Application of Statistical Methods in Determining Learning Evaluation Values and Scores

Tri Haryati[1], Isti Koma[2], Alfira Mersania Shintia[3], Julmi[4]

[1] Sekolah Tinggi Agama Islam Ibnurusyd Kotabumi Lampung; trihariyati94@gmail.com
[2] Sekolah Tinggi Agama Islam Ibnurusyd Kotabumi Lampung; istiiqomah188@gmail.com
[3] Sekolah Tinggi Agama Islam Ibnurusyd Kotabumi Lampung; alfirarasya03@gmail.com
[4] Sekolah Tinggi Agama Islam Ibnurusyd Kotabumi Lampung; julmikrisna30@gmail.com

ARTICLE INFO

ABSTRACT

Learning assessment is an important aspect of education that requires accuracy, objectivity, and fairness, as assessments that rely solely on teacher subjectivity can lead to misinterpretation of students' abilities. This study aims to analyze the application of statistical methods in determining scores and evaluating results in learning. The method used is a literature review. The analysis was conducted based on literature related to statistical methods, including mean calculation, standard deviation, z-score, Pearson Product Moment correlation, Cronbach's Alpha, and multiple linear regression. The results indicate that applying statistical methods can improve the accuracy, validity, and reliability of evaluations, help teachers understand the distribution of student scores, determine achievement categories regularly, and reduce subjectivity in assessment. Therefore, the use of these methods allows for more objective and measurable assessments, serving as a basis for informed educational decision-making. Teachers and educational institutions are encouraged to regularly apply statistical methods to enhance learning quality and provide more meaningful feedback to students

Corresponding Author:
Tri Haryati
Sekolah Tinggi Agama Islam Ibnurusyd Kotabumi Lampung; trihariyati94@gmail.com

1. INTRODUCTION

Learning evaluation is an important part of the educational process which functions to assess the achievement of student competencies (Idris & Duskri, 2024). In practice, the determination of grades and scores for learning outcomes is often done subjectively or without considering data variation, grade distribution, and the reliability of assessment instruments. This can lead to inaccuracies in academic

decision-making, such as determining graduation, remedial classes, and awarding academic awards. Therefore, an objective and measurable approach is needed to support fairness and accuracy in learning evaluation (Hasibuan et al., 2025).

The main problem that often arises in traditional assessment systems is the suboptimal use of statistical methods to analyze evaluation results (Siarkanasa et al., 2025). Most educators still use simple arithmetic averages without considering data distribution, measurement error, or instrument reliability. As a result, evaluation results do not fully reflect students' abilities. The use of appropriate statistical methods, such as regression analysis, validity and reliability testing, or item analysis, can improve the quality of assessment decisions (Nurhusain & Hadi, 2021).

Statistical methods in learning evaluation can be used for various purposes, such as determining the score weights for each assessment aspect, identifying valid test items, and calculating final scores that are representative of student abilities. This approach allows educators to understand score distribution patterns, test difficulty levels, and the relationships between learning variables (Prijowuntoro, 2020). Thus, the assessment becomes more objective, transparent, and scientifically accountable.

Although statistical methods have long been known, their application in determining grades and scores for learning evaluations in various educational institutions remains limited. Many teachers and lecturers lack adequate statistical analysis skills, so the evaluation process is still conducted manually or intuitively (Maulani et al., 2024). Furthermore, not many evaluation systems automatically integrate statistical methods into their calculations. This creates a gap between theory and practice that needs to be bridged through applied research (Damayanti et al., 2023).

Several studies have been conducted regarding the application of statistical methods in learning evaluation.Rangkuti & Albina (2025) examined the application of descriptive statistical analysis to determine the distribution of high school students' grades and found that the use of such analysis helped teachers understand the distribution of learning outcomes more objectively Fernanda & Hidayah (2021)using the Item Response Theory (IRT) method to measure the validity of exam questions, which has been proven to increase the reliability of evaluation instruments by up to 15%. Meanwhile,Marsela & Sitepu (2025) applied multiple linear regression analysis to link scores on assignments, midterm exams, and final exams with students' final grades. Their research showed that this method provided a more accurate grade prediction model than conventional calculations, thereby improving the quality and objectivity of the learning evaluation process.

Based on the description above, the problem formulation in this study includes how the application of statistical methods can increase objectivity and accuracy in determining the value and score of learning evaluation, what statistical methods are most effective to use in the process of evaluating learning outcomes, and how the application model of statistical methods can be integrated into the assessment system used by educators. The purpose of this study is to analyze the application of statistical methods in determining the value and score of learning evaluation, identify the most appropriate statistical method to increase the reliability of evaluation results, and develop a statistical-based assessment model that can be used practically by educators in the evaluation process.

2. METHOD

This study uses a qualitative approach with library research. This approach was chosen because the study focuses on the study of theories, concepts, and previous research findings related to the application of statistical methods in determining learning evaluation scores and values (Sugiyono, 2020).This library research does not involve direct field data collection, but rather is conducted by reviewing various relevant literature sources, such as books, scientific journals, undergraduate theses, dissertations, and other academic documents that discuss statistical methods and their application in evaluating learning outcomes. The goal of this approach is to gain a comprehensive understanding of the concepts, methods, and strategies for applying statistical analysis in the context of learning evaluation, both from a theoretical and practical perspective (Moleong, 2017).

The research procedure was carried out through several stages, namely: (1) library data collection, by searching and selecting relevant literature according to the research focus; (2) data management and classification, namely selecting library sources based on the suitability of the theme, the level of novelty, and the credibility of the author; (3) content analysis, namely reviewing the contents of the literature to identify the main concepts, similarities, and differences between theories; and (4) drawing conclusions, namely compiling the results of synthesis and interpretation from various sources to answer the formulation of the research problem. Data analysis in this study was carried out descriptively qualitatively by describing the results of the study in a narrative manner, so that it can provide an in-depth picture of how statistical methods can be applied in determining learning evaluation values and scores objectively and systematically.

3. RESULTS AND DISCUSSION

Results

Based on the results of a literature review from various sources, it was found that the application of statistical methods in determining learning evaluation grades and scores significantly assists educators in obtaining more objective, measurable, and fair assessment results. The use of statistical methods allows teachers to analyze the distribution of grades, calculate averages, determine instrument reliability, and predict relationships between assessment components such as assignments, mid-term exams, and final exams. In various previous studies, (Rangkuti & Albina, 2025; Fernanda & Hidayah, 2021), it was found that statistical methods contributed greatly to increasing the accuracy of evaluation results and reducing the subjectivity of teacher assessments.

As an example of application, assume there are grade data for 10 students who participated in a learning evaluation with three assessment components, namely assignments (X_1), mid-term exams (X_2), and final exams (X_3). The final grade (Y) is determined using a statistical approach based on weighting: assignments 30%, mid-term exams 30%, and final exams 40%.

Table 1. Results of Assignment, Mid-Term and Final Exam Calculations

No	Student Name	Task (X_1)	Mid-Term Exam (X_2)	Final Exam (X3)	Final score
1	Aisha	80	78	82	(0.3×80) + (0.3×78) + (0.4×82) = 80.0
2	Budi	75	70	76	(0.3×75) + (0.3×70) + (0.4×76) = 73.9
3	Image	85	88	90	(0.3×85) + (0.3×88) + (0.4×90) = 87.1
4	Dika	78	80	79	(0.3×78) + (0.3×80) + (0.4×79) = 79.1
5	Eka	90	85	88	(0.3×90) + (0.3×85) + (0.4×88) = 87.1
6	Dawn	70	65	72	(0.3×70) + (0.3×65) + (0.4×72) = 69.5
7	Gita	88	90	94	(0.3×90) + (0.3×85) +

					(0.4×88) = 91.2
8	Hani	82	79	80	(0.3×82) + (0.3×79) + (0.4×80) = 80.0
9	Indra	76	75	78	(0.3×76) + (0.3×75) + (0.4×78) = 76.9
10	Eka	92	90	95	(0.3×92) + (0.3×90) + (0.4×95) = 92.9

To find out the distribution of final values, the average (mean) formula is used.(Sugiyono, 2020):

Calculation 1. Mean (Average)

$$X = \frac{\sum X_i}{n}$$

$$X = \frac{80.0 + 73.9 + 87.1 + 79.1 + 87.1 + 69.5 + 91.2 + 80.0 + 76.9 + 92.9}{10} = \frac{817.7}{10}$$

$$= 81.77$$

So, the average class score is 81.77. Next, to determine the distribution of scores, the standard deviation (S) is used.(Sugiyono, 2020):

Calculation 2. Standard Deviation

$$S = \sqrt{\frac{\sum(X_i - X)^2}{n - 1}}$$

After calculating, it was obtained$S$= 7.23. This means that most students' scores are in the range of 74.5 to 89.0 (±1 SD from the mean). This indicates that the distribution of scores is fairly even and not too deviant.

Based on literature review (Sugiyono, 2020; Moleong, 2017), the validity of the test items can be tested using the Pearson Product Moment correlation formula:

Calculation 3. Pearson Product Moment Correlation

$$r_{xy} = \frac{n\left(\sum XY\right) - \left(\sum X\right)\left(\sum Y\right)}{\sqrt{\left[n \sum X^2 - \left(\sum X^2\right)\right]\left[n \sum Y^2 - \left(\sum Y\right)^2\right]}}$$

From the results of the theoretical test, if rxy > rtable (for example, rtable = 0.632 for n = 10), then the question is declared valid. In general, in previous research, the rxy

value ranges from 0.65–0.85, indicating good validity (Moleong, 2017). For reliability, the Cronbach's Alpha (α) formula is used:

Calculation 4. Cronbach's Alpha (α) Formula

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_t^2}\right)$$

The results of the study show that evaluation instruments with an α value ≥ 0.70 are considered reliable. For example, in the study Rangkuti & Albina (2025), obtained $\alpha = 0.81$ which indicates a high level of instrument consistency.

In several studies Fernanda & Hidayah (2021), a multiple linear regression model was used to predict the final grade based on three assessment components, with the equation(Nugraha, 2022):

Calculation 6. Results of Multiple Linear Regression Model

$$Y = \alpha + b_1 X_1 + b_2 X_2 + b_3 X_3$$

$$Y = 5.21 + 0.32 X_1 + 0.28 X_2 + 0.40 X_3$$

For example, the analysis results show that every 1 point increase in the UAS score (X_3) will increase the final score by 0.40 points, indicating that the UAS has the greatest influence on students' final scores compared to other components (Rangkuti & Albina, 2025; Fernanda & Hidayah, 2021).

In addition to calculating numerical values using statistical methods, evaluation results can also be converted into categories of student learning quality to make them easier to understand and use as a basis for providing feedback (Mulyatiningsih, 2025). Based on the literature and applicable assessment practices, the final score can be divided into a range of 0–10 with the categories in the following table.

Table 2. Final Score Assessment

Final Score	Assessment Categories	Interpretation
10-9	Very Good (A)	Very satisfactory achievement, showing maximum mastery of the material
8-7	Good (B)	Good achievement, adequate mastery of the material
6-5	Enough (C)	Achievement is sufficient, mastery of the material still needs improvement
4-3	Less (D)	Low achievement, inadequate mastery of

		material
2-0	Very Poor (E)	Very low performance, requires significant assistance and improvement

For example, the average value of the class evaluation results based on the statistical method is 7.8 with a standard deviation of 1.1. Then:

1. Students with grades 9–10 are included in the Very Good (A) category.
2. A score of 8–7 is included in the Good (B) category.
3. A score of 6–5 is in the Fair category (C), and so on.

Table 3. Example of Category Assessment

No	Student Name	Final score	Category
1	Aisha	80	Good (B)
2	Budi	73.9	Good (B)
3	Image	87.1	Good (B)
4	Dika	79.1	Good (B)
5	Eka	87.1	Good (B)

The table shows that the majority of students fall into the Good (B) category, indicating generally satisfactory learning outcomes. This category makes it easier for teachers to provide feedback, determine improvement strategies, and convey evaluation results more meaningfully to students (Rangkuti & Albina, 2025; Fernanda & Hidayah, 2021).

Discussion

In general, the application of statistical methods in learning evaluation is crucial for improving the objectivity, accuracy, and fairness of assessments. Learning evaluations that rely solely on subjective teacher assessments can lead to unfairness and misinterpretation of student abilities (Mardapi et al., 2019)By using statistical methods, teachers can systematically process grade data, analyze grade distributions, and determine student achievement categories more measurably. This enables data-driven learning decisions and supports overall improvement in learning outcomes (Soulissa et al., 2022).

Based on a literature review of various studies, the application of statistical methods in determining grades and scores for learning evaluations has been proven to significantly assist teachers in obtaining more objective and fair assessment results. This method allows teachers to analyze the distribution of grades, calculate averages, assess instrument reliability, and predict relationships between assessment components such as assignments, mid-term exams, and final exams. Previous

researchRangkuti & Albina (2025) AndFernanda & Hidayah (2021) Studies have shown that statistical methods significantly contribute to increasing the accuracy of evaluation results and reducing the subjectivity of teacher assessments. Thus, statistics is not only a calculation tool but also a basis for more scientific decision-making in the learning process.

The application of statistical methods can be seen in the calculation of students' final grades based on the weighting of the assessment components. For example, the final grade is determined by the weighting of 30% of assignments, 30% of mid-term exams, and 40% of final exams, so each component contributes proportionally to the final grade. From the data of the ten students analyzed, the average class grade was 81.77. This indicates that student achievement is generally at a fair to good level, providing an initial overview of the class's overall performance.

To determine the distribution of scores, the standard deviation was used. Based on calculations, the standard deviation of students' final scores was 7.23, indicating that the majority of students' scores fell between 74.5 and 89.0. This distribution indicates a fairly even distribution of scores, with no extreme differences between students. This allows teachers to assess more fairly and adjust learning strategies for groups of students who fall below average.

In addition, the validity and reliability of the evaluation instrument were analyzed using statistical methods. The validity of the items can be tested using Pearson Product Moment correlation, where an r value greater than r indicates a valid item. Meanwhile, the reliability of the instrument can be tested using Cronbach's Alpha, with $\alpha \geq 0.70$ indicating a high level of instrument consistency (Ahmad, 2020)This analysis ensures that the assessments conducted truly reflect students' abilities and can be used as a basis for making learning decisions (Habiby, 2017).

Statistical methods also allow the use of predictive models such as multiple linear regression to estimate students' final grades based on assessment components.(Fernanda & Hidayah, 2020)Regression analysis shows that every 1-point increase in the final exam results in a 0.40-point increase in the final grade, indicating that the final exam has the greatest impact compared to assignments or the midterm exam. In addition to numerical calculations, evaluation results can also be converted into categories of student learning quality, making it easier for teachers to provide more appropriate feedback and improvement strategies (Malik & Chusni, 2018). From the analyzed data, most students are in the Good (B) category, indicating that achievement is generally satisfactory and the learning process is running effectively.

4. CONCLUSION

Based on the results of the study and analysis, it can be concluded that the application of statistical methods in determining learning evaluation grades and scores is very effective in producing objective, accurate, and fair assessments. The use of statistical methods allows teachers to analyze the distribution of grades, calculate averages, determine the reliability and validity of instruments, and predict the relationships between assessment components such as assignments, mid-term exams, and final exams. Thus, evaluation results not only accurately reflect student abilities but also serve as a basis for more scientific and measurable learning decisions.

Based on these conclusions, it is recommended that teachers and educational institutions more routinely apply statistical methods in learning evaluations, particularly in determining final scores and student achievement categories. Furthermore, teachers need to improve their understanding of simple statistical analyses, such as averages, standard deviations, z-scores, correlations, and regressions, so that evaluation results can be optimally used for developing learning strategies and providing feedback. The application of these methods should also be supported by valid and reliable evaluation tools to ensure the quality of assessments.

REFERENCES

Ahmad, NQ (2020). Introduction to Learning Evaluation (A Concept and Practice). Takengon: Shakura.

Damayanti, AM, SH, MP, Daryono, MP, & Rayanto, YH (2023). Learning Evaluation. Pasuruan: Basya Media Utama.

Fernanda, JW, & Hidayah, N. (2020). Analysis of the Quality of Statistical Exam Questions Using Classical Test Theory and the Rasch Model. Square: Journal of Mathematics and Mathematics Education, 2(1), 49-60. https://journal.walisongo.ac.id/index.php/square/article/view/5363

Fernanda, JW, & Hidayah, N. (2021). Implementation of Statistics as an Analytical Tool in Decision-Making in Education. Journal of Community Service Innovation (JIPEMAS), 4(3), 341-350. https://riset.unisma.ac.id/index.php/jipemas/article/view/1020

Habiby, WN (2017). Educational Statistics. Surakarta: Muhammadiyah University Press.

Hasibuan, NH, Safitri, S., & Ariska, I. (2024). Evaluation Score Processing

Techniques. MUDABBIR Journal Research and Education Studies, 4(2), 460-475. https://jurnal.permapendis-sumut.org/index.php/mudabbir/article/view/646

Idris, J., & Duskri, M. (2024). Optimizing the Use of Statistics in the Evaluation of Islamic Religious Education Learning: A Quantitative Approach to Improving Assessment Quality. Malewa: Journal of Multidisciplinary Educational Research, 2(2), 88-97. https://jurnal.mymalewa.org/index.php/JoMe/article/view/119

Malik, A., & Chusni, MM (2018). Introduction to Educational Statistics: Theory and Applications. Yogyakarta: Deepubish

Mardapi, D., Hadi, S., & Retnawati, H. (2019). Determining Student-Based Minimum Completion Criteria. Journal of Educational Research and Evaluation, 19(1), 38-45. https://scholarhub.uny.ac.id/jpep/vol19/iss1/4/

Marsela, E., & Sitepu, EB (2025). The Role of Statistics in Curriculum Evaluation to Improve the Quality of Education: A Library Research Analysis. Jurnal Intelek Insan Cendikia, 2(5), 8976-8981. https://jicnusantara.com/index.php/jiic/article/view/3393

Maulani, G., Septiani, S., Susilowaty, N., Rusmayani, NGAL, Evenddy, SS, Nababan, HS, ... & Nurlely, L. (2024). Learning Evaluation. Banten: Sada Kurnia Pustaka.

Moleong, LJ (2017). Qualitative Research Methodology (Revised Edition). Bandung: Remaja Rosdakarya.

Mulyatiningsih, E. (2015). Applied Research Methods in Education. Yogyakarta: Uny Press.

Nugraha, B. (2022). Statistical Test Development: Implementation of Multiple Linear Regression Methods with Classical Assumption Test Considerations. Bandung: Pradina Pustaka.

Nurhusain, M., & Hadi, A. (2021). Good Quality Case-Based Applied Statistics Learning Design (Valid, Practical, and Effective) for Mathematics Education Students. Indonesian Journal of Educational Science (IJES), 3(2), 105-119. https://ojs.unsulbar.ac.id/index.php/ijes/article/view/951

Prijowuntato, SW (2020). Learning Evaluation. Depok: Sanata Dharma University Press.

Rangkuti, MH, & Albina, M. (2025). Assessment and Evaluation of Learning in Improving Learning Effectiveness in Islamic Educational Institutions. QOSIM: Journal of Social Education & Humanities, 3(1), 358-366.

https://ejournal.yayasanpendidikandzurriyatulquran.id/index.php/qosim/article/view/829

Siarkanasa, B., Maskhuliah, P., Lau, FM, Irman, MR, & Lulang, H. (2025). The Role of Central Statistics Measures and Data Location in the Evaluation of Educational Statistics in Schools. Journal of Management and Education Science | E-ISSN: 3062-7788, 2(2), 709-717. https://jurnal.kopusindo.com/index.php/jimp/article/view/1113

Soulisa, I., Supratman, M., Rosfiani, O., Renaldi, R., Utomo, WT, Hermawan, CM, ... & Sutisnawati, A. (2022). Learning Evaluation. West Java: Widina Publisher.

Sugiyono. (2020). Quantitative, qualitative, and R&D research methods. Bandung: Alfabeta.