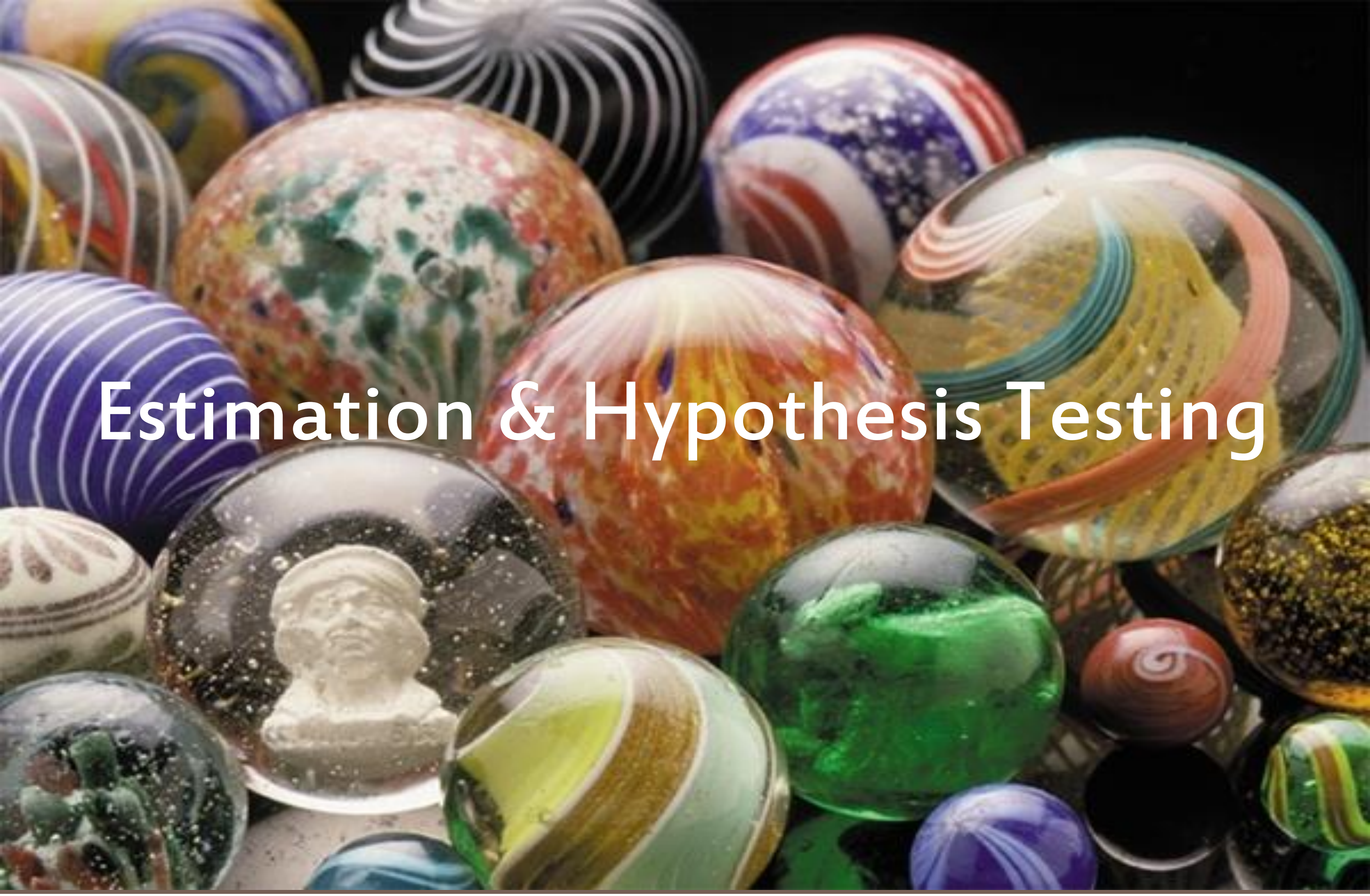
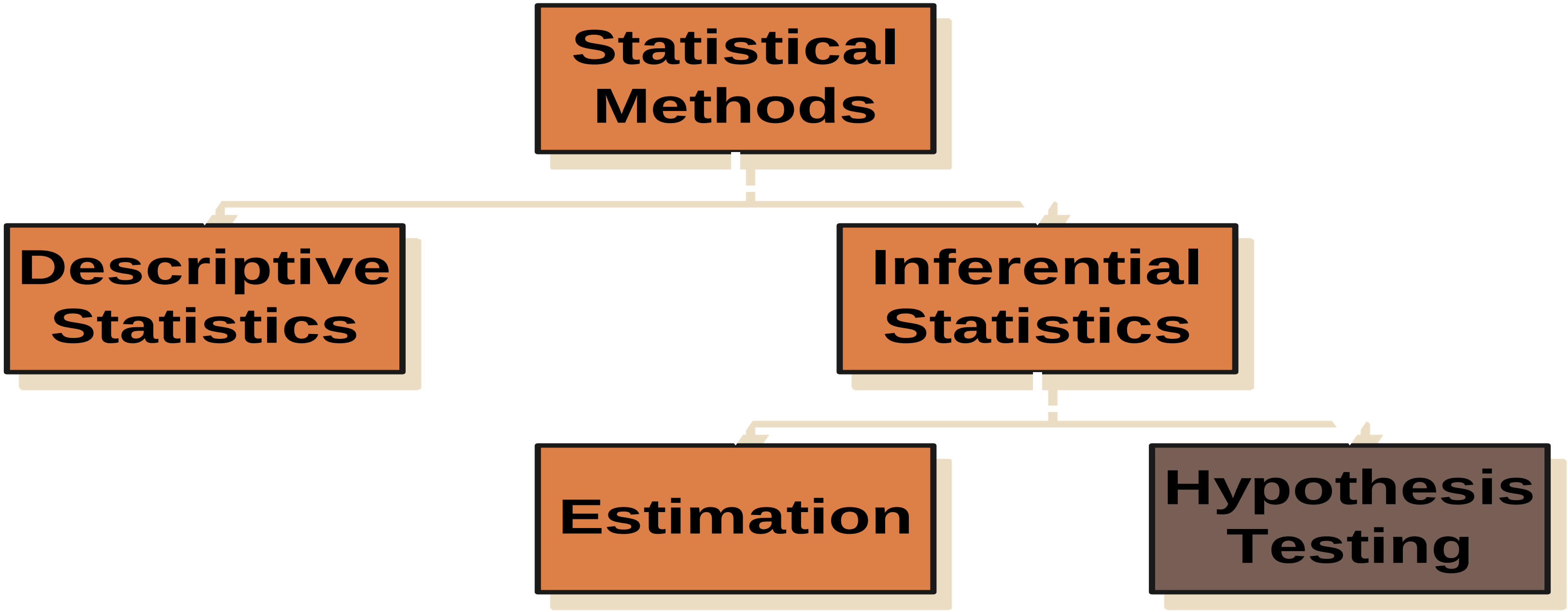


Estimation & Hypothesis Testing

INFERENTIAL STATISTICS

Statistical Methods



Descriptive Statistics Involves organization, summarization, and display of data

Inferential Statistics Involves using a sample to draw conclusions about a population.

Basic tool- Probability

Statistical Inference

- The purpose of statistical inference is to obtain information about a population from information contained in a sample.

- A population is the set of all the elements of interest.
  - Examples: All likely voters in the next election
    All parts produced today by a machine
    All sales receipts for November

- A sample is a subset of the population.
  - Examples: 1000 voters selected at random for interview
    A few parts selected for destructive testing
    Random receipts selected for audit

Why Sample?

- Less time consuming than a census
- Less costly to administer than a census
- It is possible to obtain statistical results of a sufficiently high precision based on samples.

Parameter and Statistic

- *The quantitative measures that describe the characteristics of a population such as mean, standard deviation, proportion are called **parameters**.*

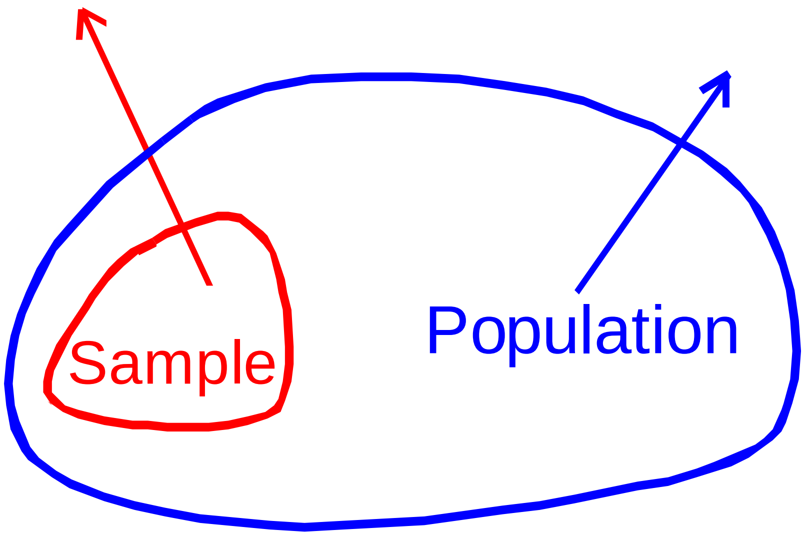
- *The corresponding measures of the sample are called **statistics***

Inferential Statistics

Sample statistics → **Inference** → Population parameters

(known)

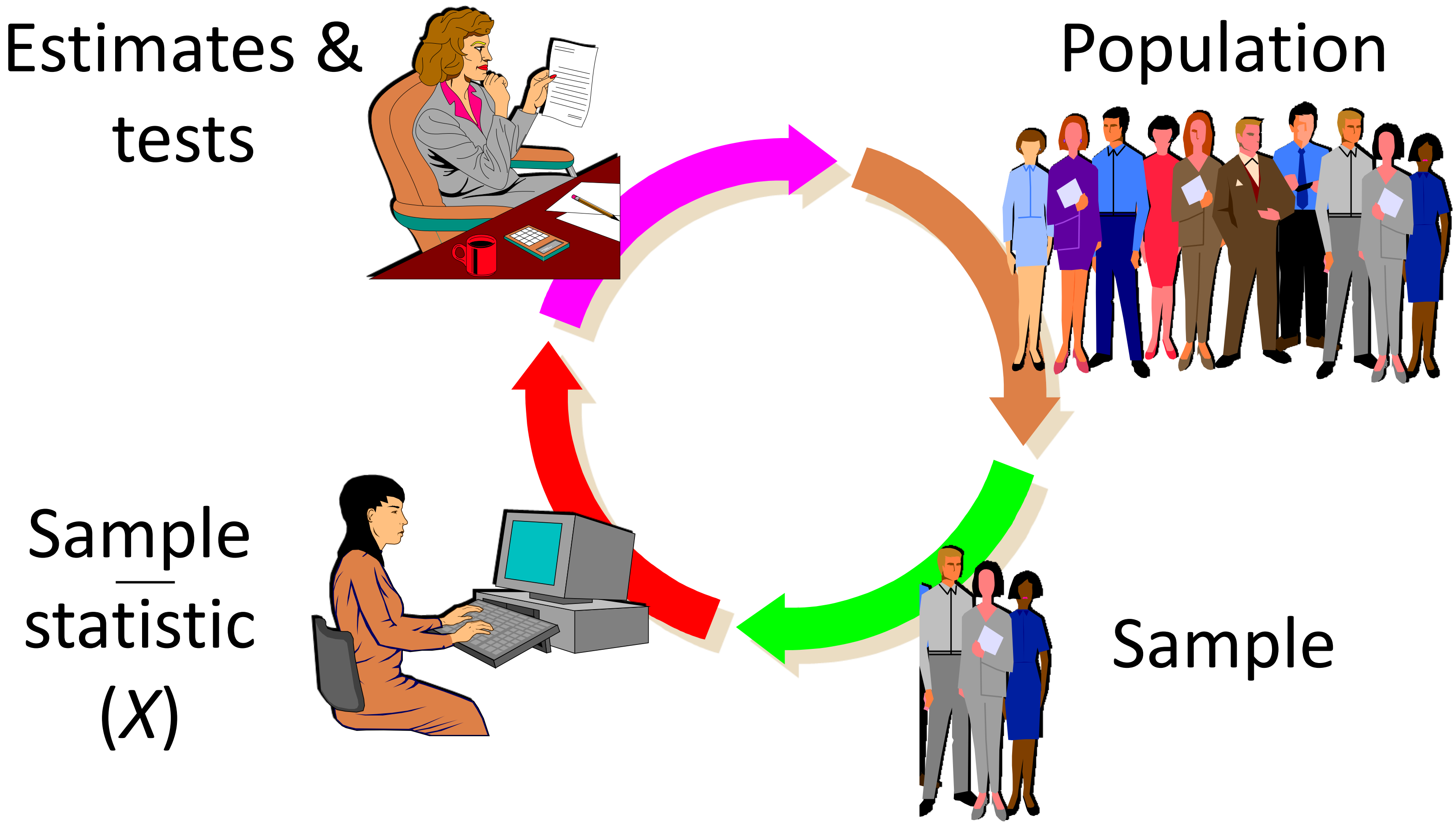
(unknown, but can be estimated from sample evidence)



- **Estimation** – A process whereby we select a **random sample** from a population and use a **sample statistic** to **estimate** a **population parameter.**

- In **Hypothesis testing**, a guess of the population parameter is made. Then the validity of a hypothesis about the population is tested using sample outcomes.

Inference Process



SAMPLING DISTRIBUTIONS

- What is a Sampling Distribution?
- Sampling Distribution of the Mean
- Central Limit Theorem
- Sampling Distribution of Proportions

What is a Sampling Distribution? I

- A ***statistic*** is a numerical quantity calculated in a sample
- A random sample should represent the population well, so sample statistics from a random sample should provide reasonable estimates of population parameters
- All sample statistics have some error in estimating population parameters
- A larger sample provides more information than a smaller sample so a statistic from a large sample should have less error than a statistic from a small sample

In real life calculating parameters of populations is prohibitive because populations are very large.

Rather than investigating the whole population, we take a sample, calculate a **statistic** related to the **parameter** of interest, and make an inference.

The **sampling distribution** of the **statistic** is the tool that tells us how close is the statistic to the parameter.

What is a Sampling Distribution? II

- The sampling distribution of a sample statistic calculated from a sample of *n* measurements is the probability distribution of the statistic.

The Probability distribution of a sample statistic that is formed when a number of samples of size ‘n’ are taken from a population is referred to as sampling distribution.

If the sample statistic is the sample mean, then it is the sampling distribution of the sample means.

Sampling Distributions		
Sampling Distribution of Sample Means	Sampling Distribution of Sample Proportions	Sampling Distribution of Sample Variances

Standard Error

Suppose we want to know the mean height of the fresher at PSG Tech. We take a series of samples and calculate the mean height of each sample. All of these sample means may not be same. There may be some variability in our observed means. This is because of the sampling.

This sampling variability is measured from the standard deviation of the sampling distribution.

The standard deviation of the sampling distribution is called the standard error.

The standard error indicates not only the size of the chance error that has been made, but also the accuracy, by which the sample statistic estimates a population parameter.

The Central Limit Theorem
(CLT)

A Wonderful Theorem

All Roads Lead to Rome…

The shape of the sampling distribution,
the information about the central tendency and dispersion,
can be stated in two theorems.

The Central Limit Theorem (CLT)

THEOREM 1

If samples are drawn from a normal population with mean μ and standard deviation σ , then the sampling distribution of the sample means will be normal with same mean μ

and standard deviation $s = \frac{\sigma}{\sqrt{n}}$

The Central Limit Theorem (CLT)

THEOREM 2

If samples are drawn from any population with mean μ and standard deviation σ, then the sampling distribution of the sample means will be normal with same mean μ and standard deviation $s = \frac{\sigma}{\sqrt{n}}$

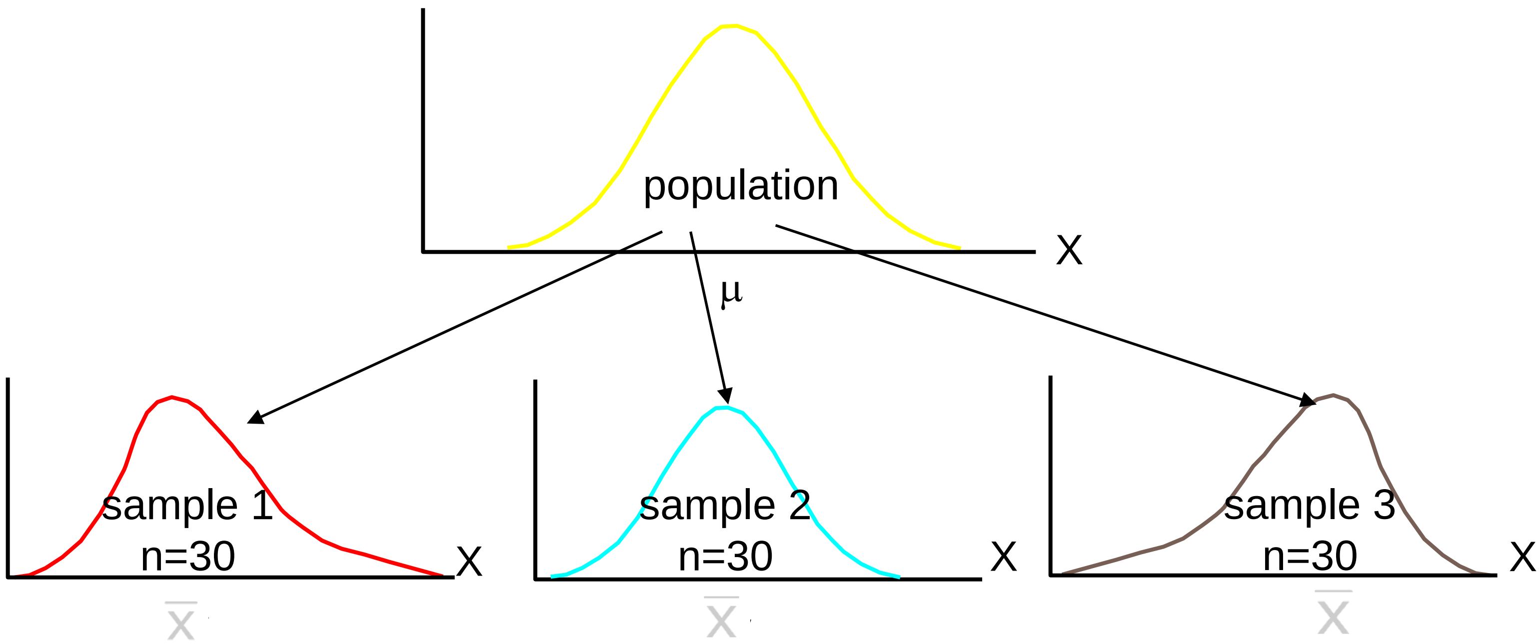
The larger the sample sıze, the better the approximation to the normal distribution.

This theorem is called the Central Limit Theorem.

This theorem removes the constraint of normality in the population.

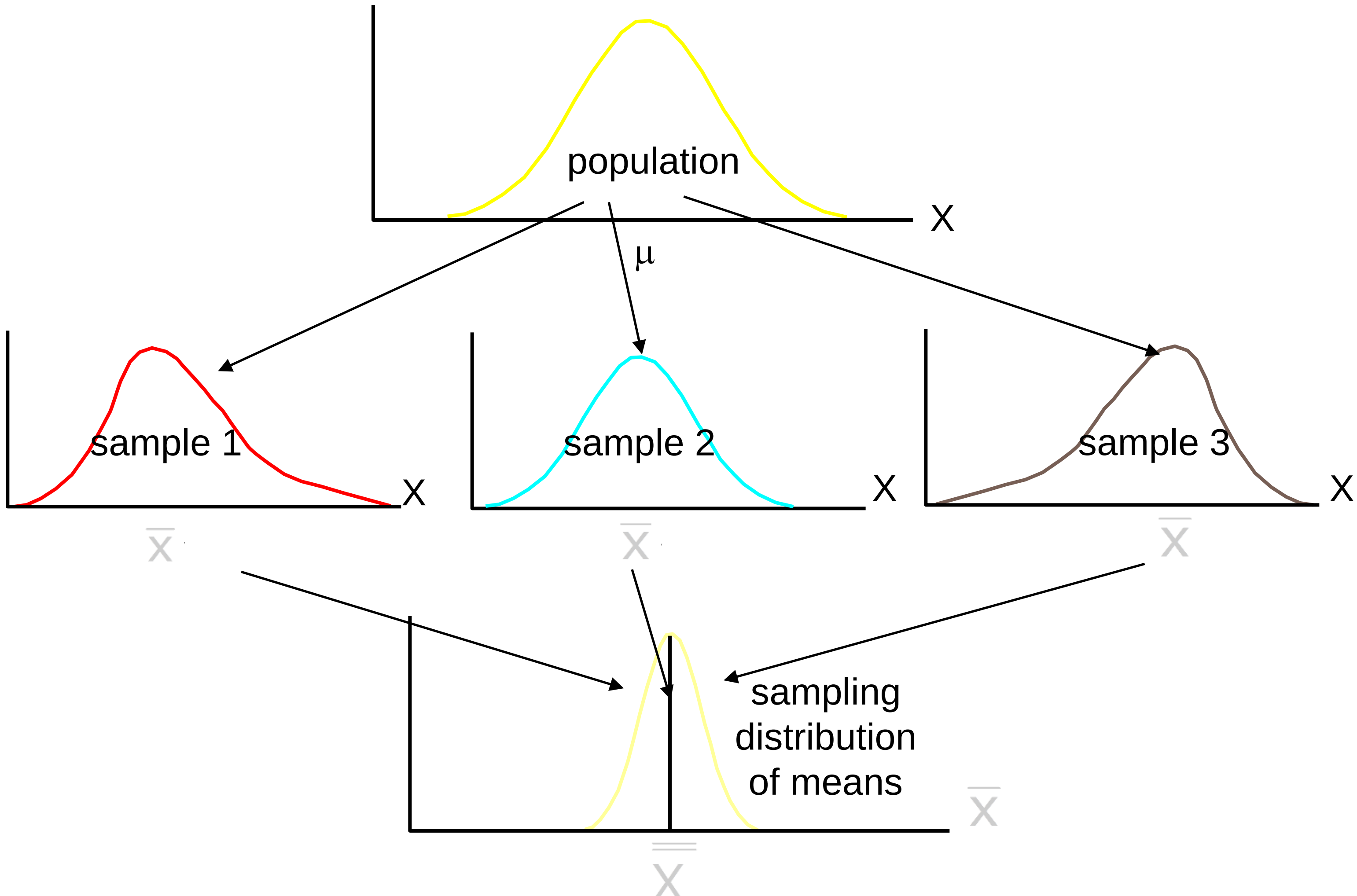
This theorem plays a vital role in inferential statistics. Hence it is called back bone of statistics.

Sampling Distribution of the Mean



If we do this many times…as many times as there are samples of size n=30…

Sampling Distribution of the Mean



Central Limit Theorem

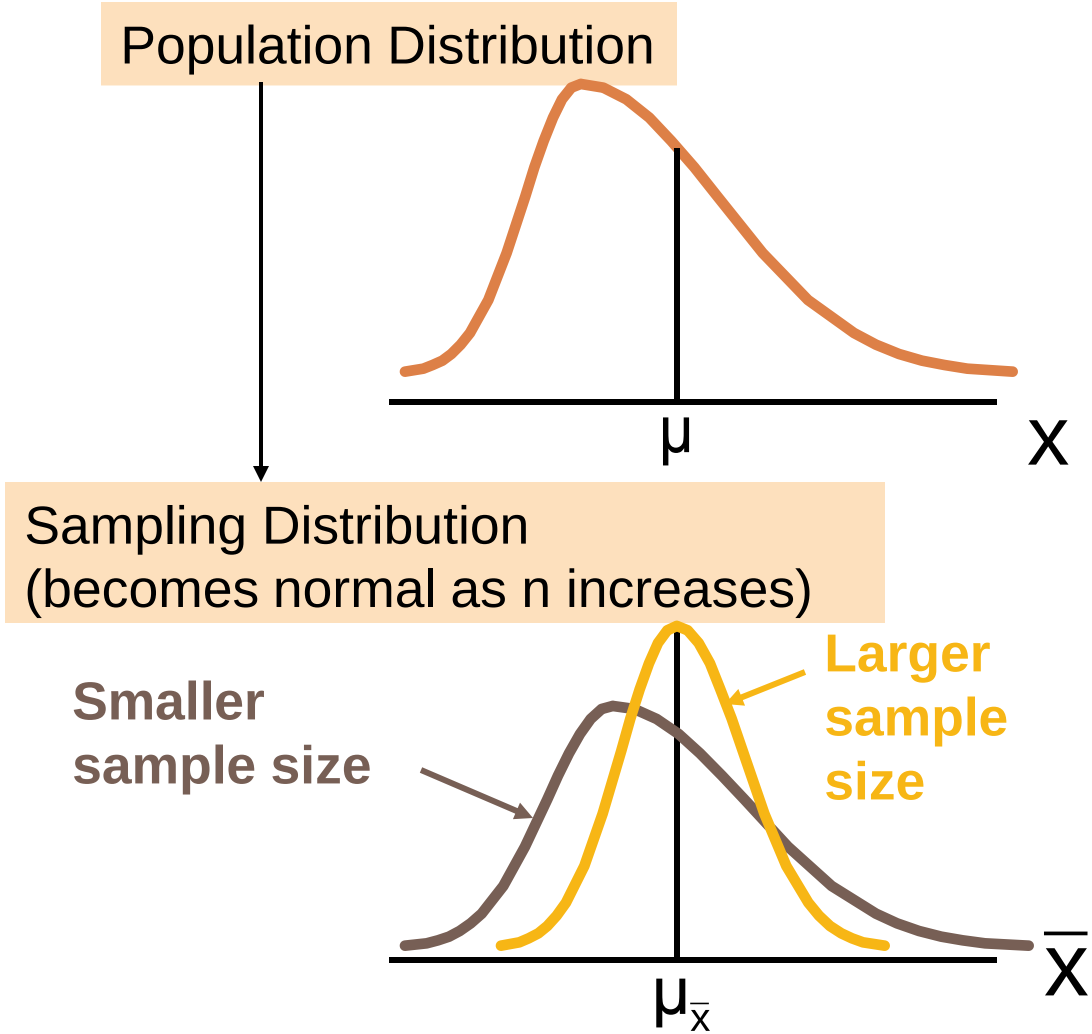
Sampling distribution properties:

Central Tendency

$$\mu_{\bar{x}} = \mu$$

Variation

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



Central Limit Theorem

- When many samples are taken from the same population, the distribution of values for the sample mean are centred around the population mean (regardless of sample size)
- As the sample size increases the mean of the means are closer to the population mean
- The standard deviation of the sample mean decreases as the sample size increases
- The distribution of the sample mean becomes more symmetrical as the sample size gets larger and becomes approximately normal for large sample sizes

Example

- Suppose a population has mean $\mu = 8$ and standard deviation $\sigma = 3$. Suppose a random sample of size $n = 36$ is selected. What is the probability that the sample mean is between 7.8 and 8.2?

Solution:

- Even if the population is not normally distributed, the central limit theorem can be used
- … so the sampling distribution of $\bar{X}$ is approximately normal
- … with mean $\mu_{\bar{X}} = 8$
- …and standard deviation $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{36}} = 0.5$

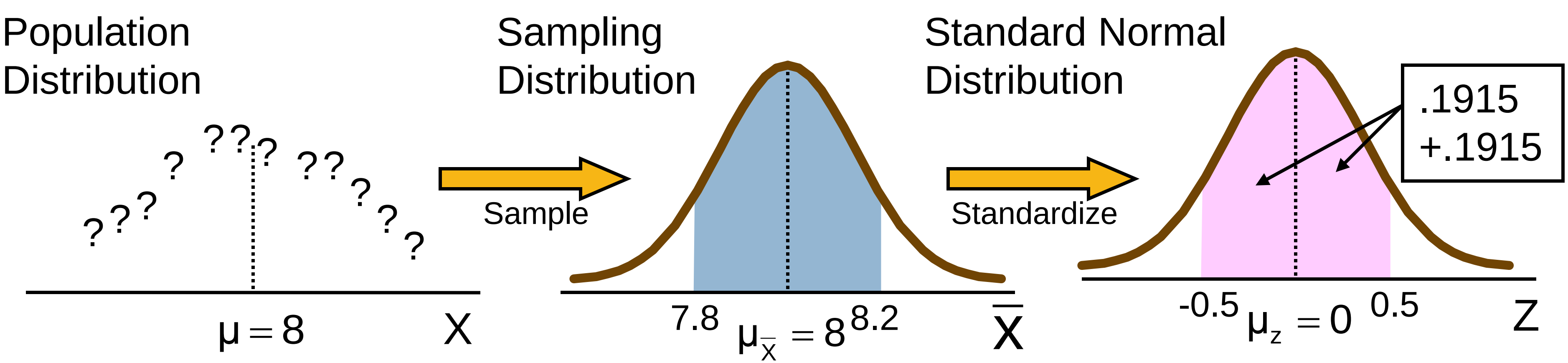
Example

(continued)

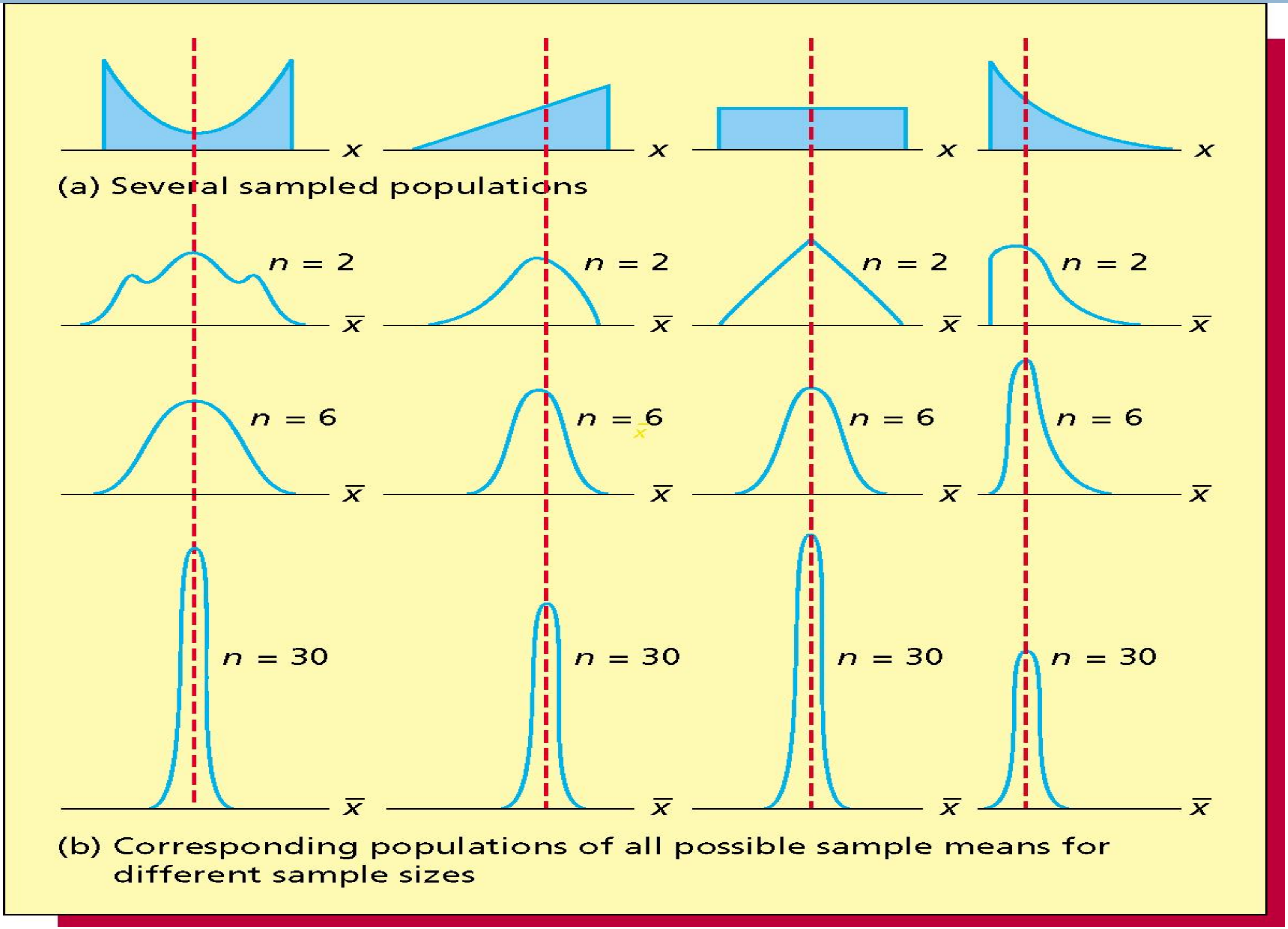
Solution (continued):

$$P(7.8 < \mu_{\bar{X}} < 8.2) = P\left(\frac{7.8 - 8}{3/\sqrt{36}} < \frac{\mu_{\bar{X}} - \mu}{\sigma/\sqrt{n}} < \frac{8.2 - 8}{3/\sqrt{36}} \right)$$

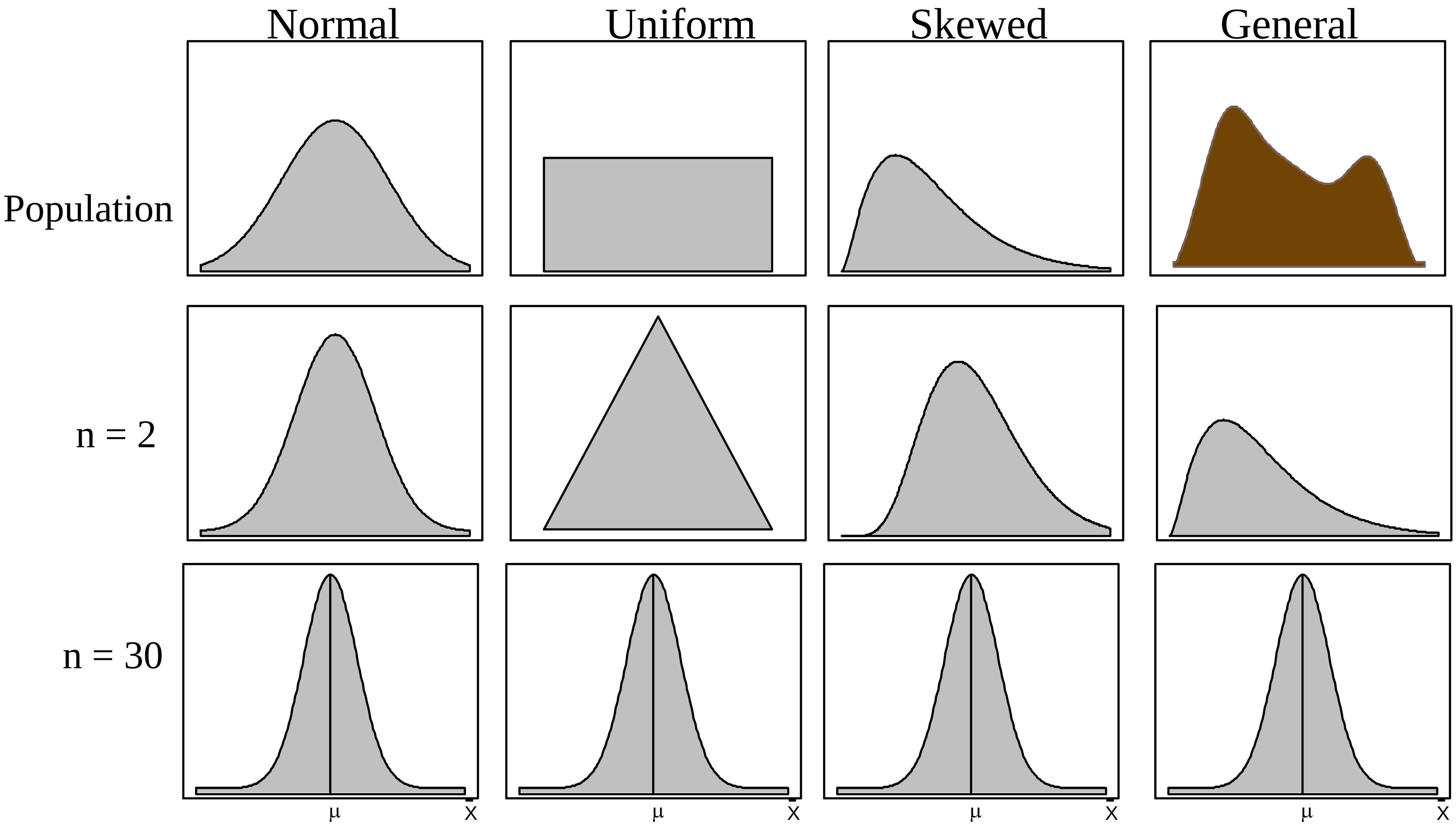
$$= P(-0.5 < Z < 0.5) = \boxed{0.3830}$$



CLT: The Effect of Sample Size I



CLT: The Effect of Sample Size II



Notations

Population Parameter	Sample Statistic
Mean – μ	Mean -
Variance – σ^2	Variance - s^2
Standard Deviation - σ	Standard Deviation – s
Proportion – p	Proportion – $\hat{p}$ = x/n

- Estimation & Estimators
- What is a Confidence Interval?
- Confidence Intervals for Means
- Confidence Intervals for Proportions

CONFIDENCE INTERVALS

Estimation

- Estimation : act of estimating a specific value in the population from the sample
- An Estimator is a function of the sample. It is a rule that tells us how to calculate an estimate of a parameter from the sample.

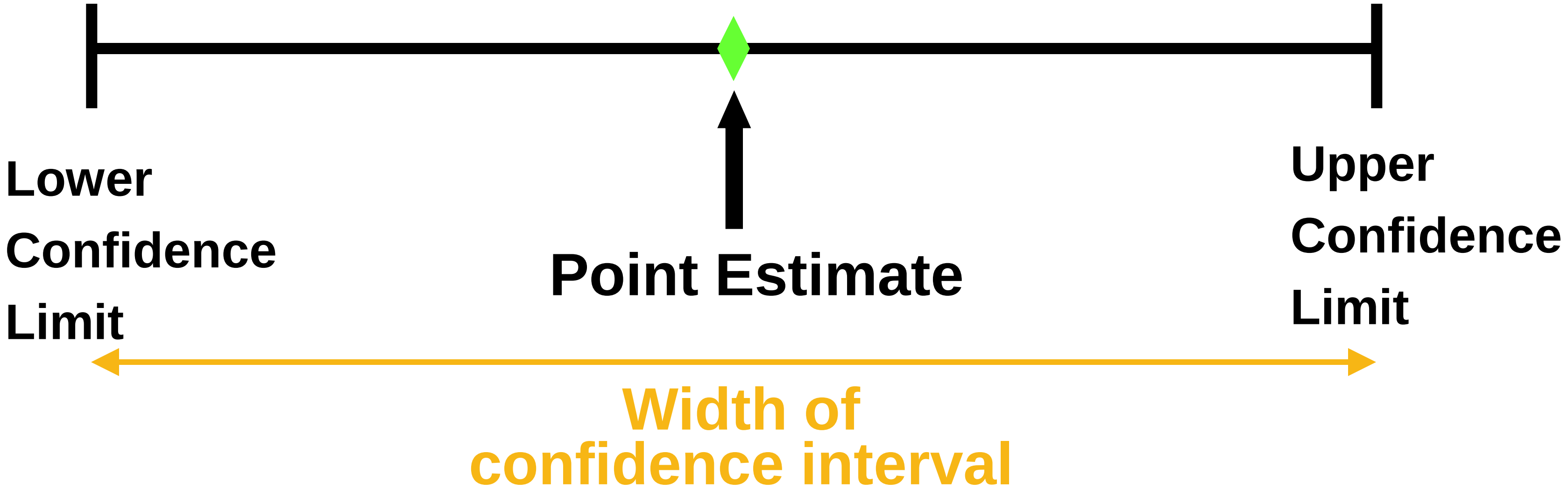
Different estimators are possible for the same parameter. We need to need find the 'good' estimators.

- An Estimate is a value of an estimator calculated from a sample

- There are two types
  - Point Estimation
  - Interval Estimation

Types of Estimates

- **Point Estimate** – A sample statistic used to estimate the **exact value** of a population parameter
- **Confidence interval (*interval estimate*)** – **A range of values** defined by the confidence level within which the population parameter is **estimated** to fall. it therefore provides additional information about variability

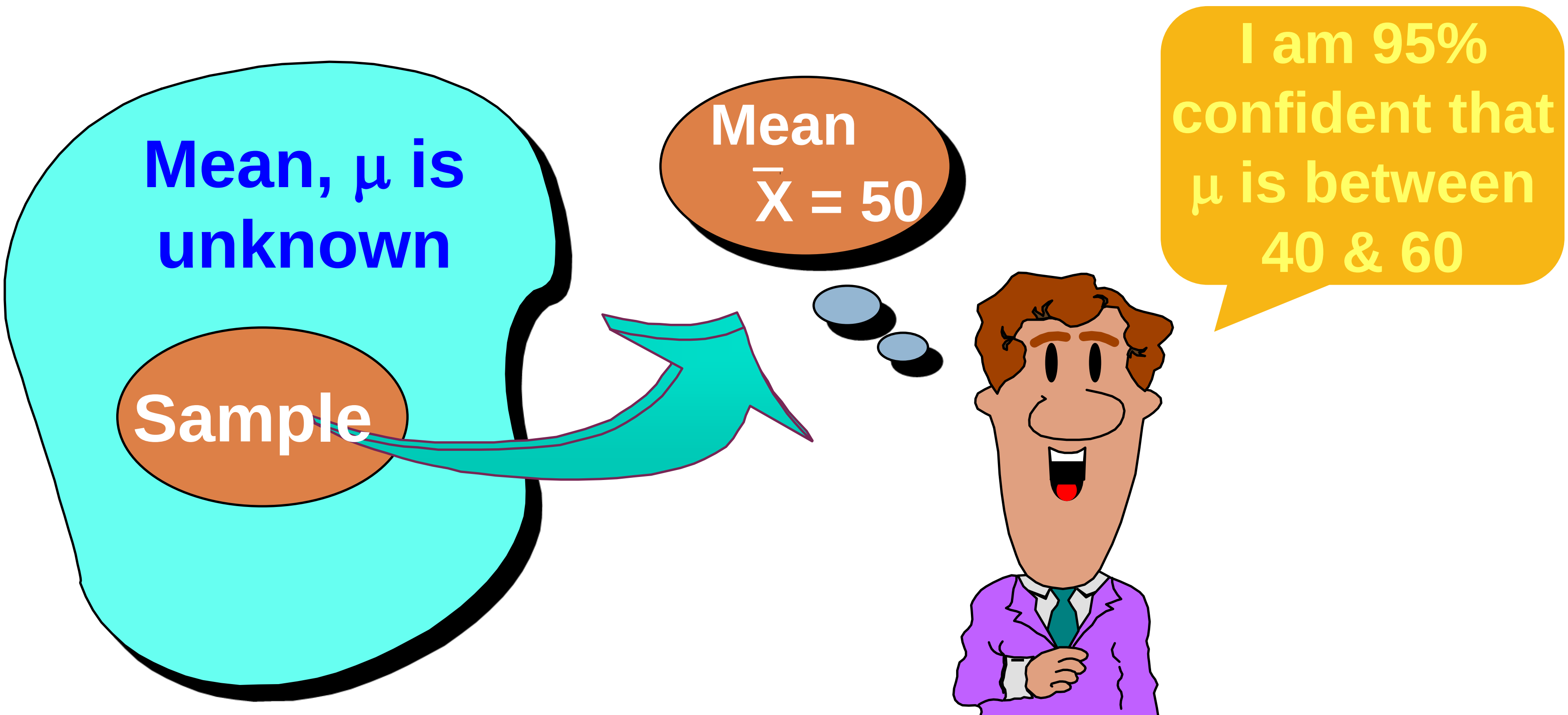


What is a Confidence Interval?

- A confidence interval is a range of guesses at a population value
  - Means
  - Proportions
- It is a type of interval estimation
- The confidence level is that chance (probability) that the range of values captures the true population value (or will contain the unknown population parameter)

Point estimate and Interval estimate

Population **Point estimate** **Interval estimate**



Parameter = Statistic ± Its Error

Point Estimate ± (Critical Value)*(Std. Deviation of the Point Estimate)

Aka: the Standard Error

Confidence Levels:

- **Confidence Level** – The likelihood, expressed as a percentage or a probability, that a specified interval will **contain** the **population parameter.**
- A 95% confidence level means that:

  -95% of the confidence intervals calculated on repeated sampling of the same population will contain μ

  - there is a 0.95 probability that a specified interval DOES contain the population mean.
- Note that the population value does *not* vary i.e. it's not a 95% chance that it falls in that specific interval[1]
- In other words, *the CI attempts to capture the true population mean, but we would have a different interval estimate for each sample drawn*

99% confidence level – there is 1 chance out of 100 that the interval ***DOES NOT*** contain the population mean.

Confidence Interval for μ

Confidence Interval Estimate:

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

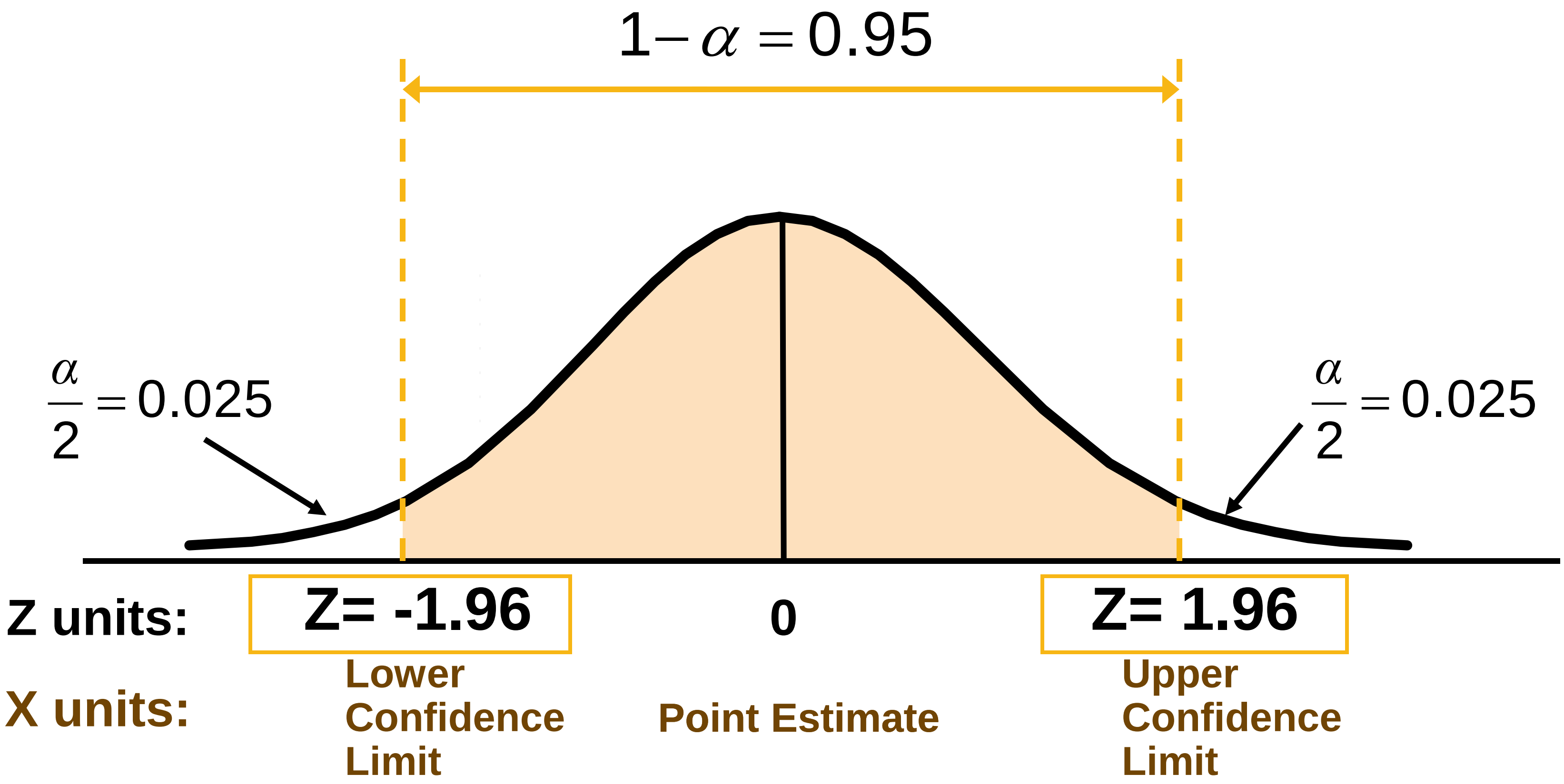
- where $\bar{X}$ is the point estimate

  $Z_{\alpha/2}$ is the normal distribution critical value for a probability of $\alpha/2$ in each tail

  $\frac{\sigma}{\sqrt{n}}$ is the standard error

Finding the Critical Value Z

- Consider a 95% confidence interval: $Z = \pm 1.96$



Common Levels of Confidence

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

Confidence Level	*Confidence Coefficient,* $1-\alpha$	*Z value*
95%	**0.95**	**1.96**
99%	**0.99**	**2.58**

Calculating Confidence Intervals for Means

- **Step 1**: Calculate the mean for the sample

- **Step 2**: Calculate the square root of the variance divided by the sample size

- **Step 3**:Calculate the critical value

- **Step 4**: Apply the formula

Confidence Interval for the Population Mean

Example -1

- A sample of 35 circuits has a mean resistance of 2.20 ohms. We know from past testing that the population standard deviation is 0.35 ohms. Determine a 95% confidence interval for the true mean resistance of the population.

Example -1

- The real estate assessor for Kingston wants to study various characteristics of single-family houses in the parish. A random sample of 70 houses reveals the following:
- Area of the house in square feet: x-bar = 1759, s = 380.
- Construct a 99% confidence interval estimate of the population mean area of the house.

Confidence Intervals for Population Proportion, p

- Upper and lower confidence limits for the population proportion are calculated with the formula

$$\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- where
  - $\hat{p}$ is the sample proportion
  - n is the sample size
  - $Z_{\alpha/2}$ is the normal distribution critical value for a probability of $\alpha/2$ in each tail

Confidence Interval for Proportions

Example 1

- A random sample of 100 people shows that 25 are left-handed. Form a 95% confidence interval for the true proportion of left-handers

Example 2

- The real estate assessor wants to study various characteristics of single-family houses in a town. A random sample of 70 houses reveals the following:
  - 42 houses have central air-conditioning
- Set up a 95% confidence interval estimate of the population proportion of houses that have central air-conditioning

Sampling Error

- The required sample size can be found to reach a desired margin of error (e) with a specified level of confidence ($1 - \alpha$)

- The margin of error is also called sampling error
  - the amount of imprecision in the estimate of the population parameter
  - the amount added and subtracted to the point estimate to form the confidence interval

Determining Sample Size - Mean

Determining Sample Size

For the Mean

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Sampling error (margin of error)

$$e = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Determining Sample Size

(continued)

Determining Sample Size

For the Mean

$$e = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Now solve for n to get

$$n = \left(\frac{Z_{\alpha/2}\,\sigma}{e}\right)^2$$

Example - Sample Size Determination (Mean)

If σ = 45, what sample size is needed to estimate the mean within ± 5 with 90% confidence?

$$n = \left(\frac{Z_{\alpha/2}\,\sigma}{e}\right)^2 = \left(\frac{1.645 \times 45}{5}\right)^2 = 219.19$$

So the required sample size is **n = 220**

(Always round up)

Example – Sample Size Determination (Mean)

- A department store wishes to estimate, with a confidence level of 95% and a maximum error of \$5, the true mean value of purchases per month of its customers. Determine the minimum size of the sample that is required to ensure this, given that the standard deviation is \$15.

- An alumni association wants to estimate the mean debt of this year's university graduates. It is known that the population standard deviation of debts of this year's college graduates is \$11,800. How large a sample should be selected so that the estimate with a 99% confidence level is within \$800 of the population mean?

Determining Sample Size - Proportion

Determining Sample Size

For the Proportion

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

Now solve for n to get

$$n = \frac{Z^2\,\hat{p}(1-\hat{p})}{e^2}$$

Examples – Sample Size Determination - Proportion

- How large a sample would be necessary to estimate the true proportion of defectives in a large population within ±3%, with 95% confidence?

(Assume a pilot sample yields $\hat{p}$ = 0.12)

- A consumer agency wants to estimate the proportion of all drivers who wear seatbelts while driving. Assume that a preliminary study has shown that 76% of drivers wear seatbelts while driving. How large should the sample be so that the 99% confidence interval for the population proportion has a maximum error of 0.03?

- A preliminary sample of 200 parts produced by a new machine showed that 7% of them are defective. How large a sample should the company select so that the 95% confidence interval for *p* is within 0.02 of the population proportion.

Hypothesis Testing

Hypothesis

On a cloudy day, you look out the window and say
" It is going to rain today".

You make a statement forecasting the weather.
You have made a hypothesis regarding the weather.

Now you have to decide whether you should take the umbrella or not?

Statistical Hypothesis is a statement or claim about the parameters of one or more populations.

Statistical Hypotheses

For example, suppose that we are interested in the CGPA of students in a class.

- Now CGPA is a random variable that can be described by a probability distribution.

- Suppose that our interest focuses on the **mean** CGPA (a parameter of this distribution).

- Specifically, we are interested in deciding whether or not the mean CGPA is 8.0.

Statistical Hypotheses

Two-sided Hypotheses

$$H_0: \mu = 8.0$$
$$H_1: \mu \neq 8.0$$

One-sided Hypotheses

$$H_0: \mu \leq 8.0 \qquad \qquad H_0: \mu \geq 8.0$$
$$H_1: \mu > 8.0 \qquad \textit{Or} \qquad H_1: \mu < 8.0$$

Statistical Hypotheses

Two-sided Hypotheses

$$H_0: \mu = 8.0$$
$$H_1: \mu \neq 8.0$$

One-sided Hypotheses

$$H_0: \mu \leq 8.0$$
$$H_1: \mu > 8.0$$

Or

$$H_0: \mu \geq 8.0$$
$$H_1: \mu < 8.0$$

Null & Alternate Hypotheses

Null Hypothesis is a hypothesis of no difference

E.g. The mean CGPA of juniors is at least 8.0 **(H_0: $\mu \geq 8.0$)**

It states the assumption (numerical) to be tested.

Hypothesis testing begins with the assumption that the null hypothesis is TRUE.

- Similar to the notion of **innocent until proven guilty**
- Refers to the '**status quo**'
- Always contains the ' **=** ' sign
- Null Hypothesis **may or may not be rejected.**

Null & Alternate Hypotheses

Alternate Hypothesis is the opposite of the null hypothesis

E.g.: The mean CGPA of juniors is less than 8.0 (H_1: μ<8.0)

- Challenges the Status Quo
- Never contains the '=' sign
- May or may not be accepted

Alternate hypothesis is generally the hypothesis that is believed to be true by the researcher
(Also called 'research hypothesis')

Either hypothesis – the null or alternative – may represent the original claim.

Hypothesis Testing

- Hypothesis testing is be used to determine whether a statement about the value of a population parameter should or should not be rejected.
- The null hypothesis*,* denoted by H_0 *,* is a tentative assumption about a population parameter.
- The alternative hypothesis, denoted by H_1, is the opposite of what is stated in the null hypothesis.
- The hypothesis testing procedure uses data from a sample to test the two competing statements indicated by H_0 and H_1.
- If sample information is *consistent* with H_0, then we will conclude that the null hypothesis is **true**; Else we will conclude that the hypothesis is **false**.

Developing Null and Alternative Hypotheses

- It is not always obvious how the null and alternative hypotheses should be formulated.
- Care must be taken to structure the hypotheses appropriately so that the test conclusion provides the information the researcher wants.
- The context of the situation is very important in determining how the hypotheses should be stated.
- In some cases it is easier to identify the alternative hypothesis first. In other cases the null is easier.
- Correct hypothesis formulation will take practice.

You make errors

No matter which hypothesis represents the claim, you always test the null hypothesis.

At the end of the test,

you make one of the two decisions.

Reject the null hypothesis

or

Do not reject the null hypothesis.

But …

B'coz your decision is based on incomplete information (a sample rather than the entire population), there is always the possibility that **you will make a wrong decision !!**

The judge errs…

If a person is really innocent,
but the jury decides (s)he is guilty.

Rejecting the null hypothesis when it is true
– **Type 1 error** *(also called rejection error)*

If a person is really guilty,
but the jury finds him/her not guilty
– a criminal is walking free on the streets.

Not rejecting null hypothesis when it is false
– **Type 2 error** *(also called acceptance error)*

Not our fault. It's by chance.

Often we don't know if an error has been made.
And therefore we can only talk about
the probability of making an error.

P(Type 1 error) = α, called 'level of significance' (LoS)
P(Type 2 error) = β, and $1-\beta$ is called 'power' of the test

If $\alpha = 0.05$, we are using 5% LoS.

There is at most 5% chance of rejecting a null hypothesis.

In another way our confidence level is $1 - \alpha = 0.95$.

We are at least 95% confident that the test results in correct decision.

Decisions in Hypothesis Testing

Jury Trial

Actual State

Verdict

	Innocent	*Guilty*
Innocent	Correct decision	Type 2 error
Guilty	Type 1 error	Correct decision

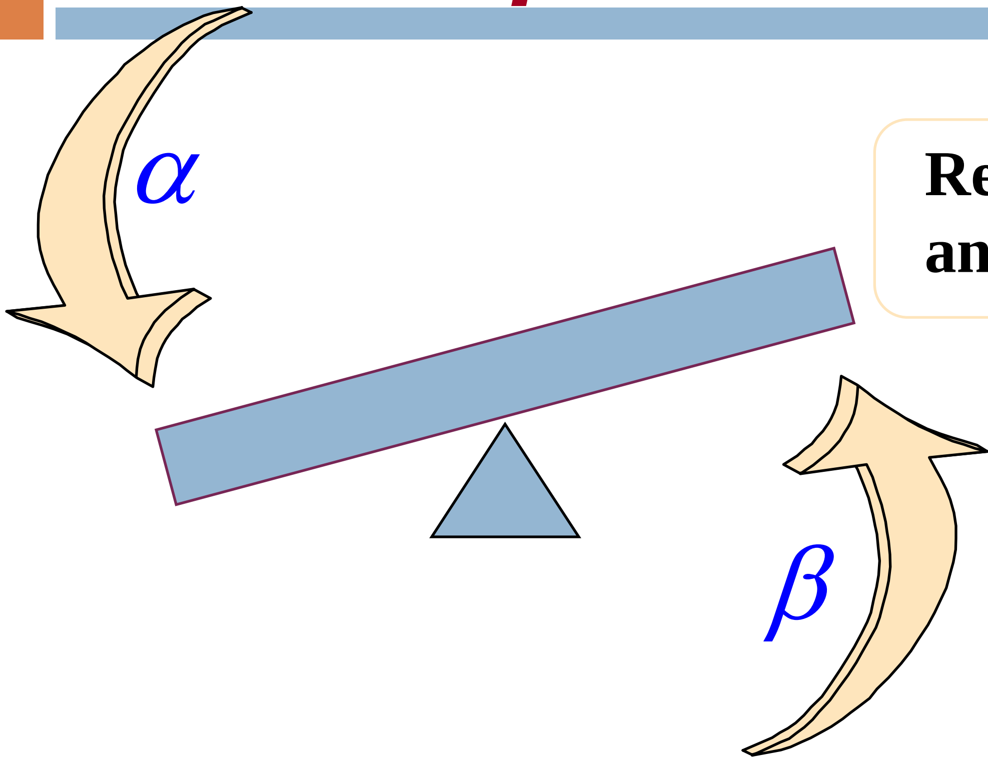
Hypothesis Testing

Actual State

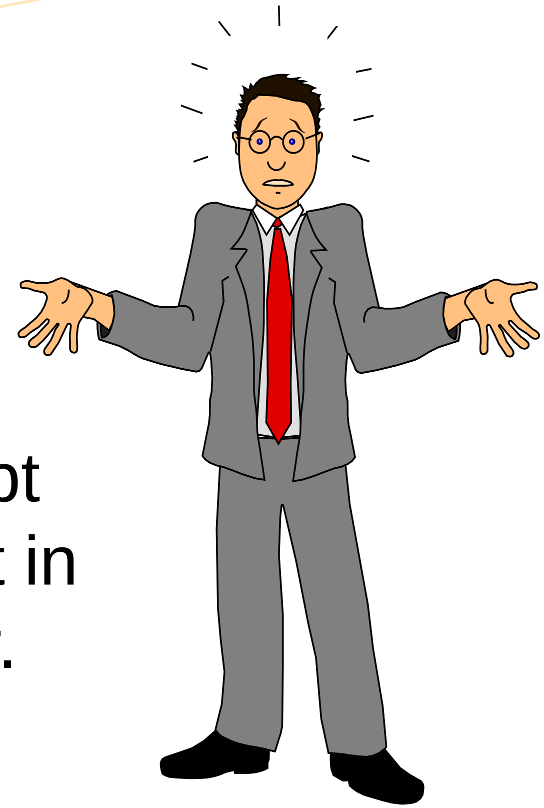
Decision from Sample

	H_0 is true	*H_0 is false*
Fail to reject H_0	Correct decision (1-α)	Type 2 error (β)
Reject H_0	Type 1 error (α)	Correct decision (1-β)

α & β - Inverse Relationship



Reduce probability of one error and the other one goes up.



For a given sample size n, an attempt to reduce one type of error will result in an increase in the other type of error.

Is it possible to use LOS $\alpha = 0$?

Which error is more serious?

It depends on the nature of the problem

The legal system assumes that more harm is done by convicting the innocent, than by not convicting the guilty.

Here Type 1 error is more serious than Type 2 error.

In some cases, even the Type 2 error is dangerous.

Type 1 error is called Producer's risk

– *rejecting a lot when it is actually good.*

Type 2 error is called Consumer's risk

– *accepting a lot when it is actually not good.*

How to choose between errors?

- Choice depends on the cost of the error
- Choose little type I error when the cost of rejecting the maintained hypothesis is high

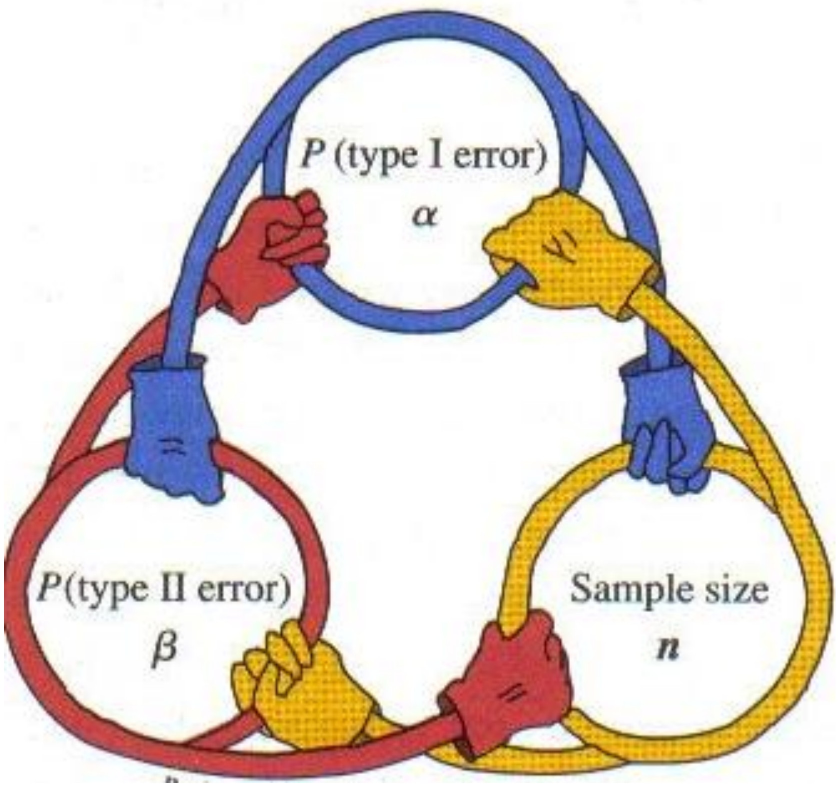
A criminal trial: Convicting an innocent person

- Choose large type I error when you have an interest in changing the status quo.

A decision in a startup company about a new piece of software

It is possible to reduce both the errors if we increase the sample size, but increasing the sample size may not be possible.

The hypothesis test is guaranteed with ‘low’ probability of rejecting the null hypothesis wrongly. i.e., by fixing type 1 error, we try to maximize the power or minimize type 2 error.



Level of Significance (LoS) α

- The significance level, a, is a probability used as a criterion for rejecting the null hypothesis.

  Choices for α are typically: α=.01, α = .05 or α =.10

  α is just the area in the tails of the distribution

  Selected by the Researcher at the beginning of the test.

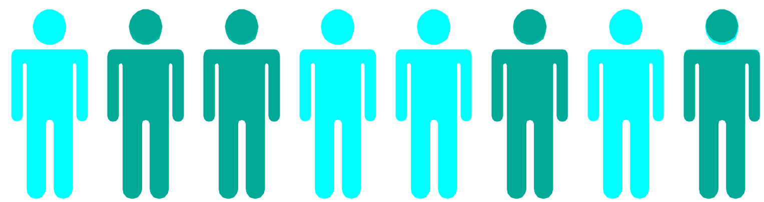
Critical Region or Rejection Region

- Set of possible values of the test statistic for which H_0 *will be rejected* is called Rejection region.

  The rest of the sample space is called Acceptance region.
- The critical or rejection region is the range of values that indicates a significant difference between the sample data and the null hypothesis parameter
- The remaining region is the *non-critical* region which indicates a difference due to chance- we "*fail to reject the null hypothesis*"

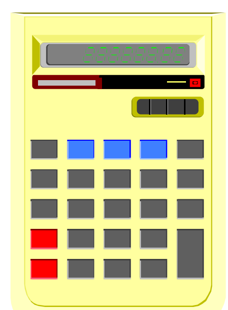
Hypothesis Testing Process

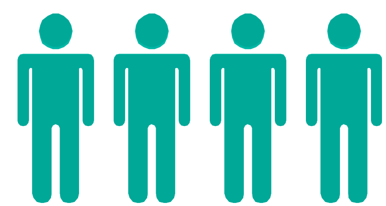
Assume the population mean age is 50. (Null Hypothesis)



Population

The Sample Mean Is 20





Sample

$$Is\ \ \overline{X} = 20 \cong \ \mu = 50\ ?$$

No, not likely!

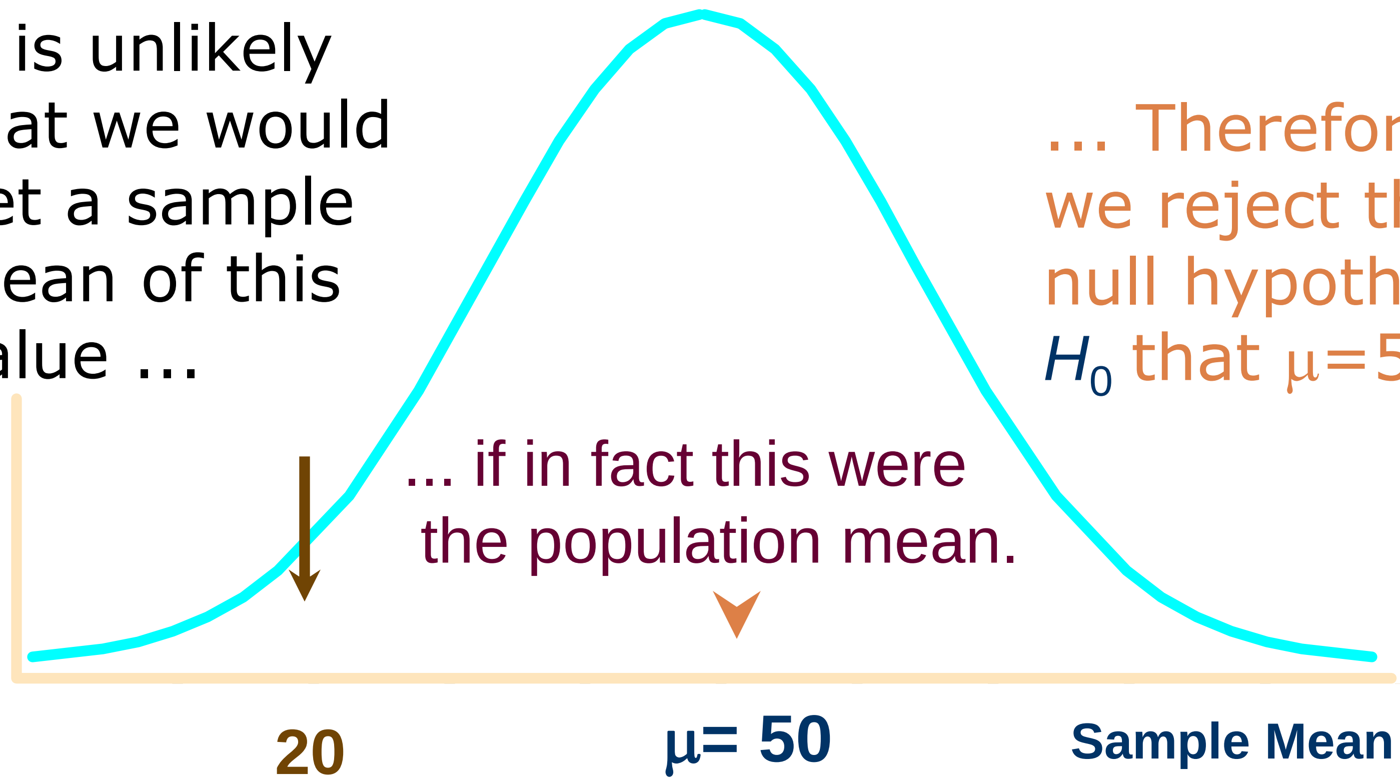
REJECT

Null Hypothesis

Reason for Rejecting H_0

Sampling Distribution

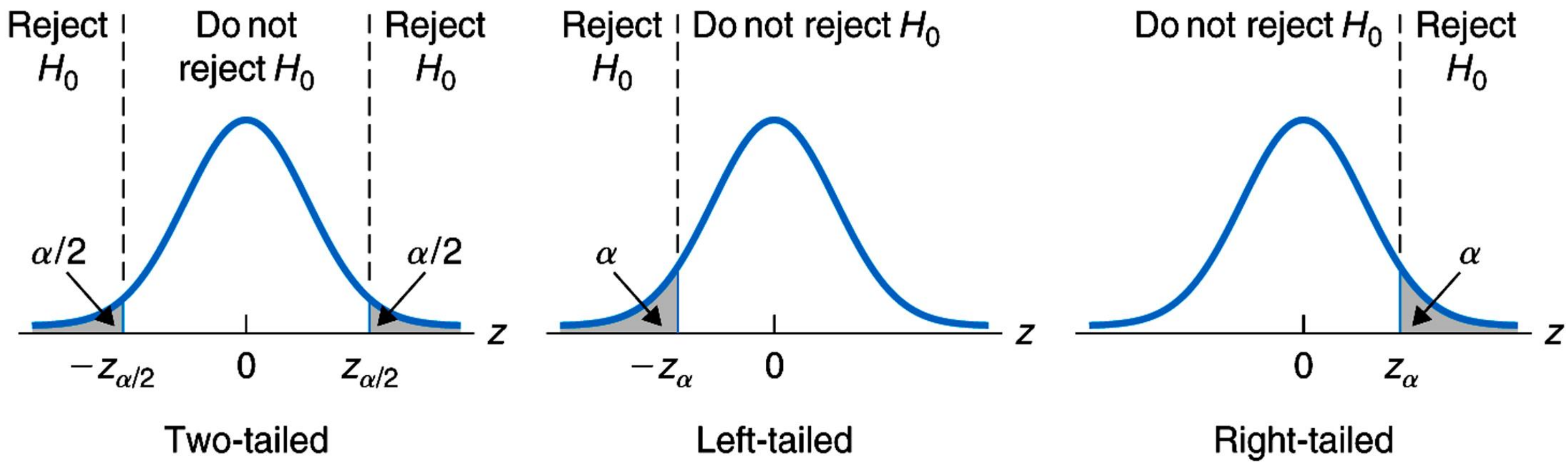
It is unlikely that we would get a sample mean of this value ...



... Therefore, we reject the null hypothesis H_0 that $\mu=50$.

Critical value

- the cutoff value(s) that separates the rejection region from non rejection region
- the nature of value depends on whether the test is right-, left-, or two-tailed.
- the area of rejection is equal to LOS.



In two-tailed test, the LOS must be divided equally among the two tails. Hence, it corresponds to $\alpha/2$

LoS α and the Rejection Region

H_0: $\mu \geq 8.0$

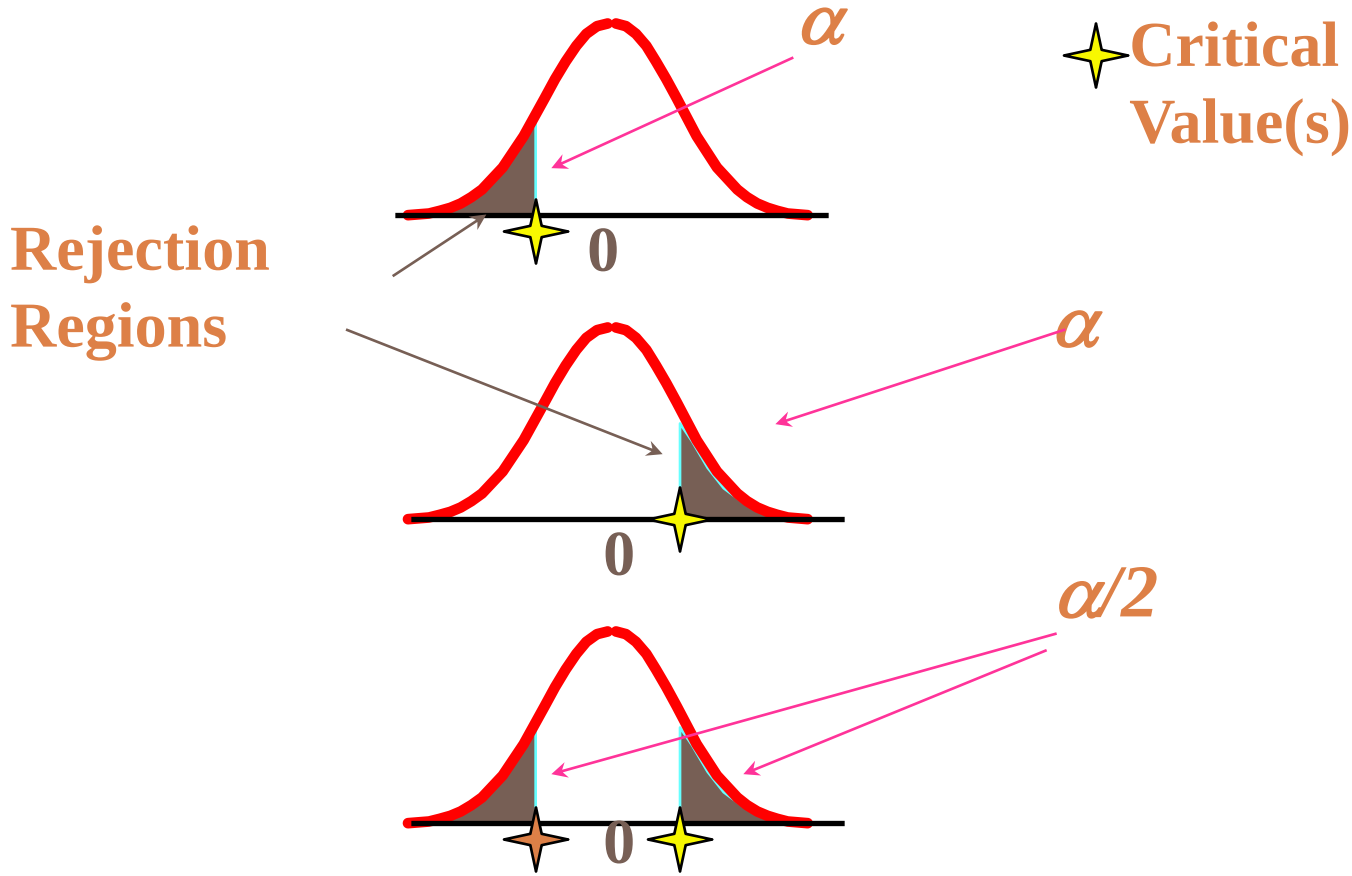
H_1: $\mu < 8.0$

H_0: $\mu \leq 8.0$

H_1: $\mu > 8.0$

H_0: $\mu = 8.0$

H_1: $\mu \neq 8.0$



Critical values for z-test:

	LOS 1%	LOS 5%
One-tailed	2.33	1.645
Two-tailed	2.58	1.96

Critical values for other tests: Refer t, F, chi-square tables

Test Statistic

N = Population size

μ = Population mean

σ = Population standard deviation

p = Population proportion

n = Sample size

$\overline{\mathrm{X}}$ = Sample mean

s = Sample standard deviation

$\hat{\mathrm{p}}$ = Sample proportion

General formula for Test Statistic:

$$Z = \frac{t - E(t)}{SE}$$

where

t – Statistic

E(t) – expected value of t

S.E – Standard error

Developing Null and Alternative Hypotheses

- Alternative Hypothesis as a Research Hypothesis

  - Example:
    A new teaching method is developed that is believed to be better than the current method.

  - Alternative Hypothesis:
    The new teaching method is better.

  - Null Hypothesis:
    The new method is no better than the old method.

Developing Null and Alternative Hypotheses

- Alternative Hypothesis as a Research Hypothesis

  - Example:
    A new sales force bonus plan is developed in an attempt to increase sales.

  - Alternative Hypothesis:
    The new bonus plan increase sales.

  - Null Hypothesis:
    The new bonus plan does not increase sales.

Developing Null and Alternative Hypotheses

- Alternative Hypothesis as a Research Hypothesis

- Example:
  A new drug is developed with the goal of lowering blood pressure more than the existing drug.

- Alternative Hypothesis:
  The new drug lowers blood pressure more than the existing drug.

- Null Hypothesis:
  The new drug does not lower blood pressure more than the existing drug.

Developing Null and Alternative Hypotheses

- Null Hypothesis as an Assumption to be Challenged
  - Example:
    The label on a soft drink bottle states that it contains 200 ml.
  - Null Hypothesis:
    The label is correct. $\mu \geq 200$ ml.
  - Alternative Hypothesis:
    The label is incorrect. $\mu < 200$ ml.

p-Value Approach to One-Tailed Hypothesis Testing

- The $\underline{p\text{-value}}$ is the probability, computed using the test statistic, that measures the support (or lack of support) provided by the sample for the null hypothesis.
- A p- value is the probability of seeing the effect(E) when the null hypothesis is true. p-value=P[E | H_0]. That is, the probability of getting the observed data, or more extreme than this data, if the null hypothesis were true.
- If the p-value is less than or equal to the level of significance α, the value of the test statistic is in the rejection region.
- Reject H_0 if the p-value $\leq \alpha$.

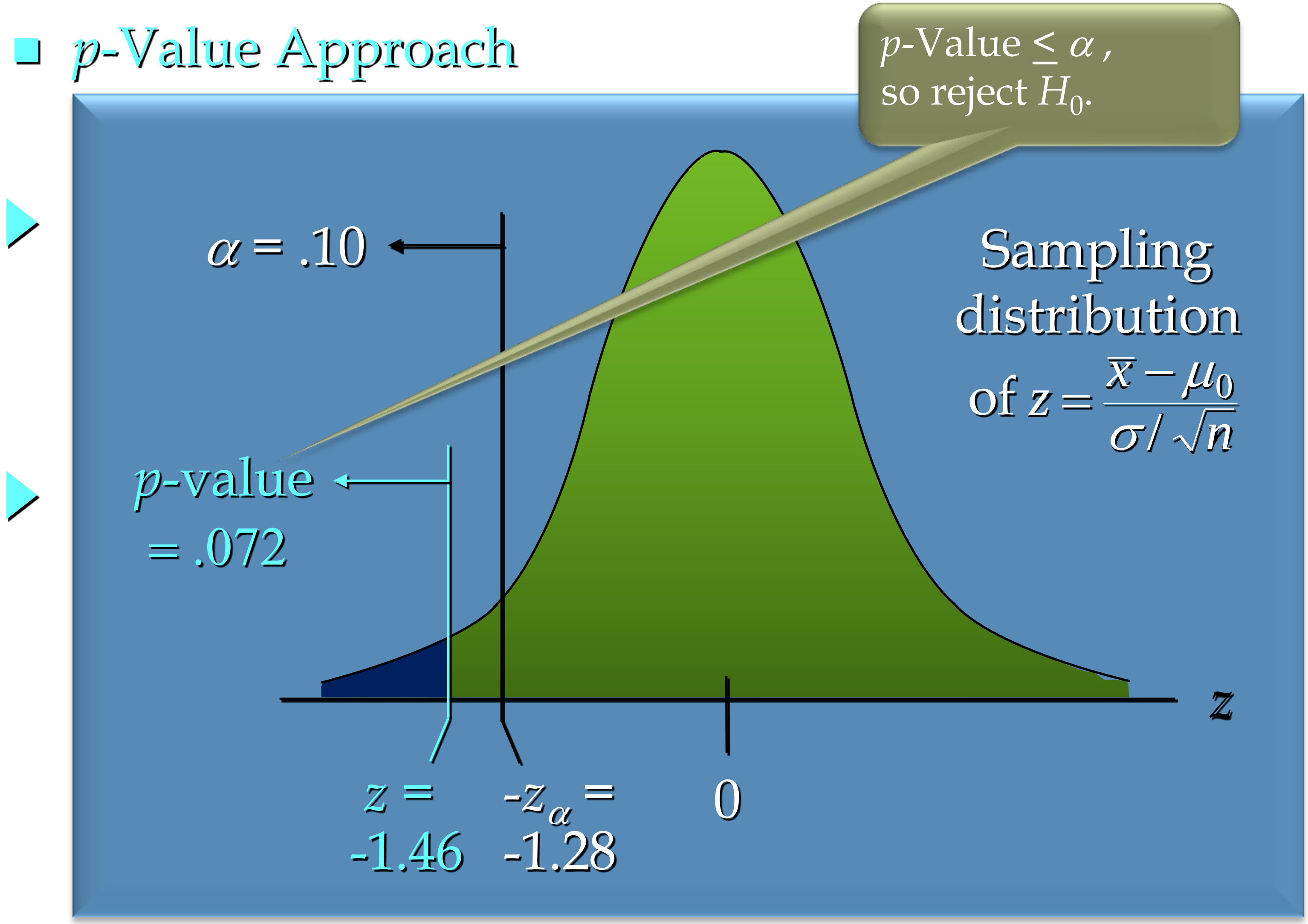
Suggested Guidelines for Interpreting *p*-Values

p - values for $\alpha = 0.1$

- Less than .01
  Overwhelming evidence to conclude H_a is true.
- Between .01 and .05
  Strong evidence to conclude H_a is true.
- Between .05 and .10
  Weak evidence to conclude H_a is true.
- Greater than .10
  Insufficient evidence to conclude H_a is true.

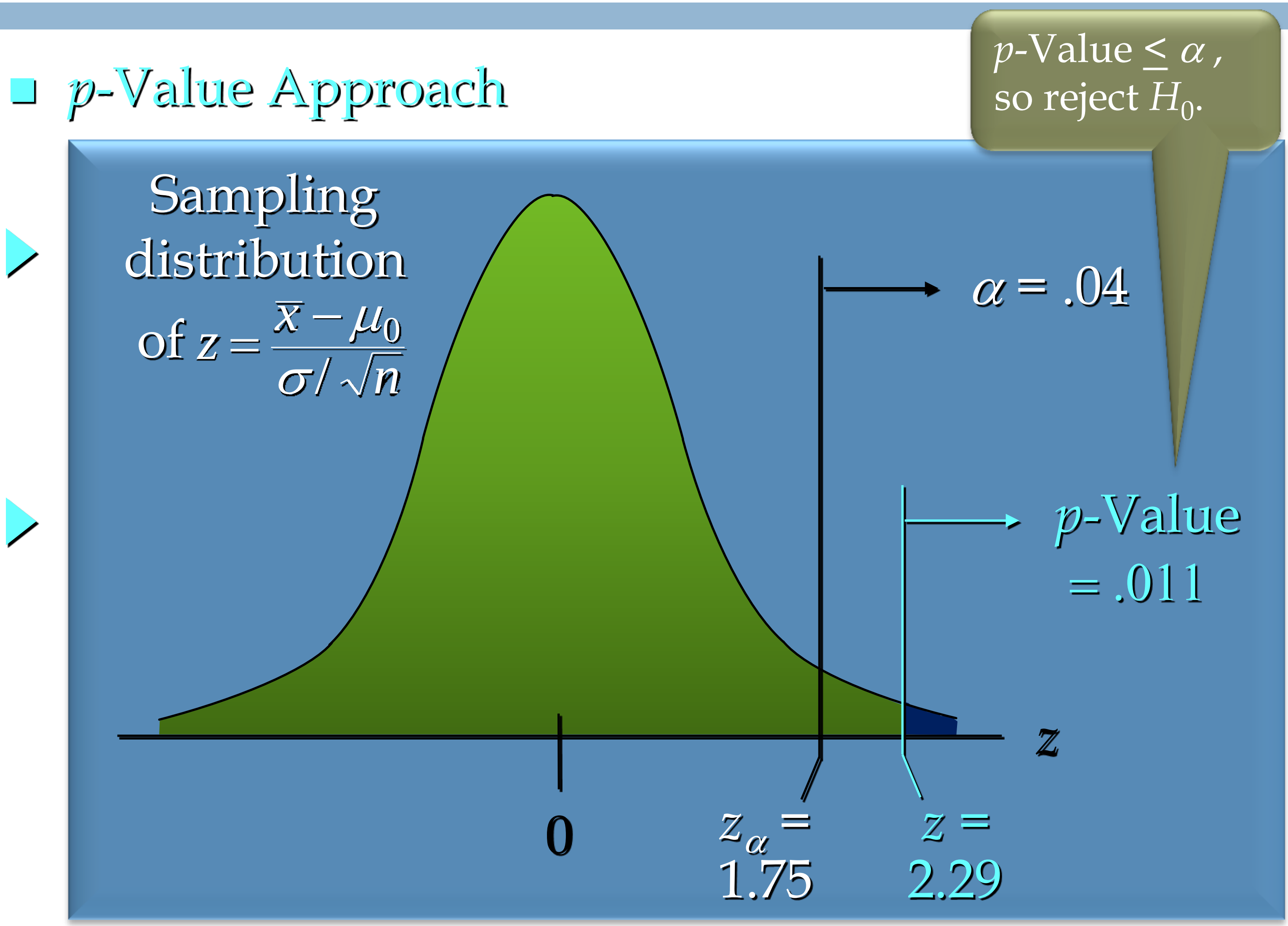
Lower-Tailed Test About a Population Mean: σ Known

- *p*-Value Approach



Upper-Tailed Test About a Population Mean: σ Known

- *p*-Value Approach



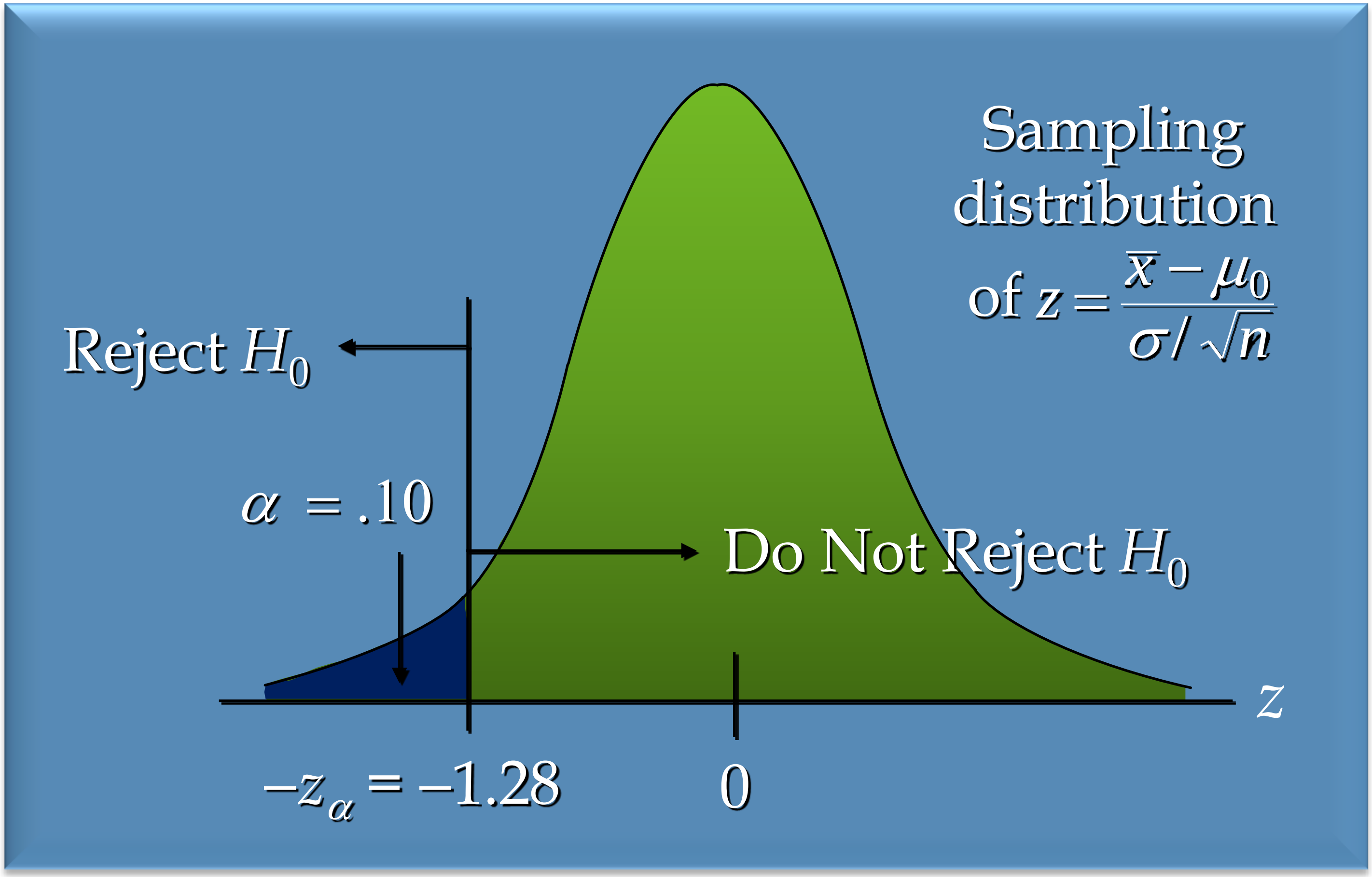
Critical Value Approach

One-Tailed Hypothesis Testing

- The test statistic z has a standard normal probability distribution.
- We can use the standard normal probability distribution table to find the z-value with an area of α in the lower (or upper) tail of the distribution.
- The value of the test statistic that established the boundary of the rejection region is called the <u>critical value</u> for the test.
- The rejection rule is:
  - Lower tail: Reject H_0 if $z \leq -z_\alpha$
  - Upper tail: Reject H_0 if $z \geq z_\alpha$

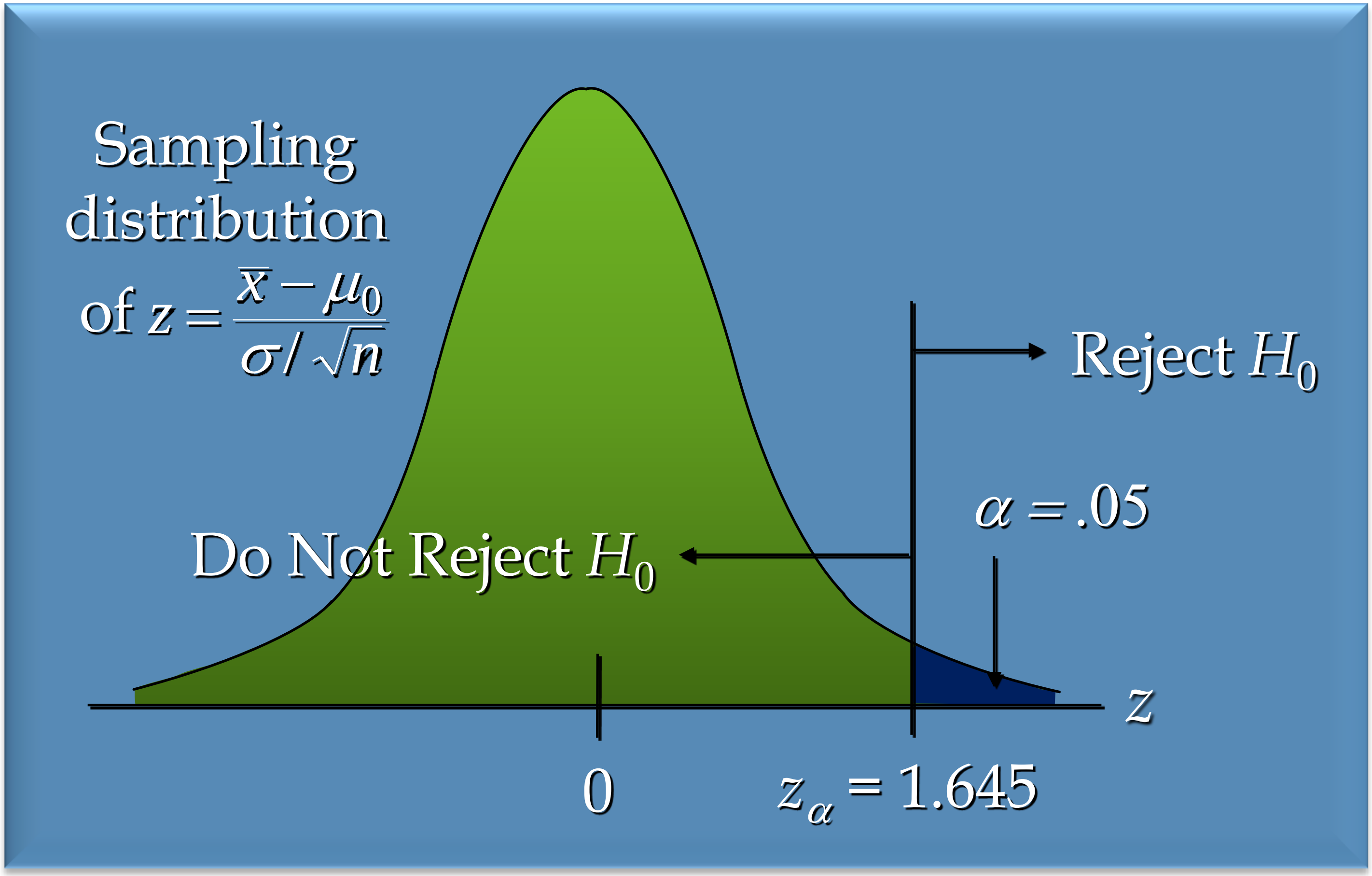
Lower-Tailed Test About a Population Mean: σ Known

- Critical Value Approach



Upper-Tailed Test About a Population Mean: σ Known

- **Critical Value Approach**



Steps of Hypothesis Testing

- Step 1. Develop the null and alternative hypotheses.
- Step 2. Specify the level of significance α.
- Step 3. Collect the sample data and compute the value of the test statistic.

p-Value Approach

- Step 4. Use the value of the test statistic to compute the *p*-value.
- Step 5. Reject H_0 if *p*-value $\leq \alpha$.

Critical Value Approach

- Step 4. Use the level of significance to determine the critical value and the rejection rule.
- Step 5. Use the value of the test statistic and the rejection rule to determine whether to reject H_0.

One-Tailed Tests About a Population Mean: σ Known

- Example: Metro EMS

  - The response times for a random sample of 40 medical emergencies were tabulated. The sample mean is 13.25 minutes. The population standard deviation is believed to be 3.2 minutes.

  - The EMS director wants to perform a hypothesis test, with a .05 level of significance, to determine whether the service goal of 12 minutes or less is being achieved.

One-Tailed Tests About a Population Mean: σ Known

- p -Value and Critical Value Approaches

1. Develop the hypotheses.

$$H_0:\ \mu \le 12$$
$$H_a:\ \mu > 12$$

2. Specify the level of significance.

$$\alpha = .05$$

3. Compute the value of the test statistic.

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{13.25 - 12}{3.2 / \sqrt{40}} = 2.47$$

One-Tailed Tests About a Population Mean: σ Known

- p –Value Approach

4. Compute the p –value.

For $z = 2.47$, cumulative probability = .9932.

p–value = 1 – .9932 = .0068

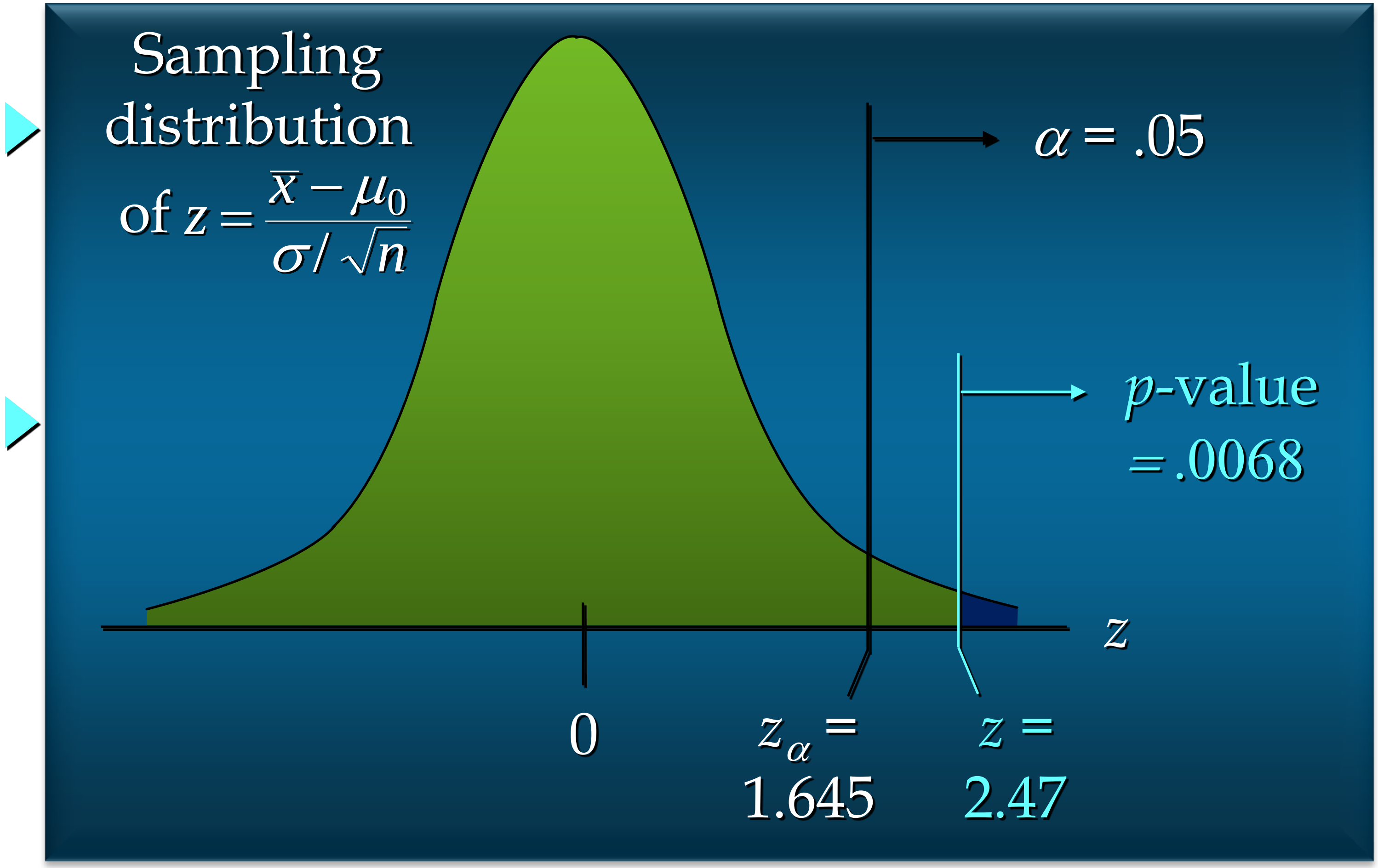
5. Determine whether to reject H_0.

Because p–value = .0068 $\leq \alpha = .05$, we reject H_0.

There is sufficient statistical evidence to infer that Metro EMS is <u>not</u> meeting the response goal of 12 minutes.

One-Tailed Tests About a Population Mean: σ Known

- p –Value Approach



One-Tailed Tests About a Population Mean: σ Known

- Critical Value Approach

4. Determine the critical value and rejection rule.

For $\alpha = .05$, $z_{.05} = 1.645$

Reject H_0 if $z \geq 1.645$

5. Determine whether to reject H_0.

Because $2.47 \geq 1.645$, we reject H_0.

There is sufficient statistical evidence to infer that Metro EMS is <u>not</u> meeting the response goal of 12 minutes.

p-Value Approach to Two-Tailed Hypothesis Testing

- Compute the p-value using the following three steps:
  1. Compute the value of the test statistic z.
  2. If z is in the upper tail ($z > 0$), compute the probability that z is greater than or equal to the value of the test statistic. If z is in the lower tail ($z < 0$), compute the probability that z is less than or equal to the value of the test statistic.
  3. Double the tail area obtained in step 2 to obtain the p –value.
- The rejection rule:

$$\text{Reject } H_0 \text{ if the } p\text{-value} \leq \alpha.$$

Critical Value Approach to Two-Tailed Hypothesis Testing

- The critical values will occur in both the lower and upper tails of the standard normal curve.
- Use the standard normal probability distribution table to find $z_{\alpha/2}$ (the z-value with an area of $\alpha/2$ in the upper tail of the distribution).
- The rejection rule is:

$$\text{Reject } H_0 \text{ if } z \leq -z_{\alpha/2} \text{ or } z \geq z_{\alpha/2}.$$

Two-Tailed Tests About a Population Mean: σ Known

Example: Glow Toothpaste

- The production line for Glow toothpaste is designed to fill tubes with a mean weight of 6 oz. Periodically, a sample of 30 tubes will be selected in order to check the filling process.
- Quality assurance procedures call for the continuation of the filling process if the sample results are consistent with the assumption that the mean filling weight for the population of toothpaste tubes is 6 oz.; otherwise the process will be adjusted.
- Assume that a sample of 30 toothpaste tubes provides a sample mean of 6.1 oz. The population standard deviation is believed to be 0.2 oz.
- Perform a hypothesis test, at the .03 level of significance, to help determine whether the filling process should continue operating or be stopped and corrected.

Two-Tailed Tests About a Population Mean: σ Known

- *p* –Value and Critical Value Approaches

1. Determine the hypotheses.

H_0: $\mu = 6$
H_a: $\mu \neq 6$

2. Specify the level of significance.

α = .03

3. Compute the value of the test statistic.

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{6.1 - 6}{.2 / \sqrt{30}} = 2.74$$

Two-Tailed Tests About a Population Mean: σ Known

- *p* –Value Approach

4. Compute the *p* –value.

For $z = 2.74$, cumulative probability = .9969

p–value = 2(1 – .9969) = .0062

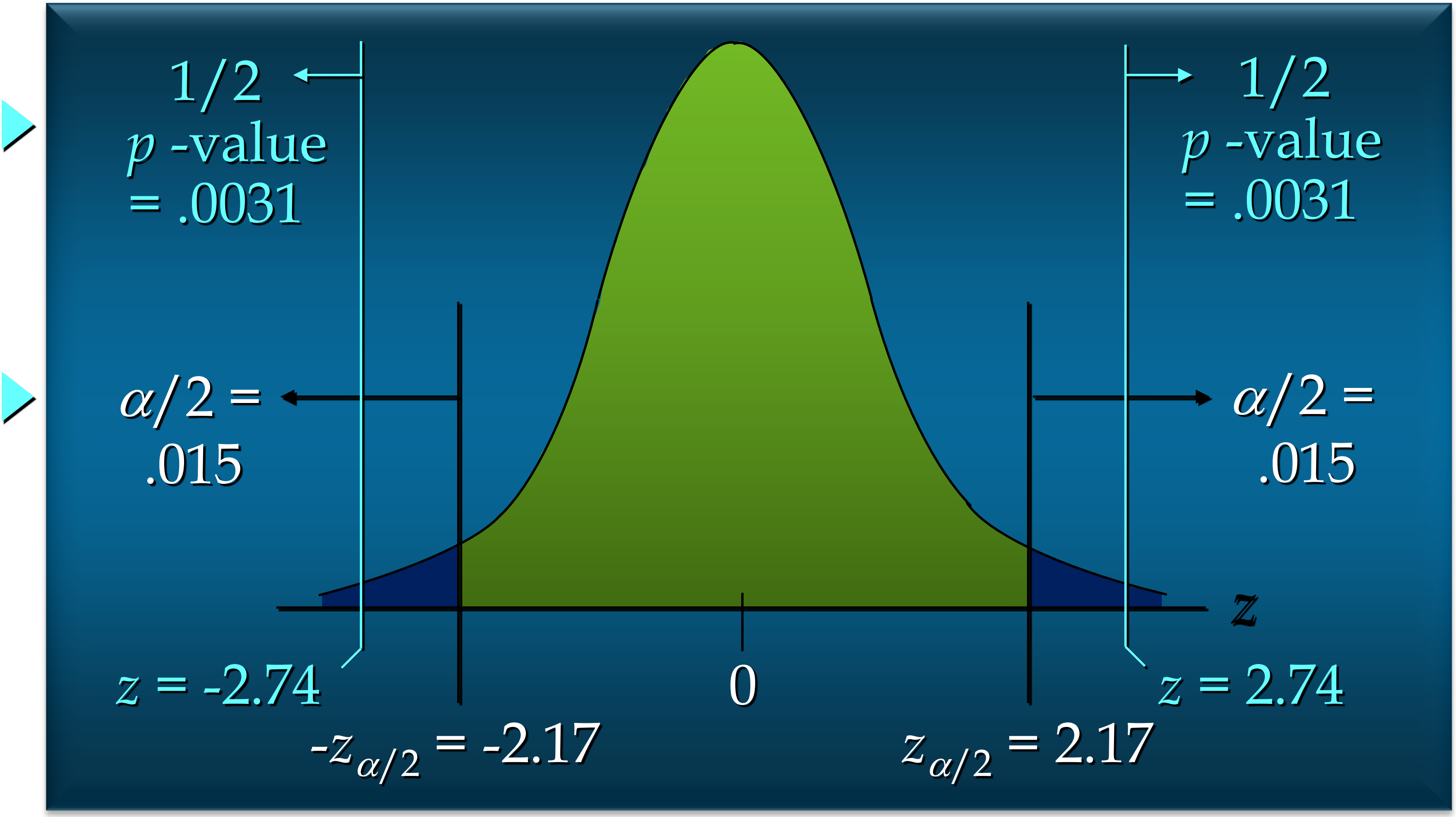
5. Determine whether to reject H_0.

Because p–value = .0062 $\leq \alpha$ = .03, we reject H_0.

There is sufficient statistical evidence to infer that the alternative hypothesis is true (i.e. the mean filling weight is not 6 ounces).

Two-Tailed Tests About a Population Mean: σ Known

- p-Value Approach



Two-Tailed Tests About a Population Mean: σ Known

- Critical Value Approach

4. Determine the critical value and rejection rule.

For $\alpha/2 = .03/2 = .015$, $z_{.015} = 2.17$

Reject H_0 if $z \leq -2.17$ or $z \geq 2.17$

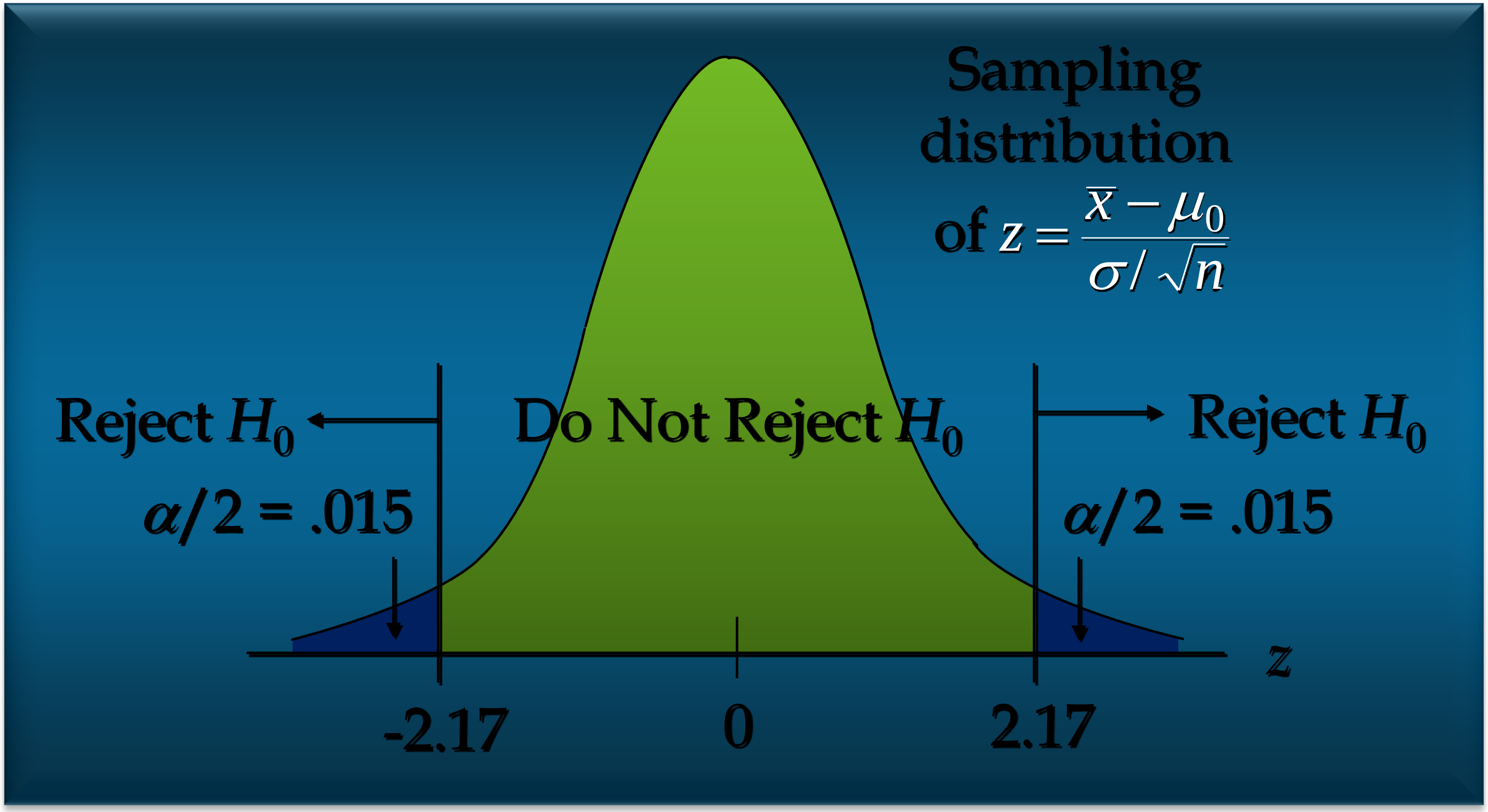
5. Determine whether to reject H_0.

Because $2.74 \geq 2.17$, we reject H_0.

There is sufficient statistical evidence to infer that the alternative hypothesis is true (i.e. the mean filling weight is not 6 ounces).

Two-Tailed Tests About a Population Mean: σ Known

- Critical Value Approach



One-Tailed Test About a Population Mean: σ Unknown

- Test Statistic

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

This test statistic has a *t* distribution with *n* - 1 degrees of freedom.

Tests About a Population Mean: σ Unknown

- Rejection Rule: p -Value Approach

  Reject H_0 if p –value $\leq \alpha$

- Rejection Rule: Critical Value Approach

  - H_0: $\mu \geq \mu_0$ Reject H_0 if $t \leq -t_\alpha$
  - H_0: $\mu \leq \mu_0$ Reject H_0 if $t \geq t_\alpha$
  - H_0: $\mu = \mu_0$ Reject H_0 if $t \leq - t_{\alpha/2}$ or $t \geq t_{\alpha/2}$

p -Values and the t Distribution

- The format of the t distribution table provided in most statistics textbooks does not have sufficient detail to determine the exact p-value for a hypothesis test.
- However, we can still use the t distribution table to identify a range for the p-value.
- An advantage of computer software packages is that the computer output will provide the p-value for the t distribution.

One-Tailed Test About a Population Mean: σ Unknown

Example: Highway Patrol

- A State Highway Patrol periodically samples vehicle speeds at various locations on a particular roadway. The sample of vehicle speeds is used to test the hypothesis H_0: $\mu \leq 65$.
- The locations where H_0 is rejected are deemed the best locations for radar traps. At Location F, a sample of 64 vehicles shows a mean speed of 66.2 mph with a standard deviation of 4.2 mph. Use $\alpha = .05$ to test the hypothesis.

One-Tailed Test About a Population Mean: σ Unknown

- p –Value and Critical Value Approaches

1. Determine the hypotheses.

H_0: $\mu \leq 65$
H_a: $\mu > 65$

2. Specify the level of significance.

$\alpha = .05$

3. Compute the value of the test statistic.

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{66.2 - 65}{4.2/\sqrt{64}} = 2.286$$

One-Tailed Test About a Population Mean: σ Unknown

- p –Value Approach

4. Compute the p –value.

For t = 2.286, the p–value must be less than .025 (for t = 1.998) and greater than .01 (for t = 2.387).

.01 < p–value < .025

5. Determine whether to reject H_0.

Because p–value $\leq \alpha$ = .05, we reject H_0.

We are at least 95% confident that the mean speed of vehicles at Location F is greater than 65 mph.

One-Tailed Test About a Population Mean: σ Unknown

- Critical Value Approach

4. Determine the critical value and rejection rule.

For $\alpha = .05$ and d.f. = 64 – 1 = 63, $t_{.05} = 1.669$

Reject H_0 if $t \geq 1.669$

5. Determine whether to reject H_0.

Because 2.286 $\geq$ 1.669, we reject H_0.

We are at least 95% confident that the mean speed of vehicles at Location F is greater than 65 mph. Location F is a good candidate for a radar trap.

One-Tailed Test About a Population Mean: σ Unknown