

CS-23334 FUNDAMENTALS OF DATA SCIENCE

ASHWIN C 240701058

Experiment 14

Date: 23.10.2025

14. Hypothetical using ANOVA-Test

Aim:

To compare the growth rates of plants under three different fertilizer treatments (Treatment A, B, and C) to determine if there is a significant difference in their mean growth.

Description:

A statistical ANOVA -Test is conducted to determine the growth rates of plants under three different treatments .

Algorithm:

1. Null Hypothesis (H_0): The mean growth rates of plants under all three fertilizer treatments are equal.

2. Alternative Hypothesis (H_1): At least one pair of mean growth rates of plants under different fertilizer treatments are not equal.

3. Samples:

Measure the growth (in centimeters) of 25 plants under Treatment A.

Measure the growth (in centimeters) of 25 plants under Treatment B.

Measure the growth (in centimeters) of 25 plants under Treatment C.

4. ANOVA: Conduct a one-way ANOVA to compare the mean growth rates of plants across the three fertilizer treatments.
5. Decision Rule: Use a significance level of $\alpha = 0.05$.

Code With Example:

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

# Set a random seed for reproducibility
np.random.seed(42)

# Generate hypothetical growth data for three treatments (A, B, C)
n_plants = 25

growth_A = np.random.normal(loc=10, scale=2, size=n_plants)
growth_B = np.random.normal(loc=12, scale=3, size=n_plants)
growth_C = np.random.normal(loc=15, scale=2.5, size=n_plants)

# Combine all data into one array
all_data = np.concatenate([growth_A, growth_B, growth_C])

# Treatment labels for each group
treatment_labels = ['A'] * n_plants + ['B'] * n_plants + ['C'] * n_plants

# Perform one-way ANOVA
f_statistic, p_value = stats.f_oneway(growth_A, growth_B, growth_C)

print("Treatment A Mean Growth:", np.mean(growth_A))
print("Treatment B Mean Growth:", np.mean(growth_B))
print("Treatment C Mean Growth:", np.mean(growth_C))
print()
print(f"F-Statistic: {f_statistic:.4f}")
print(f"P-Value: {p_value:.4f}")
```

Output:

```
Treatment A Mean Growth: 9.672983882683818
Treatment B Mean Growth: 11.137680744437432
Treatment C Mean Growth: 15.265234904828972

F-Statistic: 36.1214
P-Value: 0.0000
```

```

alpha = 0.05
if p_value < alpha:
    print("Reject the null hypothesis: There is a significant
difference in mean growth rates among the three treatments.")
else:
    print("Fail to reject the null hypothesis: There is no significant
difference in mean growth rates among the three treatments.")

if p_value < alpha:
    from statsmodels.stats.multicomp import pairwise_tukeyhsd

    tukey_results = pairwise_tukeyhsd(all_data, treatment_labels,
alpha=0.05)

print("\nTukey's HSD Post-hoc Test:")
print(tukey_results)

```

Output:

```

Reject the null hypothesis: There is a significant difference in mean
growth rates among the three treatments.

Tukey's HSD Post-hoc Test:
Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====
group1 group2 meandiff  p-adj   lower   upper  reject
-----
      A      B    1.4647  0.0877 -0.1683  3.0977  False
      A      C    5.5923    0.0   3.9593  7.2252   True
      B      C    4.1276    0.0   2.4946  5.7605   True
-----
```

Result:

The Python program to compare the growth rates of plants under three different fertilizer treatments involves conducting an ANOVA test to determine if there is a significant difference in mean growth rates among the treatments was executed.