Hypothesis Testing

```
import numpy as np
import pandas as pd
from scipy import stats
```

Z test with one mean

Suppose an e-commerce platform receives an average of 100 visitors per day (known population mean). We want to test if the average number of visitors for a recent sample of 39 days is significantly different from this known average. Verify the claim at significance level alpha = 0.05.

We aim to test:

- **Null Hypothesis (H0):** The sample mean is equal to the population mean is equal to 100.
- **Alternative Hypothesis (H1):** The sample mean is not equal to the population mean is not equal to 100.

```
# Given data
pop_mean = 100
pop_sd = 15
visitors = [
    117, 119, 132, 106, 106, 133, 121, 102, 118, 103, 103, 113, 81,
84, 101, 94, 114, 96, 88, 131, 106, 111, 88, 101,
    111, 92, 115, 100, 105, 100, 137, 109, 94, 122, 91, 113, 80, 90,
112
]

# Calculate sample mean and size
sample_mean = np.mean(visitors)
sample_size = len(visitors)

# Calculate Z-score
z_score_calculated = (sample_mean - pop_mean) / (pop_sd /
np.sqrt(sample_size))

# Find the critical Z-value for a two-tailed test at alpha = 0.05
using scipy
alpha = 0.05
z_critical = stats.norm.ppf(1 - alpha / 2)

# Print results
print(f"Calculated Z-Score: {z_score_calculated:.3f}")
print(f"Critical Z-Value: ±{z_critical:.3f}")
```

```
# Compare and conclude
if z_score_calculated > z_critical or z_score_calculated < -
z_critical:
    print("Conclusion: Reject the null hypothesis. The sample mean is
significantly different from the population mean.")
else:
    print("Conclusion: Fail to reject the null hypothesis. The sample
mean is not significantly different from the population mean.")


Calculated Z-Score: 2.551
Critical Z-Value: ±1.960
Conclusion: Reject the null hypothesis. The sample mean is
significantly different from the population mean.
```

Decision Rule:

- Reject the Null Hypothesis (H_0):
  - If Zcalculated lies outside the range [-Zcritical, Zcritical], the sample mean is significantly different from the population mean.
  - This means there is sufficient evidence to conclude that the observed difference is unlikely due to random sampling alone.
- Fail to Reject the Null Hypothesis (H_0):
  - If Zcalculated lies within the range [-Zcritical, Zcritical], the sample mean is not significantly different from the population mean.
  - This means the observed difference could reasonably be due to random sampling.

```
# Define significance level
alpha = 0.05

# Z-critical value for a right-tailed test
z_critical_right = stats.norm.ppf(1 - alpha)

# Z-critical value for a left-tailed test
z_critical_left = stats.norm.ppf(alpha)

# Z-critical value for a two-tailed test
z_critical_two = stats.norm.ppf(1 - alpha / 2)

# Display results
print(f"Z-Critical (Right-Tailed Test) for alpha={alpha}:
{z_critical_right:.3f}")
print(f"Z-Critical (Left-Tailed Test) for alpha={alpha}:
{z_critical_left:.3f}")
print(f"Z-Critical (Two-Tailed Test) for alpha={alpha}:
{z_critical_two:.3f}")
```

```
Z-Critical (Right-Tailed Test) for alpha=0.05: 1.645
Z-Critical (Left-Tailed Test) for alpha=0.05: -1.645
Z-Critical (Two-Tailed Test) for alpha=0.05: 1.960
```

```
z_score = 1.96  # Z-score (for example, for a 95% confidence interval)

# Calculate the cumulative probability (CDF)
probability = stats.norm.cdf(z_score)

# Print the result
print(f"The probability for Z = {z_score} is: {probability}")
```

```
The probability for Z = 1.96 is: 0.9750021048517795
```

Z-test with two Mean

```
campaign1_data = [157, 147, 159, 172, 146, 146, 173, 161, 142, 158,
143, 143, 153, 121, 124, 141, 134, 154, 136, 128, 171, 146, 151, 128,
141, 151, 132, 155, 146, 145, 148, 177, 149, 134, 162]

campaign2_data = [178, 167, 181, 195, 166, 166, 196, 183, 162, 179,
162, 162, 174, 137, 148, 168, 152, 175, 154, 145, 194, 166, 171, 145,
168, 171, 158, 176, 159, 165, 159, 281, 169, 152, 183]

camp1_mean = np.mean(campaign1_data)
camp1_std = np.std(campaign1_data)
n1 = len(campaign1_data)

camp2_mean = np.mean(campaign2_data)
camp2_std = np.std(campaign2_data)
n2 = len(campaign2_data)

z_score_calculated = (camp1_mean - camp2_mean) / np.sqrt(((camp1_std
** 2) / n1) + ((camp2_std ** 2) / n2))

alpha = 0.05

z_critical = stats.norm.ppf(1 - alpha / 2)

# Display results
print(f"Campaign 1 Mean: {camp1_mean}")
print(f"Campaign 2 Mean: {camp2_mean}")
print(f"Z-Score: {z_score_calculated}")
print(f"Critical Z-Value: {z_critical}")

# Conclusion based on comparing Z-score with critical Z-value
if z_score_calculated > z_critical or z_score_calculated < -
z_critical:
    print("Reject the null hypothesis: There is a significant
```

```
difference in the average number of daily visitors between the two
campaigns.")
else:
    print("Fail to reject the null hypothesis: There is no significant
difference in the average number of daily visitors between the two
campaigns.")
```

```
Campaign 1 Mean: 147.82857142857142
Campaign 2 Mean: 170.4857142857143
Z-Score: -4.9425885015566635
Critical Z-Value: 1.959963984540054
Reject the null hypothesis: There is a significant difference in the
average number of daily visitors between the two campaigns.
```

T-test with single mean

An eComm platform believes that the average number of daily visitors is 150. Test if the average number of visitors for a sample of 20 days is significantly different from this value.

```
# Number of visitors for 20 days
visitors = [
    157, 147, 159, 172, 146, 146, 173, 161, 142, 158, 143, 143, 153,
121, 124, 141, 134, 154, 136, 128
]

pop_mean = 150

sample_mean = np.mean(visitors)
sample_std = np.std(visitors, ddof=1)
#Here, ddof=1 tells NumPy to use n-1 when calculating the standard
deviation, which is necessary for sample data.
#We use ddof =0 for polulation standard deviation.
n = len(visitors)

df = n - 1

# T-Statistic calculation
t_calc, _ = stats.ttest_1samp(visitors, popmean=pop_mean)

#Note: In Python, when a function returns multiple values and you are
not interested in all of them,
#you can use _ to ignore the values that you don't need.
#It is a convention to indicate that the value is deliberately being
discarded.

alpha = 0.05
t_critical = stats.t.ppf(1 - alpha / 2, df)
```

```
print(f"Sample Mean: {sample_mean}")
print(f"Sample Standard Deviation: {sample_std}")
print(f"T-Statistic: {t_calc}")
print(f"Critical T-Value: {t_critical}")

# Conclusion based on comparing T-statistic with critical T-value
if abs(t_calc) > t_critical:
    print("Reject the null hypothesis: The sample mean is
significantly different from the population mean.")
else:
    print("Fail to reject the null hypothesis: There is no significant
difference between the sample mean and the population mean.")
```

```
Sample Mean: 146.9
Sample Standard Deviation: 14.322966467218047
T-Statistic: -0.9679294783122818
Critical T-Value: 2.093024054408263
Fail to reject the null hypothesis: There is no significant difference
between the sample mean and the population mean.
```

T test with Two Sample Mean

```
# Data
data = {
    'Gender': ['Female', 'Male', 'Female', 'Female', 'Male', 'Male',
'Male', 'Male', 'Female', 'Female', 'Male', 'Male',
               'Female', 'Male', 'Female', 'Female', 'Female', 'Male',
'Male', 'Female', 'Female', 'Male', 'Male', 'Male',
               'Male', 'Female', 'Female', 'Male', 'Male', 'Male',
'Female', 'Female', 'Female', 'Female', 'Female',
               'Male', 'Male', 'Female', 'Female', 'Male'],
    'Marks': [77, 89, 89, 91, 76, 85, 79, 78, 93, 88, 91, 77, 81, 88,
86, 80, 82, 95, 87, 83, 79
              , 94, 84, 73, 85, 85, 78, 88, 81, 82, 84, 86, 75, 83,
87, 80, 90, 90, 92, 92]
}
# Create DataFrame
df = pd.DataFrame(data)
print(df)
# Separate the data into two groups
male_marks = df[df['Gender'] == 'Male']['Marks']
female_marks = df[df['Gender'] == 'Female']['Marks']

print(male_marks)
print(female_marks)

mean_male = male_marks.mean()
mean_female = female_marks.mean()
```

```
std_male = male_marks.std(ddof=1)
std_female = female_marks.std(ddof=1)

n_male = len(male_marks)
n_female = len(female_marks)

pooled_variance = ((std_male ** 2) / n_male) + ((std_female ** 2) /
n_female)

t_statistic = (mean_male - mean_female) / np.sqrt(pooled_variance)

df = ((std_male ** 2 / n_male + std_female ** 2 / n_female) ** 2) / \
     ((std_male ** 2 / n_male) ** 2 / (n_male - 1) + (std_female ** 2
/ n_female) ** 2 / (n_female - 1))

alpha = 0.05

t_critical = stats.t.ppf(1 - alpha / 2, df)

# Display results
print(f"Male Group Mean: {mean_male}")
print(f"Female Group Mean: {mean_female}")
print(f"T-Statistic: {t_statistic}")
print(f"Degrees of Freedom: {df}")
print(f"Critical T-Value: {t_critical}")

# Conclusion
if abs(t_statistic) > t_critical:
    print("Reject the null hypothesis: There is a significant
difference in the means of Marks for males and females.")
else:
    print("Fail to reject the null hypothesis: There is no significant
difference in the means of Marks for males and females.")
```

	Gender	Marks
0	Female	77
1	Male	89
2	Female	89
3	Female	91
4	Male	76
5	Male	85
6	Male	79
7	Male	78
8	Female	93
9	Female	88
10	Male	91
11	Male	77
12	Female	81
13	Male	88

```
14  Female     86
15  Female     80
16  Female     82
17    Male     95
18    Male     87
19  Female     83
20  Female     79
21    Male     94
22    Male     84
23    Male     73
24    Male     85
25  Female     85
26  Female     78
27    Male     88
28    Male     81
29    Male     82
30  Female     84
31  Female     86
32  Female     75
33  Female     83
34  Female     87
35    Male     80
36    Male     90
37  Female     90
38  Female     92
39    Male     92
1     89
4     76
5     85
6     79
7     78
10    91
11    77
13    88
17    95
18    87
21    94
22    84
23    73
24    85
27    88
28    81
29    82
35    80
36    90
39    92
Name: Marks, dtype: int64
0     77
2     89
```

```
3     91
8     93
9     88
12    81
14    86
15    80
16    82
19    83
20    79
25    85
26    78
30    84
31    86
32    75
33    83
34    87
37    90
38    92
Name: Marks, dtype: int64
Male Group Mean: 84.7
Female Group Mean: 84.45
T-Statistic: 0.13733874423305972
Degrees of Freedom: 36.61108492606419
Critical T-Value: 2.0269192513782666
Fail to reject the null hypothesis: There is no significant difference
in the means of Marks for males and females.
```

F test

```
import pandas as pd
from scipy.stats import f, f_oneway

data = {
    "Shipping Option": ["Standard", "Express", "Same-Day", "Standard",
"Express", "Same-Day","Standard", "Express", "Same-Day"],
    "Purchase Amounts": [50, 70, 90, 55, 75, 85, 60, 80, 95]
}

df = pd.DataFrame(data)

standard_shipping = df[df["Shipping Option"] == "Standard"]["Purchase
Amounts"]
express_shipping = df[df["Shipping Option"] == "Express"]["Purchase
Amounts"]
same_day_shipping = df[df["Shipping Option"] == "Same-Day"]["Purchase
Amounts"]
```

```
t = 3


n = len(df)

df_between = t - 1  # Degrees of freedom between groups
df_within = n - t  # Degrees of freedom within groups

f_statistic, _ = f_oneway(standard_shipping, express_shipping,
same_day_shipping)


alpha = 0.05

f_critical = stats.f.ppf(1 - alpha, df_between, df_within)

# Display the F-statistic and F-critical value
print(f"F-Statistic: {f_statistic}")
print(f"F-Critical Value: {f_critical}")

# Conclusion
if f_statistic > f_critical:
    print("Reject the null hypothesis: There are significant
differences between the group means.")
else:
    print("Fail to reject the null hypothesis: There are no
significant differences between the group means.")
```

```
F-Statistic: 37.00000000000006
F-Critical Value: 5.143252849784718
Reject the null hypothesis: There are significant differences between
the group means.
```