Tests of Significance

• Z test	For a single mean or percentage Difference between two means Difference between two percentages
• X^2 test	Association between two categorical variables
• t test	Means of two small samples

Standard Error of Mean
(Standard deviation of sample means)

- Depends on

  Variability – Standard Deviation

  Sample size

Tests of Significance

- State the null hypothesis
- Find the value of the test statistic
- Refer table & find the probability of the test statistic when the null hypothesis is true
- Conclude that the data are consistent or inconsistent with the null hypothesis

In a clinical trial 75.7% (56 out of 74) of patients receiving the new treatment got better in 4 days.

Only 65.3% (45 out of 69) of patients receiving placebo got better in 4 days.

Can we conclude that the new treatment is better than placebo?

- Null hypothesis – There is no difference between the new treatment and placebo with regard to the rate of improvement

- Difference =75.7 – 65.3 = 10.5%
- SE of difference = 7.6

- Difference =75.7 – 65.3 = 10.5%
- SE of difference = 7.6
- Z = 10.5/7.6 = 1.19
- P = 0.233
- Fail to reject the null hypothesis
- In this trial there isn’t enough evidence to say that the rate of improvement in the treatment group is greater than the rate of improvement in the placebo group

- The mean haemoglobin of a random sample of 100 pregnant women was 11g/dl (standard deviation=2). The mean haemogobin of adult females is 12g/dl. Haemoglobin is known to have a Normal distribution.
  Can we conclude that pregnant women have a lower haemoglobin level than non pregnant women?

- Among a group of 250 children with hookworm infection 200 were anaemic.
- In the same school among 250 non infected children 180 were anaemic.
- Can this be taken as evidence for an association between hookworm infection and anaemia?

Standard error of difference between two percentages

$$SE_{diff} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

q=100 - p

Standard error of difference between two percentages

$$\sqrt{(76\times 24)\left(\frac{1}{250}+\frac{1}{250}\right)}=3.5$$

Null hypothesis Difference = 0

Observed difference = 80-72 = 8

$Z = 8/3.5 = 2.286$

$P = 0.022$

Marks by sex

Sex	N	Mean	SD
Female	80	61	7
Male	70	57	7

Tests of Significance

- State the null hypothesis
- Find the value of the test statistic
- Refer table & find the probability of the test statistic when the null hypothesis is true
- Conclude that the data are consistent or inconsistent with the null hypothesis

Standard error of difference between two means

$$SE_{diff} = \sqrt{\frac{SD_1^{\,2}}{n_1} + \frac{SD_2^{\,2}}{n_2}}$$

SD – Standard Deviation;   n – Sample Size

Standard error of difference between two means

$$SE_{diff} = \sqrt{\frac{7^2}{80} + \frac{7^2}{70}}$$

$$= 1.15$$

Null hypothesis Difference = 0

Observed difference = 61 - 58 = 3

Z = 3/1.15 = 2.61

P = 0.009

Errors associated with significance tests

Errors associated with significance tests

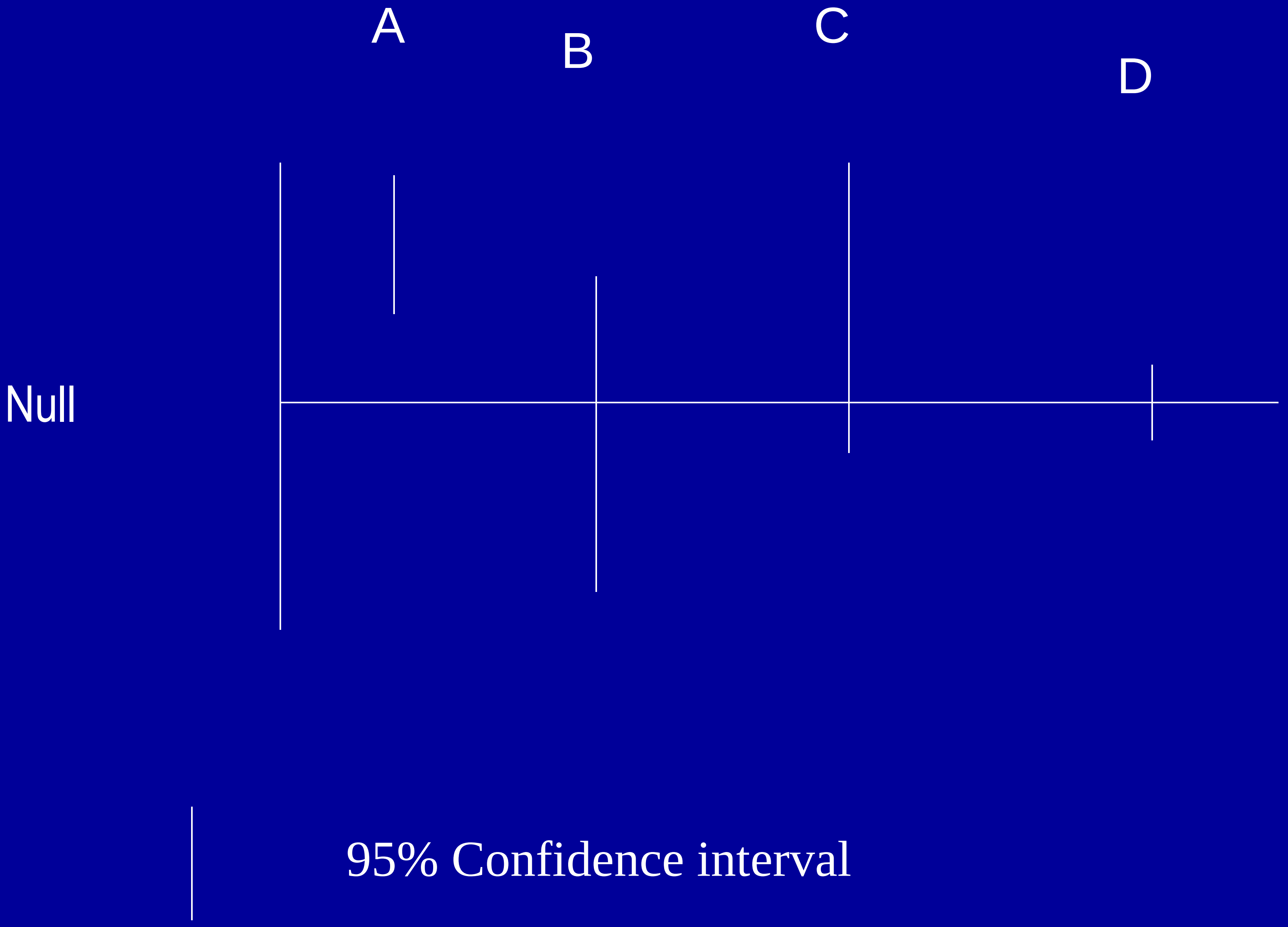
- Rejecting a true null hypothesis
  - Type I error (α error)
  - Significance
  - The probability of this is given by the P value

Errors associated with significance tests

- Failing to reject a false null hypothesis
  - Type II error (β error)
  - 1 – β is known as the power of a test

		Truth in Population	
		Association between the two variables	No association between the two variables
Results in the sample	Reject null hypothesis	Correct	Type I error
	Fail to reject null hypothesis	Type II error	Correct

Relationship between significance tests and confidence intervals


A
B
C
D
Null
95% Confidence interval

Chi squared test

$$X^2 = \sum \frac{(O-E)^2}{E}$$

df = (c –1) x (r – 1)

df – degrees of freedom

Chi squared table

	Probability		
Df	0.1	0.05	0.01
1	2.71	3.84	6.64
2	4.61	5.99	9.21
3	6.25	7.82	11.34

A trial was conducted to assess two methods of treatments X and Y for a particular disease.

Out of the 257 patients treated by X 41 died; out of the 244 treated by Y 64 died.

Can this be taken as good evidence regarding the superiority of X over Y as treatment for this disease

Outcome by mode of treatment

Treatment	Died		Survived		Total
	N	%	N	%	N
X	41	16	216	84	257
Y	64	26	180	74	244
Total	105	21	396	79	501

Outcome by mode of treatment

	Treatment X		Treatment Y	
	N	%	N	%
Survived	41	16	64	26
Died	216	84	180	74
Total	257	100	244	100

Observed and expected frequencies

	O	E	(O-E)	$(O-E)^2$	$(O-E)^2/E$
A	41	53.8	-12.8	163.8	3.0
B	64	51.3	12.7	161.3	3.1
C	216	203.1	12.9	166.4	0.8
D	180	192.9	-12.9	166.4	0.9

$$E = (RT \times CT) / GT$$

Results of Chi squared test

$X^2 = (3.0 + 0.8 + 3.1 + 0.9) = 7.8$

$Df = (2 - 1) \times (2 - 1) = 1$

P value < 0.01

Sample size and X^2 test

For the test to be valid

- All expected frequencies should exceed 1
- 80% of expected frequencies should exceed 5
- Minimum sample size 20
- Cell values should be counts

Outcome by mode of treatment

	Treatment X		Treatment Y	
	N	%	N	%
Survived	41	16	64	26
Died	216	84	180	74
Total	257	100	244	100

Alternative method using standard error of difference between two percentages

- How likely is it to get death rates of 16% and 26% when taking two samples of size 257 and 244 from the same population?

$$SE = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

$$q = 100 - p$$

$$SE = \sqrt{\frac{16 \times 84}{257} + \frac{26 \times 74}{244}}$$

SE = 3.62

Test of significance Vs Confidence Intervals

Homocystine and Alzheimer's Disease

	Mean	SD
Alzheimer's (n=21)	12.3	5.4
Controls (n=19)	8.3	2.3

Tests of Significance

- State the null hypothesis
- Find the value of the test statistic
- Refer table & find the probability of the test statistic when the null hypothesis is true
- Conclude that the data are consistent or inconsistent with the null hypothesis

Standard Error of Difference

- Standard deviation of the differences between sample means is called the standard error of difference between two means

$$SE_{diff} = \sqrt{\frac{SD_1^2}{n_1} + \frac{SD_2^2}{n_2}}$$

Z test

- Only sample means from large samples form a Normal distribution

- What is a large sample?

  Usually about 60 or more

  Absolute minimum is 30

t test

- Standard error = 1.33
- T statistic =(12.3 – 8.3)/ 1.33 = 2.99
- Degrees of freedom = 21 + 19 – 2 = 38

t table

	Probability		
DF	0.10	0.05	0.01
1	6.31	12.7	63.66
2	2.92	4.30	9.93
30	1.70	2.04	2.75
38	1.69	2.02	2.71
60	1.67	2.00	2.66
120	1.66	1.98	2.62

t test

- Standard error = 1.33
- T statistic =(12.3 – 8.3)/ 1.33 = 2.99
- Degrees of freedom = 21 + 19 – 2 = 38
- P = 0.005

Table: Anxiety scores recorded for 10 patients receiving new drug and placebo

Drug	Placebo
19	22
11	18
14	17
17	19
23	22
11	12
15	14
12	19
11	19
08	07

	Mean	SD
New Drug	14.1	4.5
Placebo	16.9	4.7

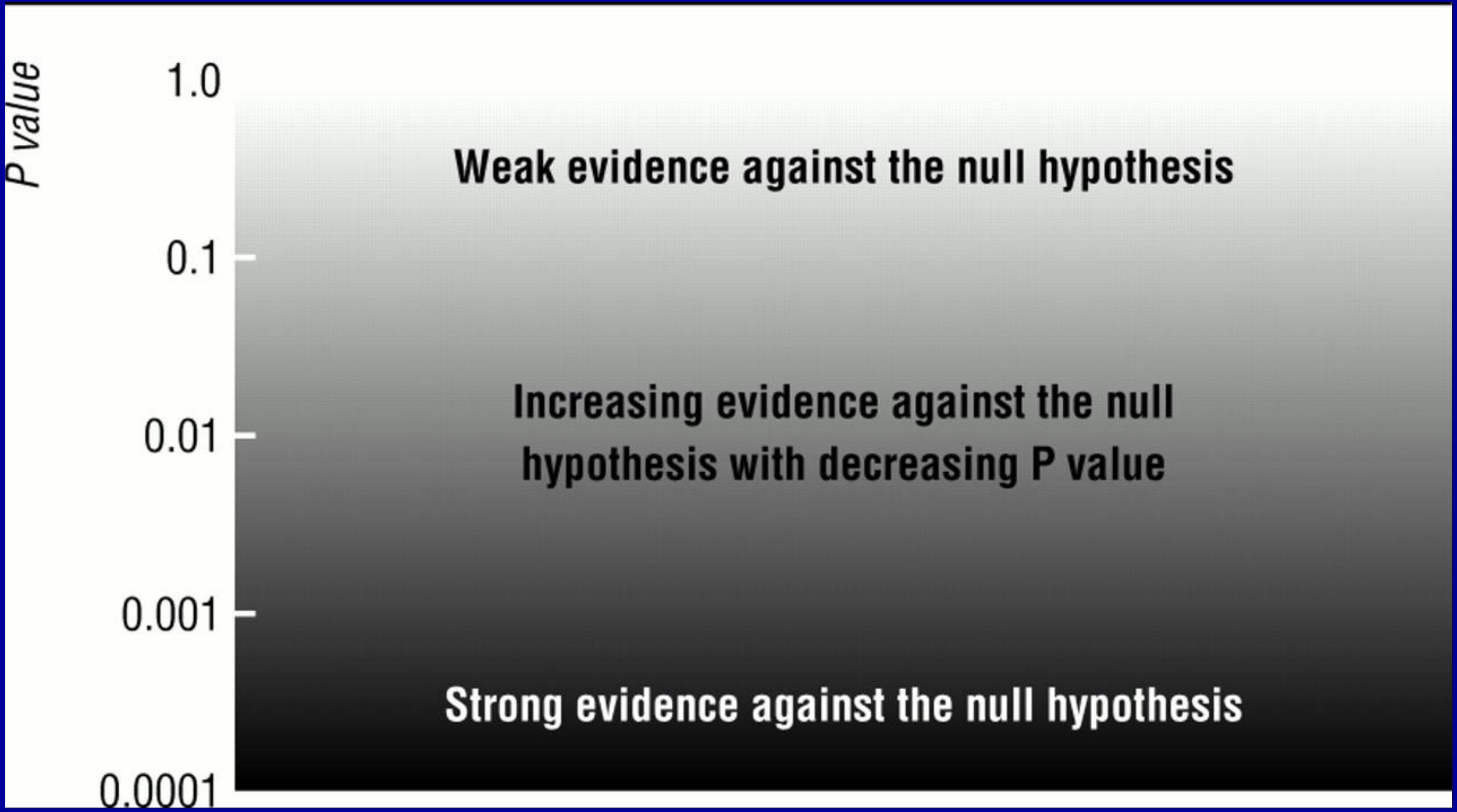
t test

- Standard error of difference = 2.05
- t statistic =(14.1 – 16.9)/ 2.05 = - 1.36
- Degrees of freedom = 10 + 10 – 2 = 18
- P = 0.19

	Mean	SD
New Drug	14.1	4.5
Placebo	16.9	4.7
Difference	-2.8	3.5

Paired t test

- Standard error = 1.10
- t statistic = -2.8/1.10 = -2.53
- Degrees of freedom = 10 – 1 = 9
- P = 0.03



Suggested interpretation of P values

Fuel consumption (mpg) of cars in the US by place of manufacture

Origin	mean	median	sd
Domestic	20	19	4.7
Foreign	25	25	6.6

Fuel consumption (mpg) of cars in the US by place of manufacture

Origin	obs	rank sum	expected
Domestic	52	1688.5	1950
Foreign	22	1086.5	825
combined	74	2775.0	2775

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

$Z = 3.101$; $p = 0.002$