Comparison of Two Population

TESTS OF MEANS

Comparing Means

- For comparing means of two populations, we can use the following two alternatives assuming that the distribution of population is Normal:
  - – Paired t-test : Matched Samples
  - – Two Samples t-test : Independent Samples

Test for Matched Samples

PAIRED T-TEST

Paired-samples Scenario

- We apply this test when we have data with matched samples

Prewt	Postwt
80.7	80.2
89.4	80.1
91.8	86.4
74.0	86.3
78.1	76.1
88.3	78.1

- In the given example, we have Weight of the patient before treatment in Prewt and weight of the same patient after treatment in Postwt.
- The data in both the columns is matched samples data. Here, we will be interested in knowing whether the average weight before treatment is significantly different from that after treatment.
- In other words, we want to analyze as : **Did treatment make any impact on weight of the patients?**

Paired t-test

- Let x_i and y_i be the paired observations under study with $\mathbf{n}$ as the sample size
- Let d_i= x_i - y_i be the difference in corresponding paired observations
- Let s_d be the sample standard deviation for the difference and $\bar{d}$ be the mean of differences in samples
- Let D be the population mean for difference
- The two tailed hypotheses for the test can be written as:

  H_0 : D = 0 against H_1 : D ≠ 0
- The test statistic of paired t-test is

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

The test statistic has n-1 degrees of freedom

Paired t-test in Python

scipy.stats.ttest_rel(*a, b, axis=0, nan_policy='propagate'*) [source]

Calculate the T-test on TWO RELATED samples of scores, a and b.

This is a two-sided test for the null hypothesis that 2 related or repeated samples have identical average (expected) values.

Parameters:

a, b : *array_like*

The arrays must have the same shape.

axis : *int or None, optional*

Axis along which to compute test. If None, compute over the whole arrays, *a*, and *b*.

nan_policy : *{'propagate', 'raise', 'omit'}, optional*

Defines how to handle when input contains nan. 'propagate' returns nan, 'raise' throws an error, 'omit' performs the calculations ignoring nan values. Default is 'propagate'.

Returns:

statistic : *float or array*

t-statistic

pvalue : *float or array*

two-tailed p-value

Example : Paired t-test

- Consider the a subset of the dataset *anorexia* from the package **MASS** with Treat = Cont.

```
In [8]: anoCont = anorexia[anorexia.Treat == "Cont"]
   ...: anoCont.head()
Out[8]:
  Treat  Prewt  Postwt
0  Cont   80.7    80.2
1  Cont   89.4    80.1
2  Cont   91.8    86.4
3  Cont   74.0    86.3
4  Cont   78.1    76.1
```

- We find here, whether there is any significant difference between Prewt and Postwt.

Example : Paired t-test

- Hypothesis*:*

  H_0 : D = 0 i.e. There is no difference in weights before and after treatment. Hence, treatment **Cont** is not effective

  against

  H_1 : D ≠ 0 i.e. There is some difference in weights before and after treatment. Hence, treatment **Cont** may be effective

R Program and Output

```
In [9]: stats.ttest_rel(anoCont.Prewt,anoCont.Postwt)
Out[9]: Ttest_relResult(statistic=0.2872253910150255,
pvalue=0.7763070622194167)
```

- We observe here that the p-value is greater than 0.05, hence we are inclined to not reject H_0.
- Conclusion: The treatment **Cont** might not be effective

Example : Paired t-test

- Let us consider some other treatment in the data namely, **FT** and perform the similar test on the data

```
In [10]: anoFT = anorexia[anorexia.Treat == "FT"]
    ...: anoFT.head()
Out[10]:
   Treat  Prewt  Postwt
55    FT   83.8    95.2
56    FT   83.3    94.3
57    FT   86.0    91.5
58    FT   82.5    91.9
59    FT   86.7   100.3
```

R Program and Output

```
In [11]: stats.ttest_rel(anoFT.Prewt,anoFT.Postwt)
Out[11]: Ttest_relResult(statistic=-4.184908135290033,
pvalue=0.0007002531056005393)
```

- We observe that p-value is less than 0.05 and even less than 0.01. Hence we reject H_0 at 1% level of significance.
- Conclusion: Treatment **FT** might be effective.

One-Tailed Test

- We can also consider here the hypothesis as

$$H_0: D \geq 0 \text{ against } H_1 : D < 0$$

```
In [11]: stats.ttest_rel(anoFT.Prewt,anoFT.Postwt)
Out[11]: Ttest_relResult(statistic=-4.184908135290033,
pvalue=0.0007002531056005393)

In [16]: pvalue2tailed = stats.ttest_rel(anoFT.Prewt,anoFT.Postwt)[1]
    ...: pvalue1tailed = stats.ttest_rel(anoFT.Prewt,anoFT.Postwt)[1]/2

In [17]: pvalue2tailed
Out[17]: 0.0007002531056005393

In [18]: pvalue1tailed
Out[18]: 0.0003501265528002697
```

- We observe that p-value is less than 0.05 and even less than 0.01. Hence we reject H_0 at 1% level of significance.
- Conclusion: Treatment FT might be effective in increasing weight.

2 INDEPENDENT SAMPLES TESTS

Two Sample Tests

- Used to test whether there is a significant difference between the means of two samples.
- Here, the two samples are independent
- Two-tailed Hypotheses for variances can be stated as:

$$H_0: \sigma_1^2 = \sigma_2^2 \text{ against } H_1: \sigma_1^2 \neq \sigma_2^2$$

- Two-tailed Hypotheses for means can be stated as:

$$H_0: \mu_1 = \mu_2 \text{ against } H_1: \mu_1 \neq \mu_2$$

Two Sample Test for Variance : Bartlett's test

- This test checks the equality of variances

$$H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 \text{ against } H_1: \sigma_i^2 \neq \sigma_j^2 \ for\ alteast\ one\ pair\ (i,j)$$

Bartlett's test is used to test the null hypothesis, H_0 that all k population variances are equal against the alternative that at least two are different.

If there are k samples with sizes n_i and sample variances S_i^2 then Bartlett's test statistic is

$$\chi^2 = \frac{(N-k)\ln(S_p^2) - \sum_{i=1}^{k}(n_i - 1)\ln(S_i^2)}{1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\left(\frac{1}{n_i-1}\right) - \frac{1}{N-k}\right)}$$

where $N = \sum_{i=1}^{k} n_i$ and $S_p^2 = \frac{1}{N-k}\sum_i (n_i - 1)S_i^2$ is the pooled estimate for the variance.

The test statistic has approximately a χ^2_{k-1} distribution. Thus the null hypothesis is rejected if $\chi^2 > \chi^2_{k-1,\alpha}$ (where $\chi^2_{k-1,\alpha}$ is the upper tail critical value for the χ^2_{k-1} distribution).

Source: [Bartlett's test - Wikipedia]

Example: Two Samples Test

- The CO_2 uptake of six plants from Quebec and six plants from Mississippi was measured at several levels of ambient CO_2 concentration.
- Half the plants of each type were chilled overnight before the experiment was conducted.
- We will see whether there is a significant difference in the variances and means of Uptake in plants

Dataset: CO2

```
In [20]: co2.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 84 entries, 0 to 83
Data columns (total 5 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Plant      84 non-null     object
 1   Type       84 non-null     object
 2   Treatment  84 non-null     object
 3   conc       84 non-null     int64
 4   uptake     84 non-null     float64
dtypes: float64(1), int64(1), object(3)
memory usage: 3.4+ KB
```

Program and Output

$H_0: \sigma^2_{chilled} = \sigma^2_{non-chilled}$ $H_1: \sigma^2_{chilled} \neq \sigma^2_{non-chilled}$

```
In [5]: import pandas as pd
   ...: from scipy import stats
   ...:
   ...: co2 = pd.read_csv("CO2.csv")
   ...: co2_chill = co2[co2.Treatment == "chilled"]
   ...: co2_nonchill = co2[co2.Treatment == "nonchilled"]
   ...:
   ...: uptake_chill = co2_chill.uptake
   ...: uptake_nonchill = co2_nonchill.uptake
   ...: stats.bartlett(uptake_chill,uptake_nonchill)
Out[5]: BartlettResult(statistic=0.5315695885641828, pvalue=0.46594771841246396)
```

- We observe that p-value is greater than 0.05. Hence we cannot reject H_0 at 5 % level of significance
- Conclusion : Variances of uptakes of two treatments might be same

T-test for two Samples

- The t-test for comparison of means between two samples can be applied under two cases:
  - Two Samples having same variance
  - Two Samples having different variance
- It has got different statistics under these two cases

Two Samples With Equal Variances

- The test statistic in this case is as follows:

$$t = \frac{\overline{x_1} - \overline{x_2}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where

$\overline{x_1}$: Sample mean of sample from population 1

$\overline{x_2}$: Sample mean of sample from population 2

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 - n_2 - 1}}$$

Two Samples With Equal Variances

- The t statistic has degrees of freedom as

$$(n_1-1)+(n_2-1)= n_1+n_2-2$$

Two Samples With Unequal Variances

- The test statistic in this case is as follows:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where

$\overline{x_1}$: Sample mean of sample from population 1

$\overline{x_2}$: Sample mean of sample from population 2

s_1^2 : Sample variance of sample from population 1

s_2^2: Sample variance of sample from population 2

Two Samples With Unequal Variances

- The t statistic has degrees of freedom as

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left({s_1^2}/{n_1}\right)^2}{n_1 - 1} + \frac{\left({s_2^2}/{n_2}\right)^2}{n_2 - 1}}$$

Example : Mean Uptake

- We had seen in the example of CO2, that the variances of the uptake values are equal.
- Let us examine the means of uptake and compare them for variable Treatment

$$H_0: \mu_{chilled} = \mu_{non-chilled} \quad H_1: \mu_{chilled} \neq \mu_{non-chilled}$$

```
In [9]: stats.ttest_ind(uptake_chill,uptake_nonchill,equal_var=True)
Out[9]: Ttest_indResult(statistic=-3.0484611149819503,
pvalue=0.0030957332525416484)
```

Conclusion : The means of uptake values may not be equal for two treatments at 5% level of significance

Example : CO2

- We can also test for the following hypothesis:

$$H_0: \mu_{chilled} \geq \mu_{non-chilled} \quad H_1: \mu_{chilled} < \mu_{non-chilled}$$

```
In [16]: stats.ttest_ind(uptake_chill,uptake_nonchill,equal_var=True)
Out[16]: Ttest_indResult(statistic=-3.0484611149819503,
pvalue=0.0030957332525416484)
In [17]: pvalue2tailed =
stats.ttest_ind(uptake_chill,uptake_nonchill,equal_var=False)[1]

In [18]: pvalue1tailed =
stats.ttest_ind(uptake_chill,uptake_nonchill,equal_var=False)[1]/2

In [19]: pvalue1tailed
Out[19]: 0.0015534684495496483
```

Conclusion : The mean of uptake values of chilled plants may be lesser than mean of uptake values of non-chilled plants at 5% level of significance

Sane's STATS Academy of Statistics

Thank You