T-TEST

INFERENTIAL STATISTICS

prepared by:

Gyro A. Madrona
Electronics Engineer

TOPIC OUTLINE

1-Sample t-Test

F-Test

2-Sample t-Test

1-SAMPLE T-TEST

1-SAMPLE T-TEST

A one-sample t-test evaluates whether a **sample mean** significantly differs from hypothesized population mean.

Assumptions

- Continuous data
- Normally distributed

One-sample t-test function

```
t_stat, p_value =
stats.ttest_1samp(sample_data,
pop_mean)
```

Null Hypothesis

$$H_o: \mu_1 = \mu_o$$

Alternative Hypothesis

$$H_a: \mu_1 \neq \mu_o \text{ (p-value} \leq \alpha)$$

EXERCISE

The dataset contains the electricity production in MWh by the following production types:

Type	μ	σ
Nuclear	1282.48	241.19
Wind	771.73	666.53
Hydroelectric	1792.64	679.98
Oil and Gas	1166.84	433.92
Coal	1138.50	329.93
Solar	171.81	257.26
Biomass	54.94	13.76

dataset

"**electricity-sample-1-dataset.csv**"

Perform a **1-sample t-test** to determine whether the mean electricity production (in MWh) for each production type in the given dataset differs significantly from known population parameters.

2-SAMPLE T-TEST

F-TEST

An **F-test** is a statistical test used to compare the **variances** of two populations and determine if they are significantly different.

Formula

$$F = s_1^2 / s_2^2$$

where:

F = F-statistic

s_1^2 = larger sample variance

s_2^2 = smaller sample variance

Cumulative Probability of F-Distribution

```
p_value =
    1 - stats.f.cdf(f_stat,dof1,dof2)
```

Null Hypothesis

$H_o: \sigma_1^2 = \sigma_2^2$

Alternative Hypothesis

$H_a: \sigma_1^2 \neq \sigma_2^2$ (p-value $\leq \alpha$)

2-SAMPLE T-TEST

A **two-sample t-test** evaluates whether the mean of **two samples** are significantly different from each other.

Assumptions

- Continuous data
- Normally distributed

Independent Samples

```
t_stat, p_value =
stats.ttest_ind(sample_1, sample_2)
```

Dependent Samples

```
t_stat, p_value =
stats.ttest_rel(sample_1, sample_2)
```

Null Hypothesis

$$H_o: \mu_1 = \mu_1$$

Alternative Hypothesis

$$H_a: \mu_1 \neq \mu_2 \quad (\text{p-value} \leq \alpha)$$

EXERCISE

The dataset contains the electricity production in MWh by the following production types:

Type	μ	σ
Nuclear	1282.48	241.19
Wind	771.73	666.53
Hydroelectric	1792.64	679.98
Oil and Gas	1166.84	433.92
Coal	1138.50	329.93
Solar	171.81	257.26
Biomass	54.94	13.76

dataset

"**electricity-sample-1-dataset.csv**"

Perform an **independent two-sample t-test** to compare the means of electricity production (in MWh) between two distinct groups in the dataset.

LABORATORY