Statistical Inference for Applied Sciences: Why and How?

Shashikant Gupta, PhD (IISc)
SoBAS, GD Goenka University

Roadmap

- What is statistics?
- Statistical Inference
- Hypothesis testing
- Type I and Type II Errors
- Parametric vs non-parametric tests
- One & Two tailed tests
- Summary
- References for further study

What is Statistics?

- As the knife is to surgery so is the statistics to research.
- Statistics is defined as the Collection, Presentation, Analysis and Interpretation of numerical data.
- Branch of Mathematics?
- Different from Mathematics!

- There are three kinds of lies: lies, damned lies, and statistics!

Does it really work?

An example from Political Science: Exit Pole

News24-Today's Chanakya (LS 2019)

- NDA: 350 ± 14
- UPA: 95 ± 9
- Others: 97 ± 11

Actual Results:

- NDA: 353
- UPA: 90
- Others: 99

Some exit polls predicted accurate results!

Descriptive Vs Inferential Statistics

Descriptive Statistics

- Organize, analyze and present data in a meaningful way.
- Result: Charts, Graphs and Tables
- Describe a situation

Inferential Statistics

- Compares, tests and predicts data
- Result: Probability
- Explain the chances of occurrence of an event

Statistical Inference: Aspects

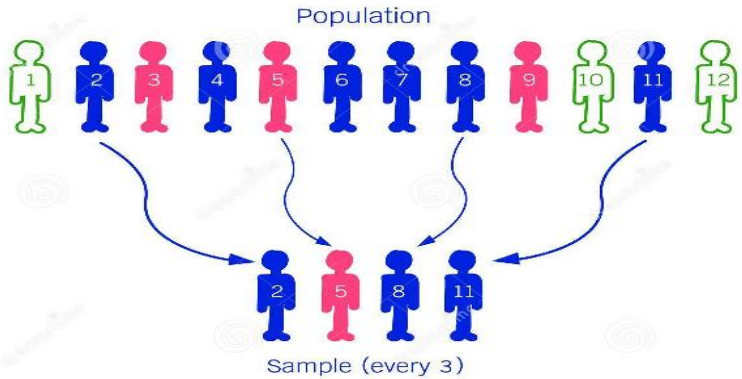
Statistical Inference

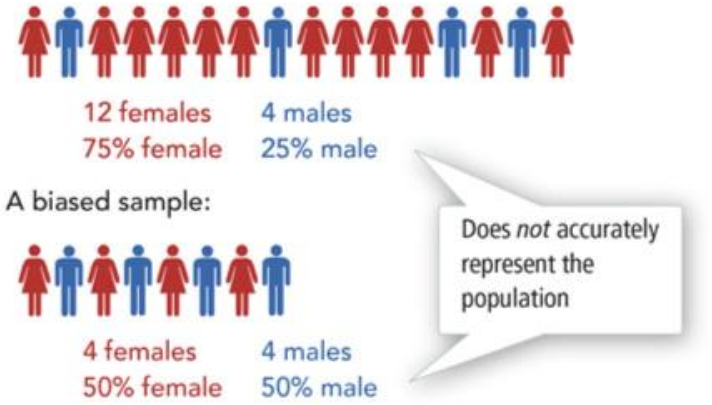
Estimation of
Population Parameters

Testing of Hypothesis

Population vs Sample!

- Que: What is the average height of all men (age between 20-50) in India?
- The set of height values of all men: population (Average = μ)
- Select some men randomly and measure their heights: sample (Average =$\bar{X}$)





Hypothesis and Hypothesis testing

- During investigation there is assumption/presumption which must be proved or disproved.
- On the basis of hypothesis we collect data.
- Two hypothesis:
    - Null Hypothesis (H_0): No difference between the groups.
    - Alternative Hypothesis (H_A): Alternative to Null hypothesis.

Test Null Hypothesis Against Alternative Hypothesis (H_0 vs H_A)

A simple example

- Question: The average height of men (μ_1) in India is same as that of women (μ_2), or it is different?
- Height of all men in India is called the population in our problem.
- Height of all women is also a population.
- H_0: $\mu_1 = \mu_2$ (Null hypothesis)
- H_A: $\mu_1 \neq \mu_2$ (Alternative hypothesis)

Can you think of a different Alternative Hypothesis?

Population vs Sample!

- Populations parameters (μ_1 and μ_2) are not known.
- Select some men **randomly** and measure their heights: Sample 1.
- Select some women **randomly** and measure their heights: samples2.
- Let $\overline{X_1} = 156.5\ cm$ and $\overline{X_2} = 152.5\ cm$
- Can not reject the Null hypothesis unless a detailed analysis is performed.

Hypothesis Testing: An example from Medical science

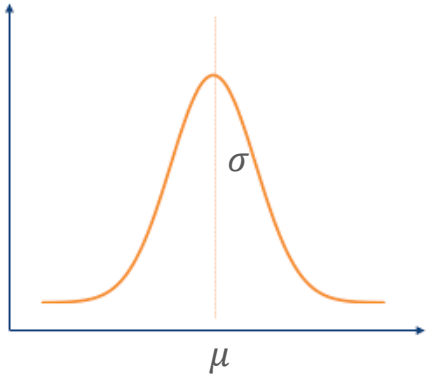
Is it true that vitamin C has the ability to cure or prevent the common cold? Or is it just a myth?

- **Null hypothesis (H_0)** - Children who take vitamin C are no less likely to become ill during flu season.
- **Alternative hypothesis (H_A)** - Children who take vitamin C are less likely to become ill during flu season.

Which test should we apply?

- Student t test: to compare means of two small groups
- Z-test: compares mean for large samples
- F-test/ANOVA: to compare variances of different groups
- In all of these tests we assume that the data obey

  Gaussian distribution!



Parametric vs non-parametric test!

- Tests mentioned in previous slide are called parametric test.
- Assumption: Data is drawn from Gaussian distribution.
- Nature prefers Gaussian distribution.
- What if the data does not obey Gaussian or we are not sure about the distribution?
- Use non-parametric tests: rank sum test, U test, K-S test, etc.

Possible Outcomes

Four possible outcomes of the test:

1. H_0 is true but the test rejects it. **X**
2. H_0 is false but the test does not reject it. **X**
3. H_0 is true and the test does not reject it. √
4. H_0 is false and the test rejects it. √

Type 1 Error: rejecting null hypothesis when it is true (α error).

Type 2 Error: Fail to reject null hypothesis when it is false (β error).

Can we avoid the errors?

- Statistics is a game of probability. It cannot be known for certain whether statistical conclusions are correct.

- All statistical hypothesis tests have a probability of making type I and type II errors.

An example from Forensic Science

Consider a criminal trial. We test the following hypothesis:

- H_0: The defendant did not commit the crime.
- H_A: The defendant committed the crime.

Following errors are possible:

- Type I: Convicting a person who, in reality did not commit the crime.
- Type II: Acquitting a person who, in reality committed the crime.

Significance level

- Probability of a Type I error, given H_0 is true, is called the significance level of the test.
- $P(\text{Type I error}|H_0 \text{ is true}) = \alpha$
- Usually, we set $\alpha = 0.05$ (5%)
- It is acceptable to have a 5% probability of incorrectly rejecting the true null hypothesis.

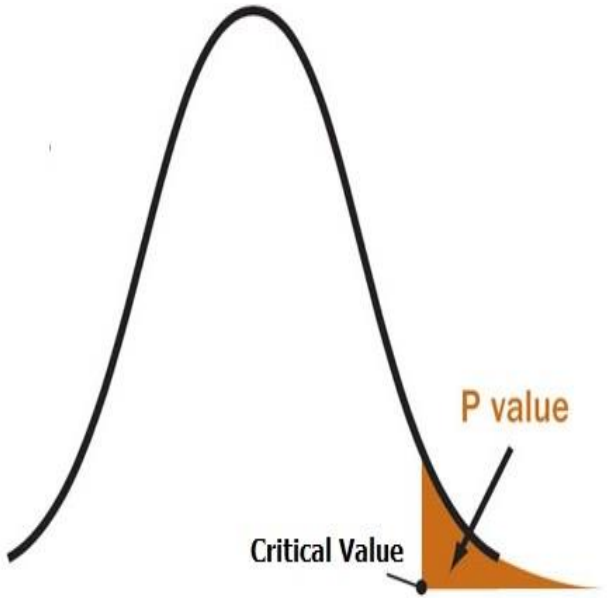
Power of test

- Probability of Type II error: P(Type II error)= β
- Depends on choice of α, sample size, and the true value of the parameter.
- Power of a test is related to type II error.
- Power = 1 – *P*(Type II error)= 1 – β
- If we choose a very small value of α, we will be making it very difficult to reject the Null hypothesis. Hence increasing the chances of Type II error.

How to decide rejection or non-rejection of H_0?

- p-value = Probability of observing a result as extreme as the current one, assuming that H_0 is true.
- Small p-value: reject H_0.
- Large p-value: Fail to reject H_0.
- How small or how large?
- Compare p-value with significance level (α).

When should we reject H_0?



- Zone of acceptance: If the results of a sample falls in the plain area: H_0 is not rejected.

- Zone of rejection: If the result of a sample falls in the shaded area it is significantly different from population value: H_0 is rejected.

Tongue twister

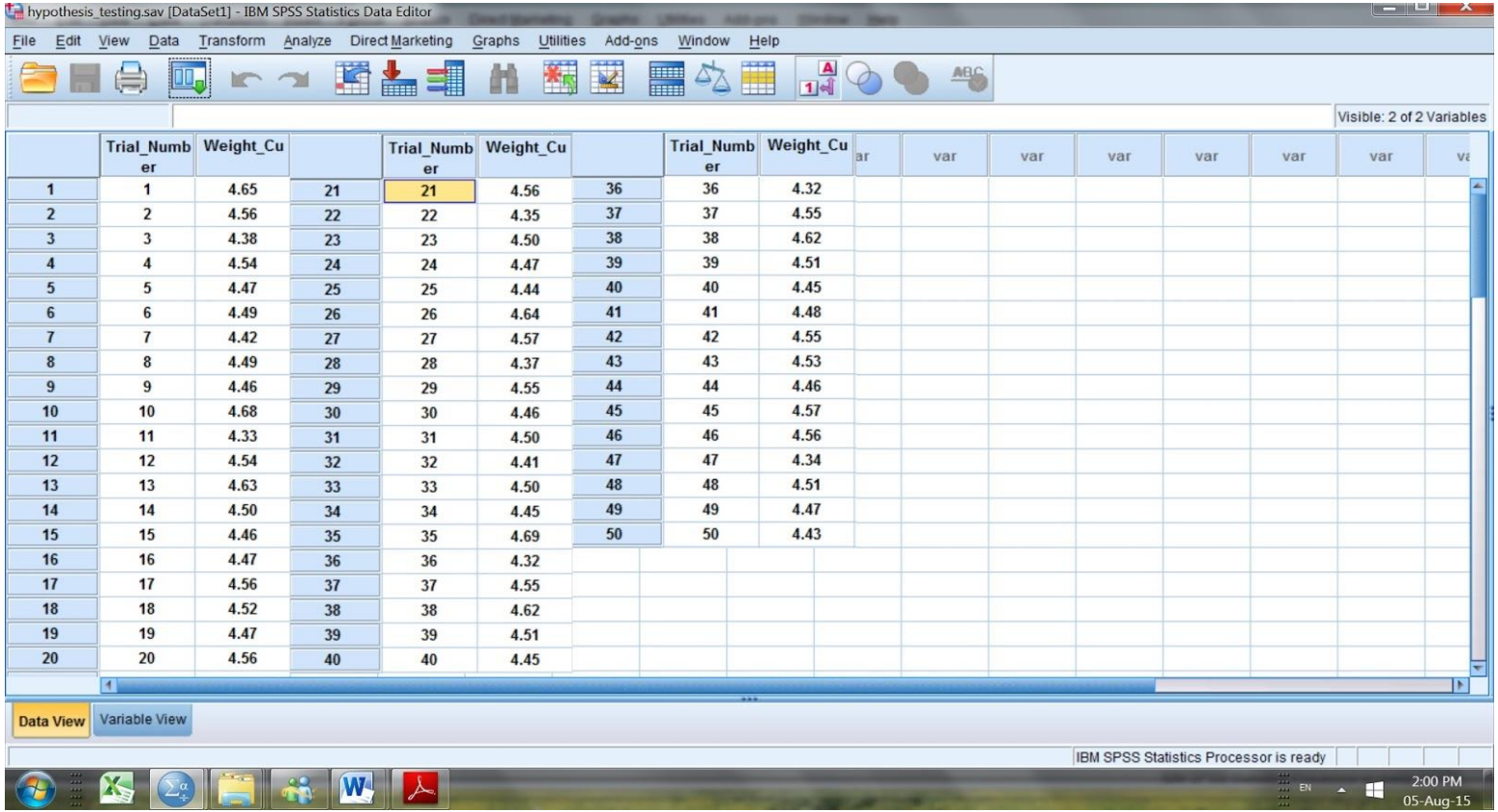
- p-value < 0.05 means that Null hypothesis is false?
- **No!**
- p-value < 0.05 means that we reject the Null hypothesis at 95% confidence level.

- Can we accept a hypothesis?
- **No!**
- We reject or fail to reject.
- Any single experiment or test in future can reject the hypothesis.

Student’s t test – an example from Chemistry

- T test is used to compare two means when sample size is small.
- A new analytical instrument is tested in chemistry lab by determining the mass of Cu in per gram of a certified reference material (CRM). Certified value in CRM is 4.54 mg in 1 gm of CRM.
- Data (on next slide) collected from the instrument.
- Does the instrument work properly?

Data for the test



	Trial_Number	Weight_Cu		Trial_Number	Weight_Cu		Trial_Number	Weight_Cu	var	var	var	var	var	var	var	var
1	1	4.65	21	21	4.56	36	36	4.32								
2	2	4.56	22	22	4.35	37	37	4.55								
3	3	4.38	23	23	4.50	38	38	4.62								
4	4	4.54	24	24	4.47	39	39	4.51								
5	5	4.47	25	25	4.44	40	40	4.45								
6	6	4.49	26	26	4.64	41	41	4.48								
7	7	4.42	27	27	4.57	42	42	4.55								
8	8	4.49	28	28	4.37	43	43	4.53								
9	9	4.46	29	29	4.55	44	44	4.46								
10	10	4.68	30	30	4.46	45	45	4.57								
11	11	4.33	31	31	4.50	46	46	4.56								
12	12	4.54	32	32	4.41	47	47	4.34								
13	13	4.63	33	33	4.50	48	48	4.51								
14	14	4.50	34	34	4.45	49	49	4.47								
15	15	4.46	35	35	4.69	50	50	4.43								
16	16	4.47	36	36	4.32											
17	17	4.56	37	37	4.55											
18	18	4.52	38	38	4.62											
19	19	4.47	39	39	4.51											
20	20	4.56	40	40	4.45											

Alternative hypothesis matters!

- H_0: The mean mass of Cu in the fifty 1 g samples of CRM analyzed equals 4.54 mg
- H_A: The mean mass of Cu in the fifty 1 g samples of CRM analyzed does not equal 4.54 mg
- Can we define HA in a different way?

- H_A: The mean mass of Cu in the fifty 1 g samples of CRM analyzed is greater than 4.54 mg
- H_A: The mean mass of Cu in the fifty 1 g samples of CRM analyzed is less than 4.54 mg

What to do with the data?

- Process in some software like SPSS, R, MATLAB, etc. You can use it like a black-box.
- Or use a calculator. You need to know the underlying formula.
- Is it safe to use software as black-box?
- Not recommended. Having understanding of the test can improve the interpretation of your results.

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

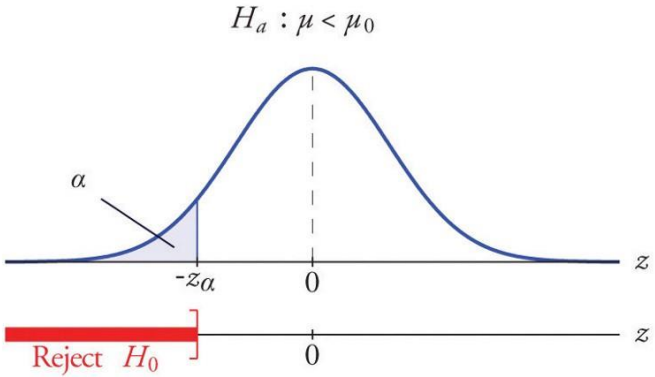
Tables are available for different tests!

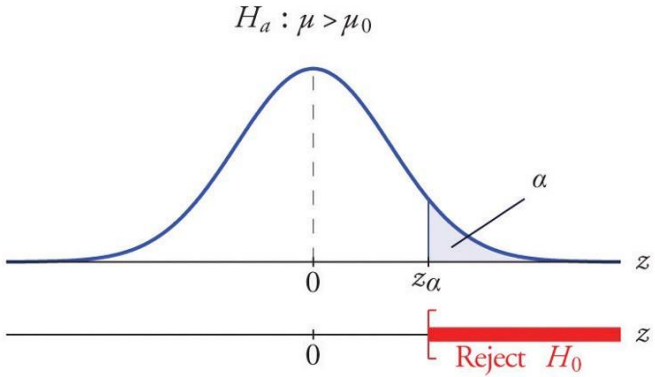
t Table

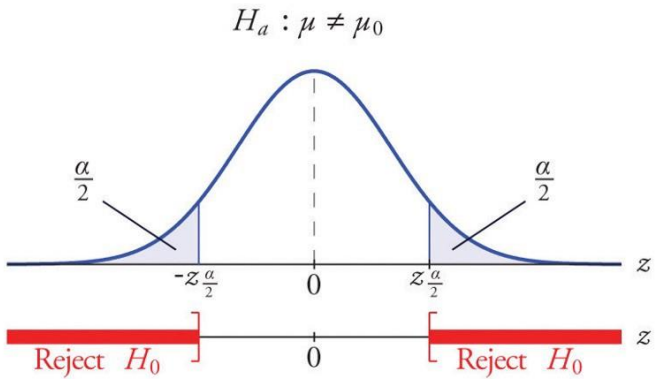
cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05
two-tails	1.00	0.50	0.40	0.30	0.20	0.10
df						
1	0.000	1.000	1.376	1.963	3.078	6.314
2	0.000	0.816	1.061	1.386	1.886	2.920
3	0.000	0.765	0.978	1.250	1.638	2.353
4	0.000	0.741	0.941	1.190	1.533	2.132
5	0.000	0.727	0.920	1.156	1.476	2.015
6	0.000	0.718	0.906	1.134	1.440	1.943
7	0.000	0.711	0.896	1.119	1.415	1.895
8	0.000	0.706	0.889	1.108	1.397	1.860
9	0.000	0.703	0.883	1.100	1.383	1.833
10	0.000	0.700	0.879	1.093	1.372	1.812

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.

One Tailed & Two Tailed Tests







Summary

- Two aspects: (i) Descriptive (ii) Inferential
- Inferential: (a) Parameter Estimation (b) Hypothesis Testing
- Hypothesis: Null (H_0) & Alternative (H_A)
- Test H_0 against H_A
- Type I & Type II errors
- Different tests available

References

Books:

- Practical Statistics for Data Scientists: 50 Essential Concepts by Peter Bruce & Andrew Bruce, O'Reilly Media, 2017.
- Biostatistics for the Biological and Health Sciences, Marc M. Triola, Mario Triola & Jason Roy, Pearson, 2017.

Software/Programming Languages:

- SPSS from IBM: https://www.ibm.com/in-en/analytics/spss-statistics-software
- R Project: https://www.r-project.org/

For further queries/suggestions, please feel free to write to shashikant.gupta@gdgoenka.ac.in

Thank You