MODULE 4

INFERENCE AND SAMPLING DISTRIBUTIONS

OUTLINE

1. Frequentist vs. Bayesian statistical paradigms
2. Sampling error
3. Estimating sampling error with computer simulation
4. Estimating sampling error with statistical theory
5. Study questions

OUTLINE

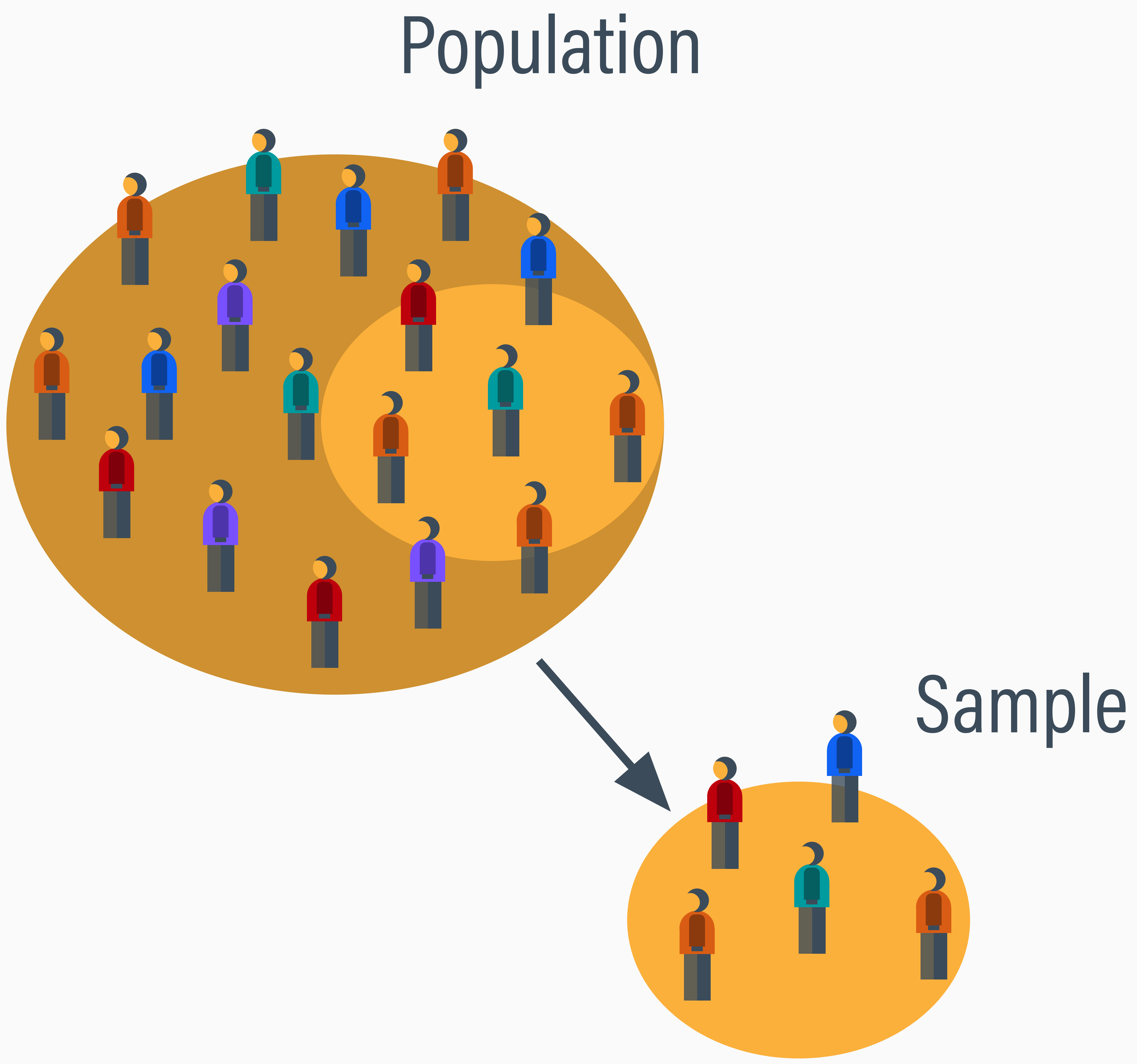
1. Frequentist vs. Bayesian statistical paradigms
2. Sampling error
3. Estimating sampling error with computer simulation
4. Estimating sampling error with statistical theory
5. Study questions

IMPORTANT TERMINOLOGY

Participants	Statistic
Population = all possible participants who share an attribute of interest	Parameter = a hypothetical statistic computed using the full population
Sample = the subset of people who participated in the study	Estimate = an observed statistic computed from the sample data

POPULATIONS VS. SAMPLES

- A population is the entire group of individuals that are of interest in a study
- Researchers almost exclusively work with a smaller subset called a sample
- The usual goal is use a sample statistic as a best guess about the population statistic

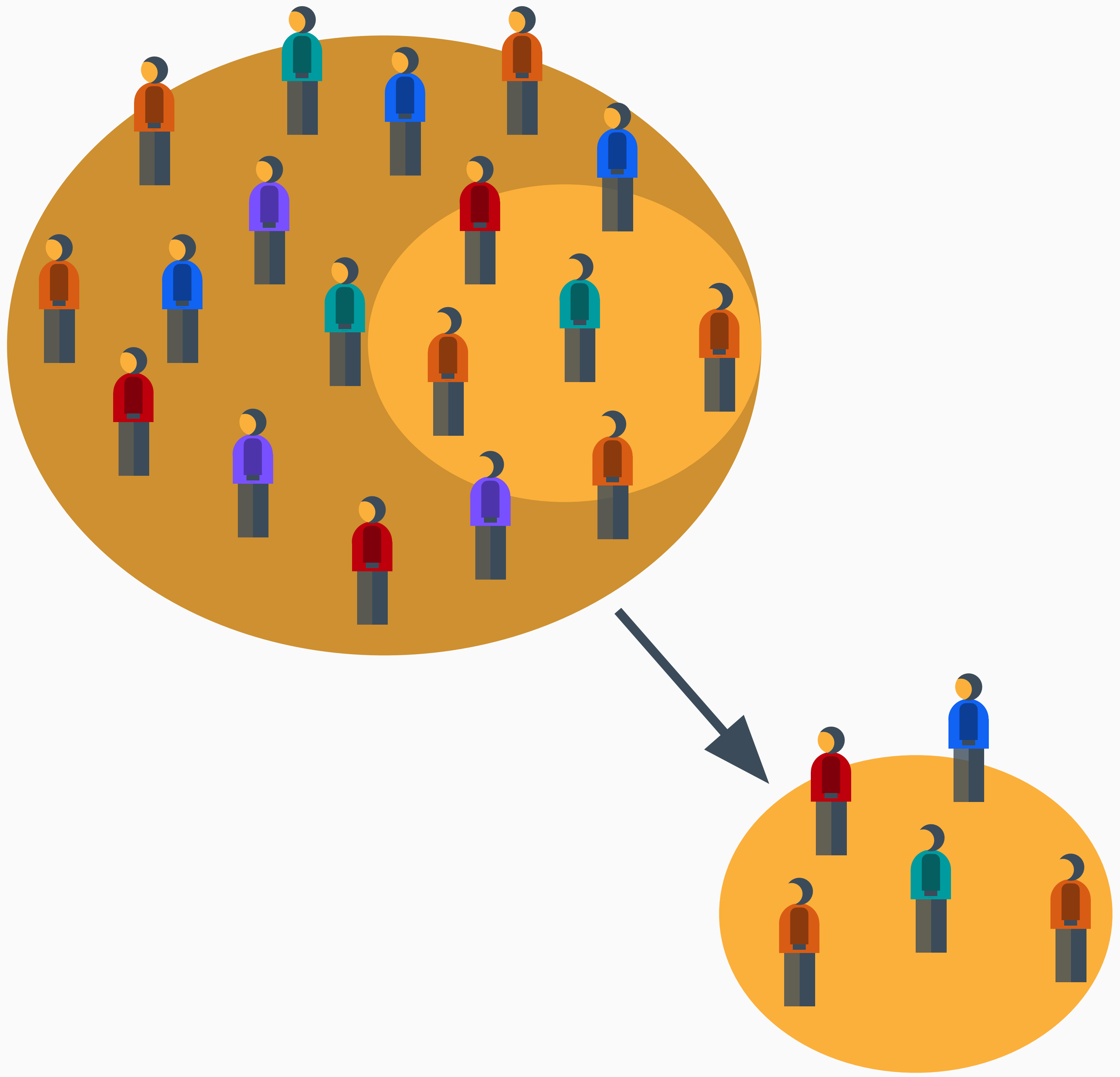


PARAMETERS VS. ESTIMATES

- Greek letters μ and σ denote the true population-level mean and standard deviation (parameters)
- Roman letters $\bar{X}$ and s reference a sample's mean and standard deviation (estimates)
- Ideally, the sample statistics will be close to the true values, but they will always differ

Population mean = μ

Population standard deviation = σ



Sample mean = $\bar{X}$

Sample standard deviation = s

CONTRASTING STATISTICAL PARADIGMS

- Two main frameworks for statistical inference (generalizing from a sample to a larger population): frequentist and Bayesian

- Frequentist: Assumes a single population that spawns many hypothetical random samples of data (one parameter, many hypothetical estimates)

- Bayesian: Assumes a many hypothetical populations that could spawn a single sample of data (many hypothetical parameters, one estimate)

FREQUENTIST VS. BAYESIAN SUMMARY

Frequentist

- The parameter is a fixed quantity, estimates vary across different samples
- Statements about probability, precision, and confidence refer to estimates
- Probability = the long run frequency of an event across many different samples

Bayesian

- There is a single sample, and parameters vary across different populations
- Statements about probability, precision, and intervals refer to the parameter
- Probability = our degree of certainty about a parameter after analyzing data

FREQUENTIST STATISTICAL PARADIGM

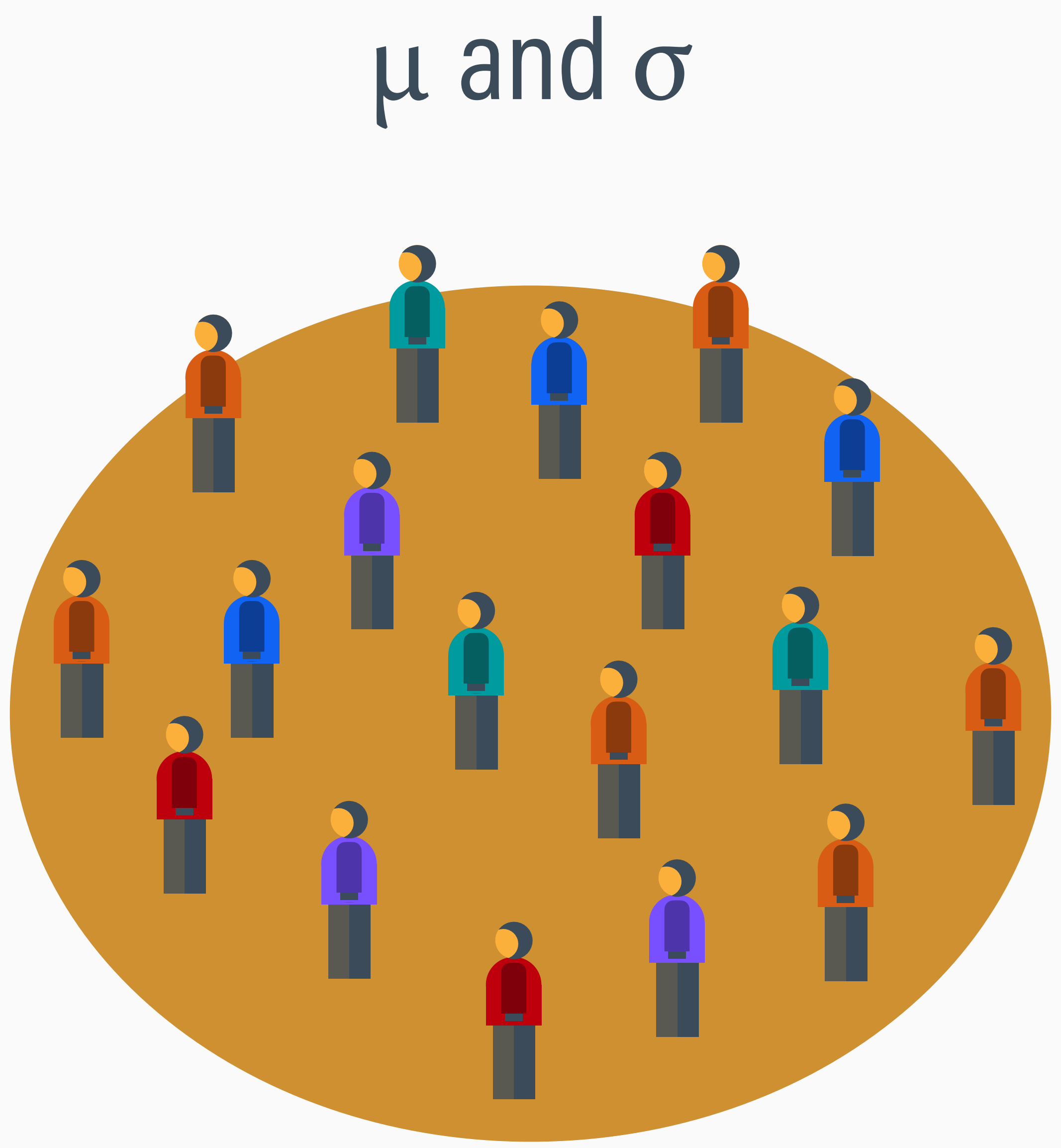
- The frequentist paradigm is defined by the idea that there is one population with unknown parameters (e.g., μ and σ)

- We imagine numerous hypothetical samples of size N from the population, each with its own unique estimates (e.g., $\bar{X}$ and s)

- The population-level statistics (parameters) are locked in at a single set of values, whereas the sample-level statistics (estimates) vary across different data sets

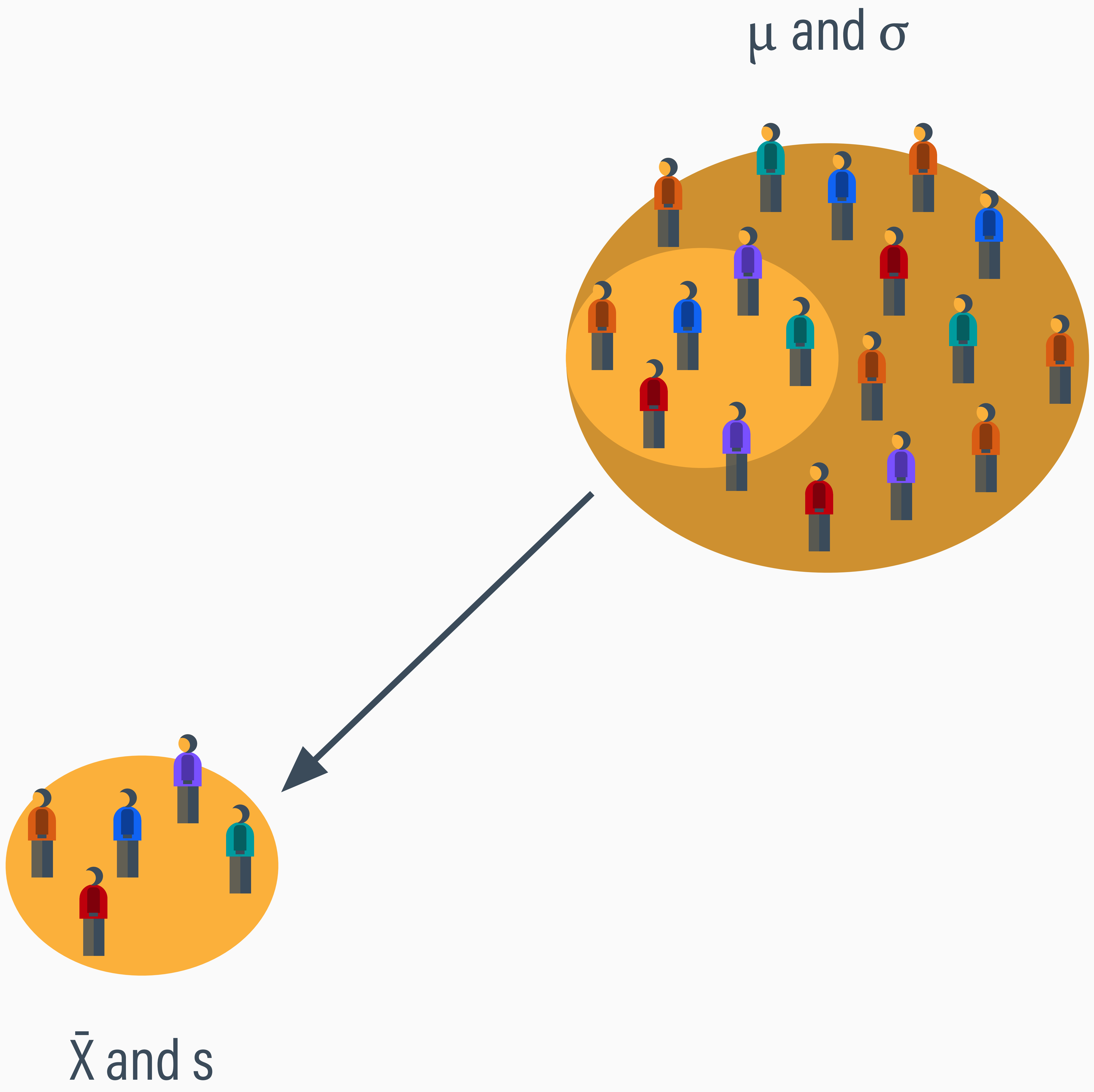
FREQUENTIST FRAMEWORK

One population with unknown parameters μ and σ



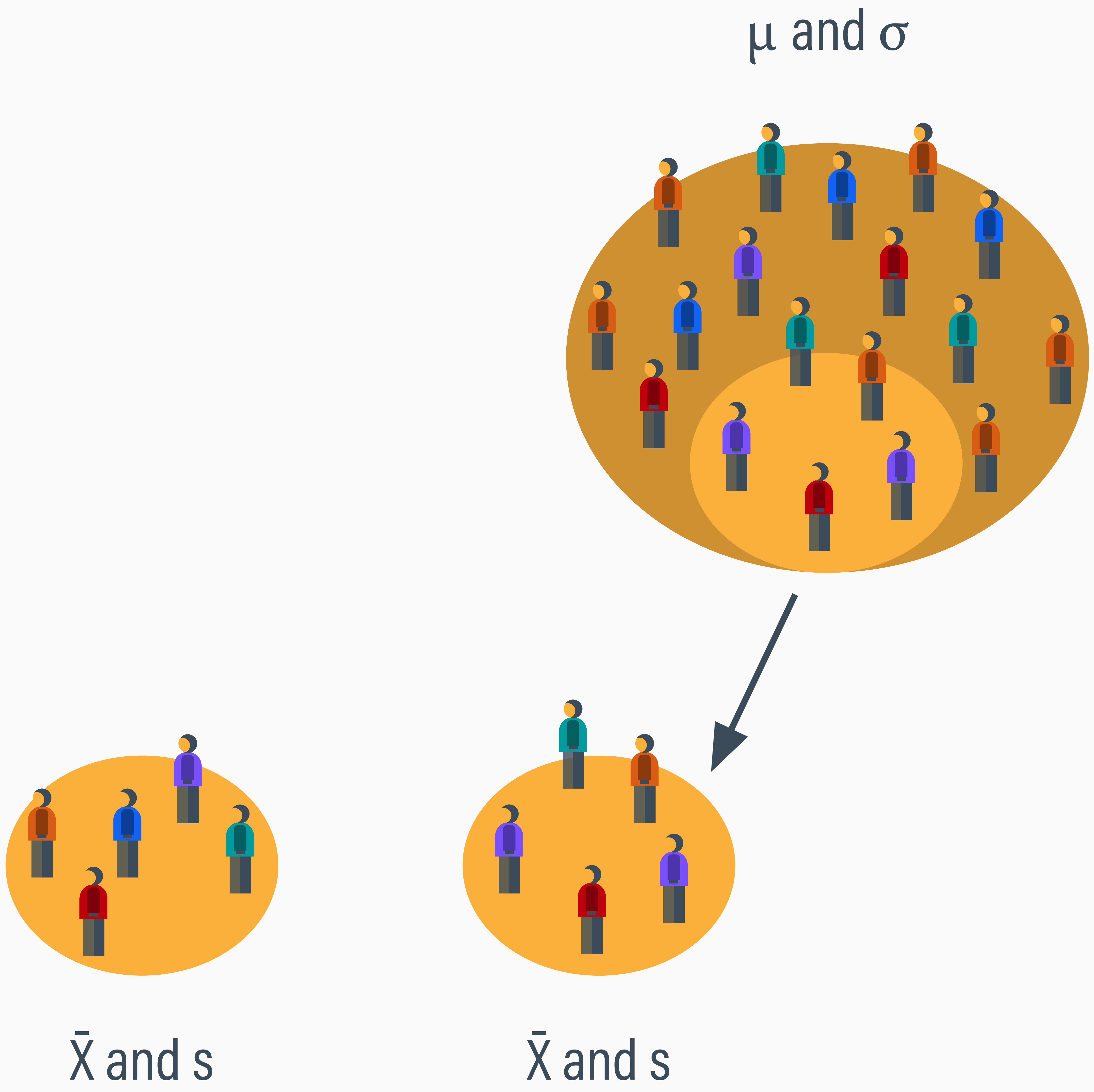
FREQUENTIST FRAMEWORK

One population with unknown parameters. Estimates vary across different hypothetical samples.



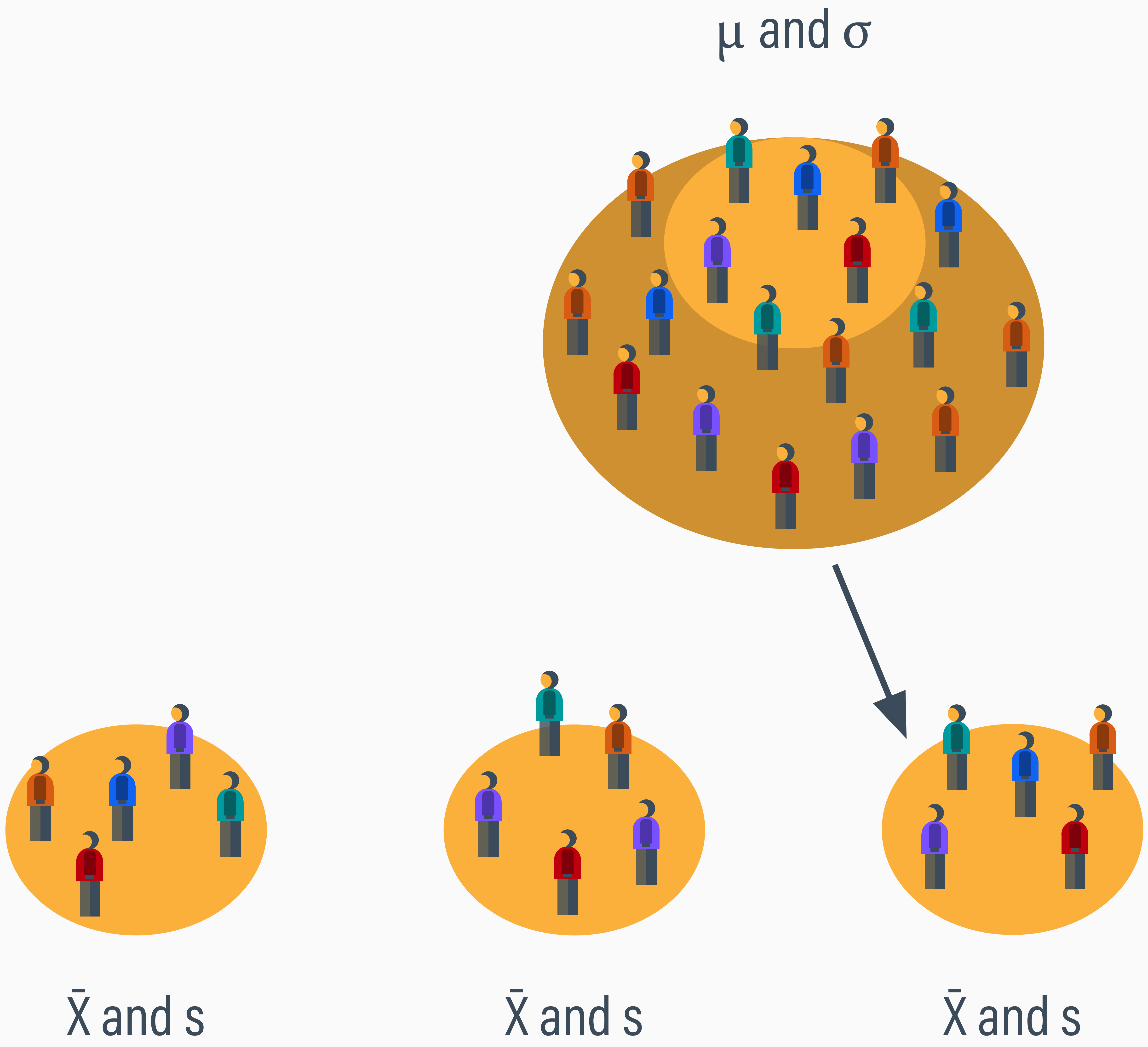
FREQUENTIST FRAMEWORK

One population with unknown parameters. Estimates vary across different hypothetical samples.



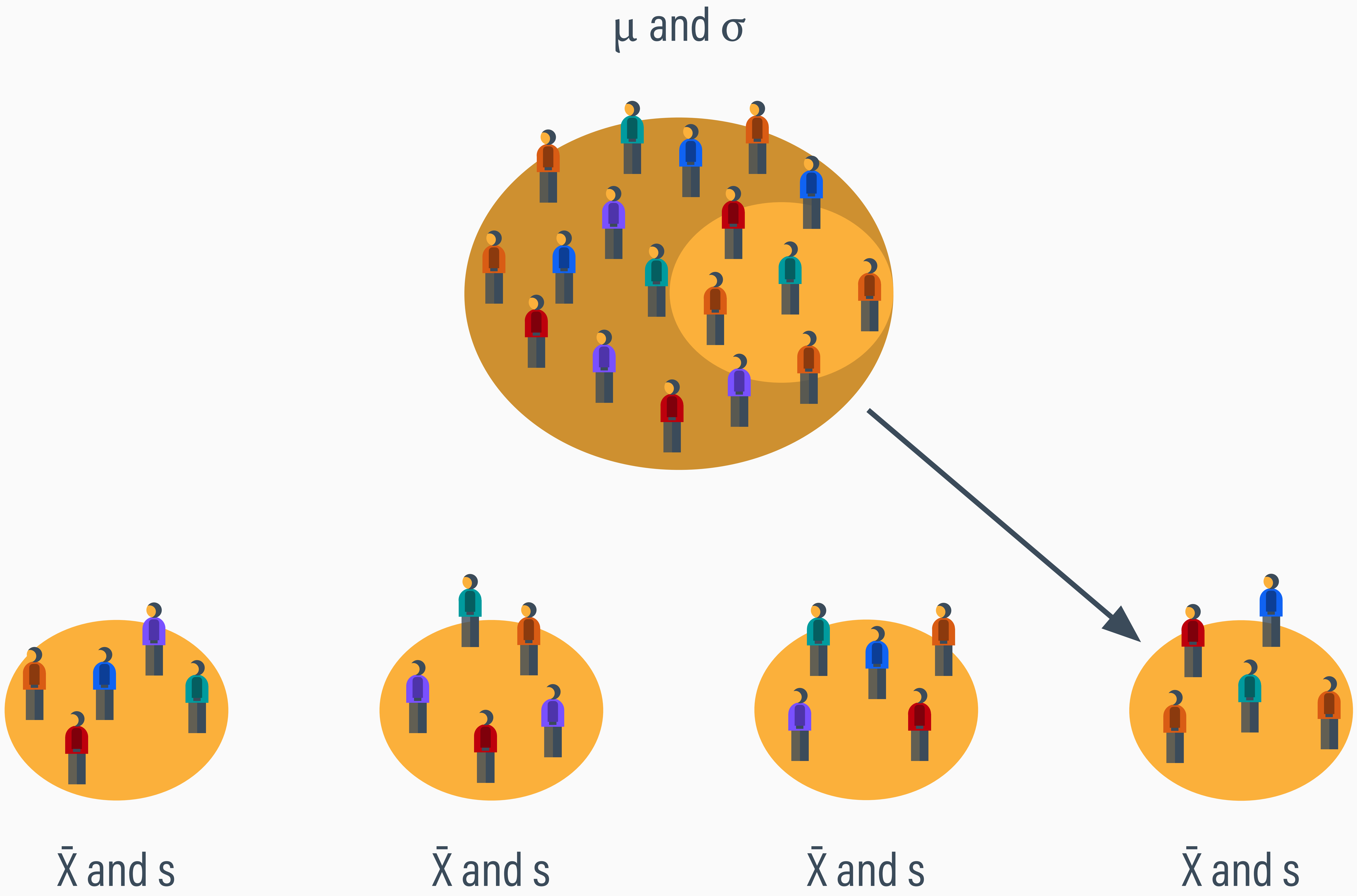
FREQUENTIST FRAMEWORK

One population with unknown parameters. Estimates vary across different hypothetical samples.



FREQUENTIST FRAMEWORK

One population with unknown parameters. Estimates vary across different hypothetical samples.

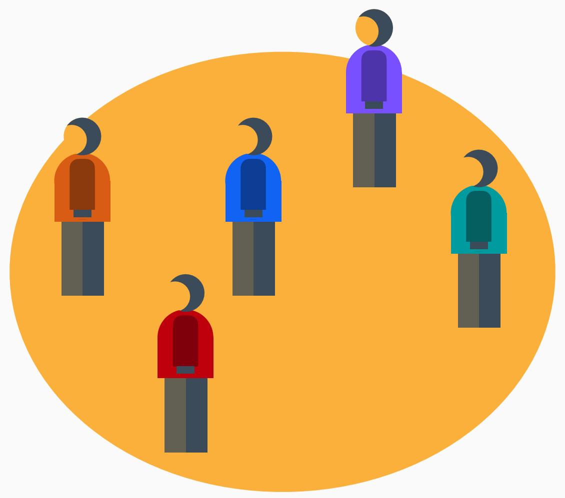


BAYESIAN STATISTICAL PARADIGM

- The Bayesian paradigm is defined by the idea that there is one sample of size N that we observe (no hypothetical samples)
- We imagine numerous hypothetical populations that could have produced our particular sample of data, each with unique parameter values (e.g., many values of μ and σ)
- The sample data are locked, and parameter values vary across different hypothetical populations

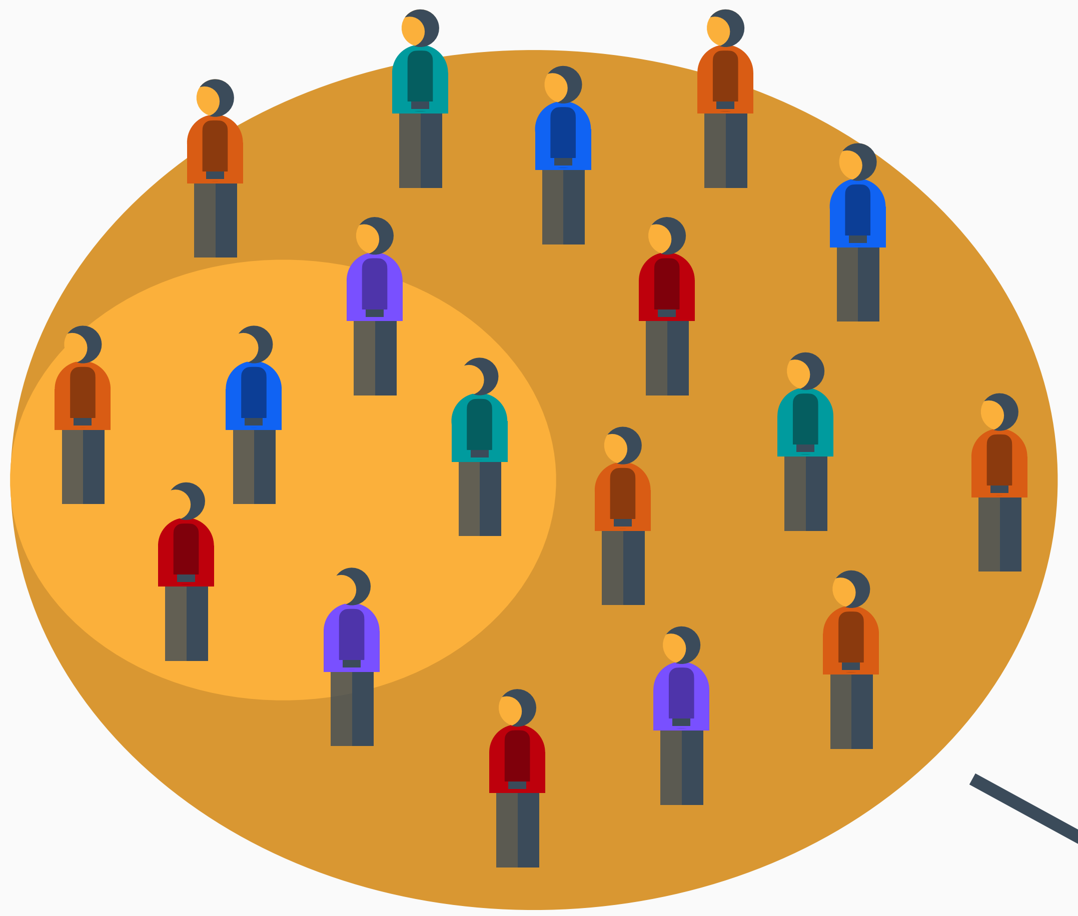
BAYESIAN FRAMEWORK

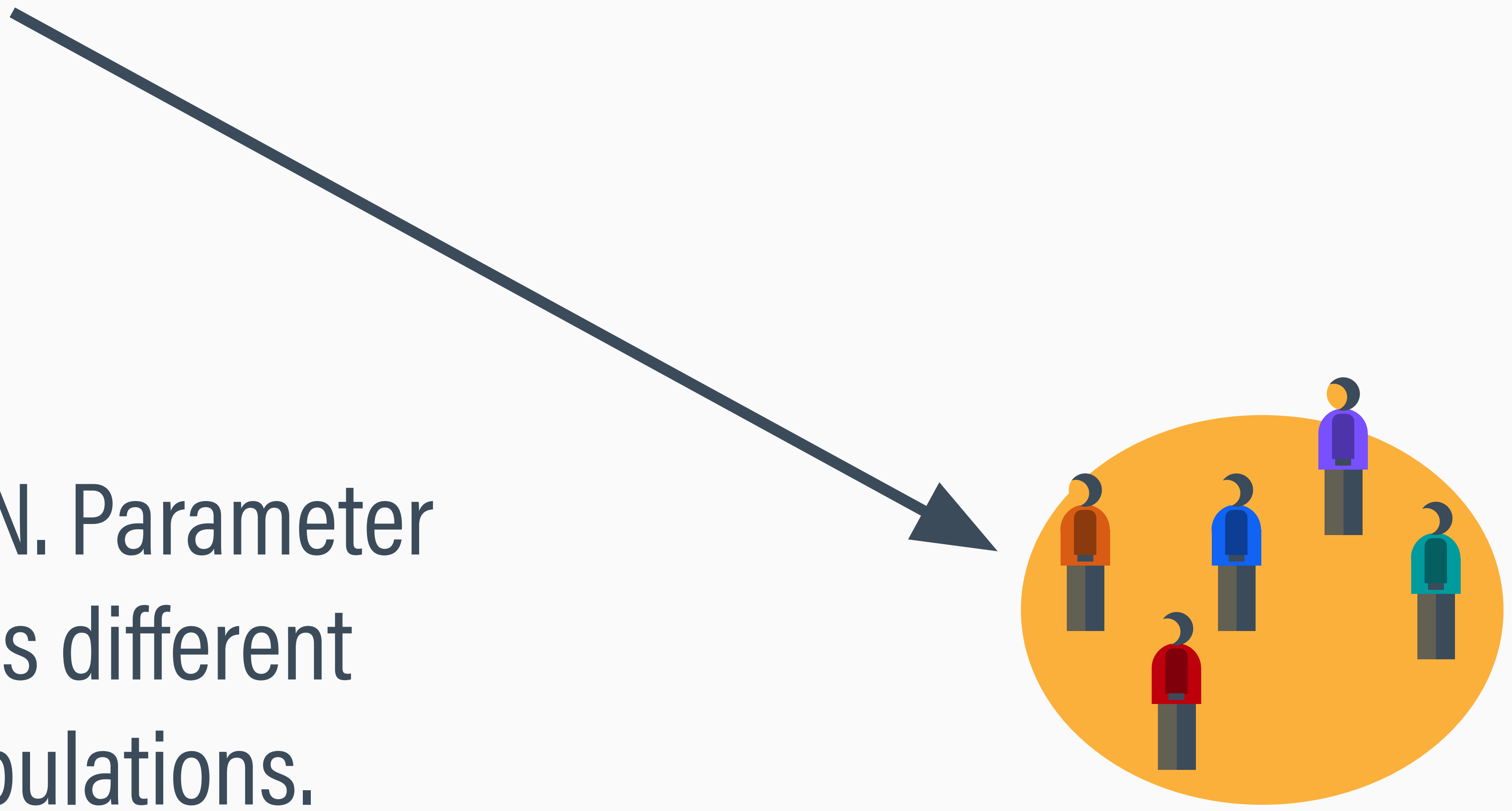
One sample of size N



BAYESIAN FRAMEWORK

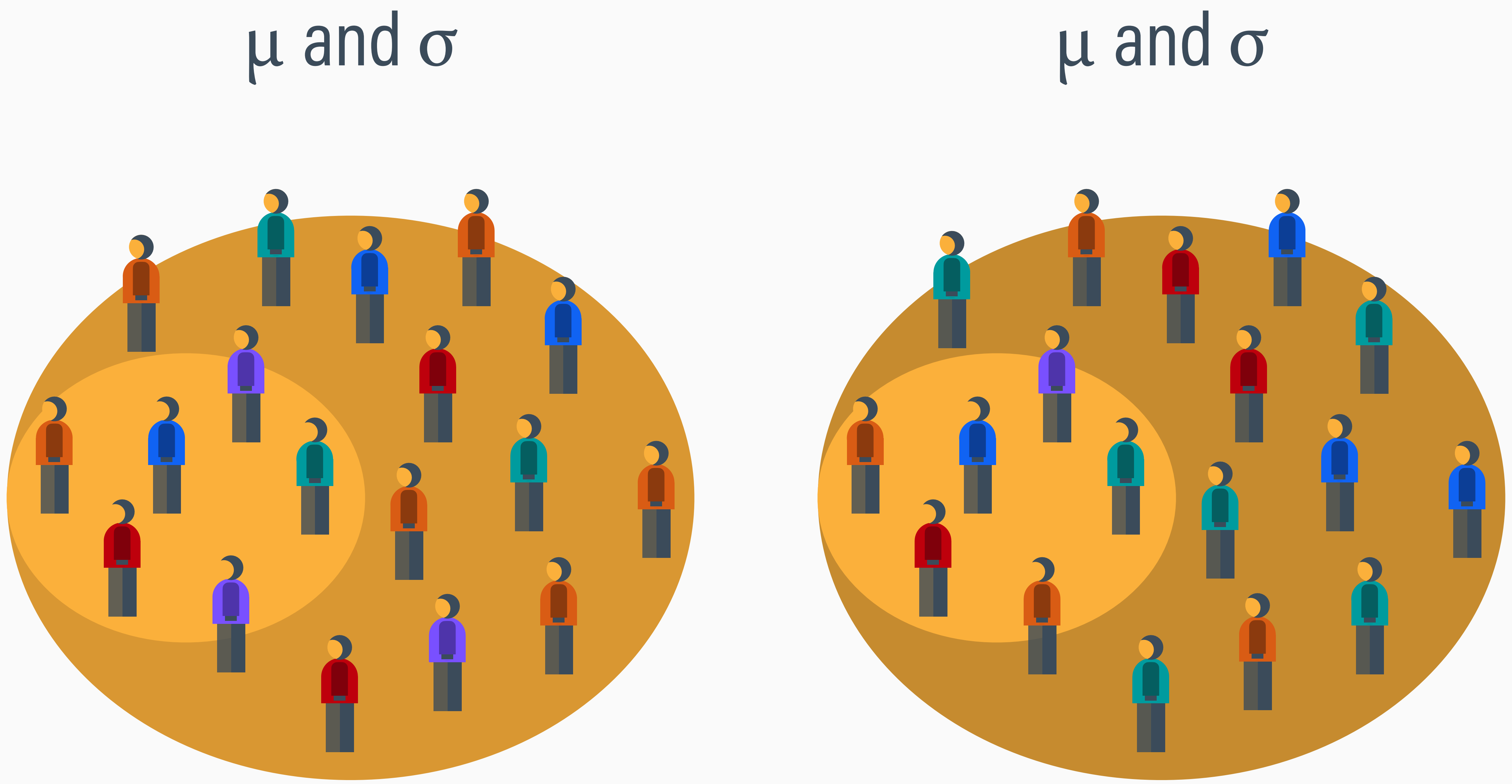
μ and σ

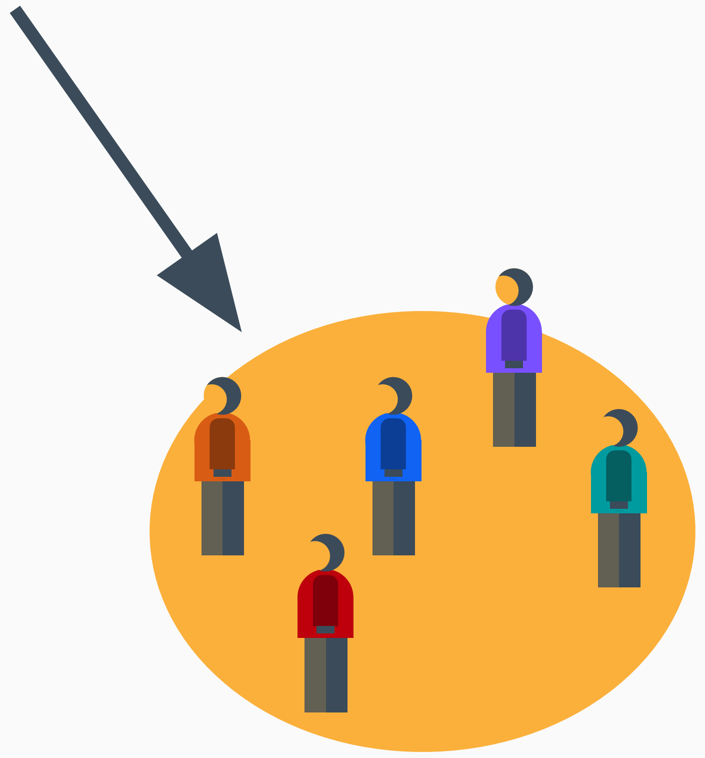




One sample of size N. Parameter values vary across different hypothetical populations.

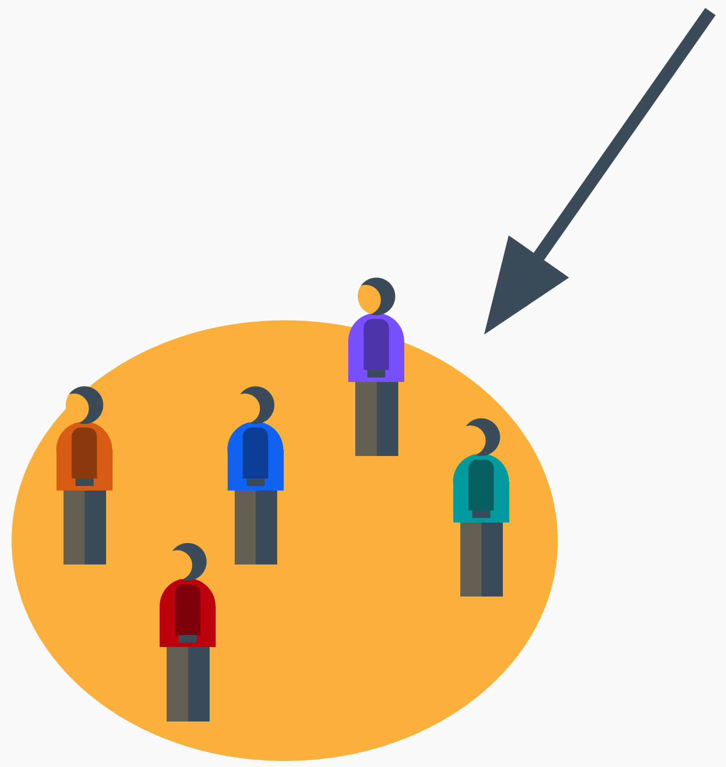
BAYESIAN FRAMEWORK





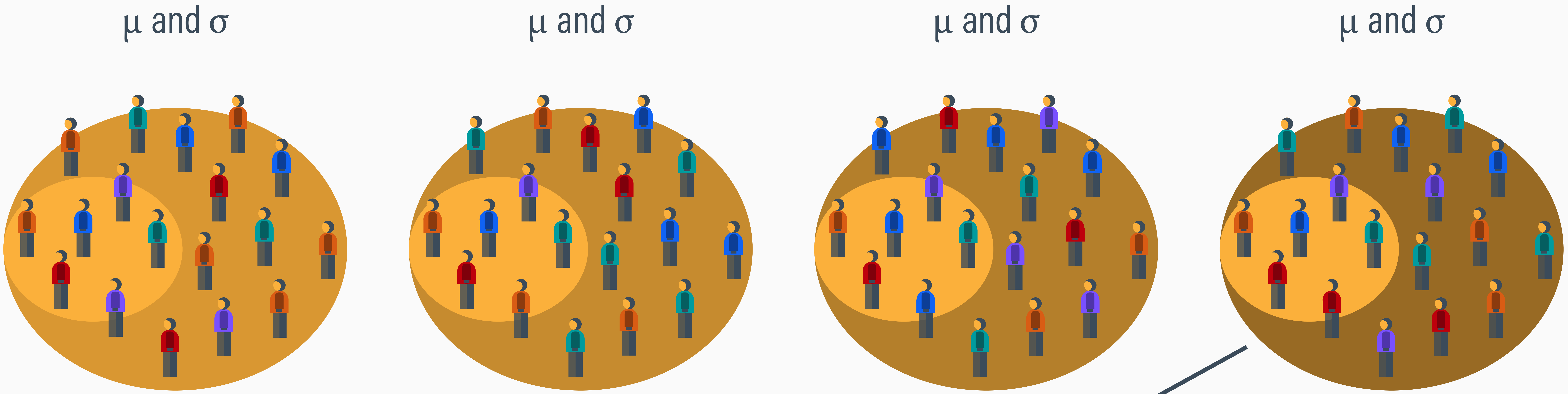
One sample of size N. Parameter values vary across different hypothetical populations.

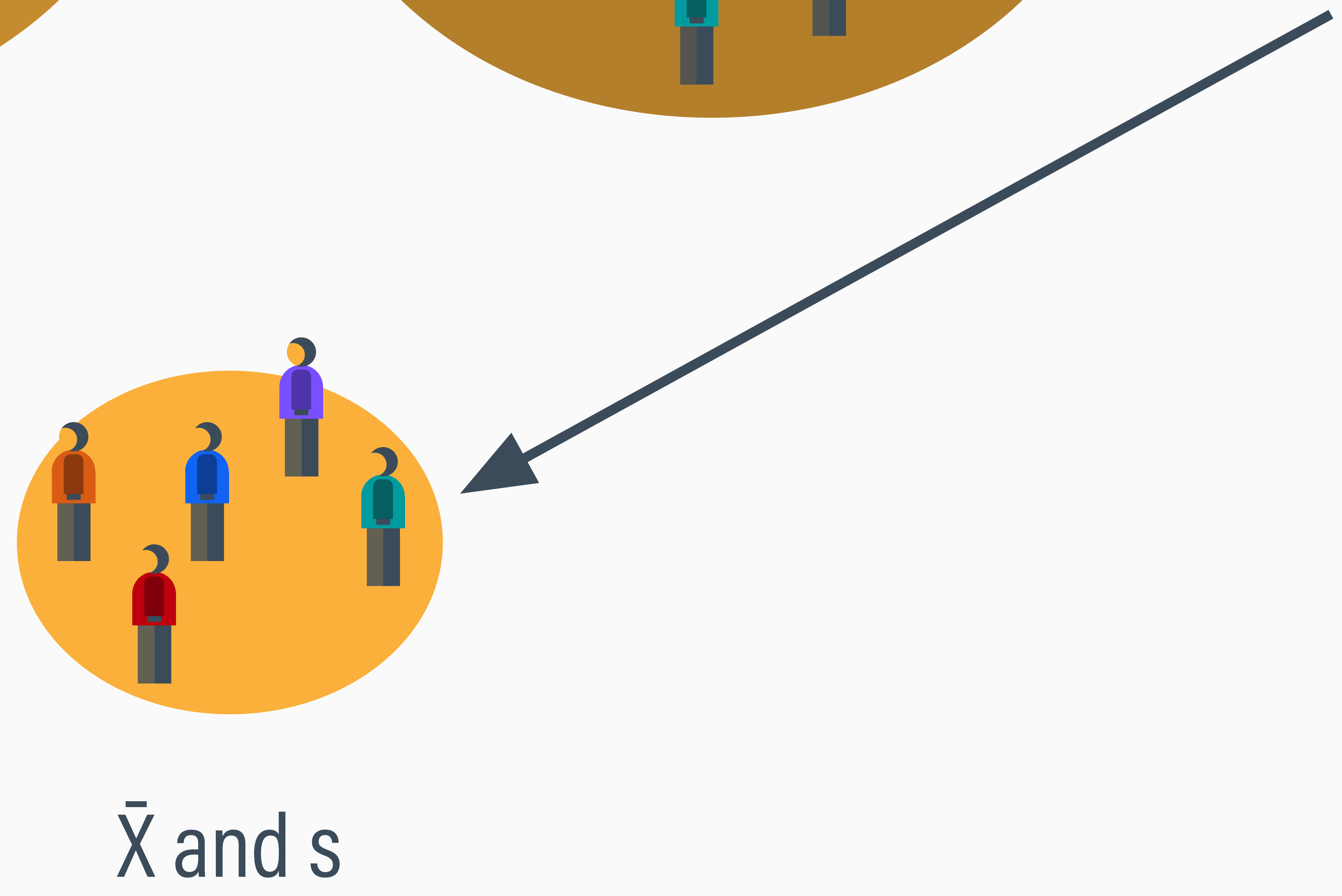
BAYESIAN FRAMEWORK



One sample of size N. Parameter values vary across different hypothetical populations.

BAYESIAN FRAMEWORK





One sample of size N. Parameter values vary across different hypothetical populations.

OUR FOCUS: FREQUENTIST INFERENCE

- The frequentist paradigm is the predominant statistical framework in psychology and the behavioral sciences
- The idea that estimates vary across different hypothetical samples is integral to foundational concepts like p-values, confidence intervals, and statistical significance
- Everything we will do going forward relies on the hypothetical process of drawing many samples of data

OUTLINE

1. Frequentist vs. Bayesian statistical paradigms
2. Sampling error
3. Estimating sampling error with computer simulation
4. Estimating sampling error with statistical theory
5. Study questions

SMOKING AND DRINKING CESSATION TRIAL

Pharmacological treatments that can concomitantly address cigarette smoking and heavy drinking stand to improve health care delivery for these highly prevalent co-occurring conditions. This superiority trial compared the combination of varenicline and naltrexone against varenicline alone for smoking cessation and drinking reduction among heavy-drinking smokers.

Lara Ray

ReJoyce Green

Ray, L.A., Green, R., Enders, C., et al. (2021). Efficacy of combining varenicline and naltrexone for smoking cessation and drinking reduction: A randomized clinical trial. *American Journal of Psychiatry*, *178*, 818–828.

KEY VARIABLES

Breath (alveolar) carbon monoxide

A measure of carbon monoxide in the lungs. Breath carbon monoxide is a biomarker of smoking behavior common in clinical trials. Higher scores reflect more frequent smoking.

Medication arm

Participants were randomly assigned to receive one of two meds: varenicline plus naltrexone or varenicline plus placebo pills

HYPOTHETICAL POPULATION

- To illustrate key concepts, suspend reality and pretend the entire population of smokers comprises only 10 individuals

Breath CO level	4	8	5	5	6	6	5	6	7	8

POPULATION PARAMETER

- The true mean of the population of all possible participants is $\mu = 6$

$$\mu = \frac{4 + 8 + 5 + 5 + 6 + 6 + 5 + 6 + 7 + 8}{10} = 6$$

4	8	5	5	6	6	5	6	7	8

MANY HYPOTHETICAL SAMPLES

- In practice, researchers must work with samples because they are unable to access the entire population
- For any given study, it is hypothetically possible to draw countless sample of size N from the population
- Thus, the participants in a study represent just one of many hypothetical samples from the population

MANY HYPOTHETICAL ESTIMATES

- Suppose the project's budget allows the researchers to run the trial with a sample of N = 3 participants
- The mean breath CO from a sample of N = 3 participants almost certainly differs from the true mean in the population
- Each hypothetical sample will produce a different estimate

HYPOTHETICAL SAMPLE 1

Population

4	8	5	5	6	6	5	6	7	8

Parameter
$\mu = 6$

Sample

4	5	6

Estimate
$\bar{X} = 5.67$

HYPOTHETICAL SAMPLE 2

Population

4	8	5	5	6	6	5	6	7	8

Parameter

$\mu = 6$

Sample

8

7	8

Estimate

$\bar{X} = 7.67$

HYPOTHETICAL SAMPLE 3

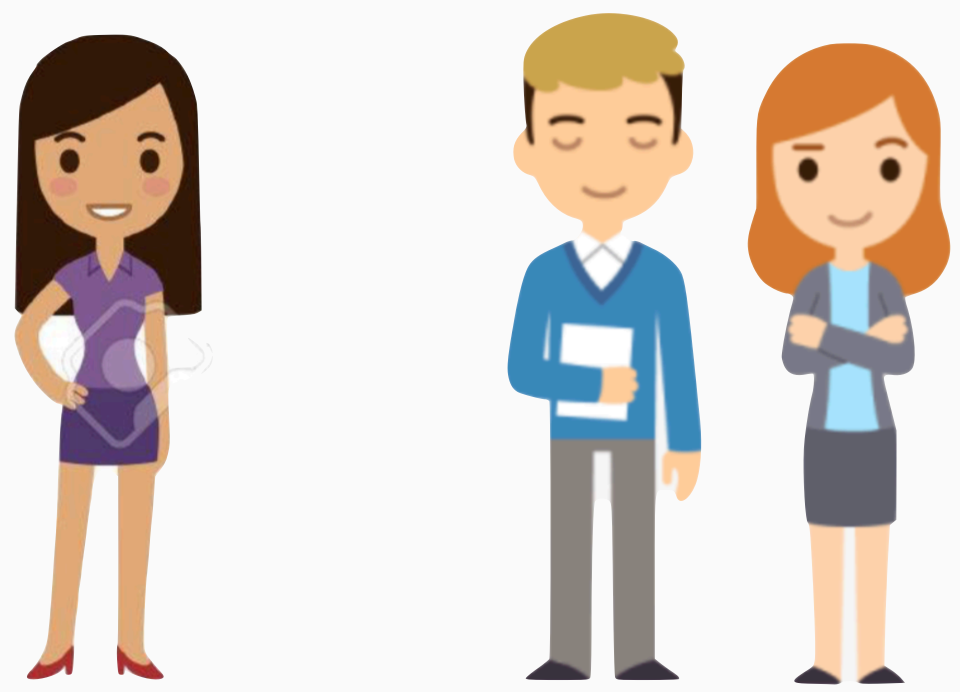
Population

4 8 5 5 6 6 5 6 7 8

Parameter

$\mu = 6$

Sample



5 6 5

Estimate

$\bar{X} = 5.33$

HYPOTHETICAL SAMPLE 4

Population

4 8 5 5 6 6 5 6 7 8

Parameter

$\mu = 6$

Sample

5

6 7

Estimate

$\bar{X} = 6.00$

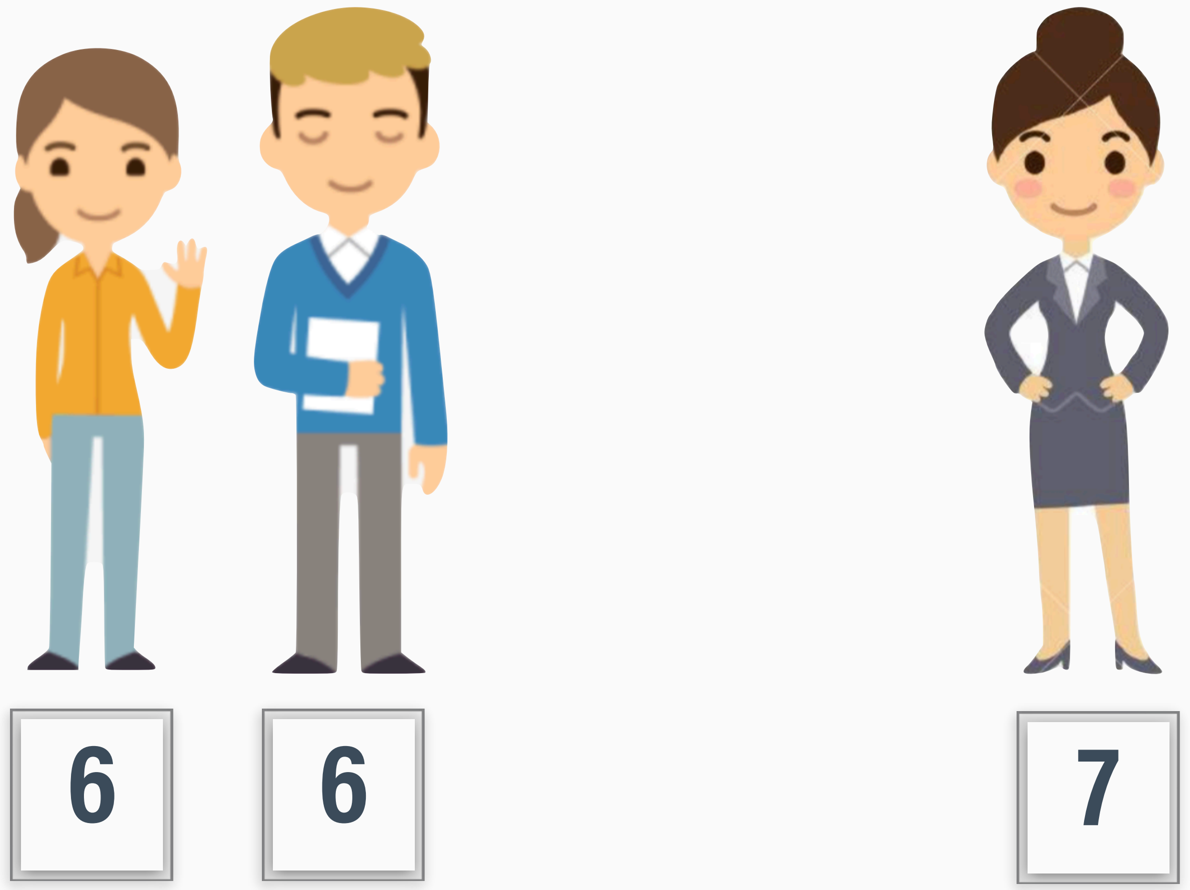
HYPOTHETICAL SAMPLE 5

Population

Parameter

$\mu = 6$

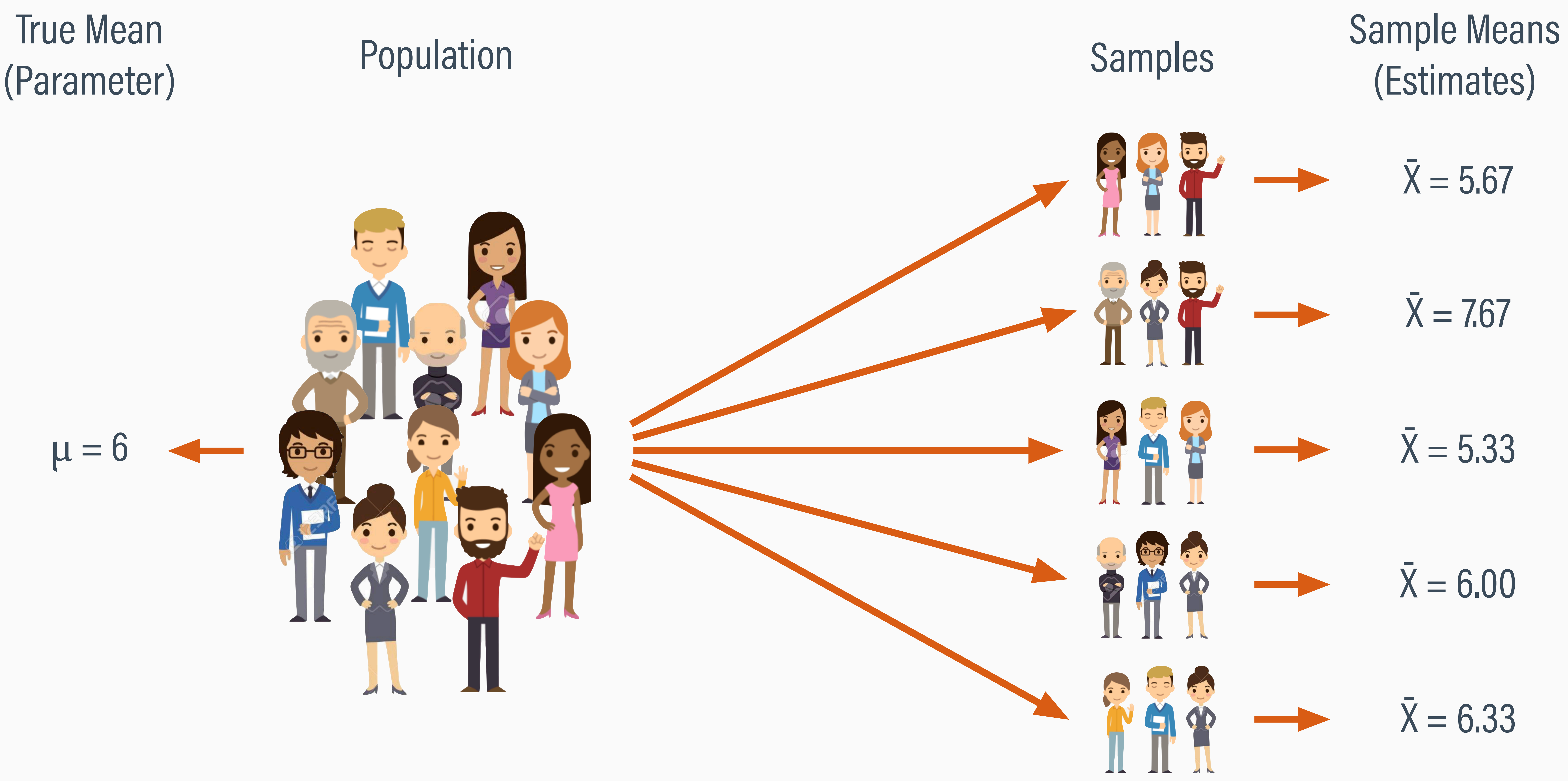
Sample



Estimate

$\bar{X} = 6.33$

RANDOM SAMPLING SUMMARY



IMPORTANT OBSERVATIONS SO FAR

- The true mean is a single value obtained by analyzing the entire population’s data (the parameter never changes)
- Each hypothetical sample produces different estimates
- Estimates can be higher or lower than the true parameter value, and some are closer to the truth than others

SAMPLING ERROR

- Sampling error is the amount by which an estimate differs from the true population statistic

$$\text{Sampling error} = \bar{X} - \mu$$

- Errors can be positive or negative, and each hypothetical sample gives an estimate with a different amount of error

SAMPLING ERROR ILLUSTRATION

- Each unique sample we could work with has a different mean and different error
- Errors can be positive or negative, and they can be large or small (or even zero)

Parameter (the truth) → μ

Estimate (best guess) → $\bar{X}$

Sample	$\bar{X}$	μ	Sampling Error
1	5.67	6	-0.33
2	7.67	6	1.67
3	5.33	6	-0.67
4	6.00	6	0
5	6.33	6	0.33

In small groups of two or three, discuss the relationship between sampling errors from different hypothetical samples and deviation scores from different individuals. How is sampling error similar to a deviation score, and how is it different?

The amount of error varies across samples. In small groups of two or three, discuss how you could use the information from the table to estimate the "typical" or "expected" amount of sampling error from a sample of N = 3.

Sample	$\bar{X}$	μ	Sampling Error
1	5.67	6	-0.33
2	7.67	6	1.67
3	5.33	6	-0.67
4	6.00	6	0
5	6.33	6	0.33

ESTIMATING SAMPLING ERROR

- Estimating sampling error is vital because it tells us how close our sample statistic is to the truth in the population
- The previous illustration was purely hypothetical—in real-world applications, we have only one sample of data, and we never know the true population statistic
- Three approaches to estimating sampling error: computer simulation, statistical theory, and bootstrapping

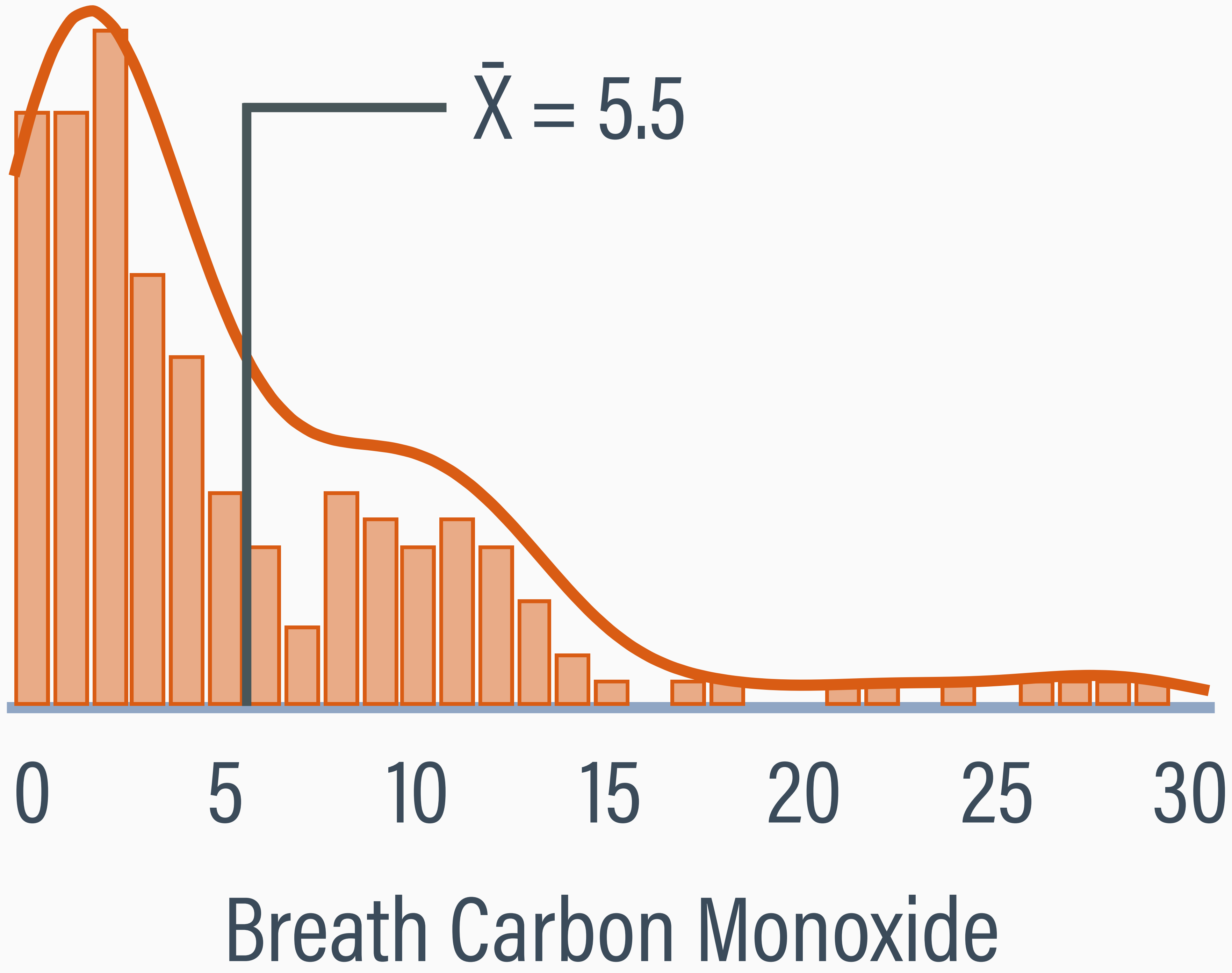
OUTLINE

1. Frequentist vs. Bayesian statistical paradigms
2. Sampling error
3. **Estimating sampling error with computer simulation**
4. Estimating sampling error with statistical theory
5. Study questions

SAMPLE STATISTICS

Mean ($\bar{X}$)	5.5
Std. Dev. (s)	6
N	165

- This sample of N = 165 produced a sample mean of $\bar{X} = 5.5$
- How different is the estimate from the true mean? How much sampling error would we expect?



COMPUTER SIMULATION STRATEGY

- The sample mean of 5.5 could be higher or lower than the true population mean, and it could be similar or different
- Sampling error differs for each hypothetical sample
- We can use computer simulation to generate artificial datasets and then calculate the average error across numerous hypothetical estimates

COMPUTER SIMULATION RECIPE

1. Provide the true mean and standard deviation in a hypothetical population (a real data set can provide good guesses)

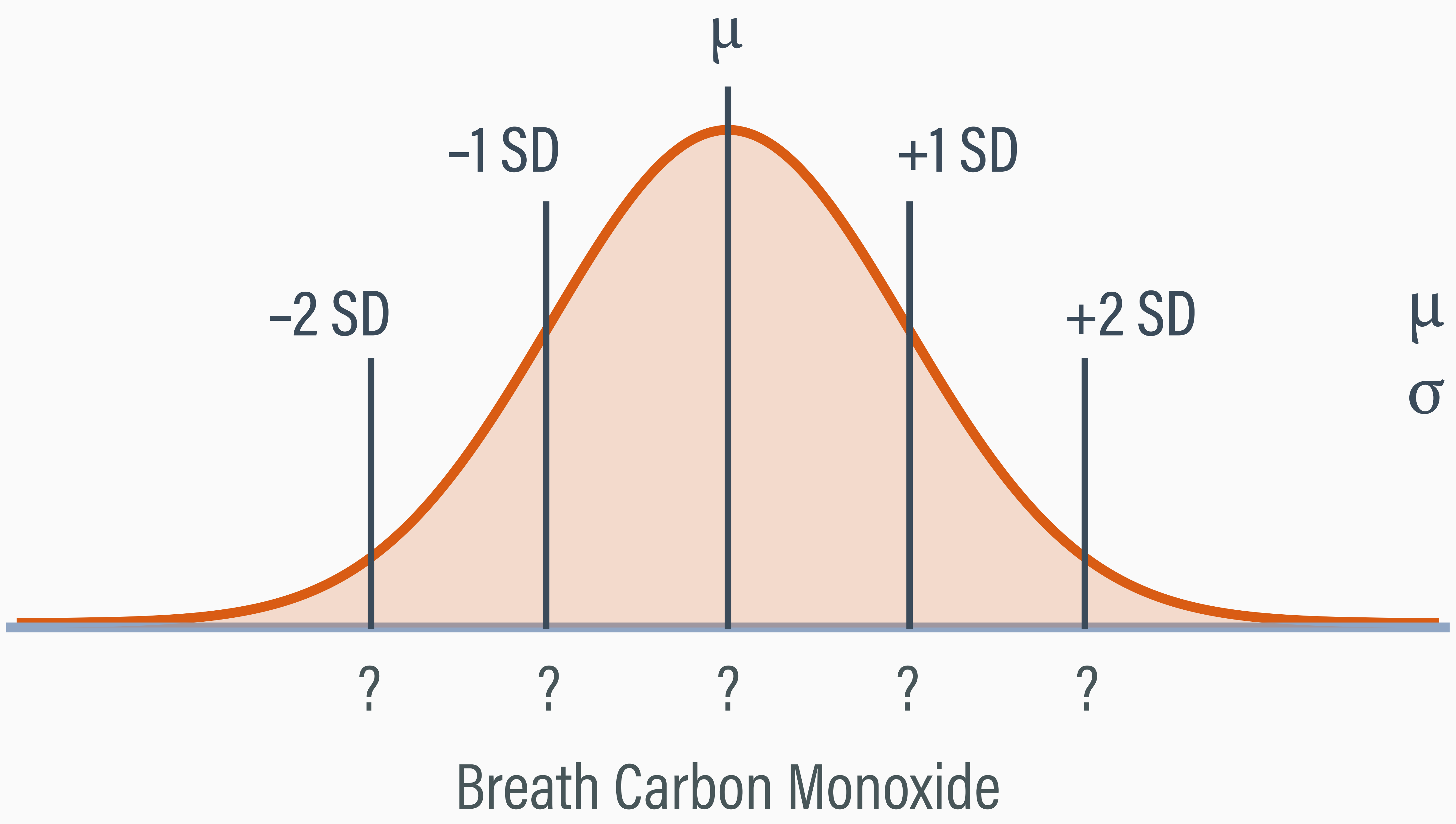
2. Use R to create many random samples of artificial data from the hypothetical population

3. Compute the mean from each artificial data set

4. Compute the sampling error for each estimate (i.e., estimate minus true mean), the average the sampling errors

STEP 1: DEFINE POPULATION PARAMETERS

- For simplicity, we can assume that population-level smoking scores follow a normal distribution

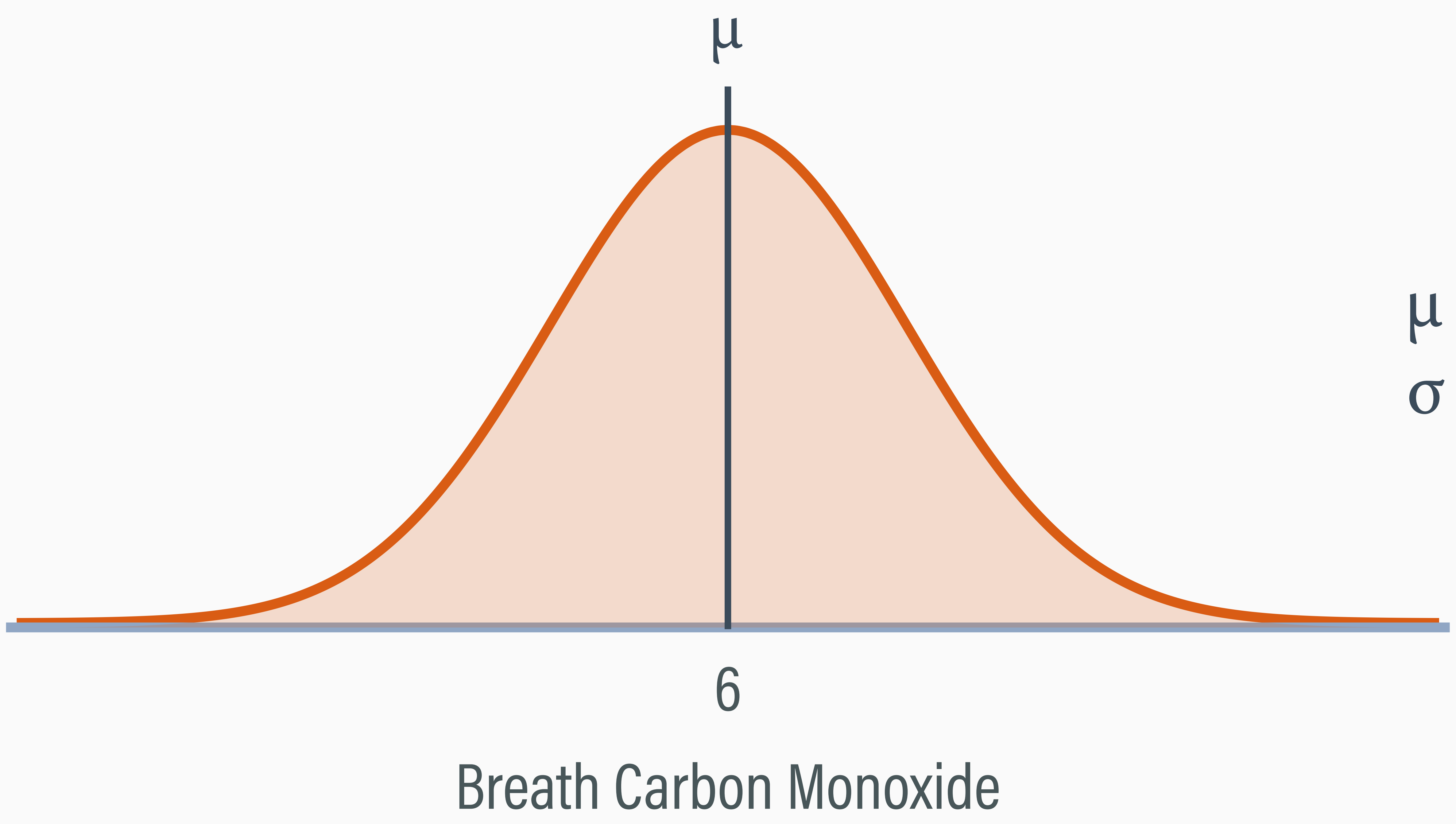


We need to specify reasonable values for the population statistics

$\mu = ?$
$\sigma = ?$

STEP 1: DEFINE POPULATION MEAN

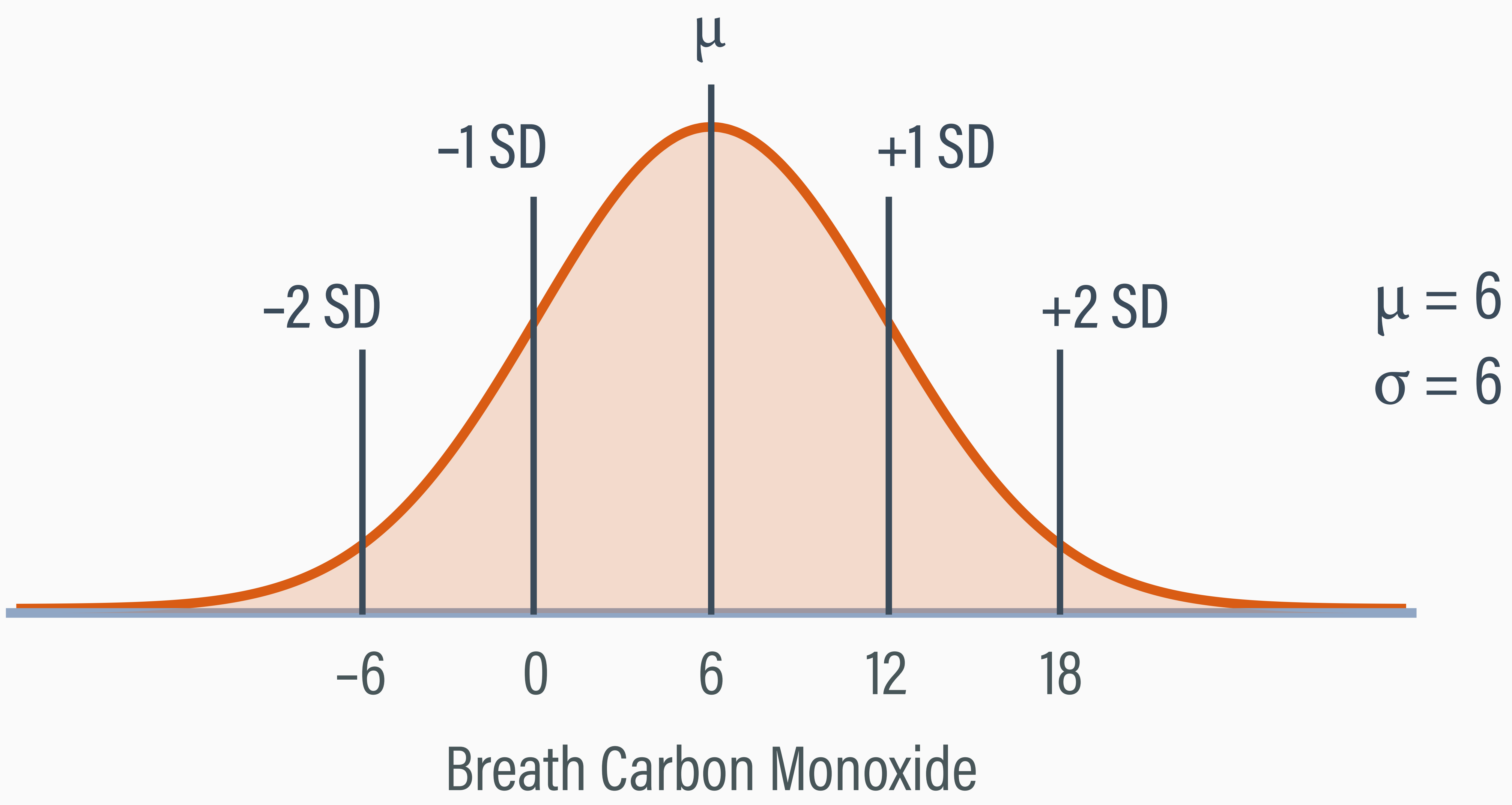
- A breath CO level of 6 is strongly indicative of smoking, so we can use that as the true mean (any value for μ works fine)



$\mu = 6$
$\sigma = ?$

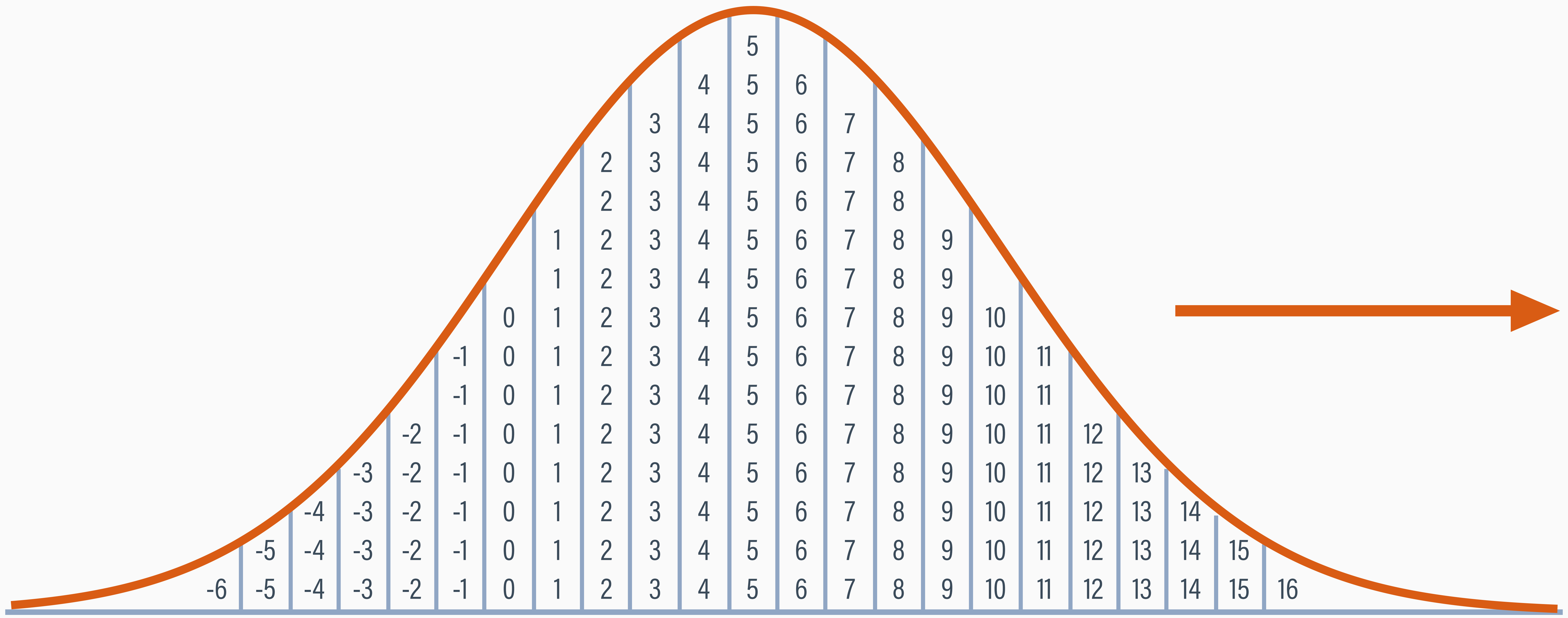
STEP 1: DEFINE POPULATION STANDARD DEVIATION

- The population standard deviation is important for assessing sampling error, so use the estimate from the data as a good guess



STEP 2: CREATE ARTIFICIAL DATA SETS

- A random number generator can produce samples of artificial data from a normally distributed population



R CODE TO SIMULATE A SAMPLE

= User-specified simulation feature

R function that creates artificial scores from a normal distribution

Population mean (μ)

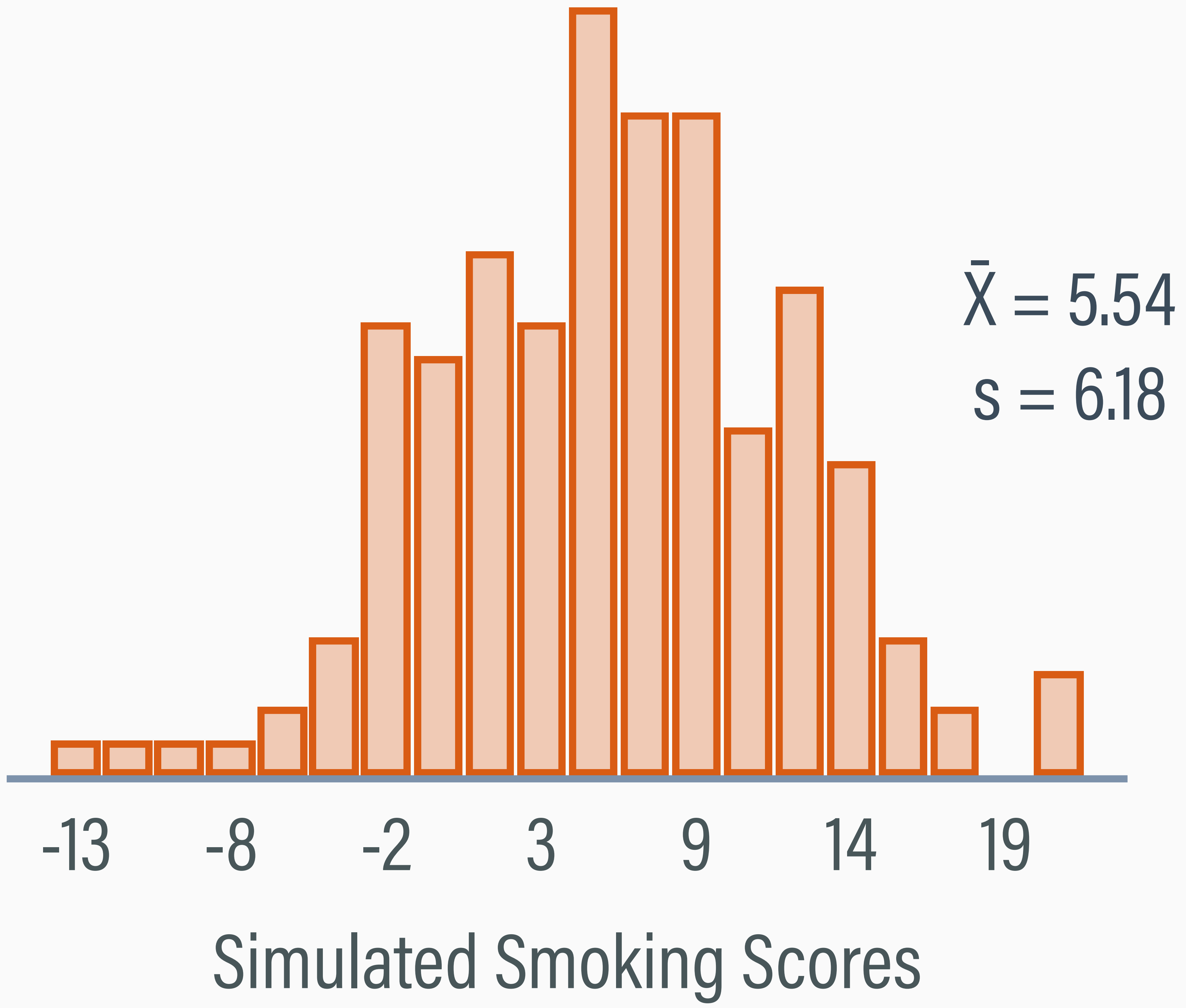
```
mysample <- rnorm( n = 165, mean = 6, sd = 6 )
```

Sample size

Population standard deviation (σ)

SIMULATED DATA SUMMARY

- A computer-generated sample of N = 165 scores produced a mean of $\bar{X} = 5.54$
- This particular sample's mean had –0.46 points of sampling error
- We need many more samples to compute the average or expected error



In small groups of two or three, each of you run the R code below, which draws a random sample of N = 165 scores from a population with a true mean of 6 and standard deviation of 6. For each of your individual samples, compute sampling error. Discuss how the direction (positive or negative) and amount of sampling error varied across your individual estimates.

```
mysample <- rnorm(n = 165, mean = 6, sd = 6)
mysample
hist(mysample)
cat("My sample mean = ", mean(mysample))
```

IMPORTANT OBSERVATIONS SO FAR

- Each hypothetical sample produces a different estimate and different amount of sampling error
- Sampling errors can be positive and negative, small or large
- To accurately estimate the average (expected) amount of sampling error, we need to summarize sampling errors from a large number of random samples (e.g., thousands)

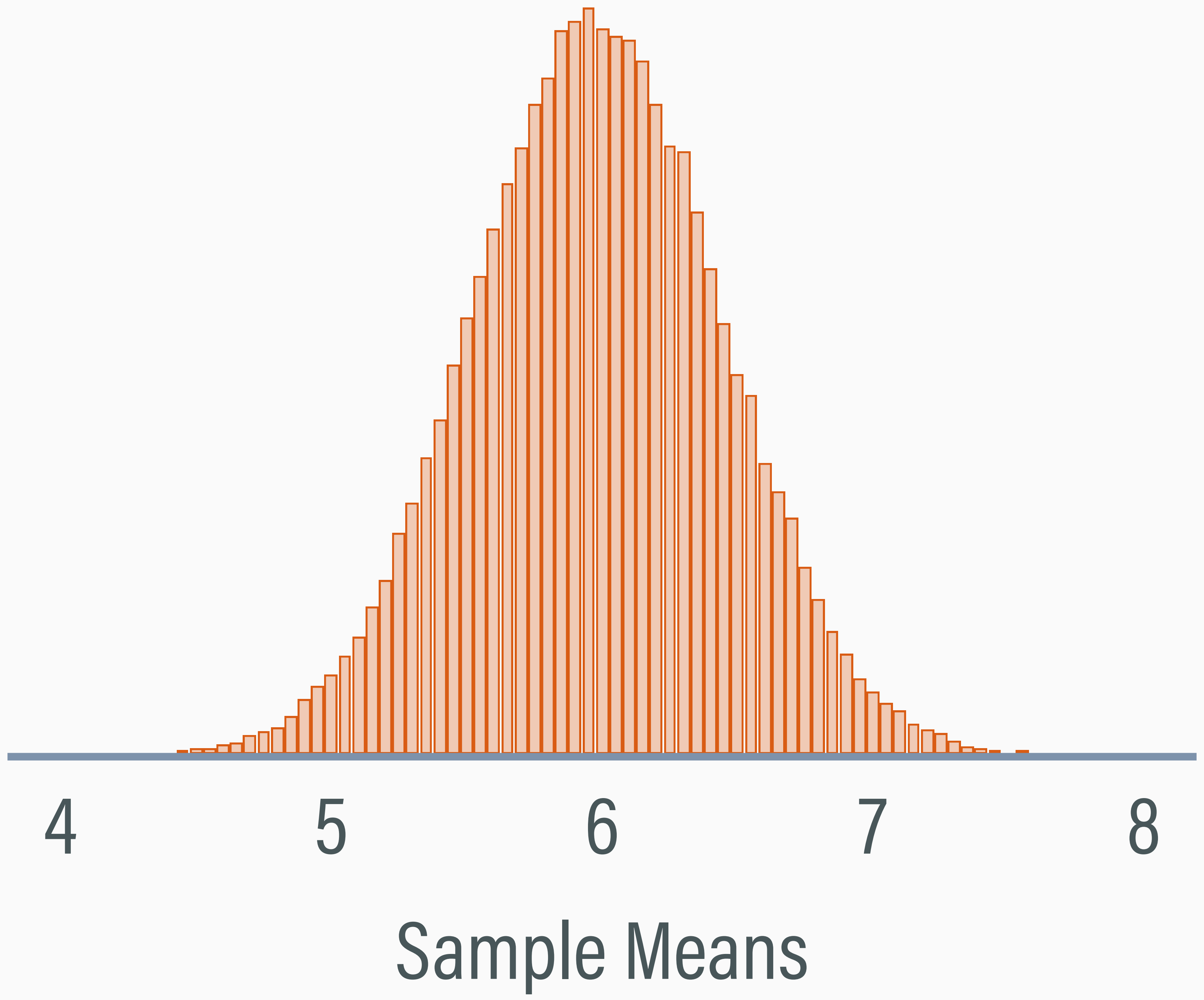
SIMULATION WITH 100,000 SAMPLES

- The R script for this week's lab generates 100,000 samples of data, computes the sample mean from each data set, then summarizes the 100,000 sample means

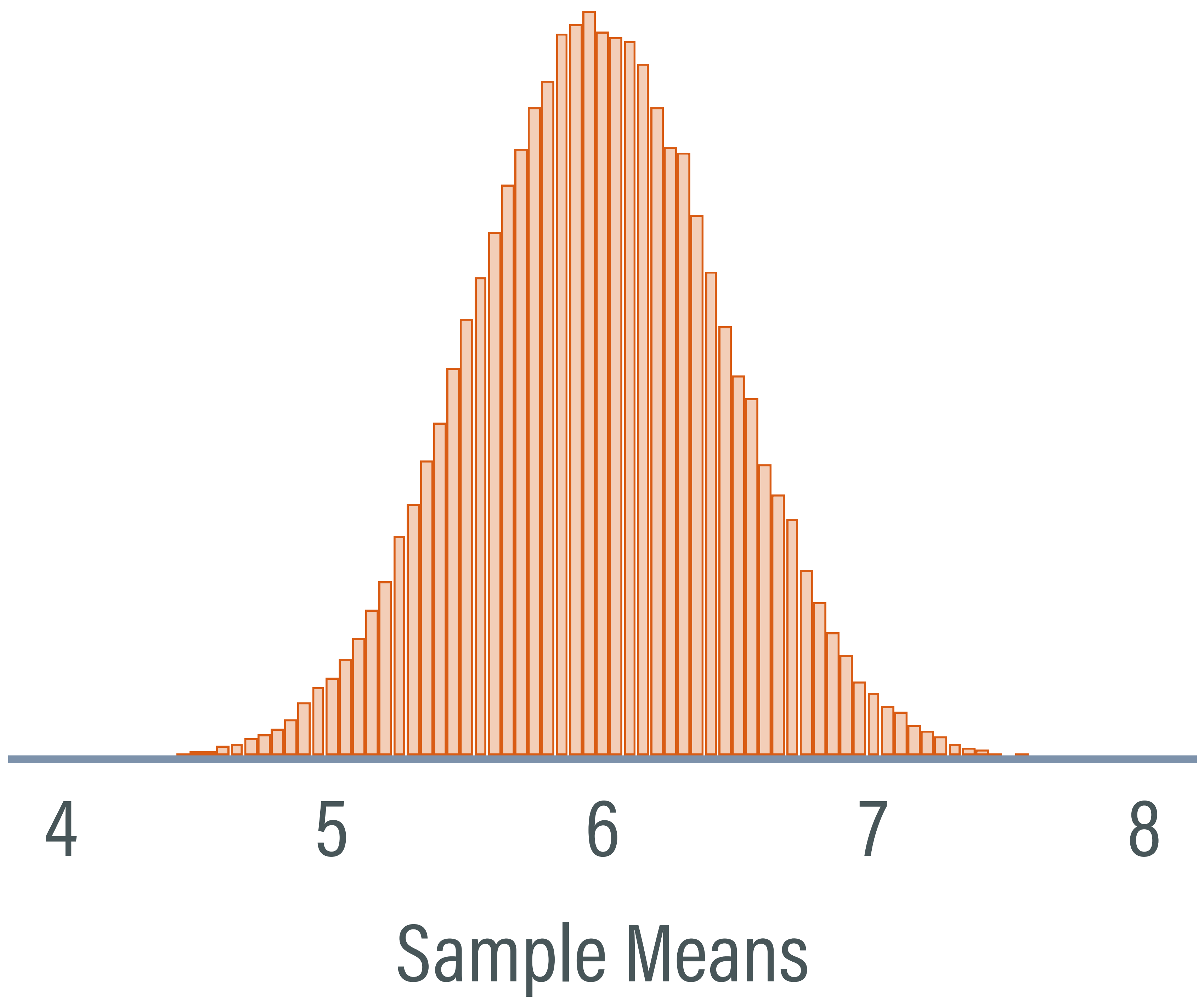
```
# SIMULATION WITH 100,000 RANDOM SAMPLES OF DATA ----

# initialize data set to hold 100,000 sample means
samplemeans <- rep(0, 100000)
# loop to draw 100,000 random samples
for(s in 1:100000){
  if(normal_data){ # normal data
    sampledata <- rnorm(n = sample_N, mean = pop_mean, sd = pop_sd)
  } else { # skewed data
    sampledata <- pop_mean + ((rchisq(sample_N, pop_mean) - pop_mean)/sqrt(2*pop_mean)) * pop_sd
  }
  # store the sample mean in element of s
  samplemeans[s] <- mean(sampledata)
}
...
```

DISTRIBUTION OF SAMPLE MEANS

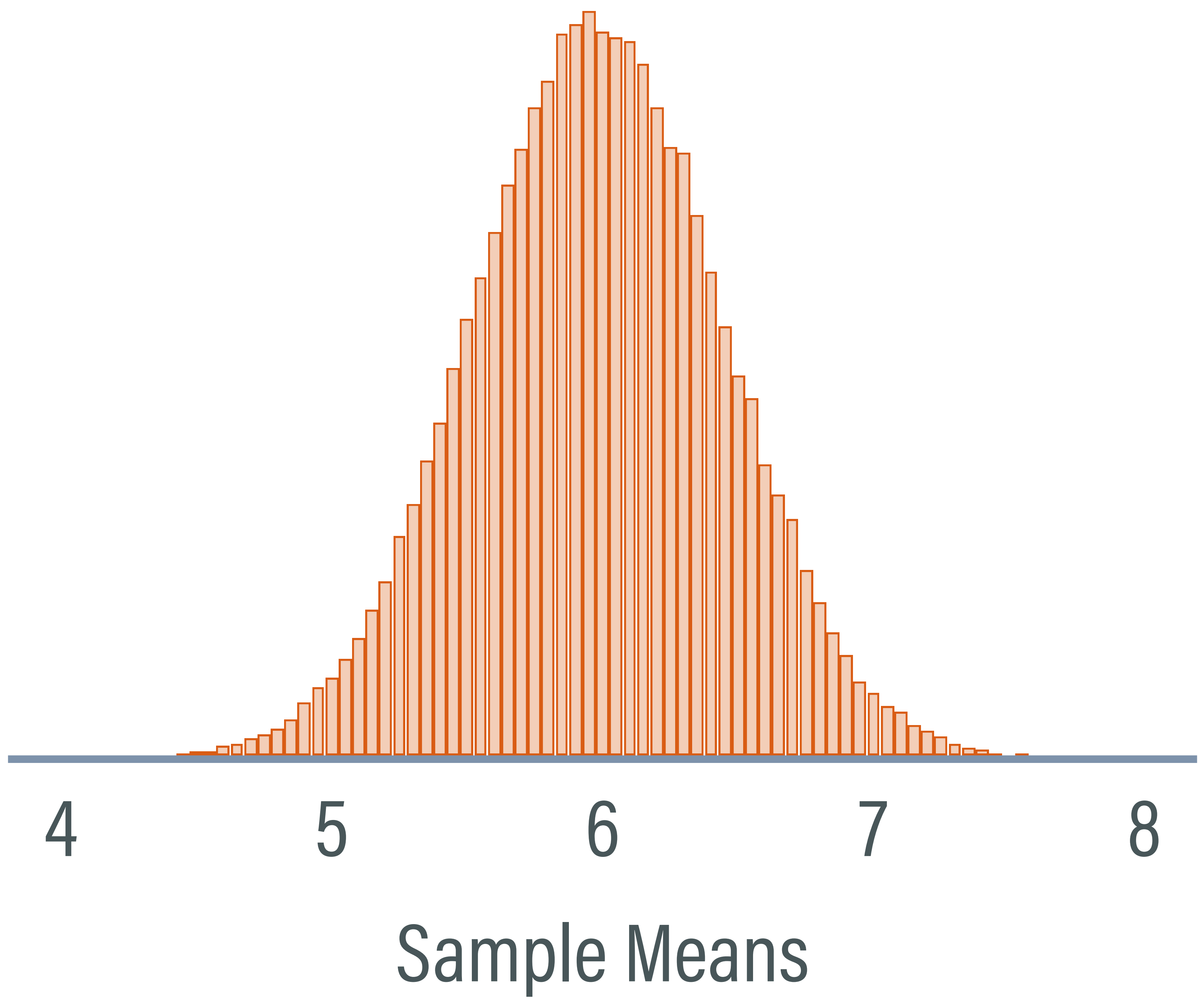
- The histogram displays 100,000 sample means from the simulation experiment
- The distribution of the estimates from numerous hypothetical samples is called a sampling distribution



In small groups of two or three, discuss the sampling distribution's main features. Focus on the shape, center, and spread of the distribution.

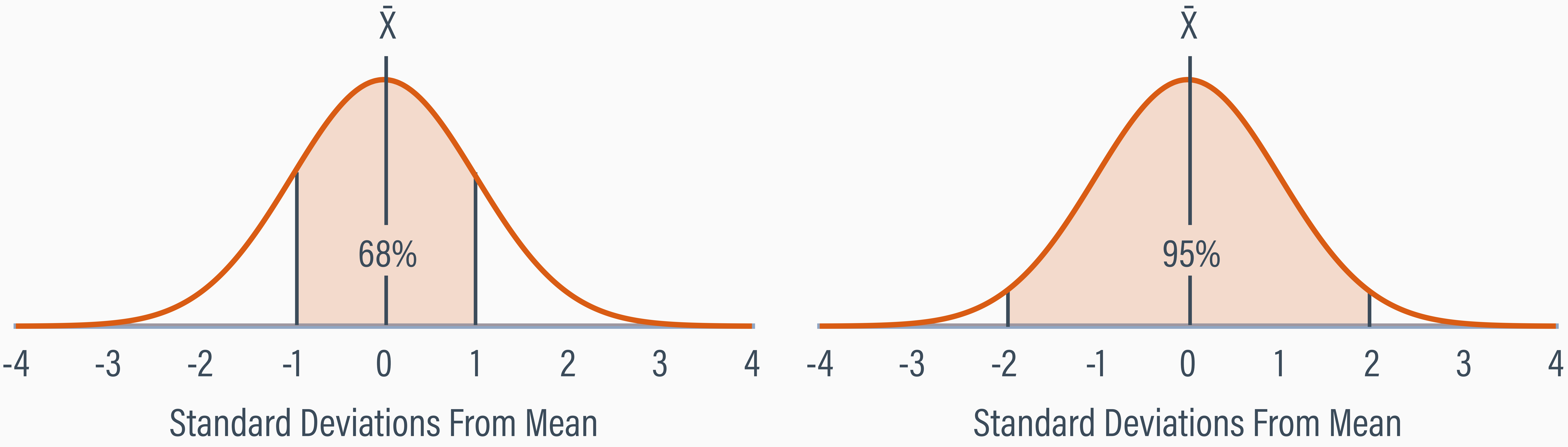


The sample mean from the clinical trial data was $\bar{X} = 5.5$. In small groups of two or three, discuss how likely it is that this sample came from a population with a true mean of 6. Which features of the sampling distribution support your conclusion?

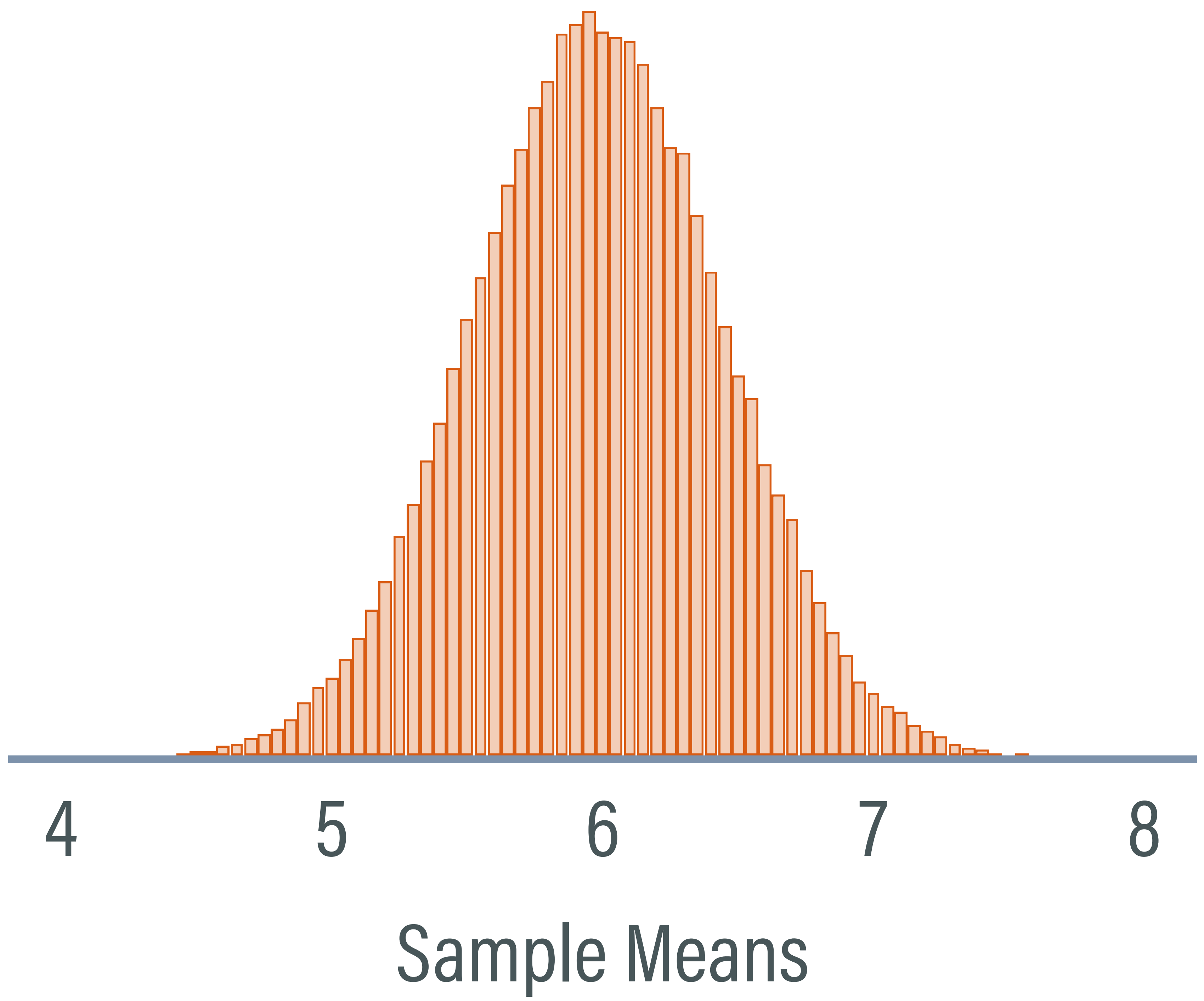


RULE OF THUMB FOR NORMAL DATA

- In a normal curve, 68% of observations are within ± 1 standard deviation of the mean, and 95% are within ≈ ± 2 standard deviations



In small groups of two or three, use the sampling distribution to estimate the standard deviation of the sample means (i.e., the average distance between the sample mean and the true mean of 6). Hint: Start by identifying the range that contains about 95% of all samples.



STANDARD DEVIATION OF SAMPLE MEANS

- The standard deviation is the average distance between a set of scores and the mean
- Sampling error is also a distance (sample mean vs. true population mean)
- We can modify the standard deviation formula to get the average (expected) amount of sampling error from the 100,000 means

Standard deviation of scores

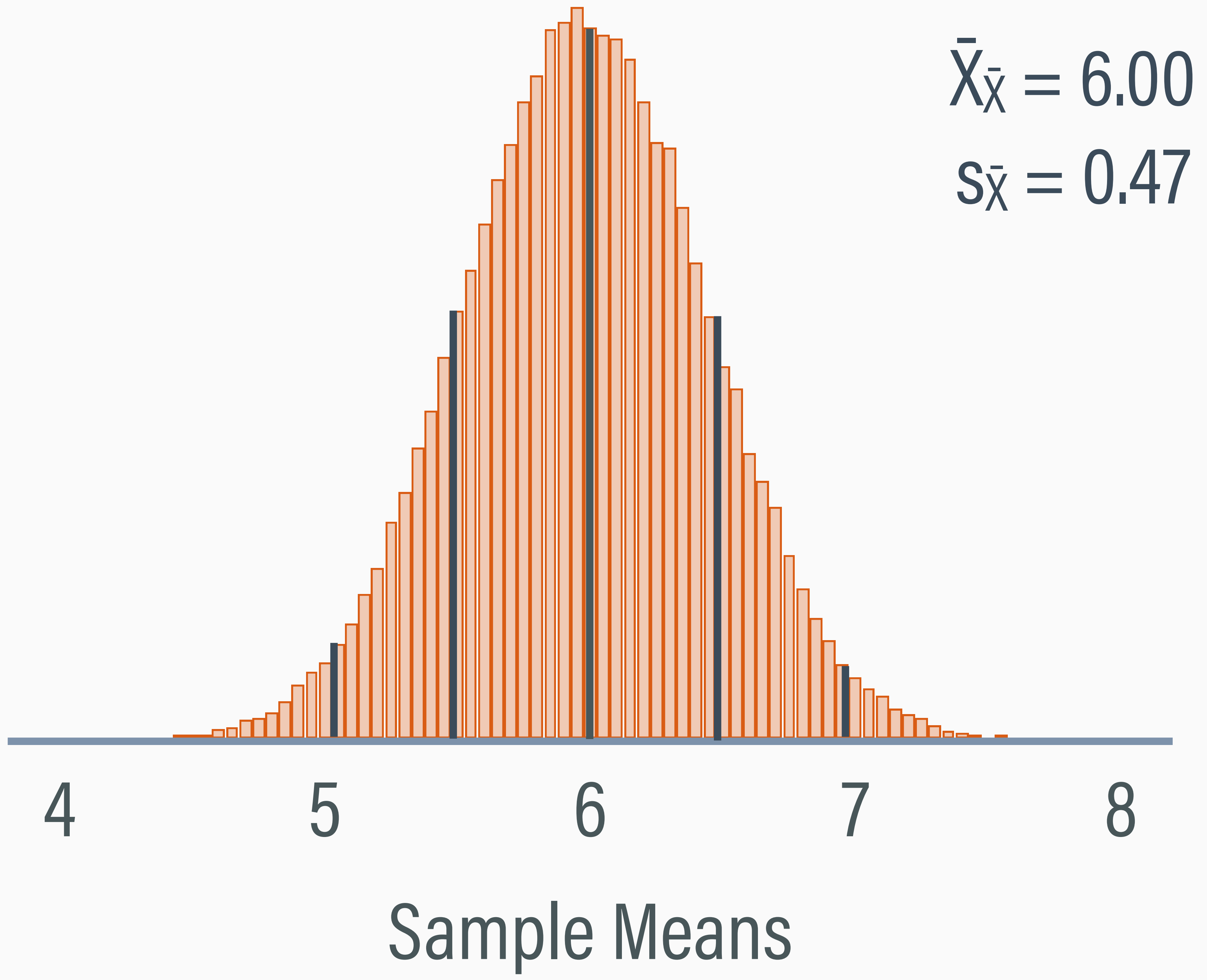
$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{N - 1}} = \sqrt{\frac{\text{sum of squared distances}}{\text{number of scores} - 1}}$$

Standard deviation of sample means

$$s_{\bar{X}} = \sqrt{\frac{\sum(\bar{X} - \mu)^2}{S - 1}} = \sqrt{\frac{\text{sum of squared distances}}{\text{number of samples} - 1}}$$

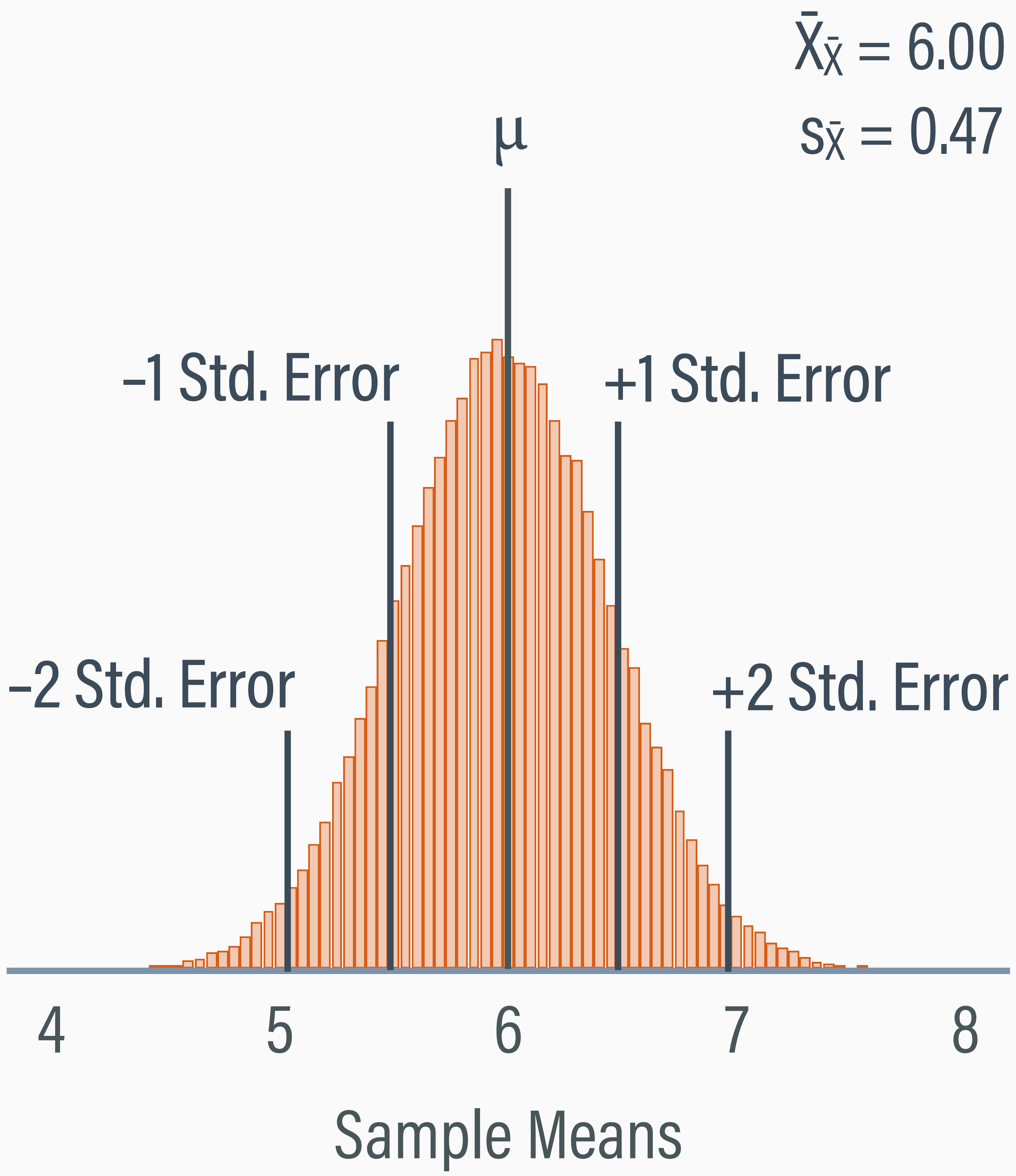
SAMPLING DISTRIBUTION SUMMARY

- The mean of the 100,000 sample means was $\bar{X}_{\bar{X}} = 6.00$ (equal to the true mean μ)
- The sample mean is unbiased: on average, it doesn't systematically miss in one direction
- The standard deviation of the 100,000 sample means (average distance from the estimates to μ) was $s_{\bar{X}} = 0.47$



STANDARD ERROR

- The standard deviation of the means from many random samples (symbolized $s_{\bar{X}}$) is called the standard error
- Standard error = standard deviation of the estimates = average sampling error
- On average, a sample of N = 165 should produce a mean that is $s_{\bar{X}} = 0.47$ points from the true population average

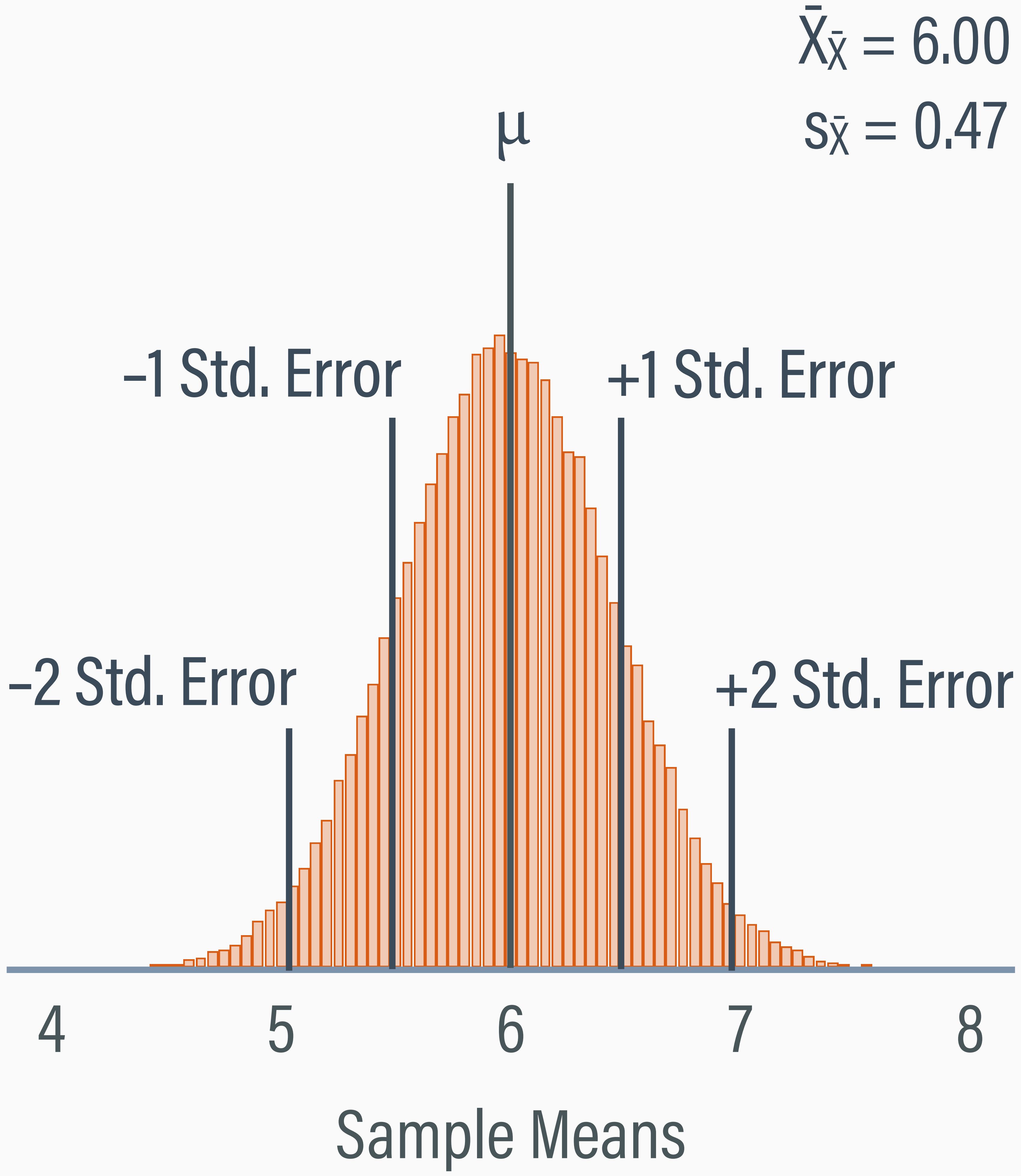


OUTLINE

1. Frequentist vs. Bayesian statistical paradigms
2. Sampling error
3. Estimating sampling error with computer simulation
4. Estimating sampling error with statistical theory
5. Study questions

COMPUTER SIMULATION SUMMARY

- Estimates from many samples were normally distributed around the true mean
- Standard deviation of estimates across many simulated data sets quantified the average sampling error (standard error)
- The expected error in the sample mean was about ± 0.47 breath CO point



USING THEORY INSTEAD OF SIMULATION

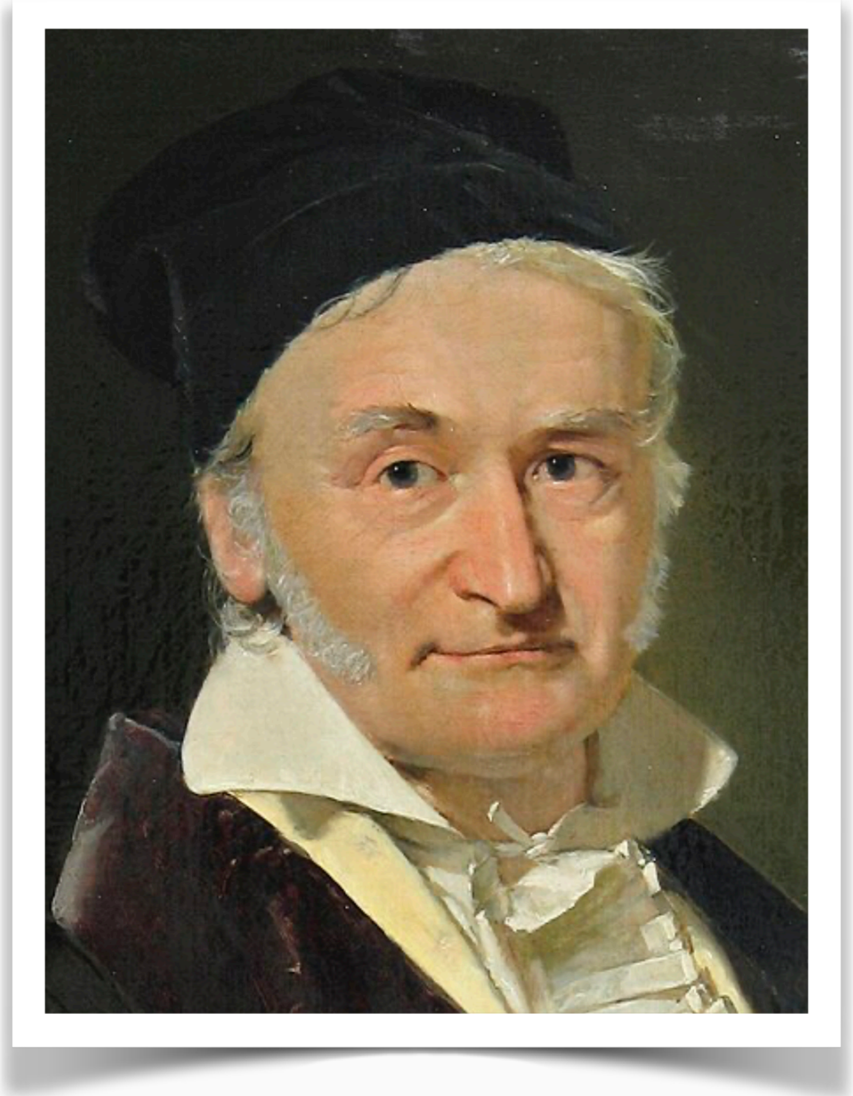
- Statistical theory explains everything we observed from the simulation, and it provides simple equations that estimate the average sampling error from a single sample of data
- Software packages like R compute standard errors (average or expected sampling error) using these theoretical formulas
- Equations and simulations usually give the same answer

STATISTICAL HISTORY: NORMAL CURVE

1809
Johann Carl Friedrich Gauss
Normal distribution

1811
Pierre-Simon Laplace
Central limit theorem

1908
William Sealy Gosset
t distribution



STATISTICAL HISTORY: CENTRAL LIMIT THEOREM

1809

Johann Carl Friedrich Gauss

Normal distribution

1811

Pierre-Simon Laplace

Central limit theorem

1908

William Sealy Gosset

t distribution

CENTRAL LIMIT THEOREM

- French mathematician Laplace was interested in how the mean fluctuated when undertaking repeated experiments

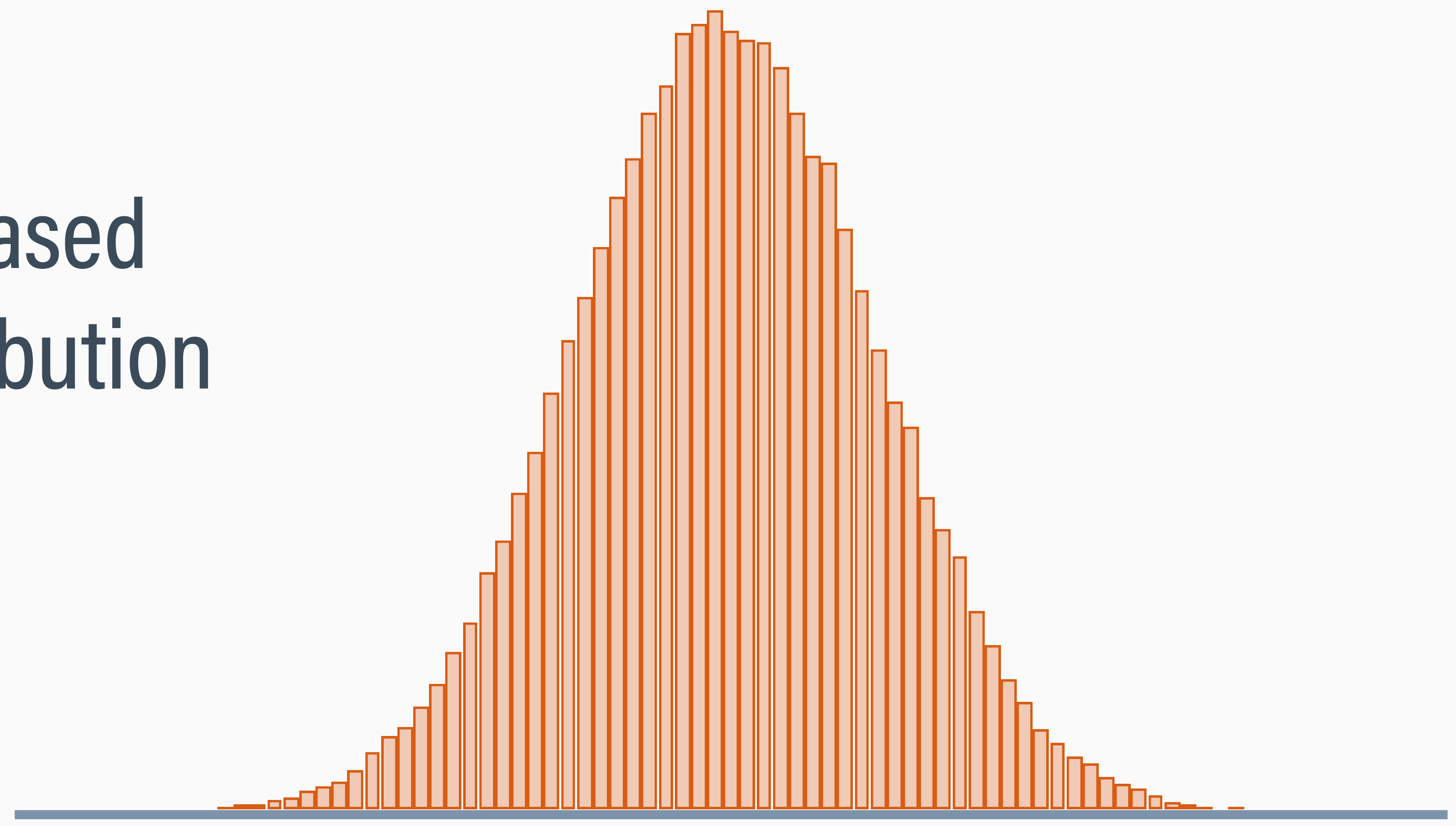
- Laplace published the first central limit theorem in 1811, which proved that means from samples of size N approximate a normal distribution if the sample size is large enough

- This is true even if the population data is not normal

CENTRAL LIMIT THEOREM PART 1

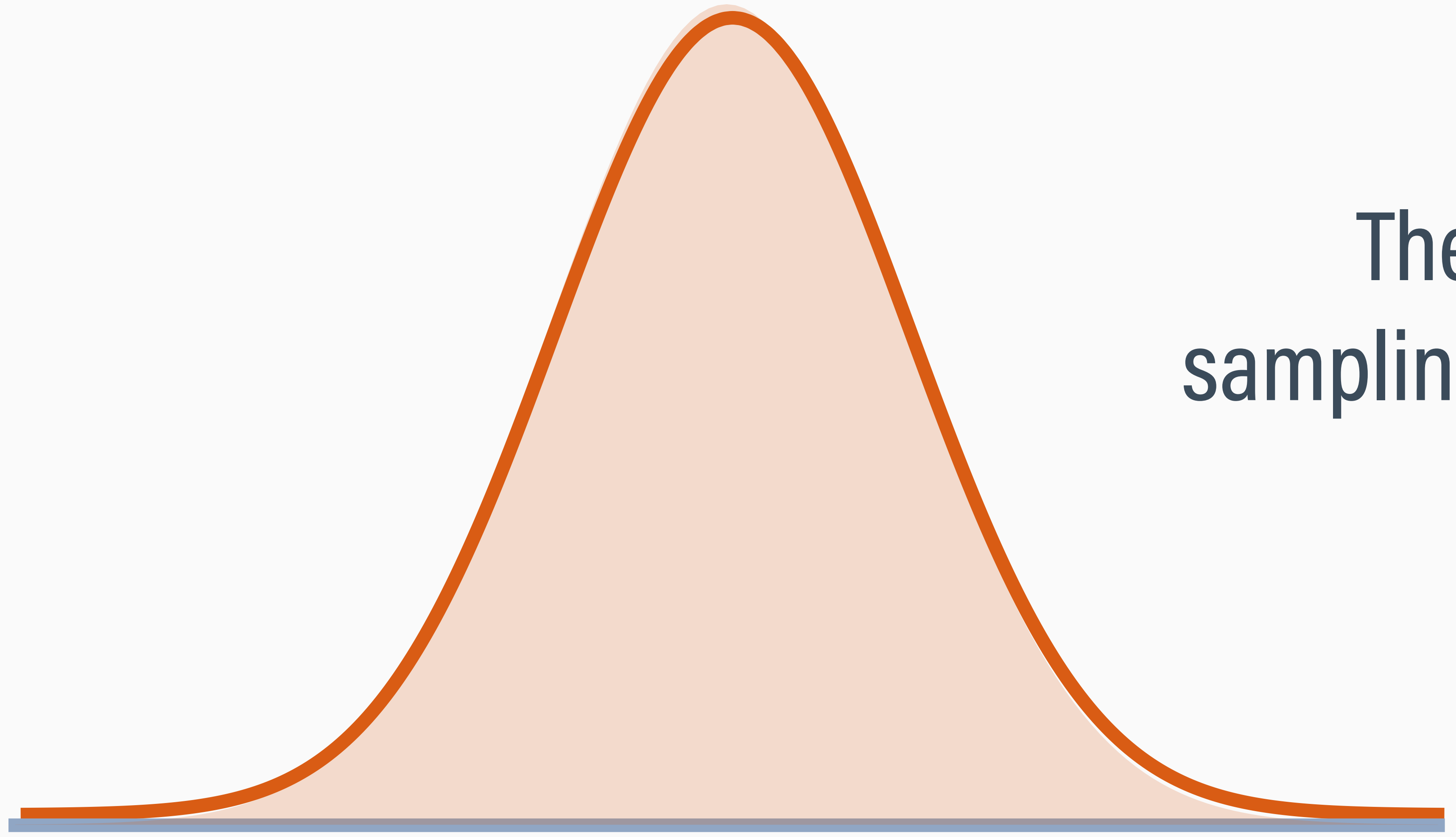
- With a large enough sample size, the means from many random samples follow a normal distribution

Simulation-based sampling distribution



Sample Means (Estimates)

Theoretical sampling distribution

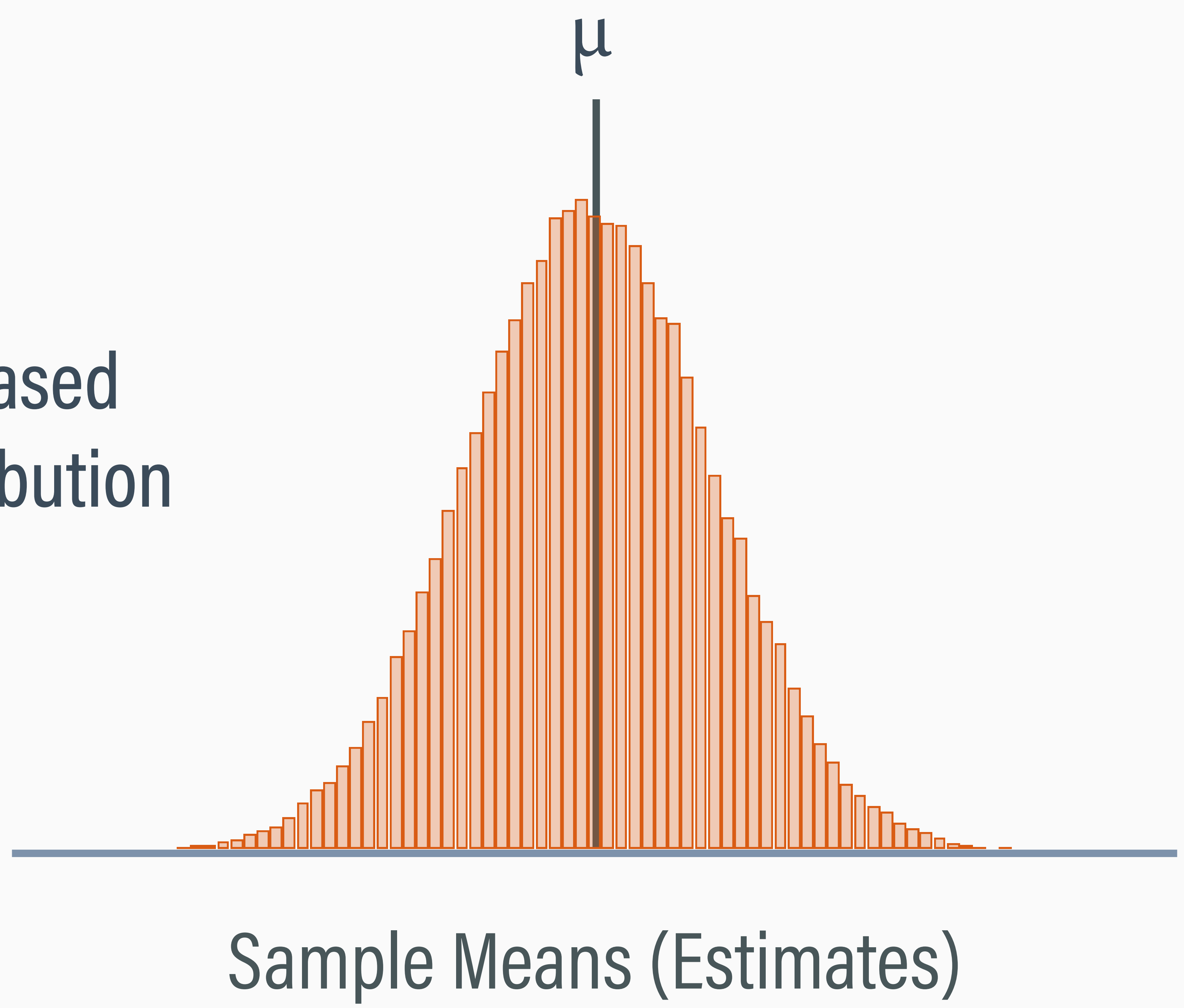


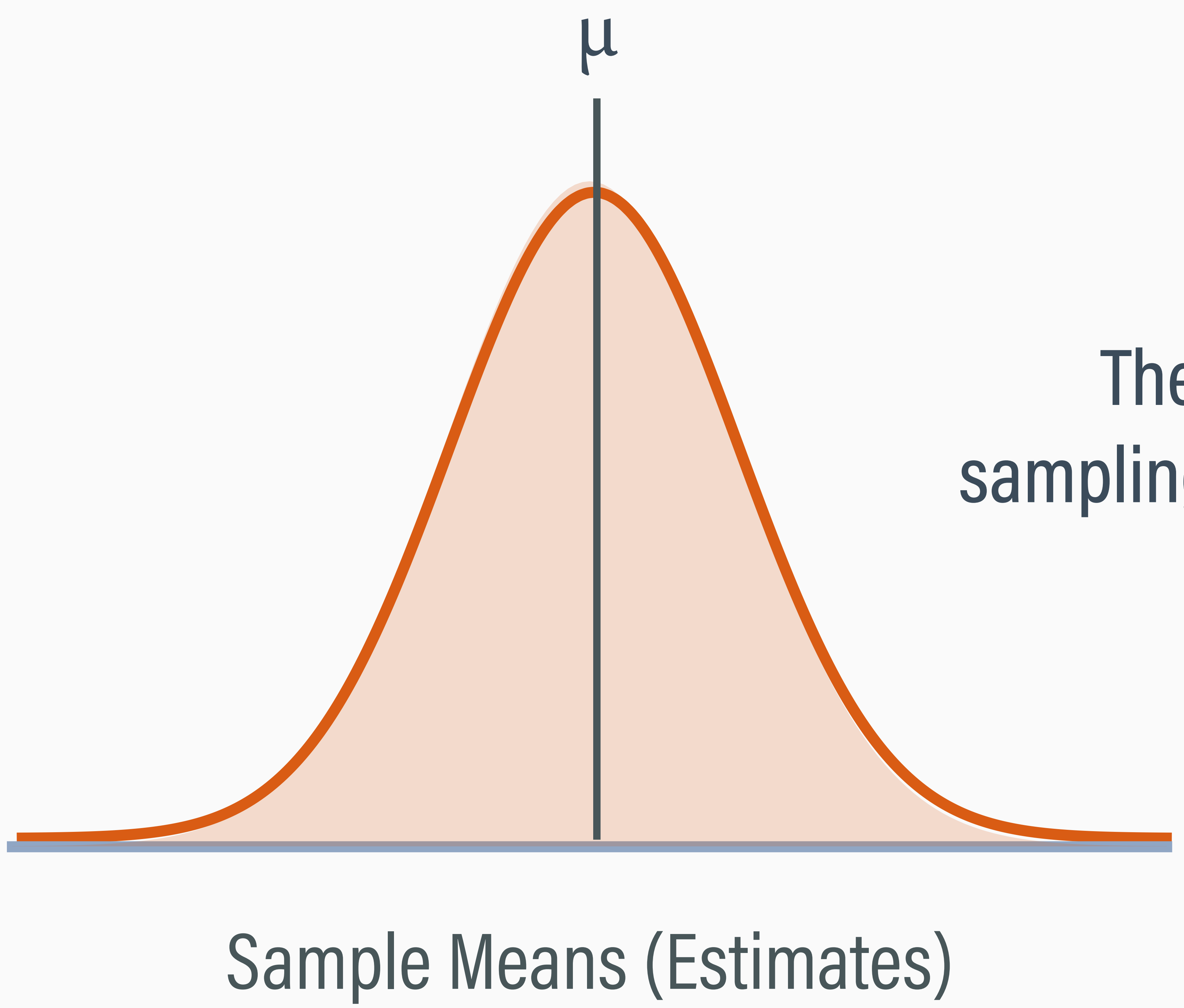
Sample Means (Estimates)

CENTRAL LIMIT THEOREM PART 2

- Estimates vary symmetrically around the true mean (across many samples, estimates do not systematically miss in one direction)

Simulation-based sampling distribution



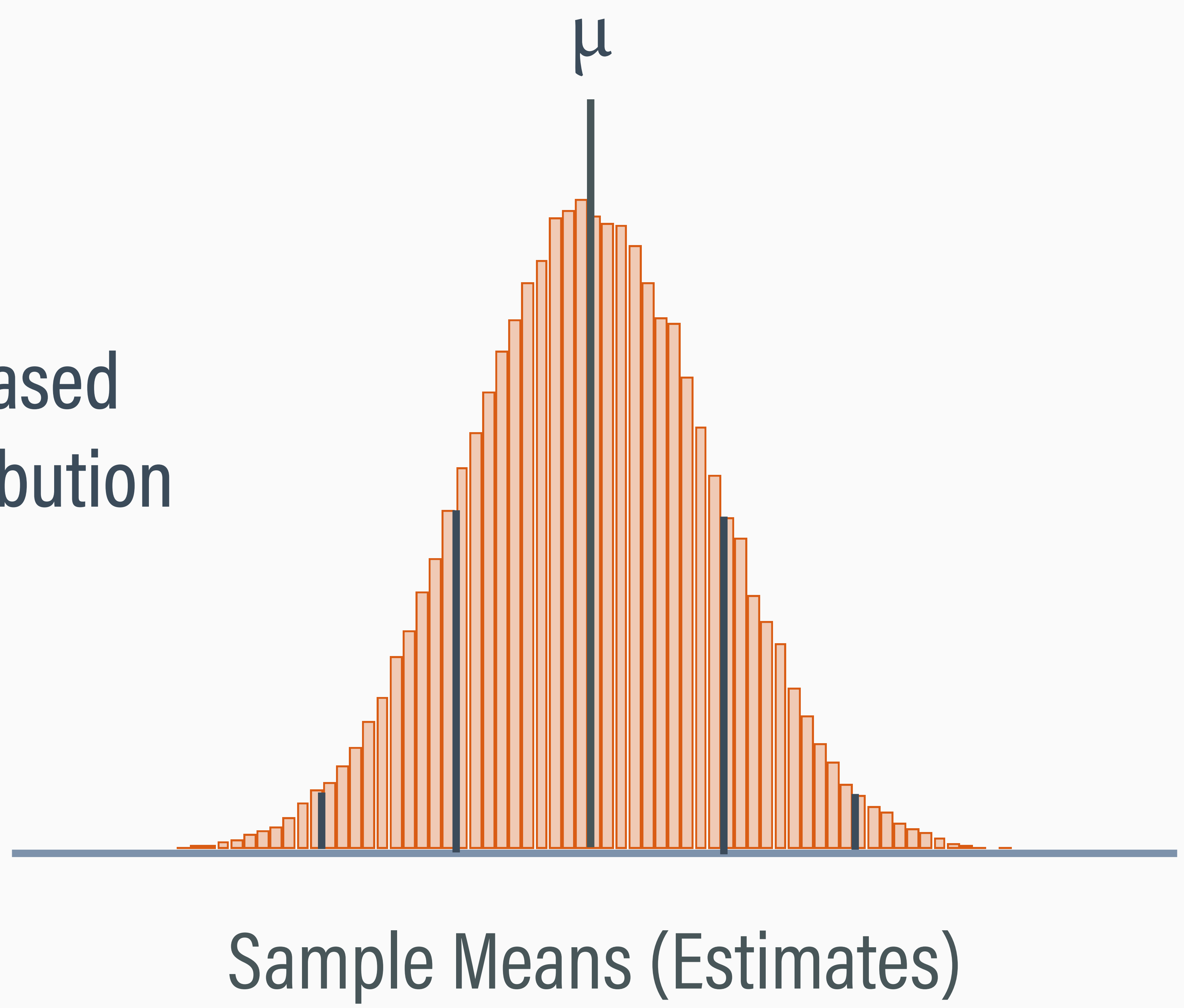


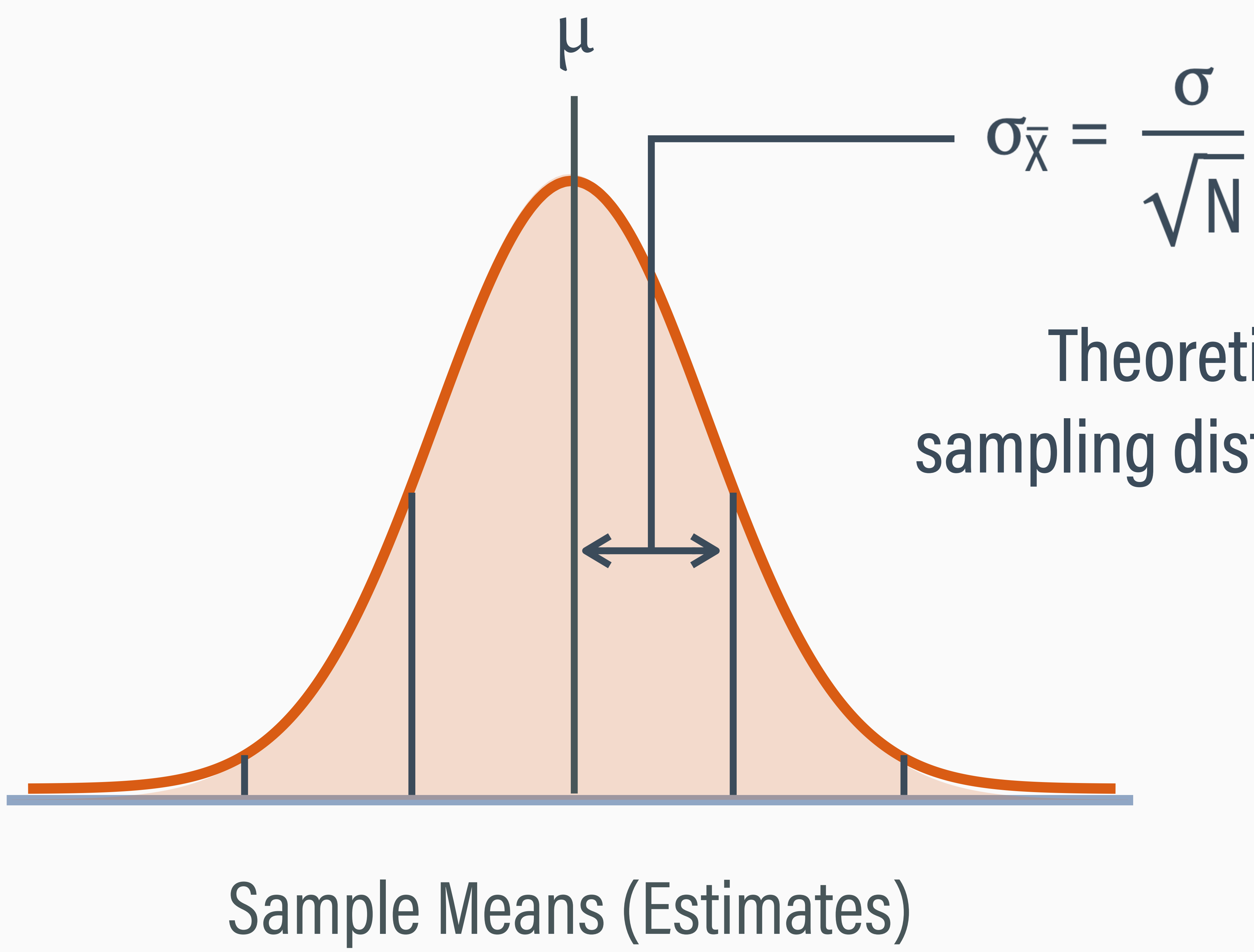
Theoretical sampling distribution

CENTRAL LIMIT THEOREM PART 3

- The standard error of the sample means can be estimated by dividing the *population* standard deviation of the scores by √N

Simulation-based sampling distribution





Theoretical sampling distribution

MEAN AND STANDARD ERROR

- The formula from the central limit theorem predicts that the average amount of sampling error is 0.47 breath CO points
- We expect this sample mean to be ±0.47 from the true mean
- The simulation-based standard error was identical, $s_{\bar{X}} = 0.47$

Mean ($\bar{X}$)	5.5
Std. Dev. (s)	6.0
N	165

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}} = \frac{6}{\sqrt{165}} = 0.47$$

R OUTPUT

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
Participant	1	165	83.00	47.78	83	83.00	60.79	1	165	164	0.00	-1.22	3.72
Condition*	2	165	1.50	0.50	2	1.50	0.00	1	2	1	-0.01	-2.01	0.04
Gender*	3	165	1.61	0.49	2	1.63	0.00	1	2	1	-0.43	-1.83	0.04
COWeek0	4	165	10.55	6.86	9	9.82	5.93	0	46	46	1.41	3.60	0.53
COWeek4	5	165	5.46	5.10	4	4.69	4.45	0	24	24	1.24	0.96	0.40
COWeek8	6	165	5.53	5.96	3	4.55	2.97	0	29	29	1.75	3.35	0.46
QuitCigsWeek4*	7	165	1.62	0.49	2	1.65	0.00	1	2	1	-0.48	-1.78	0.04
QuitCigsWeek8*	8	165	1.64	0.48	2	1.68	0.00	1	2	1	-0.59	-1.66	0.04
DrinksWeek0	9	165	6.40	4.42	6	5.77	2.97	1	35	34	2.44	10.46	0.34
DrinksWeek4	10	165	3.59	2.98	3	3.30	2.97	0	13	13	0.80	0.30	0.23
DrinksWeek8	11	165	3.23	2.68	3	2.97	2.97	0	13	13	0.89	0.67	0.21
CigsWeek0	12	165	14.22	8.22	12	13.05	5.93	3	51	48	1.59	3.20	0.64
CigsWeek4	13	165	4.18	5.70	2	3.08	2.97	0	41	41	3.01	13.13	0.44
CigsWeek8	14	165	3.16	4.75	2	2.25	2.97	0	35	35	3.44	16.81	0.37

STATISTICAL HISTORY: T DISTRIBUTION

1809

Johann Carl Friedrich Gauss

Normal distribution

1811

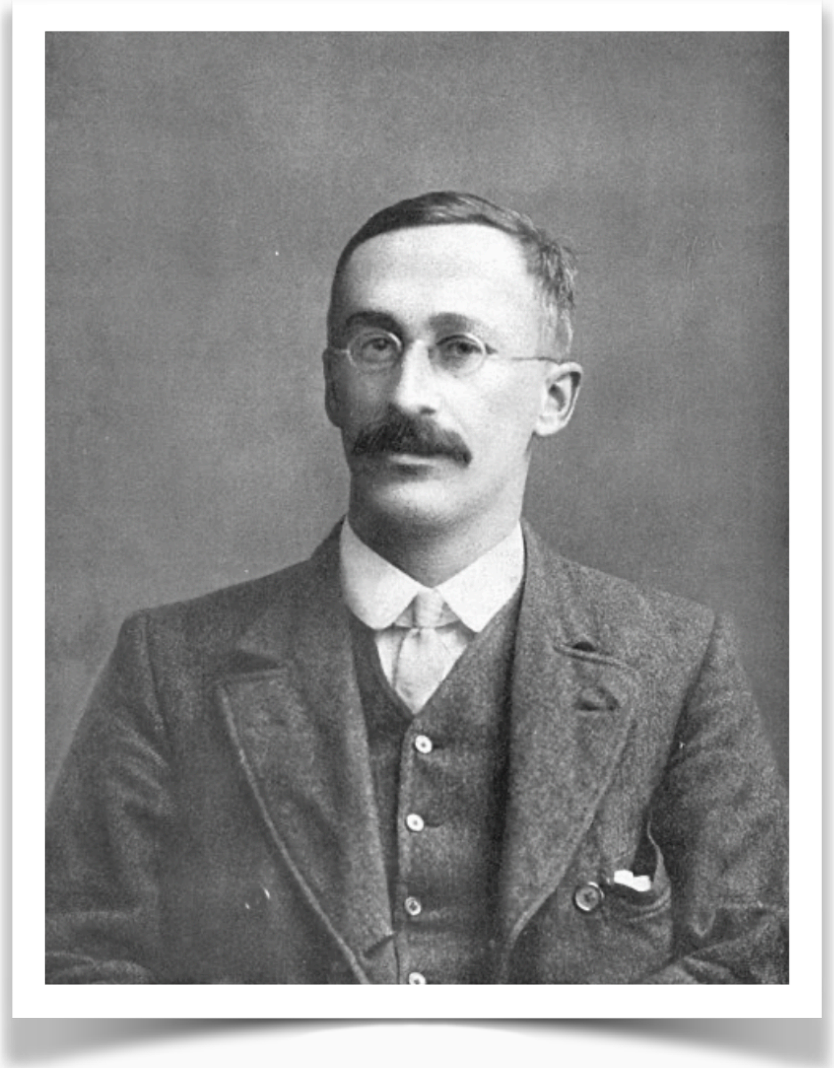
Pierre-Simon Laplace

Central limit theorem

1908

William Sealy Gosset

t distribution



WILLIAM SEALY GOSSETT

- Nearly 100 years later, a chemist at Guinness Brewery, William Gosset, noticed a problem with the central limit theorem
- Guinness was using experiments with small samples of agricultural products to draw conclusions that could be applied to a large-scale brewing process
- Gosset discovered that, when using small samples, the normal curve provided an inaccurate description of sampling error

STANDARD ERROR REVISITED

- The standard error formula requires the population standard deviation, which we never know
- We can substitute the sample standard deviation, but it too is an estimate with sampling error
- The quality of that substitution depends on the sample size (works fine when N is large enough)

The CLT requires the population standard deviation

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$$

$$s_{\bar{X}} = \frac{s}{\sqrt{N}}$$

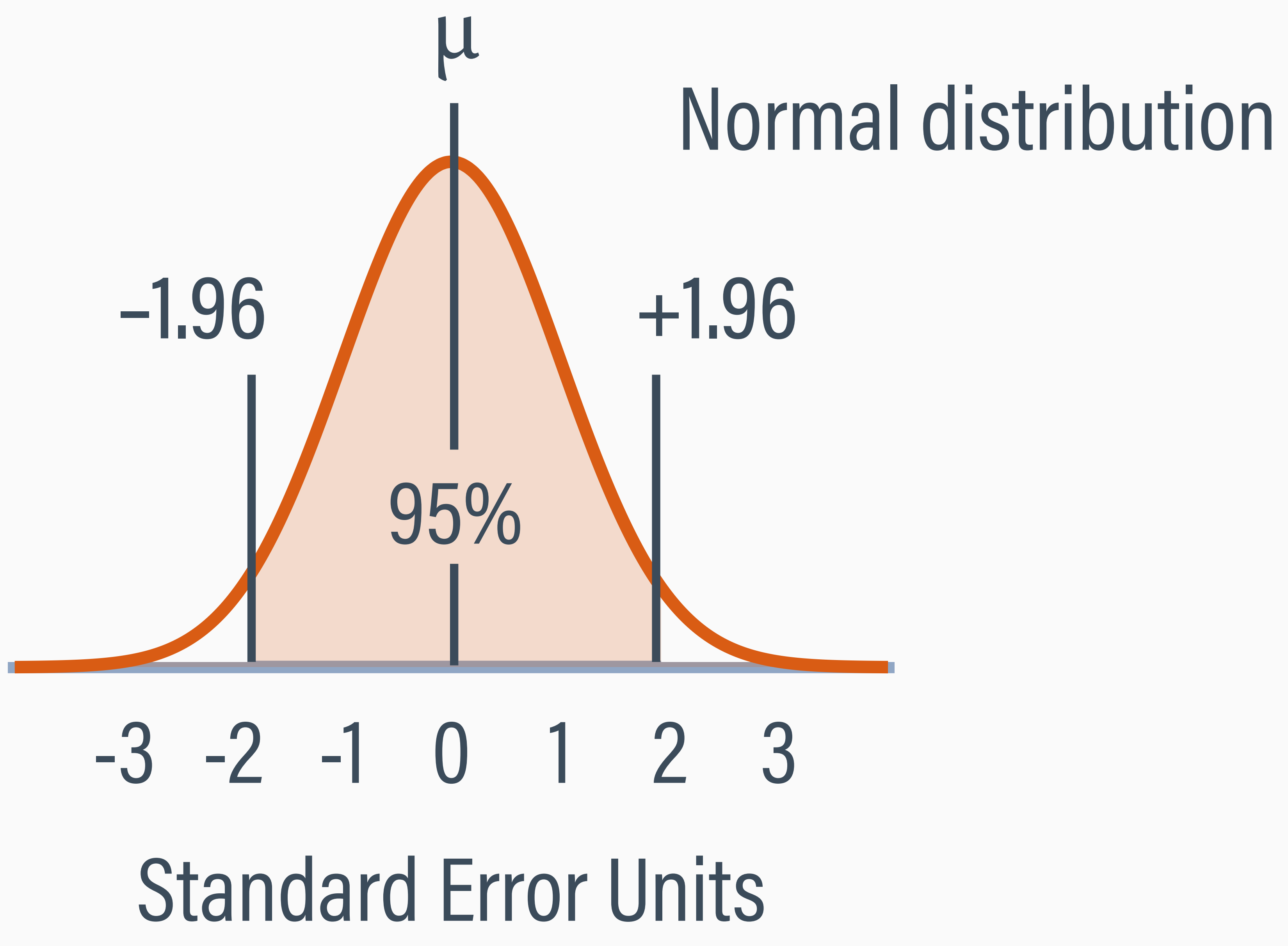
In practice, we substitute the sample standard deviation

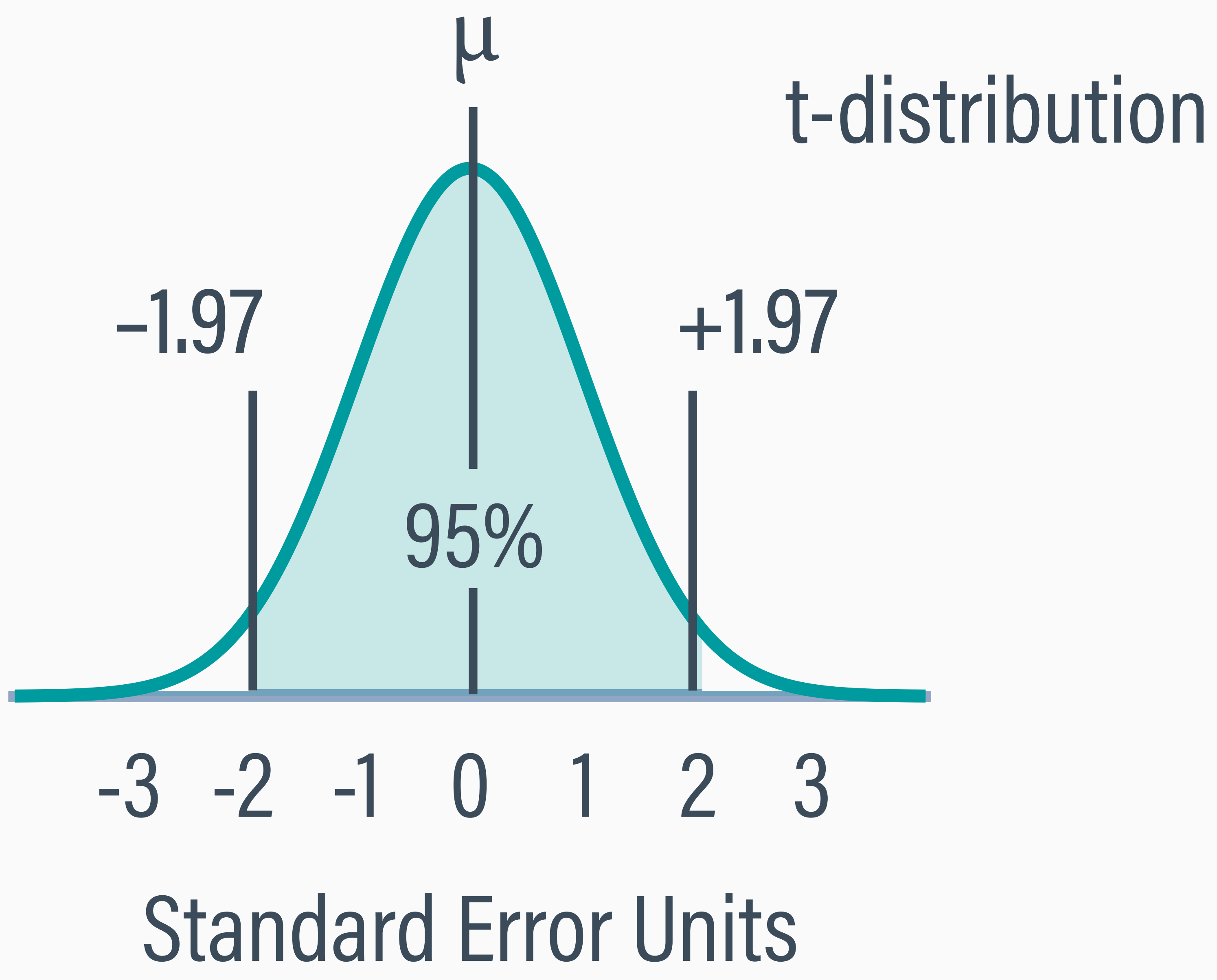
THE T-DISTRIBUTION

- Gossett derived a new distribution called the t-distribution that is a better model for sampling error when the N is small
- The t-distribution is a series of bell-shaped curves that stretch out (become more variable) as the N decreases
- Software programs typically use the t-distribution to derive statistical quantities that rely on the 95% rule of thumb

T-DISTRIBUTION VS. NORMAL CURVE

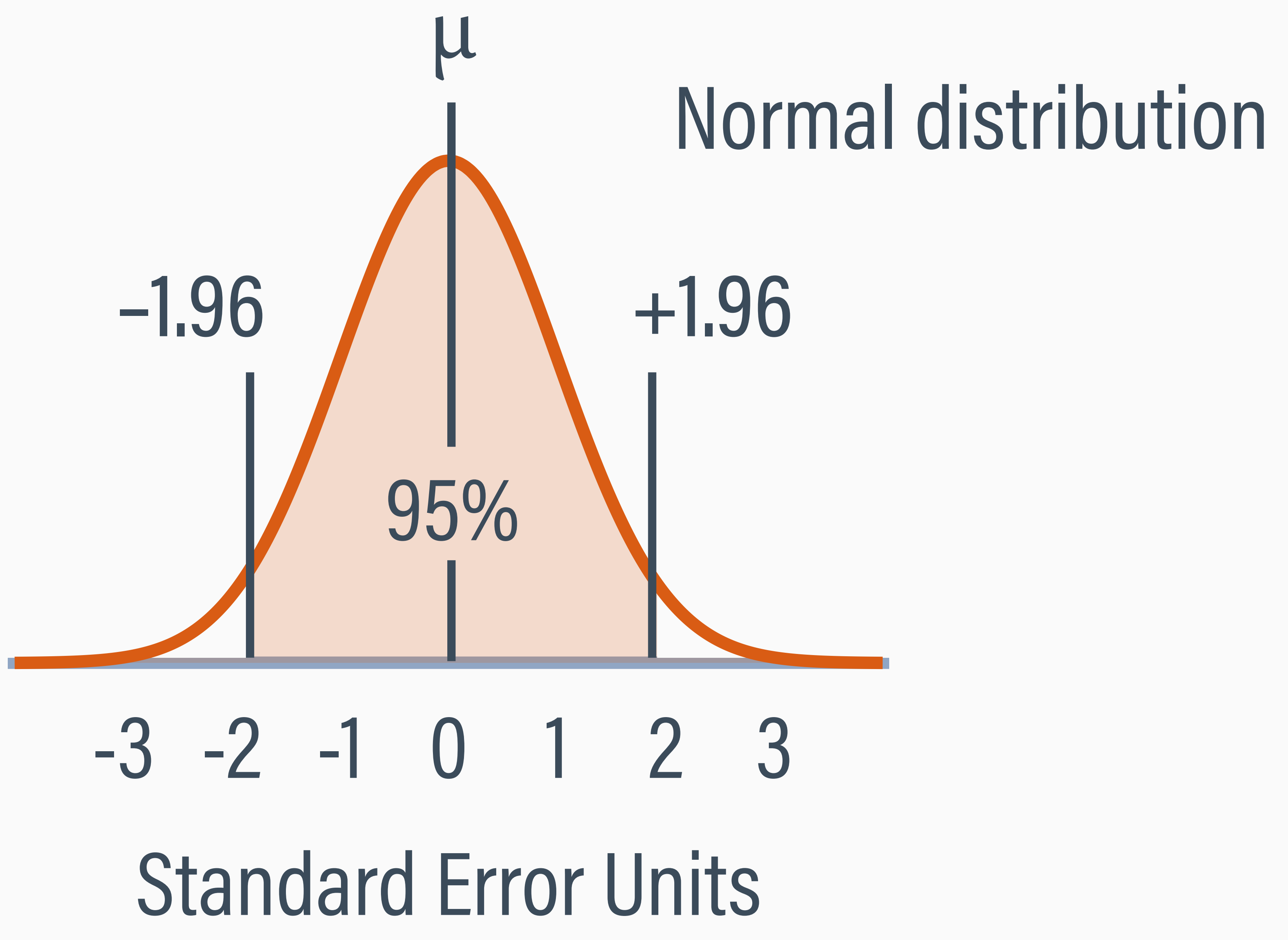
- According to the central limit theorem, 95% of the sample means are within ± 1.96 standard errors of the true mean
- When N = 165, the t-distribution predicts that 95% the sample means are within ± 1.97 standard errors of the true mean
- 95% cutoffs are virtually identical (this sample is apparently “large enough”)

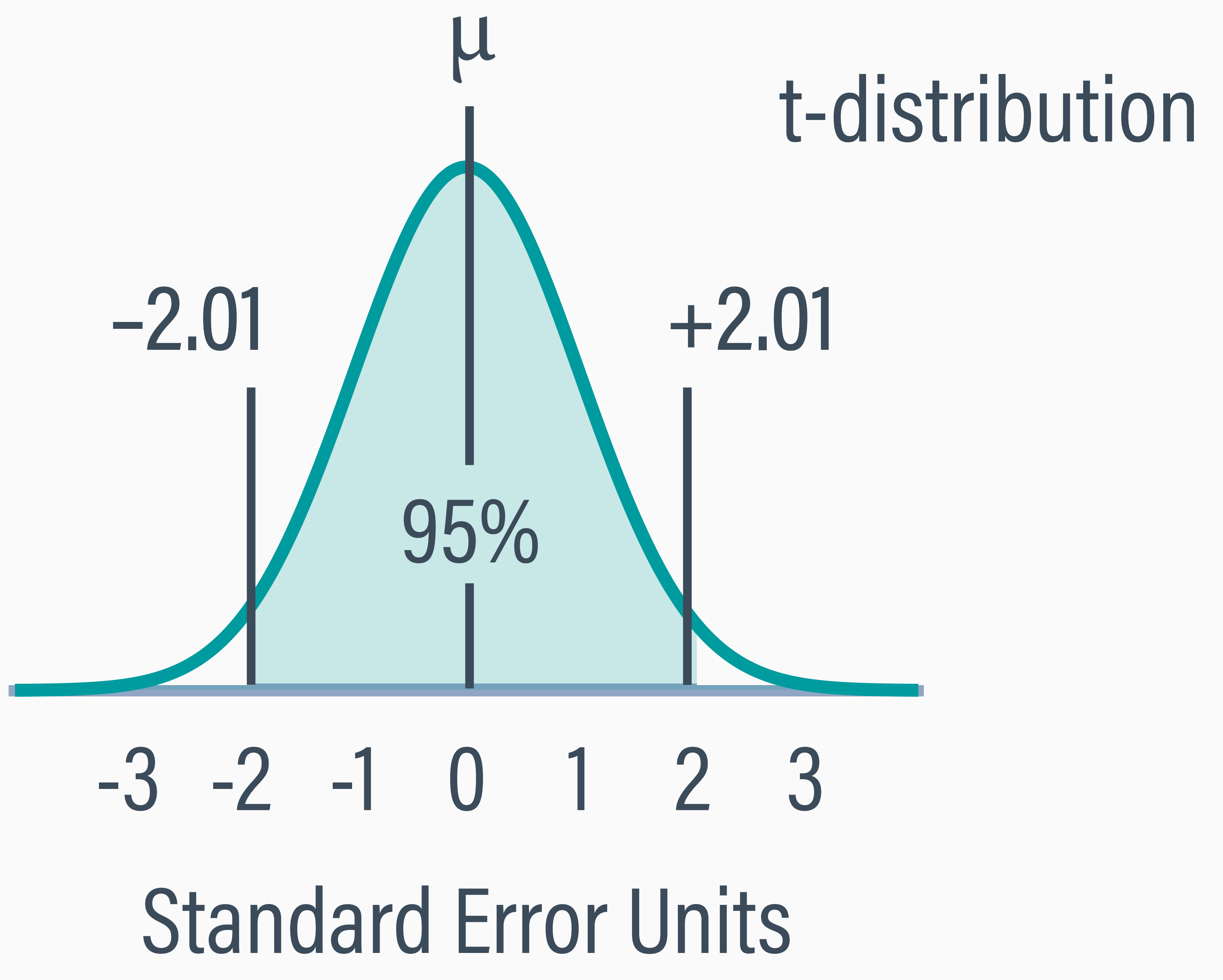




T-DISTRIBUTION VS. NORMAL CURVE

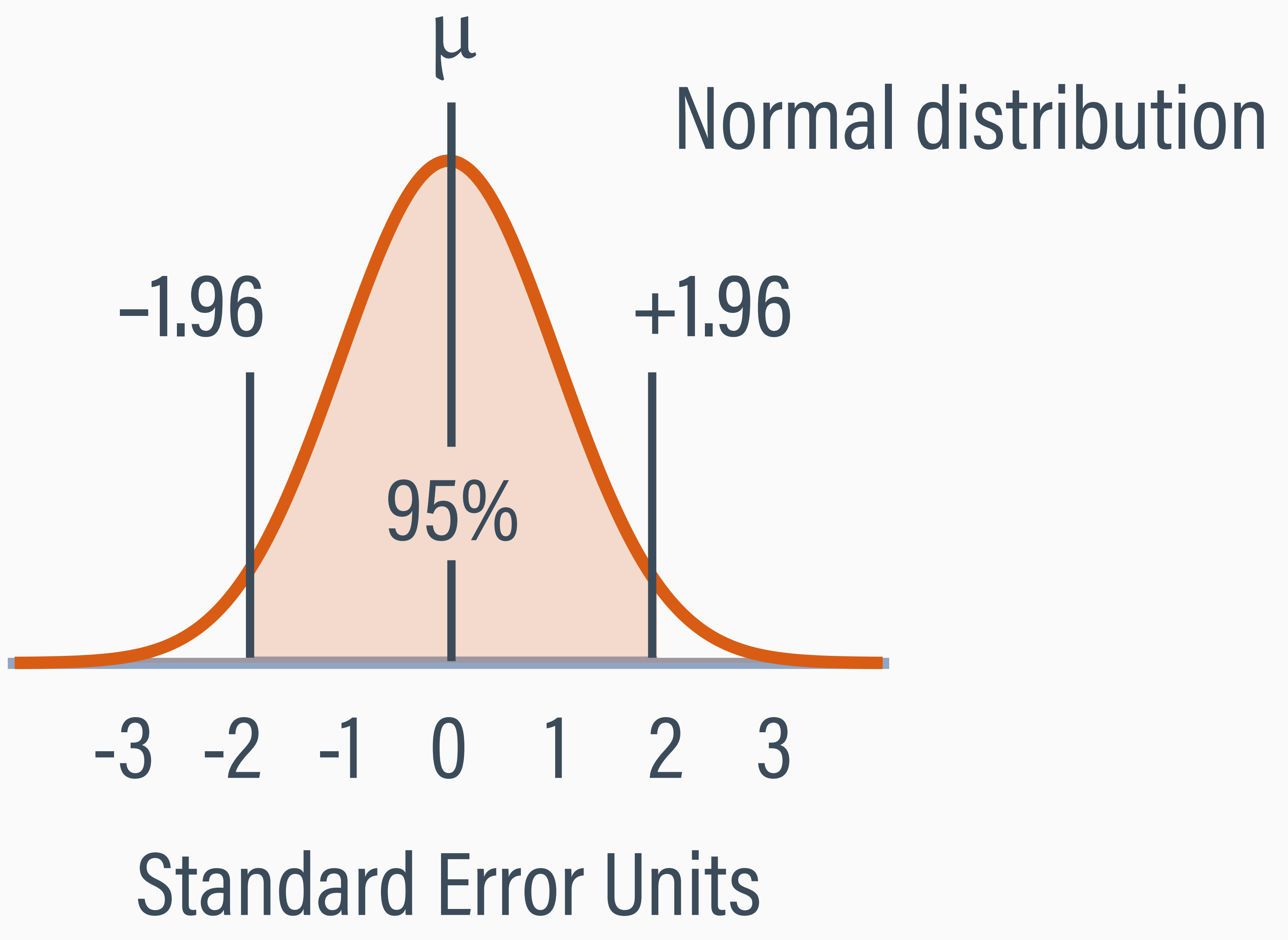
- According to the central limit theorem, 95% of the sample means are within ± 1.96 standard errors of the true mean
- When N = 50, the t-distribution predicts that 95% the sample means are within ± 2.01 standard errors of the true mean
- 95% cutoffs are similar, but the t-distribution is starting to stretch out (more outliers)

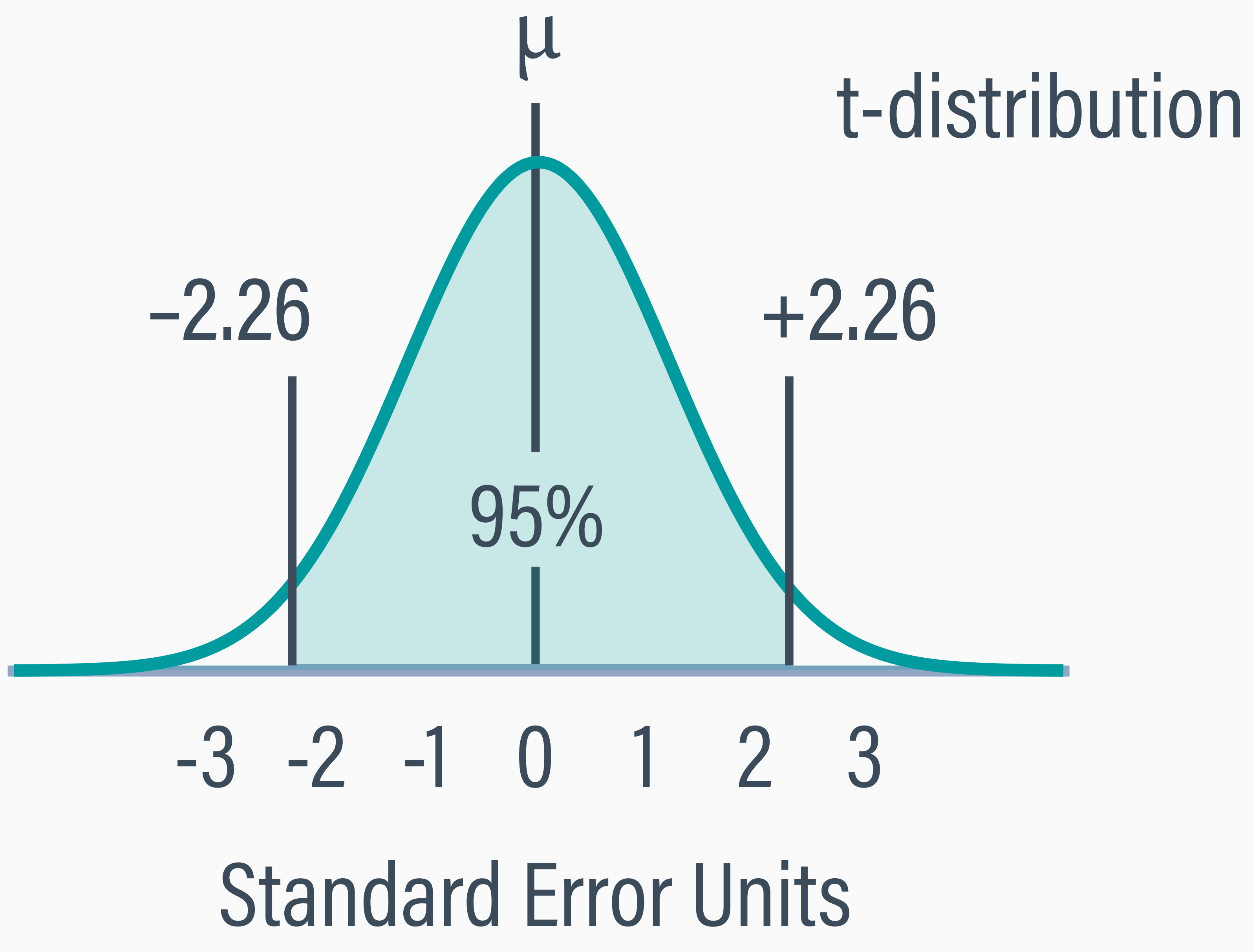




T-DISTRIBUTION VS. NORMAL CURVE

- According to the central limit theorem, 95% of the sample means are within ± 1.96 standard errors of the true mean
- When N = 10, the t-distribution predicts that 95% the sample means are within ± 2.26 standard errors of the true mean
- The t-distribution is much wider than the normal curve





OUTLINE

STUDY QUESTIONS (1)

1. Describe the key idea behind the frequentist statistical paradigm. How does that key idea differ from the Bayesian paradigm?

2. A Gallup poll of 1000 registered voters reports that 56% of respondents favor some type of government-backed universal health care. Using this as a context, describe the population, sample, parameter, and estimate.

STUDY QUESTIONS (2)

3. Still referring to the previous Gallup poll, describe the concept of random sampling error.

4. The poll reports that 56% of respondents favor some type of government-backed universal health care, and the standard error is 2%. Provide an interpretation of the standard error.

5. Hypothetically, what would the polling company have to do to obtain the information needed to construct a sampling distribution (without using simulation).

STUDY QUESTIONS (3)

A researcher conducts a study of postpartum depression in new mothers. She administers the Beck Depression Inventory to a sample of 100 participants. The sample mean and standard deviation are $\bar{X} = 20$ and $s = 6$. The standard error of the mean is 0.60.

6. Provide an interpretation of the standard error.

7. The upper cutoff for mild depression is 19. How likely is it that this sample with $\bar{X} = 20$ originated from a population with a true mean of $\mu = 19$? Explain your rationale.

STUDY QUESTIONS (4)

8. Still referring to the depression study, what would happen to the standard error if the sample size was increased from 100 to 200?

9. What would happen to the standard error if the variability among new mothers was much smaller (e.g., the standard deviation decreased from 6 to 3)?

STUDY QUESTIONS (5)

10. Describe the central limit theorem's three predictions about the means from many hypothetical samples data.

11. When and why does the central limit theorem fail to provide accurate predictions about sampling error? Describe how the t-distribution addresses the issue.