https://journal.uinsgd.ac.id/index.php/kubik/index
Vol. 10, No. 1, 2025, pp. 48-55
e-ISSN: 2686-0341 p-ISSN: 2338-0896

Parameter Estimation and Analysis of Average Years of Schooling in Merauke District with Birnbaum-Saunders Distribution Approach

Agustinus Langowuyo[1, *)], Sara Yokhu[2)], Felix Reba[3)]

[1,2,3] *Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Cenderawasih, Papua, Indonesia*

[*)]email: : agustinus.langowuyo@fmipa.uncen.ac.id

Abstract

Average years of schooling is an important indicator in assessing the success of education development in a region. This study aims to analyze data on average years of schooling in Merauke Regency, Papua Province, using the Birnbaum-Saunders (BS) Distribution approach. This distribution was chosen because of its ability to model data that has asymmetric characteristics and low variability. The parameters resulting from the analysis include a scale parameter (β) of 8.35, which reflects the average years of schooling of the population, and a shape parameter (α) of 0.0545, which indicates the low degree of dispersion of the data around the mean. The results of the analysis show that the average length of schooling in Kabupaten Merauke is at the junior high school (SMP) level, with a homogeneous data distribution. This homogeneity reflects good equity in access to education, but also indicates the potential for stagnation at certain levels of education. The Birnbaum-Saunders distribution proved to be effective in modeling education data in this region, providing a more accurate picture than traditional approaches. This research makes an important contribution in understanding the distribution pattern of average years of schooling in Merauke district. The results can be used as a basis for designing more targeted policies in improving the quality and access to education, especially at the senior secondary level. In addition, this approach can serve as a reference for analyzing education in other regions with similar geographical and socio-economic challenges.

Keywords: Average years of schooling, Merauke Regency, Maximum Likelihood Estimation, Birnbaum-Saunders Distribution
MSC2020: 62F10, 62P20, 62E15

Abstrak

Rata-rata lama sekolah merupakan indikator penting dalam menilai keberhasilan pembangunan pendidikan di suatu daerah. Penelitian ini menganalisis data rata-rata lama sekolah di Kabupaten Merauke, Provinsi Papua, menggunakan pendekatan Distribusi Birnbaum–Saunders (BS). Distribusi ini dipilih karena kemampuannya memodelkan data yang bercirikan asimetris dan memiliki dispersi rendah. Estimasi parameter dilakukan dengan metode Maximum Likelihood Estimation (MLE) untuk mendapatkan model yang sesuai dengan karakteristik data. Hasil analisis menunjukkan bahwa model BS berhasil menggambarkan distribusi data secara akurat, menghasilkan estimasi parameter yang bermakna secara substantif dan selaras dengan kondisi pendidikan di lapangan. Temuan ini memberikan gambaran yang lebih realistis dibandingkan pendekatan konvensional serta menawarkan informasi yang relevan untuk perumusan kebijakan, khususnya dalam meningkatkan akses dan

*) Corresponding Author
Received: 25-12-2024, Accepted: 09-08-2025, Published: 11-08-2025

kualitas pendidikan di jenjang menengah atas. Pendekatan ini juga berpotensi diterapkan pada analisis pendidikan di wilayah lain dengan tantangan geografis dan sosial-ekonomi yang serupa.

Kata kunci: rata-rata lama sekolah, Kabupaten Merauke, estimasi kemungkinan maksimum, distribusi Birnbaum–Saunders
MSC2020: 62F10, 62P20, 62E15

Citation: A. Langowuyo, S. Yokhu, and F. Reba, "Parameter Estimation and Analysis of Average Years of Schooling in Merauke District with Birnbaum–Saunders Distribution Approach", KUBIK J. Publ. Ilm. Mat., Vol. A, No. B, pp. 48-55, 2024.

Introduction

Education is one of the main pillars in the social and economic development of a region, especially in developing countries like Indonesia. One of the key indicators used to assess the success rate of education development is the average years of schooling, which reflects the number of years of formal education pursued by the population in a region. This indicator reflects not only access to education but also the quality of the education system as a whole [1]. Merauke Regency, as one of the largest regions in Papua Province with geographical and cultural diversity, faces significant challenges in education equity [2]. Limited infrastructure, accessibility, and socio-economic inequality are the main obstacles in increasing education participation, especially in remote areas [3].

Previous studies often rely solely on descriptive methods or conventional approaches in analyzing education data. However, the distribution of data such as average years of schooling is often asymmetrical and has low variability, so traditional approaches may not be effective enough to accurately characterize the data [4]. In this context, the Birnbaum-Saunders (BS) distribution approach becomes relevant due to its ability to handle data that is not symmetric and tends to be distributed with a certain pattern [5]. This distribution has been widely used in various fields such as system reliability, lifespan analysis, and risk management, but is rarely applied in Education analysis. Therefore, the utilization of this distribution is expected to provide new and deeper insights into the distribution pattern of average years of schooling [6].

This study aims to apply the Birnbaum-Saunders Distribution in analyzing the average years of schooling in Merauke District. By adopting this approach, the research not only provides a more accurate parameter estimation but also produces a distribution model that is relevant for understanding the condition of education in a region with complex geographical and socio-economic challenges. Furthermore, the results of this study are expected to serve as a reference in data-driven education policy planning, as well as open up opportunities for the exploration of innovative statistical approaches in education research in other regions with similar characteristics. Such a probabilistic-based approach also has the potential to make a scientific contribution to enriching the literature on non-conventional distribution-based education analysis, which is relevant to international journals with a multidisciplinary focus [7].

Basic Theory

Birnbaum – Saunders Distribution

A random variable X is said to be Birnbaum-Saunders open paren beta, gamma close paren distributed, if its probability density function is of the form [8],[9],[10],[11],[12]:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\left(\frac{\left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}}\right)^2}{2\gamma^2} \frac{\sqrt{\frac{x}{\beta}} + \sqrt{\frac{\beta}{x}}}{2\gamma x} \right); x > 0 \quad (1)$$

where:

$\gamma =$ Shape parameters $(\gamma > 0)$

$\beta =$ Scale parameter $(\beta > 0)$

Basic Concepts of MLE

Suppose that independent and identical sample data are obtained $X_1, X_2, \ldots, X_n$ which is assumed to come from a certain probability distribution $f(x, \theta)$where θ is the vector of parameters to be estimated. The likelihood function for such data is defined as [13,14, 15, 16, 24, 25].

$$L(\theta; X) = \prod_{i=1}^{n} f(X_i; \theta) \quad (2)$$

To simplify mathematical manipulation, the likelihood function is often converted to log-likelihood form:

$$\ell(\theta; X) = \ln L(\theta; X) = \sum_{i=1}^{n} \ln f(X_i; \theta) \quad (3)$$

MLE method tests to determine the value θ that maximizes the likelihood or log-likelihood function:

$$\hat{\theta} = \arg\max_{\theta} \ell(\theta; X) \quad (4)$$

Estimation Steps with MLE

Maximum Likelihood Estimation (MLE) is a parameter estimation method that aims to maximize the likelihood function, resulting in parameters that best fit the observed data. In the context of the Birnbaum-Saunders distribution, MLE is used to estimate the shape (α) and scale (β) parameters with a numerical approach, since the likelihood equation is complex. The following are systematic steps in parameter estimation using MLE [17].

a. Determine the Likelihood function $L(\theta; X)$ based on the probability distribution assumed by the data (e.g. Birnbaum - Saunders distribution)
b. Calculate the Log-Likelihood: Take the logarithm of the Likelihood function to get the log-likelihood function. $\ell(\theta; X)$. This aims to simplify the differentiation process
c. First derivative. Calculate the first derivative of the log-likelihood with respect to the parameter θparameter, and find the value θ that makes the value of the derivative equal to zero:

$$\frac{\partial \ell(\theta; X)}{\partial \theta} = 0 \quad (5)$$

d. Validate with the second derivative: To ensure that the value obtained is a maximum (not a minimum or an inflection point), check the sign of the second derivative:

$$\frac{\partial^2 \ell(\theta; X)}{\partial \theta^2} < 0 \quad (6)$$

e. Parameter Estimation: Values θ that satisfies these conditions becomes the maximum likelihood parameter estimate. $(\hat{\theta})$

Application to the Birnbaum-Saunders Distribution

In the Birbaum-Saunders distribution, the likelihood function for the parameters α (shape parameter) and β (scale parameter) are defined based on the probability density function (PDF) [18]:

$$f(t;\alpha,\beta) = \frac{1}{\sqrt{2\pi}\alpha t}\exp\left(-\frac{1}{2}\left(\frac{\beta}{t}+\frac{t}{\beta}-2\right)^2\right), t>0 \quad (7)$$

Using MLE, the parameters α and β can be estimated by maximizing the log-likehood function:

$$\ell(\alpha,\beta) = \sum_{i=1}^{n} \ln f(t;\alpha,\beta) \quad (8)$$

This estimation is often done with the help of statistical software, as the shape of the log-likelihood function for the Birnbaum - Saunders distribution tends to be complex and requires numerical algorithms to solve the nonlinear system of equations.

The Importance of Average Years of Schooling as an Education Indicator

Average years of schooling is a key indicator to measure the level of formal education attained by a region's population and is one of the components in the Human Development Index (HDI)**.** It reflects the quality and accessibility of the education system and its impact on socio-economic aspects, such as labor productivity, poverty reduction, and sustainable development [19], [1]. Factors such as geographical conditions, socio-economic disparities, limited educational infrastructure and cultural norms greatly influence the achievement of average years of schooling, especially in remote areas such as Merauke district. As a region with significant geographical and social challenges, an in-depth analysis of the distribution of education data is needed to understand education equity in the region and identify opportunities for more effective education policy development [4], [2].

Probabilistic Approach in Educational Analytics

The distribution of average years of schooling data **is** often characterized by skewed and low dispersion, requiring a flexible statistical approach. The Birnbaum-Saunders (BS) distribution**,** which was originally used for reliability analysis, proved to be effective for modeling data with this kind of pattern. The shape parameter α and scale parameter β of this distribution provide deep insight into the variability and averaging of educational data, allowing for more accurate analysis than traditional methods. The BS distribution offers an innovative approach to education analysis in areas with complex data distributions, such as Merauke, and can be used to support data-driven policy making. This approach is also relevant in a global context, in line with the Sustainable Development Goals (SDGs)**,** particularly goal 4 on quality education.

Test Statistics Value

To evaluate the fit of each probability distribution to the average years of schooling data. The Goodness of Fit (Gof) test used in this reaserch is the Kolmogorov-Smirnov (KS) test. KS test measure the maximum distanse between the emprical cumulative distribution function (CDF) of the sample and the hypothetical CDF. This statistic provides a robust measure of how closely the selected distribution matches the observed data. The KS statistic is formulated as follows [20], [21],[22],[23]:

$$D_{calculated} = sup_x|F(x) - H(x)| \quad (9)$$

where $D_{calculate}$ represents the maximum vertical distance between $F(x)$ and $H(x)$. The null hypothesis H_0 assert that the data follows the assumed distribution, while the alternatif hypothesis H_1 states that the data does follow that distribution.

Results and Discussion

Interpretation of Results

The results of the analysis using the Birnbaum-Saunders (BS) distribution of the average years of schooling data in Kabupaten Merauke provide important insights into the condition of education in the region. The value of the scale parameter (β=8.35) indicates that the average resident of Kabupaten Merauke has completed education up to the junior secondary school (SMP) level. This suggests that access to primary and junior secondary education is relatively good, but also highlights significant challenges in encouraging people to continue their education to senior secondary and higher education. This figure may reflect limitations in education infrastructure, particularly in remote areas, and socio-economic barriers such as poverty, lack of teacher resources or low motivation for education.

The shape parameter (α=0.0545) indicates that the distribution of the average years of schooling data has a low degree of dispersion, indicating high homogeneity around the mean value. This means that the majority of the population has almost uniform years of education, which may reflect success in equalizing access to basic education. However, this homogeneity may also indicate potential stagnation at certain levels of education. In other words, a large proportion of the population only completes a certain level of education without going on to higher levels. This could be due to geographical and cultural constraints, including people's perceptions of the importance of secondary education and limited job opportunities that require higher education.

Using the KS test, the success of the Birnbaum-Saunders Distribution model in modeling this data demonstrates its flexibility in capturing unsymmetrical data distribution patterns, which are often found in education data. Assuming that the average years of schooling data may not be normally distributed, the BS approach proved to be more effective than traditional parametric methods. The parameter estimates obtained provide richer information about the structure of the data, allowing for a deeper understanding of the dynamics of education in the region. The use of this approach also opens up opportunities for further analysis, such as exploring the relationship between average years of schooling and other indicators such as welfare, unemployment, or regional economic development. Figure 1 shows the probability density function (PDF) of the Birnbaum-Saunders distribution used in the analysis, and it also illustrates the dispersion pattern of the average years of schooling data.

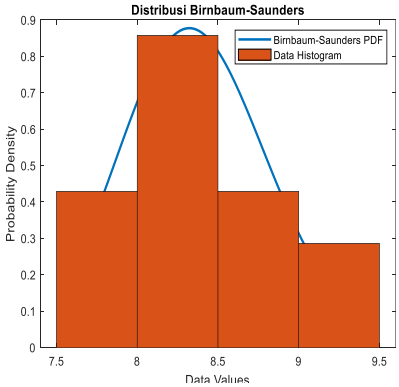

Figure 1. PDF Chart of Average Years of Schooling Data

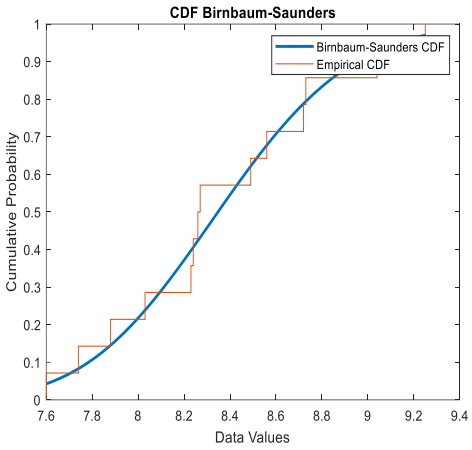

Figure 2. CDF Graph of Average Years of Schooling Data

Figure 2 visualizes the cumulative distribution function (CDF) of the analyzed data. It provides an overview of the accumulated proportion of the population based on their average years of schooling, helping in understanding the homogeneity of the data distribution. This finding has significant policy implications. The government can use these results to design a more data-driven education policy, focusing on improving access and quality of education at the upper secondary level. Efforts such as developing education infrastructure in remote areas, providing scholarships for students from underprivileged families, and training and incentives for teachers can help increase school enrollment rates at higher levels. In addition, it is important to integrate education with local labor market needs to improve the relevance of education and encourage people to continue their education to higher levels.

Overall, this interpretation suggests that although Merauke district has made progress in primary education equity, greater efforts are still needed to improve higher education levels. The results of this study not only provide empirical insights into the condition of education in the region but also reinforce the importance of innovative statistical approaches such as the Birnbaum-Saunders Distribution for understanding education data with complex characteristics. These interpretations provide important contributions to the academic literature as well as relevant recommendations for education policy makers.

Discussion

The results of the analysis show that the average years of schooling in Kabupaten Merauke, with a scale parameter value (β) of 8.35 years, reflects educational attainment at the junior secondary school level. This value provides a picture of relatively good access to primary and junior secondary education. However, this figure also reveals a major challenge in encouraging people to continue their education to higher levels, especially senior secondary school and higher education. Geographical factors, such as accessibility to educational facilities, as well as socio-economic constraints, such as poverty and lack of educators, can be significant barriers. This finding is in line with previous research that shows an education gap in the Papua region compared to other regions in Indonesia.

The value of the shape parameter (α) of 0.0545 indicates that the data distribution of the average length of schooling in Kabupaten Merauke has a low level of dispersion. This indicates that the length of education taken by the majority of the population tends to be homogeneous, with little

variation around the average. This homogeneity can be considered a success in equalizing access to basic education in Kabupaten Merauke. However, this low level of variation may also indicate the potential for stagnation at certain levels of education, where most people only complete education up to a certain level without continuing further. This requires more targeted policies to encourage people to continue their education, such as providing scholarships or incentives for students from remote areas.

The Birnbaum-Saunders (BS) Distribution approach in this study was shown to provide advantages over traditional analysis methods. This distribution is able to handle education data that often has asymmetric or non-normal characteristics. The success of this model in describing the distribution pattern of average years of schooling in Merauke district shows that this probabilistic-based approach is not only relevant for engineering or lifespan analysis but also very useful in social and educational analysis. As a flexible distribution, Birnbaum-Saunders allows the identification of more realistic data distribution patterns, particularly in situations where conventional parametric methods may fail to provide accurate results.

The results of this study have important policy implications. With an average years of schooling that tends to be close to junior secondary school and a low level of data dispersion, efforts to improve education quality and access should focus on several key areas. First, greater investment in education infrastructure is needed, especially in remote areas of Merauke district, to improve students' accessibility to higher levels of education. Secondly, the government needs to develop intervention programs that can address the socio-economic factors that are the main barriers for students to continue their education, such as financial assistance, transportation and teacher training. Thirdly, education improvement policies should include long-term strategies to integrate education with local labor market needs so that higher education can be perceived as relevant by the community.

From a methodological perspective, this study also makes an important contribution in extending the use of the Birnbaum-Saunders Distribution to the field of education. This distribution, which was originally designed for system reliability analysis, shows great flexibility in modeling social data with complex distribution patterns. This approach can be used as a basis for further analysis, such as exploring the relationship between average years of schooling and other social indicators, including unemployment rates, poverty rates, or regional economic development. In addition, this approach opens up opportunities for the application of other probabilistic distributions capable of handling more complex characteristics of education data, such as the skewed-t distribution or the log-normal distribution.

In conclusion, this discussion emphasizes the importance of using innovative statistical approaches to understand education dynamics in areas with unique challenges, such as Merauke district. This research not only enriches the literature on probabilistic distribution-based education analysis but also provides practically relevant recommendations for evidence-based education policy-making in areas with similar geographical and socio-economic conditions. Further research could focus on integrating this analysis with other factors, such as the presence of infrastructure or the level of community participation in education programs, to provide a more comprehensive picture.

Conclusion

The results of the analysis using the Birnbaum-Saunders (BS) distribution of the average years of schooling data in Kabupaten Merauke show that the average years of schooling of the population is at 8.35 years, as indicated by the scale parameter (β). This parameter reflects the estimated median value of the data distribution and illustrates that the average resident of Kabupaten Merauke has

completed education up to the junior high school (SMP) level. This value indicates good educational achievement at the primary and junior secondary levels, although it still requires attention to improve access to higher levels of education.

In addition, the shape parameter (α=0.0545) shows that the data has a low level of variability, with the distribution centered around the mean. This indicates a high degree of homogeneity in the average years of schooling of the population. In other words, most of the population has similar years of formal education, which reflects the success in equalizing access to education in the region. However, this low variability may also indicate potential stagnation, which needs to be followed up with policies to encourage continuation of education to higher levels.

This study confirms the importance of the scale parameter (β) as a measure of the average education of the population, which in this context becomes a reference for identifying successes and challenges in educational development. With the flexibility of the Birnbaum-Saunders Distribution in handling data with certain distribution patterns, the results of this analysis provide a strong basis for policy making. The government can utilize these results to focus on improving senior secondary education and addressing geographical and socio-economic barriers, thus supporting more sustainable development in Merauke district, Papua.

References

[1] Manurung, E. N., & Hutabarat, F. (2021). The Effect of Expected Years of Schooling, Average Years of Schooling, Expenditure per Capita on the Human Development Index. *Scientific Journal of Management Accounting.* https://doi.org/10.35326/jiam.v4i2.1718

[2] Putra, D. J., & Ekasari, A. (2023). Analysis of the Contribution of Education Indicators to the Human Development Index (HDI). *Journal on Teacher.*

[3] Juliarini, A., & Hatmoko, A. W. (2020). The Effect of Special Autonomy Funds on the Human Development Index in Papua. In *National Symposium* on *State Finance*.

[4] Manuhuttu, F. Y., & Hartadi, F. (2022). The Effect of Human Development Index and Economic Growth on Poverty. *Musamus Journal of Economics Development.* https://doi.org/10.35724/feb.v4i1.4161

[5] Martínez-Flórez, G., Barranco-Chamorro, I., Bolfarine, H., & Gómez, H. W. (2019). Flexible Birnbaum-Saunders distribution. *Symmetry.* https://doi.org/10.3390/sym11101305

[6] Martínez-Flórez, G., Tovar-Falón, R., & Barrera-Causil, C. (2022). Inflated Unit-Birnbaum-Saunders Distribution. *Mathematics.* https://doi.org/10.3390/math10040667

[7] de Oliveira, J. F., de Souza, A., Abreu, M. C., Nunes, R. S. C., de Sousa Nascimento, L., da Silva, S. D., Filho, W. L. F. C., & da Silva, E. B. (2023). Modeling of Dengue by Cluster Analysis and Probability Distribution Functions in the State of Alagoas in Brazilian. *Brazilian Archives of Biology and Technology.* https://doi.org/10.1590/1678-4324-2023220086

[8] Lieve, Victor, (2016). The Birnbaum – Saunders Distribution. Elsevier. London

[9] H.K.T., Kundu., Balakrishnan, (2003). Modified Moment Estimation for the Two Parameter Birnbaun – Sanders Distribution.

[10] Balakrishnan., Kundu, (2018). Birnbaun – Saunders Distribution A review of Models, Analisis and Application. Stat. ME

[11] Dupuis,D.J., Mills, J.E. (1998). Robust Estimation of the Birnbaun – Saunders Distribution. IEE Transaction On Realibility, Vol 47, No. 7.

[12] Balakrishnan, N,. Leiva, V., Sanhueza, A., Vilca,F. (2009).”Estimation in the Birnbaum – Saunders distribution based on scala-mixture of normal an the EM-algorithm”, Statistic an Opration Research Transactions, vol. 33, 171-191

[13] Sroyer, A. M., Tandiangnga, T., Reba, F., & Beno, I. S. (2018). Comparison of Parameter Estimation Methods to Determine the Frequency Data Magnitude of Aftershock in Nabire, Papua. Journal of Science & Science Education, 2(2), 17-21. http://ejournal.uksw.edu/josse/article/view/1969

[14] Jiang, H., Wang, J., Wu, J., & Geng, W. (2017). Comparison of numerical methods and metaheuristic optimization algorithms for estimating parameters for wind energy potential assessment in low wind regions. Renewable and Sustainable Energy Reviews, 69 (December 2015), 1199-1217. https://doi.org/10.1016/j.rser.2016.11.241

[15] ul Haq, M. A., Rao, G. S., Albassam, M., & Aslam, M. (2020). Marshall-Olkin Power Lomax distribution for modeling of wind speed data. Energy Reports, 6(May), 1118-1123. https://doi.org/10.1016/j.egyr.2020.04.033

[16] Balakrishnan,N., Cohen,A.C. (1991), Order Satatistics an Inference: Estimation Methods, Academic Press, San Diego.

[17] Teimouri, M., Hoseini, S. M., & Nadarajah, S. (2013). Comparison of estimation methods for the Weibull distribution. Statistics. https://doi.org/10.1080/02331888.2011.559657

[18] Parajuli, A. (2023). Evaluation of Nakagami and Birnbaum-Saunders probability distribution for wind speed and power estimation. Journal of Engineering Issues and Solutions. https://doi.org/10.3126/joeis.v2i1.49596

[19] Asmawani, & Pangidoan, E. (2021). The Effect of Life Expectancy, Average Years of Schooling, Economic Growth and Per Capita Expenditure on the Human Development Index in North Sumatra Province. Journal of Economic Sciences.

[20] Romeu,J.L. Kolmogorov - Smirnov : A Goodness of Fit Test for Small Samples, RAC START Vol. 10, No. 6

[21] Sahoo, Prasanna. (2013). Probability and Mathematical Statistics. Depertment of Mathematics University Of Louisville. Louisville.

[22] Rohatgi,V.K. (1976). An Introduction to Probability an Mathematical Statistics, Wiley, NY.

[23] Romeu, J.L. Stattistical Confidance. RAC START, Vol. 9, No. 4.

[24] Martina, A., Dewi, A. S., & Awalluddin, A. S. (2024). Peramalan Menggunakan Model Holt-Winters Exponential Smoothing Multiplikatif dengan Optimasi Parameter Menggunakan Particle Swarm Optimization (PSO). KUBIK: Jurnal Publikasi Ilmiah Matematika, 9(2), 161–171. https://doi.org/10.15575/kubik.v9i2.31368

[25] Marcelina, S., Awalluddin, A. S., Huda, A. F., Ramdani, R., & Sukaesih, E. (2022). Analisis Pengukuran Produk HKZL PT.Gradien Manufaktur Indonesia Menggunakan Multivariat Gage, Repeatability and Reproducibility (GRR) Melalui Analisis Faktor. KUBIK: Jurnal Publikasi Ilmiah Matematika, 6(2), 110–119. https://doi.org/10.15575/kubik.v6i2.14555