Statistical Inference

Nonparametric Tests

DATA SCIENCE INSTITUTE

Nonparametric Statistical Tests

- Tests based on t and F distribution assume that populations are normally distributed.
- A large body of statistical methods is available which do not make assumptions about the nature of the distribution(e.g. normality)
- These testing procedures are termed as Nonparametric tests or distribution-free tests.
- If the underlying assumptions of the parametric test are met, then the parametric test will be more powerful than nonparametric test.

Mann-Whitney test

- The Mann-Whitney test is considered as nonparametric alternative to t test for independent samples.
- The Mann-Whitney U test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed.
- H_0: The two samples come from the same population
- H_1: Not H_0.
- The test is equivalent to Wilcoxon rank-sum test (WRS).

DATA SCIENCE INSTITUTE

Mann-Whitney test...continued

- The null hypothesis is that the distributions of both groups are identical, so that there is a 50% probability that an observation randomly selected from one population exceeds an observation randomly selected from the other population.

- The Mann-Whitney test compares the distributions of ranks in two groups. If you assume that both populations have distributions with the same shape (which doesn't have to be Normal), it can be viewed as a comparison of two medians. Note that if you don't make this assumption, the Mann-Whitney test does not compare medians.

Mann-Whitney test
Procedure

1. Combine the two samples.
2. Rank all the observations from smallest to largest.
3. Keep track of the group to which each observation belongs.

\# Tied observations are assigned a rank equal to the mean of the rank positions for which they are tied.

DATA SCIENCE INSTITUTE

Mann-Whitney test Procedure..

- The test statistic is

$$U = T - \frac{m(m+1)}{2}$$

Where T is sum of the ranks of first sample in combined ordered sample, m and n are sample sizes.

$$E(U) = \frac{mn}{2} \qquad V(U) = \frac{mn(m+n+1)}{12}$$

- Standardized U is assumed to follow normal distribution.

- Compare p-value with level of significance & conclude.

DATA SCIENCE INSTITUTE

Mann-Whitney test
Example

GroupI	GroupII
11.5	11
23.6	25
12.1	26.3
14.5	12.2
16.1	18.5
19.8	15
	15.7

Ordered Sample	Group	Rank
11	GroupII	1
11.5	GroupI	2
12.1	GroupI	3
12.2	GroupII	4
14.5	GroupI	5
15	GroupII	6
15.7	GroupII	7
16.1	GroupI	8
18.5	GroupII	9
19.8	GroupI	10
23.6	GroupI	11
25	GroupII	12
26.3	GroupII	13

T is sum of the ranks of first sample in combined ordered sample
m and n are sample sizes.

T=39, m=6 and n=7

DATA SCIENCE INSTITUTE

Mann-Whitney test Example…Continued

- The test statistic is

$$U = T - \frac{m(m+1)}{2}$$

$$E(U) = \frac{mn}{2} \qquad V(U) = \frac{mn(m+n+1)}{12}$$

- Here U= 18 E(U)= 21 V(U)= 49

- p value=0.7308 (exact as obtained in R)

- Conclusion: Do not reject H0

Mobile Consumer Behaviour

- The Objective of the study is to understand factors driving buying behaviour of potential customers. One of the factor is ‘Colour’ of mobile phone.

  Factor rating is measured on Likert scale. (1-5)

  1: Least Important 5:Most Important

  Rating indicates importance given to ‘Colour’ for buying a mobile phone.

- Research question: Is there a difference between ratings given by 'male' and 'female' customers?

DATA SCIENCE INSTITUTE

Data Snapshot

resid	Gender	Color
1	M	3
2	M	3
3	M	1
4	M	1
5	F	4
6	F	4
7	F	4
8	F	5
9	M	1
10	M	5
11	M	3
12	M	3
13	F	2
14	F	4
15	F	5
16	F	1
17	F	1
18	F	2
19	M	2

Number of respondents: 73
Rating given to 'Colour' is measured on Likert scale (1-5)

DATA SCIENCE INSTITUTE

Mann-Whitney test using R

```
#import csv data set 'Mobile Consumer Behaviour'
# 'Color' is analysis variable and 'Gender' is factor
mobile_color<-read.csv(file.choose(),header=T)
wilcox.test(formula=Color~Gender,data=mobile_color)
```

```
Wilcoxon rank sum test with continuity correction

data:  Color by Gender
W = 788.5, p-value = 0.1319
alternative hypothesis: true location shift is not equal to 0
#Inference: Do not reject H0 which suggests that ratings given to
'colour' for buying mobile phone are same for males and
females
```

DATA SCIENCE INSTITUTE

Wilcoxon Signed Rank test

- The Wilcoxon Signed Rank test is considered as nonparametric alternative to paired t test .

- The Wilcoxon Signed Rank test is used to compare differences between two related or paired groups when the variable is either ordinal or continuous, but not normally distributed.

- H_0: The median of difference in the population is zero
- H_1: Not H_0.

Wilcoxon Signed Rank test Procedure

1. Define $D_i = X_i - Y_i$ which are the differences between two values for each pair.
2. Obtain $|D_i|$ which are absolute values of differences.
3. Rank all $|D_i|$ from smallest to largest.
4. Define R_i = rank of $|Di|$.
5. Obtain 'W' which is sum of the ranks associated with either positive D_i.

Wilcoxon Signed Rank test Procedure..

- The test statistic is W: which is sum of the ranks associated with positive D_i.

$$E(W) = \frac{n(n+1)}{4}$$

$$V(W) = \frac{n(n+1)(2n+1)}{24}$$

- Standardized W is assumed to follow normal distribution.

- Compare p-value with level of significance & conclude.

DATA SCIENCE INSTITUTE

Wilcoxon Signed Rank test Example

Score before and after training

Before	After
58	74
52	65
61	60
48	45
50	58
39	53
44	53
47	49
57	61
55	67
49	60
51	61

Di	Abs(Di)	Rank(Di)
-16	16	12
-13	13	10
1	1	1
3	3	3
-8	8	5
-14	14	11
-9	9	6
-2	2	2
-4	4	4
-12	12	9
-11	11	8
-10	10	7

W is sum of the ranks associated with positive D_{i}.

W=4

Wilcoxon Signed Rank test Example…Continued

- The test statistic is W :sum of the ranks associated with positive $D_{i.}$

$$E(W) = \frac{n(n+1)}{4}$$

$$V(W) = \frac{n(n+1)(2n+1)}{24}$$

- Here W= 4 E(W)= 39 V(W)= 162.5

- p value= 0.001709 (one tailed exact p value)

- Conclusion: Reject H0

DATA SCIENCE INSTITUTE

Low Back Pain
Patient's Assessment

- Patient's assessment about pain level is measured twice: before treatment and after treatment

The patient's assessment is recorded on 1-4 scale

4: Severe pain 3:Moderate pain 2: Mild pain 1:No pain

The objective is to compare pain level before and after treatment

Data Snapshot

patient_id	pain_before	pain_after
1	3	2
2	3	1
3	1	1
4	4	2
5	4	2
6	4	2
7	4	2
8	4	2
9	1	1
10	4	3
11	3	2
12	3	2
13	2	2
14	4	2
15	4	3
16	3	2
17	3	2
18	2	2
19	2	1
20	4	2

Number of patients: 35
Pain level recorded on 1-4 scale

Wilcoxon Signed Rank test using R

```
#Import CSV data set 'Pain Level Assessment'
#alternative="greater" indicates expected difference 'before-after' is greater than zero
paindata<-read.csv(file.choose(),header=T)
wilcox.test(paindata$pain_before,paindata$pain_after,paired=TRUE,
alternative="greater")
```

```
Wilcoxon signed rank test with continuity correction


data:  paindata$pain_before and paindata$pain_after
V = 406, p-value = 9.995e-07
alternative hypothesis: true location shift is greater than 0
#Inference: Reject H0 and conclude that pain level has decreased significantly after treatment.
```

Kruskal Wallis test

- The Kruskal **Wallis** test is considered as nonparametric alternative to one way analysis of variance.

- The Kruskal Wallis test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed.

- H_0: K samples come from the same population
- H_1: Not H_0.

DATA SCIENCE INSTITUTE

Kruskal Wallis test Procedure

1. Combine all the observations from k samples into a single sample of size n and arrange them in ascending order .
   Assign ranks to them from smallest to largest as 1 to n. if there is a tie at two or more places, each observation is given the mean of the ranks for which it is tied.

2. The ranks assigned to observations in each of the k groups are added separately to give k rank sums.

3. The test statistic is

$$H = \frac{12}{n(n+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(n+1)$$

$n_j = number\ of\ observations\ in\ j^{th}\ sample$

$n = number\ of\ observations\ in\ the\ combined\ sample$

$R_j = sum\ of\ the\ ranks\ in\ the\ j^{th}\ sample.$

H follows Chi Square Distribution with k-1 df

DATA SCIENCE INSTITUTE

Kruskal Wallis test Example

GroupI	GroupII	GroupIII
11.5	11	11.1
23.6	25	22.5
12.1	26.3	26
14.5	12.2	19.1
16.1	18.5	21
19.8	15	25
	15.7	22.2

Ordered Sample	Group	Rank
11	GroupII	1
11.1	GroupIII	2
11.5	GroupI	3
12.1	GroupI	4
12.2	GroupII	5
14.5	GroupI	6
15	GroupII	7
15.7	GroupII	8
16.1	GroupI	9
18.5	GroupII	10
19.1	GroupIII	11
19.8	GroupI	12
21	GroupIII	13
22.2	GroupIII	14
22.5	GroupIII	15
23.6	GroupI	16
25	GroupII	17.5
25	GroupIII	17.5
26	GroupIII	19
26.3	GroupII	20

R1=50 R2= 68 R3=92

n1=6 n2=7 n3=7

P value=0.3278

DATA SCIENCE INSTITUTE

Employee Feedback On Performance Appraisal Process

- Objective of the study is to assess feedback of employees about performance appraisal process in the large company.

  The feedback is measured on Likert scale. (1-5)

  1: Not satisfied at all   5:Very satisfied

- HR manager is interested in comparing feedback rating in 3 functions namely Marketing, Finance and IT

Data Snapshot

Empno	Satscore	Function
1	3	Marketing
2	4	Finance
3	4	IT
4	5	Marketing
5	5	IT
6	3	IT
7	2	Finance
8	1	Finance
9	4	Finance
10	2	Marketing
11	2	Marketing
12	1	Marketing
13	4	IT
14	1	IT
15	3	IT
16	5	IT
17	3	Marketing
18	5	Marketing
19	5	Marketing
20	3	Finance

Number of respondents: 78
Feedback rating is measured on Likert scale (1-5)

DATA SCIENCE INSTITUTE

Kruskal Wallis test using R

```
#Import CSV data set 'Performance Appraisal Feedback'

pa_feedback<-read.csv(file.choose(),header=T)
kruskal.test(formula=Satscore~Function,data=pa_feedback)
```

```
Kruskal-Wallis rank sum test

data:  Satscore by Function
Kruskal-Wallis chi-squared = 0.36732, df = 2, p-value = 0.8322

#Inference: Do not reject H0 which suggests feedback is similar from employees in 3 different functions
```

DATA SCIENCE INSTITUTE

Chi Square Test
Association of attributes

- The chi-square test for independence, also called Pearson's chi-square test or the chi-square test of association, is used to test if there is a relationship between two categorical variables.

- The two categorical variables can be nominal or ordinal.

- H_0: Two attributes are independent (not associated)
- H_1: Not H_0.

Chi Square test Procedure

- Assume that there are ‘r’ categories of attribute A and ‘c’ categories of attribute B. Therefore, we have a cross table r*c.

- Let R_i be the total of i^{th} row and C_j be the total of j^{th} column.

- Observed frequencies are calculated from the data.
- O_{ij}: Observed frequency in ith row and jth column

- Expected frequencies are given by E_{ij} : $(R_i * C_j)/ n$ where n is total sample size. Expected frequencies are computed under null hypothesis.

Chi Square test Example

	Recruitment Source			
Performance	Internal	Campus	Jobportal	Total
Excellent	100	150	40	290
Good	100	100	100	300
Poor	50	80	150	280
Total	250	330	290	870

Observed Frequencies

	Recruitment Source			
Performance	Internal	Campus	Jobportal	Total
Excellent	(290*250)/870=83.33	110	96.67	290
Good	86.21	113.79	100	300
Poor	80.46	106.21	(280*290)/870=93.33	280
Total	250	330	290	870

Expected Frequencies

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

O = the frequencies observed

E = the frequencies expected

$\sum$ = *the 'sum of'*

Degrees of freedom=(r-1)*(c-1)

Chi Square test using R

```
#Import CSV data set 'Recruitment Source'
#Function CrossTable is available in package 'gmodels'
association<-read.csv(file.choose(),header=T)
library(gmodels)
CrossTable(association$performance, association$source,chisq=TRUE)

Pearson's Chi-squared test
------------------------------------------------------------
Chi^2 =  107.3786     d.f. =  4     p =  2.635987e-22

#Reject the null hypothesis and conclude that 'Recruitment Source'
And 'Employee Performance' are associated.
```

ATA SCIENCE
INSTITUTE

THANK YOU!

DATA SCIENCE
INSTITUTE

THANK YOU!

THANK YOU!

DATA SCIENCE
INSTITUTE