experiment-no-6-t-test

April 8, 2024

1 T Test

[1]:
```
#Experiment no 6 To Perform Hypothesis Testing Using T-Test.
```

[2]:
```
#Name: Shravani M Karne
#Roll no.: 39
#Sec : A
#Year: 3rd Year
#Sub: Big Data Analysis (ET 2 Lab)
```

[3]:
```
ages=[10,20,32,59,28,40,55,69,16,55,30,25,43,19,67,99,85,30,28,14,24,16,17,32,35,26,27,65,18,4
```

[4]:
```
len(ages)
```

[4]:
```
35
```

[5]:
```
import numpy as np
ages_mean=np.mean(ages)
print(ages_mean)
```

```
36.57142857142857
```

[6]:
```
## Lets take sample

sample_size=10
age_sample=np.random.choice(ages,sample_size)
```

[7]:
```
age_sample
```

[7]:
```
array([28, 21, 40, 65, 65, 28, 16, 27, 14, 35])
```

[8]:
```
from scipy.stats import ttest_1samp
```

[9]:
```
ttest,p_value=ttest_1samp(age_sample,30)
```

[10]:
```
print(p_value)
```

```
0.5144771198580793
```

[11]:
```
if p_value < 0.05:    # alpha value is 0.05 or 5%
    print(" we are rejecting null hypothesis")
else:
    print("we are accepting null hypothesis")
```

```
we are accepting null hypothesis
```

[12]:
```
from numpy.random import seed
from numpy.random import rand
```

[13]:
```
seed(1)
```

[14]:
```
values = rand(45)
```

[15]:
```
print(values)
```

```
[4.17022005e-01 7.20324493e-01 1.14374817e-04 3.02332573e-01
 1.46755891e-01 9.23385948e-02 1.86260211e-01 3.45560727e-01
 3.96767474e-01 5.38816734e-01 4.19194514e-01 6.85219500e-01
 2.04452250e-01 8.78117436e-01 2.73875932e-02 6.70467510e-01
 4.17304802e-01 5.58689828e-01 1.40386939e-01 1.98101489e-01
 8.00744569e-01 9.68261576e-01 3.13424178e-01 6.92322616e-01
 8.76389152e-01 8.94606664e-01 8.50442114e-02 3.90547832e-02
 1.69830420e-01 8.78142503e-01 9.83468338e-02 4.21107625e-01
 9.57889530e-01 5.33165285e-01 6.91877114e-01 3.15515631e-01
 6.86500928e-01 8.34625672e-01 1.82882773e-02 7.50144315e-01
 9.88861089e-01 7.48165654e-01 2.80443992e-01 7.89279328e-01
 1.03226007e-01]
```

[16]:
```
from numpy.random import seed
from numpy.random import randint
```

[17]:
```
seed(1)
```

[18]:
```
values = randint(0,20,40)
print(values)
```

```
[ 5 11 12  8  9 11  5 15  0 16  1 12  7 13  6 18  5 18 11 10 14 18  4  9
 17  0 13  9  9  7  1  0 17  8 13 19 15 10  8  7]
```

[19]:
```
import numpy as np
import scipy.stats as stats
```

[20]:
```
sample_mean= 145
population_mean=120
population_std=20
sample_size=50
```

```
alpha=0.05
```

[21]:
```
z_score = (sample_mean-population_mean)/(population_std/np.sqrt(50))
print('Z-Score :',z_score)
```

```
Z-Score : 8.838834764831844
```

[22]:
```
z_critical = stats.norm.ppf(1-alpha)
print('Critical Z-Score :',z_critical)
```

```
Critical Z-Score : 1.6448536269514722
```

[23]:
```
if z_score >  z_critical:
    print("Reject Null Hypothesis")
else:
  print("Fail to Reject Null Hypothesis")
```

```
Reject Null Hypothesis
```

[26]:
```
A = np.random.normal(10, 20, 100)
B = np.random.normal(10, 20, 100)

stats.ttest_ind(A, B)
```

[26]:
```
TtestResult(statistic=1.1232131007475072, pvalue=0.26270705362179175, df=198.0)
```