Journal of Teaching and Learning

ANALYSIS OF INFERENTIAL STATISTICAL ERROR IN STUDENT THESIS

Neng Ranti*, S. Supriadi
Universitas Pendidikan Indonesia, Kampus Serang, West Java, Indonesia

ARTICLE INFO

Keywords:
Error analysis; inferential statistics; thesis

ABSTRACT

Learning inferential statistics is very important for students of the Elementary School Teacher Education study program, Universitas Pendidikan Indonesia, Serang Campus, to study a phenomenon and analyze research data. The purpose of this study was to determine the errors of inferential statistics in student thesis. Because the Inferential Statistics course is a subject that is considered difficult and cannot be applied in everyday life, the samples in this study were students of the Elementary School Teacher Education Study Program, Universitas Pendidikan Indonesia, Serang Campus, batch 2019. This study used quantitative research with descriptive methods. Also using the results of SPSS version 20 output. The research subjects are students of the 2019 Elementary School Teacher Education study program. The analytical technique used is normality test analysis, homogeneity, t test which is then followed by the Mann Whitney test. The output results of SPSS version 20 show that the t-test used by students in the thesis is not correct or an inferential statistic error occurs because the results do not meet the t-test requirements because the data is not normal, which should use the Mann Whitney test.

*corresponding author
E-*mail addres*: nengranti230401@upi.edu (N. Ranti)

1. Introduction

Statistics is the study of collecting, analyzing, interpreting, and presenting statistics in such a way that it becomes information to arrive at effective conclusions (Purwanto, 2019), Statistics itself is data or facts that are collected, analyzed, recorded, processed and presented in the form of diagrams or tables and so on. If you agree with the paragraph above, then the topic of this discussion is Statistics research. In a study, Statistics is used as a tool to analyze data as well as information. There are two types of statistics, including descriptive statistics and inferential statistics (Abdullah, 2015), descriptive statistics aims to provide descriptive data and information, while inferential statistics analyze and predict data.

According to the definition presented in the previous paragraph it clearly states that Statistical inference is different from Statistical descriptivism because the latter task includes collecting data, analyzing it, and drawing conclusions about the population (Bungin, 2013; Cooper & Schindler, 2014). Thus hindering the analysis and formulation of theories or hypotheses; alternatively, statistical inference is closely related to probability theory, significance level, correlation coefficient, and hypothesis.

Received 22 September 2022; Received in revised 3 October 2022; Accepted 7 October 2022; Available online 12 October 2022

Abdullah (2015), substantive statistical inference is intended to provide population characteristics based on the results of statistical case studies. Inferential statistics are often used in scientific writing, especially in academic works that can be written such as dissertations, theses, and journal articles.

The key techniques for inferential sampling in Statistics are probability or analytical techniques (Sugiyono, 2013). Inferential statistics performs probability, covariance, correlation, regression, hypothesis testing, and finally the analysis draws conclusions. Inferential statistics is also the science of analyzing data so that population size (parameters) can be described through the sample size with the aim of generalizing the analysis. What happens to the sample will be generalized and considered a description of a population by performing significance tests. Usually used in quantitative research.

What comes to your mind when you hear the word "statistics". Statistics is no longer a foreign word in the ears of students. Fixed numbers with appropriate values had to be 'translated', which required the use of special tools, or alternatively, appropriate types of diagrams explaining the relevant issue, according to a word that has since ceased to be heard. existence in everyday life. The development of the science of Statistics can be used to explain how Statistics currently has a very large toolbox that is used for things such as the scientific method, statistical inference, linear regression, psychology, electronics, biometry, and even technology. In order to interpret the results of the study data analysis accurately, it is highly expected that this knowledge of Statistics will be required (Haerudin, 2020).

Maryati (2017) argues that the purpose of statistical punishment is to help people understand information that affects their daily lives based on data, or ideas that help people understand how to collect, organize, analyze, and present data used in certain situations. Apart from being used in research conducted at the university level, Statistics can also be applied to a variety of research conducted in various disciplines.

Based on the results of research conducted by Siregar (2017), at the Faculty of Teacher Training and Education, South Tapanuli Muhammadiyah University, it was found that in general students had difficulty compiling data for statistical analysis, even though the data in question was included in the big data category. Siregar continued by stressing that the reason for the low enthusiasm of women when compiling statistical data was that women's education at the senior high school level usually did not come from the high school law. As a result, the process of training women in Statistics has deteriorated. In addition, students' low interest in reading results in a lack of knowledge in various ways, especially in terms of data processing in simple statistics.

The findings from the observations made by the observer to the seventh semester students who will start the thesis process, namely there are facts that reveal that most students in general experience many events and worries during writing a thesis. The reason for this concern is that the individuals involved have a need for data analysis. The analytical skills of masters students are the single most important aspect of the learning process, and they must be applied by masters students if studies are to give accurate results.

In line with the importance of the existence of Statistics in research, one of the missions of the Primary School Teacher Education study program is to carry out

excellent research, develop and publish innovative and competitive works in the field of basic education and early childhood education, as well as relevant multidisciplinary research, to achieve the mission of the school teacher education study program, the curriculum that applies in the elementary school teacher education study program is one of the obligations that must be taken by students in making a scientific work as a graduation requirement called a thesis.

Thesis is a form of scientific work in which the description in the thesis contains a phenomenon that occurs and is then studied by students, analyzed the data, so that in the end students get a conclusion from the results of the research. This is where the role of statistics is very important to master. However, during the trip, not many undergraduate students had difficulties in analyzing the data from the questionnaires they had completed. In this case, the campus as an education provider must be able to maximize the role so that students can be responsible for completing their final assignments, and one way is by minimizing the learning difficulties faced by students in learning inferential statistics, namely by analyzing what kind of difficulties they face and learn what factors influence student difficulties in learning inferential statistics so that in the future a kind of learning strategy can be carried out so that student difficulties in learning inferential statistics can be overcome.

The types of statistical data processing tools include SPSS, Ministep and so on, as we already know that these tools are solutions in data processing whose function is to facilitate researchers in processing thesis scientific work data, but in fact there are many errors in Statistical data processing on student thesis.

Inferential statistics data processing in current research should be done easily because there are many data processing tools that can be accessed such as SPSS, Ministep and so on. However, there are still many scientific works that are wrong in processing inferential statistics data, so it is necessary to do research on the analysis of students' thesis errors in elementary school teacher education in inferential statistics. In this study, the researcher will analyze the inferential statistical errors in the student thesis of the Elementary School Teacher Education study program, analyze the inferential statistics errors in the student thesis.

Based on the results of previous studies, various studies have been directed at the factors causing student errors in inferential statistics courses (Haerudin, 2020; Solihati and Hidayanti, 2021; Yuniarti, 2022). In previous research, there has been no exploration that draws explorations that analyze related to inferential statistical errors in student thesis. Based on previous research, the purpose of this study was to find out the inferential statistical errors in the thesis of students of the Indonesian Education study program, Serang Campus.

It can be concluded that inferential statistics is an activity of processing data obtained according to the sample rules, then analyzed accurately and systematically with the aim of providing conclusions from sample data to generalize population conditions.

2. Methodology

This study uses a quantitative research approach with methods using descriptive analysis. So the researcher uses factor analysis by looking for the output of thesis data

using SPSS software version 20. In descriptive analysis to draw conclusions about inferential statistical errors in student thesis.

Previous research that has carried out research related to inferential statistics (Solihati and Hidayanti, 2021), entitled Analysis of Information Systems Student Learning Independence in Inferential Statistics Course, which aims to determine how much confidence students have in studying inferential statistics, and also how much motivation the student has. In the study (Haerudin, 2020), entitled analysis of student difficulties in the inferential Statistics course, which aims to analyze what kind of difficulties they face and learn what factors influence students' difficulties in learning inferential statistics so that in the future such things can be done. learning strategies so that students' difficulties in learning inferential statistics can be overcome. In the study (Yuniarti, 2022), entitled Student Errors in the Public Administration Study Program in Solving Descriptive Statistics and Inferential Statistics Problems, which aims to analyze student errors in solving Statistics problems.

In this study, there is an analysis that is carried out in several stages of analysis on the output results of SPSS version 20 regarding the normality test, homogeneity then the t test when it does not meet the prerequisites should use the Mann Whitney test.

The population used by the researcher is the 2019 Elementary School Teacher Education student. While the sample from this study is one student of the Elementary School Teacher Education Study Program. The focus or problem variable in this study is to analyze the inferential statistical errors experienced by students in the thesis.

In determining the hypothesis Ha, there is a difference in the average of the experimental class (Sundanese ethnomathematical learning) and the control (Not Sundanese ethnomathematical learning). And Ho there is no difference in the average of the experimental class (Sundanese ethnomathematical learning) and the control (Not Sundanese ethnomathematics learning).

3. Results and Discussion

The results of the SPSS version 20 analysis can be analyzed by describing the contents of the output results. Is there an error in inferential statistics in the student thesis in research (Yuningsih, 2019), with the title "The Effect of Sundanese Ethnomathematical Learning on the Ability to Think Analytical Mathematics for Class III Elementary School Students", the study program for Elementary School Teacher Education, University of Education, Serang Campus. This analysis was carried out in the study including normality test, homogeneity t test with 20 students in experimental class and 20 students in control class.

A. Normality Test

Table 1. Results of the Posttest Data Normality Test on Thesis

Tests of Normality

	Kolmogrov-Smirnov[a]		
	Statistic	**Df**	**Sig.**
Experiment_Class	,183	20	,076
Control_Class	,095	20	,200*

From the data table above, it can be explained that in the student thesis there are normality test results that Sig> 0.05 = Ho is
accepted, Sig <0.05 = Ho is rejected. It was found in
the thesis data that posttest were 0.076 > 0.05 (normally distributed data) and the posttest class control 0.200 < 0.05 (normally distributed data).

Table 2. Results of the Posttest Data Normality Test
Tests of Normality

		The Kolmogrov-Smimov[a]			Shapiro-Wilk		
	Group	**Statistics**	**df**	**Sig.**	**Statistics**	**df**	**Sig.**
Value	Experiment Class	,190	20	0,056	,897	20	,036
	Control Class	,100	20	,200*	,952	20	,406

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

After checking again using a tool in the form of SPSS version 20 which is the same as the thesis data, it was found that *posttest*were 0.036 < 0.05 (data not normally distributed) and the *posttest* class *control* 0.406 > 0.05 (normally distributed data). Then the next researcher conducted a homogeneity test.

B. Homogeneity

Table 3. Test of Homogeneity Posttest on Thesis
Tests of Homogenety of Variance

		Levene Statistics	**df1**	**df2**	**Sig.**
Value	Based on mean	3.399	1	38	.073
	Based on median	3.167	1	38	.083
	Based on median and with adjusted df	3.167	1	35.587	.084
	Based on trimmed mean	3.318	1	38	.076

In the image data above, homogeneity test was carried out. It can be seen that Sig > 0.05 = Ho is accepted Sig < 0.05 = Ho is rejected.

Table 4. Data Homogeneity Test Posttest
Tests of Homogenety of Variance

		Levene Statistics	df1	df2	Sig.
Value	Based on mean	3.274	1	38	.078
	Based on median	3.013	1	38	.091
	Based on median and with adjusted df	3.013	1	35,932	.091
	Based on trimmed mean	3.176	1	38	.083

Based on the mean result 0.078 > 0.05 = Ho is accepted (homogeneous data) This thesis should not use the independent Sample t-test test because the requirements for the independent Sample t-test are not met, i.e. the data is not normally distributed. The requirement to use the independent sample t-test is that the data of two groups whose members are different and the data of the two samples must be normally distributed and homogeneous.

C. Independent Sample t-test with spss

Table 5. Results of t-test Data Posttest
Independent Samples Test

		t-test for Equality of Means						
		T	Df	Sig (2-tailed	Mean Difference	Std. Error Difference	95% confidence interval of the difference	
							Lower	Upper
Value	**Equal variances assumed**	5,391	38	,000	36,194	6,714	22,602	49,787
	Equal variances not assumed	5,391	32.881	,000	36,194	6,714	22.532	49,857

sig value. (2-tailed) 0.000 < 0.05 then Ho is rejected.

Table 6. Results of t-test Data Posttest
Independent Samples Test

			t-test for equality of means						
	F		Sig.	t	df	Sig (2-tailed	Mean Difference	Std. Error Difference	95% confidence interval of the difference
									Upper
Value	**Equal variances assumed**	.078	5.412	3.274	38	000	36.25000	6.69803	22.69056	49.80944
	Equal variaces not assumed			5.412	33,018	.000	36.25000	6.69803	22.62305	49.87695

From the results of the t test with SPSS version 20, it can be concluded that Ho is rejected, which means there is a difference in the average value of the class with Sundanese ethnomathematical learning (experimental class) and non-Sundanese ethnomathematical learning (control class). Because the data is not normally distributed, then there is a further test that the researcher must fulfill by conducting the Mann Whitney test.

D. Mann Whitney test with SPSS

Table 7. Ranks

Class		N	Mean Rank	Sum of Rank
Posttest	Experiment Class	20	28.08	561.50
	Control Class	20	12.93	258.50
	Total	40		

Table 8. Test Statistics[a]

	Posttest
Mann-Withney U	48,500,000
Wilcoxon W	258,500
Z	-4,106
Asymp. Sig. (2-tailed)	,000
Exact Sig.[2*(1-tailed Sig.)]	.000[b]

a. Grouping variables: Class
b. Not corrected for ties.

The results obtained from the Mann Withney test that the value of Asymp.sig (2-tailed) 0.000 <0.05 then Ho is rejected and Ha is accepted, meaning that the development of the Sundanese ethnomathematical learning class (experimental class) is better than the non-Sundanese ethnomathematical learning class (control class).

4. Conclusion

It was concluded that the inferential statistical error in the student thesis of the Elementary School Teacher Education study program at the Indonesian Pendiidkan University, Serang Campus, was an error in the author's thesis using the independent t-test while the data did not meet the prerequisites. Then the error in the normality test found significant differences in the posttest of the experimental class 0.076 > 0.05 (normally distributed data) and the posttest class control 0.200 < 0.05 (normally distributed data). After checking again the results posttest0.036 <0.05 (data not normally distributed) and the posttest class control 0.406 > 0.05 (normally distributed data). This means that there is a comparison between the thesis data of the experimental class and the control is normally distributed but different from the data from the retest by the researcher where the experimental class and the control are not normally distributed.

Then the test is continued with homogeneity when viewed based on the mean result 0.078 > 0.05 = Ho is accepted meaning (homogeneous data) This thesis should not use the independent Sample t-test test because the independent Sample t-test requirements are not met the data is not normally distributed. In the thesis Ho is rejected, which means that there is a difference in the average value of the class with Sundanese ethnomathematical learning (experimental class) and non-Sundanese ethnomathematical learning (control class). When the researcher conducted a truth test using SPSS with the same version as the writer of the SPSS version 20 thesis, because the data were not normally distributed, it was continued with the Mann Whitney test. The results of the Mann Whitney test found the value of Asymp.sig (2-tailed) 0.000 <0.05 then Ho was rejected and Ha was accepted, meaning that the development of the Sundanese ethnomathematical learning class (experimental class) was better than the non-Sundanese ethnomathematical learning class (control class).

The problem that tend to occur frequently and are often made by students are inaccuracies in determining procedural, especially the steps in determining the test used, so that re-checking is needed and also directed guidance to produce a better thesis. And make sure if you want to use the independent t-test, make sure the two groups have different members and the data for the two samples are normally distributed and homogeneous.

References

Abdullah, M. (2015). *Metode Penelitian Kuantitatif* (Cetakan 1). Aswaja Pressindo. Google Scholar

Bungin, 2013; Cooper & Schindler, 2014. (2013). *Metodologi Penelitian Kuantitatif*. Kencana Prenada Media Group. Google Scholar

Haerudin, I. R. D. N. (2020). Analisis kesulitan mahasiswa pada mata kuliah statistika inferensial. *Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, *1*(3), 208–215. Google Scholar

Maryati, I. (2017). *Peningkatan Kemampuan Penalaran Statistik Siswa Sekolah Menengah Pertama melalui Pembelajaran Kontekstual* (6(1)). Mosharofa. Google Scholar

Purwanto, S. dan. (2016). *Statistika untuk Ekonomi dan Keuangan Modem* (Edisi 3 Bu). Salemba Empat. Google Scholar

Siregar, R. F. (2017). *Analisis kesulitan mahasiswa dalam statistika ditinjau dari banyaknya penggunaan jasa pengolahan data statistik di universitas muhammadiyah tapanuli selatan.* Google Scholar

Solihati, T. I., & Hidayanti, R. K. (2021). Analisis Kemandirian Belajar Mahasiswa Sistem Informasi Pada Mata Kuliah Statistika Inferensial. *Kajian Penelitian Dan Pendidikan Dan Pembelajaran*, *6*(1), 992–1001. Google Scholar

Sugiyono. (2013). *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*. Alfabeta. Google Scholar

Yuniarti, R. (2022). Kesalahan Mahasiswa Program Studi Administrasi Publik dalam Menyelesaikan Soal Statistika Deskriptif dan Statistika Inferensial. *Sains Matematika Dan Statistika*, *8*(1), 46–58. Google Scholar

Yuningsih, R. (2019). *Pengaruh Pembelajaran Etnomatematika Sunda Terhadap Kemampuan Berfikir Analitis Matematika Siswa Kelas III Sekolah Dasar*. Universitas Pendidikan Indonesia. Google Scholar