# Tests of significance

Dr. Imran Zaheer

JRII

Dept. of Pharmacology

# outline

- Types of data
- Basic terms – Sampling Variation, Null hypothesis, P value
- Steps in hypothesis testing
- Tests  of significance and type
- SEDP
- Chi Square test
- Student t test
- ANOVA

# Types of Data

## Qualitative Data:

- Also called as enumeration data.

- Qualitative are those which can be answered as YES or NO, Male or Female, tall or short.

- Can  only measure their numbers with out unit of measurement eg: Number of males, Number of MDR-TB cases among TB patients, etc.

## Quantitative Data (measurement data)

- Quantitative are those which can be measured in numbers with measurement, like Blood pressure (mm of Hg), haemoglobin (gm%), height (cm) etc.

- A set of quantitative date can be expressed in mean and its standard deviation.

- Mean is the average of all variables in the data.

- Standard deviation is a measure of the distribution of variables around the mean

## Sampling Variation

- Research done on samples and not on populations.
- Variability of observations occur among different samples.
- This complicates whether the observed difference is due to biological or sampling variation from true variation.
- To conclude actual difference, we use tests of significance

# Null Hypothesis ($H_O$)

- 1st step in testing any hypothesis.
- Set up such that it conveys a meaning that there exists no difference between the different samples.
- Eg: Null Hypothesis – The mean pulse rate among the two groups are same (or) there is no significant difference between their pulse rates.

- By using various tests of significance we either:

–Reject the Null Hypothesis

(or)

–Accept the Null Hypothesis

- Rejecting null hypothesis $\rightarrow$ difference is significant.
- Accepting null hypothesis $\rightarrow$ difference is not significant.

# Level of Significance – "P" Value

- ***p*-value** is a function of the observed sample results (a statistic) that is used for testing a statistical hypothesis.
- It is the probability of null hypothesis being true. It can accept or reject the null hypothesis based on P value.
- Practically, $P < 0.05$ (5%) is considered significant.

- P = 0.05 implies,
  - We may go wrong 5 out of 100 times by rejecting null hypothesis.
  - Or, We can attribute significance with 95% confidence.

# Steps in Testing a Hypothesis

- **General procedure in testing a hypothesis**

1. Set up a null hypothesis ($H_O$).
2. Define alternative hypothesis ($H_A$).
3. Calculate the test statistic (t, $X^2$ Z, etc).
4. Determine degrees of freedom.
5. Find out the corresponding probability level (P Value) for the calculated test statistic and its degree of freedom.
6. This can be read from relevant tables.
7. Accept or reject the Null hypothesis depending on P value

# Test of significance

- The test which is done for testing the research hypothesis against the null hypothesis.

Why it is done?

-To assist administrations and clinicians in making decision.

- The difference is real ?
- Has it happen by chance ?

# Classification of tests of significance

**For Qualitative data:-**

1. Standard error of difference between 2 proportions ($SE_{p1-p2}$)
2. Chi-square test or $X^2$

**For Quantitative data:-**

1. Unpaired (student) ‘t’ test
2. Paired ‘t’ test
3. ANOVA

# Standard error of difference between proportions ($SE_{p1-p2}$)

- For comparing qualitative data between 2 groups.
- Usable for large samples only. (> 30 in each group).
- Calculate by ($SE_{p1-p2}$) using the formula

$$S.E.(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}.$$

- And then calculate the test statistic ‘Z Score’ by using the formula:

$$Z = \frac{P_1 - P_2}{SE_{P_1 - P_2}}$$

- If $Z > 1.96$ $P < 0.05$ SIGNIFICANT
- If $Z < 1.96$ $P > 0.05$ INSIGNIFICANT

# Example

- Consider a hypothetical study where cure rate of Typhoid fever after treatment with Ciprofloxacin and Ceftriaxone were recorded to be 90% and 80% among 100 patients treated with each of the drug.
- How can we determine whether cure rate of Ciprofloxacin is better than Ceftriaxone?

# Solution

## Step – 1: Set up a null hypothesis – H0:

– "There is no significant difference between cure rates of Ciprofloxacin and Ceftriaxone."

## Step – 2: Define alternative hypothesis – Ha:

– "Ciprofloxacin is 1.125 times better in curing typhoid fever than Ceftriaxone."

## Step – 3: Calculate the test statistic – 'Z Score'

## Step – 3: Calculating Z score

$$Z = \frac{P_1 - P_2}{SE_{P1-P2}}$$

- Here, P1 = 90, P2 = 80
- $SE_{P1-P2}$ will be given by the formula:

$$S.E.(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}.$$

So, Z = 10/5 = 2

**Step – 4: Find out the corresponding P Value**

– Since $Z = 2$ i.e., $> 1.96$, hence, $P < 0.05$

**Step – 5: Accept or reject the Null hypothesis**

– Since $P < 0.05$, So we **reject** the null hypothesis(H0)

<u>There is no significant difference between cure rates of Ciprofloxacin and Ceftriaxone</u>

– And we **accept** the alternate hypothesis (Ha)

**<u>that, Ciprofloxacin is 1.125 times better in curing typhoid fever than Ceftriaxone</u>**

# Chi square ($X^2$) test

- This test too is for testing qualitative data.
- Its advantages over SEDP are:
    - Can be applied for smaller samples as well as for large samples.

- Prerequisites for Chi square ($X^2$) test to be applied:
    - The sample must be a random sample
    - None of the observed values must be zero.
    - Adequate cell size

## Steps in Calculating ($X^2$) value

1. Make a contingency table mentioning the frequencies in all cells.
2. Determine the expected value (E) in each cell.

$$E = \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}} \qquad \frac{r \times t}{T}$$

3. Calculate the difference between observed and expected values in each cell (O-E).

- Calculate $X^2$ value for each cell

$$\chi^2 \text{ of each cell} = \frac{(O - E)^2}{E}$$

- Sum up $X^2$ value of each cell to get $X^2$ value of the table.

$$\chi^2 = \Sigma \frac{(O - E)^2}{E}$$

# Example

- Consider a study done in a hospital where cases of breast cancer were compared against controls from normal population against possession of a family history of Ca Breast.
- 100 in each group were studied for presence of family history.
- 25 of cases and 15 among controls had a positive family history.
- Comment on the significance of family history in breast cancer.

## Solution

- From the numbers, it suggests that family history is 1.66 (25/15) times more common in Ca breast.

- So is it a risk factor in population

- We need to test for the significance of this difference.

- We shall apply $X^2$ test.

# Solution

**Step – 1: Set up a null hypothesis**

– H0: “There is no significant difference between incidence of family history among cases and controls.”

**Step – 2: Define alternative hypothesis**

**– Ha: “Family history is 1.66 times more common in Ca breast”**

**Step – 3: Calculate the test statistic – $X^2$**

## Step – 3: Calculating $X^2$

1. Make a contingency table mentioning the frequencies in all cells

| Group | Risk factor (Family History) | | Total |
|---|---|---|---|
| | Present | absent | |
| Cases | 25 (a) | 75 (b) | 100 |
| Controls | 15 (c) | 85 (d) | 100 |
| Total | 40 | 160 | 200 |

## *Step – 3: Cont..*

2. Determine the expected value (E) in each cell.

$$E = \frac{\text{Row total x Column total}}{\text{Grand total}} \qquad \frac{r \times t}{T}$$

- For (a), $E_a = 100 \times 40 / 200 = 20$
- For (b), $E_b = 100 \times 160 / 200 = 80$
- For (c), $E_c = 100 \times 40 / 200 = 20$
- For (d), $E_d = 100 \times 160 / 200 = 80$

| | O | E | (O – E) | $(O-E)^2$ | $\frac{(O-E)^2}{E}$ |
|---|---|---|---|---|---|
| **a** | 25 | 20 | 5 | 25 | 1.25 |
| **b** | 75 | 80 | -5 | 25 | 0.3125 |
| **c** | 15 | 20 | 5 | 25 | 1.25 |
| **d** | 85 | 80 | -5 | 25 | 0.3125 |
| | | | | $\chi^2$ | = **3.125** |

## Step – 4: Determine degrees of freedom.

- DoF is given by the formula:

$$\mathbf{DoF = (r\text{-}1) \times (c\text{-}1)}$$

where r and c are the number of rows and columns respectively

- Here, $r = c = 2$.
- Hence, $DoF = (2-1) \times (2-1) = 1$

## Step – 5: Find out the corresponding P Value

– P values can be calculated by using the $\mathbf{X^2}$ distribution tables

| | 0.95 | 0.90 | 0.80 | 0.70 | 0.50 | 0.30 | 0.20 | 0.10 | 0.05 | 0.01 | 0.001 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.004 | 0.02 | 0.06 | 0.15 | 0.46 | 1.07 | 1.64 | 2.71 | 3.84 | 6.64 | 10.83 |
| 2 | 0.10 | 0.21 | 0.45 | 0.71 | 1.39 | 2.41 | 3.22 | 4.60 | 5.99 | 9.21 | 13.82 |
| 3 | 0.35 | 0.58 | 1.01 | 1.42 | 2.37 | 3.66 | 4.64 | 6.25 | 7.82 | 11.34 | 16.27 |
| 4 | 0.71 | 1.06 | 1.65 | 2.20 | 3.36 | 4.88 | 5.99 | 7.78 | 9.49 | 13.28 | 18.47 |
| 5 | 1.14 | 1.61 | 2.34 | 3.00 | 4.35 | 6.06 | 7.29 | 9.24 | 11.07 | 15.09 | 20.52 |
| 6 | 1.63 | 2.20 | 3.07 | 3.83 | 5.35 | 7.23 | 8.56 | 10.64 | 12.59 | 16.81 | 22.46 |

## Step – 6: Accept or reject the Null hypothesis

- In given scenario,

$$\mathbf{X^2} = 3.125$$

- This is less than 3.84 (for P = 0.05 at dof =1)
- Hence Null hypothesis is Accepted, i.e.,

**"There is no significant difference between incidence of family history among cases and controls"**

# Student’s ‘t’ test

- Very common test used in biomedical research.
- Applied to test the significance of difference between two means.
- It has the advantage that it can be used for small samples.

**Types**

—Unpaired ‘t’ test

—Paired ‘t’ test.

# Unpaired 't' test

- This is used to test the difference between two independent samples
- Test statistic is

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{(\bar{x}1 - \bar{x}2)}}$$

$S\,(\bar{x}_1 - \bar{x}_2)$ Standard error of difference between two means.

$$S(\bar{x}_1 - \bar{x}_2) = S\sqrt{\frac{1}{n1} + \frac{1}{n2}}$$

where S is the combined SD

$$\text{Combined SD } (S) = \sqrt{\frac{S_1^2\,(n_1-1) + S_2^2\,(n_2-1)}{n_1 + n_2 - 2}}$$

where $S_1$ and $S_2$ are Standard deviations and $n_1$ and $n_2$ are the sample sizes of 1st and 2nd samples respectively.

# Example

- The chest circumference (in cm) of 10 normal and malnourished children of the same age group is given below:

Normal : 42, 46, 50, 48, 50, 52, 41, 49, 51, 56

Malnourished : 38, 41, 36, 35, 30, 42, 31, 29, 31, 35

Find out whether the mean chest circumference is significantly different in the two groups (table ‘t’ value at 18 df at 0.05 level of significance is 2.02)

# Solution

- In the given example, following are the values of mean, SD and sample size

| | Mean | SD | N |
|---|---|---|---|
| First sample | 48.5 | 4.53 | 10 |
| Second sample | 34.8 | 4.57 | 10 |

By substituting the values, we get

(a) $$\text{Combined SD (S)} = \sqrt{\frac{4.53^2\ (10-1) + 4.57^2\ (10-1)}{10 + 10 - 2}}$$

$$\sqrt{20.70} = 4.55 \text{ cm}$$

Solution contd.

(b) Thus $S_{(\bar{x}_1 - \bar{x}_2)} = 4.55 \sqrt{\frac{1}{10} + \frac{1}{10}}$

$$= 4.55 \sqrt{0.2}$$

$$= 4.55 \text{ X } 0.44 = 2.03 \text{ cm}$$

Solution contd.

(c) Calculation of t value

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{(\bar{x}1 - \bar{x}2)}}$$

$$= \frac{48.5 - 34.8}{2.03}$$

$$= 6.75$$

Solution contd.

(d) Degrees of freedom (df) = n1 + n2 -2

= 10 + 10 – 2 = 20 -2 = 18.

- The maximum ‘t’ value for 18 df for 0.05 level of significance is 2.02 (from table)

Solution contd.

**(e) Conclusion:** As the calculated ‘t’ value is higher than the table value for 18 df and 0.05 level of significance,

— reject the null hypothesis and state that the difference in the mean chest circumference values between normal and malnourished groups is statistically significant.

— The difference might not have occurred due to chance (P< 0.05)

# Paired ‘t’ test

- This test is used where each person gives two observations – before and after values

- For example, if the same person gives two observations before and after giving a new drug – in such cases, this test is used.

# Example

- In a clinical trial to evaluate a new drug in reducing anxiety scores of 10 patients,
- the following data are recorded.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Before treatment | 22 | 18 | 17 | 19 | 22 | 12 | 14 | 11 | 19 | 7 |
| After treatment | 19 | 11 | 14 | 17 | 23 | 11 | 15 | 19 | 11 | 8 |

- Find out whether the drug is significantly effective in reducing anxiety score
- (Table ‘t’ value at 9 df at 0.05 level of significance is 2.21)

| Before treat. $(X_1)$ | After treat. $(X_2)$ | Differ. $(X_1 - X_2) = x$ | $\bar{x}$ | $(x - \bar{x})$ | $(x - \bar{x})^2$ |
|---|---|---|---|---|---|
| 22 | 19 | + 3 | 1.3 | + 1.7 | 2.89 |
| 18 | 11 | +7 | 1.3 | + 5.7 | 32.49 |
| 17 | 14 | +3 | 1.3 | + 1.7 | 2.89 |
| 19 | 17 | + 2 | 1.3 | + 0.7 | 0.49 |
| 22 | 23 | - 1 | 1.3 | - 2.3 | 5.29 |
| 12 | 11 | + 1 | 1.3 | - 0.3 | 0.09 |
| 14 | 15 | - 1 | 1.3 | - 2.3 | 5.29 |
| 11 | 19 | -8 | 1.3 | -9.3 | 86.49 |
| 19 | 11 | + 8 | 1.3 | + 6.7 | 44.89 |
| 7 | 8 | -1 | 1.3 | - 2.3 | 5.29 |
| 161 | 148 | + 13 | +13 | --- | 186.1 |

Solution contd.

## Calculation of Mean of difference and SD

$$\text{Mean} = \frac{\Sigma x}{n} = +13 / 10 = +1.3$$

$$\text{SD} = \sqrt{\frac{\Sigma(x - \bar{x})2}{n-1}}$$

$$= \sqrt{186.1 / 9} = 4.54$$

Calculation of Standard error of mean

(c) $SE = SD / \sqrt{n}$

$= 4.54 / \sqrt{10}$

$= 4.54 / 3.16 = 1.43$

(d) Calculation of t value

$t = x / SE = 1.3 / 1.43 = 0.90$

(e) Calculation of df = n-1 = 10 -1 = 9

# Solution contd.

**Conclusion:** As the calculated ‘t’ value is lower than the table value for df -9 at 5% level (0.05 level) of significance,

- we accept the null hypothesis and conclude that there is no statistically significant difference in the anxiety scores before and after treatment.
- The difference might have occurred because of chance ($P>0.05$; NS)

# ANOVA (*Analysis Of Variance*)

# How ANOVA works

ANOVA measures two sources of variation in the data and compares their relative sizes

- variation BETWEEN groups
  - for each data value look at the difference between its group mean and the overall mean

$$\left(\bar{x}_i - \bar{x}\right)^2$$

- variation WITHIN groups
  - for each data value look at the difference between that value and the mean of its group

$$\left(x_{ij} - \bar{x}_i\right)^2$$

The ANOVA F-statistic is a ratio of the Between Group Variaton divided by the Within Group Variation:

$$\boxed{F = \frac{Between}{Within} = \frac{MSG}{MSE}}$$

A large F is evidence *against* $H_0$, since it indicates that there is more difference between groups than within groups.

THANK YOU