**Name: Pratik Premkishor Mantri**

**NUID: 001353214**

**Module 4R-Practise**

**Appendix:**

|  |  |
| --- | --- |
| Index | Page no |
| Part-1 | 2-4 |
| Part-2 | 5-8 |
| References | 9 |

**Part-1**

**Two-sample t-test with unequal variance:**

In statistics, Welch's t-test, or unequal variances t-test, is a two-sample location test which is used to test the hypothesis that two populations have equal means.

In this example we have used a dataset ‘cats’ from package MASS.

The **unpaired two-samples t-test** is used to compare the **mean** of two independent groups.

For example, suppose that we have measured the body weight of 100 cats: 50 male (group A) and 50 females (group B). We want to know if the mean body weight of female cat is significantly different from that of male cat.

In this case, we have two independent groups of samples. Therefore, it’s possible to use an **independent t-test** to evaluate whether the means are different.

**Null hypothesis:** Male and Female cat have the same body weight**.**

**Alternate hypothesis:** Male and Female cat do not have the same body weight**.**

The printed data can be seen in the **R file**.

**Computing summary statistics by groups of male and female**

Text, letter

Description automatically generated

**Visualize our data:**

Graphical user interface, text, application

Description automatically generated

Chart, box and whisker chart

Description automatically generated

Here we can see that the mean value of the sex ‘male(M) is higher to that of female(F).

**Preliminary test to check independent t-test assumptions:**

In this, we have two assumptions:

1. Are these two samples independent?

Yes, because the male cat and female cat are not related to each other.

2. Is the data from each of two groups follow normal distribution?

Lets check that using **Shapiro-Wilk normality test** .

Here we will be using function **with ()** and**shapiro.test()** to compute Shapiro-Wilk test for each group.

Graphical user interface, text, application, email

Description automatically generated

Here while **looking at normality,** we can say that the p-value is greater than 0.05 (level of significance) for sex “male” so it is normally distributed but not for female.

**p-value for male**= 0.119>0.05(normally distributed)

**p-value for female**= 0.0003754<0.05(not normally distributed)

**Hence, if the if the data are not normally distributed, it is**

**recommended to use the non-parametric two-samples Wilcoxon rank test.**

P-value by **the non-parametric two-samples Wilcoxon rank test=8.201e-11,** which is less than 0.05 and hence we got significant evidence to reject the null hypothesis and select the alternative hypothesis which is male and female cat samples does not have the same bodyweight (“Bwt”).

**Part-2**

In this example we will be seeing that does meditation effects the sleep quality. The two samples are related groups of samples and are dependent on each other, and we know The **paired samples t-test** is used to compare the means between two related groups of samples.

As in the given problem, 10 students have done meditation workshop. We want to know whether the meditation workshop has an impact on the sleeping scores of the students. So from that we can say that both the samples are dependent on each other.

To answer to this question, the have taken 10 students and measured their sleeping score before and after a meditation.

**Null hypothesis**: meditation influences sleep quality

**Alternate hypothesis**: meditation does not influence sleep quality.

**Print the data:**

Table

Description automatically generated with medium confidence

**Compute summary statistics:**

Text

Description automatically generated with medium confidence

The mean values are

After: 6.87

Before: 6.25

**Visualize the data:**

Chart, box and whisker chart

Description automatically generated

Here, Box plots show you the increase, but lose the paired information. You can use the function **plot.paired** to plot paired data.

We should now use paired data:

Text

Description automatically generated

Chart, line chart

Description automatically generated

From the above figure we can say the **data is paired.**

**Preliminary test to check paired T-test assumptions:**

**1. Are these two samples paired?**

Yes, they are paired. Since the data have been collected from measuring twice the sleeping scores of the same students.

**2. is this a large sample?**

No, we have a sample size is less than 30.

**3. Check for normality**

Use Shapiro-Wilk to check for normality.

* Null hypothesis: the data is normally distributed.
* Alternative hypothesis: the data is not normally distributed.

Graphical user interface, text, application, email

Description automatically generated

P-value: 0.2177>0.05(level of significance).

Hence, we can not reject the null hypothesis and the data is normally distributed.

**Compute the paired T-test.**

Graphical user interface, text

Description automatically generated

In the result above:

* **t** is the **t-test statistic** value (t = -1.9481),
* **df** is the degrees of freedom (df= 9),
* **p-value** is the significance level of the **t-test** (p-value = 0.08322).
* **conf.int** is the **confidence interval** (conf.int) of the mean differences at 95% is also shown (conf.int= [-1.3399,0.099)
* **sample estimates** is the mean differences between pairs (mean = -0.62).

Here the p-value is greater than the level of significance 0.05, so,we do not have significant evidence to reject the Null hypothesis. Hence, we can say meditation influences sleep.

**If we change the level of significance from 0.05 to 0.1 then,**

We have a **p-value =0.08<0.1(level of significance)**

Hence, we can **reject** the Null hypothesis and select the alternate hypothesis which is meditation does not influence sleep quality..

**Reference:**

M, S. (2005). Paired samples t-test in r. Retrieved March 28, 2021, from <http://www.sthda.com/english/wiki/paired-samples-t-test-in-r>

R, M. (2020, December 28). T test (student's t-test): Definition and examples. Retrieved March 28, 2021, from <https://www.statisticshowto.com/probability-and-statistics/t-test/>