Description of Probability and Randomized Events Literacy for Statistics Student at Universitas Negeri Makassar

Description of Probability and Randomized Events Literacy for Statistics Student at Universitas Negeri Makassar

M A Tiro[1], Aswi[1], Sudarmin[1],
[1]*Statistics Department of UNM Makassar*

*Email: arif_tiro@unm.ac.id

Abstract. The concept of probability and randomized events is the main basis of statistics. Thus, literacy of probability and randomized events is a basis of statistical literacy. This study aimed to describe and assess the literacy of probability and randomized events for students of the statistics department of the first (2013) and the second (2014) batch at Universitas Negeri Makassar. Statistical literacy is defined in four indicators namely understanding the concept; the ability to apply; the ability to calculate; and interpretation skills of probability and randomized events. The approach used for the analysis of this survey data were descriptive quantitative and classification. Classification analysis was used to determine problems classification faced by students in probability and randomized events literacy. A confirmatory factor analysis was used to see the validity of the construct operational definitions of the variables. This study found that the level of probability and randomized events literacy achieved by statistics students at Universitas Negeri Makassar is categorized medium, pointing to the need for increased learning strategies that emphasize the four basic competencies.

1. Introduction

Schield's statement [1] was revealed [2] that curriculum echoes must be based on material from the discipline of statistics. In statistics, matter is data that is not just a number as studied by mathematicians, but data in context. Some changes are pervasive and important in the form and type of data available to the general public. Data from large-scale studies are increasingly available. For larger sample sizes, the small differences would be statistically significant. For example, a result from the National Longitudinal Survey of Youth (NLSY), which involved 12,000 subjects, found a 0.4-point difference in IQ test scores between men and women that were statistically significant at the 5% level [3]. A person who has statistical literacy is equipped to recognize the effect of sample size on the statistical significance. Thus, in large data sets, assigning persons to have statistical literacy should anticipate that statistical significance is almost insignificant because the relationship between the two variables is expected to be statistically significant. We can anticipate that understanding context and confounding effects is much more relevant than understanding random variation and statistical significance. Such knowledge and skills become increasingly important when population data (that is, studies are based on large samples), become more generalized in decision making, both for government and for business.

Based on this description, statistical literacy needs to be emphasized in the learning process. The main basis of statistics is the concept of probability and randomized events. Thus, probability literacy and randomized events provide a solid basis for statistical literacy. From this background, the researcher designed a study entitled: Description of Probability and Randomized Events Literacy for Statistics Students at Universitas Negeri Makassar.

1.1. Formulation of the Problems

Based on the observations of researchers, in general, learning statistics today still emphasizes technical procedures with many formulas. Learning like this is scary for learners. Therefore, it is necessary to formulate a learning strategy that is challenging and fun. The correct learning of statistics is to emphasize statistical literacy which views statistics not only in the presentation of numbers in the form of numerals, but numbers together with their context. Basic concepts that are very important in statistics are probability and randomized events. Thus, the research questions in this study are:

1. How do students understand the concept of probability and randomized events?
2. How is the student's ability to apply the concept of probability and randomized events in real-life contexts?
3. How is the student's ability to calculate the probability value of a given randomized events?
4. How is the student's ability to interpret the probability value of calculated randomized events?
5. Does the construction theory that builds the operational definition of probability literacy and randomized events literacy (conceptual understanding, application, calculation, and interpretation) have the support of empirical facts?
6. How is the relationship between probability literacy and randomized events literacy in the context of the four indicators (concept understanding, application, calculation, and interpretation)?

2. Research Methods

This type of research is analytical descriptive conducted on students of the Statistics Department at UNM regarding probability and randomized events literacy. Students are given a test whose results are analyzed and described. The subjects of this research were statistics students of the first batch (2013) and the second batch (2014), then they were given a test on probability and randomized event literacy

2.1. Research Variable

Two research variables can be defined operationally as follows:

1. Probability literacy is the ability to understand the concept of probability, apply the concept of probability in a real-life context, calculate the value of probability, and interpret the calculated probability values.
2. Randomized event literacy is the ability to understand the concept of randomized events, apply the concept of randomized events in the context of real-life, calculate the probability value of randomized events, and interpret the calculated randomized event probability value

Based on this operational definition, a data collection instrument was developed, the process of which is described as follows.

2.2. Instruments for Collecting Data

The data collection instruments are (1) the basic test of probability literacy in the form of multiple-choice questions, where the questions grid is given in Table 1. Then the basic randomized events literacy test is also in the form of multiple-choice questions, where the questions grid is given in Table 2. These two tests were developed utilizing trials which simultaneously collected data by analyzing the distractor functions and difficulty level of each item [4].

Table 1. Components of a basic probability literacy test

No.	Indicators of probability literacy	Item numbers
1.	Understanding of concepts probability	1, 2, 17, 26
2.	The ability to apply probability	3, 4, 24, 27
3.	The ability to calculate probability	5, 6, 7, 21
4.	Interpretation skills of probability	19, 23, 25, 28

Table 2. Components of a basic randomized events literacy test

No.	Indicators of randomized events literacy	Item numbers
1.	Understanding of concepts randomized events	9, 10, 18, 29
2.	The ability to apply concepts randomized events	11, 12, 20, 30
3.	The ability to calculate the probability of randomized events	8, 14, 22, 31
4.	Interpretation skills of randomized events.	13, 15, 16, 32

2.3. Data Analysis

The approach used for the analysis of this survey data were descriptive quantitative and classification. Classification analysis was used to determine the classification of problems faced by students in probability literacy and randomized events. Furthermore, a confirmatory factor analysis was carried out to see the validity of the construct operational definitions of the variables. In addition to the construct validity test, the relationship between the two variables was also analyzed by path analysis. Confirmation factor analysis was performed using the model of Figure 1. This figure shows the confirmation factor analysis for each probability literacy and randomized events literacy and involves a correlation analysis of both.

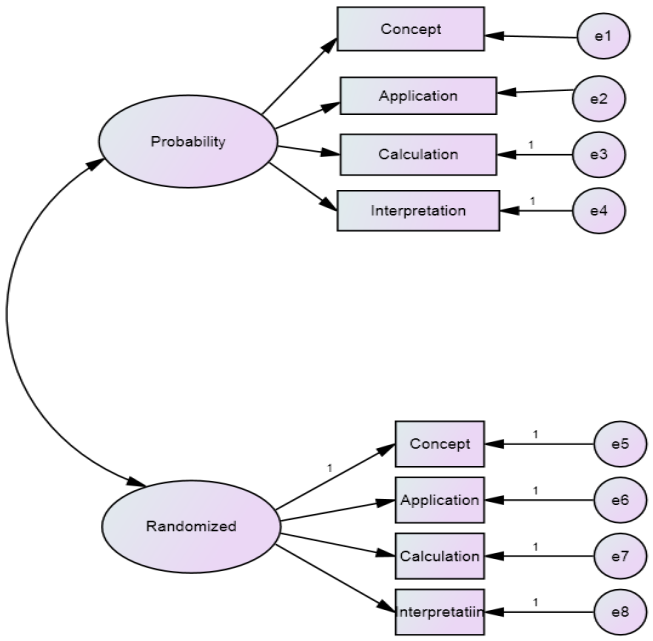

Figure 1 Diagram of probability and randomized events literacy factors

3. Results

3.1. Descriptive Analysis

In descriptive analysis, the bootstrap method is used, namely resampling to increase the size of the sample. Increasing the size of the sample (using a large sample) is carried out to avoid the demands of a normality test for the population from which the sample originates. The bootstrap method is described theoretically [5]. Furthermore, the achievement, mean, standard deviation, and 95% confidence intervals for the probability literacy are presented in Table 3, both those using the original sample data and those using bootstrap data with a sample size of 1000. From this table it can be seen that the mean and standard deviation of probability literacy for the two original and bootstrapped samples are the same, only the bootstrap method provides a 95% confidence interval for the mean. The achievement of probability literacy ranging between 60% and 87% categorized medium [4].

Table 3. Mean and standard deviation of probability literacy of sample and bootstrap data

Iindicators	Achievement (%)	sample data (*n*=38)		bootstrap data (*n*=1000)		
		Means	standard deviation	Means	standard deviation	95% Confidence Intervals
Understanding of probability concepts	60	2.39	0.917	2.39	0.917	2.13 – 2.68
The ability to apply probability	77	3.08	1.023	3.08	1.023	2.74 – 3.39
The ability to calculate probability	77	3.08	0.882	3.08	0.882	2.82- 3.34
Interpretation skills of probability	87	3.47	0.862	3.47	0.862	3.21 – 3.74

Likewise, the achievement, mean, standard deviation, and 95% confidence intervals for the mean of randomized events literacy are presented in Table 4, both using the original sample data and using the bootstrap data with a sample size of 1000. From this table, it appears that the mean and standard deviation of randomized events literacy from the two original and bootstrap samples are the same, and the bootstrap method provides a 95% confidence interval for the mean. The achievement of randomized events literacy ranging between 44% and 59% categorized medium [4].

Table 4. Mean and standard deviation of randomized events for sample bootstrapping data

Iindicators	Achievement (%)	sample data (*n*=38)		bootstrap data (*n*=1000)		
		Means	standard deviation	Means	standard deviation	95% Confidence Intervals
Understanding of randomized events concepts	44	1.74	0.724	1.74	0.724	1.53 – 1.97
The ability to apply randomized events concepts	46	1.82	0.730	1.82	0.730	1.58 – 2.05
The ability to calculate probability pfrandomized events	59	2.37	0.751	2.37	0.751	2.13- 2.58
Interpretation skills of randomized events.	58	2.32	1.233	2.32	1.233	1.89 – 2.68

3.2. Result of confirmation factor Analysis

The validity of the two literacy constructs (probability and randomized events) is given in Table 5. Then the order of the contribution of each indicator to each literacy is shown in Table 6. From these two tables, it can be seen that for probability literacy, indicators which gave the largest significant contribution was the ability to calculate the probability value with a confidence level of 97%, then the ability to apply the probability concept with a confidence level of 94%, and the lowest was the understanding concept of the probability with a confidence level of 60% (can be considered

insignificant). Then, the probability value interpretation skill is significant with a confidence level between 94% and 97%.

From these two tables (Tables 5 and 6) it can also be seen that for randomized events literacy of the students, the indicator that gives the biggest significant contribution is the ability to calculate in the context of randomized events with a 95% confidence level, then randomized events interpretation skills 93% confidence, and the lowest is the ability to apply randomized events with a confidence level of 61% (can be considered insignificant). Then, understanding the concept of randomized events was significant with a confidence level of between 61% and 96%. However, it is surprising that the correlation coefficient between probability literacy and random event literacy is very low (0.09), not significant.

Table 5. The construct validity of the probability and randomized events literacy

			Estimate	S.E.	C.R.	P
Interpretation skills of probability	<---	the probability	1,00			
The ability to calculate probability	<---	the probability	2,08	0,939	2,212	0,027
The ability to apply probability	<---	the probability	0,99	0,538	1,853	0,064
Understanding of probability concepts	<---	the probability	0,34	0,417	0,824	0,410
Understanding of randomized events concepts	<---	randomized events	1,00			
The ability to apply randomized events concepts	<---	randomized events	0,42	0,485	0,859	0,390
The ability to calculate the probability of randomized events	<---	randomized events	1,57	0,784	2,000	0,045
interpretation skills of randomized events.	<---	randomized events	1,96	1,087	1,806	0,071

Table 6. Standard estimates of the path coefficients of each probability and randomized events literacy indicator

			Estimate
Interpretation skills of probability	<---	the probability	0,46
The ability to calculate probability	<---	the probability	0,93
The ability to apply probability	<---	the probability	0,39
Understanding of probability concepts	<---	the probability	0,15
Understanding of randomized events concepts	<---	randomized events	0,43
The ability to apply randomized events concepts	<---	randomized events	0,18
the ability to calculate the probability of randomized events	<---	randomized events	0,65
interpretation skills of randomized events.	<---	randomized events	0,49

4. Conclusions

Based on the results of data analysis, it can be concluded as follows:

1. With a confidence level of 95%, students' understanding of:
   a. the concept of probability is still medium, not more than 2.68 (67%) although not less than 2.13 (53%);

   b. the concept of randomized events is also medium, not more than 1.97 (49%) although not less than 1.53 (38%);
2. With a confidence level of 95%, students' abilities in the application:
   a. the concept of probability is not more than 3.39 (85%) although not less than 2.74 (69%);
   b. the concept of randomized events is not more than 2.05 (51%) although not less than 1.58 (40%).
3. With a confidence level of 95%, the student's ability to calculate:
   a. the probability value is not more than 3.34 (85%) although not less than 2.82 (71%);
   b. the randomized events related score is not more than 2.58 (65%) although not less than 2.13 (53%).
4. With a confidence level of 95%, students' skills in interpretation of:
   a. the probability value is not more than 3.74 (94%) although not less than 3.21 (80%);
   b. the randomized events-related score is not more than 2.68 (67%) although not less than 1.85 (46%).
5. All literacy indicators (concept understanding, application capability, calculation skills, and interpretation skills) to:
   a. probability literacy is supported by empirical facts with the lowest confidence level of 60%, even the highest is up to 97%.
   b. literacy of randomized events is supported by empirical facts with the lowest confidence level of 61%, even the highest is up to 95%.
6. The correlation between probability literacy and randomized events literacy is very weak in the perspective of the four indicators (concept understanding, application capability, calculation ability, and interpretation skills).

Reference

[1] Schield, M. 2004. Statistical Literacy Curriculum Design, Curricular Development in Statistics Education. Sweden.

[2] Tiro, M. A. & Annas, S,. 2015. *Filsafat Statistika*. Makassar: Andira Publisher.

[3] *Bureau of Labor Statistics*, 1979; www.bls.gov/nls/

[4] Stanley, J. C. & Hopkins, K. D. 1972. *Educational and Psychological Measurement and Evaluation*. New Delhi: Prentice-Hall of India Private Limited.

[5] Tiro, M. A. 1991. Edgeworth Expansion and Bootstrap Approximation for M-Estimators of Linear Regression Parameters with Increasing Dimensions. *Dissertation. Unpublished*. Ames, Iowa: Iowa State University.