PMC500/PLG 500 Penaakulan Statistik dalam Pendidikan

Statistical Reasoning in Educational Research

Dr Nooraida Yakob
Dr Nor Asniza Ishak
Dr Muzirah Musa

PAIRED T-TEST

- Previously, we have dealt with testing for the difference between two population means using independent samples.
- Independent samples are samples such that the selection of the sample from the first population does not affect the selection of the sample from the second population. That is to say our samples are chosen independently of each other. Even the sample size for each, need not be equal.
- Now, instead of independent samples, we are going to use dependent samples where samples are drawn in such a way that observations in the first sample are directly related to the observations in the second sample or they occur as pairs of values.

EXAMPLES OF DEPENDENT SAMPLES (Paired-samples)

CONTOH SAMPEL BERSANDAR (sampel berpasangan)

- For example, a personal trainer want to investigate whether his "weight loss program" successful or not. To do so, suppose he select a sample of 15 members of his health club and record their weights before and after the program.

- In this example, both sets of data (weight before and weight after the program) are collected from the same 15 persons. Thus, although there are two samples, they contain the same 15 persons. This is an example of paired samples (dependent samples or matched samples)

EXAMPLES OF DEPENDENT SAMPLES (Paired-samples)

CONTOH SAMPEL BERSANDAR (sampel berpasangan)

- Skor min bagi sesuatu ujian bagi sekumpulan pelajar, sebelum dan selepas menjalani latihan tertentu

  (Mean test scores for one group of students, before and after receiving training)

- Min bacaan tekanan darah sebelum dan selepas memakan ubat

  (Mean blood pressure reading before and after receiving some medication)

T-TEST FOR DEPENDENT SAMPLES

- Ujian t bagi sampel bersandar (*t-test for dependent samples*):
  - Ujian t sampel berpasangan *(Paired samples t-test, Correlated Groups Design, Within-Subjects Design, Repeated Measures, ...)*
  - Dalam IBM SPSS, istilah yang digunakan ialah *Paired samples t-test*

Ujian *t* sampel berpasangan (*Paired-Samples t-test*)

- *Repeated-Measures Design* atau *within-subjects design* adalah rekabentuk kajian yang mana:
  - data terdiri daripada dua set yang diperolehi daripada individu yang sama dalam sampel. Ini bermakna setiap individu dalam sampel pertama adalah individu yang sepadan dalam sampel kedua

    (*When you have two sets of scores from the same person in your sample. This means that each individual in one sample has a corresponding individual in the second sample... often the same person, or organization etc)*

- Lazimnya dua kumpulan individu mengalami kedua-dua situasi variabel

  (based on groups of individuals who experience both conditions of variables)
  - sampel diukur sebelum dan selepas rawatan

    (*participants measured before a treatment is applied and after the treatment*)

In Paired samples t-test, it involved…

- One categorical independent variables (before and after)
- One dependent variable
- Remember :
  - *Data must be interval or ratio or ordinal by choice*
  - *Assumed distribution – normally distribution*

EXAMPLES OF PAIRED T-TEST

a) To determine the productivity level of each worker before and after a training program.

b) To determine students' achievement using a new teaching method by having pre-test and post-test.

c) To determine the effectiveness of a new slimming pill on a group of overweight people by weighting their weights before and after using the pills.

Paired-samples *t*-test: Assumptions and Requirements

- *Variable is interval level (ordinal with caution)*

- *Variable is normally distributed*
  - Acceptable degree of skewness and kurtosis

    or
  - Using the Central Limit Theorem
  - Additional assumption : the difference between the two scores obtained for each subject should be normally distributed (but it can be violated, if the sample size more than 30)

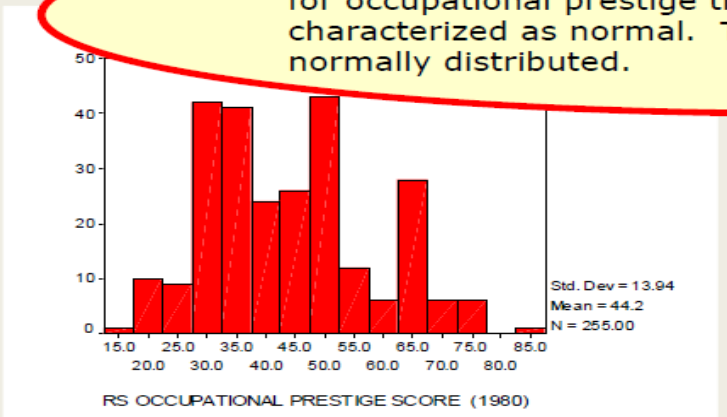
Assumption of Normality: Evaluating Normality

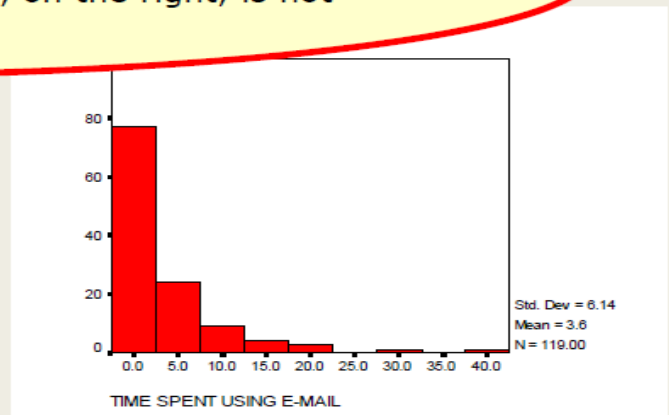
There are both graphical and statistical methods for evaluating normality.

- *Graphical methods include the histogram and normality plot.*
- *Statistical methods include diagnostic hypothesis tests for normality, and a rule of thumb that says a variable is reasonably close to normal if its skewness and kurtosis have values between –1.0 and +1.0.*

Assumption of Normality: Histograms and Normality Plots

On the left side of the slide is the histogram and normality plot for occupational prestige that could reasonably be characterized as normal. Time using email, on the right, is not normally distributed.





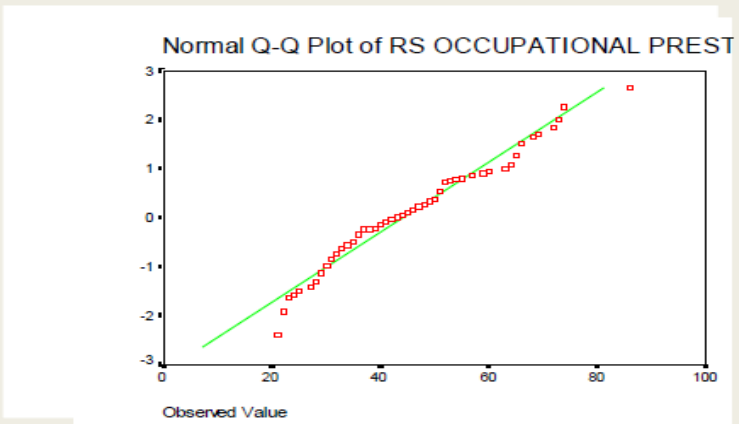


Figure 1 – Skewness and kurtosis (SPSS output)

Descriptive Statistics

	N	Skewness		Kurtosis	
	Statistic	Statistic	Std. Error	Statistic	Std. Error
Final test score	50	.256	.337	.453	.662
Valid N (listwise)	50				

Figure 2 – Histogram with normal curve plotted (SPSS output)

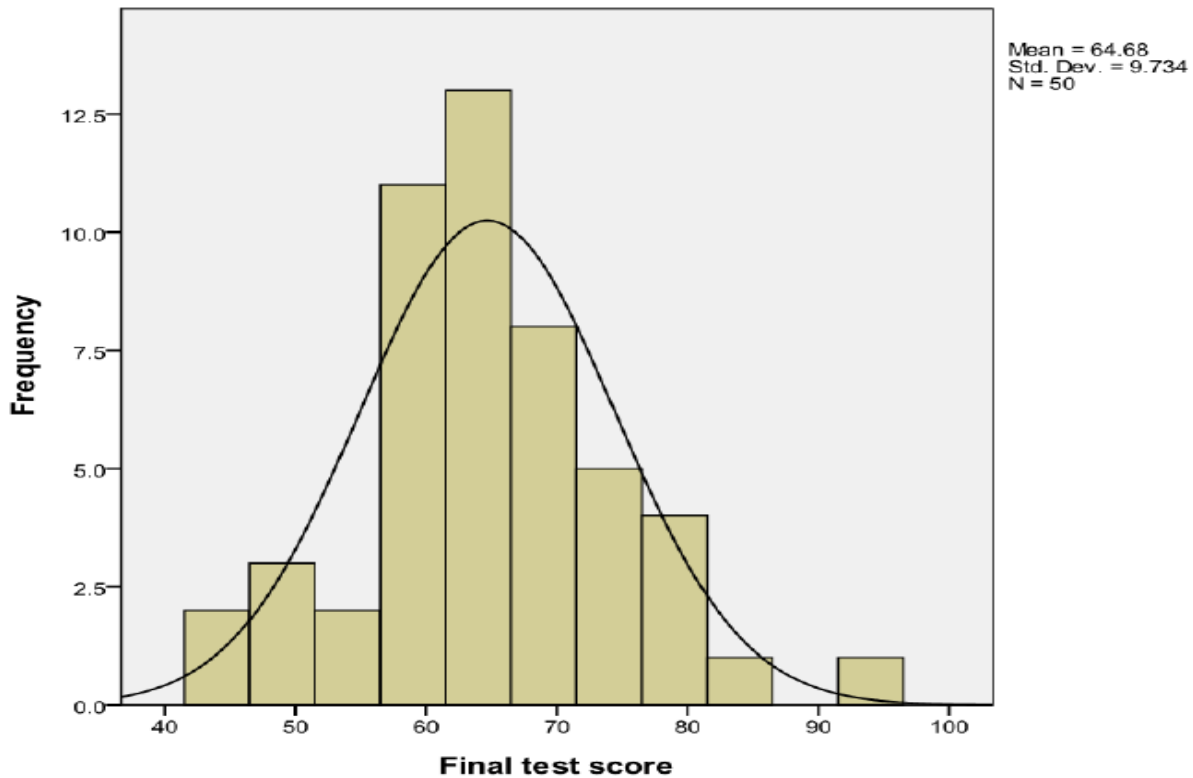


Figure 3 – SPSS output for Kolmogorov-Smirnov and Shapiro-Wilk tests of normality

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Final test score	.075	50	.200*	.986	50	.803

a. Lilliefors Significance Correction
*. This is a lower bound of the true significance.

PAIRED T-TEST

Hypothesis Testing

1. State the research question.
2. State the statistical hypothesis.
3. Set decision rule. $\alpha = .05$
4. Calculate the test statistic.
5. Decide if result is significant.
6. Interpret result as it relates to your research question.

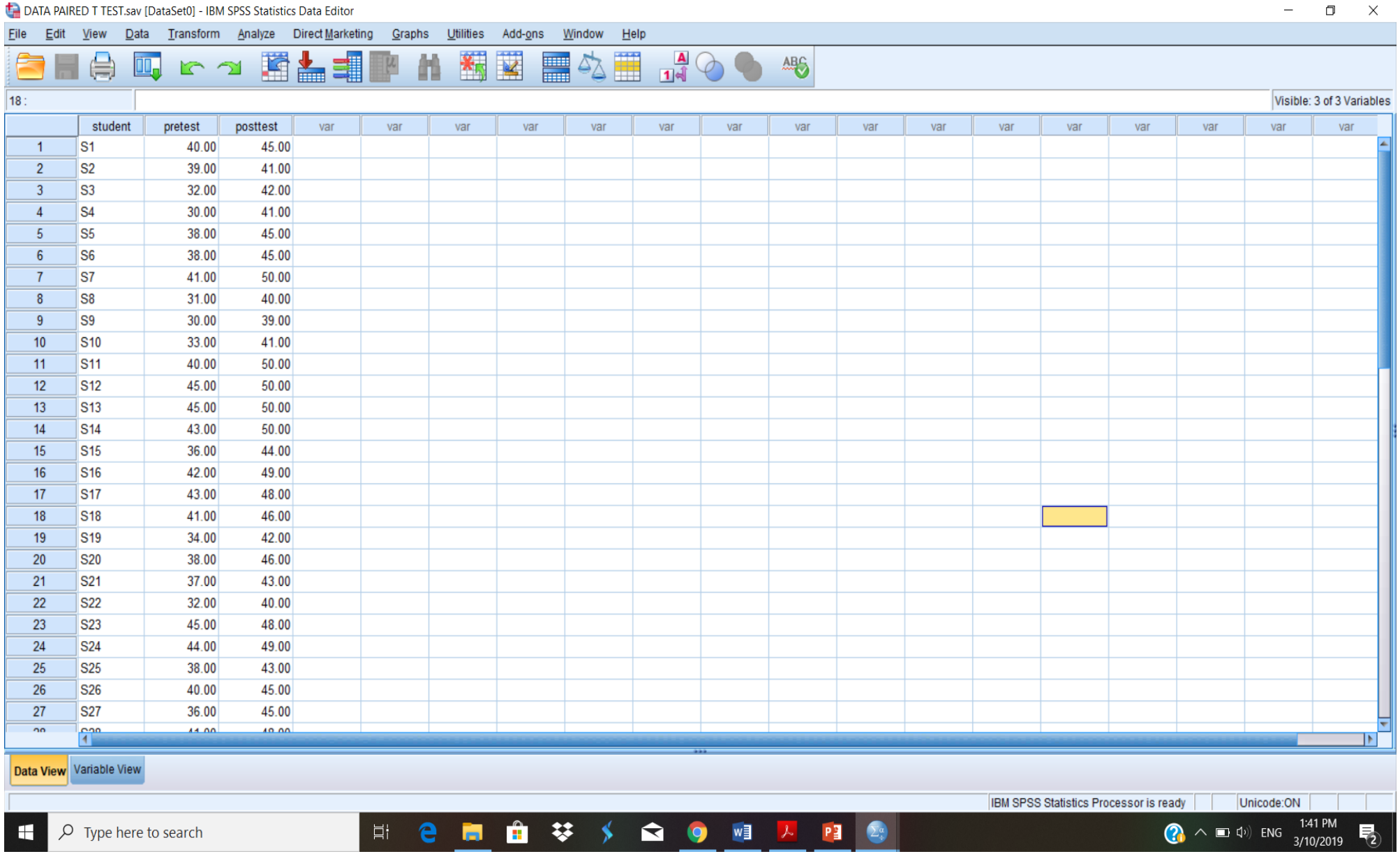
A Population of Difference Scores with a Mean of 0

- The null hypothesis in a repeated-measures design is that on the average there is *no difference* between the two groups of scores.
- This is the same as saying that the mean of the sampling distribution of difference scores is 0.

$$H_O : \mu_D = 0$$
$$H_A : \mu_D \neq 0$$

example

- A lecturer want to investigate the effect of new teaching method on the student's achievement to learn statistics. By using SPSS file, there are score (pre-test and post-test) obtained before and after the intervention.
  1. State null and alternative hypothesis.
  2. State the assumptions.
  3. What is your conclusion.



	student	pretest	posttest	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var
1	S1	40.00	45.00																
2	S2	39.00	41.00																
3	S3	32.00	42.00																
4	S4	30.00	41.00																
5	S5	38.00	45.00																
6	S6	38.00	45.00																
7	S7	41.00	50.00																
8	S8	31.00	40.00																
9	S9	30.00	39.00																
10	S10	33.00	41.00																
11	S11	40.00	50.00																
12	S12	45.00	50.00																
13	S13	45.00	50.00																
14	S14	43.00	50.00																
15	S15	36.00	44.00																
16	S16	42.00	49.00																
17	S17	43.00	48.00																
18	S18	41.00	46.00																
19	S19	34.00	42.00																
20	S20	38.00	46.00																
21	S21	37.00	43.00																
22	S22	32.00	40.00																
23	S23	45.00	48.00																
24	S24	44.00	49.00																
25	S25	38.00	43.00																
26	S26	40.00	45.00																
27	S27	36.00	45.00																

- 1. state the null and alternative hypothesis

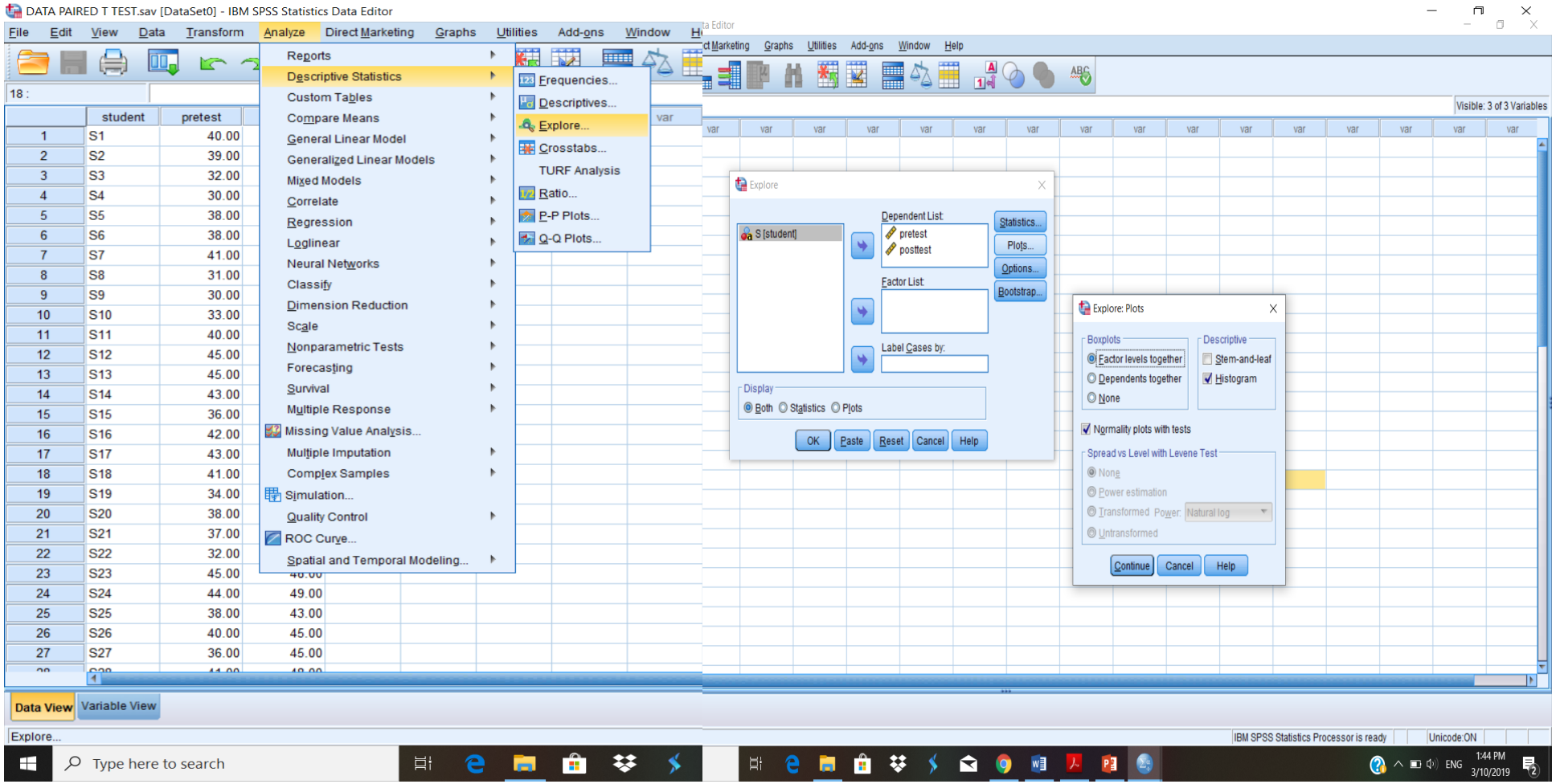
Null hypothesis:

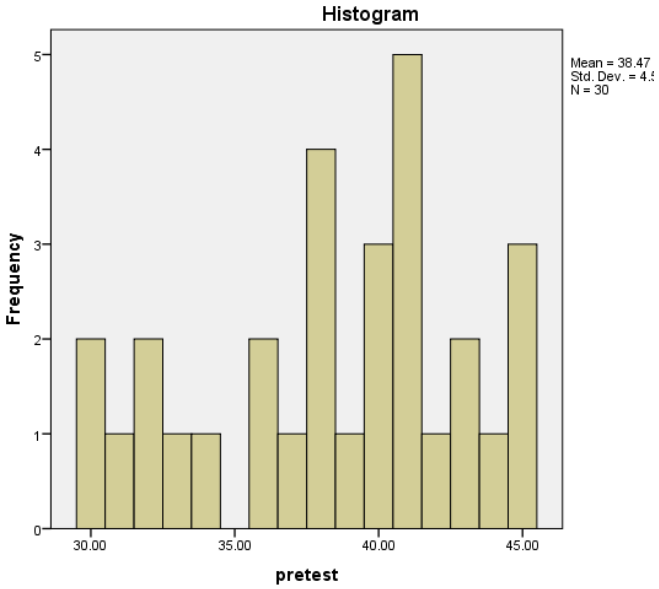
There is **no significant difference** between the mean score pre-test and post-test of student’s achievement to learn statistics.

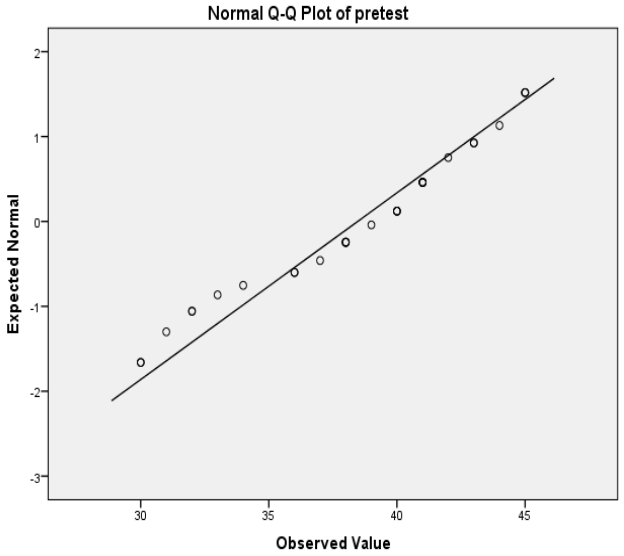
Alternative hypothesis:

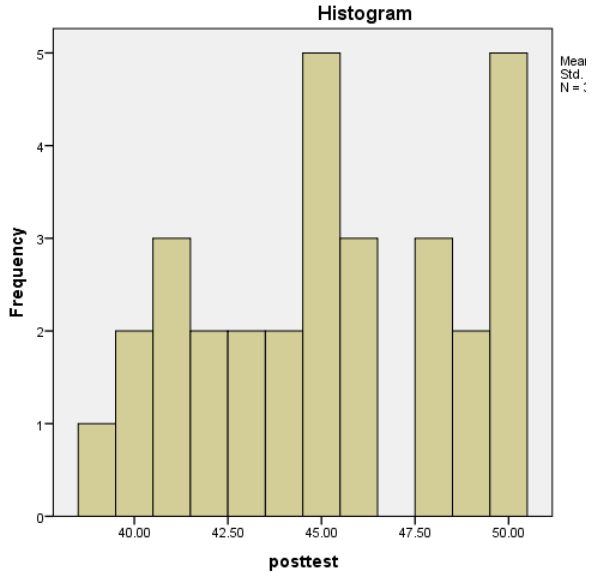
There is **a significant difference** between the mean score pre-test and post-test of student’s achievement to learn statistics.

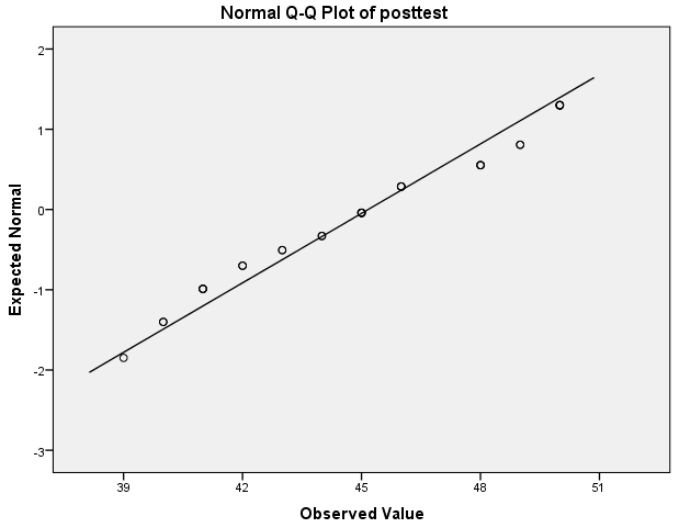
- 2. assumption –check normality











Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
pretest	.132	30	.193	.938	30	.080
posttest	.127	30	.200*	.934	30	.064

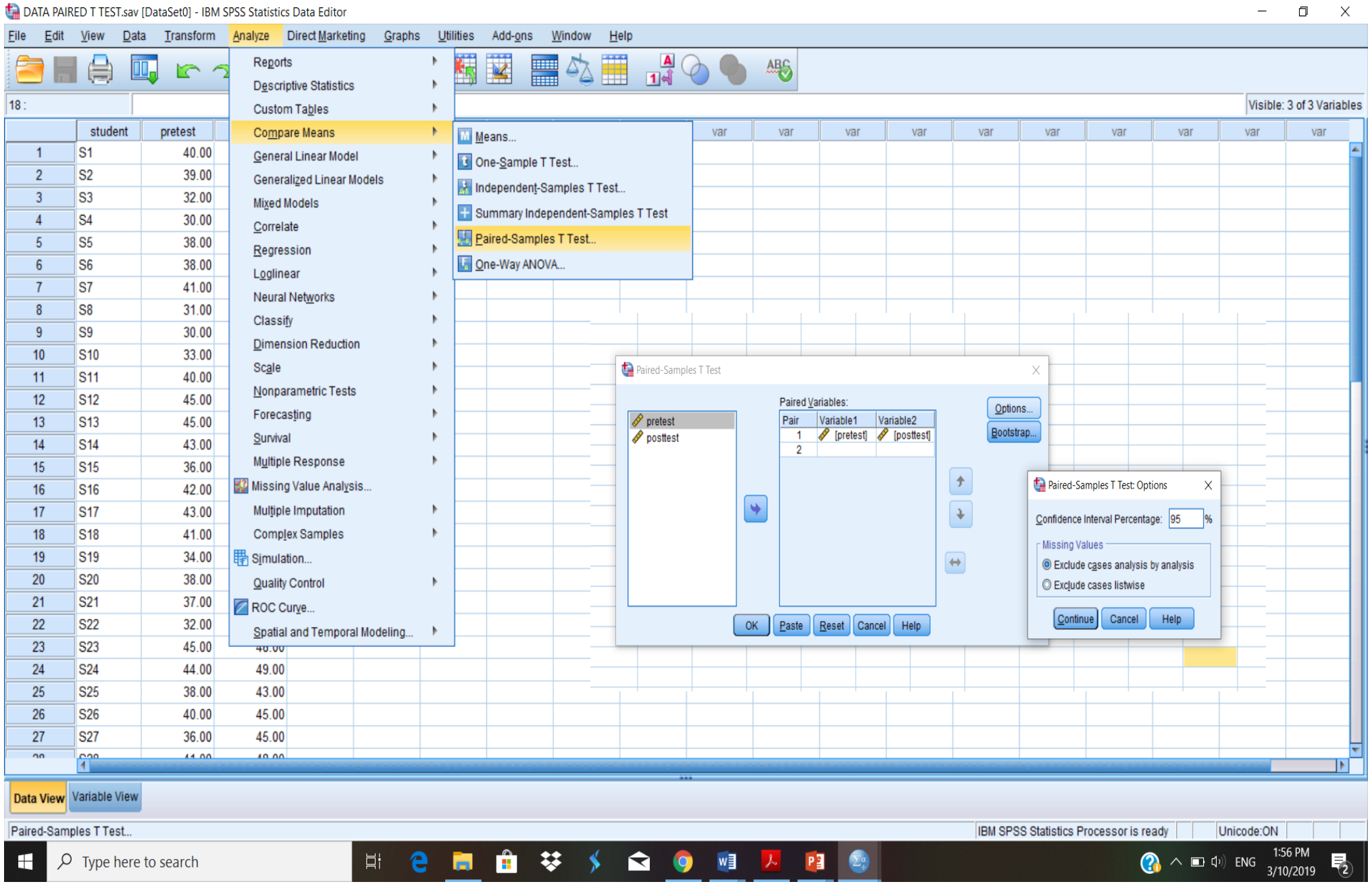
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

The p value for PRETEST is .193 (p > 0.05), then the distribution of this data is NORMAL.

The p value for POSTTEST is .200 (p > 0.05), then the distribution of this data is NORMAL.

The points form a linear pattern, thus the distribution is NORMAL.


DATA PAIRED T TEST.sav [DataSet0] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help
Reports
Descriptive Statistics
Custom Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
ROC Curve...
Spatial and Temporal Modeling...
Means...
One-Sample T Test...
Independent-Samples T Test...
Summary Independent-Samples T Test
Paired-Samples T Test...
One-Way ANOVA...
Visible: 3 of 3 Variables
student pretest
S1 40.00
S2 39.00
S3 32.00
S4 30.00
S5 38.00
S6 38.00
S7 41.00
S8 31.00
S9 30.00
S10 33.00
S11 40.00
S12 45.00
S13 45.00
S14 43.00
S15 36.00
S16 42.00
S17 43.00
S18 41.00
S19 34.00
S20 38.00
S21 37.00
S22 32.00
S23 45.00
S24 44.00 49.00
S25 38.00 43.00
S26 40.00 45.00
S27 36.00 45.00
Paired-Samples T Test
pretest
posttest
Paired Variables:
Pair Variable1 Variable2
1 [pretest] [posttest]
2
Options...
Bootstrap...
OK Paste Reset Cancel Help
Paired-Samples T Test: Options
Confidence Interval Percentage: 95 %
Missing Values
Exclude cases analysis by analysis
Exclude cases listwise
Continue Cancel Help
Data View Variable View
Paired-Samples T Test...
IBM SPSS Statistics Processor is ready
Unicode:ON
Type here to search
ENG 1:56 PM 3/10/2019

Decision…

What is your decision of this test?

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	pretest - posttest	-6.70000	2.24607	.41007	-7.53870	-5.86130	-16.339	29	.000

Since, the **p value is 0.00 (p < 0.05)**, therefore, the null hypothesis is **rejected. Hence,** There is **a significant difference** between the mean score pre-test and post-test of student's achievement to learn statistics.

Interpretation of results

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	pretest	38.4667	30	4.54657	.83009
	posttest	45.1667	30	3.46493	.63261

The mean is very important.
These means suggest that after having a new teaching method (intervention), the mean score of student's achievement was higher after the intervention.

Usually paired sample t-test automatically computes the correlation between two sets of scores.

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	pretest & posttest	30	.877	.000

The result showed that, there is a significant positive relationship between the pre test and post test score after the intervention (r=0.877 , p=0.00)

How to write a report/result (APA style)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	pretest	38.4667	30	4.54657	.83009
	posttest	45.1667	30	3.46493	.63261

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	pretest - posttest	-6.70000	2.24607	.41007	-7.53870	-5.86130	-16.339	29	.000

- A paired samples t-test was performed comparing the pre-test and post-test score of the student's achievement to learn statistics. The result showed that the mean difference is negative (M= -6.70, SD=2.25). This revealed that there is a significant difference $t(29)= -16.339$, $p < 0.05$ between the pre-test and post-test score on the student's achievement in learning statistics.

Conclusion

Thus, the finding indicated that there is an effect of new teaching method on the student's achievement to learn statistics.

Or

The new teaching method give an effect on the student's achievement to learn statistics.

Or

The student's achievement in learning statistics shows an improvement after having a new teaching method.

Thank you