THE EFFECT OF USING MATHEMATICAL SOFTWARE IN UNDERSTANDING MATERIAL IN STATISTICS COURSES

AUTHORS INFO

Akbar Nasrum
Universitas Sembilanbelas November Kolaka
akbar.nasrum@gmail.com
+6282296296096

Supratman
Universitas Sembilanbelas November
supratmanmathusnkolaka@gmail.com
+6285333327748

Suggetion for the Citation and Bibliography
Citation in Text:
Nasrum & Supratman (2019)
Bibliography:
Nasrum, A. & Supratman (2019). Effec The Effect of Using Mathematical Software in Understanding Material in Statistics Courses. *Journal of Mathematics Education, 4* (1, June), 1-7.

Abstract

This research was conducted at the Sembilanbelas November Kolaka University. This study aims to show the effect of the use of software in learning on the understanding of statistical material. The material discussed is the data normality test. To achieve that goal the researcher measures the understanding of student material before and after use of the software. Next, data analysis is done by comparing the results of the pre-test and post-test that have been obtained. The final result of the study shows that use mathematical software has a very significant effect of statistical material understanding, especially in data normality testing.

Key words: Mathematical software, learning media, Statistics

A. Introduction

Statistics Learning (3 credits) at the Sembilanbelas November Kolaka University which only relied on theory caused the students' understanding of this subject still lowest, whereas this is one of the most important course. The knowledge gained from this course will be used when writing scientific papers or writing a final assignment report (thesis). The theory (2 credits) that is equipped with practice (1 credit) will make students more understanding of the material provided.

As a mathematics student who wants to make a final assignment (proposal/thesis), the basic concepts of data analysis, both descriptive analysis and inferential analysis must be understood so that the conclusions made are correct. Before till the inferential stage, there are several data requirements that usually must be fulfilled, consist of: the data must be normal distributed, the data must be homogeneous, and the data must be linear. In quantitative study or experimental studies in the field of education often use the average parameter to get a

conclusion. Because the average parameter is not robust, it requires normal data assumptions for data analysis. Therefore the concept of data normality testing must be truly understood.

In fact, almost all of the final year students at the Sembilanbelas November Kolaka University mathematics education department did not understand the concept of testing the normality of the data itself. There are many ways that can be done to data normality test, among others using the Kolmogorov-Smirnov test, Chi-Square test, Anderson Darling test, Cramer Von Mises, Lilliefors, Shapiro Wilk and much more (Yap & Sim, 2011). In practice they only use the Kolmogorov-Smirnov test and Chi Square in SPSS. They test the data normality using statistical software (SPSS, Minitab, Easy Fit, Geogebra) but do not know what the normality test concept was like.

To overcome this problem, researchers want to use these software in learning to help an understanding of the concept of data normality test. Each software has advantages and disadvantages, therefore all the software will be used to complement each other. Hopefully to be achieved the understanding of the students about the concept of normality testing to be better.

Based on the description above, the researcher will propose research with the title "The Effect of Using Mathematical Software in Understanding Material in Statistics Courses". This research is important to be carry out so that data analysis in the preparation/ creation of student research reports (Thesis) becomes better.

B. Literature Review

Normal data is one of the requirements that must be fulfill to perform statistical inference if using parametric statistics. In educational research it is often reveal that the data of a student group forms a normal curve (Sundayana, 2015). This normal data assumption must be tested to find out whether the empirical data obtained in the field corresponds to a particular theoretical distribution. In this case the distribution is normal.

Data Normality Test can be done by various methods such as Anderson Darling test, Kolmogorov-Smirnov test, Chi-Square test, Liliefors, Sahapiro-Wilk, Kuiper, Ajne and the others (Yazici & Yolacan, 2007). These methods each have different differences and abilities in detecting deviations from normal distribution. Apart from that, the principles used in testing a data with normal distribution also vary. The following will explain two methods from the above methods, namely the Kolmogorov-Smirnov method, and the QQ-Plot.

1. Kolmogorov-Smirnov Method

The principle of the test for normality using Kolmogorv-Smirnov is to find the biggest deviation from the cumulative distribution function of observation data (empirical) on the theoretical cumulative distribution function. If the biggest deviation is not too large ($D < D_{tab}$) then the observation data can be categorized as normal distribution. Conversely, if the maximum deviation is very large ($D > D_{tab}$) then observation data said not normally distributed. The steps to test for normality using Kolmogorov Smirnov are follows:

1. Sort observation data from the smallest to the largest;
2. Create a list of cumulative data frequencies and then specify the cumulative proportion $F_{n_i}(x)$;
3. To facilitate the calculation as well as the list $F_{n_{i-1}}(x)$;
4. Convert the value of x_i value to z_i value:
5. Determine the area of the curve below the normal curve $F(z_i)$ (Theoretical Cumulative Distribution Function Value);
6. Find the value of $a_i = \left| F(z_i) - F_{n_{i-1}}(x) \right|$ and the value of $b_i = \left| F(z_i) - F_{n_i}(x) \right|$
7. Specify a value of $D = \max_{1 \le i \le n}(a_i, b_i)$
8. Make conclusion

(Nasrum, 2018)

2. QQ Plot Method

An easy way to test data normality is to use the graph method. The Quantil-Quantil normal plot or common abbreviation as QQ-Plot is the most common and effective diagnostic tool for checking data normality (Razali & Wah, 2011). The normal criteria for whether or not the data is based on the shape of the observe points according to the quintiles. If the points formed tend a straight line, then it can be said that the data is normally distributed. But line straightness is

the subjectivity of the examiner. To be more convince, a correlation test was used between the observation data and the standard normal quantil data. The steps are follows:

1. Formulating Hypotheses
   H_0 : The sample comes from a normal distribution population
   H_1 : The sample comes from a abnormal distribution population
2. Sort the initial observation data $x_1, x_2, x_3, \ldots\ldots, x_n$
3. Determine the probability value $\frac{(1-0,5)}{n}, \frac{(2-0,5)}{n}, \ldots\ldots, \frac{(n-0,5)}{n}$
4. Calculates the normal standard of quantiles $q_{(1)}, q_{(2)}, \ldots\ldots, q_{(n)}$
5. Plot the pair of observation data $(q_{(1)}, x_{(1)}), (q_{(2)}, x_{(2)}), \ldots\ldots, (q_{(n)}, x_{(n)})$ and test the significance of the lines formed through the correlation coefficient test
6. Determine Critical r according to sample size (n)
7. If $r_{hit} > r_{tab}$ then H_0 is accepted or the sample comes from a normal distribution population
   Otherwise if $r_{hit} \leq r_{tab}$ then H_0 rejected.

(Kadir, 2015)

3. Mathematical Software

The use of mathematical software is one of learning media form. Learning media is an educational tool that can be used as an intermediary in the learning process to enhance effectiveness and efficiency in achieving teaching goals (Nasrum, 2012).

The use of learning media provides many benefits. According to Kemp and Dayton (in Nasrum, 2012) there are several benefits of media in learning, including the following:

1. Submission of material can be uniformed.
2. The learning process becomes clearer and more interesting.
3. The learning process is more interactive.
4. Efficiency in time and energy.
   Teachers often spend a lot of time explaining a subject matter.
5. Improve the quality of student learning outcomes.
6. The media allows the learning process to be carry out anywhere and anytime.
7. Media can grow each student towards the material and learning process.
8. The role of the teacher adding to be more positive and productive.
   By utilizing the media well, the teacher is no longer the only source of learning for students, he can shares the role with media so that it will be easy for him to pay attention to other aspects of education such as helping students' learning difficulties, forming and motivating student learning.

Software that will be used as learning media in this study includes Geogebra, SPSS, Minitab, Easy fit, and Microsoft Excel.

C. Methodology

1. Research Design

The research method used is quantitative research methods because the research data is in numbers form and the analysis uses statistics. The type of research is pre-experimental design. The research design used was the pre-test post-test of one group design as follows:

| O_1 | X | O_2 |
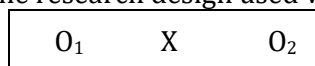
O_1 = the value of pre-test
O_2 = The value of post-test

(Sugiyono, 2014)

This research will be conducted for one year starting from June 2017 to June 2018 located at the Sembilanbelas November Kolaka University. The population in this study were all students of the Sembilanbelas November University of Kolaka mathematics education in sixth semester consist of two classes. This population collection is based on the consideration that they have obtained several statistical courses while studying at the Nineteen November University of Kolaka. Based on the experience of researchers during being a mentor in preparing the final project, most high-level students who are undergoing the thesis-making process do not understand the normality test.

To illustrate the ability of students' understanding of the concept of data normality, there are several indicators that are used as a measure tools, namely:

1. Students understand the usefulness of the data normality test
2. Students are able to write a hypothesis formula for testing data normality.

3. Students are able to use one of these two methods (Kolmogorov Smirnov, or QQ-Plot) to data normality test.
4. Students skill in using statistical software such as SPSS, Minitab, Microsoft Excel and others in performing data normality tests.
5. Students are able to explain the interpretation of the output of the software.
6. Students are able to draw conclusions from the software output.
7. Students can explain the differentiation between the two methods above.

Because the study population is quite large, researchers will only work on samples. Sampling is done by using Simple Random Sampling technique by taking a fairly representative amount.

2. Technique of Data Analysis

The type of data in this study is quantitative data. Data collection is done by giving a test. The test questions are given, written or developed based on indicators of learning achievement. Data obtained from the test results will be analyzed descriptively and to draw the conclusions on the population followed by inference stages.

The statistical inference that will be conduct is hypothesis test. It will be tested whether students' understanding of the normality data after using media is better than before? The statistical hypothesis can be written as follows:

$$H_0 : \bar{x}_d \leq 0$$

$$H_1 : \bar{x}_d > 0$$

where $\bar{x}_d$ is the average of the difference between the post test and the pre test.

Because the data will be process is pair data, then corresponding test statistics are t test statistics for paired data, namely:

$$t = \frac{\bar{x}_d}{s_d / \sqrt{n}} \qquad \text{(Sudjana, 2005, p. 242)}$$

With criteria accept H_0 if $t < t_{tab}$ and in the other cases H_0 rejected.

Acceptance or rejection criteria for H_0 can also conduct with the help of software. If the P-Value result of the software output more than the significance level $(\alpha = 0,05)$ is used, then H_0 is accepted, otherwise H_0 is rejected.

D. Finding and Discussion

1. Findings

In this section the results of the research that have been obtained will be presented. The level of understanding of students for the material given in the study was measured by giving the pre test and the post test. Pre tests are conducted so that researchers obtain data on student knowledge / understanding before the experiment conduct while the post test is conduct to measure the ability of students after the experiment. From the results of the pre test and post test data1 is obtained in Appendix 1.

Differentiation in understanding / ability before and after the use of mathematical software can be seen by compare the obtain results. Statistics from the data are present in the form of numerical extracts which can be seen in table 1 below:

Table 1. Statistics for Pre-tests (X1) And Post Tests (x2)

		x1	x2
N	Valid	31	31
	Missing	0	0
Mean		10.0403	73.7903
Median		6.2500	75.0000
Mode		6.25	71.25
Std. Deviation		11.57171	14.60409
Skewness		2.047	-.102
Std. Error of Skewness		.421	.421
Kurtosis		4.915	-.436
Std. Error of Kurtosis		.821	.821
Minimum		0.00	41.25
Maximum		51.25	100.00

From table 1 it can be seen that students' understanding ability before using teaching media and after using teaching media is very different. This can be seen from the size of the data concentration, namely the average, mode, or middle value. The average obtained from the pre test is only 10 while the average post test is 73.79. In the pre test data, most students scored 6.25 while in the post test data most students scored 71.25. The median value gives the meaning that before the use of media, 50% of students only get a value below 6.25 and the remaining above 6.25. While after using the media 50% of students get a score of 71.25 and the rest above 71.25. from the three measures of data concentration, namely the mean, mode and median, the difference in the content of material that was very far before the use of the media and after the use of media was seen. The average value, median and mode in the pre test data is very small because the scale used is a scale of 100 while in the post test data, the values of data central values have met the desired standards.

In addition to measures of data centralization, differences can also be seen from the maximum and minimum values obtained by students before and after the use of media. During the pre test, there were eight students from 31 students who could not answer all the questions so that there were eight people whose value was 0. Certainly this was the lowest score. Whereas in the post test the lowest value is 41.25. The highest score obtained by students before the use of media is 51.25 while after using the media there are two students who get a perfect score of 100.

From these statements, it is clear that differences arise due to the use of mathematical software in learning statistics, especially in data normality testing. Student grades are much better after using math software than before. To generalize the results obtained, further testing is needed. The following will test the hypothesis whether the results obtained also apply to the entire population or not.

Inferential statistics

In this section the following hypothesis will be tested:

$$H_0 : \bar{x}_d \leq 0$$

$$H_1 : \bar{x}_d > 0$$

By using SPSS assistance, the following results are obtained:

Tabel 2. Paired Sampel Test

		Paired Differences					
		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	X2 - X1	63.75000	18.98464	3.40974	18.696	30	.000

From the table above can be seen the value of Sig. or P-Value in the last column. This value is less than $\alpha = 0.05$ so that H_0 is rejected. Thus it can be concluded that there are significant differences regarding students' understanding of data normality test material before and after the use of mathematical software. From the results of the descriptive analysis, it is clear that the understanding of students after the use of the software is much better than before the use of the software.

2. *Discussion*

Empirically the data supports the hypothesis proposed because the average understanding of material after use the media is greater than before (73.79 > 10.04). Besides that the hypothesis testing shows that significantly understanding the material about testing the normality of data after using mathematical software is higher than before using the software.

This learning outcomes increased shows the level of success gained by the research team. The learning outcomes increase are influenced by many factors, but these factors are thought to arise due to the use of software. By using mathematical software when learning takes place, students become more enthusiastic and more motivated because learning becomes more interesting. The theories explained with the help of software can become clearer. The things that was abstract became more real when explained with the help of software. The use of software as assistance is one of learning media. Uniquely, this media is also a practical tool that can be used by students. Students can visualize, simulate or practice directly using this software.

Practicum in statistics courses should be exist. The theories taught without practice have a greater chance of being forgotten than learning with practice. Practicums that can more often strengthen the memory of the theory being taught. It does not possibility that the practice can think out to a theory.

In this study researchers tried to balance theory and practice. Sometimes in the explanation of the theory students are less understanding but after practice, they just understand what was explained before. In the course of Statistics, there are a lot of things that cannot be explained only by theory. For example, when explaining the area under the normal curve, if only explained with pictures on the whiteboard, there are still many do not understand, especially if only preaching on lectures. The illustrations depicted on the blackboard can make students understand if the images provided interesting / good. The problem is not all lecturers have art in drawing. Here is the role of learning media. With this software, you can explain in detail how to find the area under the normal curve with various forms of curves or wide several of area. One example is the following picture.
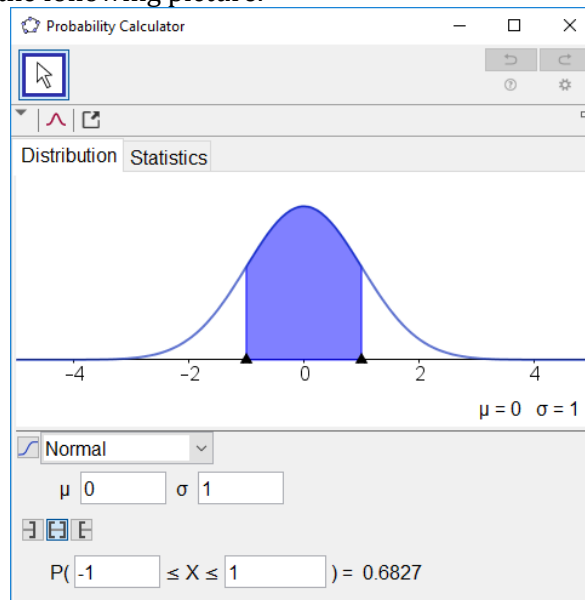

Figure 1. Example of how to calculate the area under a
normal curve using GeoGebra.

Students can practice repeatedly. With their repetitive practice, they will automatically understand the concepts in the textbook.

The learning process in the classroom when using this software is better, more interesting and more interactive. Students who are less active in learning before using the software become more active after using the software. Their activeness triggers the curiosity that is their within so that the understanding of the material being taught becomes better.

E. Conclusions

From the results of the study obtained several conclusions, namely:
1. In general, the results of this study indicate that the use of mathematical software as a media of learning as well as a practical tool has a significant influence on understanding statistical materials.
2. Empirically, the average test results after using mathematical software are much better than before using software.
3. The average understanding of the material after using the software has a significant difference compared to before using the software.

ACKNOLWLEDGEMENTS

F. REFERENCES

Kadir. (2015). *StatistikaTerapan: Konsep, Contoh dan Analisis Data dengan Program SPSS/Lisrel dalam Penelitian.* Depok: PT. Raja Grafindo Persada.

Nasrum, A. (2012). *Efektifitas Penggunaan Media Pembelajaran Berbasis Multimedia Interaktif Dalam Meningkatkan Hasil Belajar Matematika Siswa Kelas Xi Ekologi Sma Negeri 1 Kolaka.* Kolaka: Skripsi Universitas Sembilanbelas November Kolaka, Tidak dipublikasikan.

Nasrum, A. (2018). *Uji Normalitas Data untuk Penelitian.* Bali: Jayapangus Press.

Razali, N. M., & Wah, Y. B. (2011). Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors, and Anderson-Darling Tests. *Journal of Statistical Modeling and Analytics*, 21–33.

Sudjana. (2005). *Metoda Statistika.* Bandung: Tarsito.

Sugiyono. (2014). *Metode Penelitian Manajemen.* Bandung: CV. Alfabeta.

Sundayana. (2015). *Statistika Penelitian Pendidikan.* Bandung: Alfabeta.

Yap, B. W., & Sim, C. H. (2011). Comparisons of Various Types of Normality Tests. *Journal of Statistical Computation and Simulation*, 2141-2155. doi:10.1080/00949655.2010.520163.

Yazici, B., & Yolacan, S. (2007). A Comparison of Various Tests of Normality. *Journal of Statistical Computation and Simulation*, 175-183. doi:10.1080/10629360600678310.