

# Confidence Intervals

Grinnell College

November 10<sup>th</sup>, 2025

# Where We are at....

We are using normal distributions to help approximate the sampling distributions of common statistics.

- ▶ 1 mean (w/ known variance)
  - ▶ Hypothesis Test (Z-test)
  - ▶ Confidence Interval
- ▶ 1 mean (w/ unknown variance)
  - ▶ HT (t-test) and CI
- ▶ Difference for 2 means
  - ▶ HT and CI
- ▶ Regression Coefficients
  - ▶ Multiple Regression
  - ▶ HT and CI
- ▶ 1 proportion (HT and CI)
- ▶ 2 proportions (HT and CI)

We then have other tests to discuss

- ▶ ANOVA (testing difference of more than 2 means)
- ▶  $\chi^2$ -test (testing for association between two categorical variables)

# Motivation

On the dice lab, I talked about rolling the dice twice being “safer” than doubling the value of the dice.

The idea is that I’m going to have a lower probability of rolling low (and high) values and instead get more rolls in the middle (over the long run)

# Goals

We can talk about the same thing using our sampling distribution for our sample mean, given our assumptions are met (which are...?)

- ▶  $N(\mu, \sigma^2/n)$
- ▶ Still assuming we know  $\sigma^2$
- ▶ Instead of final dice roll value we have the sample mean
- ▶ Can talk about where likely sample means are going to be
- ▶ Our population is all possible sample means
  - ▶ Similar to permutation test's all possible juries
  - ▶ Here we assume the number of possible combinations is infinite

# Recall the Sampling Distribution

If our assumptions are met (which are....?) then we can talk about the long term behavior of the means from random samples.

- ▶ We have already seen calling BS via hypothesis tests

We will now talk about where most of the sample means will be

- ▶ Report things like where middle 95% of sample means will be
  - ▶ 68-95-99.7% rule
- ▶ Uses the normal probability distribution
- ▶ And we will formalize it with math and interpretations

# Motivating Example

My niece's first birthday is next month (12/3) and I \*might\* get her some Duplo legos.

How much is that going to cost me?

# Motivating Example Solution 1

My niece's first birthday is next month (12/3) and I \*might\* get her some Duplo legos.

How much is that going to cost me on average?

One strategy is to just report the mean for the cost of the Duplo legos which is \$34.26.....but I probably want pay that exactly

## Motivating Example Solution 2

My niece's first birthday is next month (12/3) and I \*might\* get her some Duplo legos.

How much is that going to cost me on average?

A better strategy is to report where we think any given sample's mean is most likely going to be (in the long run)

- ▶ Need an upper and lower value
- ▶ Calculated off of the sampling distribution (which is normal)



# Confidence Interval

This is called a confidence interval (often abbreviated CI). A **confidence interval** gives the location of the middle (BLANK) % of the sampling distribution.

- ▶ Suggests where we think most sample means will be
- ▶ Also suggests where we think the population parameter (mean in this slide deck) will be
- ▶ The sampling distribution requires the same assumptions as the hypothesis test

# Sampling Distribution's location

Wait, wasn't there a whole thing about not knowing where the sampling distribution was actually located?

# Sampling Distribution's location

Wait, wasn't there a whole thing about not knowing where the sampling distribution was *\*actually\** centered?

Correct! Under hypothesis testing we assumed we knew the population's mean's true value ( $H_0$ )

# Sampling Distribution's location

Wait, wasn't there a whole thing about not knowing where the sampling distribution was \*actually\* centered?

Correct! Under hypothesis testing we assumed we knew the population's mean's true value ( $H_0$ )

For confidence intervals we actually just plug in our sample mean

- ▶ Sample Dist  $\approx$  Pop Dist (we hope)
- ▶ Sample mean  $\approx$  Pop mean
  - ▶  $\bar{X} = \mu$

## Example 1: Penguins!!

Body measurements were taken on 340 of penguins, randomly selected, on some islands off of Antarctica with a mean flipper length of 200mm. Find a 95% confidence interval for the mean flipper length



## Example 1: Penguin Assumptions

Are the observations random?

- ▶ Yes, it says so

Are the observations independent?

- ▶ Maybe? There'll be genetic relatives in each colony
- ▶ And each colony could be different with different means

Is the population normally distributed or is the sample size large?

- ▶ The population is not normally distributed (because the sample isn't)
- ▶ The sample size is....large with over 300 observations

Despite possible concerns with independence let's move forward

## Ex: Penguin Sampling Distribution

The resulting sampling distribution then is  $N(\mu, \sigma^2/n)$  which is

$$\bar{X} \sim N(200, 180/340)$$

where

- ▶  $\mu$  is the mean of our population
  - ▶ Plug in our estimated sample mean (200)
  - ▶ We can hypothesize it's value (last slide deck)
- ▶ 180 is our (pop) variance
  - ▶ Assumed known before we ever started talking about penguins
  - ▶ Very unrealistic short of an ornithologist savant
- ▶ 340 is our sample size

Alright, let's see the confidence interval formula!

# CI Formula

A  $(1-\alpha)$  **confidence interval** is calculated as....

$$\text{estimate} \pm (\text{distributional value})(\text{standard error})$$

$$\text{sample mean} \pm z_{\alpha/2}(\sqrt{\sigma^2/n})$$

$$200 \pm z_{\alpha/2}(\sqrt{180/340})$$

$$200 \pm z_{\alpha/2}(0.7276)$$

Okay but what is  $z_{\alpha/2}$ ?



This value comes from the standard normal distribution and represents the  $\alpha/2$ -percentile.

- ▶  $\alpha/2$  because we want  $\alpha$  probability in both tails combined
- ▶ So we cut  $\alpha$  in half for symmetry
- ▶ Can look it up with `qnorm()` in R
- ▶ Common choices given below
- ▶ No choice is more common than 95%

Confidence Level (= $1 - \alpha$ )	80	90	95	99
$z_{\alpha/2}$	-1.28	-1.645	-1.96	-2.58

## Example 1: Confidence Interval

So, we wanted a 95% confidence interval so we need to plug in our values...

$$200 \pm z_{\alpha/2}(0.7276)$$

$$200 \pm 1.96^*(0.7276)$$

$$200 \pm 1.43$$

---

$$\text{Lower Bound} = 200 - 1.43 = 198.57$$

$$\text{Upper Bound} = 200 + 1.43 = 201.43$$

# CI Interpretation

“ We are (BLANK) % confident the true (STATISTIC) is between (LOWER BOUND) and (UPPER BOUND)”

- ▶ BLANK = confidence level
  - ▶ Here 95%
- ▶ STATISTIC = statistic we are interested in
  - ▶ Here it's a mean
- ▶ Lower and Upper Bounds that were calculated last slide
  - ▶ Lower gets reported first!!
  - ▶ Safety check: should both be same distance from the mean

## Example 1: CI Interpretation

“ We are 95% confident the true mean flipper length is between 198.57 and 201.43 mm”

- ▶ BLANK = 95% =  $(1 - \alpha)$
- ▶ STATISTIC = mean flipper length
- ▶ Lower and Upper Bounds: 198.57 and 201.43 from our calculations

# Okay but what is confidence? And what does it tell us?

**Confidence** does NOT mean probability. Instead, its the idea that if

- ▶ We take infinite number of samples
- ▶ And build CI at 95% confidence (say)
- ▶ About 95% of our intervals will include the “truth”

There is a famous picture for this.....

## CI Interpretation: What it's not

A super common mistake is to think that the next observation is most inside our interval. It's not.

- ▶ The CI are built using our sampling distribution for the sample means
- ▶ It doesn't talk about our sample directly per se
- ▶ The distribution was created with means, then we talk about means
- ▶ CI are most useful for explanation/description of distributions
- ▶ and almost useless for predictions

## Another Example

Lumber from cherry trees can be quite beautiful running from a pinkish brown to a reddish color. A common goal for woodworkers is to get long planks which is a function of the height of the tree. In a random sample of 31 ( $= n$ ) trees from around a forest, the average height was 73 ( $= \bar{x}$ ) feet with a known population variance of 45 ( $= \sigma^2$ )

Make a 90% confidence interval for the mean height of cherry trees

(HINT: You'll need to check assumptions to build a sampling distribution)

# Another Example: Assumptions

From z-tests we know we are good but to reiterate...

- ▶ Random?
  - ▶ Yes, it says so
- ▶ Independent and Identically Distributed?
  - ▶ Yes although trees could compete maybe? Relatives?
  - ▶ I'm not too concerned
- ▶ Population distribution is normal or large  $n$ ?
  - ▶ Sample distribution is normal(-ish) so we are good to go



## Another Example: Visualize



## Another Example: CI

$$\text{estimate} \pm (\text{distributional value})(\text{standard error})$$

which leads to...

$$73 \pm (1.645)(\sqrt{45/31})$$

$$73 \pm 1.98$$

$$(71.02, 74.98)$$

## Another Example: CI Interpretation

“ We are (BLANK) % confident the true (STATISTIC) is between (LOWER BOUND) and (UPPER BOUND)”

- ▶ BLANK = 90%
- ▶ STATISTIC = mean cherry tree height
- ▶ Lower = 71.02
- ▶ Upper = 74.98

“We are 90% confident the true mean cherry tree height is between 71.02 and 74.98 inches”

# Conclusion

Broadly, for any statistic there are three main types of inferences we can do

- ▶ Hypothesis test

- ▶ We can see if a current working hypothesis seems reasonable
- ▶ Check the probability we'd see a result as extreme as we saw IF THE NULL HYPOTHESIS IS TRUE
- ▶ We assume the population mean in  $H_0$

- ▶ Confidence Intervals

- ▶ Gives the range that we have confidence in that the true parameter is located there
- ▶ We assume the population mean \*is\* our sample mean for building the sampling distribution

- ▶ Prediction Intervals

- ▶ Introduced in your next lab
- ▶ Gives a reasonable range for the next observation