Stochastic Processes

Week 08 (Version 1.0)

Hypothesis Testing

Hamid R. Rabiee

Fall 2024

Introduction to Hypothesis Testing

- A statistical hypothesis test is a method of statistical inference used to determine a possible conclusion from two different, and likely conflicting, hypotheses.
- A hypothesis is an assumption about the population parameter:
  - A parameter is a population mean or proportion.
  - The parameter must be identified before analysis.
- For example a hypothesis could be: The mean GPA of this class is 17.5.

The Null and Alternative Hypothesis

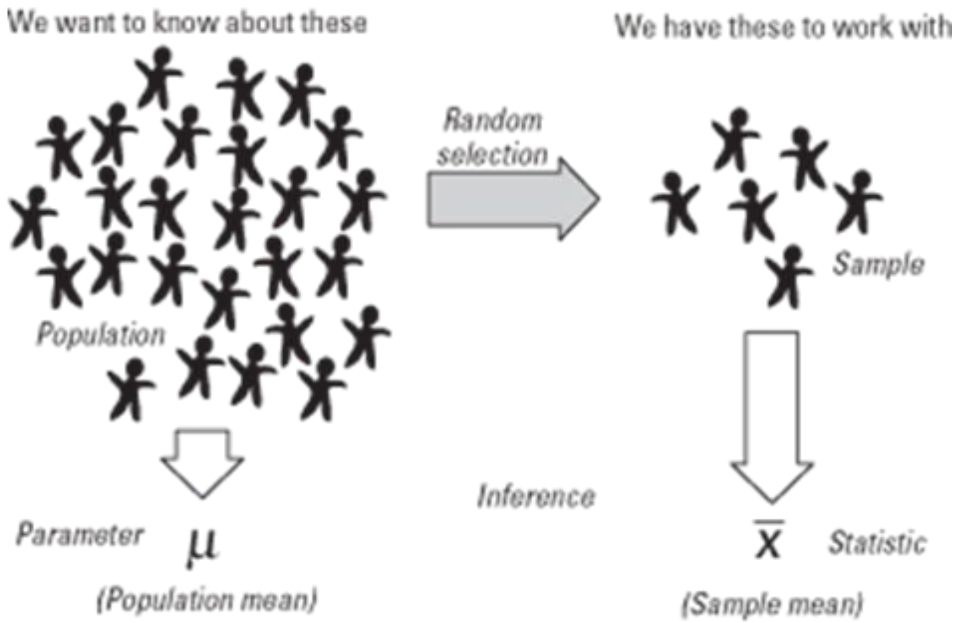
- The Null Hypothesis (H_0) states the assumption (numerical) to be tested.
- e.g. the average number of mobiles in Iranian homes is at least 3 (H_0: $\mu \geq 3$).
- We begin with the assumption that the null hypothesis is TRUE (Similar to the notion of innocent until proven guilty).
- The Null Hypothesis may or may not be rejected.
- The Alternative Hypothesis (H_1) is the opposite of the null hypothesis.

The Null and Alternative Hypothesis

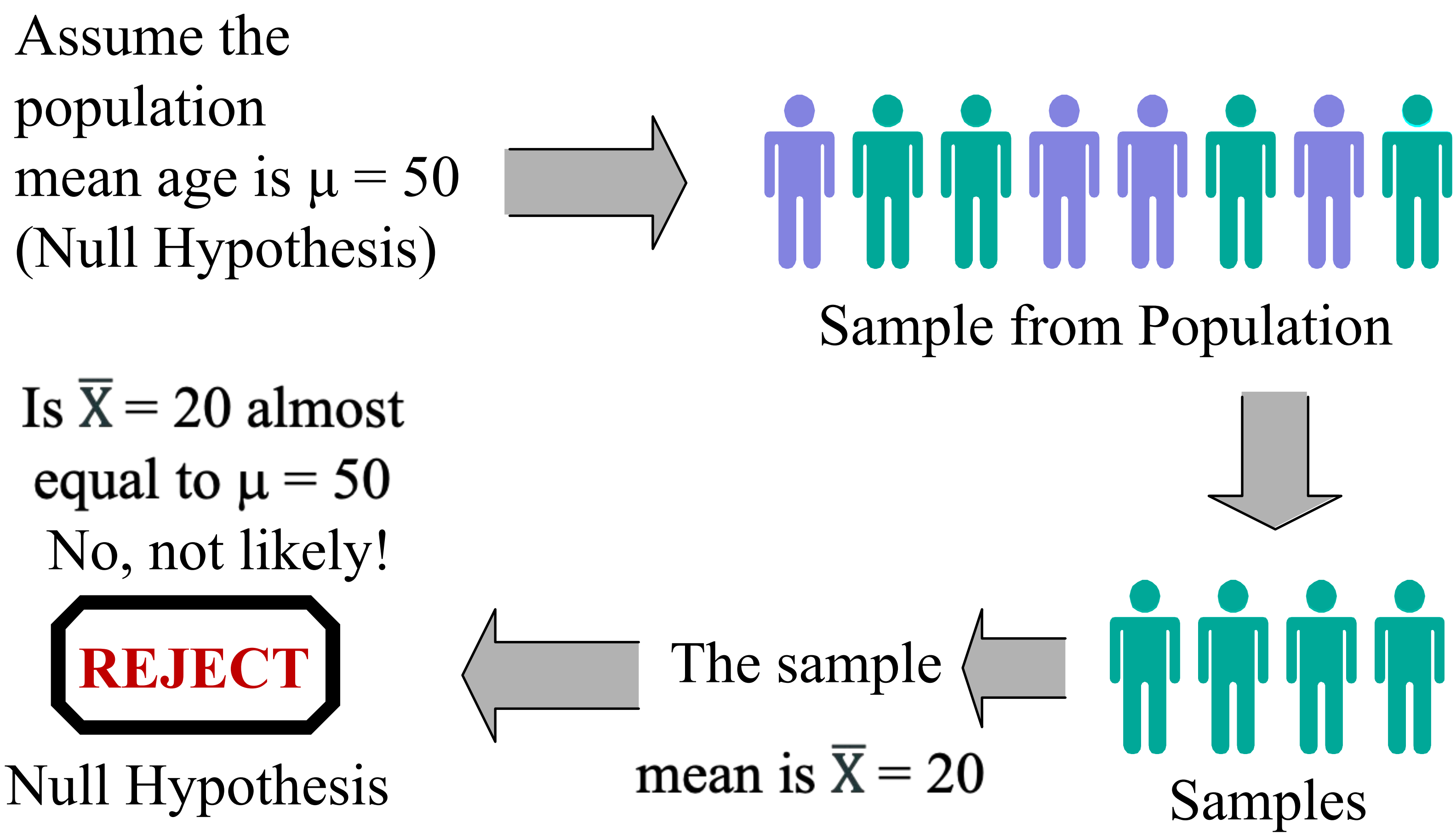
- e.g. the average number of mobiles in Iranian homes is less than 3 (H_1: $\mu < 3$)
- The Alternative Hypothesis may or may not be accepted.
- Hypothesis testing steps:
  1. Define your hypotheses (null, alternative)
  2. Specify your null distribution
  3. Do an experiment by sampling
  4. Calculate the test statistics of what you observed
  5. Reject or Accept the null hypothesis

The Null and Alternative Hypothesis

- Recall: Sample data ‘represents’ the whole population:



Hypothesis Testing Process by Example



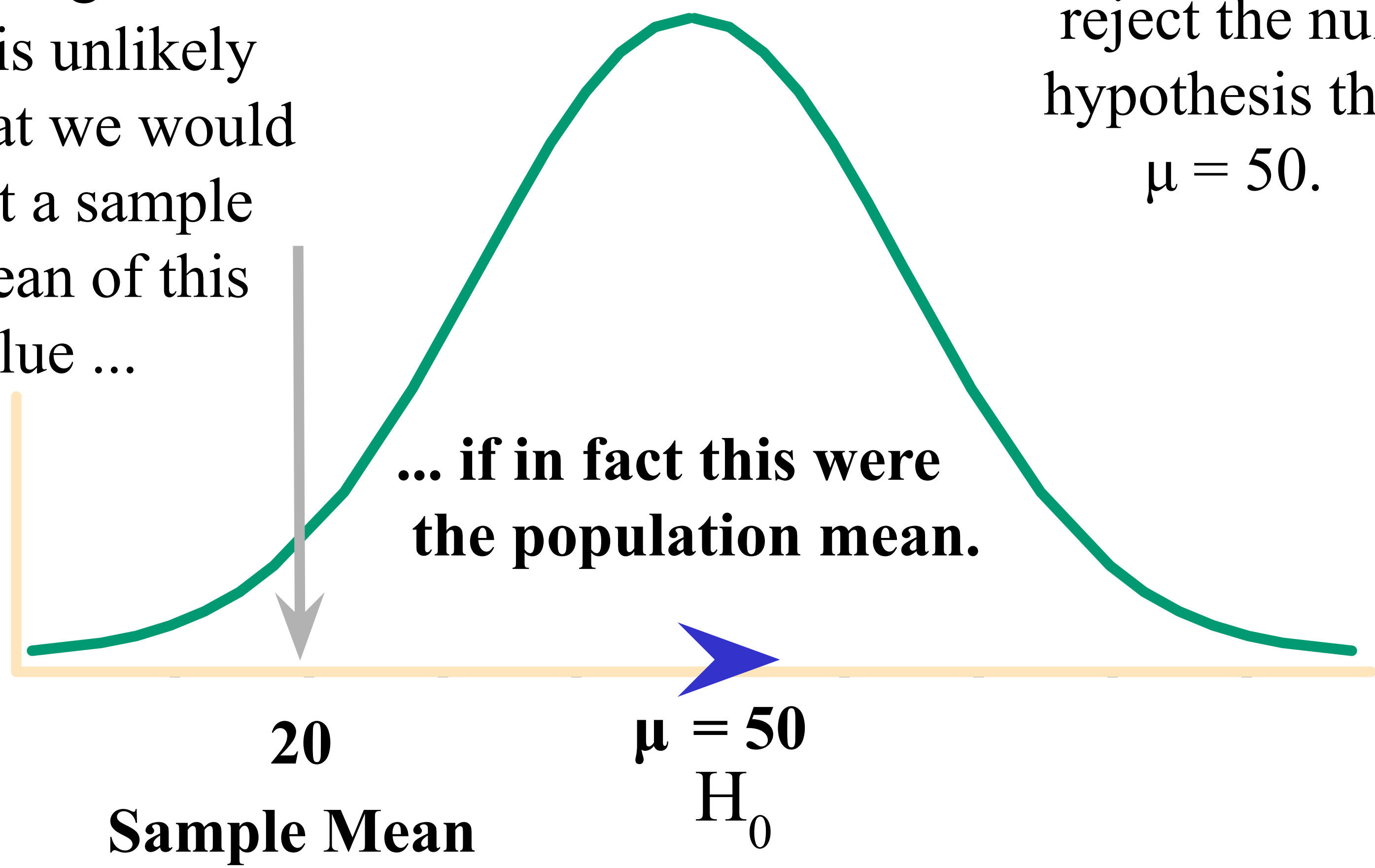
Hypothesis Testing Process

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 that $\mu = 50$.



Level of Significance: α

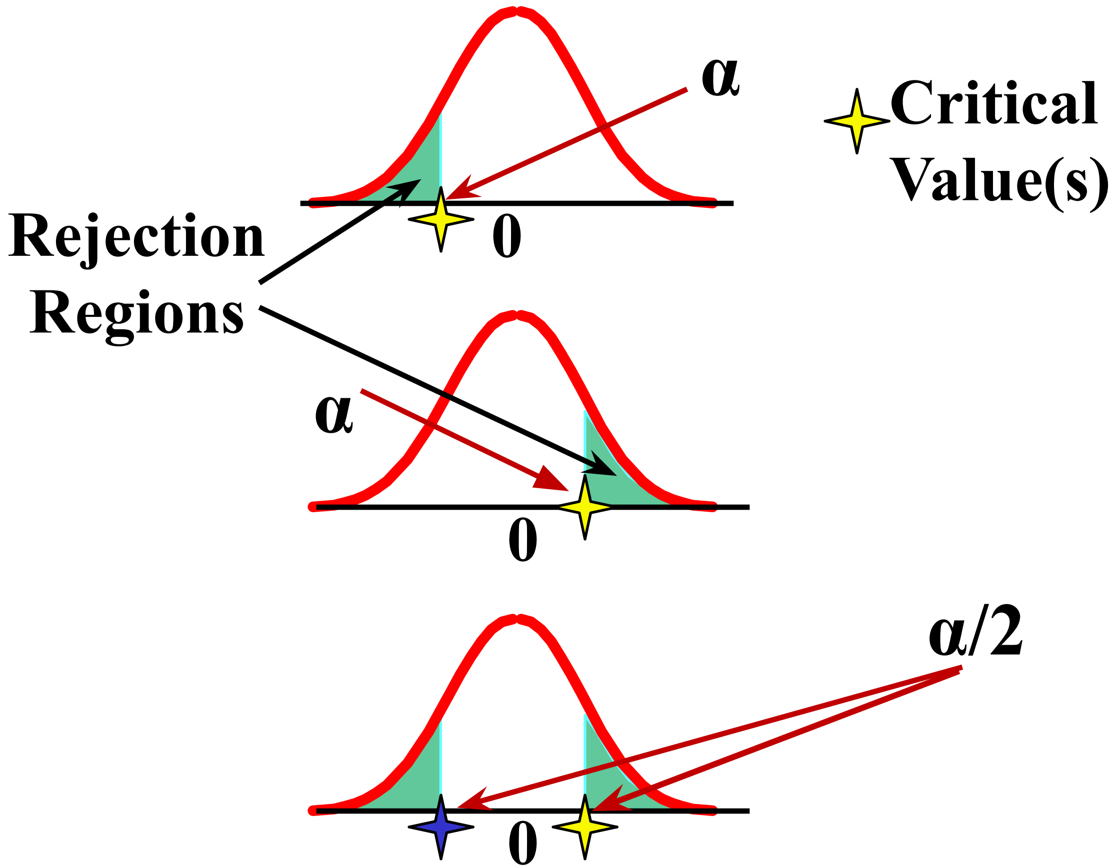
- Level of significance (α) defines unlikely values of sample statistic if Null Hypothesis is true.
  - It defines the Rejection Region of sampling distribution.
  - Typical values are 0.01, 0.05, 0.10.
  - It provides the Critical Value(s) of the test.

Level of Significance: α

$H_0: \mu \geq 3$
$H_1: \mu < 3$

$H_0: \mu \leq 3$
$H_1: \mu > 3$

$H_0: \mu = 3$
$H_1: \mu \neq 3$



Errors When Making Decisions

- Type I Error
  - Rejecting a true null hypothesis.
  - Has serious consequences.
  - Probability of Type I Error is α,
    (Called level of significance).
- Type II Error
  - Do not reject false null hypothesis.
  - Probability of Type II Error is β.

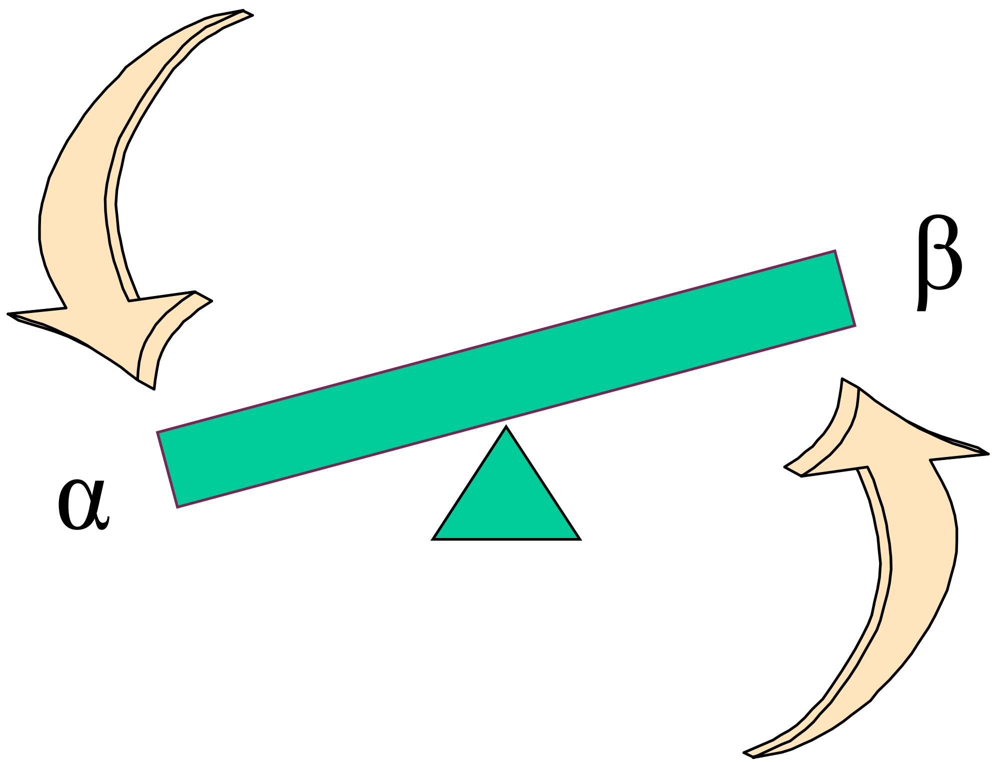
Decisions Possibilities: Court Example

H_0: Innocent

Jury Trial			Hypothesis Test		
	Actual Situation			Actual Situation	
Verdict	**Innocent**	**Guilty**	**Decision**	**H_0**	**H_0 False**
Innocent	Correct	Error	**Do Not Reject H_0**	True $1 - \alpha$	Type II Error (β)
Guilty	Error	Correct	**Reject H_0**	Type I Error (α)	Power $(1 - \beta)$

Errors When Making Decisions

- α & β Have an inverse relationship.
- Reducing probability of one error causes the other one going up.



Z-Test Statistics (σ known)

- Convert sample statistic (e.g., $\overline{X}$) to standardized Z variable:

$$Z = \frac{(\bar{X} - \mu\ )}{s}$$

- If the observed data X_1, ..., X_n are i.i.d. with mean μ, and variance σ^2, then the sample average $\overline{X}$ has mean μ and variance $s^2 = \frac{\sigma^2}{n}$.
- If Z test statistic falls in the critical (rejection) region, Reject H_0; Otherwise do not Reject H_0.

The P-Value Test

- Probability of obtaining a test statistic more extreme ($\leq$ or $\geq$) than actual sample value given H_0 is true.
- Used to make rejection decision:
  - If p-value $\geq \alpha$, do not Reject H_0
  - If p-value $< \alpha$, reject H_0
- A very small p-value means that such an extreme observed outcome would be very unlikely under the null hypothesis.

Z-Test Statistics vs P-Value Test

- The Z score is a test of statistical significance that helps you decide whether or not to reject the null hypothesis. The p-value is the probability that you have falsely rejected the null hypothesis.

- Z scores are measures of standard deviation. For example, if a tool returns a Z score of +2.5 it is interpreted as "+2.5 standard deviations away from the mean".

- P-values are probabilities. Both statistics are associated with the standard normal distribution. Very high or a very low (negative) Z scores, associated with very small p-values, are found in the tails of the normal distribution.

Hypothesis Testing: Example

Test the Assumption that the true mean # of mobiles in Iranian homes is at least 3.

1. State H_0 $H_0: \mu \geq 3$
2. State H_1 $H_1: \mu < 3$
3. Choose α $\alpha = .05$
4. Choose n $n = 100$
5. Choose Test: *Z Test (or p-value)*

Hypothesis Testing: Example

Test the Assumption that the true mean # of mobiles in Iranian homes is at least 3.

6. Set Up Critical Value(s) — Z= -1.645
7. Collect Data — 100 households surveyed
8. Compute Test Statistic — Computed Test Stat.= -2
9. Make Statistical Decision — Reject Null Hypothesis
10. Express Decision — The true mean # mobiles is less than 3 in the Iranian households.

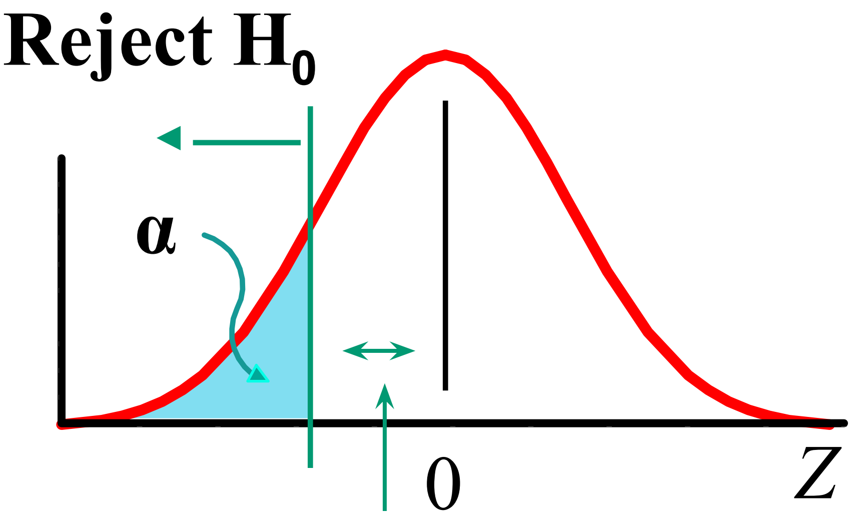
One-Tail Z Test for Mean (σ Known)

- Assumptions:
  - Population is normally distributed
  - If not normal, use large samples (CLT)
  - Null Hypothesis Has ≤ or ≥ Sign Only
- Z Test Statistic:

$$Z = \frac{(\bar{X} - \mu\ )}{s}$$

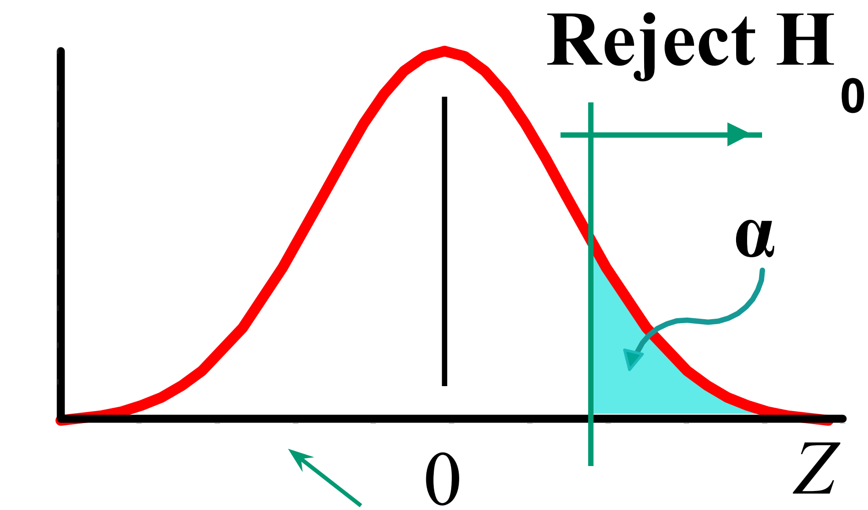
Rejection Region

H_0: $\mu \geq 0$
H_1: $\mu < 0$



Must Be *Significantly* Below μ = 0

H_0: $\mu \leq 0$
H_1: $\mu > 0$



Small values don't contradict H_0
Don't Reject H_0.

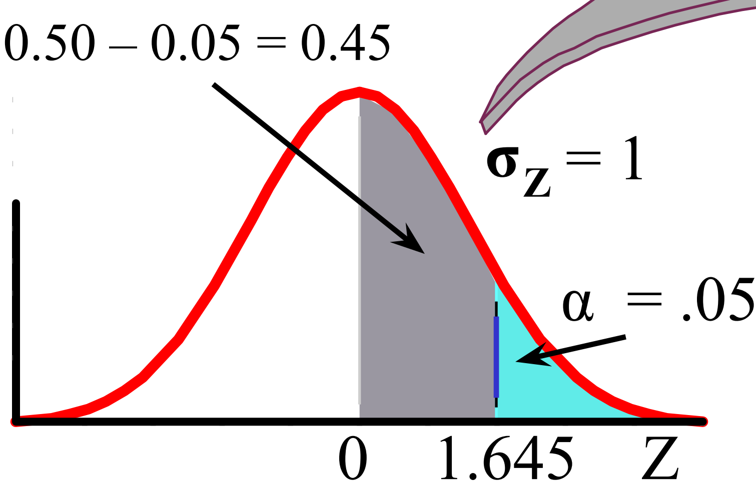
Example: One Tail Test

- Does an average box of cereal contain more than 368 grams of cereal?
- A random sample of 25 boxes showed $\bar{X} = 372.5$ grams.
- The company has specified σ to be 15 grams. Test at the $\alpha = 0.05$ level.

H_0: $\mu \leq 368$
H_1: $\mu > 368$

Finding Critical Values: One Tail

What is Z given $\alpha = 0.05$?



Critical Value
= 1.645

Probability Table

	.04	.05	.06
1.6	0.4495	0.4505	.4515
1.7	.5591	.5599	.5608
1.8	.5671	.5678	.5686
1.9	.5738	.5744	.5750

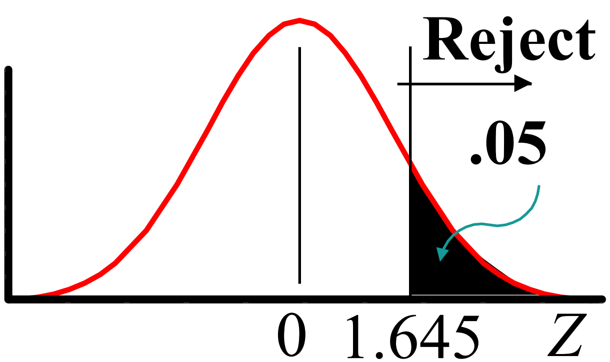
Example Solution: One Tail

H_0: $\mu \leq$ 368
H_1: $\mu >$ 368

α = 0.05

n = 25

Critical Value: 1.645



Test Statistic:

$$Z = \frac{(\bar{X} - \mu\ )}{s} = 1.5$$

(372.5-368)/(15/5)=1.5

Decision:

Do Not Reject at α = .05

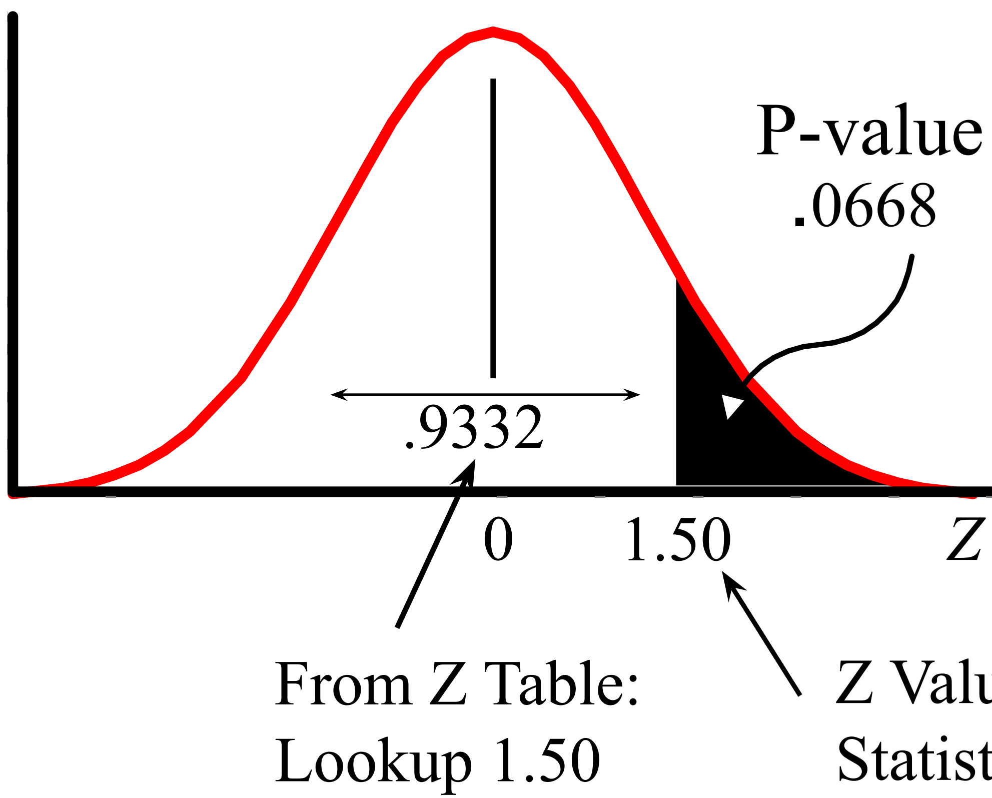
Conclusion:

No evidence true mean is more than 368.

P-Value Solution

P-Value is $P(Z \geq 1.50) = 0.0668$

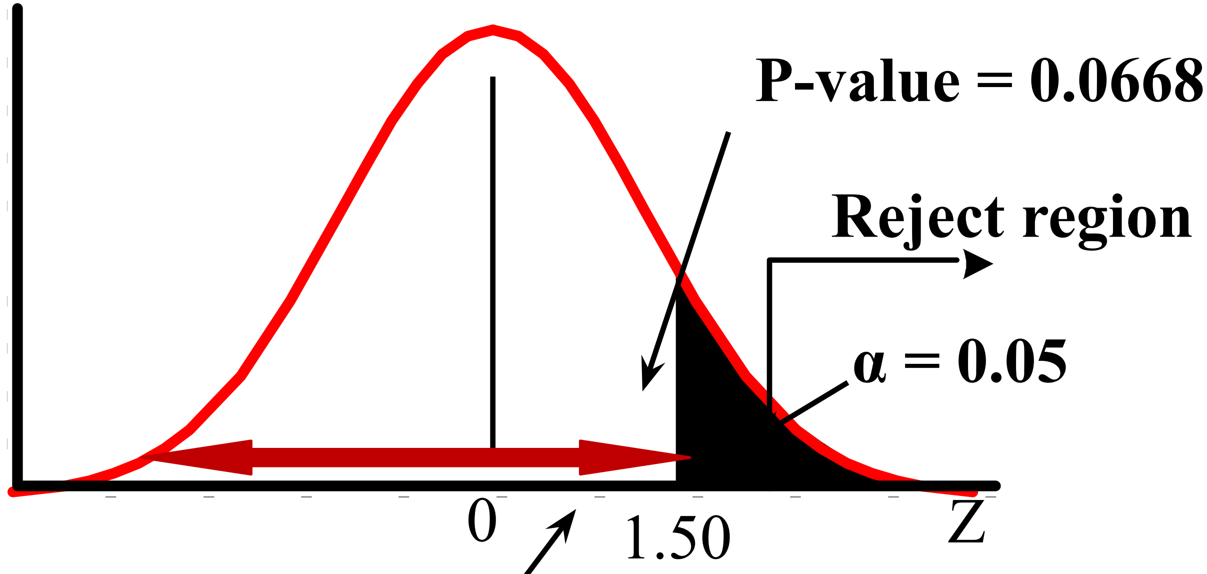
Use the alternative hypothesis to find the direction of the test.



$$\begin{array}{r} 1.0000 \\ -\ .9332 \\ \hline .0668 \end{array}$$

P-Value Solution

(P-value = 0.0668) ≥ (α = 0.05).
Do Not Reject.



Test statistic is in the Do Not Reject region

Example: Two Tail Test

- Does an average box of cereal contain 368 grams of cereal?
- A random sample of 25 boxes showed $\bar{X} = 372.5$ grams.
- The company has specified σ to be 15 grams. Test at the $\alpha = 0.05$ level.

H_0: $\mu = 368$

H_1: $\mu \neq 368$

Example Solution: Two Tail

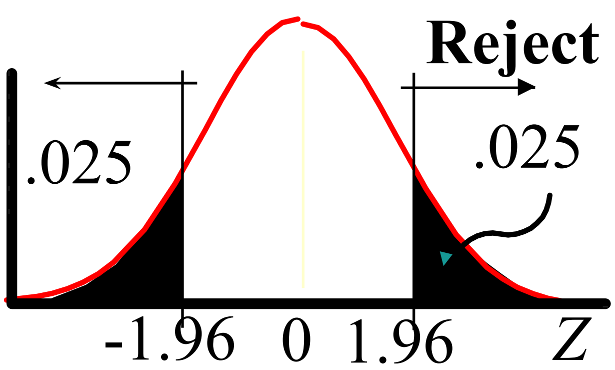
H_0: μ = 386

H_1: μ ≠ 386

α = 0.05

n = 25

Critical Value: ±1.96



Test Statistic:

$$Z = \frac{(\bar{X} - \mu\ )}{s} = 1.5$$

(372.5-368)/(15/5)=1.5

Decision:

Do Not Reject at α = .05

Conclusion:

No evidence that true mean is not 368.

Connection to Confidence Intervals

For $\overline{X} = 372.5$, $\sigma = 15$ and $n = 25$,

The 95% Confidence Interval is:

$$372.5 - (1.96)\,(15)/(5) \text{ to } 372.5 + (1.96)\,(15)/(5)$$

Or

$$366.62 \le \mu \le 378.38$$

If this interval contains the Hypothesized mean (368), we do not reject the null hypothesis.
Since it does, do not reject.

t-Test: σ Unknown

- t-tests are used to compare two population means.
- Assumptions:
  - Population is normally distributed
  - If not normal, only slightly skewed & a large sample taken (CLT)
- Use parametric test procedure
- t-test statistic:

$$t = \frac{Z}{s} = \frac{\bar{X} - \mu}{\widehat{\sigma}/\sqrt{n}}$$

Example: A Coin Toss

- You have a coin and you would like to check whether it is fair or not. Let θ be the probability of heads, $\theta=P(H)$
  - You have two hypotheses:
    - H_0 (the null hypothesis): The coin is fair i.e., $\theta=1/2$.
    - H_1 (the alternative hypothesis): The coin is not fair, $\theta\neq 1/2$.

Example: A Coin Toss

- We need to design a test to either accept H_0 or H_1
- We toss the coin 100 times and record the number of heads.
- Let X be the number of heads that we observe:

$$X \sim \text{Binomial}(100, \theta)$$

Solution: A Coin Toss

- if H_0 is true, then $\theta=\theta_0=½$
  - we expect the number of heads to be close to 50
- We suggest the following criteria: If $|X-50|$ is less than or equal to some threshold, we accept H_0.
- On the other hand, if $|X-50|$ is larger than the threshold we reject H_0.
- Let's call that threshold **t**.

If $|X-50| \leq t$, accept H_0.

If $|X-50| > t$, accept H_1

Solution: A Coin Toss

- We need to define more parameters, e.g. Error Probability.
- Type I Error: Wrongly reject H_0 when it is true.
- P(Type I Error) = $P(|X-50| > t \mid H_0) <= \alpha$
  - α : level of significance
- Knowing that P is a binomial distribution we can now calculate t.

Solution: A Coin Toss

- X ~ Binomial (n, θ =½)
  - Can be estimated by a normal distribution since n is large enough: Y ~ N(0, 1)
- Y = $\dfrac{X - n\theta}{\sqrt{n\theta(1-\theta)}}$
- $P(|X-50| > t \mid H_0) = P(|Y| > t/5 \mid H_0)$
- if $c = t/5$:
  - $|Y| > c$, accept H_0
  - o.w. accept H_1

Solution: A Coin Toss

- Since Y = $\frac{X-5}{50}$, the conclusion can be rewritten as:
  - if $|X - 50| \leq 9.8$, accept H_0
  - else if $|X - 50| > 9.8$, accept H_1

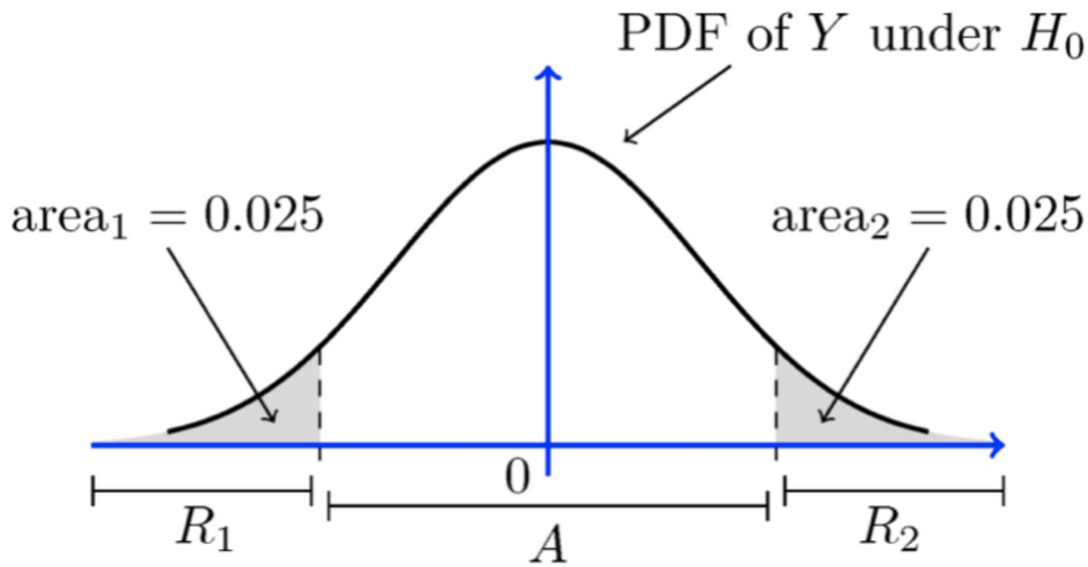
- if X in$\{41, 42, \ldots, 59\}$, accept H_0

Solution: A Coin Toss

- $P(|Y| > c) = 1 - P(-c \leq Y \leq c)$
  - Assuming $Y \sim Normal(0, 1)$
- $P(|Y| > c) = 2 - 2\phi(c) = 0.05$
  - using the z-table: $c = \phi^{-1}(0.975) = 1.96$

- $|Y| \leq 1.96$, accept H_0, o.w. accept H_1
  - Acceptance Region = [-1/96, 1.96]
  - Rejection Region = ?

Visualization: A Coin Toss



A = Acceptance Region

$R = R_1 \cup R_2$ = Rejection Region

$\alpha = \text{P(type I error)} = \text{area}_1 + \text{area}_2 = 0.05$

Next Week:

Markov Chains

Have a good day!