

## Practical No 4 T084 Jayesh Mali

### Hypothesis Testing

#### A) Formulate null and alternative hypotheses for a given problem.

```
import numpy as np
import scipy.stats as stats

#Create a Dumpy dataset of 10 year old children weight
data =np.random.randint(20,30,10)
print("Data:",data)

Data: [28 28 20 26 21 27 20 23 27 24]

#Define A null hypothesis
H0="The average weight of 10 year old children is 32 kg"

#Define A alternative hypothesis
H1="The average weight of 10 year old children is more than 32 kg"

#Calculate the statistic
t_stat,p_value = stats.ttest_1samp(data,32)

#Print the result
print("Test Statistic:",t_stat)
print("P-Value:",p_value)

Test Statistic: -7.4207679259069605
P-Value: 4.014173535583541e-05

#Conclusion
if p_value<0.05:
    print("Reject the null hypothesis.")
else:
    print("fail to reject the null hypothesis")

Reject the null hypothesis.
```

#### Two-Sample Test (Independent)

```
from scipy.stats import ttest_ind
import numpy as np

week1=np.genfromtxt("week1.csv", delimiter=",")
week2=np.genfromtxt("week2.csv", delimiter=",")

print("week1 data :-\n")
print(week1)
print("week2 data :-\n")
print(week2)
```

```
week1 data :-
```

```
[23.90506824 26.67631982 27.27433886 24.25757125 32.40423483
39.56919357
 23.07010059 29.82068109 27.59433809 28.05639569 27.06757262
30.41192979
 25.71358554 24.94294823 28.23123807 24.95337555 18.51231639
27.46234762
 28.38016611 13.91205901 29.02615866 26.90746774 22.8677726
24.8938289
 25.96947935 26.86869621 20.72676456 27.35988314 20.68408581
21.19846143
 16.25800931 23.92517681 24.47923229 29.47050863 27.28425372
26.93339272
 28.61026924 18.88377042 33.65468651 25.69470077 20.98291356
22.69700387
 28.60278855 21.36000443 30.77685156 20.83415999 23.79367158
19.7556718
 29.54421084 20.1433138 ]
```

```
week2 data :-
```

```
[18.63431907 31.28788036 34.96797943 21.81678117 28.21619974
39.39313736
 35.52223207 27.54222109 33.64395433 25.31673581 28.81392191
30.7358016
 26.37241881 26.0945555 26.34073477 19.42196017 32.58797652
24.84001926
 28.93348335 20.43667584 22.72495967 32.31728012 35.384306
29.66709637
 24.53512973 30.91406007 19.56117513 24.90816833 30.13163726
31.47466199
 27.77683598 16.51307462 35.0770162 31.74818107 36.36053496
27.70500593
 29.49869936 27.65575346 37.18504075 25.16055104 29.26553553
38.22163057
 28.92102091 24.8215439 38.30155495 34.76020645 22.26869162
28.82593733
 32.00975127 36.46437665]
```

```
week1_mean=np.mean(week1)
```

```
week2_mean=np.mean(week2)
```

```
print(week1_mean)
```

```
print(week2_mean)
```

```
25.4480593952
```

```
29.0215681076
```

```
week1_std=np.std(week1)
```

```
week2_std=np.std(week2)
```

```

print("Week1 std value",week1_std)
print("Week2 std value",week2_std)

Week1 std value 4.531693386680561
Week2 std value 5.497966708987187

ttest,pval=ttest_ind(week1,week2)
print("p-value",pval)

p-value 0.000676767690454633

if pval<0.05:
    print("reject null hypothesis")
else:
    print("Accept null hypothesis")

reject null hypothesis

```

Paired Sample test (Relative Or Not )

```

import pandas as pd
from scipy import stats

df=pd.read_csv("blood_pressure.csv")
df[['bp_before','bp_after']].describe()

   count  bp_before  bp_after
count   120.000000   120.000000
mean    156.450000   151.358333
std      11.389845    14.177622
min     138.000000   125.000000
25%     147.000000   140.750000
50%     154.500000   149.500000
75%     164.000000   161.000000
max     185.000000   185.000000

ttest,pval = stats.ttest_rel(df['bp_before'],df['bp_after'])
print(pval)

0.0011297914644840823

if pval<0.05:
    print("reject null hypothesis")
else:
    print("Accept null hypothesis")

reject null hypothesis

```

Z-Test for one Variable

```
import pandas as pd
from scipy import stats
from statsmodels.stats import weightstats as stests

ztest,pval = stests.ztest(df['bp_before'],x2=None,value=156)
print(float(pval))

0.6651614730255063

if pval<0.05:
    print("reject null hypothesis")
else:
    print("Accept null hypothesis")

Accept null hypothesis
```

#### Z-Test for Two Variable

```
ztest,pval1=stests.ztest(df['bp_before'],x2=df['bp_after'],value=0,alt
ernative='two-sided')
print(pval1)

0.002162306611369422

if pval1<0.05:
    print("reject null hypothesis")
else:
    print("Accept null hypothesis")

reject null hypothesis
```

#### Chi-Test