternative hypothesis is a “greater than”, then a lower confidence bound is desired and Z^* has an area equal to the level of confidence GREATER THAN it (Figure 14.3-Right). This definition produces a negative Z^* so that the scaling factor will be $-Z^*$.

◇ Confidence intervals can be constructed for any level of confidence and for nearly every parameter.

◇ When finding Z^* for a confidence bound, the level of confidence always represents an area shaded in the same direction as the sign in H_A .

The following are three examples for calculating confidence regions.

1. For the Square Lake example where $H_A : \mu < 105$ and $\alpha = 0.05$ a 95% upper confidence bound is constructed. The corresponding $Z^* = +1.645$ is found with²

```
> ( distrib(0.95,type="q") )
[1] 1.644854
```

Thus, with the $\bar{x}=100.04$ from $n = 50$ (Table 2.2), the 95% upper confidence bound is $100.04 + 1.645 \frac{31.49}{\sqrt{50}}$, $100.04 + 7.33$, or 107.37. Thus, one is 95% confident that the mean total length of ALL fish in Square Lake is less than 107.4 mm. By confident, it is meant that 95% of all 95% confidence regions will contain μ .

2. Suppose that the mouse water consumption data from Table 5.1 was tested with $H_A : \mu \neq 10$ and $\alpha = 0.01$. In this case, a 99% confidence interval is computed with $Z^* = \pm 2.576$ as found with

²Note that the default settings for `distrib()` are to use `mean=0` and `sd=1`. As these are the defaults, they can be omitted when finding a Z^* .

```
> ( distrib(0.995,type="q") )
[1] 2.575829
```










Thus, assuming that $\sigma = 2$ ml and $\bar{x} = 14.04$ (from Section 5.5) the 99% confidence interval is $14.04 \pm 2.576 \frac{2}{\sqrt{30}}$, 14.04 ± 0.94 , or (13.10, 14.98). Thus, one is 99% confident that the mean level of water consumption by ALL mice is between 13.1 and 15.0 ml. By confident, it is meant that 99% of all 99% confidence regions will contain μ .


3. Suppose that the second example hypothesis test in Module 13 about battery life (i.e., $H_A : \mu > 36$ vs $H_0 : \mu = 36$) is being tested with $\alpha = 0.10$. Further suppose that $\sigma = 7$ and that $\bar{x} = 45$ from $n = 40$. In this case, a 90% lower confidence bound is constructed using $Z^* = -1.282$ as found with


```
> ( distrib(0.90,type="q",lower.tail=FALSE) )
[1] -1.281552
```


Thus, the 90% lower confidence bound is $45 - 1.282 \frac{7}{\sqrt{40}}$, $45 - 1.42$, or 43.58 months. Thus, one is 90% confident that the mean life for ALL batteries with the additive is more than 43.58 months. By confident, it is meant that 90% of all 90% confidence regions will contain μ .

Review Exercises

- 14.6  What is Z^* for a 99% confidence interval? [Answer](#)
- 14.7  What is Z^* for a 92% lower confidence bound? [Answer](#)
- 14.8  What is Z^* for a 90% upper confidence bound? [Answer](#)
- 14.9  What is Z^* for a 98% confidence interval? [Answer](#)
- 14.10  What is Z^* for a 95% lower confidence bound? [Answer](#)
- 14.11  What is Z^* for a 70% upper confidence bound? [Answer](#)
- 14.12  Construct and interpret (including describing what is meant by “confidence”) a proper confidence region for the mean BOD level presented in Review Exercise 13.8. [Answer](#)
- 14.13  Construct and interpret (including describing what is meant by “confidence”) a proper confidence region for the mean grade point average presented in Review Exercise 13.9. [Answer](#)
- 14.14  Construct and interpret (including describing what is meant by “confidence”) a proper confidence region if H_A is a “not equals” and $\alpha = 0.05$ for the population mean gage height on the Bois Brule River presented in Review Exercise 5.13 assuming that the population standard deviation is 0.20 feet and the sampling distribution is approximately normal. [Answer](#)

14.15  Construct and interpret (including describing what is meant by “confidence”) a proper confidence region if H_A is a “less than” and $\alpha=0.10$ for the mean population density of all counties in Wisconsin using the data presented in Review Exercise 5.14 assuming that $\sigma = 125$ people/land acre and the sampling distribution is approximately normal. [Answer](#)

14.16  Construct and interpret (including describing what is meant by “confidence”) a proper confidence region if H_A is a “greater than” and $\alpha=0.05$ for the population mean creatine phosphokinase value using the data presented in Review Exercise 5.5 assuming that $\sigma = 40$. [Answer](#)

14.17  Hebblewhite (2000) reported the mean snow pack height (in cm) for Banff (data are below). These data were strongly right-skewed with a possible outlier at the maximum. Assume that it is known that $\sigma=15$ cm. (A) Compute a 99% confidence interval for the mean snow pack height. (B) In addition, comment on whether or not a confidence interval should be computed for these data (note: compute the CI in (A) regardless of your answer here). [Answer](#)

29.00, 45.51, 30.18, 45.83, 39.54, 80.39, 32.64, 32.89,
46.84, 45.79, 62.92, 67.24, 30.96, 46.08, 33.28

14.3 Hypothesis Tests and Confidence Region Relationship

An alternative conceptualization of confidence intervals can show how confidence regions and hypothesis tests are related. This conceptualization rests on considering the sample means that would be “reasonable to see” from populations with various values of μ . A graphic is constructed below using the Square Lake population as an example and assuming that σ is known ($=31.49$), $n = 50$, and 95% CIs are used.

First, compute the most common 95% of sample means assuming that $\mu = 70$; i.e., $70 \pm 1.960 \frac{31.49}{\sqrt{50}}$, 70 ± 8.73 , or (61.27, 78.73). This range is plotted as a vertical rectangle centered $\mu = 70$ (left-most rectangle) in Figure 14.4-Left). Next, compute and plot the same range for a slightly larger μ (e.g., with $\mu = 71$, plot (62.27, 78.73)). Then repeat these steps for sequentially larger values of μ until a plot similar to Figure 14.4 is constructed.

Consider very carefully what Figure 14.4 represents. The vertical rectangles represent the ranges of the most common 95% of sample means (values read from the y-axis) that will be produced for a particular population mean (value read from the x-axis). In essence, each vertical line represents the sample means that are likely to be observed from a population with a given population mean (x-axis).

Now suppose that $\bar{x}=100.04$ is observed as in Table 2.2. Draw a horizontal line across Figure 14.4 at this value and then draw vertical lines down from where the horizontal line first enters and last leaves the band of possible sample means (Figure 14.4-Right). The x-axis values that these vertical lines intercept are an approximate 95% CI for μ . The approximation is only as close as the intervals used to construct the rectangles (i.e., 1.0 mm were used here). However, the results from this graphical approach (i.e., (92, 108)) compare favorably to the previous results from using the CI formula (i.e., (91.27, 108.73)).

Surely, the CI formula (Section 14.2) is a more efficient and precise way to construct confidence intervals. However, this conceptualization illustrates that a confidence interval (or region, more generally) consists of population means that are likely to produce the observed sample mean. Thus, a confidence region represents possible null hypothesized population means that WOULD NOT BE rejected during hypothesis testing.

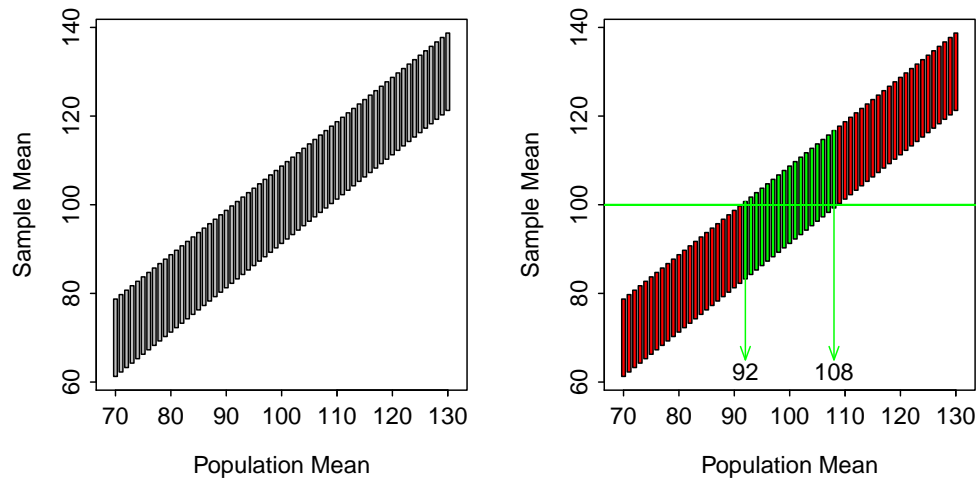


Figure 14.4. Range (95%) of sample means that would be produced by particular population means in the Square Lake fish length example (**Left**) and the ranges intercepted by $\bar{x} = 100.04$ mm (**Right**).

◇ A confidence region represents population means that would likely have produced the observed sample mean.

◇ A confidence region represents null hypothesized population means that would NOT be rejected.

14.4 Precision and Sample Size

The width of a confidence interval explains how precisely the parameter is estimated. For example, narrow intervals represent precise estimates of the parameter. The width of a confidence interval depends on the margin-of-error which depends on (1) the standard error and (2) the scaling factor. As either of these two items gets smaller (while holding the other constant), the width of the confidence interval gets smaller.

◇ The width of a confidence interval is a measure how precisely the parameter is estimated.

◇ The width of a confidence interval depends on the standard error of the statistic and the scaling factor used.

A small standard error means that sampling variability is low and the parameter is precisely estimated by the statistic. Smaller standard errors are obtained only by increasing the sample size. A smaller standard deviation would also result in a smaller SE, but the standard deviation cannot be made smaller (i.e., it is an inherent characteristic of the population).

◇ Confidence intervals can be made narrower by increasing the sample size.

A smaller scaling factor is obtained by reducing the level of confidence. For example, a 90% confidence interval uses a $Z^* = \pm 1.645$ whereas a 95% confidence interval uses a $Z^* = \pm 1.960$ (as shown previously). Thus, decreasing the confidence level narrows the CI. However, reducing the level of confidence will also increase the number of confidence intervals that do not contain the parameter. Thus, reducing the level of confidence may not be the best choice for narrowing the confidence interval.

◇ **Confidence intervals can (but generally should not) be made narrower by decreasing the level of confidence.**

The relationship between the margin-of-error (m.e.; i.e., width of a confidence interval) and sample size (n) provides a means for computing the n required to estimate μ within $\pm m.e.$ units with C% confidence assuming that σ is known. Specifically, the margin-of-error formula can be solved for n .

$$\begin{aligned} m.e. &= z * \frac{\sigma}{\sqrt{n}} \\ \sqrt{n} &= \frac{z * \sigma}{m.e.} \\ n &= \left(\frac{z * \sigma}{m.e.} \right)^2 \end{aligned}$$

For example, suppose that one wants the sample size required to estimate the mean length of fish in Square Lake to within 5 mm with 90% confidence knowing that the population standard deviation is 34.91. From this, $m.e.=5$, $\sigma=34.91$, and $Z^*=1.645$ (found previously for 90% confidence).³ Thus, $n = \left(\frac{1.645 * 34.91}{5} \right)^2 = 131.91$. Therefore, a sample of at least 132 fish from Square Lake should be taken to meet these constraints. Note that sample size calculations are always rounded up to the next integer because rounding down would produce a sample size that does not meet the desired criteria.

◇ **Always round sample size calculations up to the next integer.**

The margin-of-error and confidence level in these calculations need to come from the researcher's beliefs about how much error they can live with (i.e., chance that a confidence interval does not contain the parameter) and how precise their estimate of the mean needs to be. Values for σ are rarely known in practice (because it is a parameter) and estimates from preliminary studies, previous similar studies, similar populations, or best guesses are often used instead. In practice, a researcher will often prepare a graph with varying values of σ to make an informed decision of what sample size to choose (Figure 14.5).

³Strictly, $Z^* \pm 1.645$, but the sign is inconsequential due to squaring in the sample size formula.

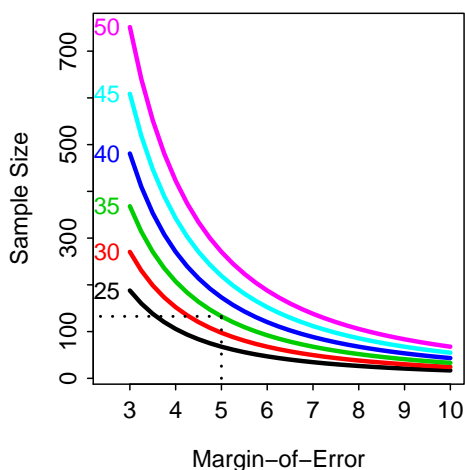


Figure 14.5. Desired sample size versus margin-of-error for constant values of σ (shown to the left of each line) and $C = 90$. The desired sample size for $\text{m.e.} = 5$, $\sigma = 35$, and $C = 90$ is illustrated with the black dotted lines.

Review Exercises

- 14.18** If two populations have the same standard deviation and a sample of size 30 is taken from population A and a sample of size 50 from population B, which will have a narrower CI? [Answer](#)
- 14.19** If the same size of sample is taken from two populations, but Population C has a smaller standard deviation than Population D, which will have a narrower CI? [Answer](#)
- 14.20** From the same data, is a 95% or a 99% CI narrower? [Answer](#)
- 14.21** Describe how the margin of error will change as each of the following change (all others held constant): confidence level (C), z^* , n , σ , μ , and \bar{x} (in the case of CIs for μ). Make sure to explain your reasoning for each. [Answer](#)
- 14.22** Geographers measure the longest axis of pebbles to determine “grain” sizes. If the standard deviation of pebble long-axis length for a particular site is known to be 4 mm, how many pebbles must be measured in order to determine the average pebble length within 0.1 mm with 99% confidence? [Answer](#)
- 14.23** An investment group wants to start an Internet Service Provider (ISP) and, for their business plan and model, needs to estimate the average Internet usage of households. How many households must be randomly selected to be 95% sure that the sample mean is within 1 minute of the population mean? Assume that a previous survey of household usage had a standard deviation of 6.95 minutes. [Answer](#)

Part IV

Specific Hypothesis Tests