BAREKENG: Journal of Mathematics and Its Applications
March 2024 Volume 18 Issue 1 Page 0313–0322
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol18iss1pp0313-0322

PREDICTIVE DISTRIBUTION TO DETERMINE LEARNING MODEL AT THE STRATEGIC COMPETENCE LEVEL OF STUDENTS IN STATISTICS GROUP COURSE

Mokhammad Ridwan Yudhanegara[1*], **Karunia Eka Lestari**[2]

[1,2]*Departement of Mathematics Education, Faculty of Teacher Training and Education, Universitas Singaperbangsa Karawang, St. H. S. Ronggowaluyo Teluk Jambe Timur, Karawang, 41361, Indonesia*

*Corresponding author's e-mail: ** ***mridwan.yudhanegara@staff.unsika.ac.id***

Article History:

Received: 25th August 2023
Revised: 3rd December 2023
Accepted: 2nd January 2024

Keywords:

Learning Model;
Multinomial;
Statistical Test;
Statistics Course.

ABSTRACT

The problem of this research comes from a situation or condition that is not static. The description of these problems is the condition of the learning system, which tends to change due to the COVID-19 pandemic, causing learning conditions to be dynamic. From a statistical perspective, the dynamic situation can be modeled using a predictive distribution approach so its characteristics can be studied. The purpose is to provide policy recommendations on appropriate learning models for lecturers in improving students' strategic competence, which is an ability that students need to master in solving various mathematical problems. The main discussion of this paper consists of three parts: clustering, predictive distribution, and statistical inference. The purpose of clustering is to group students based on test results to determine the level of strategic competence. In addition, clustering is also used as an initial process to predict students' strategic competence level if the learning used is still the same. The benefits of statistical inference in the distribution procedure in this study are used to determine the type of data distribution from each arrival of new information or data. The results of the statistical inference determine whether or not it is necessary to update the learning model of the lecturer. This research produces a new alternative statistical inference needed to make decisions. Based on the simulation results and discussion, the use of a predictive distribution approach to predict dynamic data is very appropriate. The distribution approach can be used for detecting changes in new data distribution with historical data for the dynamic condition. If the changes are insignificant, direct instruction can still be used for the learning model in statistics courses. A new learning model is recommended for the statistics group course at a higher level when the changes are significant.

How to cite this article:

M. R. Yudhanegara and K. E. Lestari., "PREDICTIVE DISTRIBUTION TO DETERMINE LEARNING MODEL AT THE STRATEGIC COMPETENCE LEVEL OF STUDENTS IN STATISTICS GROUP COURSE," *BAREKENG: J. Math. & App.*, vol. 18, iss. 1, pp. 0313-0322, March, 2024.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

To provide sufficient understanding to students about the importance of applying statistics, it is necessary to provide statistics courses. Statistics is a branch of mathematics concerned with developing and studying methods for collecting, analyzing, interpreting, and presenting empirical data [1].

Current research issues in academic circles focus on Sustainable Development Goals (SDGs) [2][3]. SDGs are the main focus of countries that are members of the United Nations in state management procedures [3]. These SDGs exist thanks to the awareness that natural resources originating from the environment are believed to be depleted. So, we humans who live in it must use the environment rationally and protect it for the sake of the survival of humanity and all living creatures on earth.

SDGs are a series of goals set by the United Nations to achieve a better and more sustainable life for everyone globally. SDGs are global and national commitments to improve society, including 17 global goals and targets for 2030 [4], declared by both developed and developing countries at the United Nations General Assembly in September 2015 [3].

One of the important competencies related to achieving the SDGs is mastery of statistics [5]. Mastering statistics enables individuals to understand and analyze data, contributing to improving education quality in Indonesia. Mastery of Statistics courses also contributes to achieving the SDGs indicators, namely the fourth indicator related to Education Quality.

Statistics and mathematics are closely related fields and often intersect in various ways. Mathematics provides the theoretical foundation for statistical methods and concepts like probability theory, calculus, and linear algebra. Statistical analysis usually involves mathematical calculations, such as finding the mean, median, and standard deviation and using mathematical models to make predictions and draw conclusions from data. Statistics also contributes to mathematics by providing real-world applications for mathematical theories and concepts.

The relationship between statistics and mathematics is symbiotic, with each field informing and improving the other. Discussion of mathematics learning means not also discussing statistics learning. Likewise, mathematical knowledge contains statistical knowledge.

Kilpatrick et al. [6] stated the results of their research on knowledge, mathematical knowledge, understanding, and people's skills for successful mathematics learning. Mathematical proficiency is an aspect of achieving successful mathematics learning. Mathematical proficiency has five strands: conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition.

At the level of higher education in university, strategic competence is considered the most important for successful mathematics learning. Strategic competence is the ability to formulate, represent, and solve mathematical problems. The indicators understand the problem, choose the suitable formulation, present a problem in various forms of mathematical representation, choose a strategy to solve the problem, use or develop problem-solving strategies, solve the problem, and interpret answers.

The problem of this research comes from a situation or condition that is not static. The description of these problems is the condition of the learning system, which tends to change due to the COVID-19 pandemic, causing learning conditions to be dynamic. From a statistical perspective, the dynamic situation can be modeled using a predictive distribution approach so its characteristics can be studied.

This paper uses predictive distribution to analyze data in making decisions. Another term for predictive distribution is posterior predictive distribution. A predictive distribution is a distribution of possible unobserved values conditional on the observed values. For example, we want to look for the conditional probability distribution of observation y (unobserved) given observation $\mathbf{x}$, then the conditional probability function is

$$f(y|\mathbf{x}) = \int f(y|\theta) f(\theta|\mathbf{x})\, d\theta. \qquad (1)$$

To understand predictive distributions, it is necessary to understand prior and posterior distributions, which will be discussed in the next section.

The aim of this research is to produce a new alternative statistical inference needed to make decisions for learning model in statistics course. The purpose is to provide policy recommendations on appropriate learning models for lecturers in improving students' strategic competence, which is an ability that students need to master in solving various mathematical problems.

2. RESEARCH METHODS

2.1 Statistics Course

In the statistics group course, students are required to reason statistically, which includes understanding the data, making hypotheses about what the data might show, investigating hypotheses, and writing valid conclusions [7]. Strategic competence is required for achieving, as statistical reasoning is only one part of strategic competence. The five strands are illustrated in the intertwined rope in **Figure 1**[6].

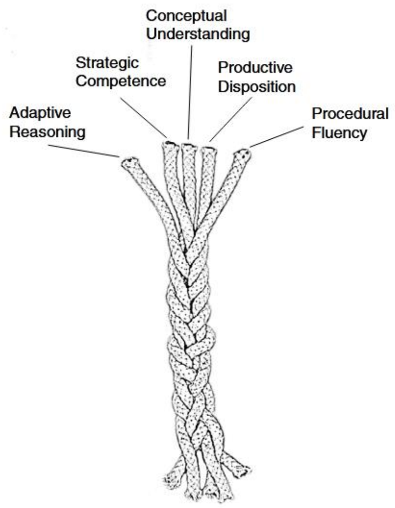

Figure 1. Intertwined Strands of Proficiency

In our department, the statistics group course consists of three courses. There are basic statistics in the third semester, probability theory in the sixth semester, and inferential statistics in the seventh semester. The basic statistics course is considered the lowest level; the probability theory course is the intermediate level, and the inferential statistics course is at the top level of the statistics group course for strategic competence level, see **Figure 2**.

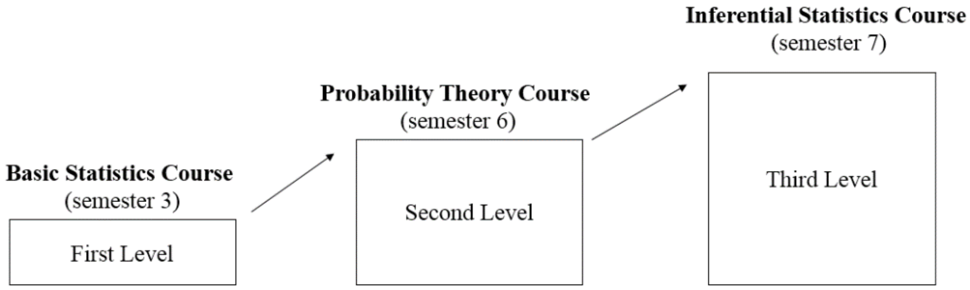

Figure 2. Strategic Competence Level in Statistics Group Course

An appropriate learning model is needed to improve students' strategic competence. The learning model is a pattern of interaction between students and teachers. It consists of learning strategies, approaches, methods, and techniques applied in implementing in-class learning activities. The decision to use a learning model was taken based on the data obtained from the learning outcomes.

Generally, if the resulting data does not change the distribution, the learning model tends not to change [8] [9]. In general, if the data obtained does not change the distribution, then the learning model tends not to change. For example, lecturers will continue to use the direct instruction model if the data distribution does not change. However, if distribution changes, e.g., project-based learning becomes an alternative model used. For this reason, the role of predictive distribution as a method in statistics to deal with dynamic data. This case is illustrated in **Figure 3**.

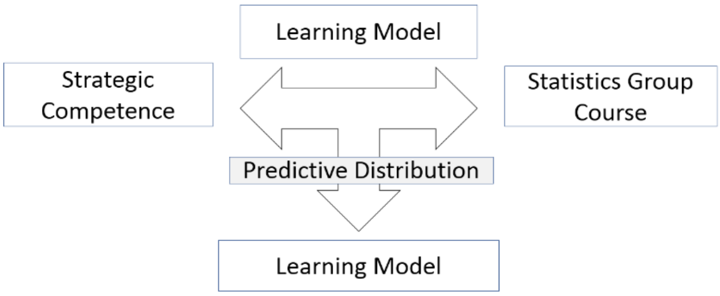

Figure 3. Predictive Distribution to Determine Learning Model

2.2 Multinomial Distribution

In general, multinomial data is the number of frequencies in several data categories. Characteristics of a multinomial distribution include that each experiment has more than two possible events (outcomes) that will occur; the number of experiments was carried out n times, and each experiment was statistically independent so that the events resulting from one experiment did not affect the next experiment; and the probability of each event on each trial does not change.

If an experiment has m possible outcomes, say $\kappa_1, \kappa_2, \cdots, \kappa_m$ with each probability $\theta_1, \theta_2, \cdots, \theta_m$, then the probability distribution of the random variables $X_1, X_2, \cdots, X_m$ which describes the number of occurrences of $\kappa_1, \kappa_2, \cdots, \kappa_m$ in n independent trials will follow a multinomial distribution. The multinomial distribution is denoted $\mathrm{Mult}(\theta_1, \theta_2, \cdots, \theta_m, n)$, with a probability mass function (PMF) as follows.

$$p(x_1, x_2, \cdots, x_m; \theta_1, \theta_2, \cdots, \theta_m) = \binom{n}{x_1\, x_2 \cdots x_m} \theta_1^{x_1} \theta_2^{x_2} \cdots \theta_m^{x_m} \quad (2)$$

or

$$p(\mathbf{x}; \boldsymbol{\theta}) = \frac{n!}{\prod_{i=1}^{m} x_i!} \prod_{i=1}^{m} \theta_i^{x_i} \quad (3)$$

where $\mathbf{x} = [x_1\ x_2 \cdots x_m]^T, \sum_{i=1}^{m} x_i = n, n \in \mathbb{N}, \boldsymbol{\theta} = [\theta_1\ \theta_2\ \cdots\ \theta_m]^T$, $\sum_{i=1}^{m} \theta_i = 1, \theta_i > 0$, and $\theta_1, \theta_2, \cdots, \theta_m$ are parameters of probability.

2.3 Predictive Distribution for Multinomial Case

The prior distribution is an initial distribution that provides information about parameters and must be determined first before formulating the posterior distribution. Prior distributions have an important role in estimating unknown parameters. In general, a prior distribution is a probability distribution of parameter values [10][11].

Suppose the random variable X has a probability distribution with unknown parameter $\theta \in \Omega$. These parameters are assumed to be random variables, denoted Θ, with a probability distribution called the prior distribution. Next, x is the observed value of X, and θ is the observed value of Θ. The distribution of X depends on Θ. The probability density function (pdf) of Θ is denoted $h(\theta)$, with $h(\theta) = 0$ if $\theta \notin \Omega$. The function $h(\theta)$ is called the prior pdf of Θ so that $X|\theta \sim f(x|\theta)$, and $\Theta \sim h(\theta)$. The prior distribution for the parameters of a multinomial distribution is the Dirichlet distribution.

The posterior distribution is the conditional probability distribution of parameters given observational data [10][11]. Suppose the random variables $X_1, X_2, \ldots, X_l$ are mutually independent of the conditional distribution X with $\Theta = \theta$ and pdf $f(x|\theta)$. For random vectors $\mathbf{X} = [X_1\ X_2 \cdots X_l]^T$ and $\mathbf{x} = [x_1\ x_2, \cdots\ x_l]^T$, so conditional pdf $\mathbf{X}$ if we have $\Theta = \theta$ is

$$f(\mathbf{x}|\theta) = f(x_1|\theta) f(x_2|\theta) \cdots f(x_l|\theta). \quad (4)$$

The conditional pdf $\mathbf{X}$ and Θ is

$$f(\mathbf{x}, \theta) = f(\mathbf{x}|\theta) h(\theta), \quad (5)$$

and pdf of marginal $\mathbf{X}$ is

$$h(\mathbf{x}) = \int_{-\infty}^{\infty} f(\mathbf{x},\theta)d\theta. \tag{6}$$

For conditional pdf Θ, if we have $\mathbf{X}$ is

$$f(\theta|\mathbf{x}) = \frac{f(\mathbf{x},\theta)}{h(\mathbf{x})} = \frac{f(\mathbf{x}|\theta)h(\theta)}{\int_{-\infty}^{\infty} f(\mathbf{x}|\theta)h(\theta)d\theta}, \tag{7}$$

hereinafter called the posterior distribution function.

Let $\mathbf{X}_t = \left[X_{1,t}\ X_{2,t} \cdots X_{m,t}\right]^T$ be a random vector representing the number of occurrences in each time t, where $X_{i,t}$ is a random variable that represents the number of independent events at the i-th time. The random vector $\mathbf{X}_t$ is assumed to have a multinomial distribution with the parameter $\boldsymbol{\theta}$. In this study, the vector $\boldsymbol{\theta}$ is the observed value of the random vector $\boldsymbol{\Theta}$. The random vector $\boldsymbol{\Theta}$ has a distribution that acts as the prior of the multinomial, with $\boldsymbol{\Theta}$ coming from the Dirichlet distribution [8]. Then look for the conditional probability distribution for the observation $\mathbf{x}_{t+k}$ if given the observation $\mathbf{D}_{t+(k-1)} = \left[\mathbf{x}_1\ \mathbf{x}_2 \cdots \mathbf{x}_{t+(k-1)}\right]^T$. The predictive distribution function is

$$p\left(\mathbf{x}_{t+k}\middle|\mathbf{D}_{t+(k-1)}\right) = \int p(\mathbf{x}_{t+k}|\boldsymbol{\theta})f\left(\boldsymbol{\theta}\middle|\mathbf{D}_{t+(k-1)}\right) d\theta. \tag{8}$$

For multinomial case, we have **Equation (9)** [8][9],

$$p\left(\mathbf{x}_{t+k}\middle|\mathbf{D}_{t+(k-1)}\right) = \frac{n!\Gamma(\sum_{i=1}^{m}\alpha_i')}{\Gamma(n+\sum_{i=1}^{m}\alpha_i')}\prod_{i=1}^{m}\frac{\Gamma(x_i+\alpha_i')}{x_i!\Gamma(\alpha_i')}. \tag{9}$$

The equation is a predictive distribution probability function which is a probability mass function of the Dirichlet-multinomial. In the Dirichlet-multinomial case, we can take the value $\alpha_0 = \sum_{i=1}^{m}\alpha_i$ dan $\theta_i = \frac{\alpha_i}{\sum_{i=1}^{m}\alpha_i} = \frac{\alpha_i}{\alpha_0}$ [12], so that the following relationship can be seen:

$$E\left[X_{i,t+k}\right] = n\,\theta_i = n\,\frac{\alpha_i}{\alpha_0}. \tag{10}$$

Thus, we can find θ_i from the multinomial estimated parameter, namely:

$$\hat{\theta}_i = \frac{\sum_{t=1}^{l} x_{i,t}}{l\,n}, i = 1,2,..,m, \tag{11}$$

where $\sum_{i=1}^{m} x_i = n, n \in \mathbb{N}$.

In this case, the number of events indicates successive strategic competence. The indicators of strategic competence, i.e.

a. To understand the problem, seen as the first occurrence;

b. To choose the relevant formulation, seen as a second occurrence;

c. To present a problem in various forms of mathematical representation, seen as a third occurrence;

d. To choose a strategy to solve the problem, seen as the fourth event;

e. To use or develop problem-solving strategies, seen as the fifth occurrence;

f. To solve the problem, seen as the sixth occurrence;

g. To interpret answers, seen as the seventh occurrence.

For example, in the basic statistics course, the number of students who only reached indicator 4 was 30, while the total number of students was 170. The 40 students certainly succeeded in achieving indicators 1, 2, 3, and 4. In this case, the 40 students were defined as the number of occurrences in indicator 4.

2.4 Statistical Test

Usually, to test the suitability of a model for data with a multinomial distribution, chi-squared is used. This statistical form is derived from the multinomial likelihood ratio. The derivation process is carried out using an approximation process to obtain the statistical form [13][14]. This approximation results in the level of accuracy of a statistical test being reduced. The statistical form presented in this research was analytically obtained without approximation. Based on these reasons, research does not use existing statistical forms to make decisions. The strategy used is to find other forms of statistics.

The new method in this research is built from the logarithmic distribution of the likelihood function. Through this method, it is possible to determine statistical inferences using a normal statistical test approach iteratively.

Theorem 1 [9] provides information that the transformation of multinomial case data produces a normal distribution for random variables $L_l(\boldsymbol{\theta}) = \prod_{t=1}^{l}\left(\frac{n!}{\prod_{i=1}^{m} x_{i,t}!}\prod_{i=1}^{m}\theta_i^{x_{i,t}}\right)$.

Theorem 1. Let $L_l(\boldsymbol{\theta}) = \prod_{t=1}^{l}\left(\frac{n!}{\prod_{i=1}^{m} x_{i,t}!}\prod_{i=1}^{m}\theta_i^{x_{i,t}}\right)$ be a multinomial likelihood function, for $n, m, l \gg 1$, then the random variable $log\, L_l(\boldsymbol{\theta})$ has a normal distribution with mean:

$$\ddot{\mu} = \sum_{t=1}^{l}\sum_{i=1}^{m} x_{i,t}\log\theta_i - \sum_{t=1}^{l}\sum_{i=1}^{m}\log x_{i,t}! + \sum_{t=1}^{l}\log n! \quad (12)$$

and variance:

$$\ddot{\sigma}^2 = \frac{m}{(m-1)}\sum_{t=1}^{l}\sum_{i=1}^{m}\left(x_{i,t}\log\theta_i - \frac{\sum_{i=1}^{m} x_{i,t}\log\theta_i}{m}\right)^2 - \sum_{t=1}^{l}\sum_{i=1}^{m}\left(\log x_{i,t}! - \frac{\sum_{i=1}^{m}\log x_{i,t}!}{m}\right)^2. \quad (13)$$

After **Theorem 1**, we have **Corollary 1** [9]

Corollary 1. Let $L_l(\boldsymbol{\theta}) = \prod_{t=1}^{l}\left(\frac{n!}{\prod_{i=1}^{m} x_{i,t}!}\prod_{i=1}^{m}\theta_i^{x_{i,t}}\right)$ be a multinomial likelihood function, for $n, m, l \gg 1$, then random variable $\left(log\, L_l(\boldsymbol{\theta}) - \sum_{t=1}^{l} log\, n!\right)$ has a normal distribution with mean:

$$\ddot{\mu} = \sum_{t=1}^{l}\sum_{i=1}^{m} x_{i,t}\log\theta_i - \sum_{t=1}^{l}\sum_{i=1}^{m}\log x_{i,t}! \quad (14)$$

and variance $\ddot{\sigma}^2 = \ddot{\sigma}^2$ based on **Equation (13)**.

The statistical form for multinomial goodness-of-fit test, it applies the following hypothesis test:

H₀: $\left(\log L_{l+1}(\boldsymbol{\theta}) - \sum_{t=1}^{l+1}\log n!\right)$ *follows a normal distribution with the mean* $\ddot{\mu}$ *and variance* $\ddot{\sigma}^2$,
H₁: $\left(\log L_{l+1}(\boldsymbol{\theta}) - \sum_{t=1}^{l+1}\log n!\right)$ *does not follow a normal distribution with the mean* $\ddot{\mu}$ *and variance* $\ddot{\sigma}^2$.

The steps for testing new data based on historical data are described in **Algorithm 1**.

Algorithm 1

1) *input* $(\ddot{\mu}, \ddot{\sigma}^2)$;
2) *find the estimated parameter* $(\boldsymbol{\theta})$ *of the predictive distribution (based on historical data);*
3) *enter the new observed value under the assumption* H_0;
4) *perform the Z statistical test at the significance level α.* H_0 *is not rejected, it means that* $(l+1) \sim N(\ddot{\mu}, \ddot{\sigma}^2)$;
5) *update the parameters* $(\boldsymbol{\theta})$ *and* $(\ddot{\mu}, \ddot{\sigma}^2)$ *if* H_0 *is rejected;*
6) *return to step 3.*

Furthermore, the flow of the problem-solving approach is presented in **Figure 4**. Starting from the data from the evaluation of the statistical group lectures, they are faced with two static and dynamic conditions.

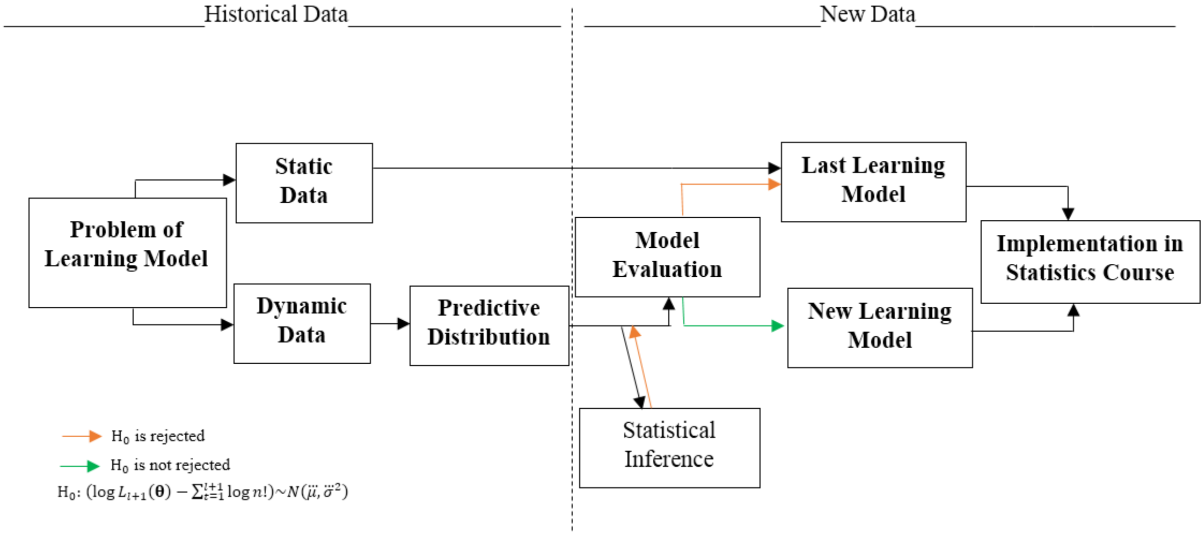

Figure 4. Predictive Distribution to Determine Learning Model

Based on **Figure 4**, it can be seen that the problem in deciding the learning model to be taken at a higher level in group course statistics is when dynamic data occurs.

Therefore, an approach is needed to overcome these dynamic conditions, namely predictive distribution. If the condition is static, the last learning model (e.g., direct instruction) can still be used for learning models in the statistics group course at a higher level. If the condition is dynamic, we use a distribution approach to detect changes in new data distribution with historical data (based on predictive distribution). If the changes are insignificant, direct instruction can still be used for the next learning model. However, if the changes are significant, it is recommended to use a new learning model (e.g., project-based learning) in the statistics group course at a higher level.

2.5 Prediction Accuracy

According to Lewis [15], MAPE scores can be interpreted into four categories, see **Table 1**.

Table 1. Criteria of MAPE

Percentage	Interpretation
MAPE score $< 10\%$	excellent prediction
$10\% \leq$ MAPE score $< 20\%$	good prediction
$20\% \leq$ MAPE score $< 50\%$	fair prediction
MAPE score $\geq 50\%$	not accurate prediction

The smaller the MAPE score, the smaller the prediction error; on the contrary, the larger the MAPE score, the greater the prediction error. The results of a predictive model have the excellent predictive ability if the MAPE score is less than 10% and good predictive ability if the MAPE score is between 10% and 20%.

The **Table 1** shows the error percentage score on MAPE; namely, the MAPE score can still be used if it does not exceed 50%, and if the MAPE score is above 50%, then the prediction model can no longer be used. The MAPE formula is as follows,

$$\text{MAPE} = \frac{1}{m} \sum_{i=1}^{m} \left| \left(\frac{x_{i,t} - \hat{x}_i}{x_{i,t}} \right) \right| \times 100\%, \tag{15}$$

where $\hat{x}_i = n \frac{\alpha_i}{\alpha_0}$, $\alpha_0 = \sum_{i=1}^{m} \alpha_i$, $\frac{\alpha_i}{\sum_{i=1}^{m} \alpha_i} = \theta_i$, and $\sum_{i=1}^{m} \theta_i = 1$, $\theta_i > 0$.

3. RESULTS AND DISCUSSION

Let $\mathbf{X}_t = [X_{1,t} X_{2,t} \cdots X_{m,t}]^T$ be a random vector that states the number of students who achieve strategic competence indicators in each semester t, where $X_{i,t}$ is a random variable that states the number of students who achieve the i-th indicator in semester t. Furthermore, the data is generated 40 times under the assumption of a multinomial distribution with the pmf $\mathbf{X}_t \sim \text{Mult}\,(\theta_1, \theta_2, \cdots, \theta_7, 140)$ namely

$$p(\mathbf{x}; \boldsymbol{\theta}) = \frac{140!}{\prod_{i=1}^{7} x_i!} \prod_{i=1}^{7} \theta_i^{x_i}, \tag{16}$$

where $\sum_{i=1}^{7} \theta_i = 1$ and $\sum_{i=1}^{7} x_i = 140$.

Simulations of 40 data were carried out to make prediction based on historical data. The 24 data are used as historical data or training, and 16 data are used as new data or testing. The statistic used to analyze the data is the Z statistic test, which has been formulated in **Equation (17)** [9], with the following hypothesis.

$\mathbf{H_0}$: $\left(\log L_{l+1}(\boldsymbol{\theta}) - \sum_{t=1}^{l+1} \log n!\right)$ follows a normal distribution with the mean $\ddot{\mu}_l$ and variance $\ddot{\sigma}_l^2$,
$\mathbf{H_1}$: $\left(\log L_{l+1}(\boldsymbol{\theta}) - \sum_{t=1}^{l+1} \log n!\right)$ does not follow a normal distribution with the mean $\ddot{\mu}_l$ and variance $\ddot{\sigma}_l^2$.

The Z statistic test formula is

$$z_{statistical} = \frac{\left(\log L_{l+1}(\boldsymbol{\theta}) - \sum_{t=1}^{l+1} \log n!\right) - \ddot{\mu}_l}{\ddot{\sigma}_l}, \tag{17}$$

where $\log L_{l+1}(\boldsymbol{\theta}) - \sum_{t=1}^{l+1} \log n!$ as a random variable under the assumption $\mathbf{H}_0$. The $\mathbf{H_0}$ is not rejected when $z_{\left(\frac{\alpha}{2}\right)} \le z_{statistical} \le z_{\left(1-\frac{\alpha}{2}\right)}$, and $\mathbf{H}_0$ is rejected when $z_{statistical} < z_{\left(\frac{\alpha}{2}\right)}$ or $z_{\left(1-\frac{\alpha}{2}\right)} < z_{statistical}$. With $\alpha = 0{,}05$, and $z_{critical} = 1{,}96$ or $z_{critical} = -1{,}96$, we have a summary of the hypothesis test described in **Table 2**. The Python program is used to process the data.

Table 2. Summary of Hypothesis Testing

Cases	Parameter $(\boldsymbol{\theta})$ and $(\ddot{\mu}, \ddot{\sigma}^2)$	Data	$z_{statistical}$	Decision	Learning model changes	Parameter $(\ddot{\mu}_l, \ddot{\sigma}_l^2)$	Error $(\boldsymbol{\varepsilon})$
1	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{25}	$-0{,}222$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}24}$	0,0139
2	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{26}	$-0{,}224$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}25}$	0,0122
3	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{27}	$-0{,}332$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}26}$	0,0211
4	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{28}	$-0{,}475$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}27}$	0,0133
5	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{29}	$-0{,}999$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}28}$	0,0111
6	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{30}	$-0{,}923$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}29}$	0,0241
7	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{31}	$-0{,}803$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}30}$	0,0258
8	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{32}	$-1{,}253$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}31}$	0,0148
9	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{33}	$-1{,}365$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}32}$	0,0239
10	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{34}	$-1{,}671$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}33}$	0,0133
11	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{35}	$-1{,}762$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}34}$	0,0229
12	$T_{1\text{-}24}$	$T_{1\text{-}24}$ vs T_{36}	$-2{,}113$	$\mathbf{H_0}$ is rejected	yes	$D_{1\text{-}35}$	0,0317
13	$T_{1\text{-}36}$	$T_{1\text{-}36}$ vs T_{37}	$-0{,}009$	$\mathbf{H_0}$ is not rejected, with updating parameter $(\boldsymbol{\theta})$ and $(\ddot{\mu}, \ddot{\sigma}^2)$	no	$D_{1\text{-}36}$	0,0194
14	$T_{1\text{-}36}$	$T_{1\text{-}36}$ vs T_{38}	$-0{,}151$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}37}$	0,0210
15	$T_{1\text{-}36}$	$T_{1\text{-}36}$ vs T_{39}	$-0{,}339$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}38}$	0,0240
16	$T_{1\text{-}36}$	$T_{1\text{-}36}$ vs T_{40}	$-0{,}733$	$\mathbf{H_0}$ is not rejected	no	$D_{1\text{-}39}$	0,0202

T_{25} *is the 25th observation data, and* $T_{1\text{-}24}$ *is the 1st to 24th observation data.*

Table 2 shows the conditions for obtaining grades, namely data on the latest learning outcomes for each case in lower-level statistics courses. The latest results reference learning models used in higher-level statistics courses. In the 12th case, if no changes are made to the learning model (ignoring that H_0 is rejected), the last learning model will no longer be effective. It happens because the data obtained in the form of statistical values used are no longer suitable if the last learning model is still applied. The result is that there is the same treatment in lectures, which may result in small grades in higher-level statistics courses.

Then, it is done using MAPE in **Equation (15)** to evaluate the prediction results from the predictive distribution model. The results of the MAPE calculations are described in **Table 3**. Based on the MAPE value criteria in **Table 1**, the MAPE value between the observations and the expected value compared to the new data in **Table 3** indicates that the model used is accurate because the MAPE value is below 10%.

Table 3. MAPE Score

Observation	MAPE score	Interpretation
$P_{1\text{-}24}$ vs T_{25}	0,75%	Very accurate prediction
$P_{1\text{-}24}$ vs T_{26}	0,95%	Very accurate prediction
$P_{1\text{-}24}$ vs T_{27}	1,23%	Very accurate prediction
$P_{1\text{-}24}$ vs T_{28}	1,54%	Very accurate prediction
$P_{1\text{-}24}$ vs T_{29}	2,92%	Very accurate prediction
$P_{1\text{-}24}$ vs T_{30}	2,48%	Very accurate prediction
$P_{1\text{-}24}$ vs T_{31}	3,53%	Very accurate prediction
$P_{1\text{-}24}$ vs T_{32}	2,77%	Very accurate prediction
$P_{1\text{-}24}$ vs T_{33}	8,72%	Very accurate prediction
$P_{1\text{-}24}$ vs T_{34}	8,18%	Very accurate prediction
$P_{1\text{-}24}$ vs T_{35}	6,18%	Very accurate prediction
$P_{1\text{-}36}$ vs T_{37}	5,13%	Very accurate prediction
$P_{1\text{-}36}$ vs T_{38}	5,67%	Very accurate prediction
$P_{1\text{-}36}$ vs T_{39}	3,25%	Very accurate prediction
$P_{1\text{-}36}$ vs T_{40}	4,85%	Very accurate prediction

T_{25} *is the 25th observation data, and* $P_{1\text{-}24}$ is *expected value* of *the 1st to 24th observation data*.

The predictive distribution model for determining learning models has not been carried out by other researchers. The predictive distribution method for decision-making in determining learning models in statistics lectures is an implementation of previous research conducted by Indratno et al. [9]. The difference lies in the implementation of different case studies. They [9] discussed a case study about item delivery with a networking system and in the case of online decisions. The procedures Indratno et al. [9] recommended also apply to this study. The results can be seen in **Table 3**; the model used is accurate.

For further research, we will use clustering techniques to determine the level of strategic competence of students in other courses as described in Yudhanegara and Lestari [12], Yudhanegara et al. [16], [17], [18]. In addition, we will also combine research using correspondence analysis for categorical data [19], a combination of correspondence analysis with clustering [17] [19] [20] [21], and a regression model [22] [23].

4. CONCLUSIONS

Based on the simulation results and discussion, the use of a predictive distribution approach to predict dynamic data is very appropriate. It is measured through the MAPE score with the overall prediction results are very accurate. In addition, we can also analyze the errors generated from the testing algorithm, which are generally small error values when $\mathbf{H_0}$ is not rejected.

The limitation of the research, i.e., predictive distribution, is built based on data under the assumption of a multinomial distribution with independent and identically distributed. From the research result, we can use a distribution approach to detect changes in new data distribution with historical data for the dynamic condition. If the changes are insignificant, direct instruction can still be used for the learning model. However, if the changes are significant, it is recommended to use a new learning model for the statistics group course at a higher level.

ACKNOWLEDGMENT

This research was supported by LPPM Universitas Singaperbangsa Karawang through the HIPKA Universitas Singaperbangsa Karawang grant. The authors thank the research assistant team, Mr. Ikhsan Baihaqi, Mr. Hamdhani, and Mrs. Dini Gustiani.

REFERENCES

[1] R. V. Hogg, E. A. Tanis, and D. L. Zimmerman, *Acceptability and statistical inference*, vol. 61, no. 2. New York: Pearson, 1974.

[2] P. Jones, M. Wynn, D. Hillier, and D. Comfort, "The Sustainable Development Goals and Information and Communication Technologies," *Indones. J. Sustain. Account. Manag.*, vol. 1, no. 1, p. 1, 2017, doi: 10.28992/ijsam.v1i1.22.

[3] S. Fukuda-parr, "CDP Policy Review Series Sustainable Development Goals (SDGs) and the Promise of a Transformative Agenda," *CDP Policy Rev.*, vol. 2022, no. 11, pp. 1–14, 2022.

[4] C. Kroll, A. Warchold, and P. Pradhan, "Sustainable Development Goals (SDGs): Are we successful in turning trade-offs into synergies?," *Palgrave Commun.*, vol. 5, no. 1, pp. 1–11, 2019, doi: 10.1057/s41599-019-0335-5.

[5] UNECE, *Conference of European Statisticians' Road Map on Statistics for Sustainable Development Goals*, no. June. New York: United Nations, 2017.

[6] S. J. Kilpatrick Jeremy, *Adding It Up: Helping Children Learn Mathematics*. Washington DC: National Academy Press, 2005.

[7] K. E. Lestari, M. R. Utami, and M. R. Yudhanegara, "Exploratory Analysis on Adaptive Reasoning of Undergraduate Student in Statistical Inference," *Int. J. Intruction*, vol. 15, no. 4, pp. 535–554, 2022, doi: 10.4135/9780857020123.n561.

[8] M. R. Yudhanegara, S. W. Indratno, and R. K. N. Sari, "Dynamic items delivery network: prediction and clustering," *Heliyon*, vol. 7, no. 5, p. e06934, 2021, doi: 10.1016/j.heliyon.2021.e06934.

[9] S. W. Indratno, K. N. Sari, and M. R. Yudhanegara, "Optimization in Item Delivery as Risk Management: Multinomial Case Using the New Method of Statistical Inference for Online Decision," *Risks*, vol. 10, no. 6, p. 122, 2022, doi: 10.3390/risks10060122.

[10] S. A. Klugman, H. H. Panjer, and G. E. Willmot, *Loss Models: From Data to Decisions, 5th Edition*. 2019.

[11] R. V. Hogg and A. T. Craig, *Introduction to Mathematical and Statistics.*, vol. 67, no. 7. Pearson, 2019.

[12] M. R. Yudhanegara and K. E. Lestari, "Clustering for multi-dimensional data set: A case study on educational data," *J. Phys. Conf. Ser.*, vol. 1280, no. 4, 2019, doi: 10.1088/1742-6596/1280/4/042025.

[13] J. Oosterhoff, "The Likelihood Ratio Test for the Multinomial Distribution," *Theory Stat.*, pp. 31–50, 2020, doi: 10.1525/9780520325883-004.

[14] L. A. B. Owen, "Lecture 10 : Generalized likelihood ratio test Generalized likelihood ratio test Multinomial," vol. 2018, pp. 1–5, 2018.

[15] C. D. Lewis, *Industrial and Business Forecasting Methods*. London: Butterworths, 1982.

[16] M. R. Yudhanegara, S. W. Indratno, and R. R. K. N. Sari, "Clustering for Items Distribution Network," *J. Phys. Conf. Ser.*, vol. 1496, no. 1, 2020, doi: 10.1088/1742-6596/1496/1/012019.

[17] M. R. Yudhanegara, S. W. Indratno, and R. K. N. Sari, "Clustering for Item Delivery Using Rule-K-Means," *J. Indones. Math. Soc.*, vol. 26, no. 2, pp. 185–191, 2020, doi: 10.22342/jims.26.2.871.185-191.

[18] M. R. Yudhanegara, S. W. Indratno, and K. N. Sari, "Role of clustering method in items delivery optimization," *J. Phys. Conf. Ser.*, vol. 2084, no. 1, 2021, doi: 10.1088/1742-6596/2084/1/012011.

[19] K. E. Lestari, U. S. Pasaribu, S. W. Indratno, and H. Garminia, "Generating roots of cubic polynomials by Cardano's approach on correspondence analysis," *Heliyon*, vol. 6, no. 6, p. e03998, 2020, doi: 10.1016/j.heliyon.2020.e03998.

[20] K. E. Lestari, M. R. Utami, and M. R. Yudhanegara, "Sequential Exploratory Design by Performing Correspondence Analysis to Investigate Procedural Fluency of Undergraduate Student," *AIP Conf. Proc.*, vol. 2588, no. January, 2023, doi: 10.1063/5.0111974.

[21] K. E. Lestari, M. R. Utami, and M. R. Yudhanegara, "Simple Algorithm To Construct Circular Confidence Regions in Correspondence Analysis Using R," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 16, no. 1, pp. 065–074, 2022, doi: 10.30598/barekengvol16iss1pp065-074.

[22] S. Vantika, M. R. Yudhanegara, and K. E. Lestari, "Poisson Regression Modelling of Automobile Insurance Using R," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 16, no. 4, pp. 1399–1410, 2022, doi: 10.30598/barekengvol16iss4pp1399-1410.

[23] I. Syahzaqi, S. A. P. A., and K. Fithriasari, "Modelling of Poverty Percentage in East Java Province With Semiparametric Regression Approach," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 17, no. 2, pp. 0727–0734, 2023, doi: 10.30598/barekengvol17iss2pp0727-0734.