Launching into Inference

Common Core State Standards in Mathematics

Making Inferences and Justifying Conclusions S-IC

Understand and evaluate random processes underlying statistical experiments
1. Understand statistics as a process for making inferences about population parameters based on a random sample from that population.
2. Decide if a specified model is consistent with results from a given data-generating process, e.g., using simulation. *For example, a model says a spinning coin falls heads up with probability 0.5. Would a result of 5 tails in a row cause you to question the model?*

Make inferences and justify conclusions from sample surveys, experiments, and observational studies
3. Recognize the purposes of and differences among sample surveys, experiments, and observational studies; explain how randomization relates to each.
4. Use data from a sample survey to estimate a population mean or proportion; develop a margin of error through the use of simulation models for random sampling.
5. Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant.
6. Evaluate reports based on data.

HOW TO ORGANIZE A STATISTICAL INFERENCE PROBLEM: A 4-STEP PROCESS

From *The Practice of Statistics,* 5th edition, by Starnes, Tabor, Yates, and Moore, W. H. Freeman and Co., in production

	Confidence intervals (CIs)	Significance tests
STATE:	What *parameter* do you want to estimate, and at what *confidence level*?	What *hypotheses* do you want to test, and at what significance level? Define any *parameters* you use.
PLAN:	Choose the appropriate inference *method*. Check *conditions*.	Choose the appropriate inference *method*. Check *conditions*.
DO:	If the conditions are met, perform *calculations*.	If the conditions are met, perform *calculations*. • Compute the **test statistic.** • Find the ***P*-value.**
CONCLUDE:	*Interpret* your interval in the context of the problem.	*Interpret* your test result in the context of the problem.

Inference in a Nutshell

Inference: Using sample data to draw conclusions about populations or treatments

Numerical summary	Parameter	Statistic	Type of Data
Proportion	p	$\hat{p}$	Categorical
Mean	μ	$\bar{x}$	Quantitative
Standard deviation	σ	s_x	Quantitative

Two main types of inference:

- Estimating — Confidence intervals
- Testing claims — Hypothesis tests/significance tests

Lesson – Idea of a Confidence Interval

Confidence Interval

A **confidence interval** gives an interval of plausible values for a parameter.

Confidence Level

The **confidence level *C*** gives the long-run success rate of confidence intervals calculated with *C*% confidence. That is, in *C*% c
all possible samples, the interval computed from the sample data wi
capture the true parameter value.

How to Interpret a Confidence Interval

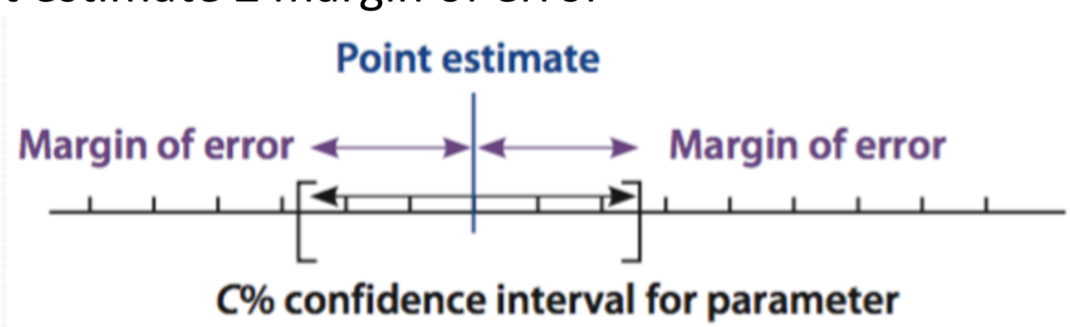
To interpret a *C*% confidence interval for an unknown parameter, say
"We are *C*% confident that the interval from _____ to _____ captures the
[parameter in context]."

To create an interval of plausible values for a parameter, we need two components:

- a **point estimate** to use as the midpoint of the interval and
- a **margin of error** to account for sampling variability.

The structure of a confidence interval is:

point estimate ± margin of error



Point Estimate, Margin of Error

A **point estimate** is a single-value estimate of a population parameter.

The **margin of error** of an estimate describes how far, at most, we expect the estimate to vary from the true population value. That is, i[n] a *C*% confidence interval, the distance between the point estimate and the true parameter value will be less than the margin of error in *C*% of all samples.

Example

Tim purchased a random sample of clementines at a local grocery store. The 95% confidence interval for the mean weight of all clementines at this store is 76.6 grams to 90.1 grams.

1. Interpret this confidence interval.
2. Calculate the point estimate and margin of error used to create this confidence interval.
3. The nutritional label on a bag of clementines says a typical clementine weighs 74 grams. Does this interval provide convincing evidence that the mean weight of clementines from this store is larger than the nutritional label suggests? Explain.

Solution

1.

We are 95% confident that the interval from 76.6 to 90.1 captures the true weight of all clementines at this store.

2.

point estimate = (76.6+90.1)/2=83.35 grams

margin of error = 90.1-83.35=6.75 or 83.35-76.6=6.75

3.

Yes. Because all the values in the confidence interval are greater than 74 grams, there is convincing evidence that mean weight of clementines from this store is larger than the nutritional label suggests.

Lesson – What Affects the Margin of Error?

The confidence level reveals how likely it is that the method we are using will produce an interval that captures the population parameter *if we use it many times.*

How to Interpret a Confidence Level

To interpret the confidence level *C*, say, "If we were to select many random samples from a population and construct a *C*% confidence interval using each sample, about *C*% of the intervals would capture the [parameter in context]."

- **The confidence level reveals how likely it is that the method we are using will produce an interval that captures the population parameter *if we use it many times.***
- **Caution! The confidence level does not tell us the probability that a particular confidence interval captures the population parameter**.

In general, we prefer short confidence intervals—that is, confidence intervals with a small margin of error.

To reduce the margin of error, we can change two factors:

- sample size and
- confidence level.

We can get a more precise estimate of a parameter by increasing the sample size. Larger samples yield narrower confidence intervals at any confidence level.

- When we calculate a confidence interval, we include the margin of error because we expect the value of the point estimate to vary somewhat from the parameter.
- However, the margin of error accounts for *only* the variability we expect from random sampling.
- **Caution: The margin of error does *not* account for any sources of bias in the data collection process.**

Example

The puzzle editor of a game magazine asked 43 randomly selected subscribers how long it took them to complete a certain crossword puzzle. Based on this sample, a 99% confidence interval for the mean completion time for all subscribers is 15.2 to 18.6 minutes.

1. Interpret the confidence level.

2. Name two things the puzzle editor could do to reduce the margin of error. What drawbacks do these actions have?

3. The puzzle editor originally sent requests for completion times to 150 subscribers, only 43 of whom reported their times. Describe one potential source of bias in the study that is not accounted for by the margin of error?

Solution

1. The puzzle editor took many random samples of subscribers and constructed a 99% confidence interval for each sample, about 99% of these intervals would capture the true mean completion time for all subscribers.

2. The puzzle editor could reduce the margin of error by decreasing the confidence level. The drawback is that we can't be as confident that our interval will capture the true mean. The puzzle editor could also reduce the margin of error by increasing the sample size. The drawback is that larger samples cost more time and money to obtain.

3. People embarrassed by their puzzle completion time might be less likely to respond (nonresponse). The mean puzzle completion time for all subscribers may be even more than 18.6 minutes. The bias due to nonresponse is not accounted for by the margin of error, because the margin of error accounts only for variability we expect from random sampling.

Lesson – Estimating a Proportion

How to Check the Conditions for Constructing a Confidence Interval for *p*

To construct a confidence interval for *p*, check the following condition

- **Random**: The data come from a random sample from the population of interest.
- **Large Counts**: Both $n\hat{p}$ and $n(1 - \hat{p})$ are at least 10.

Critical Value

The **critical value** is a multiplier that makes the interval wide enough have the stated capture rate.

Standard Error of $\hat{p}$

The **standard error of $\hat{p}$** is an estimate of the standard deviation of th sampling distribution of $\hat{p}$.

$$SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

The standard error of $\hat{p}$ estimates how much $\hat{p}$ typically varies from *p*.

How to Calculate a Confidence Interval for *p*

When the Random and Large Counts conditions are met, a *C*% confidence interval for the population proportion *p* is

point estimate ± margin of error

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where z* is the critical value for the standard normal curve with C% c its area between −z* and z*.

Example

An online retailer wants to estimate p = the proportion of all orders in a given week that are from new customers. He takes an SRS of 40 orders and finds that 12 of them are from new customers.

1. Show that the conditions for calculating a confidence interval for a proportion are satisfied.

2. Calculate a 99% confidence interval for the proportion of the week's order s that are from new customers.

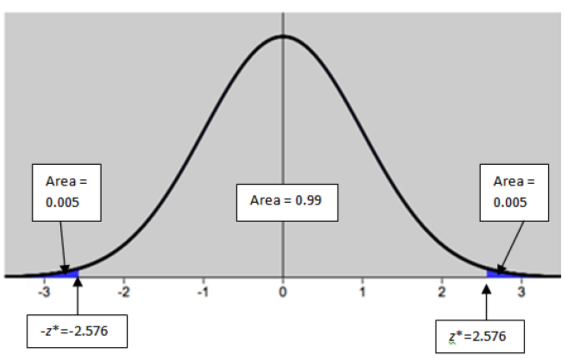
3. Interpret the interval from Question 2.

Solution

1.

- Random? The online retailer take an SRS of orders.
- Large counts?
  $n\hat{p} = 40\frac{12}{40} = 12 \geq 10$ and $n(1-\hat{p}) = 40\left(1-\frac{12}{40}\right) = 28 \geq 10$

2. Calculate the critical value *z** for 99% confidence using the standard normal curve.



Because $\frac{1-0.99}{2} = 0.005$, *z** for a 99% confidence interval can be found by looking for a left-tail area of 0.005. The closest area is 0.0049, corresponding to a critical value of *z** = 2.58.
Using technology: invNorm(left-tail area:0.005, mean:0, SD:1) gives z*= 2.576.

$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \rightarrow$ $0.30 \pm 2.576\sqrt{\frac{0.30(1-0.30)}{40}} \rightarrow 0.30 \pm 0.187 \rightarrow 0.113$ to 0.487

3.
We are 99% confident that the interval from 0.113 to 0.487 captures the true proportion of all orders in a given week that are from new customers.

Lesson 7.4 – Confidence Interval for a Proportion

How to Use the Four-Step Process: Confidence Intervals

- **STATE**: State the parameter you want to estimate and the confidence level.
- **PLAN**: Identify the appropriate inference method and check conditions.
- **DO**: If the conditions are met, perform calculations.
- **CONCLUDE**: Interpret your interval in the context of the problem.

How to calculate the confidence interval for a proportion:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

How to Calculate a Sample Size for a Desired Margin of Error

To determine the sample size n that will yield a *C*% confidence interv for a population proportion *p* with a maximum margin of error *ME*, solve the following inequality for *n*:

$$z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq ME$$

where $\hat{p}$ is a guessed value for the sample proportion. The margin of error will always be less than or equal to *ME* if you use $\hat{p}$= 0.5.

Two possible values of the sample proportion to use are:

1. Use a guess based on a preliminary study or past experience.
2. Use p = 0.5. The margin of error is largest when p = 0.5, so this guess is conservative.

Example

A recent CNN poll found that in a random sample of 508 U.S. residents, 31% answered "Yes" to the question "Do you think life has ever existed on Mars?"

1. Construct and interpret a 99% confidence interval for the proportion of all United States residents who believe life ever existed on Mars.

2. Based on the interval, is there convincing evidence that more than 35% of U.S. residents think life ever existed on Mars? Explain.

3. If the polltakers wanted to reduce the margin of error to at most 4%, about how many *additional* U.S. ressidents do they need to randomly select? Use $\hat{p}=0.31$.

Solution

1.

STATE: We want to estimate *p* = the true proportion of all U.S. residents who think life ever existed on Mars.

PLAN: One-sample *z* interval for *p*.

Random? CNN took a random sample of508 U.S. residents.

Large counts? $n\hat{p}=508(0.31)=157\geq 10$ and $n(1-\hat{p})=508(1-0.31)=351\geq 10$

$\hat{p}\pm z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\rightarrow$ $0.31\pm 2.576\sqrt{\frac{0.31(1-0.31)}{508}}\rightarrow 0.31\pm 0.053\rightarrow 0.257$ to 0.363

CONCLUDE: We are 99% confident that the interval from 0.257 to 0.363 captures the true proportion of all U.S. residents who think life ever existed on Mars.

2.

Yes. All the values of the confidence interval are greater than 0.35, so there is convincing evidence that more than 35% of all U.S. residents think life ever existed on Mars.

3.

$$z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\leq ME$$

$2.576\sqrt{\frac{0.31(1-0.31)}{n}}\leq 0.04\rightarrow n\geq 887.12$

We need to survey at least 888 U.S. residents. This would require an *additional* 888-508=380 U.S. residents be randomly selected.

Lesson – Estimating a Mean

How to Check the Conditions for Constructing a Confidence Interval for μ

- **Random**: The data come from a random sample from the population of interest.
- **Normal/Large Sample**: The data come from a normally distribute population or the sample size is large ($n \geq 30$).

- Confidence intervals are made up of two parts, the point estimate and the margin of error:

  point estimate ± margin of error

- The confidence interval for a population mean is:

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

- Unfortunately, if we don't know the true value of μ, we rarely know the true value of ⍰ either. We can use s_x as an estimate for ⍰, but things don't work out as nicely as we might like.
- When calculating a confidence interval for a population mean, we use a t^* critical value rather than a z^* critical value whenever we use s_x to estimate ⍰.

How to Find t^*

1. Using Table B, find the correct confidence level at the bottom c the table.
2. On the left side of the table, find the correct number of *degrees freedom* (df). For this type of confidence interval, df = n - 1.
3. In the body of the table, find the value of t^* that corresponds to confidence level and df.
4. If the correct df isn't listed, use the greatest df available that is than the correct df.

To calculate 99% confidence intervals with *n* =4, use df = 4-1 =3.

	Tail probability *p*			
df	**0.02**	**0.01**	**0.005**	**0.0025**
1	15.89	31.82	63.66	127.3
2	4.849	6.965	9.925	14.09
3	3.482	4.541	5.841	7.453
⋮	⋮	⋮	⋮	⋮
z^*	2.054	2.326	2.576	2.807
	96%	**98%**	**99%**	**99.5%**
	Confidence level *C*			

For 99% confidence and 3 degrees of freedom, $t^* = 5.841$. That is, the interval should extend 5.841 standard deviations on both sides of the point estimate to have a capture rate of 99%.

Standard error of $\bar{x}$

The **standard error of $\bar{x}$** is an estimate of the standard deviation of th
sampling distribution of $\bar{x}$.

$$SE_{\bar{x}} = \frac{s_x}{\sqrt{n}}$$

The standard error of $\bar{x}$ estimates how much $\bar{x}$ typically varies from μ.

How to Calculate a Confidence Interval for μ

When the Random and Normal/Large Sample conditions are met,
C% confidence interval for the unknown mean μ is

$$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$$

where t^* is the critical value for a t distribution with df = n - 1 and
C% of its area between $-t^*$ and t^*.

Example

A study of commuting times reports the travel times to work of a random sample of 36 employed adults in New York State. The mean is ▯*=31.25* ▯▯▯▯▯▯▯*,* and the standard deviation is ▯▯*=21.88* ▯▯▯▯▯▯▯*.*

1. Verify that the conditions are met for constructing a confidence interval for ▯ .

2. Construct a 95% confidence interval for ▯ .

3. Last year's study found that the mean commuting time for employed adults was 27 minutes. Does the interval in #2 provide convincing evidence that the commuting time for this year is more than last year? Explain.

Solution

1.

Random? Random sample of 36 employed adults.

Normal/Large Sample? We don't know the shape of the population distribution, but the sample size is large: n=36>30 .

2.

95% confidence and df = 36 – 1 = 35 (use *df* = 30) gives *t** = 2.042.

$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}} \rightarrow$ *31.25±2.04221.8836→31.25±7.45 →23.8* ▯▯ *38.7*

We are 95% confident that the interval from 23.8 to 38.7 minutes captures the true mean travel time to work by employed adults in the state of New York.

3.

No. 27 minutes is included in the interval, so there is not convincing evidence that the mean travel time to work is more than last year.

Lesson – Confidence Interval for a Mean

What do we do if the sample size is small ($n < 30$) and the population distribution is unknown?

How to Check the Normal/Large Sample Condition

There are three ways that the Normal/Large Sample condition can be met:

1. The data come from a normally distributed population.
2. The sample size is large ($n \geq 30$).
3. When the sample size is small and the shape of the population distribution is unknown, a graph of the sample data shows no strong skewness or outliers.

Caution: Make sure to include the graph of sample data when the sample size is small and you are checking the Normal/Large Sample condition.

It doesn't matter whether you use a dotplot, stemplot, histogram, or boxplot to check for strong skewness or outliers.

How to Use the Four-Step Process: Confidence Intervals

- **STATE**: State the parameter you want to estimate and the confidence level.
- **PLAN**: Identify the appropriate inference method and check conditions.
- **DO**: If the conditions are met, perform calculations.
- **CONCLUDE**: Interpret your interval in the context of the problem.

A confidence interval for a population mean is often referred to as a *one-sample t interval for* μ.

How to Calculate a Confidence Interval for μ

When the Random and Normal/Large Sample conditions are met, a
C% confidence interval for the unknown mean μ is

$$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$$

where t^* is the critical value for a t distribution with df = n - 1 and
C% of its area between $-t^*$ and t^*.

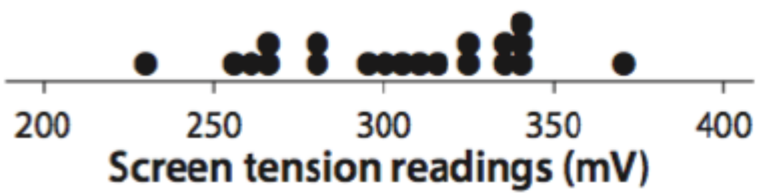
Example

A manufacturer of high-resolution video terminals must control the tension on the mesh of fine wires that lies behind the surface of the viewing screen. Too much tension will tear the mesh, and too little will allow wrinkles. The tension is measured by an electrical device with output readings in millivolts (mV). Some variation is inherent in the production process. Here are the tension readings from a random sample of 20 screens from a single day's production:

269.5	297.0	269.6	283.3	304.8	280.4	233.5	257.4	317.5	327.4
264.7	307.7	310.0	343.3	328.1	342.6	338.8	340.1	374.6	336.1

1. Construct and interpret a 90% confidence interval for the mean tension m of all the screens produced on this day.
2. The manufacturer's goal is to produce screens with an average tension of 300 mV. Based on the interval, is there convincing evidence that the screens produced this day don't meet the manufacturer's goal? Explain.

Solution
1.
STATE: We want to estimate ⍰= the true mean screen tension of all the screens produced on this day at a 90% confidence level.
PLAN: One-sample *t* interval for ⍰.
Random? Random sample of 20 screens.
Normal/Large Sample? The sample size is small, but the dotplot doesn't show any outliers



or strong skewness.
DO: For these data, $\bar{x} = 306.32$, $s_x = 36.21$, and $n = 20$. With 90% confidence and $df = 20 - 1 = 19$, $t^* = 1.729$.

$$306.32 \pm 1.729 \frac{36.21}{\sqrt{20}} \rightarrow 306.32 \pm 14.00 \rightarrow 292.32 \text{ to } 320.32$$

CONCLUDE: We are 90% confident that the interval from 292.32 mV to 320.32 mV captures the true mean screen tension of all the screens produced on this day.

2.
No. 300 mV is captured within the confidence interval and is therefore a plausible value for the mean screen tension of all the screens produced on this day. We don't have convincing evidence that the screens produced this day don't meet the manufacturer's goals.