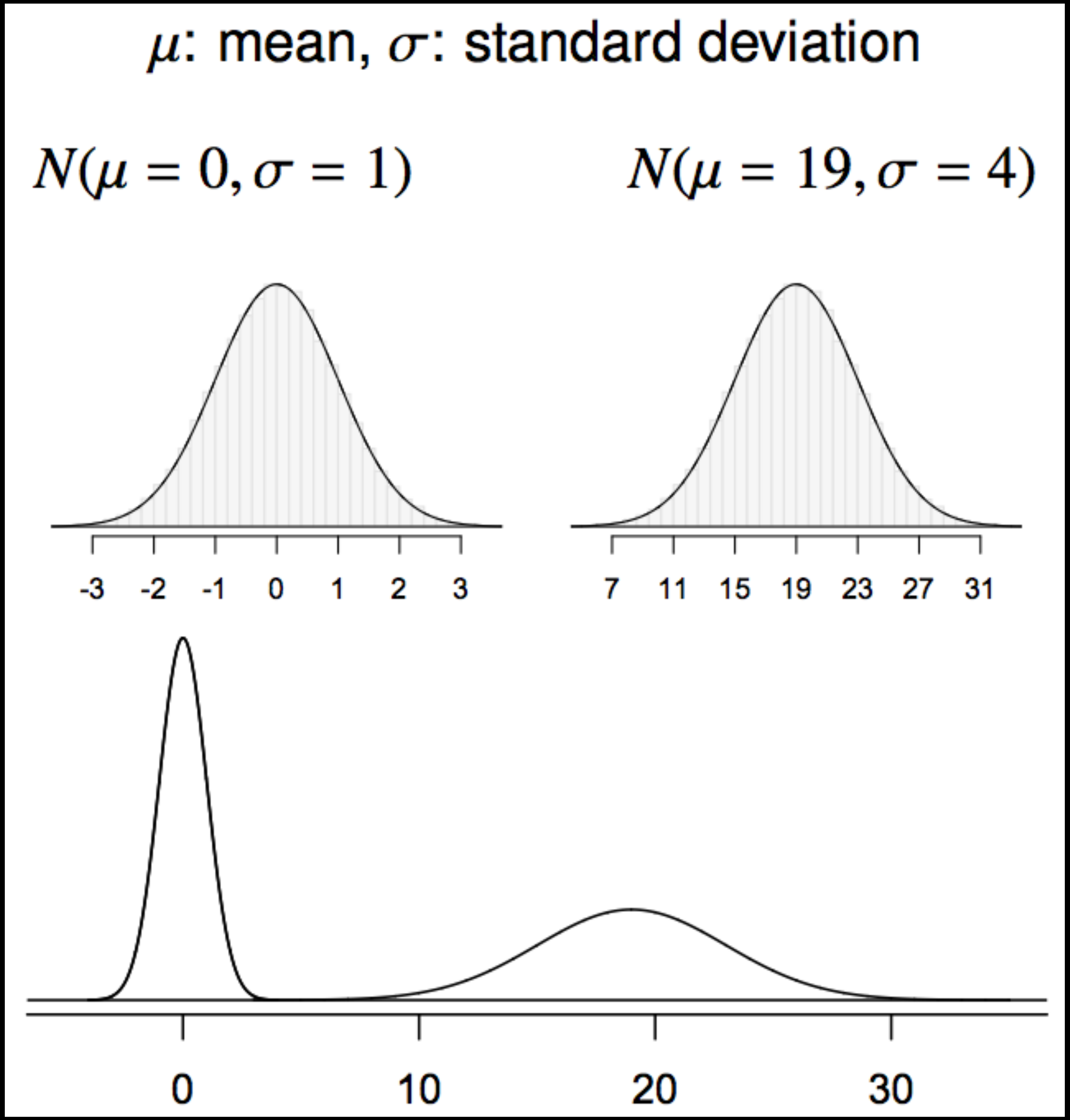
Welcome to Week #7!

Reminder: Midterm check-in due this Friday

Additionally: feedback survey

Also: Nov 3 office hours

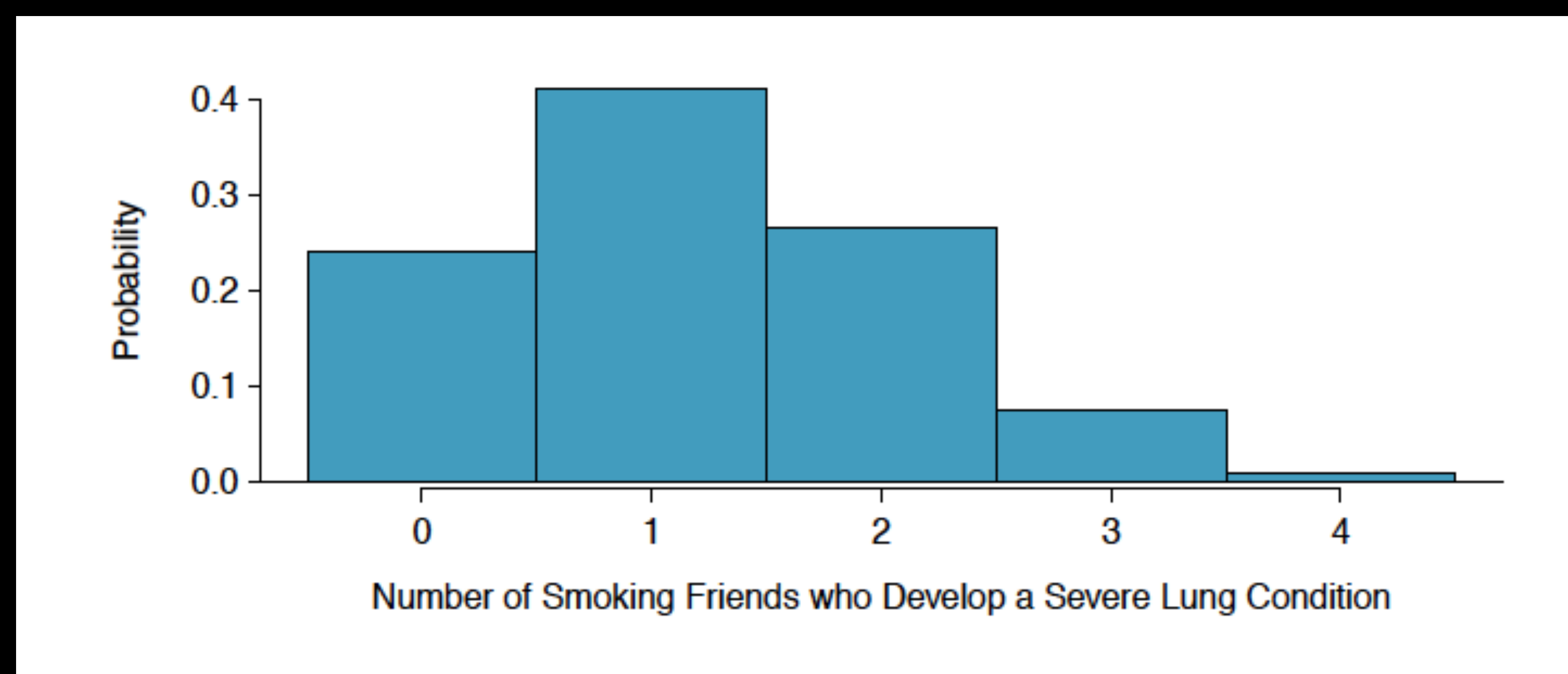
Normal distributions with different parameters



The Binomial distribution (cont.)

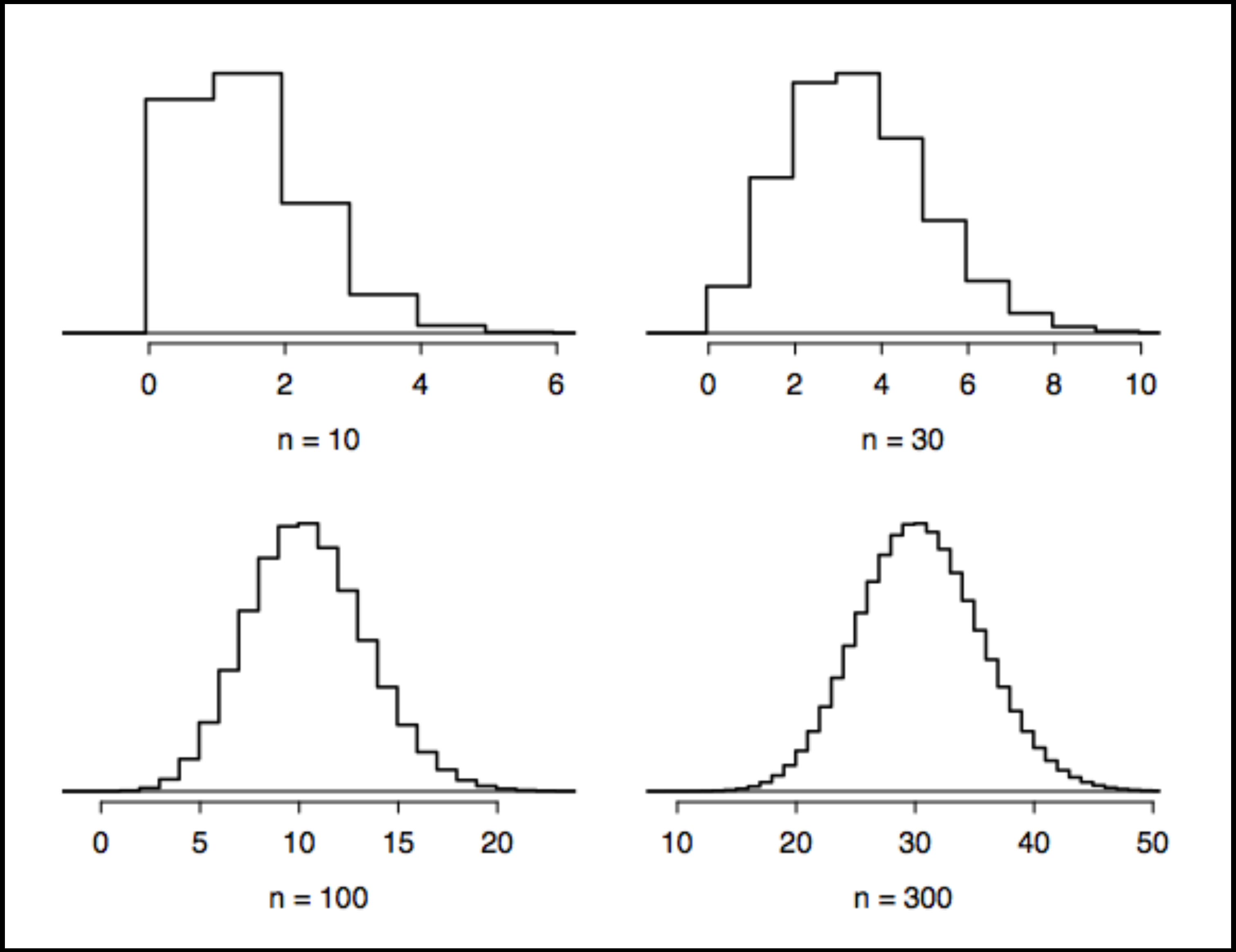
Once the probabilities of each value are calculated using the binomial formula, a probability histogram can be drawn in order to visualize the distribution. Like any distribution, the binomial distribution has a mean and a standard deviation.

x_i	p_i
0	$\binom{4}{0}(0.3)^0(0.7)^4 = 0.240$
1	$\binom{4}{1}(0.3)^1(0.7)^3 = 0.412$
2	$\binom{4}{2}(0.3)^2(0.7)^2 = 0.265$
3	$\binom{4}{3}(0.3)^3(0.7)^1 = 0.076$
4	$\binom{4}{4}(0.3)^4(0.7)^0 = 0.008$



Distributions of number of successes

Hollow histograms of samples from the binomial model where **p = 0.10** and n = 10, 30, 100, and 300.
What happens as n increases?



See this applet with sliders for n and p to see how shape binomial distribution changes as n and p change:

http://www.stat.berkeley.edu/~stark/Java/Html/BinHist.htm

Note: the scales on the histograms are different!

Foundations for Inference - How well can we really know anything?

Normal approximation of the binomial distribution

The binomial distribution with probability of success p is nearly normal when the sample size n is sufficiently large that np and $n(1-p)$ are both at least 10. The approximate normal distribution has parameters corresponding to the mean and standard deviation of the binomial distribution:

$$\mu = np \qquad \sigma = \sqrt{np(1-p)}$$

CENTRAL LIMIT THEOREM AND THE SUCCESS-FAILURE CONDITION

When observations are independent and the sample size is sufficiently large, the sample proportion $\hat{p}$ will tend to follow a normal distribution with the following mean and standard error:

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

In order for the Central Limit Theorem to hold, the sample size is typically considered sufficiently large when $np \geq 10$ and $n(1-p) \geq 10$, which is called the **success-failure condition**.

https://openstax.org/books/introductory-business-statistics/pages/7-3-the-central-limit-theorem-for-proportions

CENTRAL LIMIT THEOREM FOR THE SAMPLE MEAN

When we collect a sufficiently large sample of n independent observations from a population with mean μ and standard deviation σ, the sampling distribution of $\bar{x}$ will be nearly normal with

$$\text{Mean} = \mu \qquad\qquad \text{Standard Error } (SE) = \frac{\sigma}{\sqrt{n}}$$

RULES OF THUMB: HOW TO PERFORM THE NORMALITY CHECK

There is no perfect way to check the normality condition, so instead we use two rules of thumb:

n $<$ **30:** If the sample size n is less than 30 and there are no clear outliers in the data, then we typically assume the data come from a nearly normal distribution to satisfy the condition.

n $\geq$ **30:** If the sample size n is at least 30 and there are no *particularly extreme* outliers, then we typically assume the sampling distribution of $\bar{x}$ is nearly normal, even if the underlying distribution of individual observations is not.

TO R!

Key Insights about how well we know the "average" number representing a sample:

Let's say we want to know the average observation (sample mean) from a random sample taken from a population (usually the case, very rarely can we sample the entirety of the population):

IF the samples are independent (e.g. randomly sampled)
IF the sample size is "large enough" (typically > 30 observations)
IF the underlying population distribution is not strongly skewed

This is admittedly a bit "hand-wavy" and not rigorous (stay tuned for your future stats classes!)

THEN

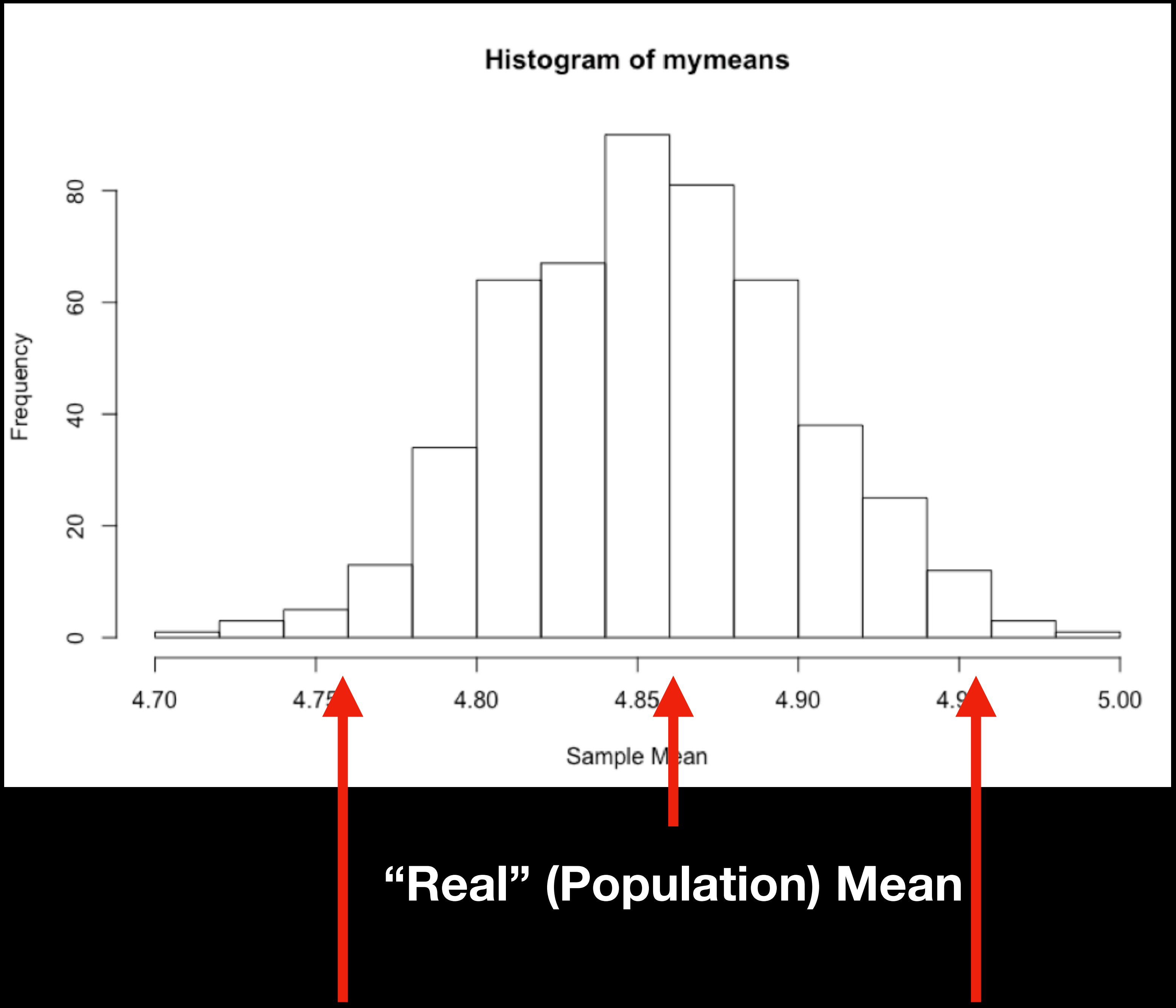
1. The "average" value of this population mean is the sample mean
2. The error on the measurement of the mean is given by the "standard error": $SE = s/n^{1/2}$

If you are curious, this comes from:

$$\mathrm{Var}\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n^2}\sum \mathrm{Var}(X_i) = \frac{1}{n^2} \times \sum \sigma^2 = \frac{n}{n^2}\sigma^2 = \frac{\sigma^2}{n}$$

Where "s" is the standard deviation of the sample & n is the number of samples

Confidence Intervals

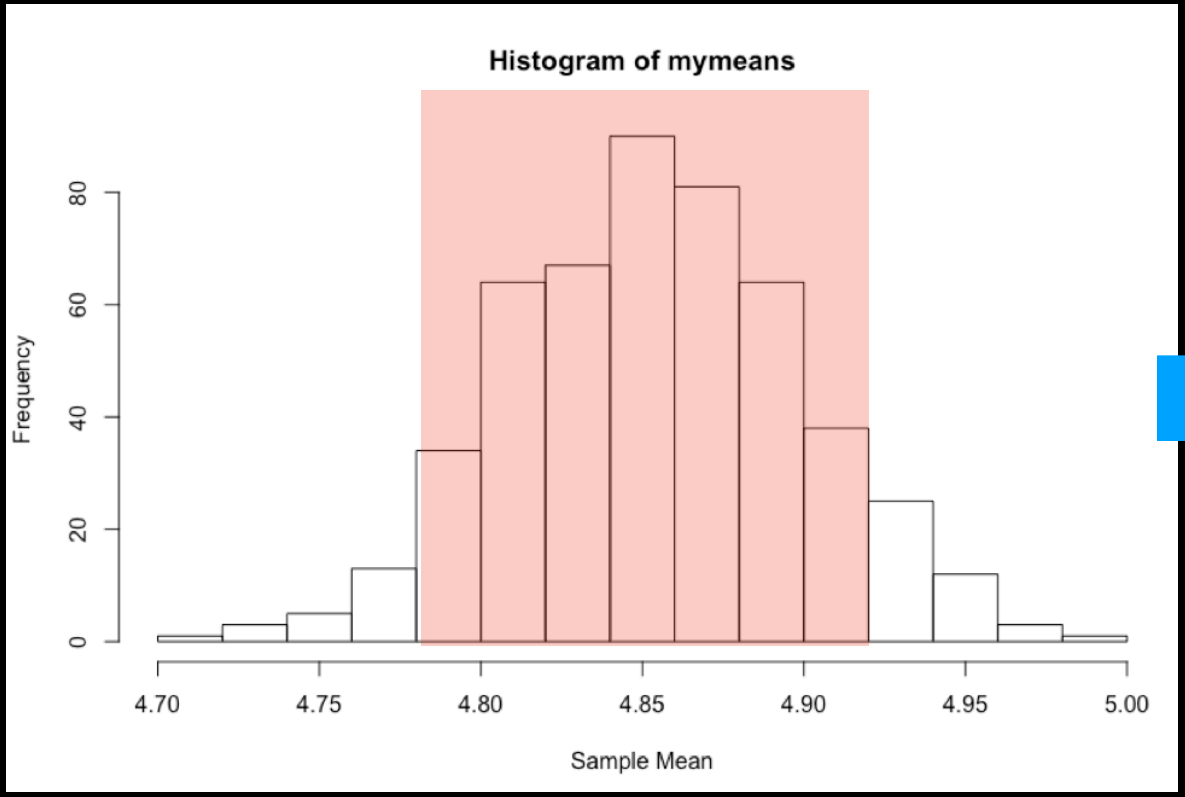


"Real" (Population) Mean

But I'm pretty sure the mean is somewhere in this interval

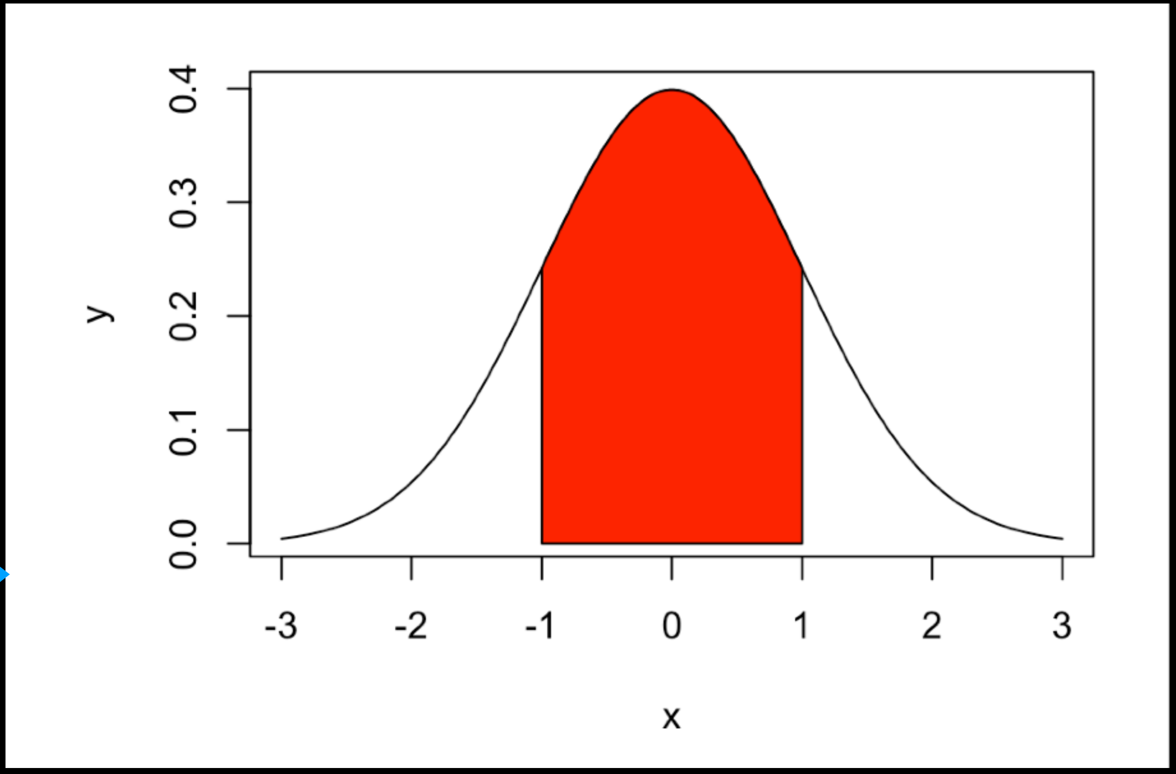
Confidence Intervals

~Normal

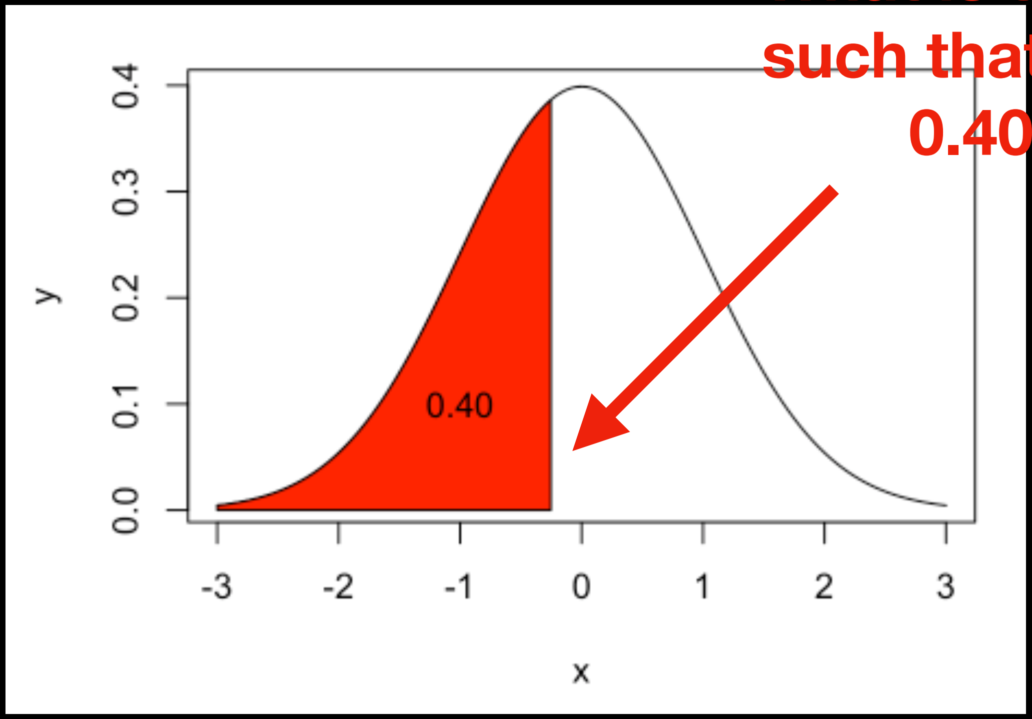


Where are we 95% confident the population mean is in between?

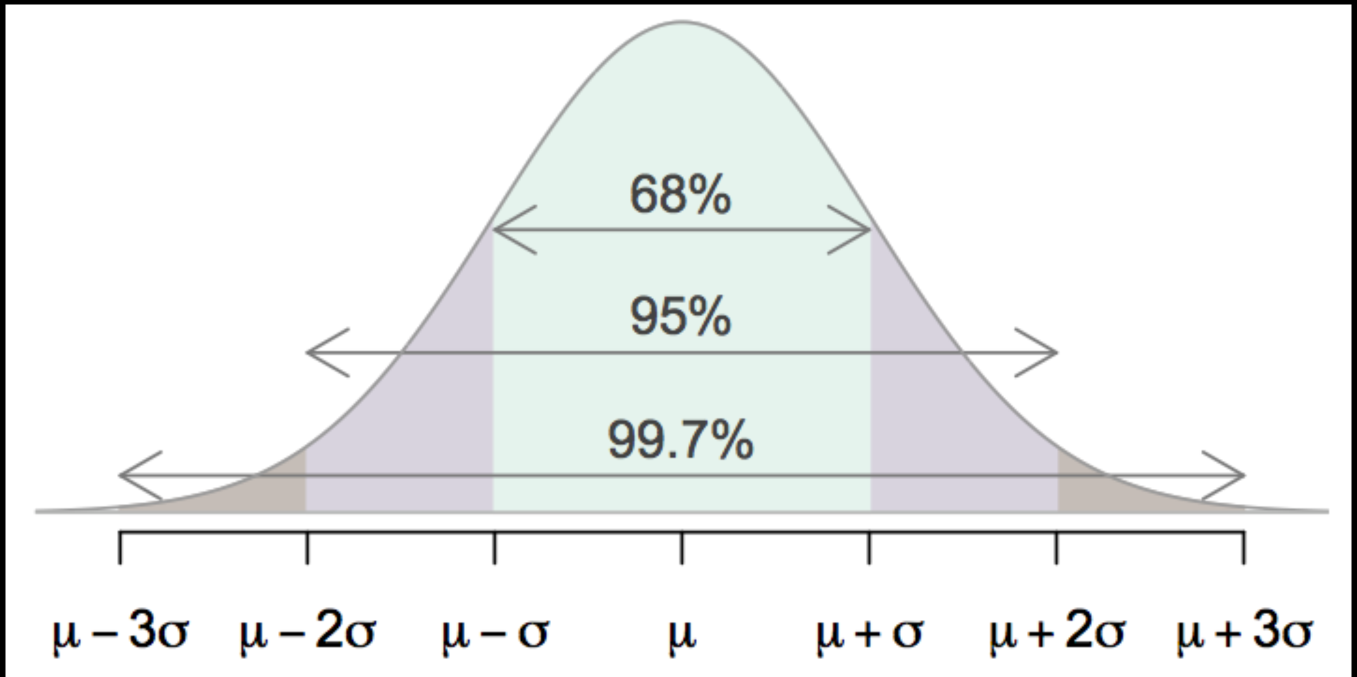
Recall:



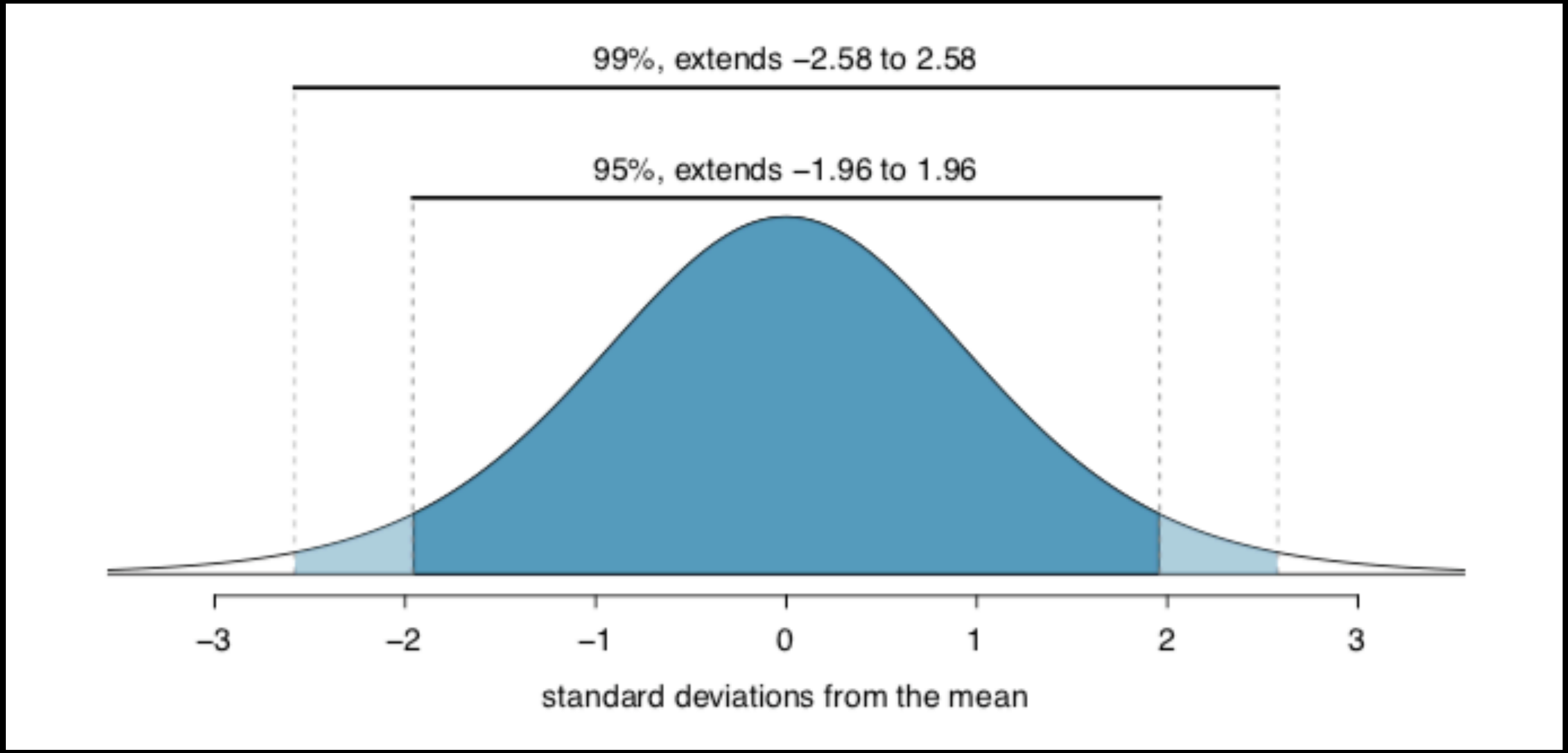
What is this number such that red area = 0.40 (40%)



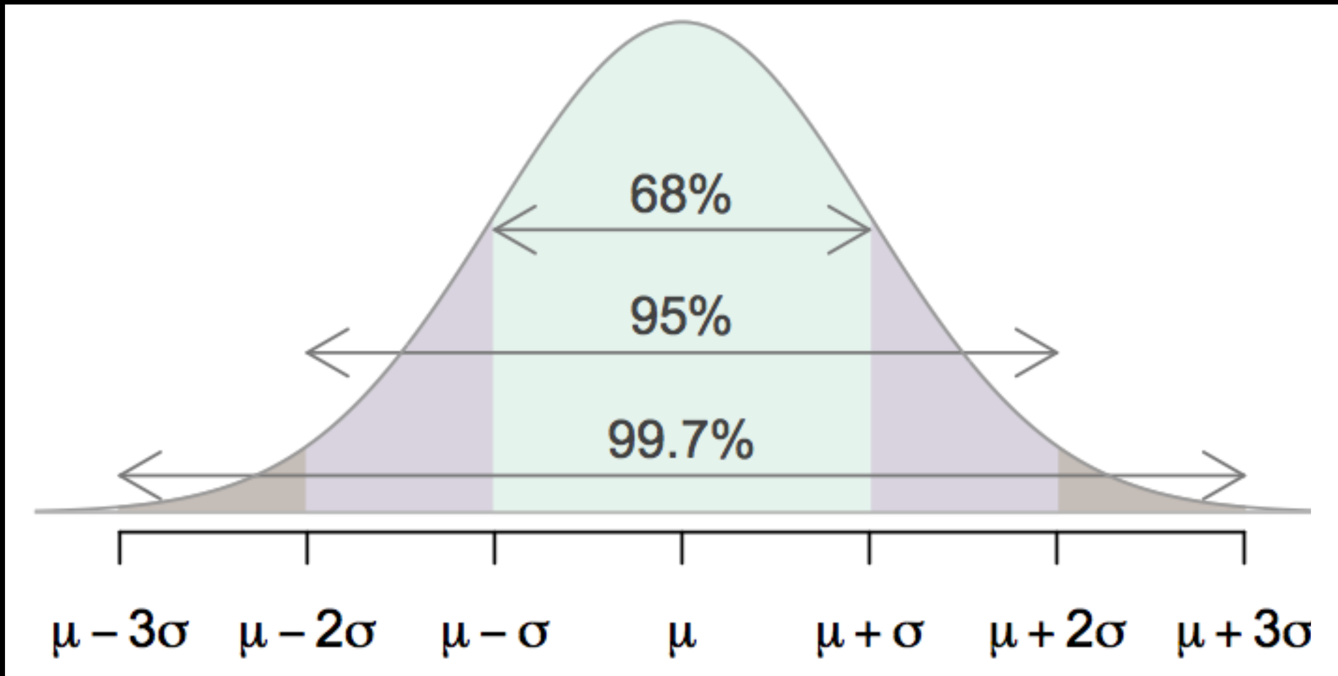
Confidence Intervals

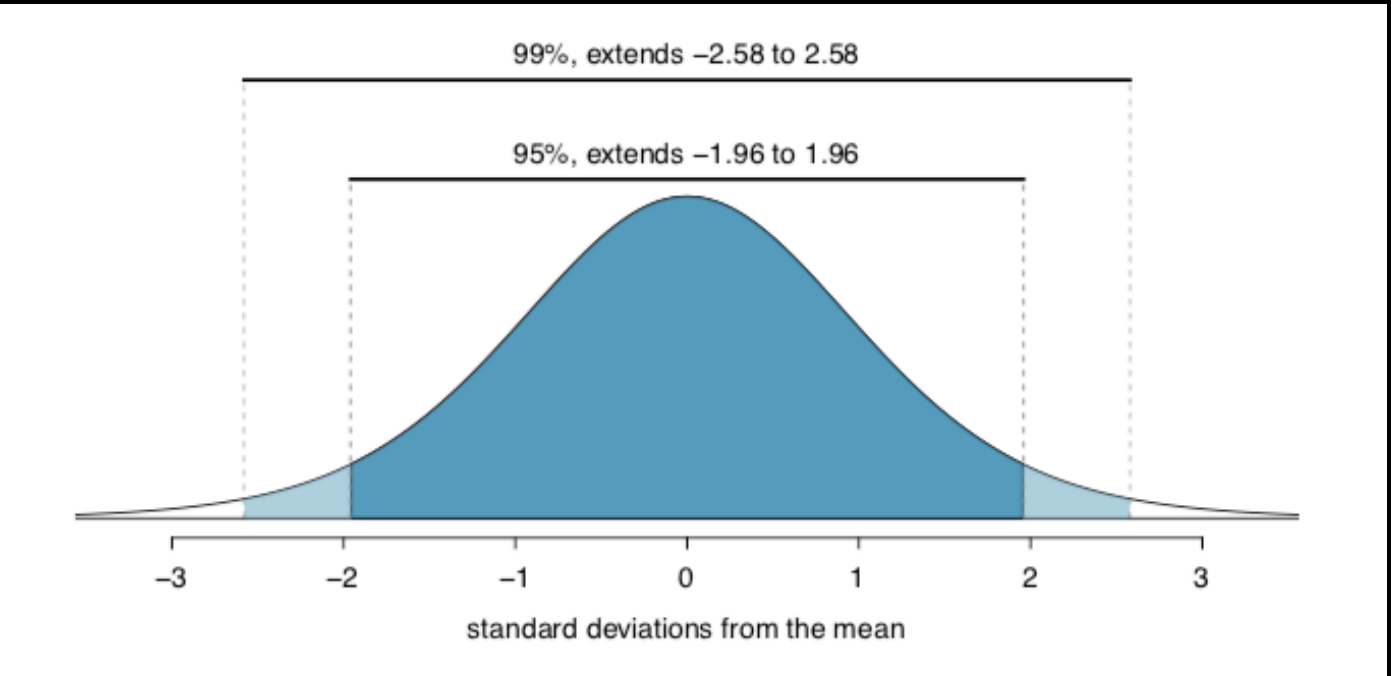


95% of the distribution is within approximately 2 SE from the mean



Confidence Intervals





Practically: We say the 95% confidence interval for a population's mean is:

sample mean +/- 1.96 X SE

Indeed, we can do this for any confidence interval requested in R.

Confidence Intervals: In general

point estimate ± z* x SE

"Point estimate" can be a mean, proportion, difference of means…

- In a confidence interval, *z* x SE* is called the margin of error, and for a given sample, the margin of error changes as the confidence level changes.

- In order to change the confidence level we need to adjust z* in the above formula.

Confidence Intervals: In general

point estimate ± z^* x SE

- In a confidence interval, $z^* \times SE$ is called the margin of error, and for a given sample, the margin of error changes as the confidence level changes.

- In order to change the confidence level we need to adjust z^* in the above formula.

- Commonly used confidence levels in practice are 90%, 95%, 98%, and 99%.

- For a 95% confidence interval, $z^* = 1.96$.

- However, using the standard normal (z) distribution, it is possible to find the appropriate z^* for any confidence level.

Overview of next 2 Classes

Hypothesis Testing Framework (Ch. 5-7)

The general outline of the process:

1. Set the hypotheses. ?
   For a single mean this will look like:
   ? H_0: μ = null value ?
   H_A: μ < or > or ≠ null value
2. Check assumptions and conditions
3. Calculate a test statistic ? and a p-value ?
4. Make a decision, and interpret it in context

- If p-value < α, reject H_0,
  ? there is sufficient evidence for [H_A]
- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Hypothesis Testing Framework (Ch. 5-7)

The general outline of the process:

English

1. Set the hypotheses.
    For a single mean this will look like:
        H_0: μ = null value
        H_A: μ < or > or $\neq$ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

Provides a rigorous way to determine the answer with a specific level of confidence.

Math

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

English

Hypothesis Testing Framework (Ch. 5-7)

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or ≠ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

How do we frame our question into the "null" and "alternative" hypothesis framework? What are these different hypotheses?

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Hypothesis Testing Framework (Ch. 5-7)

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or ≠ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]
- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

What distributions can we use to explore our sample?

Is our sample large or small?

e.g. if we are asking a question about sample means, do we expect our sample means to be normally distributed?

Use a normal distribution (Ch 5)? t-distribution (Ch 7)? Chi-square (Ch 6)?

Hypothesis Testing Framework (Ch. 5-7)

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or ≠ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

Calculate a number using our chosen distribution (e.g. the normal distribution) to see how “weird” a parameter of our sample is.

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Hypothesis Testing Framework (Ch. 5-7)

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or ≠ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Draw a "hard line" to determine if we can reject or we fail to reject the "null hypothesis"

Hypothesis Testing Framework (Ch. 5-7)

The general outline of the process:

We have actually been doing this mathematically already, you just didn't know!

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or $\neq$ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Hypothesis Testing Framework (Ch. 5-7)

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or $\neq$ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Let's look at some examples!

Hypotheses: Definition

In statistics a hypothesis means a very specific thing (slightly different then, for example, a science definition): it is a claim to be tested

H_0, Null Hypothesis: the "default", "standard" or currently accepted claim, the currently accepted value for a parameter. We start this process by assuming this is true.

H_A, Alternative Hypothesis: the "research" hypothesis, or claim we need to test

Possible Outcomes:

(1) We say we "reject the null hypothesis" - i.e. H_A is *more* true than H_0

(2) We say we "fail to reject the null hypothesis"

Note: we *cannot* say that H_A or H_0 is true, only that one is *more likely* to be true than the other.

Examples of stating Hypothesis: Practice #1

It is believed a candy machine makes peanut butter cups that are on average 5g. After maintenance, a worker claims the machine no longer makes the cups at a weight of 5g. What are H_0 and H_A? How do we write them in a statistical format?

The "default" or "previously assumed" claim is the null hypothesis

The alternative hypothesis is the claim to be tested

with math

$$H_0: \mu = 5g$$

$$H_A: \mu \neq 5g$$

population mean

Examples of stating Hypothesis: Practice #2

A company has stated their ping-pong machine makes ping-pongs that are 6mm in diameter. A worker believes the machine no longer makes ping-pongs of this size and samples 100 ping-pongs to perform a hypothesis test with 99% confidence. What are H_0 and H_A?

Think on it for a moment!

Examples of stating Hypothesis: Practice #3

Doctors believe that the average teen sleeps on average no longer than 10 hours per day. A researcher believes that teens on average sleep longer. What are H_0 and H_A?

Examples of stating Hypothesis: Practice #4

The school board claims that at least 55% of students bring an iPhone to school. A teacher believes this number is too high and randomly samples 25 students to test. What are H_0 and H_A?

Examples of stating Hypothesis: Practice #5

A super fan of shopping says that on average buying socks on ebay is cheaper than in person at their local shop. A price comparison study has shown that prices for new socks are on average the same or more expensive on ebay as in their local store. Our shopper wants to setup a statistical test to see if their intuition is right. What are H_0 and H_A?

Summary: Set the hypothesis

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or ≠ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

How do we frame our question into the "null" and "alternative" hypothesis framework? What are these different hypotheses?

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Summary: Set the hypothesis

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or ≠ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

tell us something about what tests we will perform (more in examples)

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Summary: Set the hypothesis

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or ≠ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Are we interested in a hypothesis about the population mean (μ)?

Proportion (p)? (Ch. 5)

Difference of 2 means and/or paired data (μ_1 - μ_2)? (Ch. 7)

Difference between observations and theorized results? (Ch. 6)

(more in examples)

Hypothesis Testing: Where we are going

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or $\neq$ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Picking appropriate distributions and applying - Rest of Ch 5, and 6 & 7

Hypothesis Testing: Where we are going

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or ≠ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

(a) normal, large sample
(b) normal?, small sample
(c) observations & theory

Test Statistics
(a) Z-score -> P(Z)
(b) T-Score -> P(T)
(c) χ^2 -> P(χ^2)

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

Hypothesis Testing: Where we are going

The general outline of the process:

1. Set the hypotheses.
   For a single mean this will look like:
   H_0: μ = null value
   H_A: μ < or > or ≠ null value
2. Check assumptions and conditions
3. Calculate a test statistic and a p-value
4. Make a decision, and interpret it in context

- If p-value < α, reject H_0,
  there is sufficient evidence for [H_A]

- If p-value > α, do not reject H_0,
  there is not sufficient for evidence for [H_A]

<u>Test Statistics</u>

(a) Z-score -> P(Z)

(b) T-Score -> P(T)

(c) χ^2 -> P(χ^2)

Compare Z-score, T-score or χ^2 to our <u>level of significance - α</u> - to see if we can reject the null hypothesis (if the p-value of our test statistic < α)

Anatomy of a test statistic

The general form of a test statistic is

$$\frac{\text{point estimate} - \text{null value}}{\text{SE of point estimate}}$$

Only tricks are:
(1) picking what the point and null values are based on our hypotheses

(2) what the form of the standard errors is based on what our underlying distribution looks like (normal, t-distribution, χ^{2})

This construction is based on

- identifying the difference between a point estimate and an expected value if the null hypothesis was true, and
- standardizing that difference using the standard error of the point estimate.

These two ideas will help in the construction of an appropriate test statistic for count data.