Journal of Mathematics and Mathematics Education
https://jurnal.uns.ac.id/jmme
p-ISSN 2089-8878; e-ISSN 2715-8276
Vol. 14, No. , june 2024

NORMAL DISTRIBUTION APPROXIMATION THROUGH NEGATIVE BINOMIAL DISTRIBUTION

Citra Tiara Ningati[1,*]
[1]Mathematics, Faculty of Teacher Training and Education, Sebelas Maret University, Surakarta, Indonesia

* Correspondence purposes, email: citratiara03@student.uns.ac.id

Abstrak: Tujuan dari penelitian ini adalah untuk membuktikan pendekatan distribusi normal melalui distribusi binomial negatif secara teoritis dan secara simulasi menggunakan *software minitab*. Penelitian ini menggunakan metode studi literatur dengan mengumpulkan berbagai data terkait literatur distribusi normal dan distribusi binomial negatif. Secara teoritis, pembuktian pendekatan distribusi normal melalui distribusi binomial negatif menggunakan pendekatan metode pembangkit momen, dimana distribusi binomial negatif akan mendekati distribusi normal ketika n menuju tak hingga. Sedangkan secara simulasi menggunakan *software minitab* dapat dilihat pada poligon distribusi binomial negatif dengan n semakin besar mendekati tak hingga maka akan semakin mendekati kurva normal.
Kata kunci : *Distribusi normal, distribusi binomial negatif, metode pembangkit momen*

Abstract: The purpose of this study is to prove that the normal distribution approach through binomial distributions is theoretically negative and simulatively uses minitab software. This study used a method of literature study by collecting various data related to the literature of normal distribution and negative binomial distribution. Theoretically proving the normal distribution approach via a negative binomial distribution using the approach of the moment generating method, where the negative binomial distribution approaches the normal distribution as it goes to infinity. Simulations using the minitab software can be seen in negative binomial distribution polygons with the larger approaching infinity and the closer the normal curve.
Keywords: *Normal distribution, negative binomial distribution, moment generating method*

INTRODUCTION

Mathematics is a science in everyday life that is considered difficult and definite. One very important branch of mathematics that needs to be mastered optimally is the science of statistics. In fact, there are so many problems in life that need solutions with proper analysis that cannot be separated from the use of statistics such as in the world of lectures that are used to calculate research on a particular distribution, calculation on scripting research. calculations at math seminars, etc. There are many statistical uses that can be realized in real life.

Accepted: december 28, 2023 Approved: June 01, 2024 Published: June 31, 2024
DOI: https://doi.org/10.20961/jmme.v14i1.86761

Statistics is the branch of mathematics that deals with the collection, analysis, interpretation, presentation, and organization of data. According to Budiyono (2009), statistics is knowledge related to how data is prepared, data presentation, and conclusions are drawn about a whole called population based on the data that exists in a part of the whole, part of the whole (population) called a sample. According to Supranto (2008), statistics is a science that studies how to collect, process grouping, presentation, and analysis data and conclusions by taking into account uncertainty elements.

Statistics have important materials for calculating an experiment that is the basis of probability theory that has the most important material object, the probability distribution or probability distribution. In the opportunity distribution, there is a discrete opportunity distribution and a continuous opportunity distribution. There are several kinds of distributions on discrete opportunity distributions such as the poisson distribution, binomial distribution, negative binomial distribution, and others. Then, for some of the distributions that exist in continuous distributions include normal distributions, exponential distributions, and others.

In its application the frequently used distribution is the binomial distribution and the normal distribution. However, it is rare to use a negative binomial distribution, so one of the discrete opportunity distributions to be used in this study is a negative binomial distribution while the type of continuous opportunity distribution to be used is a normal distribution.

The binomial distribution is a theoretical probability distribution with discrete random variables of mutually independent numbers of successful n experiments, with each experimental result generating opportunities. In statistics, the binomial distribution becomes the basis of the binomial test in statistical significance testing. This distribution is often used to model the number of successes x times in n experiments. A random variable X with a binomial distribution denoted by $B(x,n,\theta)$ is defined by:

$$B(x,n,\theta) = f(x) = \begin{cases} \binom{n}{x}\,\theta^{x}(1-\theta)^{n} - x \,, x = 1,2,3,\ldots n \\ 0 \qquad\qquad\qquad\qquad , x\ other \end{cases}$$

n and θ are parameters,

Description:

x = number of successful events,

n = number of trials,

θ = chance of success in a single experiment.

The probability of the expected outcome of an experiment as much as X out of n experiments is

$$P(X=x)=\binom{n}{x}\theta^{x}(1-\theta)^{n-x}$$

Theorem 1.1

The random variable X is binomially distributed with parameters n and θ, hence:

1. Moment Generator Function
   $M_x(t)=(\theta e^{t}+(1-\theta))^{n}$
2. Average
   $\mu=n\theta$
3. Variance
   $\sigma^{2}=n\theta(1-\theta)$

As the population of N increases, normal distributions and binomial distributions are often used, so it is rare to use negative binomial distributions. A negative binomial distribution is a distribution of opportunities related to the number of experiments that are repeated several times freely to achieve the $-n$ th success. This illustrates the number of bernoulli experiments used to obtain positive results with the characteristic parameters of the negative binomial distribution of n and p with n as the number of success results desired in an experiment while p as the chance of success for each experiment to be performed. If population N gets larger often using normal distributions and rarely anyone uses negative binomial distributions, it is rarely found in books that demonstrate this.

If a randomizer X specifies the number of trials needed until a successful r occurs, then the probability distribution of the randomizer X with a probability function follows. A discrete random variable X on a Negative Binomial distribution denoted by $BN\sim(x,r,p)$ can be defined by:

$$P(X=x)=f(x,r,p)=\begin{cases}\binom{x-1}{r-1}p^{r}(1-p)^{x-r}, & x=r,r+1,r+2,..\\ 0\qquad\qquad\qquad\qquad\quad , & x\ other\end{cases}$$

The probability distribution of a random modifier X can be denoted using the $Y=X-r$ transform with a probability function as follows. A discrete Y random variable on a Negative Binomial distribution denoted by $BN\sim(y,r,p)$ can be defined by:

$$P(Y=y)=f(y,r,p)=\begin{cases}\binom{y+r-1}{r-1}p^{r}(1-p)^{y}, & y=0,1,2,\ldots\\ 0\qquad\qquad\qquad\qquad\quad , & y\ other\end{cases}$$

r and p are parameters,

Description:

p = chance of success,

r = number of successful to -r,

x, y = number of attempts until successful.

The $Y = X - r$ transformation is obtained from $P(A \cap B) = P(A).P(B)$

With this information:

$A = \{experiment\ (x-1)\ at\ (r-1)\ succes\ \}$

$B = \{succes\ on\ experiment\ x\ \}$

$P(A) = \binom{x-1}{r-1} p^{r-1}$ and $P(B) = (1-p)^{(x-1)-(r-1)}$

Then, $P(A \cap B) = P(A).P(B)$

$$= \binom{x-1}{r-1} p^{r-1}\ (1-p)^{(x-1)-(r-1)}$$

$$= \binom{x-1}{r-1} p^{r-1}(1-p)^{(x-r)}.p$$

$$= \binom{x-1}{r-1} p^{r}(1-p)^{(x-r)}$$

$$P(X = x) = \binom{x-1}{r-1} p^{r}(1-p)^{(x-r)}$$

The probability distribution of a random modifier X can be denoted into another form by using the $Y = X - r$ transform so that $X = Y + r$ with Y expresses the number of failures before a successful r occurs. The probability distribution of a Y-randomizer is as follows:

$$P(X = x) = \binom{x-1}{r-1} p^{r}(1-p)^{(x-r)}$$

$A\ Random\ Variable\ Y = X - r$

$P(Y = y)\ \ = P(X - r = y)$

$$= P(X = y + r)$$

$$= \binom{y + r - 1}{r - 1} p^r (1 - p)^y$$

With $y = 0,1,2,3, .. and\ 0 \leq p \leq 1$

Description:

y = number of failures r successful

X = $y + r$

Thus obtained for the chance that the expected result of an experimental negative binomial distribution is

$$P(X = x) = \binom{x - 1}{r - 1} p^r (1 - p)^{(x-r)}$$

or

$$P(Y = y) = \binom{y + r - 1}{r - 1} p^r (1 - p)^y$$

Theorem 1.2

The random variable X is a negative binomial distribution with parameters r and p, hence:

1. Moment Generator Function

   $M_{x(t)} = \left(\frac{pe^t}{1-(1-p)e^t}\right)^r$
2. Average

   $\mu = \frac{r}{p}$
3. Variance

   $\sigma^2 = \frac{r(1-p)}{p^2}$

The normal distribution is the most important distribution in the field of statistics. A normal distribution is a basic distribution often used to denote the distribution of a variable. The normal distribution has a symmetrical bell-shaped graph formed with infinite n and the equation was first discovered in 1733 by Abraham DeMoivre. Normal distributions play an important role as a reference in drawing conclusions based on the results of the sample. In addition, the normal distribution also plays a role in determining the average sampling distribution that will be close to normal or in other words all natural events will form a normal distribution.

A random variable X can be said to be normally distributed with parameters μ and σ^2 denoted by $X \sim N(\mu, \sigma^2)$ if the density function is:

$$X \sim N(\mu, \sigma^2) = f(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, -\infty < x < \infty \\ 0 \qquad\qquad , \ x \ other \end{cases}$$

Description:

μ = average of population

σ^2 = variance of population

σ^2 = standard deviation of the population

$e \approx 2.718$ a constant

$\pi \approx 3.141$ a constant

x = value of continuous random variables that satisfy $-\infty < x < \infty$

If the normal distributed random X variables are standardized with $\mu = 0$ and $\sigma^2 = 1$, then they can be denoted by $X \sim N(0,1)$ which is usually also called the Z distribution. The standard normal distribution density function forms as follows.

$$X \sim N(0,1) = f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x)^2}, -\infty < x < \infty \\ 0 \qquad\qquad , x \ other \end{cases}$$

Theorem 1.3

The area bounded by the normal curve and the X-axis is 1 unit. This theorem can be denoted in the form of an unnatural integral, as follows

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Theorem 1.4

If the continuous random variable X has a normal distribution of $N(\mu, \sigma^2)$, then:

1. Moment Generator Function

   $M_x(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$
2. Average

   $\mu = \mu$
3. Variance

   $\sigma^2 = \sigma^2$

In practice statistics is not only to prove a theorem of distributions on statistics and work on problems of a theorem or statistical data but also requires implementation in processing data from a distribution with the help of Minitab's software. which is software used to help process statistical data. This study aims to prove the normal distribution approach through a theoretically negative binomial distribution and simulatively using minibab software. In addition, this study used a method of literature study by collecting various data related to the literature of the normal distribution and the negative binomial distribution. From the above background description, the author takes the title, "Normal Distribution Approximation Through Negative Binomial Distribution".

RESEARCH METHOD

This study is a study using a literature study method in which researchers use various research data collection, reading, and processing data in steps according to the procedure used. The literature used in this study is related to approximating the normal distribution through a negative binomial distribution. Later, researchers studied distribution materials that began with definitions and theorems of collected literature on normal distributions and negative binomial distributions. Then, it proceeds by proving the normal distribution approach through a negative binomial distribution theoretically and simulatively by using minitab software and terminated by drawing conclusions related to research obtained on the normal distribution approach through a negative binomial distribution.

RESULTS AND DISCUSSION

Proof of Normal Distribution Approach by Theoretically Negative Binomial Distribution

A random variable X with a negative binomial distribution is known

$$P(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, r+2, \dots.$$

or

$$P(Y = y) = \binom{y+r-1}{r-1} p^r (1-p)^y, y = 0,1,2, \dots.$$

Thus, according to the above theorem, a negative binomial function has a moment-generating function $M_x(t) = E[e^{tx}] = \left(\frac{pe^t}{(1-(1-p)e^t)}\right)^r$. To approach the normal distribution, it is necessary to transform the random variable X into the standard normal distributed random variable Z with $Z = \frac{X-\mu}{\sigma}$. It is known that the binomial distribution has an average of $\mu = \frac{r}{p}$ and variance $\sigma^2 = \frac{n(1-p)}{p^2} \rightarrow \sigma = \frac{\sqrt{r(1-p)}}{p}$, then $Z = \frac{x - \frac{r}{p}}{\frac{\sqrt{r(1-p)}}{p}} = \frac{xp-r}{\sqrt{r(1-p)}}$ is the default normal distributed random variable.

Proof of approximating the normal distribution through a negative binomial distribution can be done using the method of the moment generating function by giving the limit to infinity and the random variable Z to the moment generating function of the negative binomial distribution. The moment generating function of a negative binomial distribution with a random variable Z can be seen as follows,

$$\begin{aligned}
\lim_{r\to\infty} M_z(t) &= \lim_{r\to\infty} [e^{tz}] \\
&= \lim_{r\to\infty} E\left[e^{t\left(\frac{xp-r)}{\sqrt{r(1-p)}}\right)}\right] \\
&= \lim_{r\to\infty} E\left[e^{t\left(\frac{xp}{\sqrt{r(1-p)}}\right)-t\left(\frac{r}{\sqrt{r(1-p)}}\right)}\right] \\
&= \lim_{r\to\infty} E\left[e^{t\left(\frac{xp}{\sqrt{r(1-p)}}\right)} e^{-t\left(\frac{r}{\sqrt{r(1-p)}}\right)}\right]
\end{aligned}$$

By noticing that $e^{-t\left(\frac{r}{\sqrt{r(1-p)}}\right)}$ is a function of constant value. Based on a theorem on mathematical expectations that says that $E[cX] = cE[X]$, then it's earned

$$\lim_{r\to\infty} E\left[e^{t\left(\frac{xp}{\sqrt{r(1-p)}}\right)} e^{-t\left(\frac{r}{\sqrt{r(1-p)}}\right)}\right] = \lim_{r\to\infty} e^{-t\left(\frac{r}{\sqrt{r(1-p)}}\right)} E\left[e^{t\left(\frac{xp}{\sqrt{r(1-p)}}\right)}\right].$$

Watch for $E\left[e^{t\left(\frac{xp}{\sqrt{r(1-p)}}\right)}\right]$ is a negative binomial distribution moment generating function with a random variable X, hence $E\left[e^{t\left(\frac{xp}{\sqrt{r(1-p)}}\right)}\right] = E\left[e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)x}\right] = \left(\frac{pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{1-(1-p)e^{\frac{pt}{\sqrt{r(1-p)}}}}\right)^{r}$

$$E\left[e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)x}\right] = \left(\frac{1}{pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}} - \left(\frac{(1-p)e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}\right)\right)^{-r}$$

Continue with the use of $\lim_{r\to\infty} e^{-t\left(\frac{r}{\sqrt{r(1-p)}}\right)} E\left[e^{t\left(\frac{xp}{\sqrt{r(1-p)}}\right)}\right]$ obtained,

$$= \lim_{r\to\infty} e^{-t\left(\frac{r}{\sqrt{r(1-p)}}\right)} \left(\frac{pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{1-(1-p)e^{\frac{pt}{\sqrt{r(1-p)}}}}\right)^{r}$$

$$= \lim_{r\to\infty} e^{-t\left(\frac{r}{\sqrt{r(1-p)}}\right)} \left(\frac{1}{pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}} - \left(\frac{(1-p)e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}\right)\right)^{-r}$$

$$= \lim_{r\to\infty} e^{-t\left(\frac{r}{\sqrt{r(1-p)}}\right)} \left(\frac{1}{pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}} - \left(\frac{(1-p)}{p}\right)\right)^{-r}$$

$$= \lim_{r\to\infty} e^{-t\left(\frac{r}{\sqrt{r(1-p)}}\right)} \left(\frac{1- e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)} + pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}\right)^{-r}$$

$$= \lim_{r\to\infty} \frac{e^{-t\left(\frac{r}{\sqrt{r(1-p)}}\right)}}{\left(e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}\right)^{-r}} \left(\frac{1- e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)} + pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{p}\right)^{-r}$$

$$= \lim_{r\to\infty} \frac{e^{\left(\frac{-tr}{\sqrt{r(1-p)}}\right)}}{e^{\left(\frac{-rpt}{\sqrt{r(1-p)}}\right)}} \left(\frac{1- e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)} + pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{p}\right)^{-r}$$

$$= \lim_{r\to\infty} e^{\left(\frac{-tr-(-rpt)}{\sqrt{r(1-p)}}\right)} \left(\frac{1- e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)} + pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{p} \right)^{-r}$$

$$= \lim_{r\to\infty} e^{\left(\frac{-tr(1-p)}{\sqrt{r(1-p)}}\right)} \left(\frac{1- e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)} + pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{p} \right)^{-r}$$

$$= \lim_{r\to\infty} \left(e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} \right)^{-r} \left(\frac{1- e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)} + pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{p} \right)^{-r}$$

$$= \lim_{r\to\infty} \left(\frac{e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} - e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} e^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)} + e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} pe^{\left(\frac{pt}{\sqrt{r(1-p)}}\right)}}{p} \right)^{-r}$$

$$= \lim_{r\to\infty} \left(\frac{e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} - e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)} + pe^{\left(\frac{t}{\sqrt{r(1-p)}}\right)}}{p} \right)^{-r}$$

$$= \lim_{r\to\infty} \left(\frac{e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} - e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)}(1-p)}{p} \right)^{-r}$$

$$= \lim_{r\to\infty} \left(\frac{1}{p}\, e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} - \frac{1-p}{p} e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)} \right)^{-r}$$

Thus, based on the taylor series theorem, then note for $e^{t\left(\frac{1-p}{\sqrt{r(1-p)}}\right)}$ can be changed to

$$f(t) = e^{t\left(\frac{1-p}{\sqrt{r(1-p)}}\right)} \to f(0) = e^0 = 1$$

$$f'^{(t)} = \left(\frac{1-p}{\sqrt{r(1-p)}}\right) e^{t\left(\frac{1-p}{\sqrt{r(1-p)}}\right)} \to f'(0) = \frac{1-p}{\sqrt{r(1-p)}}$$

$$f''(t) = \left(\frac{1-p}{\sqrt{r(1-p)}}\right)^2 e^{t\left(\frac{1-p}{\sqrt{r(1-p)}}\right)} \to f''(0) = \left(\frac{1-p}{\sqrt{r(1-p)}}\right)^2$$

$$f'''(t) = \left(\frac{1-p}{\sqrt{r(1-p)}}\right)^3 e^{t\left(\frac{1-p}{\sqrt{r(1-p)}}\right)} \to f'''(0) = \left(\frac{1-p}{\sqrt{r(1-p)}}\right)^3$$

…

$$f^{(n)}(t) = \left(\frac{1-p}{\sqrt{r(1-p)}}\right)^n e^{t\left(\frac{1-p}{\sqrt{r(1-p)}}\right)} \to f^{(n)}(t) = \left(\frac{1-p}{\sqrt{r(1-p)}}\right)^n$$

So that it gets,

$$f(t) = f(0) + f'(0)(t-o) + \frac{f''(0)}{2!}(t-0)^2 + \frac{f'''(0)}{3!}(t-0)^3 + \cdots + \frac{f^{(n)}(0)}{n!}(t-0)^2$$

$$e^{t\left(\frac{1-p}{\sqrt{r(1-p)}}\right)} = 1 + \frac{1-p}{\sqrt{r(1-p)}}(t) + \frac{\left(\frac{1-p}{\sqrt{r(1-p)}}\right)^2}{2!}(t)^2 + \frac{\left(\frac{1-p}{\sqrt{r(1-p)}}\right)^3}{3!}(t)^3 + \cdots + \frac{\left(\frac{1-p}{\sqrt{r(1-p)}}\right)^n}{n!}(t)^n$$

$$= 1 + \frac{(1-p)t}{\sqrt{r(1-p)}} + \frac{(1-p)^2 t^2}{2!r(1-p)} + \frac{(1-p)^3 t^3}{3!(r(1-p))^{\frac{3}{2}}} + \cdots + \frac{\left(\frac{1-p}{\sqrt{r(1-p)}}\right)^n}{n!\left(\sqrt{r(1-p)}\right)^n}(t)^n$$

Then attention is also to $e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)}$ can be changed to

$$f(t) = e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)} \to f(0) = e^0 = 1$$

$$f'(t) = \left(\frac{1}{\sqrt{r(1-p)}}\right) e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)} \to f'(0) = \frac{1}{\sqrt{r(1-p)}}$$

$$f''(t) = \left(\frac{1}{\sqrt{r(1-p)}}\right)^2 e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)} \to f''(0) = \left(\frac{1}{\sqrt{r(1-p)}}\right)^2$$

$$f'''(t) = \left(\frac{1}{\sqrt{r(1-p)}}\right)^3 e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)} \to f'''(0) = \left(\frac{1}{\sqrt{r(1-p)}}\right)^3$$

...

$$f^{(n)}(t) = \left(\frac{1}{\sqrt{r(1-p)}}\right)^n e^{t\left(\frac{t}{\sqrt{r(1-p)}}\right)} \to f^{(n)}(t) = \left(\frac{1}{\sqrt{r(1-p)}}\right)^n$$

So that it gets,

$$f(t) = f(0) + f'(0)(t-o) + \frac{f''(0)}{2!}(t-0)^2 + \frac{f'''(0)}{3!}(t-0)^3 + \cdots + \frac{f^{(n)}(0)}{n!}(t-0)^n$$

$$e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)} = 1 + \frac{1}{\sqrt{r(1-p)}}(t) + \frac{\left(\frac{1}{\sqrt{r(1-p)}}\right)^2}{2!}(t)^2 + \frac{\left(\frac{1}{\sqrt{r(1-p)}}\right)^3}{3!}(t)^3 + \cdots + \frac{\left(\frac{1}{\sqrt{r(1-p)}}\right)^n}{n!}(t)^n$$

$$e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)} = 1 + \frac{t}{\sqrt{r(1-p)}} + \frac{t^2}{2!r(1-p)} + \frac{t^3}{3!(r(1-p))^{\frac{3}{2}}} + \cdots + \frac{(1)^n}{n!(r(1-p))^n}(t)^n$$

So that's for $\frac{1}{p}\ e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)}$ obtained

$$\frac{1}{p}\, e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} = \frac{1}{p}\left(1 + \frac{(1-p)t}{\sqrt{r(1-p)}} + \frac{(1-p)^2t^2}{2!r(1-p)} + \frac{(1-p)^3t^3}{3!\left(r(1-p)\right)^{\frac{3}{2}}} + \cdots + \frac{\left(\frac{1-p}{\sqrt{r(1-p)}}\right)^n}{n!}(t)^n\right)$$

So that's for $\frac{1-p}{p}\, e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)}$ obtained

$$\frac{1-p}{p} e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)} = \frac{1-p}{p}\left(1 + \frac{t}{\sqrt{r(1-p)}} + \frac{t^2}{2!r(1-p)} + \frac{t^3}{3!\left(r(1-p)\right)^{\frac{3}{2}}} + \cdots + \frac{\left(\frac{1}{\sqrt{r(1-p)}}\right)^n}{n!}(t)^n\right)$$

So that's for $\frac{1}{p}\, e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} - \frac{1-p}{p} e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)}$ obtained

$$\frac{1}{p}\, e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} - \frac{1-p}{p} e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)} = \frac{1}{p}\left(1 + \frac{(1-p)t}{\sqrt{r(1-p)}} + \frac{(1-p)^2t^2}{2!r(1-p)} + \frac{(1-p)^3t^3}{3!\left(r(1-p)\right)^{\frac{3}{2}}} + \cdots + \right.$$

$$\left.\frac{\left(\frac{1-p}{\sqrt{r(1-p)}}\right)^n}{n!}(t)^n\right) - \frac{1-p}{p}\left(1 + \frac{t}{\sqrt{r(1-p)}} + \frac{t^2}{2!r(1-p)} + \frac{t^3}{3!\left(r(1-p)\right)^{\frac{3}{2}}} + \cdots + \frac{\left(\frac{1}{\sqrt{r(1-p)}}\right)^n}{n!}(t)^n\right)$$

$$= \left(\frac{1}{p} + \frac{1-p}{p}\right) - \frac{t^2}{2r}\left(\frac{1}{p} - \frac{1-p}{p}\right) + \frac{t^3\left((1-p)^3-1\right)}{3!\left(r(1-p)\right)^{\frac{3}{2}}} + \cdots$$

Next acquired,

$$\lim_{r\to\infty}\left(\frac{1}{p}\, e^{\left(\frac{t(1-p)}{\sqrt{r(1-p)}}\right)} - \frac{1-p}{p} e^{\left(\frac{t}{\sqrt{r(1-p)}}\right)}\right)^{-r}$$

$$= \lim_{r\to\infty}\left(\left(\frac{1}{p} + \frac{1-p}{p}\right) - \frac{t^2}{2r}\left(\frac{1}{p} - \frac{1-p}{p}\right) + \frac{t^3\left((1-p)^3-1\right)}{3!\left(r(1-p)\right)^{\frac{3}{2}}} + \cdots\right)^{-r}$$

For example $\psi(r) = \frac{t^3\left((1-p)^3-1\right)}{3!\left(r(1-p)\right)^{\frac{3}{2}}} + \cdots$ and $\{\psi(n), n \geq 1\}$ is a sequence of real numbers.

So it's earned, $\lim_{r\to\infty}\left(1 - \left(\frac{t^2}{2!(r)}\right) + \psi(r)\right)^{-r}$.

From the limit of the sequence of real numbers above, there is a lemma e.g. "$\{\psi(n), n \geq 1\}$" is a sequence of real numbers so that $\lim_{r\to\infty} \psi(r) = 0$. Then $\lim_{r\to\infty}\left(1 + \left(\frac{a}{n}\right) + \frac{\psi(r)}{r}\right)^{bn} = e^{ab}$, with a and b not dependent on n.”.

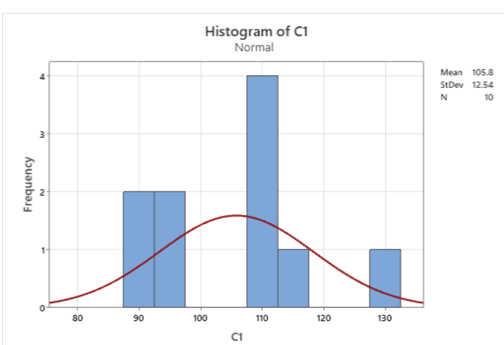
Based on the lemma above, it is obtained $\lim_{r\to\infty}\left(1 - \left(\frac{t^2}{2!(r)}\right) + \psi(r)\right)^{-r} = e^{\frac{-t^2}{2!r}(-r)} = e^{\frac{t^2}{2}}$

It has $e^{\frac{t^2}{2}}$, where $e^{\frac{t^2}{2}}$ is the moment generating function of the standard normal distribution $N(0,1)$. Thus it is proven that the negative binomial distribution will approach the normal distribution when the sum of the data is very large is close to infinity.

Proof of Normal Distribution Approach by Negative Binomial Distribution Using Minitab Software

The proof of the normal distribution approach through a negative binomial distribution can be seen through its curve. To visualize the curve of a negative binomial distributed function would be very difficult if done manually, especially if the data possessed is so much that the minitab software is not available.

In this proof, according to the previous theory that the negative binomial distribution will get closer to the normal distribution when N is very large, so it will also be compared to small to enlarged N which causes the curve to get closer to the normal curve. To prove it requires so much data that it will be difficult to collect data in real terms. Therefore, very much data can be created by generating the minitab software assisted data then compared to curves in the normal distribution.



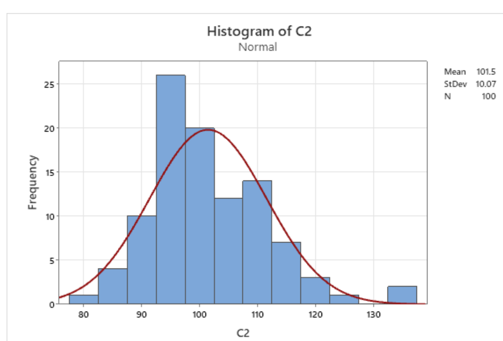


Figure 1. For $N = 10$

Figure 2. For $N = 100$

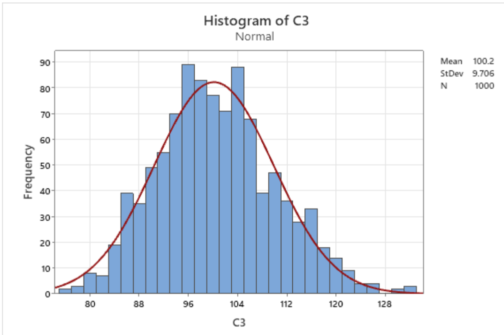


Figure 3. For $N = 1000$

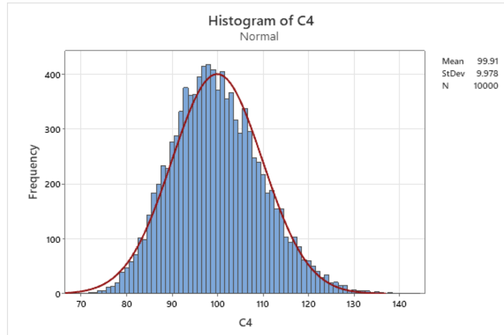


Figure 4. For $N = 10000$

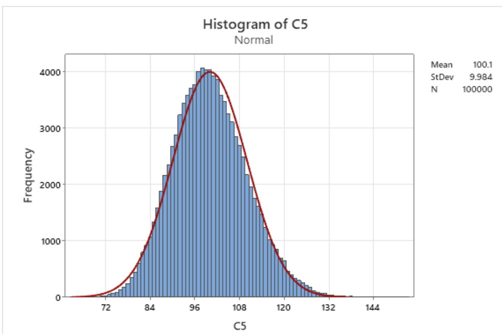


Figure 5. For $N = 100000$

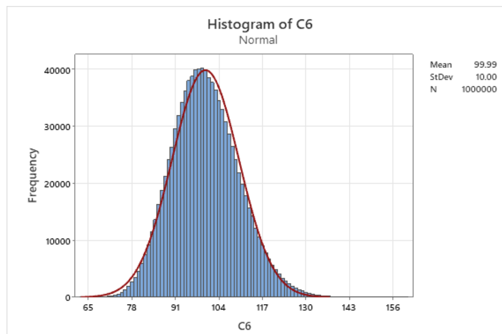


Figure 6. For $N = 1000000$

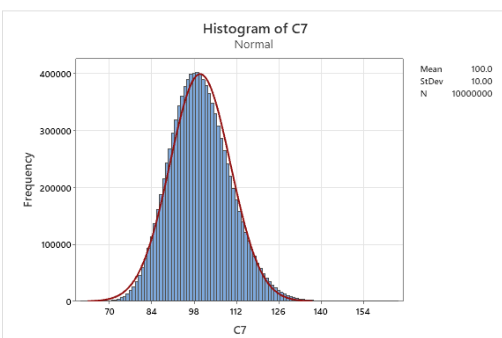


Figure 7. For $N = 10000000$

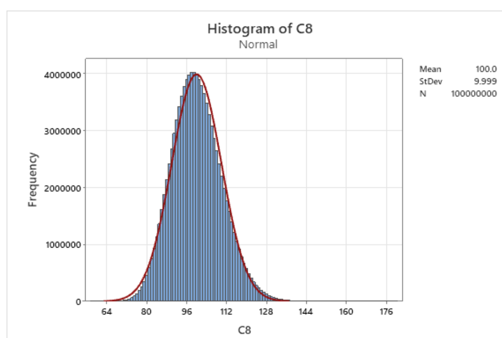


Figure 8. For $N = 100000000$

From the above polygons it can be seen that the larger the value of N with θ fixed, the closer the polygon is to the normal curve value. It is proved that a negative binomial distribution polygon will approach a normal curve when its very large data approach infinity.

CONCLUSIONS AND SUGGESTIONS

Based on the results of the discussion, it is found that normal distribution approaches through negative binomial distributions can be proved theoretically or simulated using minibab software.

Theoretically, a negative binomial distribution would approach a normal distribution by applying the limit of n approaches infinity of the moment generating function of the negative binomial distribution which would subsequently derive the moment generating function from the standard normal distribution.

Simulation using the minibab software has been done by comparing normal curves with negative binomial distribution polygons that have large amounts of data. In addition, it is also possible to visualize normal curves and polygons on negative binomial distributions with large amounts of data and fixed values by seeing that these polygons approach the normal curve when very large amounts of data approach infinity.

The work of this study deals with approximating the normal distribution through a negative binomial distribution. Further research is expected to address even more varied distribution approaches, given that there are still many specialized opportunity distributions on statistics.

ACKNOWLEDGMENTS

Researchers want to thank you for the opportunity to write this journal as well as everyone who has contributed to the preparation of this journal. The researcher also thanked our mentor professor, Henny Ekana Chrisnawati, S.Si., M.pd, Dr. Triyanto, S.Si., M.Si., my parents, Dilla Aulia, kost ijo friends, and other friends. We hope that readers can benefit from reading this journal.

REFERENCES

Bagui, S & Mehra, K, L. (2019). *On the Convergence of Negative Binomial Distribution. International Mathematical Forum*, 9(1), 44-50. https://doi.org/10.5923/j.ajms.20190901.06.

Bagui, S & Mehra, K, L. (2016). *Convergence of Binomial, Poisson, Negative-Binomial and Gamma to Normal Distribution: Moment Generating Function Technique. American Journal of Mathematics and Statistics*, 6(3), 117-118. https://doi.org/10.5923/j.ajms.20160603.05.

Bain, L, & Engelhardt. (1992). Introduction to Probability and Mathematical Statistics.

California: Wadsworth Publishing Company.

Budiyono. (2009). Statistika untuk Penelitian Edisi ke 2. Surakarta: UNS Press.

Musadi & Kurniawati, Ira. (2023). *Normal Distribution Approximation Through Binomial and Poisson Distribution. Journal of Mathematics and Mathematics Education*, Vol. 13, 109.

Purcell, Edwin J., dan Dale Verberg. (1987). Calculus with Analytic Geometry, ed 5. Terjemahan Susila, I Nyoman, dkk. Kalkulus dan Geometri Analitis. Penerbit Erlangga.

Spiegel, Murray R., John, S. dan R, Alu, S. (1975). Schaum's Outlines of Probabilitas dan Statistik Edisi Kedua. Terjemahan Refina, I. (2004). Jakarta: Penerbit Erlangga.

Supranto, J. (2008). Statistik Teori dan Aplikasi. Jakarta: Erlangga.

Sudaryono. (2012). Statistika Probabilitas - Teori & Aplikasi. Yogyakarta: ANDI.

Triyanto.(2014). Teknik Analisis Data dengan Minitab. Surakarta: FKIP UNS.

Walpole, R.E., et al. (2012). Probability & Statistics for Engineers & Scientists, 9th ed. Boston: Pearson Education, Inc.