Confidence Intervals

Prof. Wells

Math 209, 3/10/23

Outline

In this lecture, we will. . .

Outline

In this lecture, we will. . .

- Introduce confidence intervals as a method for estimating a parameter
- Interpret confidence intervals

Section 1

Confidence Intervals

Point Estimates

- To estimate a population *parameter*, we can use a sample *statistic*.

Point Estimates

- To estimate a population *parameter*, we can use a sample *statistic*.
  - Suppose you are hosting a pizza party for 200 people, and need to know what proportion p of attendees prefer vegetarian pizza.

Point Estimates

- To estimate a population *parameter*, we can use a sample *statistic*.
  - Suppose you are hosting a pizza party for 200 people, and need to know what proportion p of attendees prefer vegetarian pizza.
  - Before ordering, you can ask a random sample of attendees about their pizza preference and use the proportion $\hat{p}$ in the sample as an estimate for the total proportion p.

Point Estimates

- To estimate a population *parameter*, we can use a sample *statistic*.
  - Suppose you are hosting a pizza party for 200 people, and need to know what proportion p of attendees prefer vegetarian pizza.
  - Before ordering, you can ask a random sample of attendees about their pizza preference and use the proportion $\hat{p}$ in the sample as an estimate for the total proportion p.
  - Suppose after polling 9 pizza party attendees, you find $\hat{p} = \frac{1}{3}$.

Point Estimates

- To estimate a population *parameter*, we can use a sample *statistic*.
  - Suppose you are hosting a pizza party for 200 people, and need to know what proportion p of attendees prefer vegetarian pizza.
  - Before ordering, you can ask a random sample of attendees about their pizza preference and use the proportion $\hat{p}$ in the sample as an estimate for the total proportion p.
  - Suppose after polling 9 pizza party attendees, you find $\hat{p} = \frac{1}{3}$.
- The sample statistic is your best guess for the population parameter, but...

Point Estimates

- To estimate a population *parameter*, we can use a sample *statistic*.
  - Suppose you are hosting a pizza party for 200 people, and need to know what proportion p of attendees prefer vegetarian pizza.
  - Before ordering, you can ask a random sample of attendees about their pizza preference and use the proportion $\hat{p}$ in the sample as an estimate for the total proportion p.
  - Suppose after polling 9 pizza party attendees, you find $\hat{p} = \frac{1}{3}$.
- The sample statistic is your best guess for the population parameter, but...
  - How certain should you be in this estimate? Would you be willing to claim that the true proportion p is **exactly** 0.33?

Point Estimates

- To estimate a population *parameter*, we can use a sample *statistic*.
  - Suppose you are hosting a pizza party for 200 people, and need to know what proportion p of attendees prefer vegetarian pizza.
  - Before ordering, you can ask a random sample of attendees about their pizza preference and use the proportion $\hat{p}$ in the sample as an estimate for the total proportion p.
  - Suppose after polling 9 pizza party attendees, you find $\hat{p} = \frac{1}{3}$.
- The sample statistic is your best guess for the population parameter, but...
  - How certain should you be in this estimate? Would you be willing to claim that the true proportion p is **exactly** 0.33?
  - In fact, it's numerically impossible for $p = \frac{1}{3}$. (Why?)

Point Estimates

- To estimate a population *parameter*, we can use a sample *statistic*.
  - Suppose you are hosting a pizza party for 200 people, and need to know what proportion p of attendees prefer vegetarian pizza.
  - Before ordering, you can ask a random sample of attendees about their pizza preference and use the proportion $\hat{p}$ in the sample as an estimate for the total proportion p.
  - Suppose after polling 9 pizza party attendees, you find $\hat{p} = \frac{1}{3}$.
- The sample statistic is your best guess for the population parameter, but. . .
  - How certain should you be in this estimate? Would you be willing to claim that the true proportion p is **exactly** 0.33?
  - In fact, it's numerically impossible for $p = \frac{1}{3}$. (Why?)
- Moreover, if you were to collect another random sample, it's likely you would obtain a different value of $\hat{p}$.

Point Estimates

- To estimate a population *parameter*, we can use a sample *statistic*.
  - Suppose you are hosting a pizza party for 200 people, and need to know what proportion p of attendees prefer vegetarian pizza.
  - Before ordering, you can ask a random sample of attendees about their pizza preference and use the proportion $\hat{p}$ in the sample as an estimate for the total proportion p.
  - Suppose after polling 9 pizza party attendees, you find $\hat{p} = \frac{1}{3}$.
- The sample statistic is your best guess for the population parameter, but...
  - How certain should you be in this estimate? Would you be willing to claim that the true proportion p is **exactly** 0.33?
  - In fact, it's numerically impossible for $p = \frac{1}{3}$. (Why?)
- Moreover, if you were to collect another random sample, it's likely you would obtain a different value of $\hat{p}$.
  - In fact, due to the randomness of sampling, it's fairly unlikely that your sample **exactly** matches the population.

Point Estimates

- To estimate a population *parameter*, we can use a sample *statistic*.
  - Suppose you are hosting a pizza party for 200 people, and need to know what proportion p of attendees prefer vegetarian pizza.
  - Before ordering, you can ask a random sample of attendees about their pizza preference and use the proportion $\hat{p}$ in the sample as an estimate for the total proportion p.
  - Suppose after polling 9 pizza party attendees, you find $\hat{p} = \frac{1}{3}$.
- The sample statistic is your best guess for the population parameter, but. . .
  - How certain should you be in this estimate? Would you be willing to claim that the true proportion p is **exactly** 0.33?
  - In fact, it's numerically impossible for $p = \frac{1}{3}$. (Why?)
- Moreover, if you were to collect another random sample, it's likely you would obtain a different value of $\hat{p}$.
  - In fact, due to the randomness of sampling, it's fairly unlikely that your sample **exactly** matches the population.
  - So if it's unlikely that $p = \hat{p}$, what do we do?

Interval Estimates

- Instead of providing a single estimate for p, it might be preferable to estimate p as an *interval* of plausible values.

Interval Estimates

- Instead of providing a single estimate for p, it might be preferable to estimate p as an *interval* of plausible values.
  - Having observed $\hat{p} = 0.33$, we might guess that the true value of p is *close* to $\hat{p}$

Interval Estimates

- Instead of providing a single estimate for p, it might be preferable to estimate p as an *interval* of plausible values.
    - Having observed $\hat{p} = 0.33$, we might guess that the true value of p is *close* to $\hat{p}$
    - Perhaps we might say p is within 0.33 ± 0.2; that is, between 0.13 and 0.53.

Interval Estimates

- Instead of providing a single estimate for p, it might be preferable to estimate p as an *interval* of plausible values.
    - Having observed $\hat{p} = 0.33$, we might guess that the true value of p is *close* to $\hat{p}$
    - Perhaps we might say p is within 0.33 ± 0.2; that is, between 0.13 and 0.53.
- Of course, the length of our interval estimate will probably depend on the sample size (and other factors)

Interval Estimates

- Instead of providing a single estimate for p, it might be preferable to estimate p as an *interval* of plausible values.
  - Having observed $\hat{p} = 0.33$, we might guess that the true value of p is *close* to $\hat{p}$
  - Perhaps we might say p is within 0.33 ± 0.2; that is, between 0.13 and 0.53.
- Of course, the length of our interval estimate will probably depend on the sample size (and other factors)
  - Larger samples will allow you to provide a smaller lengths, while smaller samples will require you to provide a larger lengths.

Interval Estimates

- Instead of providing a single estimate for p, it might be preferable to estimate p as an *interval* of plausible values.
  - Having observed $\hat{p} = 0.33$, we might guess that the true value of p is *close* to $\hat{p}$
  - Perhaps we might say p is within 0.33 ± 0.2; that is, between 0.13 and 0.53.
- Of course, the length of our interval estimate will probably depend on the sample size (and other factors)
  - Larger samples will allow you to provide a smaller lengths, while smaller samples will require you to provide a larger lengths.
  - More variability in the population might necessitate larger lengths as well.

Interval Estimates

- Instead of providing a single estimate for p, it might be preferable to estimate p as an *interval* of plausible values.
  - Having observed $\hat{p} = 0.33$, we might guess that the true value of p is *close* to $\hat{p}$
  - Perhaps we might say p is within 0.33 ± 0.2; that is, between 0.13 and 0.53.
- Of course, the length of our interval estimate will probably depend on the sample size (and other factors)
  - Larger samples will allow you to provide a smaller lengths, while smaller samples will require you to provide a larger lengths.
  - More variability in the population might necessitate larger lengths as well.
  - With just $n = 9$ people, you might estimate that p is 0.33 ± 0.3, or 0.03 to 0.63.

Interval Estimates

- Instead of providing a single estimate for p, it might be preferable to estimate p as an *interval* of plausible values.
  - Having observed $\hat{p} = 0.33$, we might guess that the true value of p is *close* to $\hat{p}$
  - Perhaps we might say p is within 0.33 ± 0.2; that is, between 0.13 and 0.53.
- Of course, the length of our interval estimate will probably depend on the sample size (and other factors)
  - Larger samples will allow you to provide a smaller lengths, while smaller samples will require you to provide a larger lengths.
  - More variability in the population might necessitate larger lengths as well.
  - With just $n = 9$ people, you might estimate that p is 0.33 ± 0.3, or 0.03 to 0.63.
  - But with $n = 60$, you might instead estimate that p is 0.33 ± 0.13, or 0.2 to 0.46.

Interval Estimates

- Instead of providing a single estimate for p, it might be preferable to estimate p as an *interval* of plausible values.
  - Having observed $\hat{p} = 0.33$, we might guess that the true value of p is *close* to $\hat{p}$
  - Perhaps we might say p is within 0.33 ± 0.2; that is, between 0.13 and 0.53.
- Of course, the length of our interval estimate will probably depend on the sample size (and other factors)
  - Larger samples will allow you to provide a smaller lengths, while smaller samples will require you to provide a larger lengths.
  - More variability in the population might necessitate larger lengths as well.
  - With just $n = 9$ people, you might estimate that p is 0.33 ± 0.3, or 0.03 to 0.63.
  - But with $n = 60$, you might instead estimate that p is 0.33 ± 0.13, or 0.2 to 0.46.
- We'll discuss how to obtain the precise sizes of these intervals later today.

Confidence Interval Estimates

A **confidence interval estimate** for a parameter takes the form

$$\text{Statistic} \;\pm\; \text{Margin of Error}$$

Confidence Interval Estimates

A **confidence interval estimate** for a parameter takes the form

$$\text{Statistic} \pm \text{Margin of Error}$$

- The confidence interval gives a range of plausible values for the parameter.

Confidence Interval Estimates

A **confidence interval estimate** for a parameter takes the form

$$\text{Statistic} \ \pm \ \text{Margin of Error}$$

- The confidence interval gives a range of plausible values for the parameter.
- The **Margin of Error** indicates the precision of our estimate

Confidence Interval Estimates

A **confidence interval estimate** for a parameter takes the form

$$\text{Statistic} \pm \text{Margin of Error}$$

- The confidence interval gives a range of plausible values for the parameter.
- The **Margin of Error** indicates the precision of our estimate
    - When sampling pizza preferences with $n = 60$, we estimate p using the interval

      $$0.33 \pm 0.13 \qquad \text{or } 0.2 \text{ to } 0.46$$

    - In this case, the *margin of error* is 0.13.

Confidence Interval Estimates

A **confidence interval estimate** for a parameter takes the form

$$\text{Statistic} \ \pm \ \text{Margin of Error}$$

- The confidence interval gives a range of plausible values for the parameter.
- The **Margin of Error** indicates the precision of our estimate
  - When sampling pizza preferences with $n = 60$, we estimate p using the interval

    $$0.33 \pm 0.13 \qquad \text{or } 0.2 \text{ to } 0.46$$

  - In this case, the *margin of error* is 0.13.
- We also report a success rate (or **confidence level**) for the estimation technique.

Confidence Interval Estimates

A **confidence interval estimate** for a parameter takes the form

$$\text{Statistic} \pm \text{Margin of Error}$$

- The confidence interval gives a range of plausible values for the parameter.
- The **Margin of Error** indicates the precision of our estimate
  - When sampling pizza preferences with $n = 60$, we estimate p using the interval

$$0.33 \pm 0.13 \qquad \text{or } 0.2 \text{ to } 0.46$$

  - In this case, the *margin of error* is 0.13.
- We also report a success rate (or **confidence level**) for the estimation technique.
  - The confidence level is the percentage of sample statistics in the sampling distribution within the margin of error distance of the true parameter.

Confidence Interval Estimates

A **confidence interval estimate** for a parameter takes the form

$$\text{Statistic} \ \pm\ \text{Margin of Error}$$

- The confidence interval gives a range of plausible values for the parameter.
- The **Margin of Error** indicates the precision of our estimate
    - When sampling pizza preferences with $n = 60$, we estimate p using the interval

      $$0.33 \pm 0.13 \qquad \text{or } 0.2 \text{ to } 0.46$$

    - In this case, the *margin of error* is 0.13.
- We also report a success rate (or **confidence level**) for the estimation technique.
    - The confidence level is the percentage of sample statistics in the sampling distribution within the margin of error distance of the true parameter.
    - At 95% confidence and a margin of error of 0.13, we are suggesting that 95% of all sample proportions are within a distance of 0.13 of p.

Confidence Interval Estimates

A **confidence interval estimate** for a parameter takes the form

$$\text{Statistic} \pm \text{Margin of Error}$$

- The confidence interval gives a range of plausible values for the parameter.
- The **Margin of Error** indicates the precision of our estimate
  - When sampling pizza preferences with $n = 60$, we estimate p using the interval

    $$0.33 \pm 0.13 \qquad \text{or } 0.2 \text{ to } 0.46$$

  - In this case, the *margin of error* is 0.13.
- We also report a success rate (or **confidence level**) for the estimation technique.
  - The confidence level is the percentage of sample statistics in the sampling distribution within the margin of error distance of the true parameter.
  - At 95% confidence and a margin of error of 0.13, we are suggesting that 95% of all sample proportions are within a distance of 0.13 of p.
- To get the margin of error and confidence level, we use the sampling distribution

The Sampling Distribution

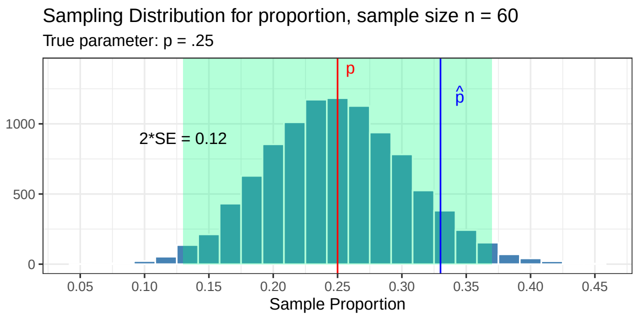
- Suppose that in truth, $p = .25$ (i.e. 25% of all party attendees prefer vegetarian)

The Sampling Distribution

- Suppose that in truth, $p = .25$ (i.e. 25% of all party attendees prefer vegetarian)
- For approximately bell-shaped sampling distributions, 95% of all sample statistics are within 2 standard errors of the parameter.

The Sampling Distribution

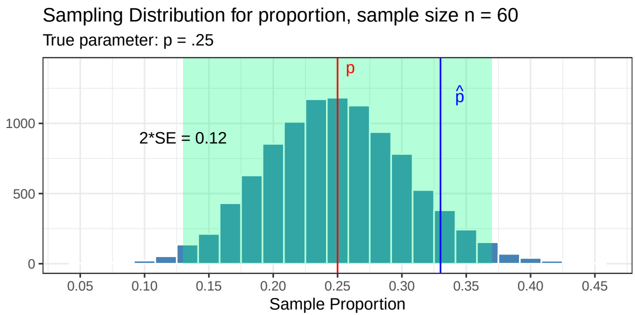
- Suppose that in truth, $p = .25$ (i.e. 25% of all party attendees prefer vegetarian)
- For approximately bell-shaped sampling distributions, 95% of all sample statistics are within 2 standard errors of the parameter.



- Note that our previously observed $\hat{p} = 0.33$ falls within the range of typical sample proportions.

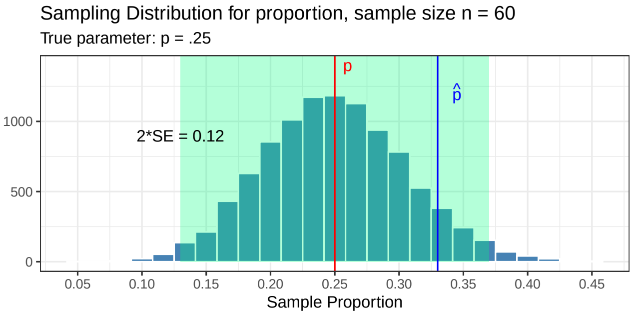
Pivoting the Center

- Consider again the sampling distribution for $\hat{p}$:



Pivoting the Center

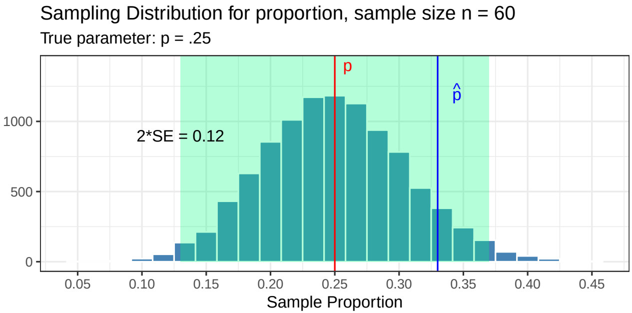
- Consider again the sampling distribution for $\hat{p}$:



- For 95% of all samples, the sample proportion is within a distance of 2 SE of the true population proportion.

Pivoting the Center

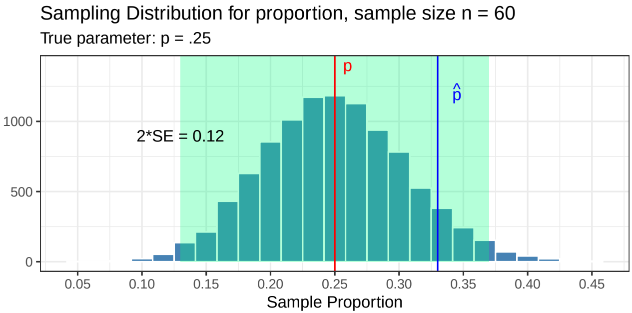
- Consider again the sampling distribution for $\hat{p}$:



- For 95% of all samples, the sample proportion is within a distance of 2 SE of the true population proportion.
- But this also means that for 95% of all samples, the true *parameter* will be within a distance of 2 SE of the sample *statistic*

Pivoting the Center

- Consider again the sampling distribution for $\hat{p}$:



- For 95% of all samples, the sample proportion is within a distance of 2 SE of the true population proportion.
- But this also means that for 95% of all samples, the true *parameter* will be within a distance of 2 SE of the sample *statistic*
  - Every sample in the green region has this property.

Interval Estimates

- To estimate the parameter, build an interval centered at the sample statistic, with a margin of error of $2 \cdot SE$:

$$\hat{p} \pm 2 \cdot \mathrm{SE}$$

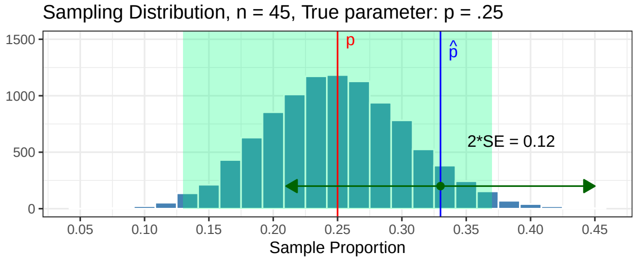
  - This interval will contain the parameter p for 95% of all samples

Interval Estimates

- To estimate the parameter, build an interval centered at the sample statistic, with a margin of error of $2 \cdot SE$:

$$\hat{p} \pm 2 \cdot \mathrm{SE}$$

  - This interval will contain the parameter p for 95% of all samples



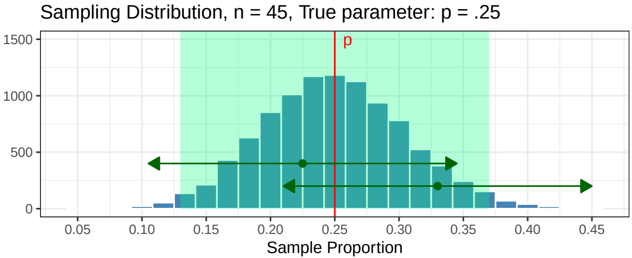
- The interval for our sample was $.33 \pm .12$, which *does* contain the parameter p

Interval Estimates

- To estimate the parameter, build an interval centered at the sample statistic, with a margin of error of $2 \cdot SE$:

$$\hat{p} \pm 2 \cdot \mathrm{SE}$$

  - This interval will contain the parameter p for 95% of all samples



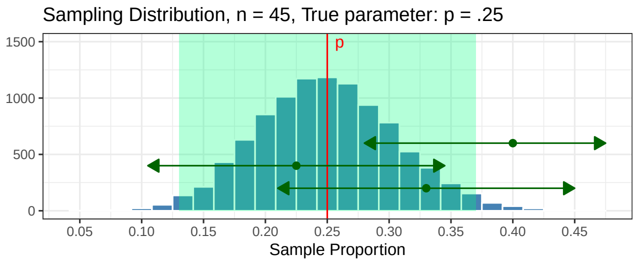
- Samples with $\hat{p}$ in the green region have intervals that also contain the parameter p

Interval Estimates

- To estimate the parameter, build an interval centered at the sample statistic, with a margin of error of $2 \cdot SE$:

$$\hat{p} \pm 2 \cdot \mathrm{SE}$$

  - This interval will contain the parameter p for 95% of all samples



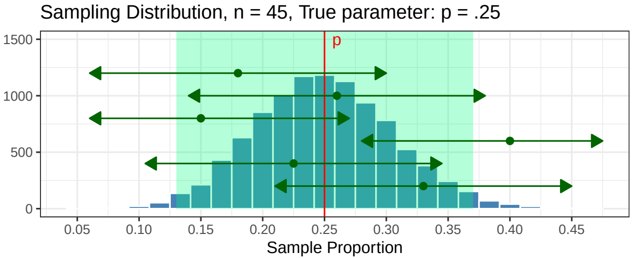
- Samples with $\hat{p}$ outside the green region have intervals that don't contain p

Interval Estimates

- To estimate the parameter, build an interval centered at the sample statistic, with a margin of error of $2 \cdot SE$:

$$\hat{p} \pm 2 \cdot \mathrm{SE}$$

  - This interval will contain the parameter p for 95% of all samples



- But 95% of all samples will have intervals that do contain p

Confidence Level

- Confidence intervals consists of both an interval estimate and a confidence level.

Confidence Level

- Confidence intervals consists of both an interval estimate and a confidence level.
  - Based on our pizza party sample, we estimated the true proportion of vegetarian pizza-eaters was between 0.2 and 0.46, with 95% confidence.

Confidence Level

- Confidence intervals consists of both an interval estimate and a confidence level.
  - Based on our pizza party sample, we estimated the true proportion of vegetarian pizza-eaters was between 0.2 and 0.46, with 95% confidence.
- What does confidence mean?

Confidence Level

- Confidence intervals consists of both an interval estimate and a confidence level.
  - Based on our pizza party sample, we estimated the true proportion of vegetarian pizza-eaters was between 0.2 and 0.46, with 95% confidence.
- What does confidence mean?
  - It gives the success rate for our method. For 95% of all possible samples, the interval we construct will actually contain the population parameter.

Confidence Level

- Confidence intervals consists of both an interval estimate and a confidence level.
  - Based on our pizza party sample, we estimated the true proportion of vegetarian pizza-eaters was between 0.2 and 0.46, with 95% confidence.
- What does confidence mean?
  - It gives the success rate for our method. For 95% of all possible samples, the interval we construct will actually contain the population parameter.
- The problem?
  - We only have 1 sample, and we don't know if it belongs to the 95% of "good" samples, or the 5% of "bad" ones

Confidence Level

- Confidence intervals consists of both an interval estimate and a confidence level.
  - Based on our pizza party sample, we estimated the true proportion of vegetarian pizza-eaters was between 0.2 and 0.46, with 95% confidence.
- What does confidence mean?
  - It gives the success rate for our method. For 95% of all possible samples, the interval we construct will actually contain the population parameter.
- The problem?
  - We only have 1 sample, and we don't know if it belongs to the 95% of "good" samples, or the 5% of "bad" ones
- The consolation?

Confidence Level

- Confidence intervals consists of both an interval estimate and a confidence level.
  - Based on our pizza party sample, we estimated the true proportion of vegetarian pizza-eaters was between 0.2 and 0.46, with 95% confidence.
- What does confidence mean?
  - It gives the success rate for our method. For 95% of all possible samples, the interval we construct will actually contain the population parameter.
- The problem?
  - We only have 1 sample, and we don't know if it belongs to the 95% of "good" samples, or the 5% of "bad" ones
- The consolation?
  - If I go through my life constructing 95% confidence intervals, I will be telling the truth about 95% of the time (I'll take that rate!)

Section 2

Confidence Interval Misunderstandings

Common Confidence Interval Misunderstandings

Suppose we wish to estimate the average point value in a custom deck of cards. We obtain a sample of size 20, with mean $\bar{x} = 8.4$, producing the 95% confidence interval of $(7.9, 8.9)$

Common Confidence Interval Misunderstandings

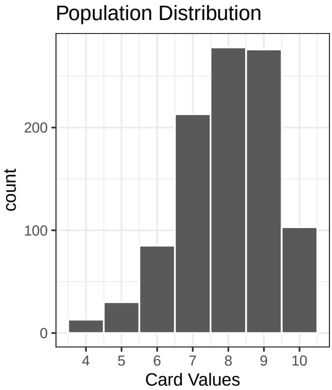
Suppose we wish to estimate the average point value in a custom deck of cards. We obtain a sample of size 20, with mean $\bar{x} = 8.4$, producing the 95% confidence interval of $(7.9, 8.9)$

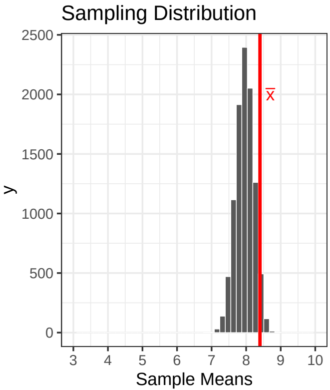
1. A 95% confidence interval **does not** contain 95% of observations in the population.

Common Confidence Interval Misunderstandings

Suppose we wish to estimate the average point value in a custom deck of cards. We obtain a sample of size 20, with mean $\bar{x} = 8.4$, producing the 95% confidence interval of $(7.9, 8.9)$

1. A 95% confidence interval **does not** contain 95% of observations in the population.

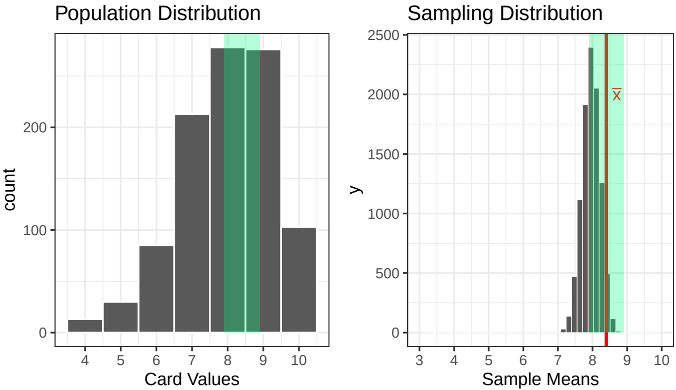




Common Confidence Interval Misunderstandings

Suppose we wish to estimate the average point value in a custom deck of cards. We obtain a sample of size 20, with mean $\bar{x} = 8.4$, producing the 95% confidence interval of $(7.9, 8.9)$

1. A 95% confidence interval **does not** contain 95% of observations in the population.

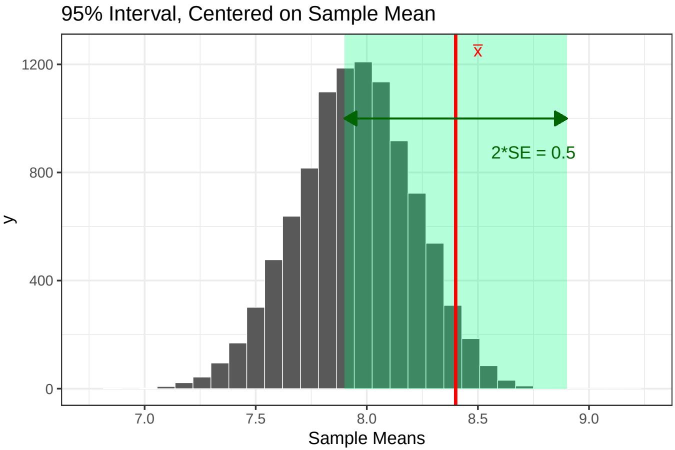


Common Confidence Interval Misunderstandings

2. A 95% confidence interval **does not** imply that 95% of all sample means fall within the given range.

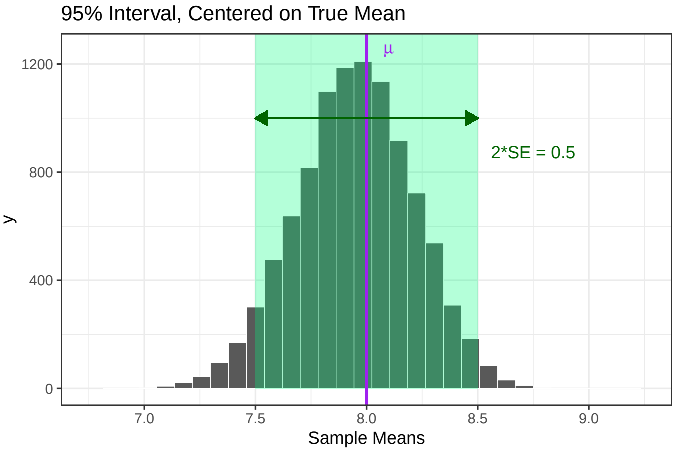
Common Confidence Interval Misunderstandings

2. A 95% confidence interval **does not** imply that 95% of all sample means fall within the given range.



Common Confidence Interval Misunderstandings

2. A 95% confidence interval **does not** imply that 95% of all sample means fall within the given range.



Common Confidence Interval Misunderstandings

3. A 95% confidence interval **does not** mean that there is a 95% probability that the true parameter falls in particular confidence interval

Common Confidence Interval Misunderstandings

3. A 95% confidence interval **does not** mean that there is a 95% probability that the true parameter falls in particular confidence interval

- Suppose we collect the random sample of cards, finding that $\bar{x} = 8.4$, which we use to construct the confidence interval $(7.9, 8.9)$.

Common Confidence Interval Misunderstandings

3. A 95% confidence interval **does not** mean that there is a 95% probability that the true parameter falls in particular confidence interval

- Suppose we collect the random sample of cards, finding that $\bar{x} = 8.4$, which we use to construct the confidence interval $(7.9, 8.9)$.
- At this point, once the sample has been observed and the interval calculated, there is no more randomness in the process. We cannot make probabilistic statements about the outcome.

Common Confidence Interval Misunderstandings

3. A 95% confidence interval **does not** mean that there is a 95% probability that the true parameter falls in particular confidence interval

- Suppose we collect the random sample of cards, finding that $\bar{x} = 8.4$, which we use to construct the confidence interval $(7.9, 8.9)$.
- At this point, once the sample has been observed and the interval calculated, there is no more randomness in the process. We cannot make probabilistic statements about the outcome.
  - The interval $(7.9, 8.9)$ either does or does not contain the fixed parameter.

Common Confidence Interval Misunderstandings

3. A 95% confidence interval **does not** mean that there is a 95% probability that the true parameter falls in particular confidence interval

- Suppose we collect the random sample of cards, finding that $\bar{x} = 8.4$, which we use to construct the confidence interval $(7.9, 8.9)$.
- At this point, once the sample has been observed and the interval calculated, there is no more randomness in the process. We cannot make probabilistic statements about the outcome.
    - The interval $(7.9, 8.9)$ either does or does not contain the fixed parameter.
- This is why we use the language "The confidence level for the interval is 95%" or "the success rate for the procedure is 95%"

Confidence Interval Conundrum

Confidence Interval Conundrum

- We do have one more problem:

Confidence Interval Conundrum

- We do have one more problem:
  - To make a confidence interval, we need the sampling distribution in order to compute the standard error.

Confidence Interval Conundrum

- We do have one more problem:
  - To make a confidence interval, we need the sampling distribution in order to compute the standard error.
  - But in order to visualize the sampling distribution, and compute it's standard deviation, we would need to obtain thousands of samples.

Confidence Interval Conundrum

- We do have one more problem:
  - To make a confidence interval, we need the sampling distribution in order to compute the standard error.
  - But in order to visualize the sampling distribution, and compute it's standard deviation, we would need to obtain thousands of samples.
- In practice, we just have 1 sample! And if we had the time / funding to obtain thousands of samples, we could probably just conduct a census of the population.

Confidence Interval Conundrum

- We do have one more problem:
  - To make a confidence interval, we need the sampling distribution in order to compute the standard error.
  - But in order to visualize the sampling distribution, and compute it's standard deviation, we would need to obtain thousands of samples.
- In practice, we just have 1 sample! And if we had the time / funding to obtain thousands of samples, we could probably just conduct a census of the population.
- Miraculously, it turns out we can assess the variability and shape of the sampling distribution using just a single sample!

Confidence Interval Conundrum

- We do have one more problem:
  - To make a confidence interval, we need the sampling distribution in order to compute the standard error.
  - But in order to visualize the sampling distribution, and compute it's standard deviation, we would need to obtain thousands of samples.
- In practice, we just have 1 sample! And if we had the time / funding to obtain thousands of samples, we could probably just conduct a census of the population.
- Miraculously, it turns out we can assess the variability and shape of the sampling distribution using just a single sample!
  - This process is called **Bootstrapping**, which we'll investigate next week!