Paired sample t-test

A Paired/dependent sample t-test compares means from the same group at different times.

Here we will use a sample set of 10 performance scores of kids in the morning and evening on a certain game.

Load data

.import excel "student_t_test\data\t-test2.xls", sheet("Paired-sample t-test") firstrow

//Browse the imported dataset
.Browse

Data Editor (Edit) - [Untitled]

File Edit View Data Tools

Morning[1] 13.2

	Morning	Evening
1	13.2	14
2	8.2	8.8
3	10.9	11.2
4	14.3	14.2
5	10.7	11.8
6	6.6	6.4
7	9.5	9.8
8	10.8	11.3
9	8.8	9.3
10	13.3	13.6

Let's summarize data

```
. summarize Morning Evening

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
     Morning |         10       10.63    2.451326        6.6       14.3
     Evening |         10       11.04    2.518465        6.4       14.2
```

As seen above, the mean for morning performance is 10.63 and that for the evening is 11.04.

Question

Is there a statistically significant difference between the morning and evening performance?

Hypothesis

H0: There's no difference between morning and evening performance

Ha: There's a statistically significant difference between morning and evening performance

The level of significance

alpha = 0.05

Assumptions

Determine if data meets requirements to perform an dependent samples t-test.

Assumption #1: Your dependent variables should be measured on a continuous scale.

Assumption #2: You should have a dependence on observations.

Assumption #3: There should be no significant outliers.

Assumption #4: Your dependent variables should be approximately normally distributed.

Assumption #5: There needs to be homogeneity of variances.

CHECK FOR OUTLIERS

Check outliers by plotting a boxplot

. graph hbox Morning Evening

As seen from the above boxplot we don't have outliers from our sample.

<u>NORMALITY TEST</u>

Normality Law test using Skewness Kurtosis test for normality

H0: The data follows a normal distribution.

Ha: The data does not follow a normal distribution.

. sktest Morning Evening

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	joint adj chi2(2)	joint Prob>chi2
Morning	10	0.9784	0.6027	0.27	0.8730
Evening	10	0.5688	0.8643	0.34	0.8418

Normality Law test using Shapiro-Wilk W test for normal data

H0: The data follows a normal distribution.

Ha: The data does not follow a normal distribution.

. swilk Morning Evening

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
Morning	10	0.96240	0.579	-0.888	0.81286
Evening	10	0.94815	0.799	-0.376	0.64665

Normality Law test using Shapiro-Francia W' test for normal data

H0: The data follows a normal distribution.

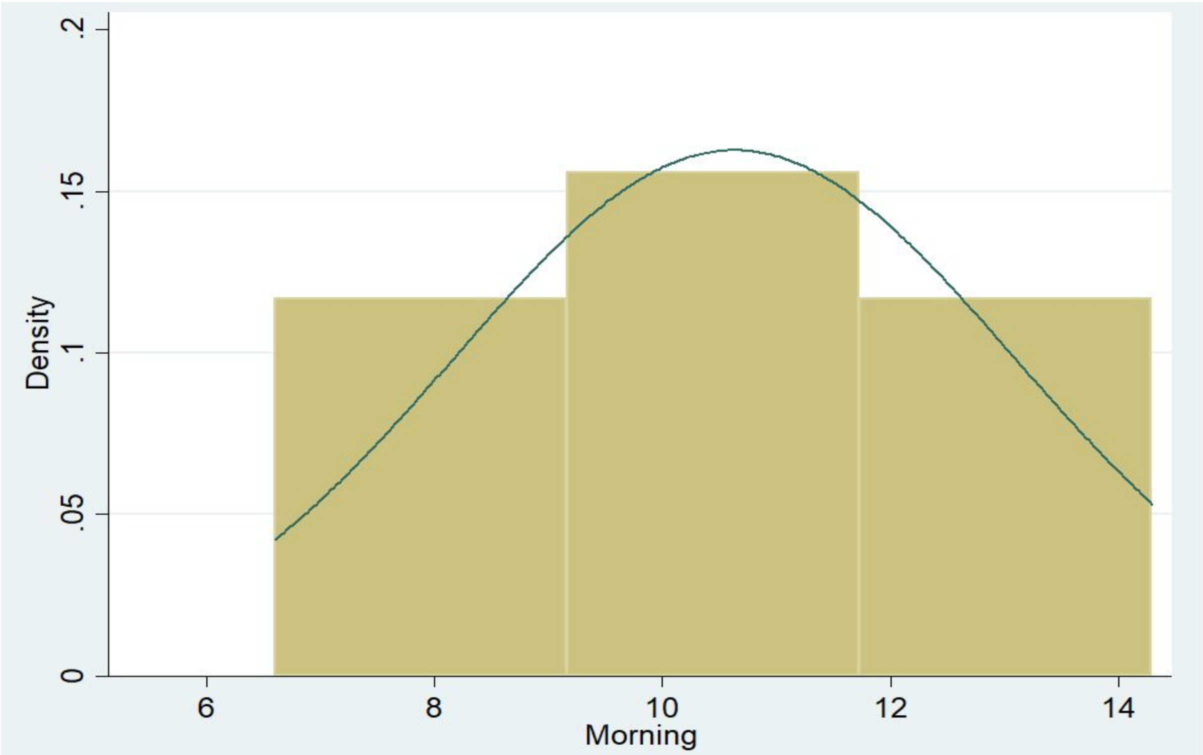
Ha: The data does not follow a normal distribution.

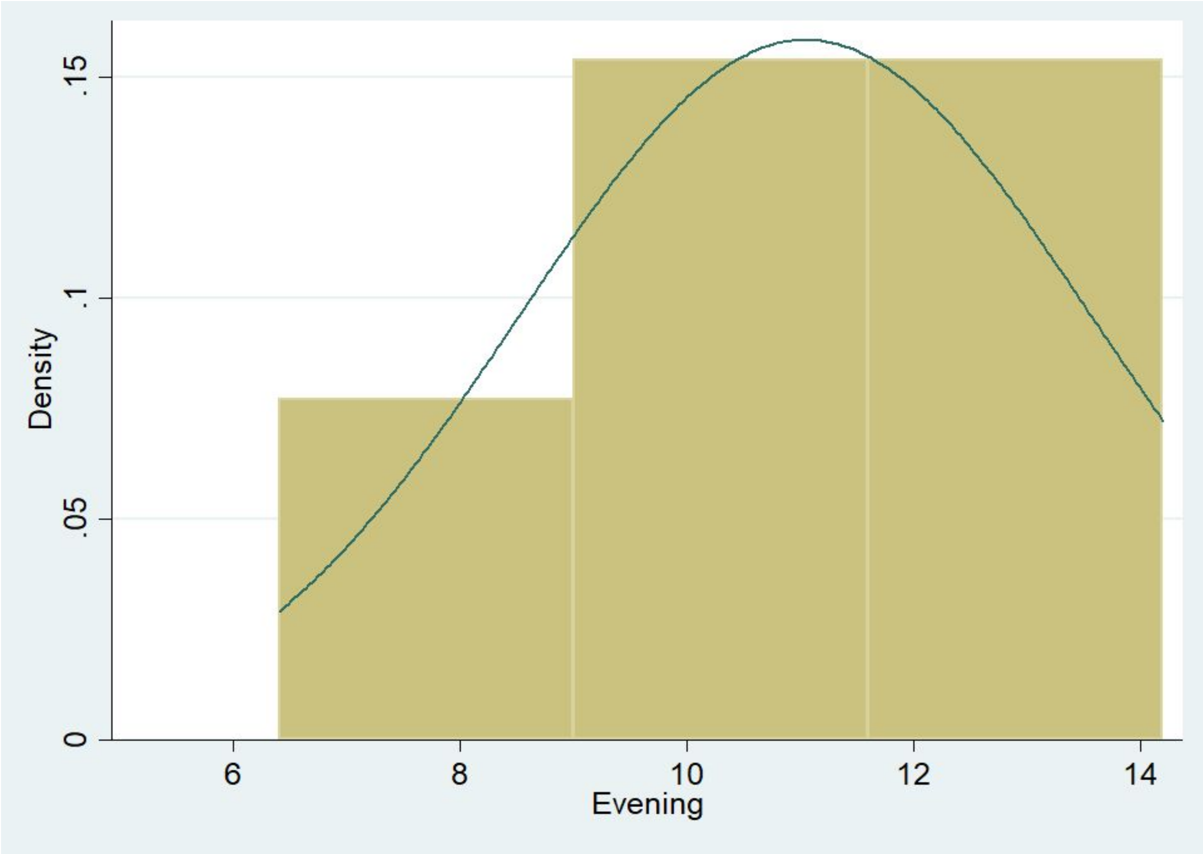
```
. sfrancia Morning Evening

                  Shapiro-Francia W' test for normal data

    Variable |       Obs       W'          V'        z       Prob>z
-------------+------------------------------------------------------
     Morning |        10    0.97022      0.498    -1.211    0.88698
     Evening |        10    0.95685      0.722    -0.566    0.71421
```

As you see in all the above 3 tests (off course you don't need all the 3), they all are NOT significant and so we have no evidence to reject the H0, that states that the data follow a normal distribution. We can plot a histogram to visualize;





Paired Sample T-test

With all data requirements for Paired Sample T-test satisfied, let us not run the test.

```
. ttest Evening == Morning

Paired t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Evening |      10       11.04    .7964086    2.518465    9.238399     12.8416
 Morning |      10       10.63    .7751774    2.451326    8.876427    12.38357
---------+--------------------------------------------------------------------
    diff |      10         .41    .1224291    .3871549    .1330461    .6869539
------------------------------------------------------------------------------
     mean(diff) = mean(Evening - Morning)                         t =   3.3489
 Ho: mean(diff) = 0                              degrees of freedom =        9

 Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
 Pr(T < t) = 0.9957         Pr(|T| > |t|) = 0.0085          Pr(T > t) = 0.0043
```

A paired sample t-test was used to analyze the performance of kids in the morning and evening to test if there was a significant difference. The Performance in the evening was higher (11.04 ± 2.52 units) compared to the morning performance (10.63 ± 2.45 units). The test was significant at alpha=0.05 and so there was enough evidence to reject the null hypothesis and conclude that there was a statistically significant improvement in performance (t(9)=-3.35, p= 0.0085) of 0.41(95% CI, 0.1330461 to 0.6869539) units.