

Non-Parametric Tests:

1. Wilcoxon Test-

- To compare 2 paired samples(Paired samples: same person/object but at 2 different situations)
- E.g – Calcium level in cancer patients :

sample1 – In the initial week

sample2 – After 1 week of treatment

H0: There is no significant difference in Calcium levels of initial week and 1 week after the treatment.

H1: There is a significant difference in Calcium levels of initial week and 1 week after the treatment

From scipy.stats import Wilcoxon

Stats,p = Wilcoxon(dataset.sample1,dataset.sample2)

- Stats – **critical**, statistical or calculated value (t or z value)

P – Probability approach value

If $p < 0.05$ -> reject H0

$p \geq 0.05$ -> accept H0.

2. Friedmann Test –

- To compare more than 2 paired samples(Paired samples: same person/object but at 2 different situations)
- E.g – Calcium level in cancer patients :

sample1 – In the initial week

sample2 – After 1 week of treatment

sample3 – After 2 weeks

H0: There is no significant difference in Calcium levels of initial week, 1 week after the treatment, and 2 weeks after the treatment.

H1: There is a significant difference in Calcium levels of initial week and 1 week after the treatment, and 2 weeks after the treatment.

```
from scipy.stats import friedmanchisquare
```

```
Stats,p=friedmanchisquare(dataset.sample1,dataset.sample2,dataset.sample3)
```

- Then same steps as Wilcoxon.

3.Mann-Whitney Test –

- To compare 2 independent samples
- E.g – Sales of Design1 and Design2 of shoes of different sizes in a store

Sample1 – Sales of all sizes of Design1

Sample2 – Sales of all sizes of Design2

H0: There is no significant difference in sales of design 1 and design2.

H1: There is a significant difference in sales of design 1 and design2.

```
from scipy.stats import mannwhitneyu
```

```
Stats,p=mannwhitneyu(dataset.sample1,dataset.sample2)
```

4.Kruskal Wallis test –

- To compare more than 2 independent samples
- E.g – Sales of Design1 ,Design2 and Design3 of shoes of different sizes in a store

```
from scipy.stats import kruskal
```

```
Stats,p=kruskal(dataset.sample1,dataset.sample2,dataset.sample3)
```

5.Chi-square test –

- To check dependency of categorical variables.
- E.g – Checking dependency of gender and smoking.

H0: There is no dependency between Gender and Smoking.

H1: There is dependency between Gender and Smoking.

- Step -1: Drop null variables: `dataset.dropna()`

- Step-2: `from scipy.stats import chi2_contingency`
- Step-3 : `chitable = pd.crosstab(dataset.Gender,dataset.Smoking)`
- This will present it in this form:

| | Current Smoker | Past Smoker | Non smoker |
|---------|----------------|-------------|------------|
| Males | | | |
| Females | | | |

- Step-4: `stats,p,dof,expeted=chi2_contingency(chitable)`
- `Print(stats,p)`
- Step-5 comparing p value.