

Parametric and Non-Parametric Tests

Learning objectives

- ▶ To Understand the difference between Parametric and non-parametric tests.
- ▶ To perform and interpret one-sample sign test
- ▶ To perform and interpret Wilcoxon Signed-Rank Test for paired samples
- ▶ To perform and interpret Mann-Whitney U Test for Independent Samples
- ▶ To perform and interpret Kruska-Wallis Test

Parameters of distribution

- ▶ Mean
- ▶ Standard deviation
- ▶ Proportion
- ▶ Non-parametric test are also called distribution free test.

Parametric and Non-Parametric Tests

- ▶ Both can be applied for hypothesis testing.
- ▶ Both are statistical tests applied for drawing inferences .
- ▶ The assumptions are different for applying these test.

	Parametric Tests	Non – Parametric Tests
Assumptions	Normality is required	Normality is not required
	Uses the metric data	Nominal or Ordinal scale data is used.
	Can be applied for both small and large samples	Can be applied for small samples
Applications	One sample using Z or t statistics	One sample using sign test
	Two independent samples using a t or Z-test .	Two independent samples using the Mann-Whitney U statistics
	Two paired samples using a t test.	Two paired samples using the Wilcoxon Signed-rank test .
	Several independent samples using F test in ANOVA	Several independent samples using Kruskal -Wallis test

One Sample Sign Test

- ▶ As data is converted to plus and minus sign, this test is called sign test.
- ▶
- ▶ This test is used to estimate the median of a population and compare it to a reference value or target value.
- ▶ can be used to identify the preference related to brand or any other product/goods or services.

Illustration

- ▶ A survey was conducted to understand the preference for fast food by the inhabitants of a small town . A sample of 100 respondents indicated that 54 do not prefer fast food whereas 46 have a preference for the fast food . At 5 % level of significance test the hypothesis that half of the inhabitants of the town prefer fast food.
- ▶ Solution :
- ▶ $H_0 = p=1/2$ p = proportion not preferring fast food
- ▶ $H_1 = p \neq 1/2$
- ▶ Plus sign (those not preferring fast food, 54 plus sign)
- ▶ Minus sign (those preferring fast food ,46 minus sign)
- ▶ $Z = \frac{54 - .5 \times 100}{.5 \sqrt{100}} = 0.7$

- ▶ As the critical value of z at 5% level of significance is ± 1.96 and the calculated value is less than the critical value ,thus null hypothesis cannot be rejected .
- ▶ Interpretation :The proportion of inhabitants not preferring fast food is not significantly different from the ones preferring fast food.

Wilcoxon Signed-Rank Test for paired samples

- ▶ Test is applied to compare paired samples.
- ▶ Analogous to dependent t-test or paired t-test.
- ▶ The rank sum of the positive and negative differences are calculated, the smallest of these is used as the test statistics to test the hypothesis .

Illustration:

S.No. (Sales men)	Score before Training	Score after Training	Difference	Absolute difference	Rank of absolute difference	Negative rank	Positive rank
1	85	82	-3	3	7.5	7.5	
2	76	79	3	3	7.5		7.5
3	64	68	4	4	11		11
4	59	52	-7	7	13.5	13.5	
5	72	75	3	3	7.5		7.5
6	68	69	1	1	2.5		2.5
7	43	40	-3	3	7.5	7.5	
8	54	53	-1	1	2.5	2.5	

S.No. (Sales men)	Score before Training	Score after Training	Difference	Absolute difference	Rank of absolute difference	Negative rank	Positive rank
9	57	50	-7	7	13.5	13.5	
10	61	67	6	6	12		12
11	71	74	3	3	7.5		7.5
12	82	83	1	1	2.5		2.5
13	39	54	15	15	16		16
14	51	59	8	8	15		15
15	54	51	-3	3	7.5	7.5	
16	57	58	1	0	2.5		2.5

- ▶ H_0 = There is no difference in the appraisal score because of training
- ▶ H_1 = There is a difference in the appraisal score because of training.

- ▶ Total positive Ranks = 84 = T_+
- ▶ Total Negative Ranks = 52 = T_-
- ▶ $T = \min(T_+, T_-) = 52$

- ▶ $\mu = n(n+1)/4 = 16 \times 17 / 4 = 68$

$$\sigma = \sqrt{16 \times 17 \times 33 / 24}$$

- ▶ = 19.34
- ▶ Test statistics $z = -0.83$

$$\sigma = \sqrt{n(n+1)(2n+1)/24}$$

$$Z = T - \mu T / \sigma T$$

- ▶ As the value of Z is -0.83 which is less than the critical value at 5 % level of significance that is ± 1.96 hence the null hypothesis cannot be rejected .
- ▶ Interpretation : There is no change in the performance appraisal score because of training.
- ▶ Explanations about the interpretation is very important .

Mann-Whitney U test for independent samples

- ▶ **Also known as Mann–Whitney–Wilcoxon (MWW), Wilcoxon rank-sum test, or Wilcoxon–Mann–Whitney test**
- ▶ It is a non-parametric counterpart of the t-test used to compare the means of two independent populations.
- ▶ The assumptions underlie the use of this test
 - ▶ - The samples are independent
 - ▶ - The level of data is at least ordinal
- ▶ H_0 = Two samples come from identical populations.
- ▶ H_1 = Two samples come from different populations .

Illustration

- ▶ Data on weekly expenditure on entertainment by 14 BA students of college A and 16 of college B. Check at 1 % level of significance that there is no difference in the average expenditure of the students of college A and B.

College A	250	300	350	180	280	260	400	190	320
College B	380	130	400	450	360	270	500	480	450

340	370	160	500	550					
470	500	550	575	470	480	220			

Weekly Expenditure (in INR)	College	Rank
130	B	1
160	A	2
180	A	3
190	A	4
220	B	5
250	A	6
260	A	7
270	B	8
280	A	9
300	A	10
320	A	11
340	A	12
350	A	13
360	B	14
370	A	15

Weekly Expenditure (in INR)	College	Rank
380	B	16
400	A	17.5
400	B	17.5
450	B	19.5
450	B	19.5
470	B	21.5
470	B	21.5
480	B	23.5
480	B	23.5
500	A	26
500	B	26
500	B	26
550	A	28.5
550	B	28.5
575	B	30

- ▶ R1= sum of Ranks of college A=164
- ▶ R2= Sum of Ranks of college B = 301
- ▶ n1= 14

Using this formula Z –statistics is calculated

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

where

$$U = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - R$$

- Interpretation – the calculated z value in this case is -2.203 which is less than the critical value at 1 % ± 2.57 . Hence it can be concluded that there is no significant difference in the average expenditure on entertainment by the students of two colleges.

Kruskal –Wallis Test

- ▶ Like on way ANOVA , this test is used to determine the $C \geq 3$ samples come from the same population or different.
- ▶ This test determines whether all the groups come from the same population or whether at least one group comes from a different population.
- ▶ The process of computing this test begins with ranking the data in all groups together as though they were from one group .
- ▶ The hypothesis tested by this test is as
 - ▶ H_0 = The C populations are identical
 - ▶ H_1 = atleast two of the populations are different

Decision to choose test

- ▶ Are the samples Independent ?
- ▶ If yes apply Mann whitney U test
- ▶ If No;
- ▶ Are the data at least interval ?
- ▶ If yes apply Wilcoxon Signed Rank Test
- ▶ If No;
- ▶ apply Sign Test

Chi- square test

- ▶ Data is required in the form of frequencies.
- ▶ Majority of the applications of chi-square are with the discrete data .
- ▶ It can also be applied to continuous data, provided it is reduced to certain categories and tabulated in such a way that the chi square may be applied.
- ▶ The Chi square (χ^2) is not a measure of the degree of relationship between the variables under study.

Applications of chi-square

- ▶ Test for goodness of fit –
Examples : preference, occurrence of an event .
- ▶ Test for independence of variables-
Examples : Training and performance , extra classes and results .
- ▶ Test for equality of more than two population proportions –
Examples : Proportion of satisfied employees in various class categories

Advantages and Disadvantages of Non-Parametric Test

- Need not involve population parameters
 - Easy to compute
 - Results may be exact as parametric procedures.
 - Make fewer assumptions.
-
- Tables not widely available
 - Difficulty to compute by hand for large samples .

- ▶ It is important to analyse data but equally it is important to interpret it wisely and in detail.
- ▶ Thus select the statistical analysis tool with proper knowledge and in relation with the objectives your study is aiming to achieve .

Thank You

- ▶ <https://www.slideshare.net/AnchalGarg8/non-parametric-statistics-70145722>
- ▶ Research Methodology Concepts and cases by Dr. Deepak Chawla and Dr. Neena Sondhi ,vikas publishing house pvt. Ltd. ,2015
- ▶ <https://www.chegg.com/homework-help/using-mann-whitney-u-test-mann-whitney-u-test-equivalent-wil-chapter-13.4-problem-13bb-solution-9780321694508-exc>