Unit 2: Statistical Inference

Learning objectives

- By the end of this unit, you should be able to:
- create an initial image of the field of statistics.
- introduce several basic vocabulary words used in studying statistics: *population*, *variable*, *statistic*.
- learn how to obtain sample data.

What is Statistics?

Statistics: The science of collecting, describing, and interpreting data.

Two areas of statistics:

Descriptive Statistics: collection, presentation, and description of sample data.

Inferential Statistics: making decisions and drawing conclusions about populations.

Example: A recent study examined the math and verbal SAT scores of high school seniors across the country. Which of the following statements are descriptive in nature and which are inferential.

- The mean math SAT score was 492.
- The mean verbal SAT score was 475.
- Students in the Northeast scored higher in math but lower in verbal.
- 80% of all students taking the exam were headed for college.
- 32% of the students scored above 610 on the verbal SAT.
- The math SAT scores are higher than they were 10 years ago.

Introduction to Basic Terminologies

Population: A collection, or set, of individuals or objects or events whose properties are to be analyzed.

Two kinds of populations: *finite* or *infinite*.

Sample: A subset of the population.

Variable: A characteristic about each individual element of a population or sample.

Data (singular): The value of the variable associated with one element of a population or sample. This value may be a number, a word, or a symbol.

Data (plural): The set of values collected for the variable from each of the elements belonging to the sample.

Experiment: A planned activity whose results yield a set of data.

Parameter: A numerical value summarizing all the data of an entire population.

Statistic: A numerical value summarizing the sample data.

Example: A college dean is interested in learning about the average age of faculty. Identify the basic terms in this situation.

The *population* is the age of all faculty members at the college.

A *sample* is any subset of that population. For example, we might select 10 faculty members and determine their age.

The *variable* is the “age” of each faculty member.

One *data* would be the age of a specific faculty member.

The *data* would be the set of values in the sample.

The *experiment* would be the method used to select the ages forming the sample and determining the actual age of each faculty member in the sample.

The *parameter* of interest is the "average" age of all faculty at the college.

The *statistic* is the "average" age for all faculty in the sample.

Two kinds of variables:

Qualitative, or Attribute, or Categorical, Variable: A variable that categorizes or describes an element of a population.

Note: Arithmetic operations, such as addition and averaging, are *not* meaningful for data resulting from a qualitative variable.

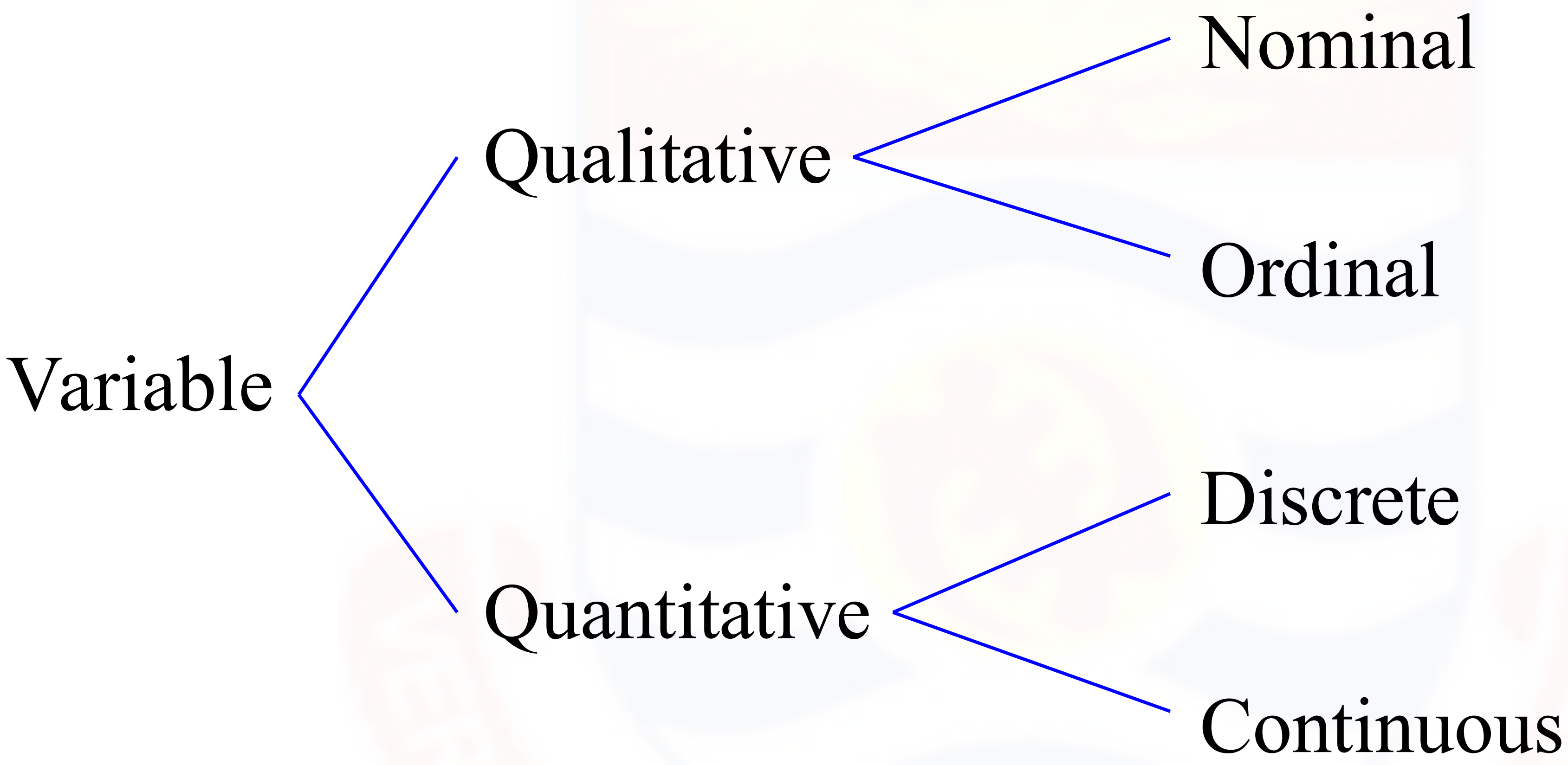
Quantitative, or Numerical, Variable: A variable that quantifies an element of a population.

Note: Arithmetic operations such as addition and averaging, are meaningful for data resulting from a quantitative variable.

Example: Identify each of the following examples as attribute (qualitative) or numerical (quantitative) variables.

1. The residence hall for each student in a statistics class. (Attribute)
2. The amount of gasoline pumped by the next 10 customers at the local Unimart. (Numerical)
3. The amount of radon in the basement of each of 25 homes in a new development. (Numerical)
4. The color of the baseball cap worn by each of 20 students. (Attribute)
5. The length of time to complete a mathematics homework assignment. (Numerical)
6. The state in which each truck is registered when stopped and inspected at a weigh station. (Attribute)

Qualitative and quantitative variables may be further subdivided:



Nominal Variable: A qualitative variable that categorizes (or describes, or names) an element of a population.

Ordinal Variable: A qualitative variable that incorporates an ordered position, or ranking.

Discrete Variable: A quantitative variable that can assume a countable number of values. Intuitively, a discrete variable can assume values corresponding to isolated points along a line interval. That is, there is a gap between any two values.

Continuous Variable: A quantitative variable that can assume an uncountable number of values. Intuitively, a continuous variable can assume any value along a line interval, including every possible value between any two values.

Note:

1. In many cases, a discrete and continuous variable may be distinguished by determining whether the variables are related to a count or a measurement.
2. Discrete variables are usually associated with counting. If the variable cannot be further subdivided, it is a clue that you are probably dealing with a discrete variable.
3. Continuous variables are usually associated with measurements. The values of discrete variables are only limited by your ability to measure them.

Example: Identify each of the following as examples of qualitative or numerical variables:

1. The temperature in Barrow, Alaska at 12:00 pm on any given day.
2. The make of automobile driven by each faculty member.
3. Whether or not a 6 volt lantern battery is defective.
4. The weight of a lead pencil.
5. The length of time billed for a long distance telephone call.
6. The brand of cereal children eat for breakfast.
7. The type of book taken out of the library by an adult.

Example: Identify each of the following as examples of (1) nominal, (2) ordinal, (3) discrete, or (4) continuous variables:

1. The length of time until a pain reliever begins to work.
2. The number of chocolate chips in a cookie.
3. The number of colors used in a statistics textbook.
4. The brand of refrigerator in a home.
5. The overall satisfaction rating of a new car.
6. The number of files on a computer's hard disk.
7. The pH level of the water in a swimming pool.
8. The number of staples in a stapler.

Measure and Variability

- No matter what the response variable: there will always be **variability** in the data.
- One of the primary objectives of statistics: measuring and characterizing variability.
- Controlling (or reducing) variability in a manufacturing process: statistical process control.

Example: A supplier fills cans of soda marked 12 ounces. How much soda does each can really contain?

- It is very *unlikely* any one can contains exactly 12 ounces of soda.
- There is variability in any process.
- Some cans contain a little more than 12 ounces, and some cans contain a little less.
- On the average, there are 12 ounces in each can.
- The supplier hopes there is little variability in the process, that most cans contain *close* to 12 ounces of soda.

Data Collection

- First problem a statistician faces: how to obtain the data.
- It is important to obtain *good*, or *representative*, data.
- Inferences are made based on statistics obtained from the data.
- Inferences can only be as good as the data.

Biased Sampling Method: A sampling method that produces data which systematically differs from the sampled population. An **unbiased sampling method** is one that is not biased.

Sampling methods that often result in biased samples:

1. **Convenience sample**: sample selected from elements of a population that are easily accessible.
2. **Volunteer sample**: sample collected from those elements of the population which chose to contribute the needed information on their own initiative.

Process of data collection:

1. Define the objectives of the survey or experiment.

   Example: Estimate the average life of an electronic component.

2. Define the variable and population of interest.

   Example: Length of time for anesthesia to wear off after surgery.

3. Defining the data-collection and data-measuring schemes. This includes sampling procedures, sample size, and the data-measuring device (questionnaire, scale, ruler, etc.).

4. Determine the appropriate descriptive or inferential data-analysis techniques.

Methods used to collect data:

Experiment: The investigator controls or modifies the environment and observes the effect on the variable under study.

Survey: Data are obtained by sampling some of the population of interest. The investigator does not modify the environment.

Census: A 100% survey. Every element of the population is listed. Seldom used: difficult and time-consuming to compile, and expensive.

Sampling Frame: A list of the elements belonging to the population from which the sample will be drawn.

Note: It is important that the sampling frame be representative of the population.

Sample Design: The process of selecting sample elements from the sampling frame.

Note: There are many different types of sample designs. Usually they all fit into two categories: judgment samples and probability samples.

Judgment Samples: Samples that are selected on the basis of being "typical."

Items are selected that are representative of the population. The validity of the results from a judgment sample reflects the soundness of the collector's judgment.

Probability Samples: Samples in which the elements to be selected are drawn on the basis of probability. Each element in a population has a certain probability of being selected as part of the sample.

Random Samples: A sample selected in such a way that every element in the population has a equal probability of being chosen. Equivalently, all samples of size *n* have an equal chance of being selected. Random samples are obtained either by sampling with replacement from a finite population or by sampling without replacement from an infinite population.

Note:

1. Inherent in the concept of randomness: the next result (or occurrence) is not predictable.
2. Proper procedure for selecting a random sample: use a random number generator or a table of random numbers.

Example: An employer is interested in the time it takes each employee to commute to work each morning. A random sample of 35 employees will be selected and their commuting time will be recorded.

There are 2712 employees.

Each employee is numbered: 0001, 0002, 0003, etc. up to 2712.

Using four-digit random numbers, a sample is identified: 1315, 0987, 1125, etc.

Systematic Sample: A sample in which every *k*th item of the sampling frame is selected, starting from the first element which is randomly selected from the first *k* elements.

Note: The systematic technique is easy to execute. However, it has some inherent dangers when the sampling frame is repetitive or cyclical in nature. In these situations the results may not approximate a simple random sample.

Stratified Random Sample: A sample obtained by stratifying the sampling frame and then selecting a fixed number of items from each of the strata by means of a simple random sampling technique.

Proportional Sample (or Quota Sample): A sample obtained by stratifying the sampling frame and then selecting a number of items in proportion to the size of the strata (or by quota) from each strata by means of a simple random sampling technique.

Cluster Sample: A sample obtained by stratifying the sampling frame and then selecting some or all of the items from some of, but not all, the strata.

Comparison of Probability and Statistics

Probability: Properties of the population are assumed known. Answer questions about the sample based on these properties.

Statistics: Use information in the sample to draw a conclusion about the population.

Example: A jar of M&M's contains 100 candy pieces, 15 are red. A handful of 10 is selected.

Probability question: What is the probability that 3 of the 10 selected are red?

Example: A handful of 10 M&M's is selected from a jar containing 1000 candy pieces. Three M&M's in the handful are red.

Statistics question: What is the proportion of red M&M's in the entire jar?

Statistics and the Technology

- The electronic technology has had a tremendous effect on the field of statistics.
- Many statistical techniques are repetitive in nature: computers and calculators are good at this.
- Lots of statistical software packages: MINITAB, SYSTAT, STATA, SAS, Statgraphics, SPSS, and calculators.

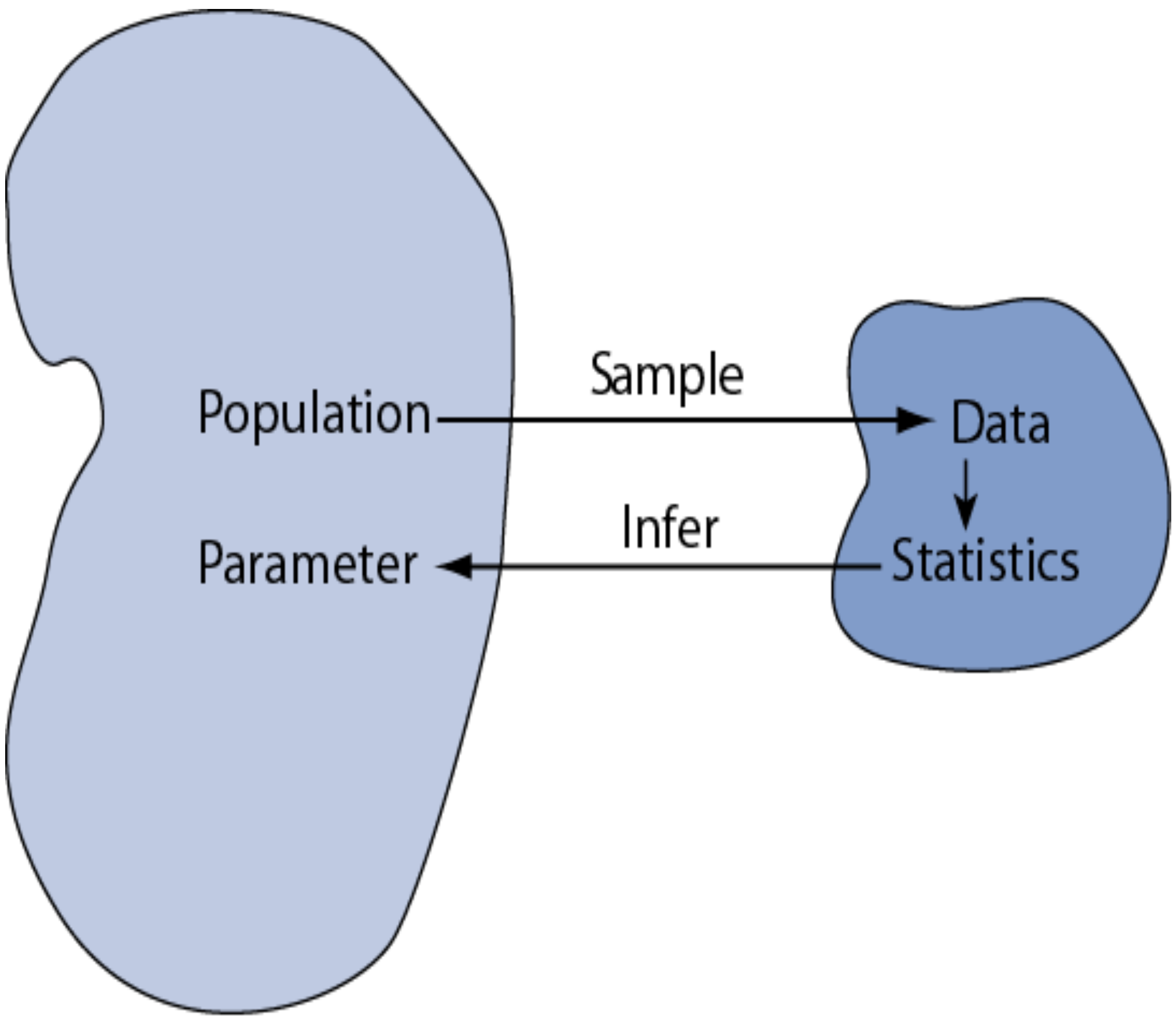
Remember: Responsible use of statistical methodology is very important. The burden is on the user to ensure that the appropriate methods are correctly applied and that accurate conclusions are drawn and communicated to others.

Note: The textbook illustrates statistical procedures using MINITAB, EXCEL 97, and the TI-83.

Statistical Inference

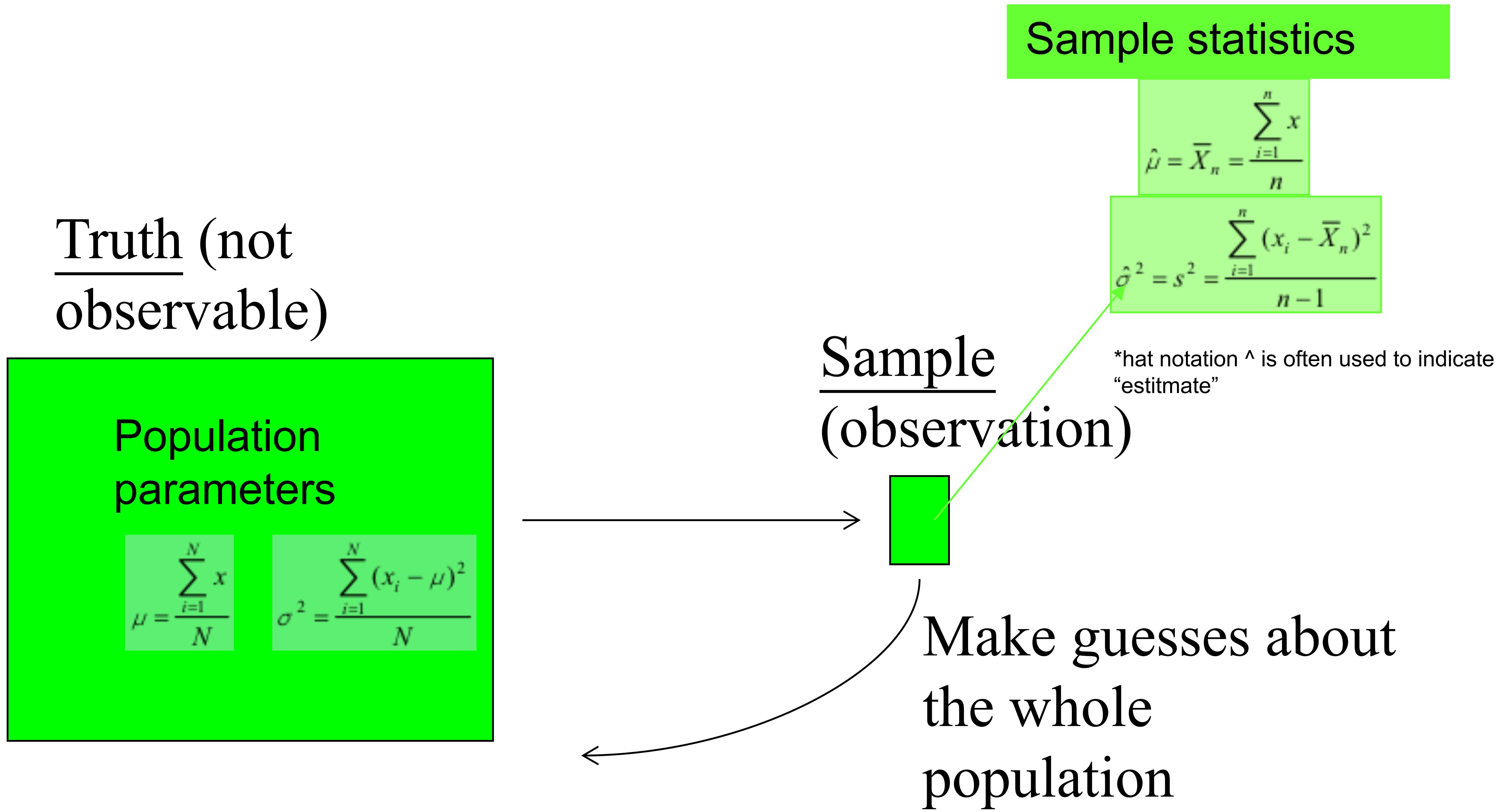
Statistical inference is the act of generalizing from a **sample** to a **population** with calculated degree of certainty.

We want to learn about population *parameter s…*



using statistics calculated in the *sample*

Again, <u>Statistical Inference</u> is the process of making guesses about the truth from a sample.



Parameters and Statistics

We MUST draw distinctions between parameters and statistics

	Parameters	**Statistics**
Source	Population	Sample
Calculated?	No	Yes
Constants?	Yes	No
Examples	μ, σ, *p*	$\bar{x}, s, \hat{p}$

Parameters and Statistics Cont'd.

- **Sample Statistic** – any summary measure calculated from data; e.g., could be a mean, a difference in means or proportions, an odds ratio, or a correlation coefficient
  - E.g., the mean vitamin D level in a sample of 100 men is 63 nmol/L
  - E.g., the correlation coefficient between vitamin D and cognitive function in the sample of 100 men is 0.15

- **Population parameter** – the true value/true effect in the entire population of interest
  - E.g., the true mean vitamin D in all middle-aged and older European men is 62 nmol/L
  - E.g., the true correlation between vitamin D and cognitive function in all middle-aged and older European men is 0.15

Examples of Sample Statistics:

Single population mean

Single population proportion

Difference in means (ttest)

Difference in proportions (Z-test)

Odds ratio/risk ratio

Correlation coefficient

Regression coefficient

…

Statistical Inference

STATISTICAL INFERENCE

Statistical inference provides methods for drawing conclusions about a population from sample data.

There are two forms of statistical inference:

- **Hypothesis ("significance") tests**
- **Confidence intervals**

Example: NEAP math scores

- Young people have a better chance of good jobs and wages if they are good with numbers.
- NEAP math scores
  - Range from 0 to 500
  - Have a Normal distribution
  - Population standard deviation σ is known to be 60
  - Population mean μ *not* known
- We sample *n* = 840 young mean
- Sample mean (“x-bar”) = 272
- Population mean μ unknown
- We want to estimate population mean NEAP score μ

Reference: Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults. *Journal of Human Resources*, **27,** 313-328.

Conditions for Inference

1. Data acquired by **S**imple **R**andom **S**ample
2. Population distribution is **Normal** or large sample
3. The value of **σ is known**
4. The value of **μ is NOT known**

The distribution of potential sample means: **The Sampling Distribution of the Mean**

- Sample means will vary from sample to sample
- In theory, the sample means form a **sampling distribution**
- The sampling distribution of means is Normal with mean μ and **standard deviation** equal to population standard deviation σ divided by the square root of *n*:

This relationship is known as the **square root law**

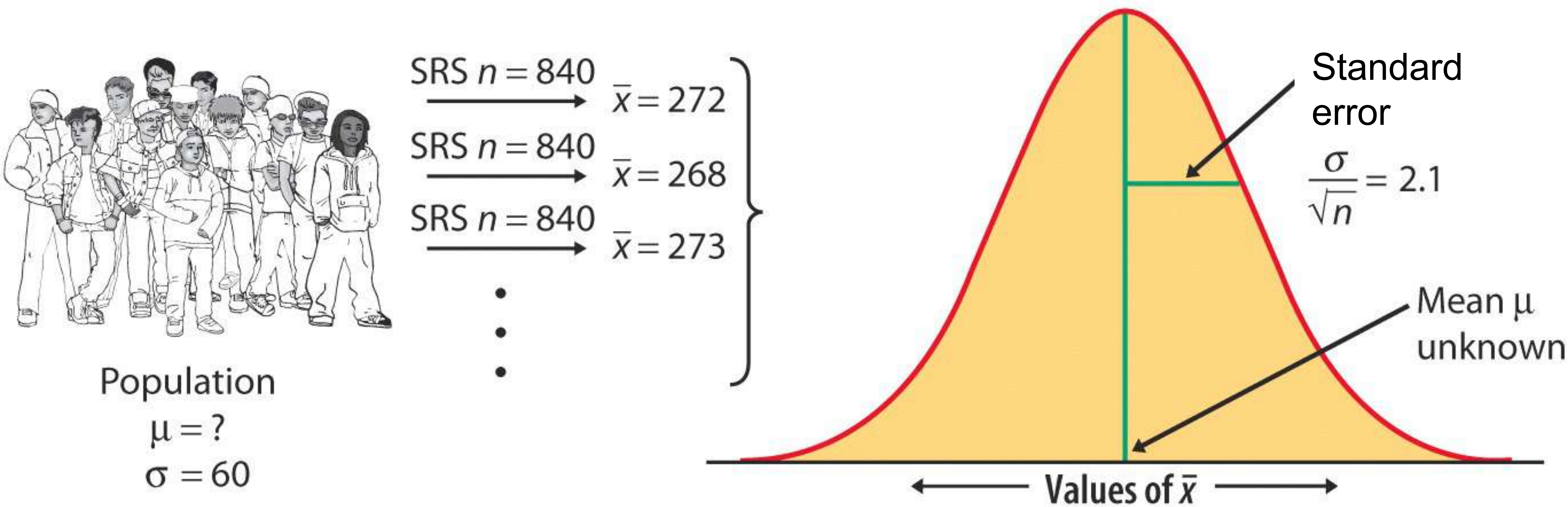
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = SE_{\bar{x}}$$

This statistic is known as the **standard error**

Standard Error of the mean

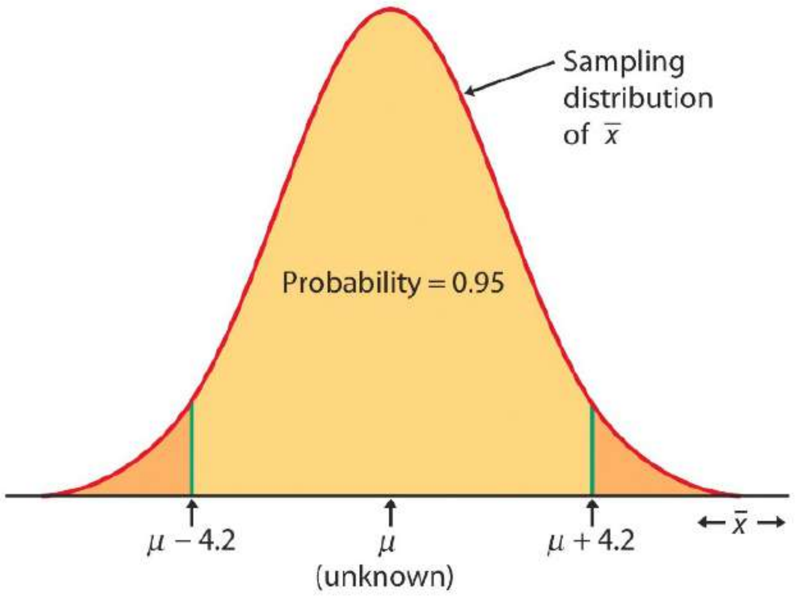
For our example, the population is Normal with σ = 60 (given). Since n = 840,

$$SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{60}{\sqrt{840}} = 2.1$$



Margin of Error *m* for 95% Confidence

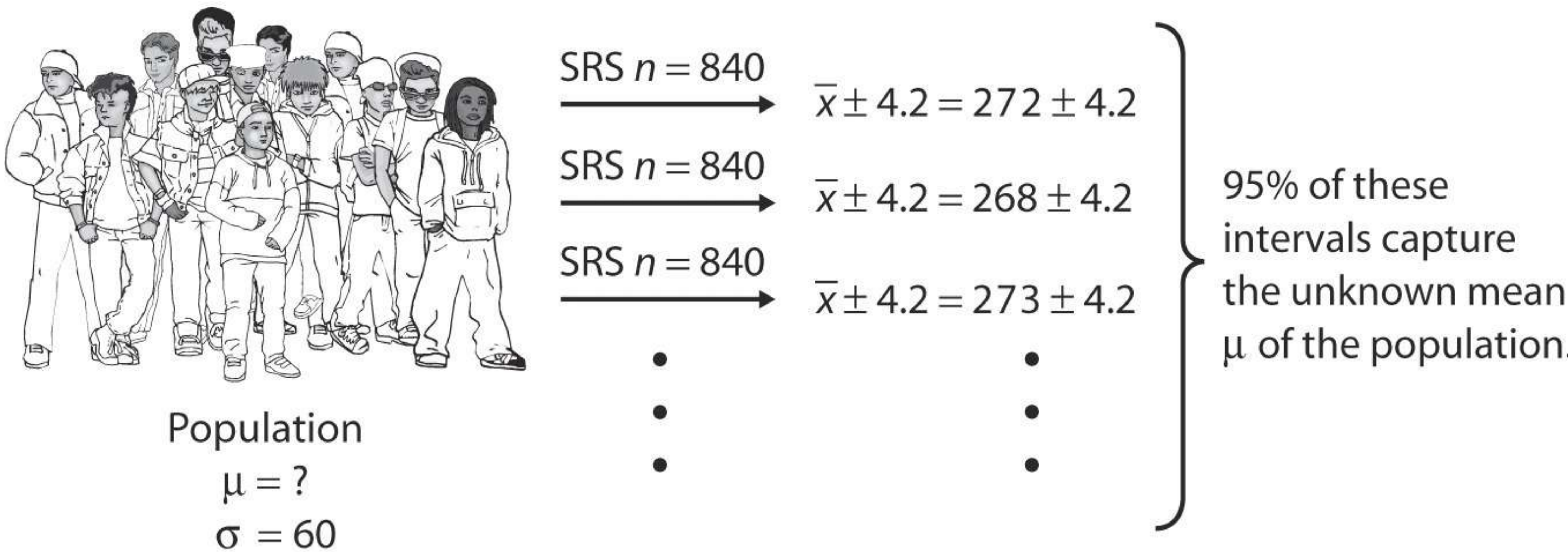
- The 68-95-99.7 rule says 95% of x-bars will fall in the interval $\mu \pm 2 \cdot SE_{xbar}$
- More accurately, 95% will fall in $\mu \pm 1.96 \cdot SE_{xbar}$
- $1.96 \cdot SE_{xbar}$ is the **margin of error *m*** for 95% confidence



For thedata example

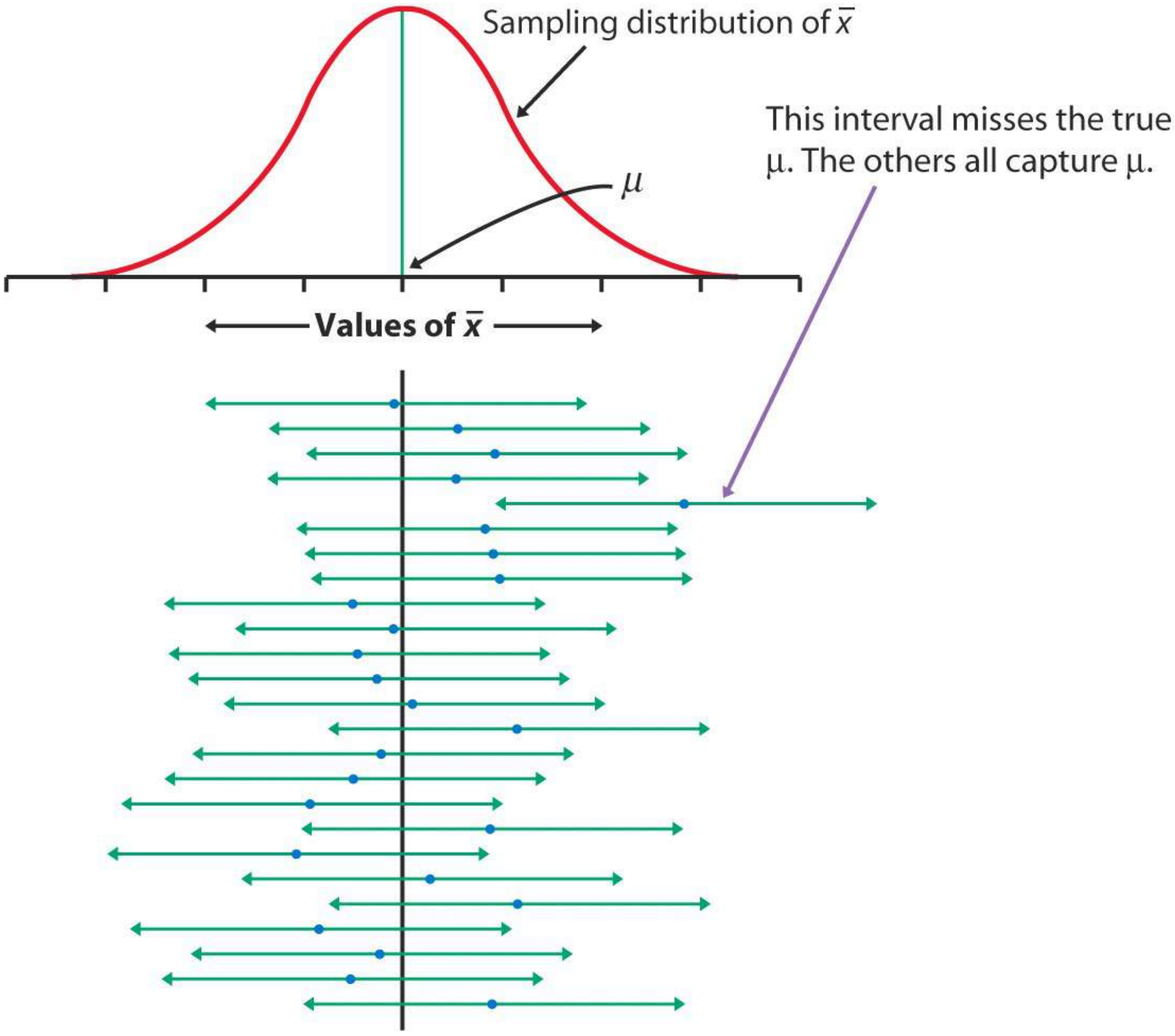
$$m = 1.96 \cdot SE_{\bar{x}}$$
$$= 1.96 \cdot 2.1$$
$$= 4.2$$

In repeated independent samples:



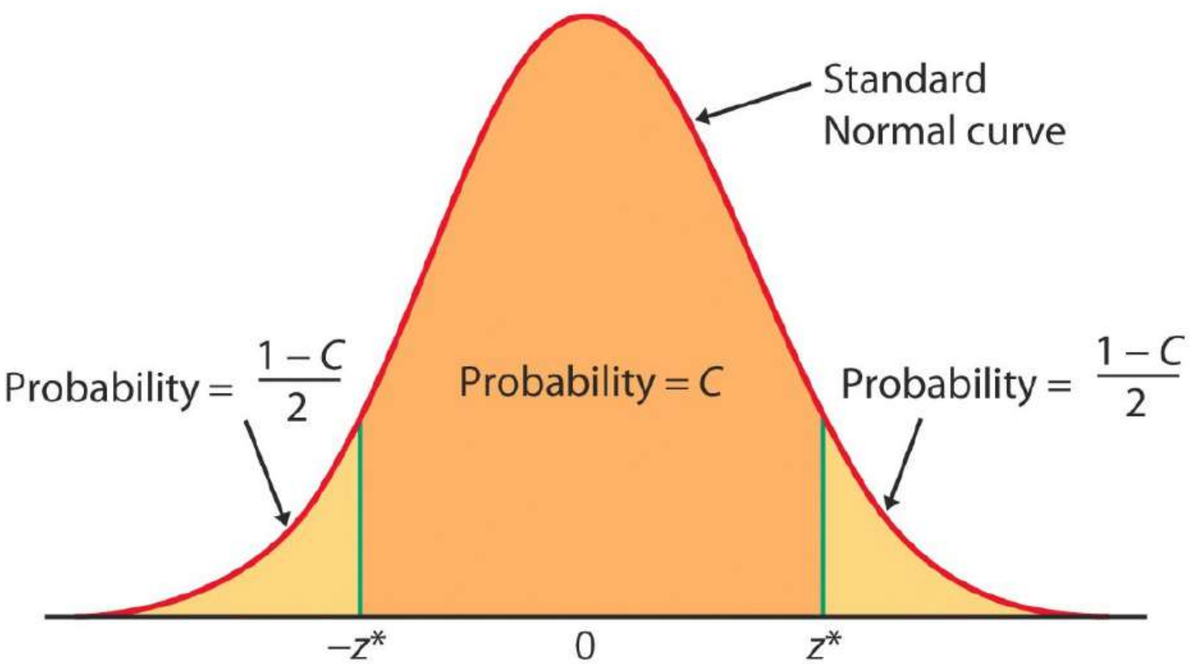
We call these intervals
Confidence Intervals (CIs)

How Confidence Intervals Behave



Other Levels of Confidence

Confidence intervals can be calculated at various levels of confidence by altering coefficient $\mathbf{z_{1-\alpha/2}}$



α ("lack of confidence level")	.10	.05	.01
Confidence level (1–α)100%	90%	95%	99%
$z_{1-\alpha/2}$	1.645	1.960	2.576

(1–α)100% Confidence Interval for μ when σ known

$$\bar{x} \pm z_{1-\frac{\alpha}{2}} \cdot SE_{\bar{x}}$$

$$\text{where } SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Confidence Intervals give:

*A plausible range of values for a population parameter.

*The precision of an estimate.(When sampling variability is high, the confidence interval will be wide to reflect the uncertainty of the observation.)

*Statistical significance (if the 95% CI does not cross the null value, it is significant at .05)

Confidence Intervals

The value of the statistic in my sample (eg., mean, odds ratio, etc.)

point estimate ± (measure of how confident we want to be) × (standard error)

From a Z table or a T table, depending on the sampling distribution of the statistic.

Standard error of the statistic.

Common "Z" levels of confidence

- Commonly used confidence levels are 90%, 95%, and 99%

Confidence Level	*Z value*
80%	1.28
90%	1.645
95%	1.96
98%	2.33
99%	2.58
99.8%	3.08
99.9%	3.27

99% confidence intervals…

- 99% CI for mean vitamin D:
  - 63 nmol/L ± 2.6 *x* (3.3) = 54.4 – 71.6 nmol/L
- 99% CI for the correlation coefficient:
  - 0.15 ± 2.6 *x* (0.1) = -.11 – .41

Testing Hypotheses

- 1. Is the mean vitamin D in middle-aged and older European men lower than 100 nmol/L (the "desirable" level)?
- 2. Is cognitive function correlated with vitamin D?

Is the mean vitamin D different than 100?

- Start by assuming that the mean = 100
- This is the "null hypothesis"
- This is usually the "straw man" that we want to shoot down
- Determine the distribution of statistics assuming that the null is true…

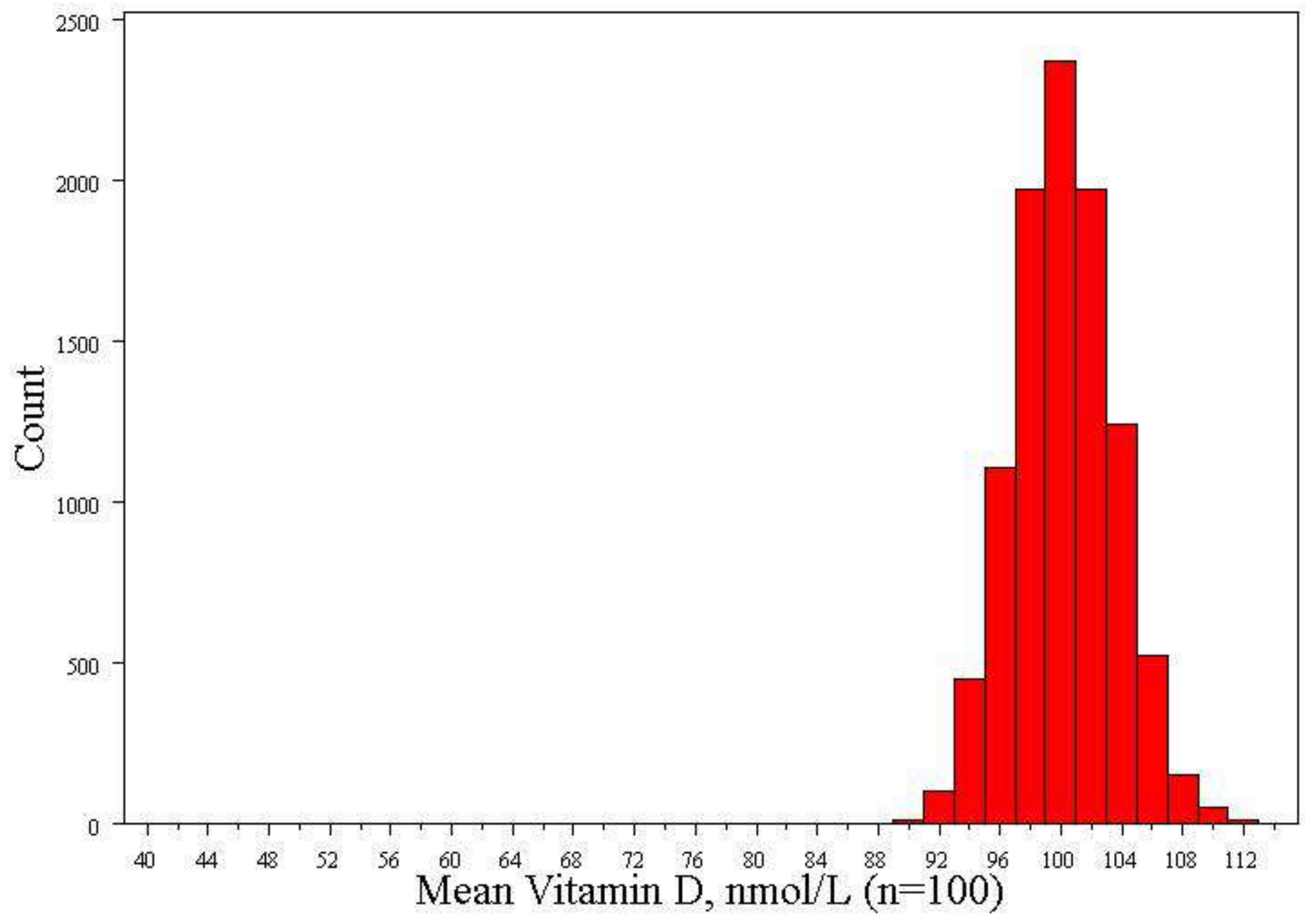
Computer simulation (10,000 repeats)…

This is called the null distribution!

Normally distributed

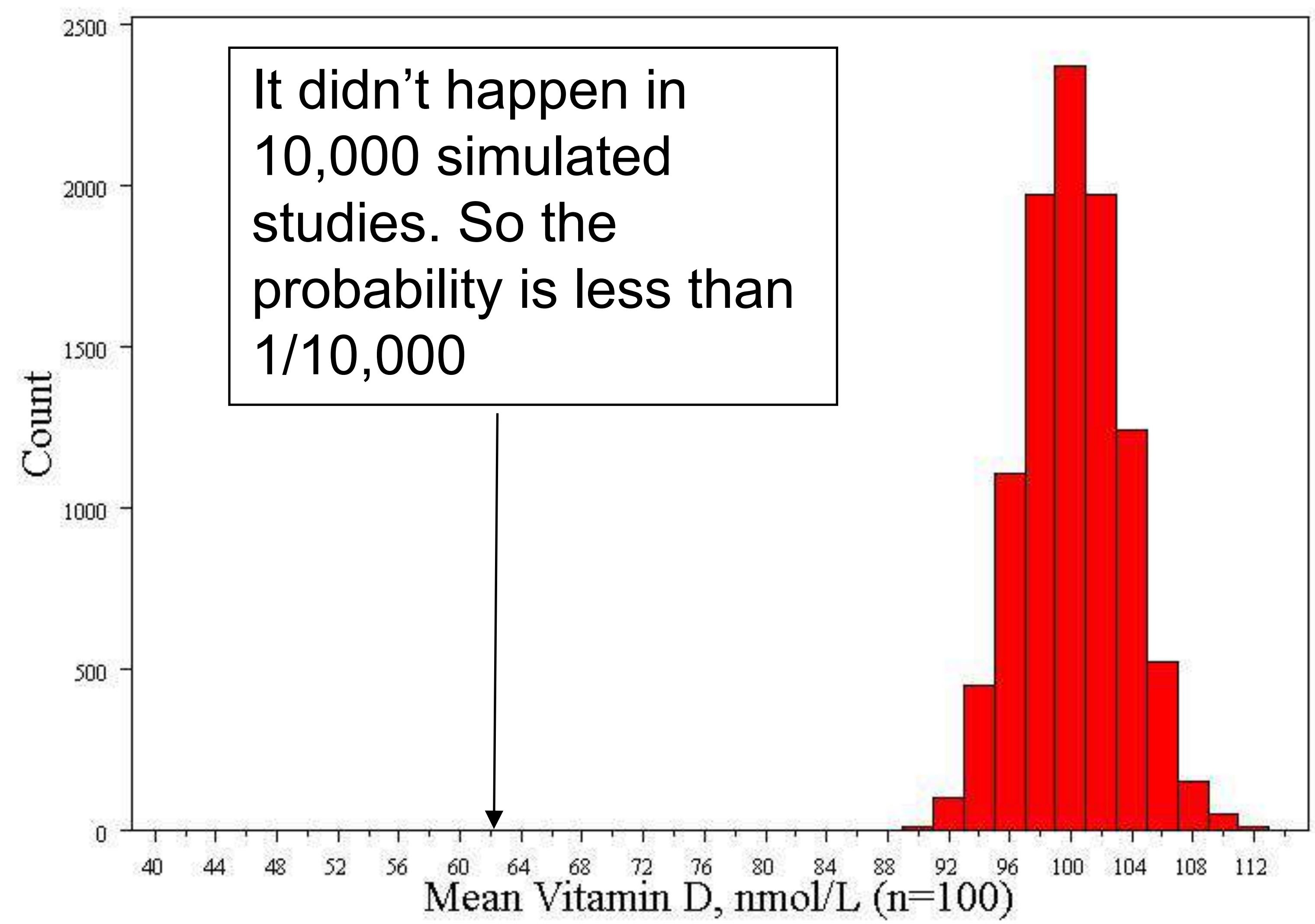
Std error = 3.3

Mean = 100

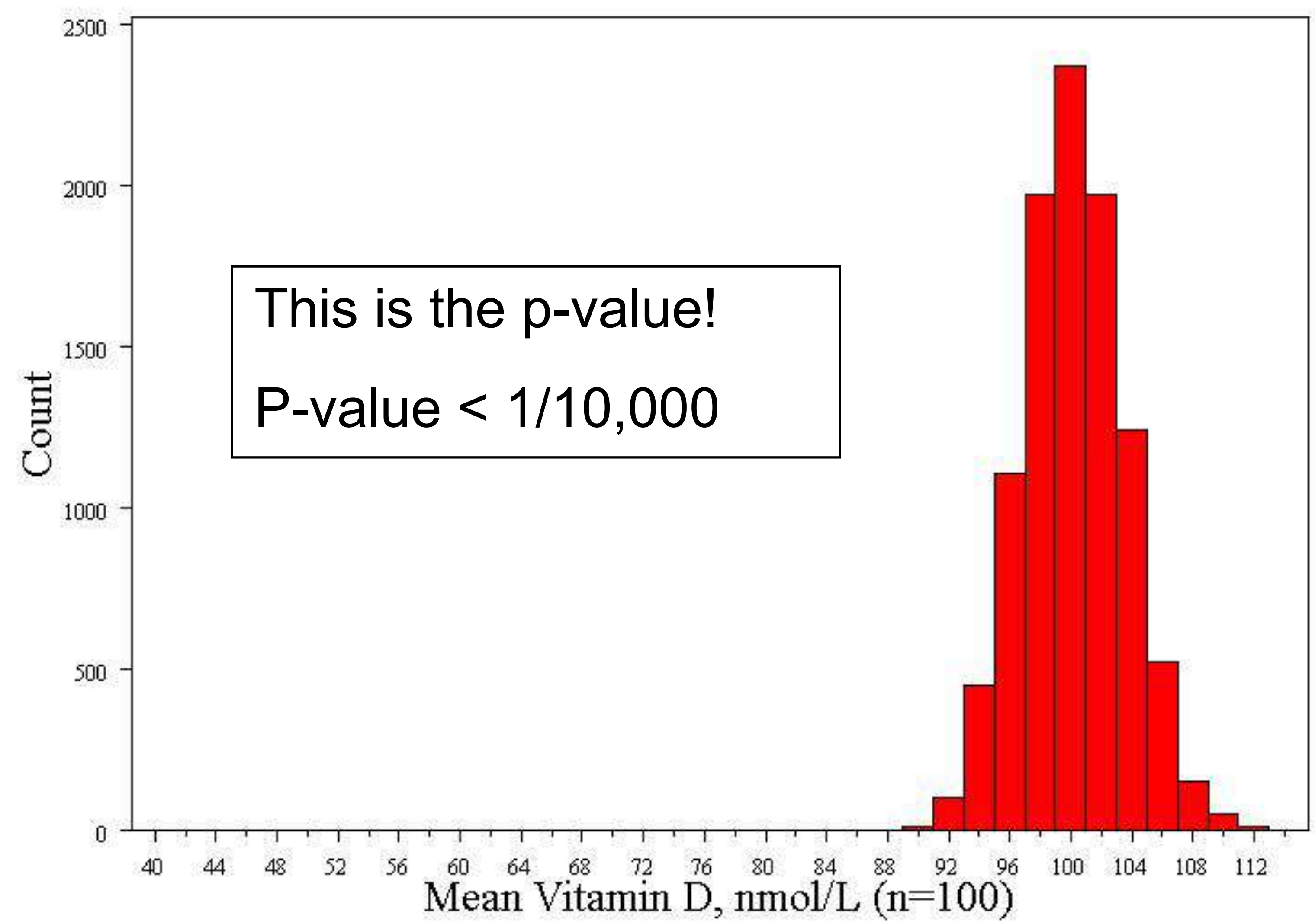


Compare the null distribution to the observed value…

What's the probability of seeing a sample mean of 63 nmol/L if the true mean is 100 nmol/L?



Compare the null distribution to the observed value…



Calculating the p-value with a formula…

Because we know how normal curves work, we can exactly calculate the probability of seeing an average of 63 nmol/L if the true average weight is 100 (i.e., if our null hypothesis is true):

$$Z = \frac{63 - 100}{3.3} = 11.2$$

Z= 11.2, P-value << .0001

The P-value

P-value is the probability that we would have seen our data (or something more unexpected) just by chance if the null hypothesis (null value) is true.

Small p-values mean the null value is unlikely given our data.

Our data are so unlikely given the null hypothesis (<<1/10,000) that I'm going to reject the null hypothesis! (Don't want to reject our data!)

P-value<.0001 means:

The probability of seeing what you saw or something more extreme *if the null hypothesis is true (due to chance)*<.0001

P(empirical data/null hypothesis) <.0001

The P-value

- By convention, p-values of <.05 are often accepted as “statistically significant” in the medical literature; but this is an arbitrary cut-off.

- A cut-off of p<.05 means that in about 5 of 100 experiments, a result would appear significant just by chance (“Type I error”).

Summary: Hypothesis Testing

The Steps:

1. Define your hypotheses (null, alternative)
2. Specify your null distribution
3. Do an experiment
4. Calculate the p-value of what you observed
5. Reject or fail to reject (~accept) the null hypothesis

Hypothesis Testing

The Steps:

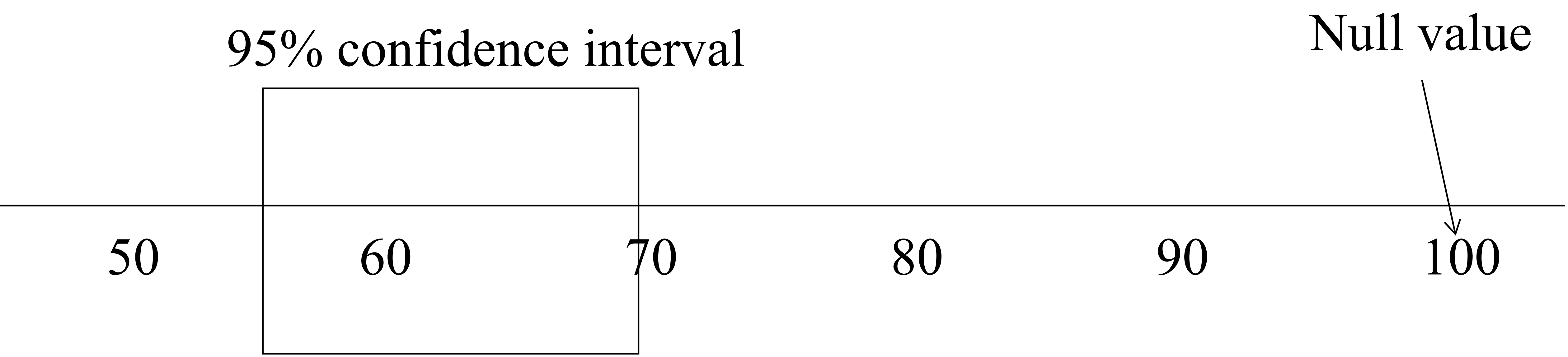
1. Define your hypotheses (null, alternative)
   - The null hypothesis is the "straw man" that we are trying to shoot down.
   - Null here: "mean vitamin D level = 100 nmol/L"
   - Alternative here: "mean vit D < 100 nmol/L" (one-sided)
2. Specify your sampling distribution (under the null)
   - If we repeated this experiment many, many times, the mean vitamin D would be normally distributed around 100 nmol/L with a standard error of 3.3

$$33/\sqrt{100} = 3.3$$

3. Do a single experiment (observed sample mean = 63 nmol/L)
4. Calculate the p-value of what you observed ($p<.0001$)
5. Reject or fail to reject the null hypothesis (reject)

- Confidence intervals give the same information (and more) than hypothesis tests…

Duality with hypothesis tests.

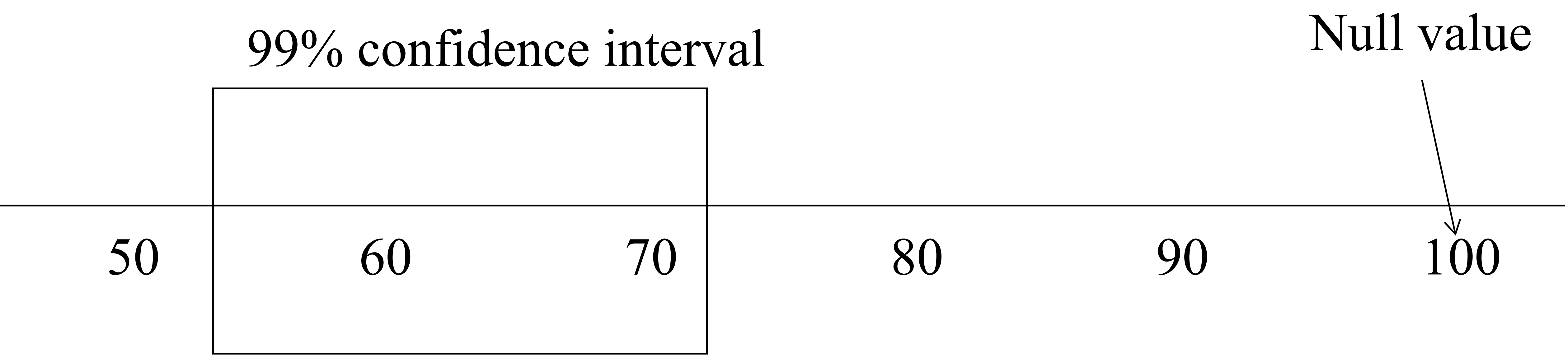


Null hypothesis: Average vitamin D is 100 nmol/L

Alternative hypothesis: Average vitamin D is not 100 nmol/L (two-sided)

P-value < .05

Duality with hypothesis tests.



Null hypothesis: Average vitamin D is 100 nmol/L

Alternative hypothesis: Average vitamin D is not 100 nmol/L (two-sided)

P-value < .01

Summary: Single population mean (large n)

- Hypothesis test:

$$Z = \frac{\text{observed mean} - \text{null mean}}{\frac{s}{\sqrt{n}}}$$

- Confidence Interval

$$\text{confidence interval} = \text{observed mean} \pm Z_{a/2} * (\frac{s}{\sqrt{n}})$$

Single population mean (small n, normally distributed trait)

- Hypothesis test:

$$T_{n-1} = \frac{\text{observed mean} - \text{null mean}}{\frac{s}{\sqrt{n}}}$$

- Confidence Interval

$$\text{confidence interval} = \text{observed mean} \pm \mathrm{T}_{n-1, a/2} * (\frac{s}{\sqrt{n}})$$

2. Is cognitive function correlated with vitamin D?

- Null hypothesis: $r = 0$
- Alternative hypothesis: $r \neq 0$
  - Two-sided hypothesis
  - Doesn't assume that the correlation will be positive or negative.

Examples of CIs for μ

$\sigma = 60 \text{ (given)}$

$\text{Data}: \bar{x} = 272; n = 840$

$$SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{60}{\sqrt{840}} = 2.1$$

Formula

$(1-\alpha)100\%$ CI for $\mu =$

$$\bar{x} \pm z_{1-\frac{\alpha}{2}} \cdot SE_{\bar{x}}$$

For **99%** confidence: $z_{1-(\alpha/2)} = z_{1-(.01/2)} = z_{.995} = 2.576$

$$\bar{x} \pm (2.576)(2.1) = 272 \pm 5.4 = (266.6 \text{ to } 277.4)$$

For **90%** confidence, $z_{1-\alpha/2} = z_{1-.10/2} = z_{.95} = 1.645$

$$\bar{x} \pm (1.645)(2.1) = 272 \pm 3.5 = (268.5 \text{ to } 275.5)$$

Margin of Error (*m*)

Margin of error (*m*) quantifies the precision of the sample mean as an estimatir of μ. The direct formula for *m* is:

$$m = z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

Note that *m* is a function of

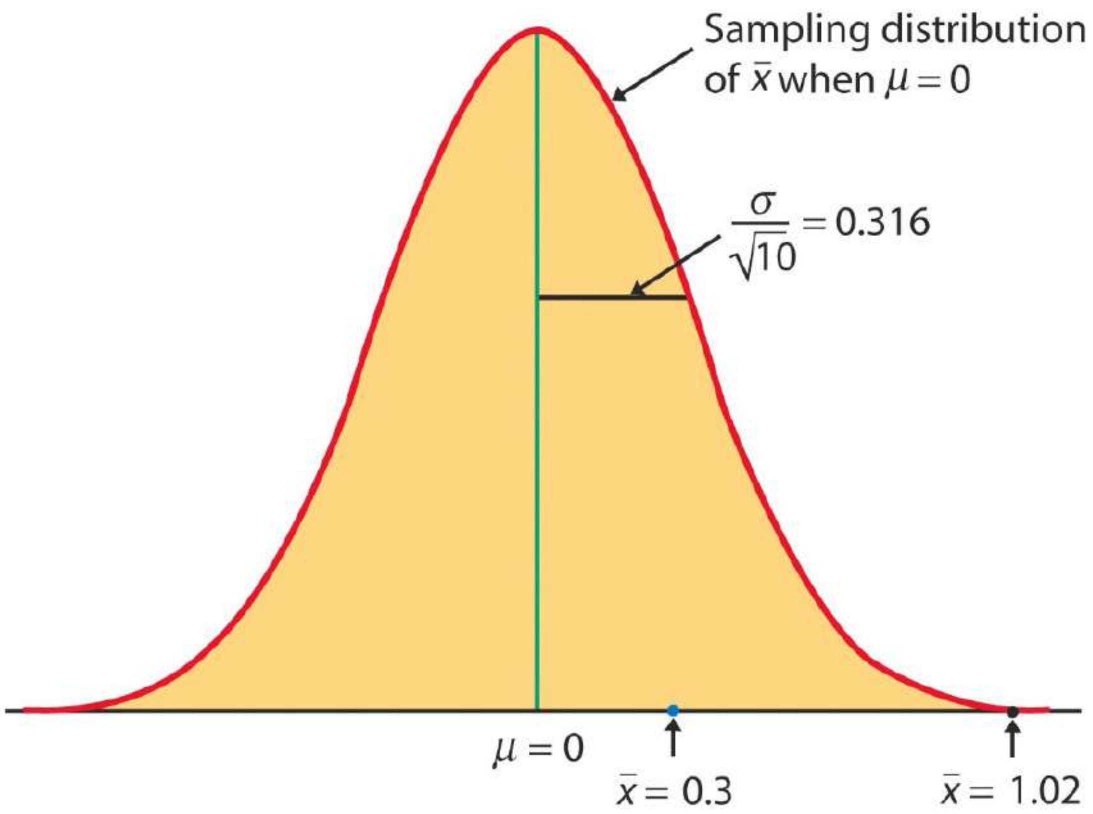
- confidence level 1 – α (as confidence goes up, z increase and m increases)
- population standard deviation σ (this is a function of the variable and cannot be altered)
- sample size *n* (as *n* goes up, *m* decreases; to decrease *m*, increase *n*!)

Tests of Significance

- Recall: two forms of statistical inference
  - Confidence intervals
  - Hypothesis tests of statistical significance
- Objective of confidence intervals: to estimate a population parameter
- Objective of a test of significance: to weight the evidence against a “claim”

Tests of Significance: Reasoning

- As with confidence intervals, we ask what would happen if we repeated the sample or experiment many times
- Let X ≡ weight gain
  - Assume population standard deviation σ = 1
  - Take an SRS of *n* = 10,
  - $SE_{xbar} = 1 / \sqrt{10} = 0.316$
  - Ask: Has there weight gain in the population?



Tests of Statistical Significance: Procedure

A. The claim is stated as a **null hypothesis** $\boldsymbol{H_0}$ and **alternative hypothesis** $\boldsymbol{H_a}$

B. A **test statistic** is calculated from the data

C. The test statistic is converted to a probability statement called a ***P*-value**

D. The *P*-value is **interpreted**

Test for a Population Mean – Null Hypothesis

- ***Example***: We want to test whether data in a sample provides reliable evidence for a population weight gain
- The **null hypothesis** $\boldsymbol{H_0}$ is a statement of "no weight gain"
- In our the **null hypothesis** is H_0: $\mu = 0$

Alternative Hypothesis

- The **alternative hypothesis H_a** is a statement that contradicts the null.
- In our weight gain example, the **alternative hypothesis** can be stated in one of two ways
  - **One-sided alternative** H_a: $\mu > 0$ ("positive weight change in population")
  - **Two-sided alternative** H_a: $\mu \neq 0$ ("weight change in the population")

Test Statistic

$$z_{\text{stat}} = \frac{\bar{x} - \mu_0}{SE_{\bar{x}}}$$

where

$\bar{x} \equiv$ the sample mean

$\mu_0 \equiv$ the value of the parameter under the null hypothesis

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

Test Statistic, Example

$$\text{Given } : \sigma = 1$$

$$\text{Data } : \bar{x} = 1.02; n = 10$$

$$SE_{\bar{x}} = \frac{1}{\sqrt{10}} = 0.3126$$

$$z_{\text{stat}} = \frac{\bar{x} - \mu_0}{SE_{\bar{x}}} = \frac{1.02 - 0}{0.3162} = 3.23$$

P-Value from z table

Convert z statistics to a P-value:

- For H_a: $\mu > \mu_0$
  P-value = $\Pr(Z > z_{stat})$ = right-tail beyond z_{stat}
- For H_a: $\mu < \mu_0$
  P-value = $\Pr(Z < z_{stat})$ = left tail beyond z_{stat}
- For H_a: $\mu \neq \mu_0$
  P-value = 2 × one-tailed P-value

P-value: Interpretation

- ***P*-value (definition)** ≡ the probability the sample mean would take a value as extreme or more extreme than observed test statistic *when* H_0 is true
- ***P*-value (interpretation)** Smaller-and-smaller *P*-values → stronger-and-stronger evidence *against* H_0
- ***P*-value (conventions)**

  $.10 < P < 1.0 \Rightarrow$ evidence against H_0 not significant

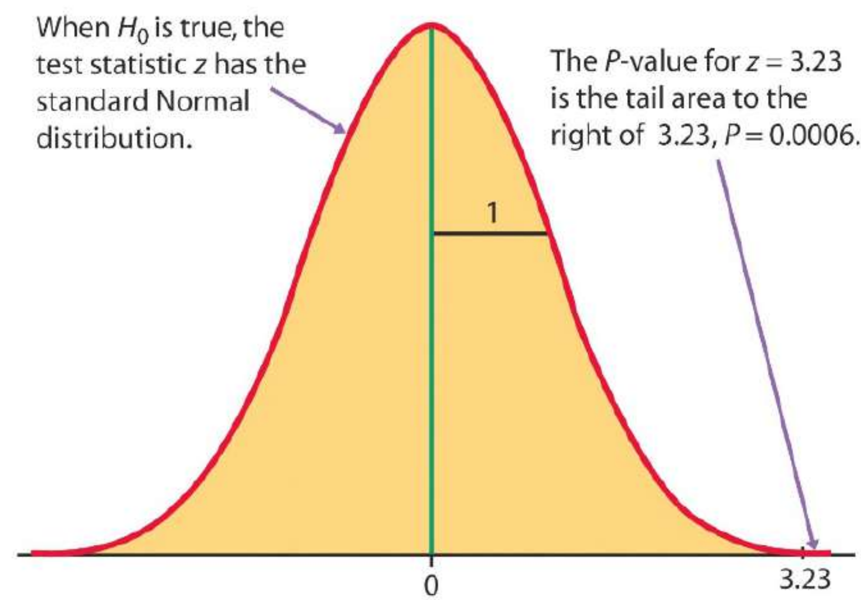
  $.05 < P \leq .10 \Rightarrow$ marginally significant

  $.01 < P \leq .05 \Rightarrow$ significant

  $0 < P \leq .01 \Rightarrow$ highly significant

P-value: Example

- z_{stat} = 3.23
- **One-sided *P*-value**
  = Pr(Z > 3.23)
  = 1 − 0.9994
  = 0.0006
- Two-sided *P-value*
  = 2 × one-sided *P*
  = 2 × 0.0006
  = 0.0012



Conclude: P = .0012
Thus, data provide highly significant evidence against H_0

Significance Level

- **α** ≡ threshold for "significance"
- *If* we choose α = 0.05, we require evidence so strong that it would occur no more than 5% of the time when H_0 is true
- **Decision rule**

  P-value ≤ α ⇒ evidence is significant

  P-value > α ⇒ evidence *not* significant
- For example, let α = 0.01

  P-value = 0.0006

  Thus, *P* < α ⇒ evidence is significant

Summary

z TEST FOR A POPULATION MEAN

Draw an SRS of size n from a Normal population that has unknown mean μ and known standard deviation σ. To test the null hypothesis that μ has a specified value,

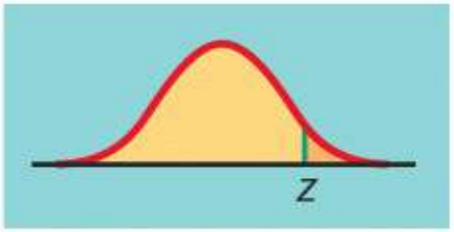
$$H_0\colon \mu = \mu_0$$

calculate the **one-sample z statistic**

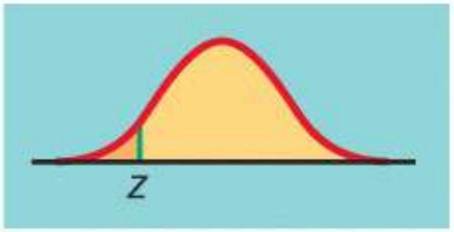
$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

In terms of a variable Z having the standard Normal distribution, the P-value for a test of H_0 against

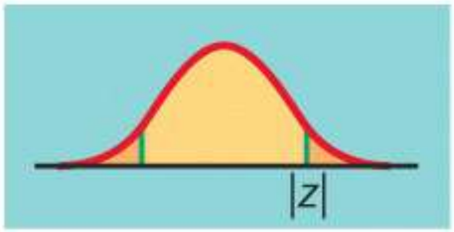
$H_a\colon \mu > \mu_0$ is $P(Z \geq z)$



$H_a\colon \mu < \mu_0$ is $P(Z \leq z)$



$H_a\colon \mu \neq \mu_0$ is $2P(Z \geq |z|)$



Relation Between Tests and CIs

- The value of μ under the null hypothesis is denoted μ_0
- Results are significant at the α-level of when μ_0 falls outside the (1–α)100% CI
- For example, when α = .05 ⇒ (1–α)100% = (1–.05)100% = 95% confidence
- When we tested H_0: μ = 0, two-sided *P* = 0.0012. Since this is significant at α = .05, we expect "0" to fall outside that 95% confidence interval

Relation Between Tests and CIs

Data: x-bar = 1.02, n = 10, σ = 1

$$95\% \text{ CI for } \mu = \bar{x} \pm z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 1.02 \pm 1.96 \frac{1}{\sqrt{10}}$$

$$= 1.02 \pm 0.62 = 0.40 \text{ to} 1.64$$

Notice that 0 falls outside the 95% CI, showing that the test of H_0: μ = 0 will be significant at α = .05