DATA SCIENCE WITH R

DATA SCIENCE WITH R

HYPOTHESIS TESTING

Introduction to Hypothesis Testing

Basic Framework of a Hypothesis Test

Distance Measures

Central Limit Theorem

Types of Hypothesis Tests

Basic Set up of a Hypothesis Test

Null Hypothesis (H^O):

Basic Set up of a Hypothesis Test

Null Hypothesis (H^O):

- Example: Production process is fine

Basic Set up of a Hypothesis Test

Null Hypothesis (H^O):

- Example: Production process is fine

Alternative Hypothesis (H^A):

Basic Set up of a Hypothesis Test

Null Hypothesis (H^O):

- Example: Production process is fine

Alternative Hypothesis (H^A):

- Production process not fine; producing > weight than specified

Basic Set up of a Hypothesis Test

Null Hypothesis (H^O):

- Example: Production process is fine

Alternative Hypothesis (H^A):

- Production process not fine; producing > weight than specified

Test Distribution:

Basic Set up of a Hypothesis Test

Null Hypothesis (H^O):

- Example: Production process is fine

Alternative Hypothesis (H^A):

- Production process not fine; producing > weight than specified

Test Distribution:

- Appropriate distribution to be used to calculate probability of outcome

Basic Set up of a Hypothesis Test

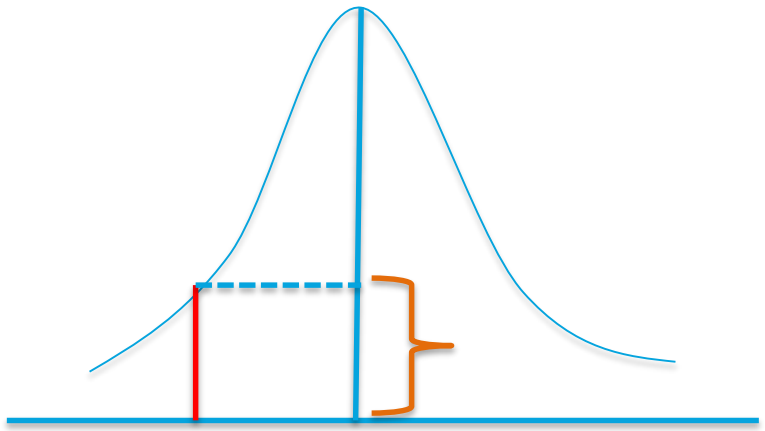
Significance Level (alpha : α)

Basic Set up of a Hypothesis Test

Significance Level (alpha : α)

Criterion used for rejecting the null hypothesis

Red line in figure

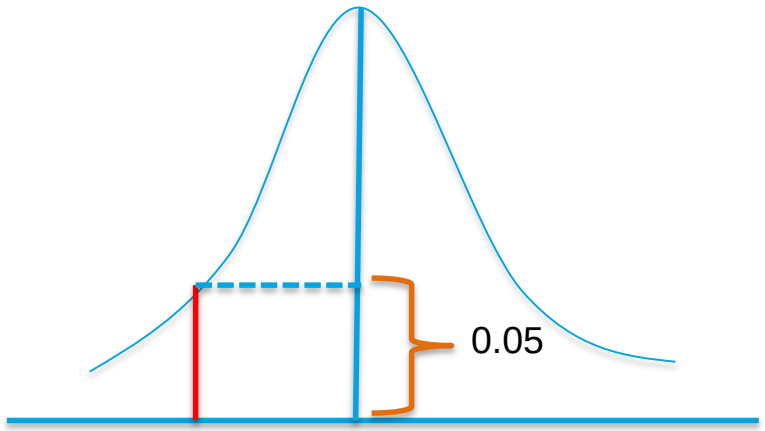


Basic Set up of a Hypothesis Test

Significance Level (alpha : α)

Criterion used for rejecting the null hypothesis

Red line in figure

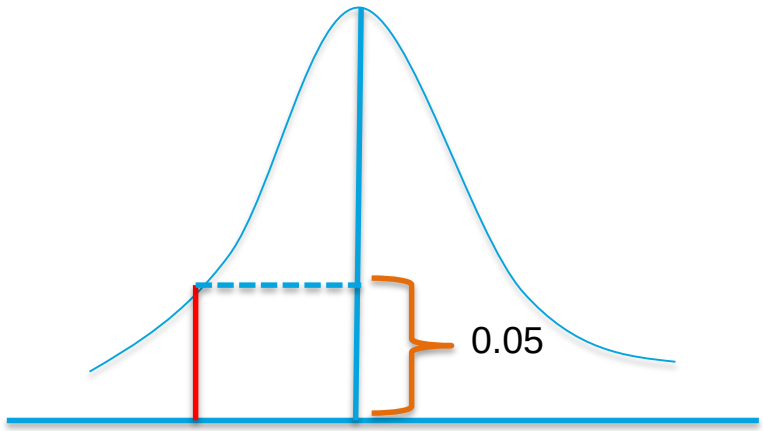


Basic Set up of a Hypothesis Test

Significance Level (alpha : α)

Criterion used for rejecting the null hypothesis

Red line in figure



P-Value:

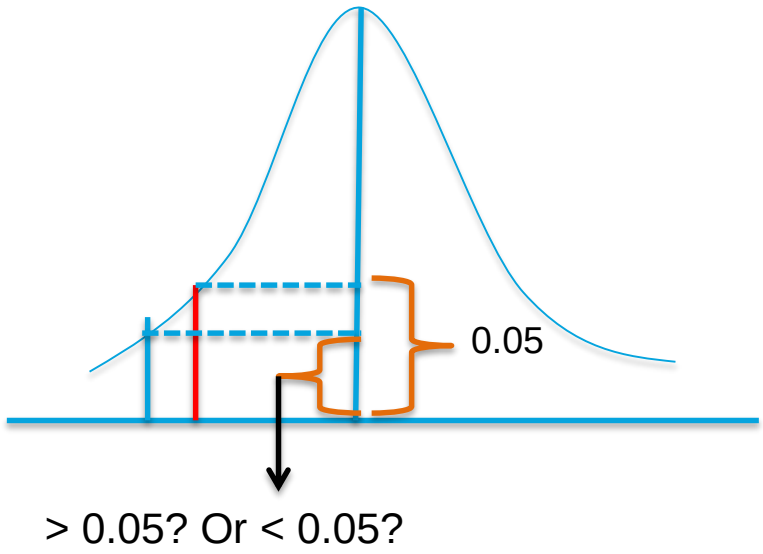
The probability of outcomes more extreme than the observed outcome, assuming the null is true

Basic Set up of a Hypothesis Test

Significance Level (alpha : α)

Criterion used for rejecting the null hypothesis

Red line in figure



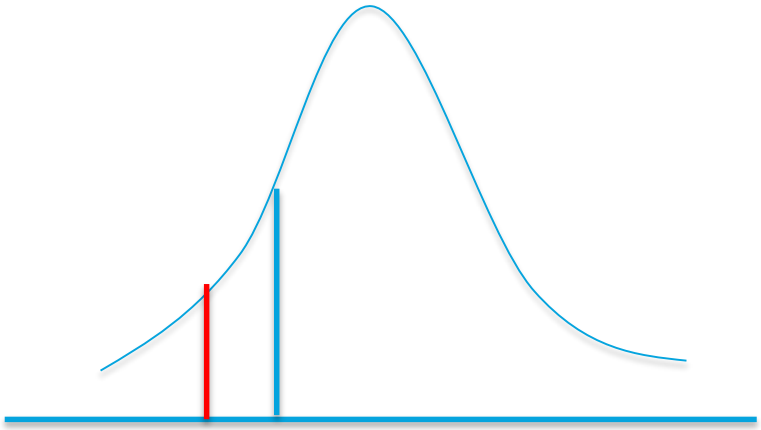
P-Value:

The probability of outcomes more extreme than the observed outcome, assuming the null is true

Area to the left of the blue line

Basic Set up of a Hypothesis Test

What if p-value > significance level?



Conclusion?

Recap

Recap

- **Null**: Ho: Process is fine, sample not different from population

Recap

- **Null**: Ho: Process is fine, sample not different from population
- **Alternate:** H1: Process not fine, sample wts > population wt

Recap

- **Null**: Ho: Process is fine, sample not different from population
- **Alternate:** H1: Process not fine, sample wts > population wt
- **Test Distribution**: Normal

Recap

- **Null**: Ho: Process is fine, sample not different from population
- **Alternate:** H1: Process not fine, sample wts > population wt
- **Test Distribution**: Normal
- **Significance value:** 5%

Recap

- **Null**: Ho: Process is fine, sample not different from population
- **Alternate:** H1: Process not fine, sample wts > population wt
- **Test Distribution**: Normal
- **Significance value:** 5%
- **P –value** (using norm.dust) function = 0.07

Recap

- **Null**: Ho: Process is fine, sample not different from population
- **Alternate:** H1: Process not fine, sample wts > population wt
- **Test Distribution**: Normal
- **Significance value:** 5%
- **P –value** (using norm.dust) function = 0.07
- **Conclusion?** Fail to reject the null

THANK YOU