

Course: Artificial Intelligence and Machine Learning**Code: 20CS51I, WEEK - 4****SESSION – 4, Univariate analysis tests**

- Hypothesis testing
- Error, Test statistic, type, interpreting test statistics.
- Understanding p-value

Univariate analysis tests

Tests of statistical hypotheses are widely used in quality of life research. The expression “univariate tests” is typically used as a shorthand for “univariate statistical tests.” Univariate statistical tests are those tests that involve one dependent variable. Examples include t-tests of means, analysis of variance (ANOVA), analysis of covariance, linear regression, and generalized linear models such as binary logistic regression. In all of these cases, there is only one dependent variable. By contrast, methods such as multivariate analysis of variance (MANOVA), multivariate analysis of covariance (MANCOVA), and canonical correlation are all examples of multivariate tests used in multivariate statistical analysis.

Hypothesis testing

Hypothesis testing is a statistical method that is used in making statistical decisions using experimental data. Hypothesis Testing is basically an assumption that we make about the population parameter.

Ex : you say avg student in class is 40 or a boy is taller than girls.

All those example we assume need some statistic way to prove those. we need some mathematical conclusion what ever we are assuming is true.

Hypothesis testing is an essential procedure in statistics. A hypothesis test evaluates two mutually exclusive statements about a population to determine which statement is best supported by the sample data. When we say that a finding is statistically significant, it's thanks to a hypothesis test.

Five Steps in Hypothesis Testing:

- Specify the Null Hypothesis.
- Specify the Alternative Hypothesis.
- Set the Significance Level (α)
- Calculate the Test Statistic and Corresponding P-Value.
- Drawing a Conclusion.

Steps Involved in Hypothesis Testing

Step 1: Define the null hypothesis, alternate hypothesis and the level of significance.

Step 2: Calculate the probability of getting the observed data (P value) assuming the null hypothesis to be true. This involves two intermediate steps as given below.

Step 2.1: Calculate the Z statistic.

Step 2.2: Calculate the P value.

Step 3: Conclude whether to reject the null hypothesis or not based on the P value i.e.

If $P \text{ value} < \text{significance level}$, then reject the null hypothesis

If $P \text{ value} \geq \text{significance level}$, the null hypothesis cannot be rejected

Step 4: State the conclusion.

Which are important parameter of hypothesis testing ?

Null hypothesis :- In inferential statistics, the null hypothesis is a general statement or default position that there is no relationship between two measured phenomena, or no association among groups

In other words it is a basic assumption or made based on domain or problem knowledge.

Example : a company production is = 50 unit/per day etc.

Alternative hypothesis :-

The alternative hypothesis is the hypothesis used in hypothesis testing that is contrary to the null hypothesis. It is usually taken to be that the observations are the result of a real effect (with some amount of chance variation superposed)

Example : a company production is \neq 50 unit/per day etc.

Null and Alternate hypothesis.

Level of significance: Refers to the degree of significance in which we accept or reject the null-hypothesis. 100% accuracy is not possible for accepting or rejecting a hypothesis, so we therefore select a level of significance that is usually 5%.

This is normally denoted with alpha(maths symbol α) and generally it is 0.05 or 5% , which means your output should be 95% confident to give similar kind of result in each sample.

Type I error: When we reject the null hypothesis, although that hypothesis was true. Type I error is denoted by alpha. In hypothesis testing, the normal curve that shows the critical region is called the alpha region

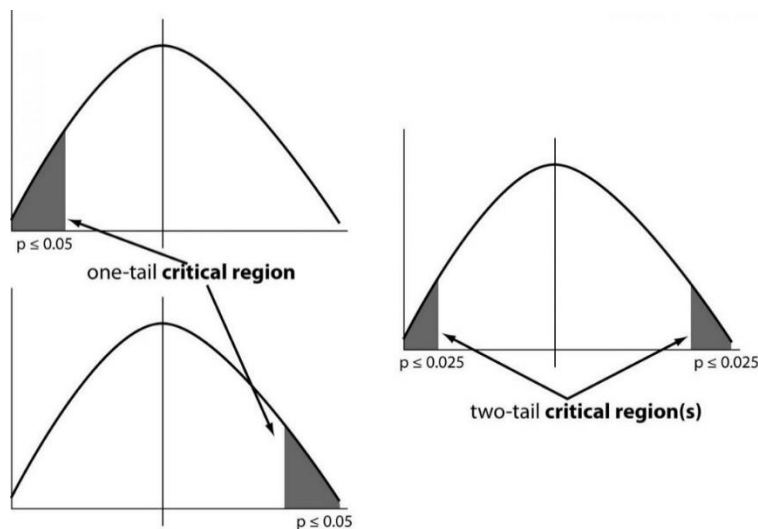
Type II errors: When we accept the null hypothesis but it is false. Type II errors are denoted by beta. In Hypothesis testing, the normal curve that shows the acceptance region is called the beta region.

One tailed test :- A test of a statistical hypothesis , where the region of rejection is on only one side of the sampling distribution , is called a one-tailed test.

Example :- a college has ≥ 4000 student or data science $\leq 80\%$ org adopted.

Two-tailed test :- A two-tailed test is a statistical test in which the critical area of a distribution is two-sided and tests whether a sample is greater than or less than a certain range of values. If the sample being tested falls into either of the critical areas, the alternative hypothesis is accepted instead of the null hypothesis.

Example : a college $\neq 4000$ student or data science $\neq 80\%$ org adopted



one and two-tailed images

P-value :- The P value, or calculated probability, is the probability of finding the observed, or more extreme, results when the null hypothesis (H_0) of a study question is true — the definition of ‘extreme’ depends on how the hypothesis is being tested.

If your P value is less than the chosen significance level then you reject the null hypothesis i.e. accept that your sample gives reasonable evidence to support the alternative hypothesis. It does NOT imply a “meaningful” or “important” difference; that is for you to decide when considering the real-world relevance of your result.

Example : you have a coin and you don’t know whether that is fair or tricky so let’s decide null and alternate hypothesis

H_0 : a coin is a fair coin.

H_1 : a coin is a tricky coin. and $\alpha = 5\%$ or 0.05

Now let’s toss the coin and calculate p- value (probability value).

Toss a coin 1st time and result is tail- P-value = 50% (as head and tail have equal probability)

Toss a coin 2nd time and result is tail, now p-value = $50/2 = 25\%$

and similarly we Toss 6 consecutive time and got result as P-value = 1.5% but we set our significance level as 95% means 5% error rate we allow and here we see we are beyond that level i.e. our null- hypothesis does not hold good so we need to reject and propose that this coin is a tricky coin which is actually.

Degree of freedom :- Now imagine you’re not into hats. You’re into data analysis. You have a data set with 10 values. If you’re not estimating anything, each value can take on any number, right? Each value is completely free to vary. But suppose you want to test the population mean with a sample of 10 values, using a 1-sample t test. You now have a constraint — the estimation of the mean. What is that constraint, exactly? By definition of the mean, the following relationship must hold: The sum of all values in the data must equal $n \times \text{mean}$, where n is the number of values in the data set.

So if a data set has 10 values, the sum of the 10 values must equal the mean $\times 10$. If the mean of the 10 values is 3.5 (you could pick any number), this constraint requires that the sum of the 10 values must equal $10 \times 3.5 = 35$.

With that constraint, the first value in the data set is free to vary. Whatever value it is, it’s still possible for the sum of all 10 numbers to have a value of 35. The second value is also free to vary,

because whatever value you choose, it still allows for the possibility that the sum of all the values is 35.

Now Let's see some of widely used hypothesis testing type :-

1. T Test (Student T test)
2. Z Test
3. ANOVA Test
4. Chi-Square Test

T- Test :- A t-test is a type of inferential statistic which is used to determine if there is a significant difference between the means of two groups which may be related in certain features. It is mostly used when the data sets, like the set of data recorded as outcome from flipping a coin a 100 times, would follow a normal distribution and may have unknown variances. T test is used as a hypothesis testing tool, which allows testing of an assumption applicable to a population.

T-test has 2 types : 1. one sampled t-test 2. two-sampled t-test.

One sample t-test : The One Sample t Test determines whether the sample mean is statistically different from a known or hypothesised population mean. The One Sample t Test is a parametric test.

Example :- you have 10 ages and you are checking whether avg age is 30 or not.

```
from scipy.stats import ttest_1samp
import numpy as np
ages = np.genfromtxt("ages.csv")
print(ages)
ages_mean = np.mean(ages)
print(ages_mean)
tset, pval = ttest_1samp(ages, 30)
print("p-values", pval)
if pval < 0.05:
    # alpha value is 0.05 or 5%
    print(" we are rejecting null hypothesis")
else:
    print("we are accepting null hypothesis")
```

Output for above code is :

```

script.py
1 from scipy.stats import ttest_1samp
2 import numpy as np
3
4 ages = np.genfromtxt("ages.csv")
5
6 print(ages)
7
8 ages_mean = np.mean(ages)
9 print("mean ages", ages_mean)
10 tset, pval = ttest_1samp(ages, 30)
11
12 print("p-values", pval)
13
14 * if pval < 0.05:
15     print(" we are rejecting null hypothesis")
16 * else:
17     print("we are accepting null hypothesis")

```

```

[ 32.  34.  29.  29.  22.  39.  38.  37.  38.  36.  30.  26.
 22.  22.]
('mean ages', 31.0)
('p-values', 0.56051558881713792)
we are accepting null hypothesis

```

one-sample t-test output

Two sampled T-test :-The Independent Samples t Test or 2-sample t-test compares the means of two independent groups in order to determine whether there is statistical evidence that the associated population means are significantly different. The Independent Samples t Test is a parametric test. This test is also known as: Independent t Test.

Example : is there any association between week1 and week2 (code is given below in python)

```

from scipy.stats import ttest_ind
import numpy as np
week1 = np.genfromtxt("week1.csv", delimiter=",")
week2 = np.genfromtxt("week2.csv", delimiter=",")
print(week1)
print("week2 data :-\n")
print(week2)
week1_mean = np.mean(week1)
week2_mean = np.mean(week2)
print("week1 mean value:", week1_mean)
print("week2 mean value:", week2_mean)
week1_std = np.std(week1)
week2_std = np.std(week2)
print("week1 std value:", week1_std)
print("week2 std value:", week2_std)
ttest, pval = ttest_ind(week1, week2)
print("p-value", pval)
if pval < 0.05:
    print("we reject null hypothesis")
else:
    print("we accept null hypothesis")

```

2-sampled t-test output

```

script.py
from scipy.stats import ttest_ind
import numpy as np

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

print(week1)
print("week2 data :-\n")
print(week2)
week1_mean = np.mean(week1)
week2_mean = np.mean(week2)

print("week1 mean value:", week1_mean)
print("week2 mean value:", week2_mean)

week1_std = np.std(week1)
week2_std = np.std(week2)

print("week1 std value:", week1_std)
print("week2 std value:", week2_std)

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

* if pval < 0.05:
    print("we reject null hypothesis")
* else:
    print("we accept null hypothesis")

```

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 value:', 25.448059395144654)			
('week2 mean value:', 29.021568107746155)			
('week1 std value:', 4.5316933870843146)			
('week2 std value:', 5.4979667086536512)			
('p-value', 0.00067676769000677567)			
we reject null hypothesis			

Interpreting test Statistics

The decision of confirming or rejecting the null hypothesis is made by interpreting the result of the test. The result of the hypothesis testing can be interpreted using p-values or critical values. The p-value is the probability of deducing the observed value, given the assumption. On the other hand, critical values are cut-off values that define regions where the test statistic is unlikely to lie. How to interpret the p-value?

A statistical hypothesis test may return a p-value. A p-value is defined as the probability of making the observation made, given the null hypothesis is true. It is calculated using the sample distribution of the test statistic, under the assumption i.e. null hypothesis.

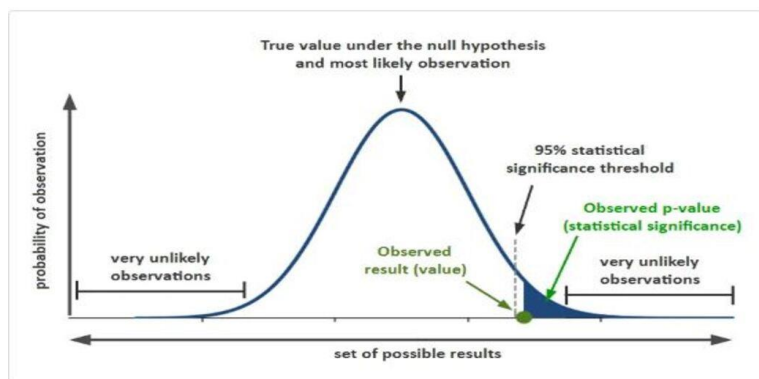
The p-value is used to quantify the result of the test given the null hypothesis. This is done by comparing the p-value to the threshold value also known as significance level referred to by the Greek letter alpha.

Typically, the alpha value is 0.05 or 5%.

The p-value is compared to the pre-defined alpha value. The result of the experiment is significant when the p-value is less than, equal to the alpha value signifying that a change was detected, rejecting the null hypothesis.

$P(\text{test statistic} \mid H_0) > \alpha (5\%)$: failed to reject H_0

$P(\text{test statistic} \mid H_0) \leq \alpha (5\%)$: reject H_0



Let us assume, we performed a statistical hypothesis test of whether the data sample is normally distributed and calculated a p-value of 0.9, we can say that the hypothesis test found that the sample is normally distributed, failing to reject the null hypothesis at a 5% significance level.

Understanding p-value

- A p-value, or probability value, is a number describing how likely it is that your data would have occurred under the null hypothesis of your statistical test.
- P values are most often used by researchers to say whether a certain pattern they have measured is statistically significant.

- Statistical significance is another way of saying that the p value of a statistical test is small enough to reject the null hypothesis of the test.
- The most common threshold is $p < 0.05$; that is, when you would expect to find a test statistic as extreme as the one calculated by your test only 5% of the time. But the threshold depends on your field of study – some fields prefer thresholds of 0.01, or even 0.001.
- The threshold value for determining statistical significance is also known as the alpha value.

How to calculate P Value?

<https://www.geeksforgeeks.org/how-to-calculate-p-value/>

P-value, also referred to as probability value is a statistical measure used to determine whether to accept or reject the Null Hypothesis, considering the Null Hypothesis to be True. For calculating the p-value, we perform an experiment and based on the observations of the test-statistic we make decisions, i.e., if the measure is statistically below the significance level, the Null Hypothesis is rejected, specifying that the Null Hypothesis is very less likely to occur.

How it works?

Consider, there is a magazine that says, on an average there are 12% Indian employees in a company in US.

Assumptions:

- Test-statistic: An average company has 12% Indian employees, i.e., μ (Population) = 0.12.
- Null Hypothesis (H_0) : There are 12% of Indian employees in a company.
- Alternative Hypothesis (H_a) : There are more than 12% of Indian employees in a company in US.
- Significance Level: $\alpha = 0.05$ (generally a low value)
- Sample statistic: We take a company X and we check for 50 employees and find that more than 20% are Indian.

We perform n experiments, and if probability of getting the sample statistic, given there are 12% Indian employees, for n observations (p-value) is less than 0.05 we Reject the Null Hypothesis and Accept the Alternative Hypothesis, otherwise we do not Reject the Null Hypothesis.

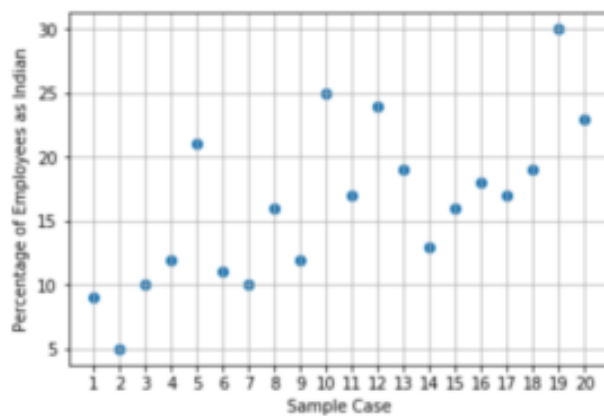
$$p\text{-value} = P(\mu(\text{sample}) > 20\% \mid H_0 \text{ is True})$$

Methods to determine P Value

By Simulation

We simulate the situation and take samples that provide our Null Hypothesis to be True. In the above example, we take 20 samples of 50 employees each, and in each sample, there are 12% of Indian employees (Null Hypothesis). We re-sample 20 samples of 50 employees each from the above-collected samples.

Below are the results of the simulation:



There are 5 samples out of 20 which had more than 20% employees as Indian.

Hence, the p-value would be $5/20 = 0.25$.

Since the significance level was 5% and after simulation, the value obtained 25%, we cannot reject the Null Hypothesis.

z Statistic

To perform z-test, three conditions are required to be met for the samples we collect:

- Random: The sampling of data to be purely random.
- Normal: The data needs to be roughly normally distributed.
- Independent: The sample must be independent from the previous sample, i.e., we need to perform the sampling with replacement, or, we can check if the sample is less than 10% of its population.

Assuming the three conditions are met, z value is calculated by looking at the standard deviation between the sample statistic and the population proportion.

For z test, the mean is not considered, instead, we take the proportions to calculate p value.

$$z = \frac{\rho(\text{Population}) - \rho(\text{Sample})}{\sqrt{\frac{\rho(\text{Sample})(1 - \rho(\text{Sample}))}{n}}}$$

Here, $\rho(\text{Population})=12\%$, $\rho(\text{Sample})=20\%$ and $n=50$ (Considering the ρ , i.e., proportion to be same as mean)

We get, $z = -0.004$

The p value is obtained from z table for above z value, which is 0.4840, i.e., roughly 48%.

Since the significance level was 5% and we obtained 48%, we cannot reject the Null Hypothesis.

t Statistic

In t test, all the three conditions are as required by z-test, i.e., the sample should be random, normal, and independent.

$$t = \frac{\mu(\text{Population}) - \mu(\text{Sample})}{\frac{\sigma(\text{Sample})}{\sqrt{n}}}$$

Here, $\mu(\text{Population})=12\%$, $\mu(\text{Sample})=20\%$, $\sigma(\text{Sample})$, i.e., the standard deviation of sample to be 5, and $n=50$.

Then, $t = -0.113$

For t-Test we look into the t table to find the p-value, the degree of freedom (df) is $n-1$, i.e., 49 and we look for a value in row 49 to be equal or greater than t, and obtain the corresponding y value to get p-value to be roughly 45%.

Since the p-value is 45% and we have a significance level to be 5%, we cannot reject the Null Hypothesis.

Note: When considering means, t-test is used, and when considering proportions, z test is used.

<https://www.analyticsvidhya.com/blog/2021/07/a-simple-guide-to-hypothesis-testing-for-dummies/#:~:text=The%20result%20of%20the%20hypothesis,statistic%20is%20unlikely%20to%20lie.>

Some More Examples

Consider the age of students in a college and in Class A

In [349]:

```
import numpy as np
import pandas as pd
import scipy.stats as stats
import math
np.random.seed(6)
school_ages=stats.poisson.rvs(loc=18,mu=35,size=1500)
classA_ages=stats.poisson.rvs(loc=18,mu=30,size=60)
```

In [350]:

```
classA_ages.mean()
```

Out[350]:

```
46.9
```

In []:

In [351]:

```
_,p_value=stats.ttest_1samp(a=classA_ages,popmean=school_ages.mean())
```

In [353]:

```
p_value
```

Out[353]:

```
1.139027071016194e-13
```

In [352]:

```
school_ages.mean()
```

Out[352]:

```
53.303333333333335
```

In [354]:

```
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 rejecting null hypothesis

Two-sample T-test With Python

The Independent Samples t Test or 2-sample t-test compares the means of two independent groups in order to determine whether there is statistical evidence that the associated population means are significantly different. The Independent Samples t Test is a parametric test. This test is also known as: Independent t Test

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$s^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 + \sum_{j=1}^{n_2} (x_j - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

In [355]:

```
np.random.seed(12)
ClassB_ages=stats.poisson.rvs(loc=18,mu=33,size=60)
ClassB_ages.mean()
```

Out[355]:

50.63333333333333

In [356]:

```
_,p_value=stats.ttest_ind(a=classA_height,b=ClassB_ages,equal_var=False)
```

In [357]:

```
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 rejecting null hypothesis

Paired T-test With Python

When you want to check how different samples from the same group are, you can go for a paired T-test

In [203]:

```
weight1=[25,30,28,35,28,34,26,29,30,26,28,32,31,30,45]
weight2=weight1+stats.norm.rvs(scale=5,loc=-1.25,size=15)
```

In [204]:

```
print(weight1)
print(weight2)

[25, 30, 28, 35, 28, 34, 26, 29, 30, 26, 28, 32, 31, 30, 45]
[30.57926457 34.91022437 29.00444617 30.54295091 19.86201983 37.57873174
 18.3299827  21.3771395  36.36420881 32.05941216 26.93827982 29.519014
 26.42851213 30.50667769 41.32984284]
```

In [205]:

```
weight_df=pd.DataFrame({"weight_10":np.array(weight1),
                        "weight_20":np.array(weight2),
                        "weight_change":np.array(weight2)-
np.array(weight1)})
```

In [206]:

```
weight_df
```

Out[206]:

	weight_10	weight_20	weight_change
0	25	30.579265	5.579265
1	30	34.910224	4.910224
2	28	29.004446	1.004446
3	35	30.542951	-4.457049
4	28	19.862020	-8.137980
5	34	37.578732	3.578732
6	26	18.329983	-7.670017
7	29	21.377139	-7.622861

	weight_10	weight_20	weight_change
8	30	36.364209	6.364209
9	26	32.059412	6.059412
10	28	26.938280	-1.061720
11	32	29.519014	-2.480986
12	31	26.428512	-4.571488
13	30	30.506678	0.506678
14	45	41.329843	-3.670157

In [358]:

```
_,p_value=stats.ttest_rel(a=weight1,b=weight2)
```

In [359]:

```
print(p_value)
```

```
0.5732936534411279
```

In [360]:

```
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
```

Correlation

In [323]:

```
import seaborn as sns
df=sns.load_dataset('iris')
```

In [326]:

```
df.shape
```

```
(150, 5)
```

Out[326]:

```
df.corr()
```

In [341]:

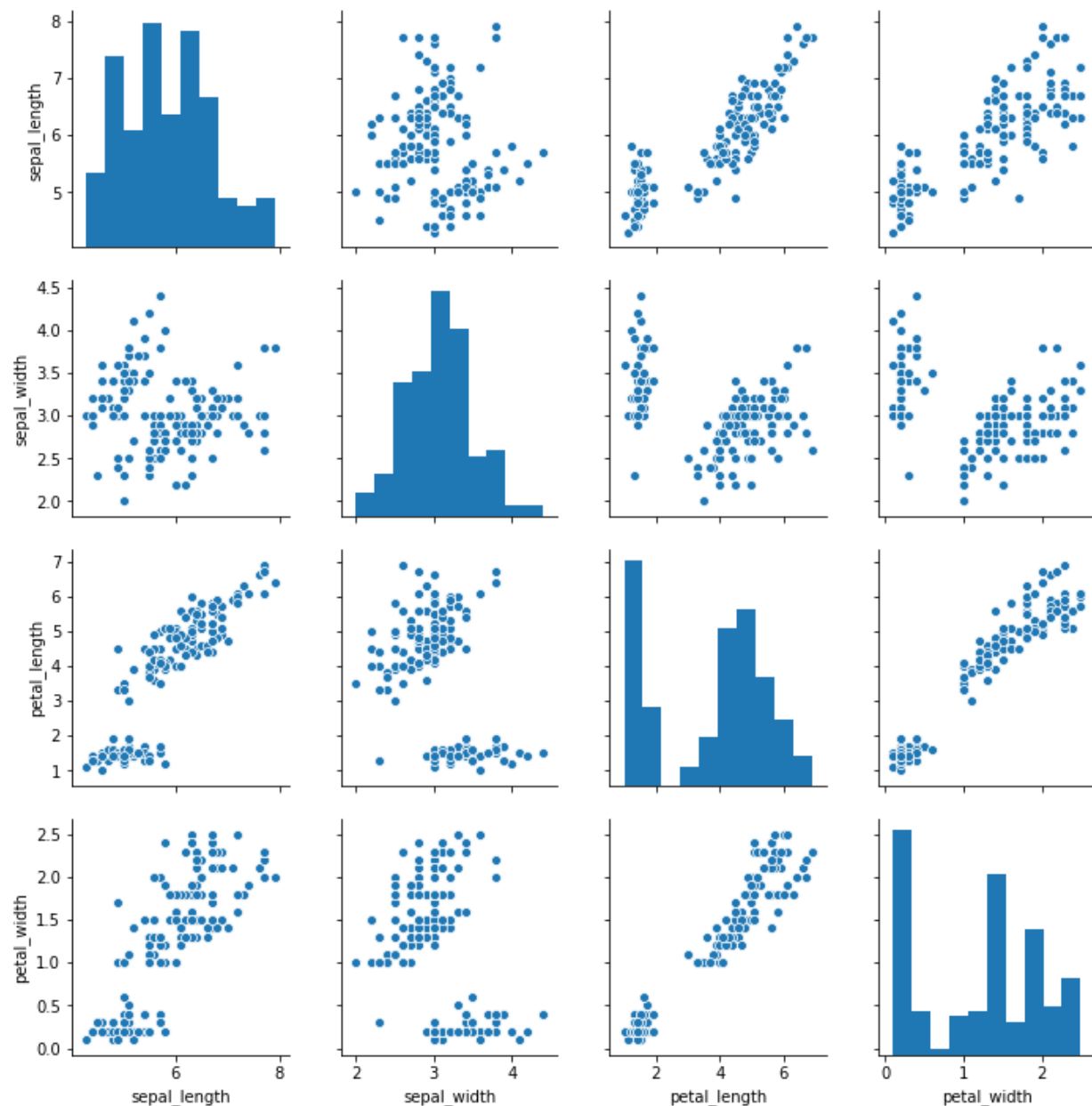
Out[341]:

	sepal_length	sepal_width	petal_length	petal_width
sepal_length	1.000000	-0.117570	0.871754	0.817941
sepal_width	-0.117570	1.000000	-0.428440	-0.366126
petal_length	0.871754	-0.428440	1.000000	0.962865
petal_width	0.817941	-0.366126	0.962865	1.000000

In [328]:

```
sns.pairplot(df)
```

Out[328]:



Anova Test(F-Test)

The t-test works well when dealing with two groups, but sometimes we want to compare more than two groups at the same time.

For example, if we wanted to test whether petal_width age differs based on some categorical variable like species, we have to compare the means of each level or group the variable

One Way F-test(Anova) :-

It tell whether two or more groups are similar or not based on their mean similarity and f-score.

Example : there are 3 different category of iris flowers and their petal width and need to check whether all 3 group are similar or not

In [283]:

```
import seaborn as sns
df1=sns.load_dataset('iris')
```

In [285]:

```
df1.head()
```

Out[285]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

In [286]:

```
df_anova = df1[['petal_width', 'species']]
```

In [288]:

```
grps = pd.unique(df_anova.species.values)
```

In [289]:

```
grps
```

Out[289]:

```
array(['setosa', 'versicolor', 'virginica'], dtype=object)
```

In [291]:

```
d_data = {grp:df_anova['petal_width'][df_anova.species == grp] for grp in
grps}
```

In [292]:

d_data

Out[292]:

```
{'setosa': 0      0.2
1      0.2
2      0.2
3      0.2
4      0.2
5      0.4
6      0.3
7      0.2
8      0.2
9      0.1
10     0.2
11     0.2
12     0.1
13     0.1
14     0.2
15     0.4
16     0.4
17     0.3
18     0.3
19     0.3
20     0.2
21     0.4
22     0.2
23     0.5
24     0.2
25     0.2
26     0.4
27     0.2
28     0.2
29     0.2
30     0.2
31     0.4
32     0.1
33     0.2
34     0.2
35     0.2
36     0.2
37     0.1
38     0.2
39     0.2
40     0.3
41     0.3
42     0.2
43     0.6
44     0.4
45     0.3
```

```
46    0.2
47    0.2
48    0.2
49    0.2
Name: petal_width, dtype: float64, 'versicolor': 50    1.4
51    1.5
52    1.5
53    1.3
54    1.5
55    1.3
56    1.6
57    1.0
58    1.3
59    1.4
60    1.0
61    1.5
62    1.0
63    1.4
64    1.3
65    1.4
66    1.5
67    1.0
68    1.5
69    1.1
70    1.8
71    1.3
72    1.5
73    1.2
74    1.3
75    1.4
76    1.4
77    1.7
78    1.5
79    1.0
80    1.1
81    1.0
82    1.2
83    1.6
84    1.5
85    1.6
86    1.5
87    1.3
88    1.3
89    1.3
90    1.2
91    1.4
92    1.2
93    1.0
94    1.3
95    1.2
96    1.3
```

```
97      1.3
98      1.1
99      1.3
Name: petal_width, dtype: float64, 'virginica': 100      2.5
101     1.9
102     2.1
103     1.8
104     2.2
105     2.1
106     1.7
107     1.8
108     1.8
109     2.5
110     2.0
111     1.9
112     2.1
113     2.0
114     2.4
115     2.3
116     1.8
117     2.2
118     2.3
119     1.5
120     2.3
121     2.0
122     2.0
123     1.8
124     2.1
125     1.8
126     1.8
127     1.8
128     2.1
129     1.6
130     1.9
131     2.0
132     2.2
133     1.5
134     1.4
135     2.3
136     2.4
137     1.8
138     1.8
139     2.1
140     2.4
141     2.3
142     1.9
143     2.3
144     2.5
145     2.3
146     1.9
147     2.0
```

```
148     2.3
149     1.8
Name: petal_width, dtype: float64}
```

In [293]:

```
F, p = stats.f_oneway(d_data['setosa'], d_data['versicolor'],
d_data['virginica'])
```

In [294]:

```
print(p)
```

```
4.169445839443116e-85
```

In [295]:

```
if p<0.05:
    print("reject null hypothesis")
else:
    print("accept null hypothesis")
```

```
reject null hypothesis
```