Hypothesis Testing

Hypothesis testing Part 1

- Hypothesis testing is a common statistical tool used in research and data science to support the certainty of findings.
- A hypothesis is often described as an “educated guess” about a specific parameter or population. Once it is defined, one can collect data to determine whether it provides enough evidence that the hypothesis is true.

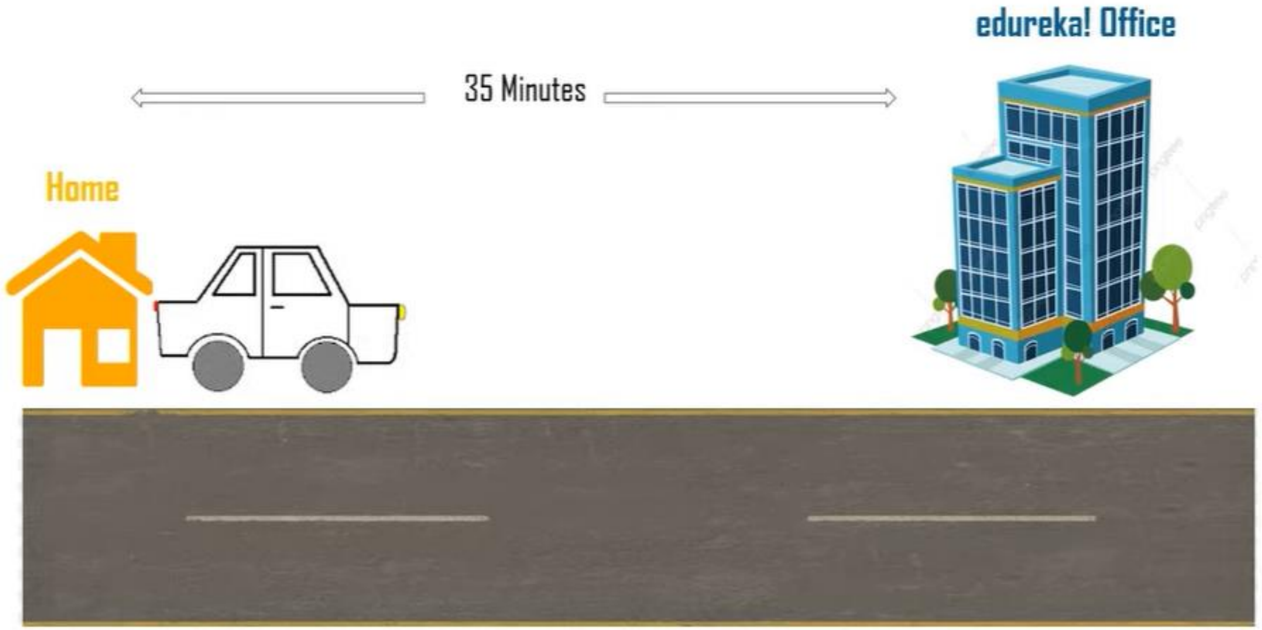
Introduction of Hypothesis Testing

What is a Hypothesis?

Lead content <= 2.5 PPM

Sample = 10 Thousand

Introduction of Hypothesis Testing



Hypothesis testing is an act in statistics whereby an analyst tests an assumption regarding a population parameter. The methodology employed by the analyst depends on the nature of the data used and the reason for the analysis.

Null & Alternate Hypotheses

Null & Alternate Hypotheses

(H_0) Null Hypothesis:

- Prevailing belief about the population
- Assumes that status quo is true

(H_1) Alternate Hypothesis

- Claim that opposes the null hypothesis

Null & Alternate Hypotheses

Outcome of Hypothesis Testing:

Null Hypothesis: H_0 – Accused is Innocent

Alternate Hypothesis: H_1 – Accused is not Innocent

Rejection of Null Hypothesis - Guilty

Failure to reject Null Hypothesis – Not Guilty

Examples

- *Null Hypothesis:* H_0: There is no difference in the salary of factory workers based on gender.
  Alternative Hypothesis: H_a: Male factory workers have a higher salary than female factory workers.
- *Null Hypothesis*: H_0: There is no relationship between height and shoe size.
  Alternative Hypothesis: H_a: There is a positive relationship between height and shoe size.
- *Null Hypothesis*: H_0: Experience on the job has no impact on the quality of a brick mason's work.
  Alternative Hypothesis: H_a: The quality of a brick mason's work is influenced by on-the-job experience.

Null & Alternate Hypotheses

Formulation of Null & Alternate Hypothesis:

If your claim statement has words like "at least", "at most", "less than", or "greater than", you cannot formulate the null hypothesis just from the claim statement

Rule to formulate the null and alternate hypotheses:

- Null hypothesis signs: = OR ≤ OR ≥
- Alternate hypothesis signs: ≠ OR > OR <

Situation 1:

Equal Sign

Total Valuation >= $20 Billion → Null Hypothesis

Total Valuation < $20 Billion → Alternate Hypothesis

No Equal Sign

Situation 2:

No Equal Sign

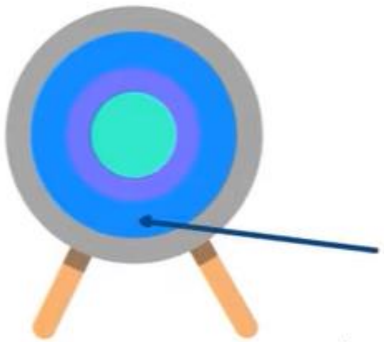
Total Valuation > $20 Billion → Alternate Hypothesis

Total Valuation <= $20 Billion → Null Hypothesis

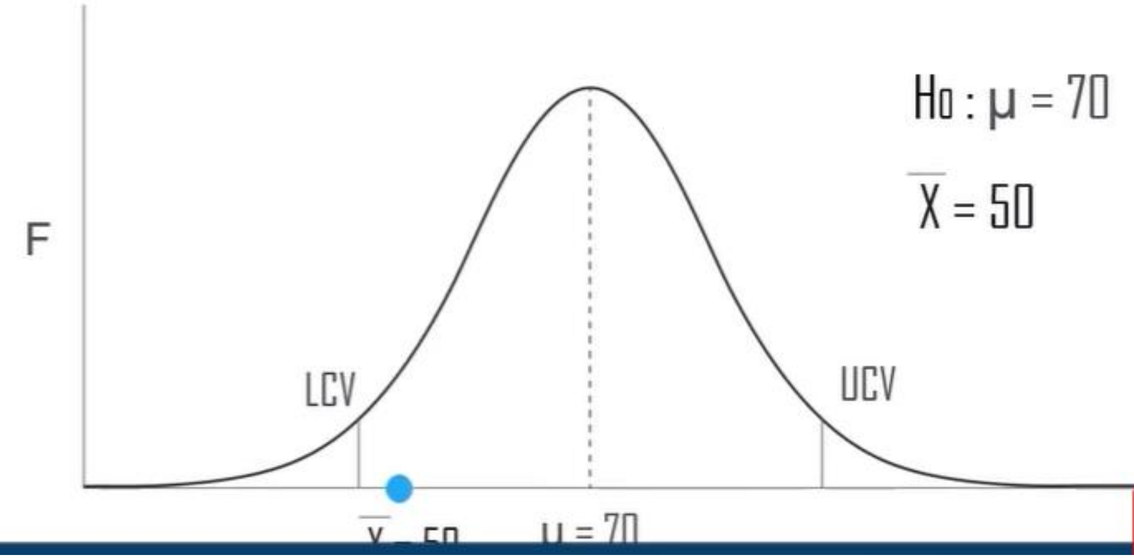
Equal Sign

Critical Value Method

Aviral's Claim: Average Score in archery = 70



Average Archery Score for 5 Games

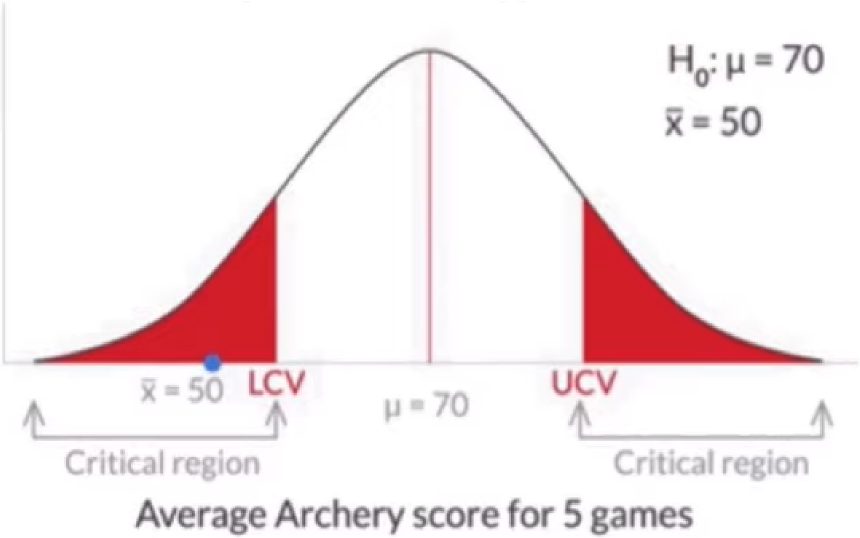


Over 5 Games of Archery:

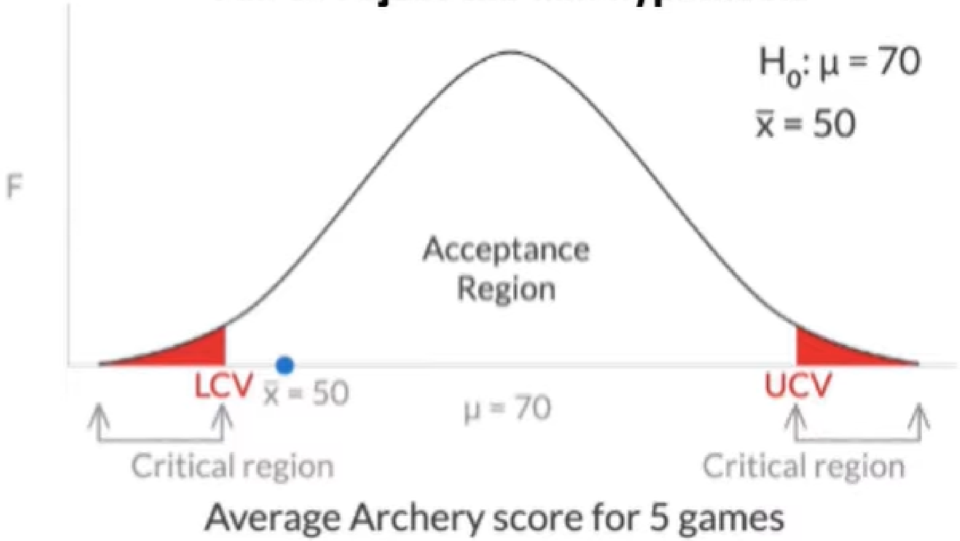
- Average Score = 20 → Less likely to believe his claim
- Average Score = 65 → More likely to believe his claim

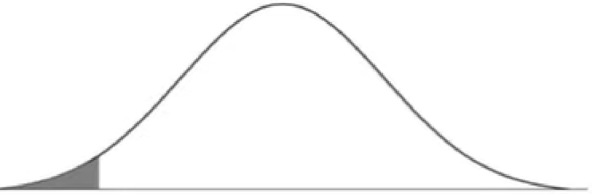
Critical Value Method

Reject the null hypothesis



Fail to reject the null hypothesis

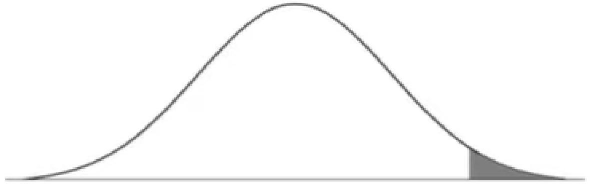




$\neq$ in H_1 $\longrightarrow$ Two-tailed test $\longrightarrow$ Rejection region on both sides of distribution

$<$ in H_1 $\longrightarrow$ Lower-tailed test $\longrightarrow$ Rejection region on left side of distribution

$>$ in H_1 $\longrightarrow$ Upper-tailed test $\longrightarrow$ Rejection region on right side of distribution



Critical Value Method

Distribution of Average AC Sales Data for 36 Stores

H_0

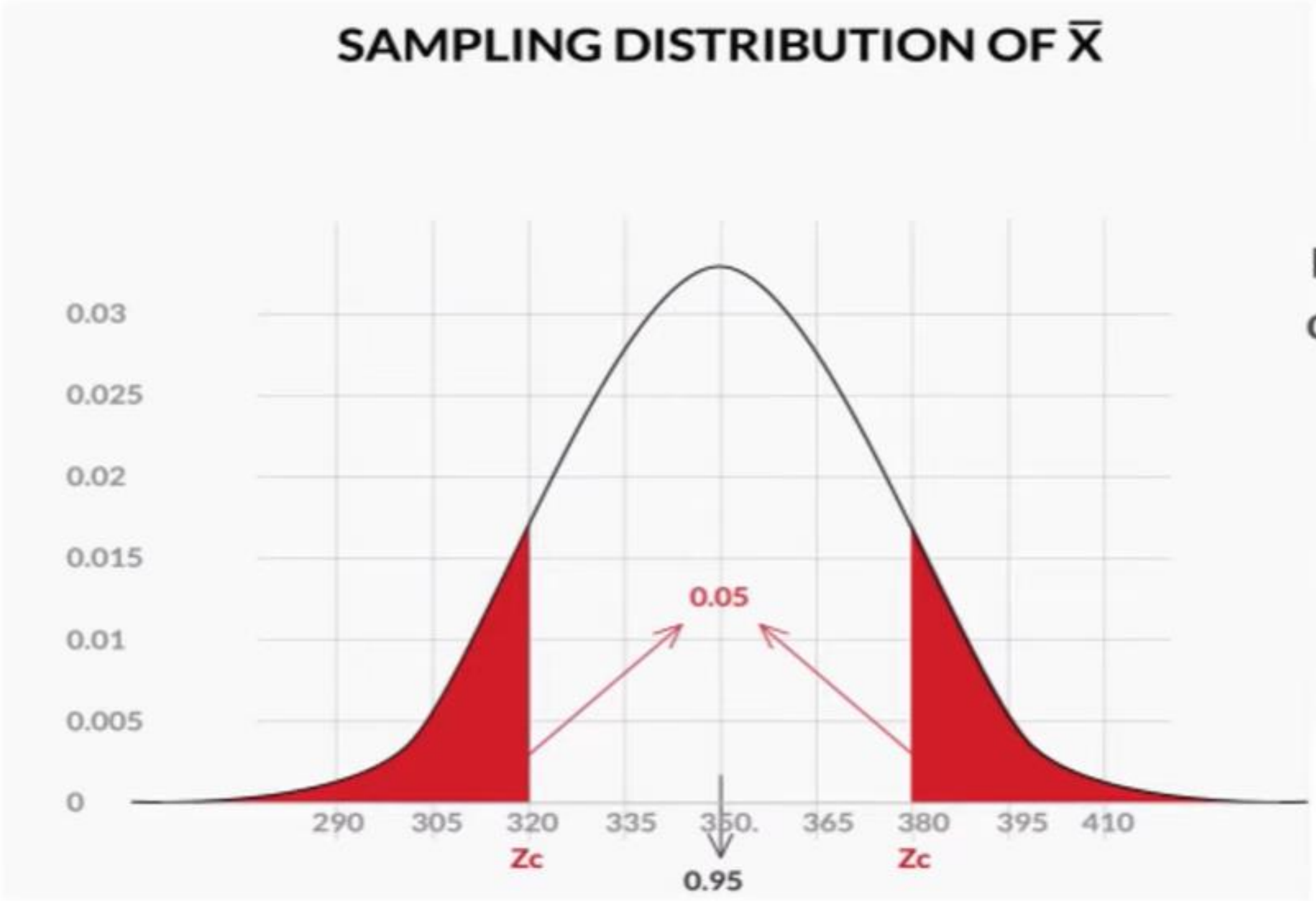
Sample Mean very far from 350

Reject H_0 and Accept H_1

Sample Mean very close to 350

H_0 cannot be rejected

Critical Value Method



$$\mu_{\bar{x}} = \mu = 350$$

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

$$= \frac{90}{\sqrt{36}}$$

$$\sigma = 15$$

$$\alpha = 0.05$$

Critical Value Method

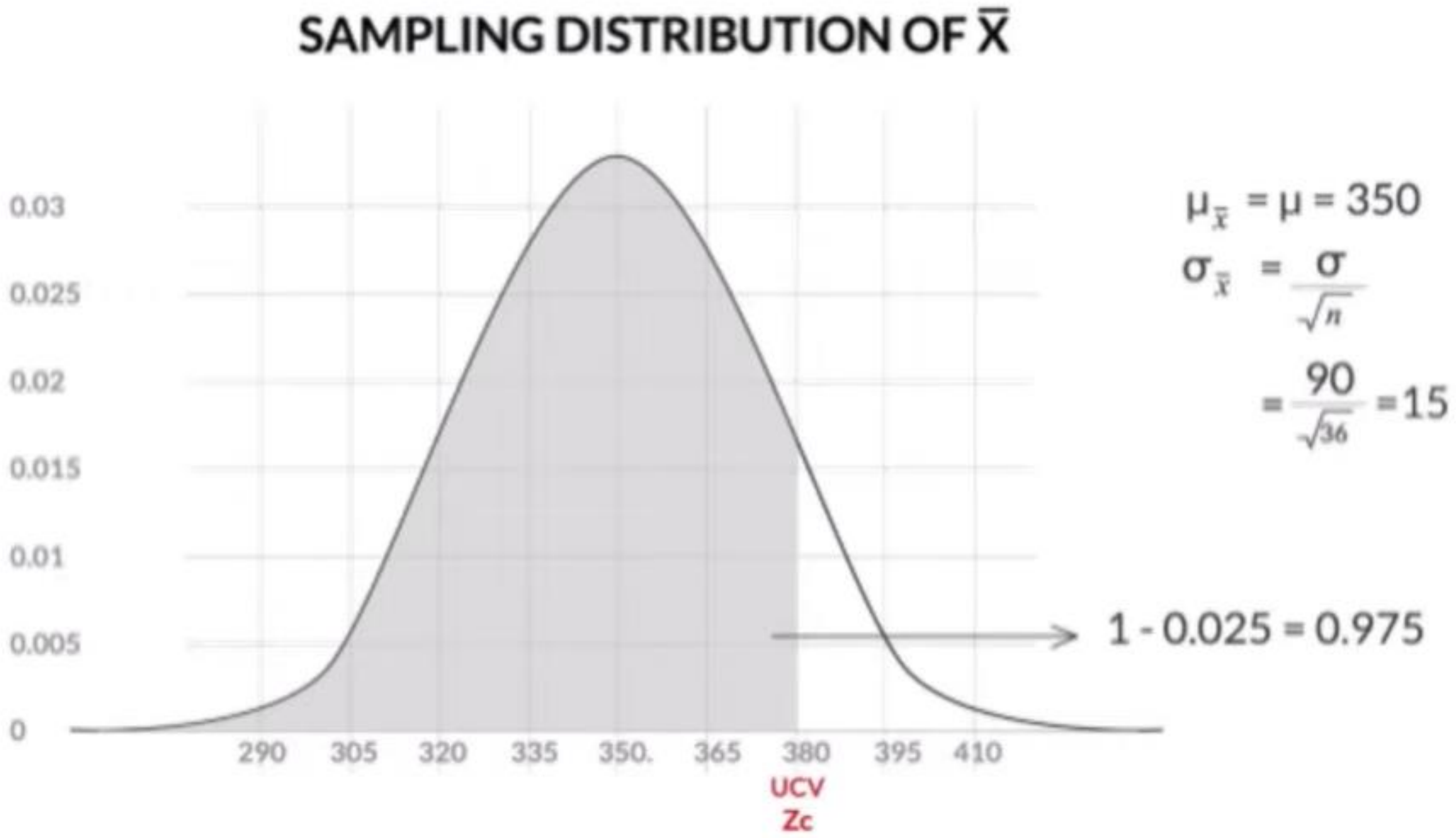
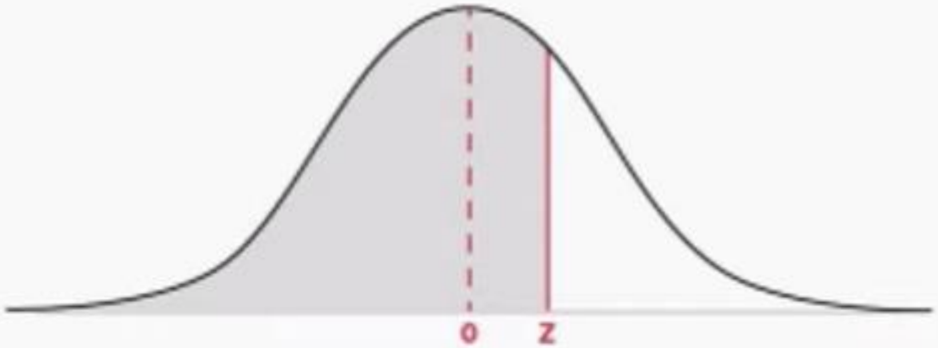


TABLE OF STANDARD NORMAL PROBABILITY FOR POSITIVE Z-SCORE

1.96

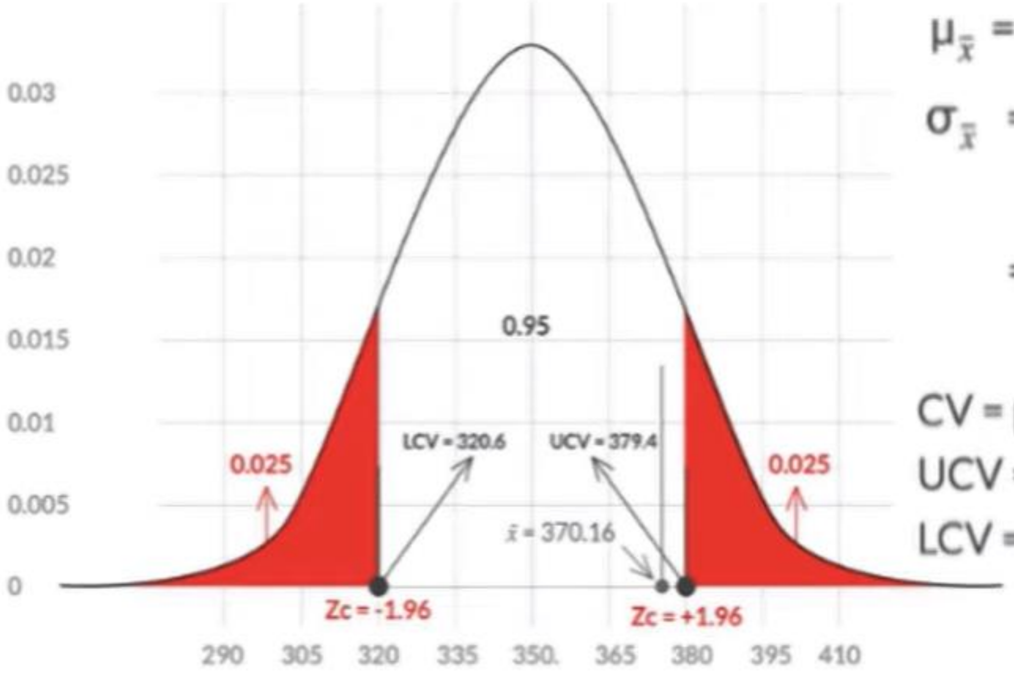


z	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750

The first step would be to formulate the hypotheses:

$H_0 : \mu = 350$ (There is no change in Status Quo)

$H_1 : \mu \neq 350$ (The Status Quo has changed)



$$\mu_{\bar{x}} = \mu = 350$$

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

$$= \frac{90}{\sqrt{36}} = 15$$

$$CV = \mu + (Zc \times \sigma_{\bar{x}})$$

$$UCV = 350 + 1.96 \times 15 = 379.4$$

$$LCV = 350 - 1.96 \times 15 = 320.6$$

Critical Value Method

After formulating the hypothesis, the steps you must follow for making a decision using the critical value method are as follows:

01

Calculate the value of Zċ from the given value of α (significance level). Take it a 5% if not specified in the problem.

02

Calculate the critical values (UCV and LCV) from the value of Zċ.

03

3. Make the decision on the basis of the value of the sample mean x with respect to the critical values (UCV AND LCV).

Part 2 Hypothesis Testing

Who makes a claim or a statement?

- A company:
  - The amount of cereal advertised on a box.
- An everyday person:
  - One person claims she is a more accurate basketball shooter than another person.
- A researcher:
  - They claim that their new drug is better than a presently used drug on the market.

How do we decide if the claim is believable? Or unbelievable?

We collect data!!!!

- A company:
  - Measure the amount in a sample of boxes.
- An everyday person:
  - Record the number of baskets made and missed.
- A researcher:
  - Assign some patients to each of the drugs and see how the patients fare on the drugs.

How do we decide if the claim is believable? Or unbelievable?

We collect data!!!!

- … and then we do a **hypothesis test**.
- What is a **hypothesis test**?
  - It is a standard procedure for testing a claim about the value of a population parameter.
  - It is a way to make a decision based on evidence (collected data).

Components of a Formal Hypothesis Test

Null Hypothesis: H_0

- The null hypothesis (denoted by H_0) is a statement that the value of a population parameter (such as proportion, mean, or standard deviation) is equal to some claimed value.
- We test the null hypothesis directly.
- Either reject H_0 or fail to reject H_0.

Alternative Hypothesis: H_1

- The **alternative hypothesis** (denoted by H_1 or H_a or H_A) is the statement that the parameter has a value that somehow differs from the null hypothesis.

- The symbolic form of the alternative hypothesis must use one of these symbols: $\neq$, $<$, $>$.

If you are conducting a study and want to use a hypothesis test to support your claim, the claim must be worded so that it becomes the alternative hypothesis

Identify the Null and Alternative Hypothesis.

- The mean height of professional basketball players is at most 7 ft.

- Step 1 express “a mean of at most 7 ft” in symbols as $\mu \leq 7$
- Step 2 if $\mu \leq 7$ is false, then $\mu > 7$ must be true.
- Step 3, expression $\mu > 7$ does not contain equality, so we let the alternative hypothesis H1 be $\mu > 7$, and we let H0 be $\mu = 7$.

Identify the Null and Alternative Hypothesis.

- The standard deviation of IQ scores of actors is equal to 15

- Step 1 express the given claim as $\sigma = 15$.
- Step 2, if $\sigma = 15$ is false, then $\sigma \neq 15$ must be true.
- Step 3, we let the alternative hypothesis H1 be $\sigma \neq 15$, and we let H0 be $\sigma = 15$.

Test Statistic

The test statistic is a value used in making a decision about the null hypothesis, and is found by converting the sample statistic to a score with the assumption that the null hypothesis is true.

Test Statistic - Formulas

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Test statistic for proportions

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

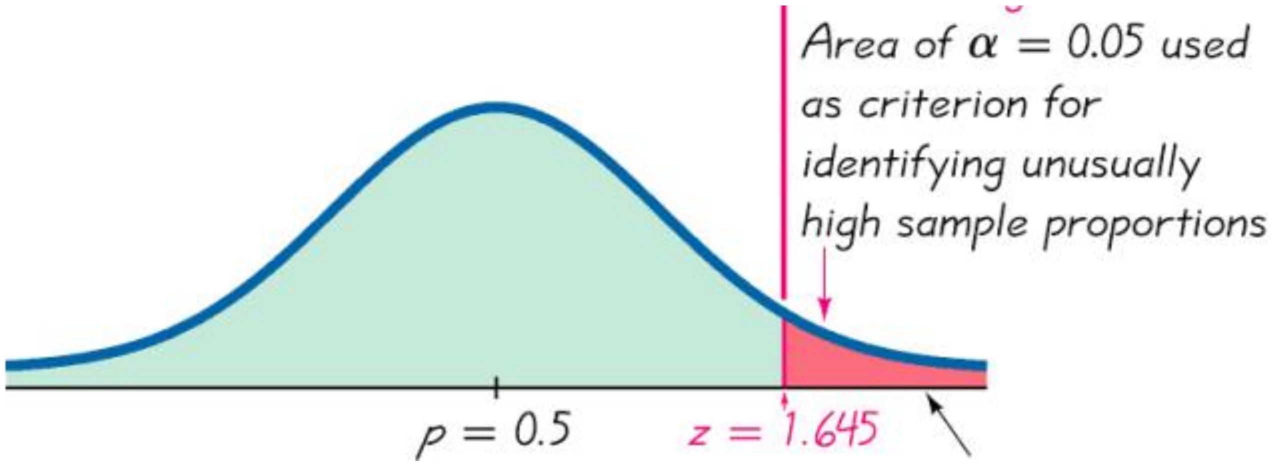
Test statistic for mean

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

Test statistic for standard deviation

Critical Region

The critical region (or rejection region) is the set of all values of the test statistic that cause us to reject the null hypothesis. For example, see the red-shaded region in the previous figure.

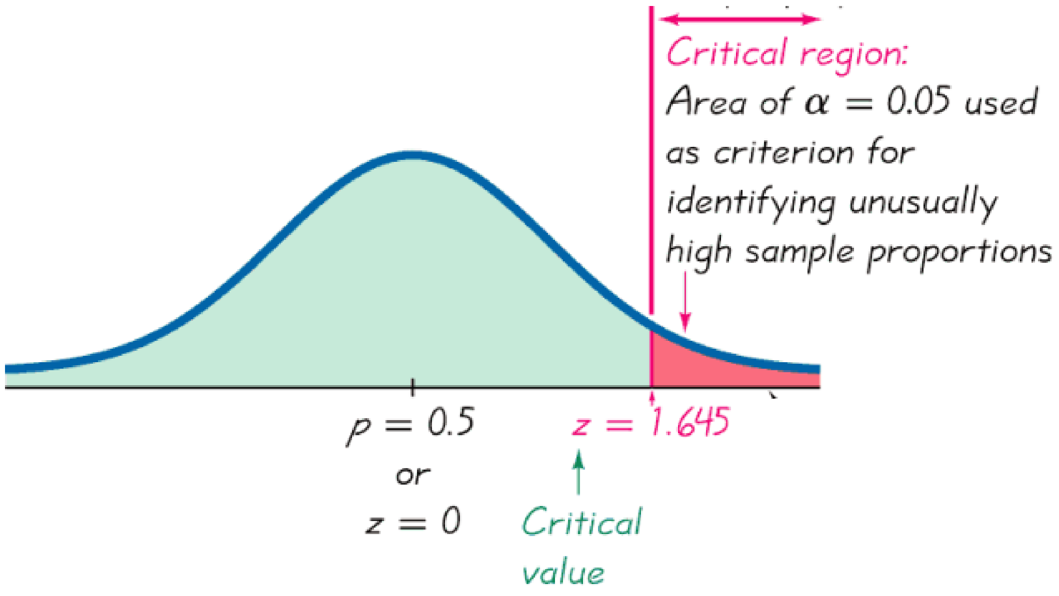


Significance Level

- The significance level (denoted by α) is the probability that the test statistic will fall in the critical region when the null hypothesis is actually true.

Critical Value

- A critical value is any value that separates the critical region (where we reject the null hypothesis) from the values of the test statistic that do not lead to rejection of the null hypothesis.
- The critical values depend on the nature of the null hypothesis, the sampling distribution that applies, and the significance level α.
- critical value of z = 1.645 corresponds to a significance level of α = 0.05.



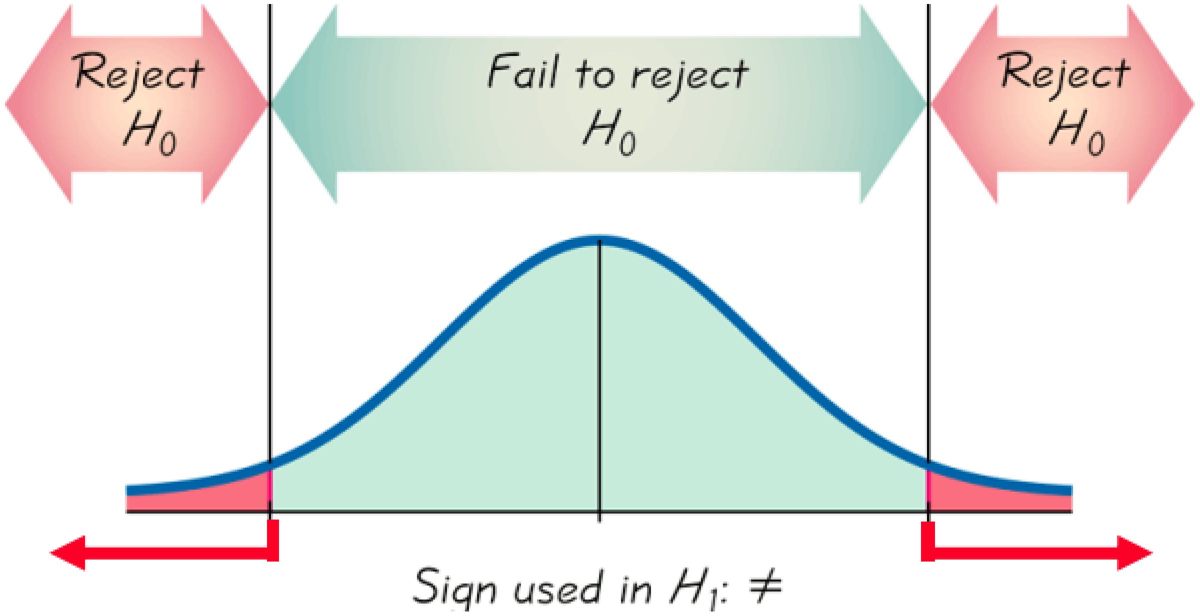
Two-tailed Test

H_0: $=$

H_1: $\neq$

Means less than or greater than

α is divided equally between the two tails of the critical region

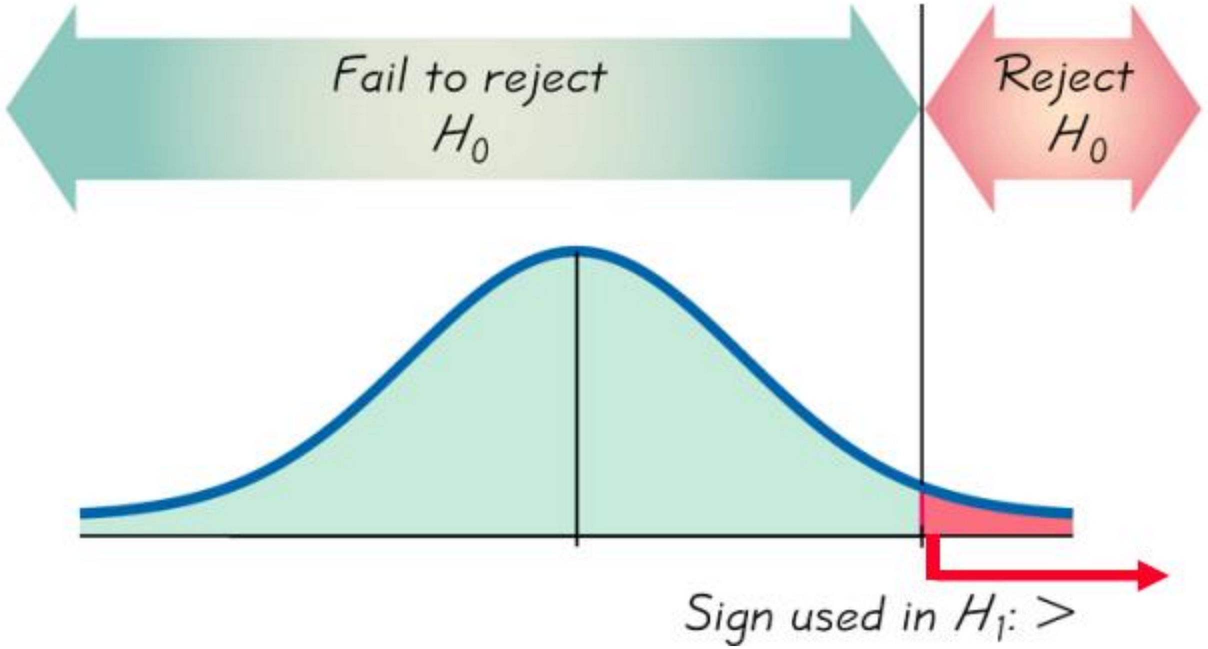


Right-tailed Test

H_0: =

H_1: >

Points Right

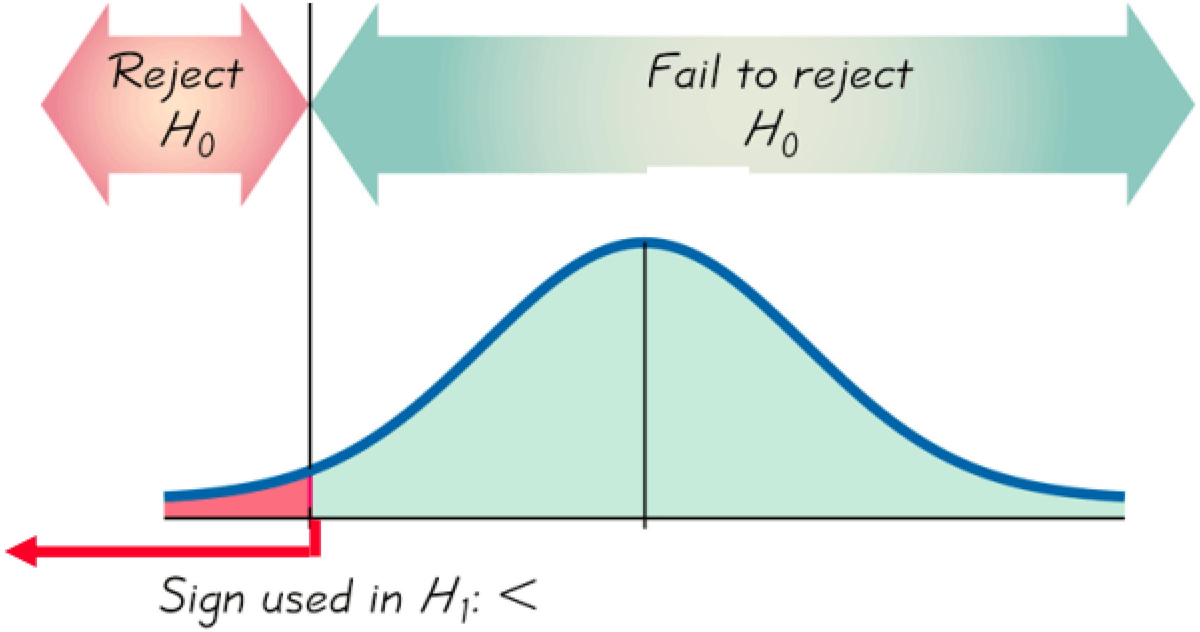


Left-tailed Test

H_0: =

H_1: <

Points Left



Conclusions in Hypothesis Testing

We always test the null hypothesis. The initial conclusion will always be one of the following:

1. Reject the null hypothesis.

2. Fail to reject the null hypothesis.

Decision Criterion

Traditional method:

Reject H_0 if the test statistic falls within the critical region.

Fail to reject H_0 if the test statistic does not fall within the critical region.

6-steps of Hypothesis Testing (using Critical Value Method)

1) State the Null and Alternative Hypotheses, and decide which one is the claim.
2) Determine the level of significance to be used.
3) Calculate the test statistic.
4) Determine (based on the level of significance) the critical value(s).
5) Make a decision regarding the Null Hypothesis.
6) State the conclusion of the hypothesis test in terms of the claim.

The alkaline batteries of a toy car were designed to last 30 hours, on average, with a known population standard deviation of 2.95 hours. However, lots of customers complained that the batteries were lasting less than 30 hours. You decide to randomly sample 38 of the manufacturer's batteries, whose mean life was 29.3 hours. Is there sufficient evidence at the 5% level of significance to support the claim that the mean battery life is less than 30 hours?

The alkaline batteries of a toy car were designed to last 30 hours, on average, with a known population standard deviation of 2.95 hours. However, lots of customers complained that the batteries were lasting less than 30 hours. You decide to randomly sample 38 of the manufacturer's batteries, whose mean life was 29.3 hours. Is there sufficient evidence at the 5% level of significance to support the claim that the mean battery life is less than 30 hours?

Step 1: Define Hypotheses

H_0: u = 30 hours

H_1: u < 30 hours

Step 2: Level of significance

$\alpha = 0.05$

The significance level (denoted by α) is the probability that the test statistic will fall in the critical region when the null hypothesis is actually true.

Step 3: Test statistic

$$z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma / \sqrt{n}}$$

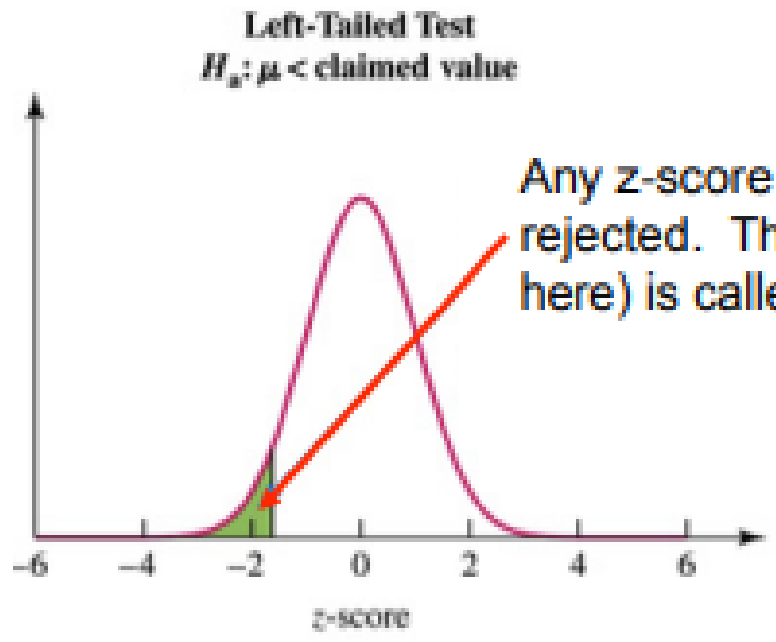
test statistic = -1.46

The alkaline batteries of a toy car were designed to last 30 hours, on average, with a known population standard deviation of 2.95 hours. However, lots of customers complained that the batteries were lasting less than 30 hours. You decide to randomly sample 38 of the manufacturer's batteries, whose mean life was 29.3 hours. Is there sufficient evidence at the 5% level of significance to support the claim that the mean battery life is less than 30 hours?

Step 4 Critical Value

Critical Values for Statistical Significance

- Significance level of **0.05**
  - One-sided left-tailed test H_a: $\mu < \mu_0$
    - Critical value is $z = -1.645$



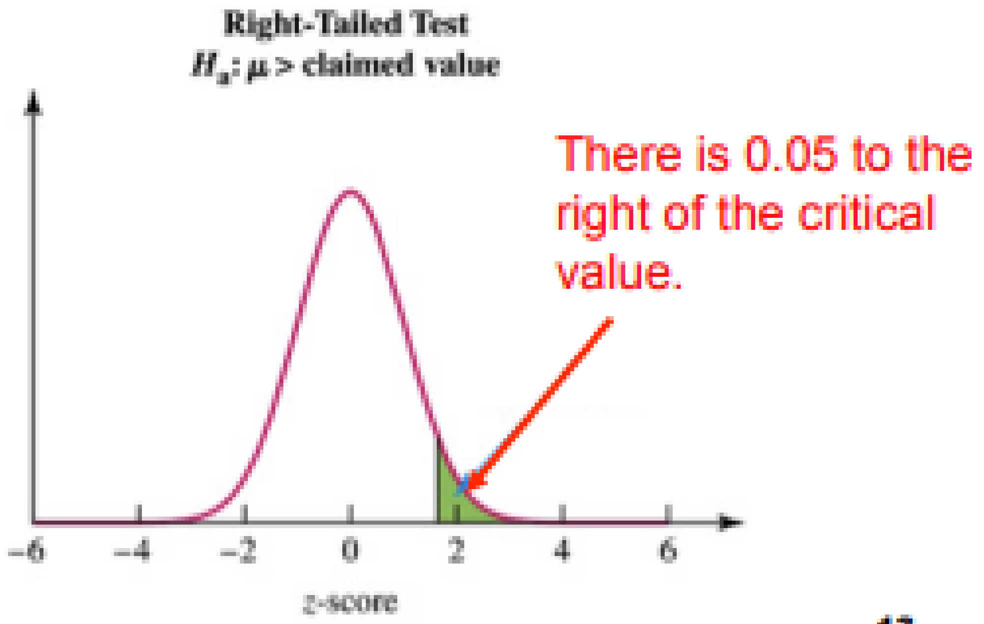
Any z-score to the left of -1.645 will be rejected. This zone (shown in green here) is called the **Rejection Region**.

If your z-score falls in the Rejection Region, you will reject the null.

Critical Values for Statistical Significance

- Significance level of **0.05**
  - One-sided right-tailed test $H_a: \mu > \mu_0$
    - Critical value is $z = 1.645$

A sample mean with a z-score greater than or equal to the critical value of 1.645 is significant at the 0.05 level.

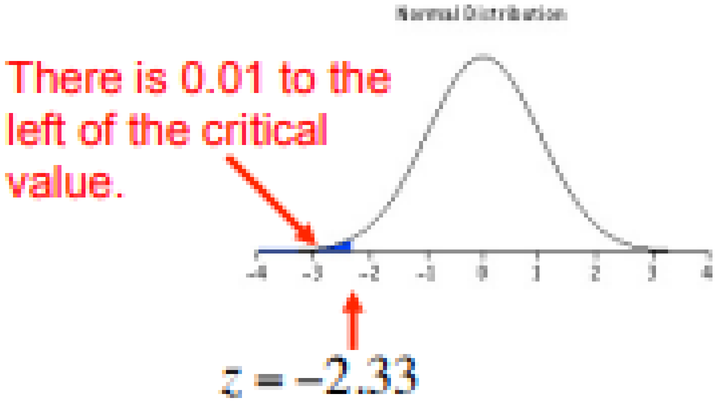


Critical Values for Statistical Significance

- Significance level of **0.01**
  - The same concept applies, but the critical values are farther from the mean.

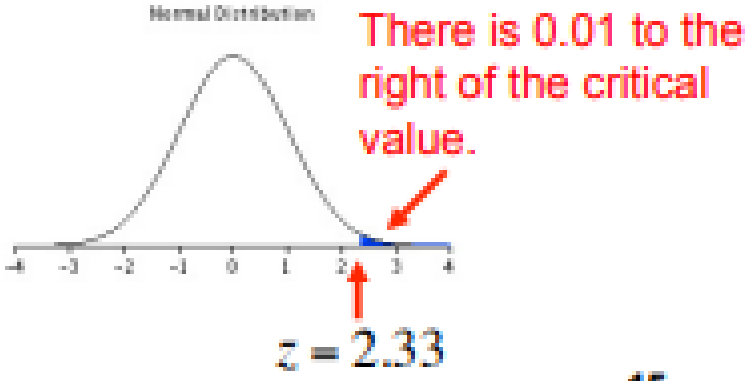
H_0: $\mu = \mu_0$

H_a: $\mu < \mu_0$ (one-sided test)



H_0: $\mu = \mu_0$

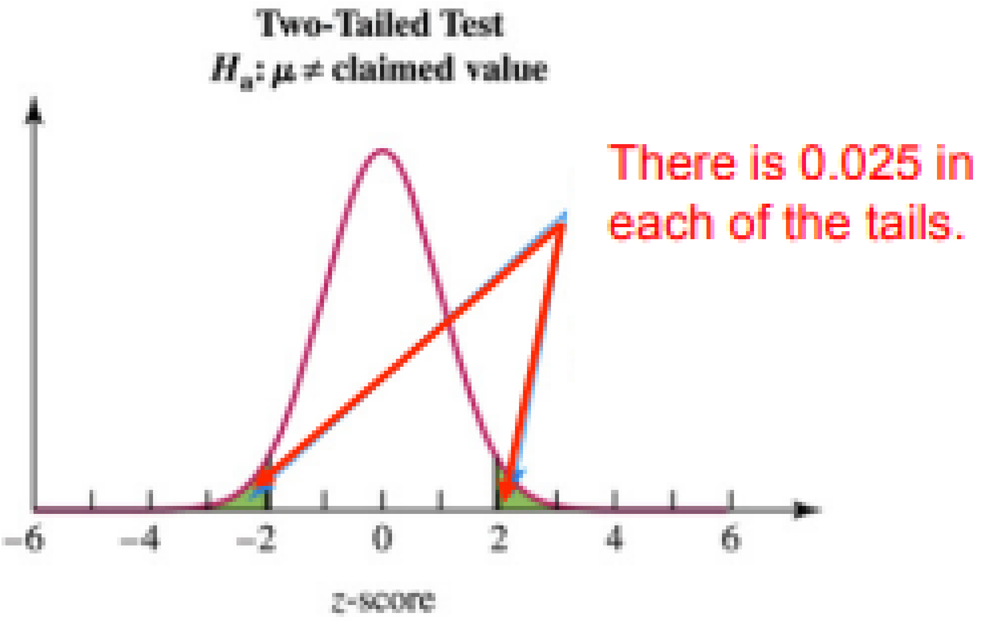
H_a: $\mu > \mu_0$ (one-sided test)



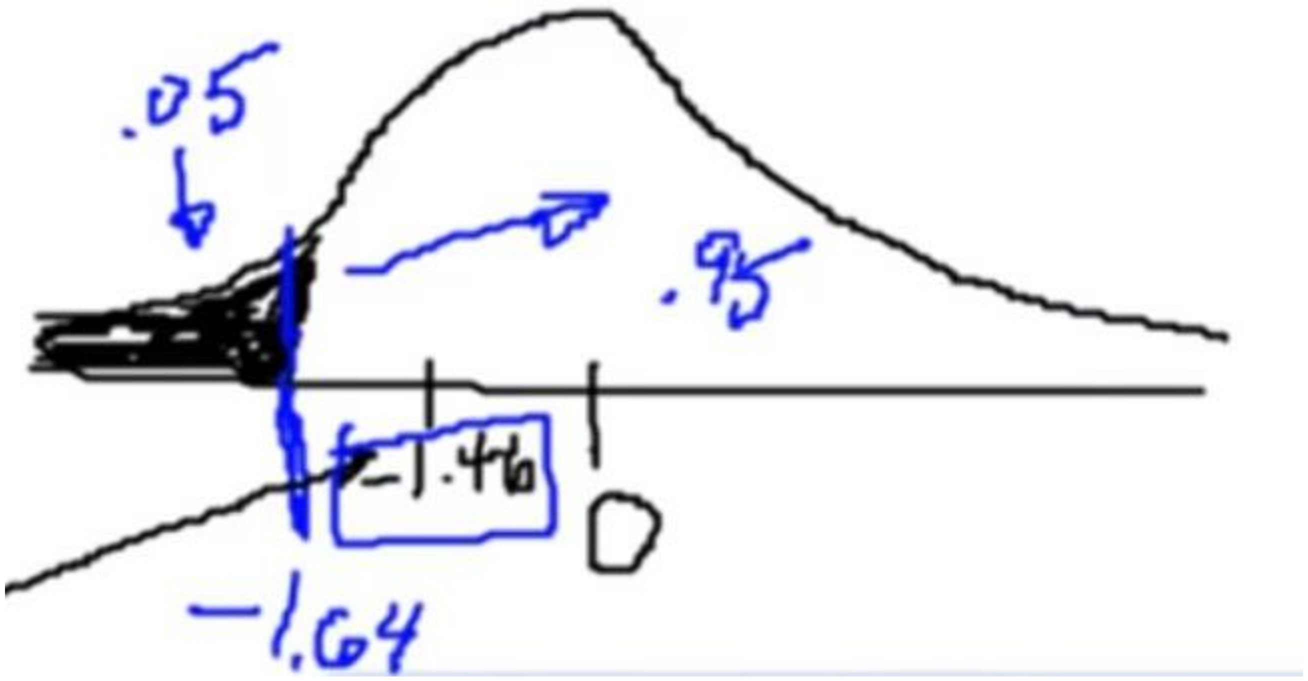
Critical Values for Statistical Significance

- Significance level of **0.05**
  - Two-sided test $H_a: \mu \neq \mu_0$ (two critical values)
    - Critical values are $z = -1.96 \; and \; z = 1.96$

A sample mean with a z-score in the rejection region (shown in green) is significant at the 0.05 level.



Step 4 Critical Value



Step 5: Null hypothesis?

Fail to reject the Null hyp.

Step 6: Conclusion

There is NOT sufficient evidence to support the claim,

Types of Errors

Type I error occurs when the null hypothesis is true, but we reject it

Type 2 Error occurs when the null hypothesis is false, but we fail to reject it.

	The null hypothesis is true	The null hypothesis is false
We decide to reject the null hypothesis	Type I error (rejecting a true null hypothesis) α	Correct decision
We fail to reject the null hypothesis	Correct decision	Type II error (failing to reject a false null hypothesis) β

P-Value

The ***P*-value** (or ***p*-value** or **probability value**) is the probability of getting a value of the test statistic that is **at least as extreme** as the one representing the sample data, assuming that the null hypothesis is true. The null hypothesis is rejected if the *P*-value is very small, such as 0.05 or less.

Decision Criterion - cont

P-value method:

Reject H_0 if the *P*-value $\leq \alpha$ (where α is the significance level, such as 0.05).

Fail to reject H_0 if the *P*-value $> \alpha$.

Procedure for Finding *P*-Values

Figure 8-6