Characteristics of Bayes Estimator in the Geometric Distribution with Prior Beta

T Susilo¹, Widiarti¹,*, D Kurniasari¹, D Aziz¹

¹Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Lampung, Indonesia, Jl. Soemantri Brojonegoro No. 1, Bandar Lampung, Indonesia.

email: widiarti08@gmail.com¹

Abstract. This study aims to examine the unbiased, minimum variance (efficient), and consistent characteristics of Bayes estimator in the Geometric distribution with prior Beta. Based on the results of simulation studies it is found that the Bayes estimator in the Geometric distribution with prior Beta are asymptotically unbiased estimator for values $\theta < 0.5$ and is biased for others, are efficient for the number of samples sizes large and values $\theta \leq 0.6$ and not efficient for others and consistent when value $\theta \leq 0.5$ and inconsistent for other.

Keyword: Bayes estimator, unbiased, efficient, consistent, geometric distribution.

1. Introduction
Bayes method is a method that combines the likelihood function and prior distribution of parameters to get the posterior distribution which is the basis in the estimation parameters [1]. Estimation of parameters using the Bayes method has been widely applied in various fields of research, namely spatial analysis, small area estimation, survival analysis and others such as those carried out by Widiarti et al., Rizki et al., Hartono et al., Yanuar et al., Fulop and Li, Kinyanjui and Korir, Pusponegoro and Rachmawati, Aw and Cabral [2-10]. In the Bayes method, parameter estimation is done by looking at unknown parameters as random variables that have an initial distribution of these parameters, this distribution is called a prior distribution[11].

Prior distributions are usually obtained based on the subjective beliefs of the researchers themselves. However, for the distribution of samples from the Exponential distribution family, one way is to use the prior conjugate [11]. Geometric distribution is a probability distribution with a success event symbolized by $\theta$ and a failure event is denoted by $(1 - \theta)$ and is a distribution that comes from the family of Exponential distribution. The random variable of this distribution is the amount of effort needed to get the first success. One of the prior conjugates for the Exponential distribution family with parameters stating probability is the Beta distribution. The Beta distribution is a continuous distribution at interval $(0,1)$ and has two parameters namely $a$ and $b$ [12]. Characteristics of estimators need to be examined in order to see the quality of an estimator. Estimator is called a good estimator if it meets certain characteristics. Some characteristics of estimators are
unbiased, efficient, consistent, sufficient statistics and complete statistics. An estimator is consistent for its parameters if the sample size gets larger then the estimator will approach the parameter and if the sample size becomes infinite then the estimator will be equal to the parameter or the Mean Square Error (MSE) of the estimator is equal to zero [13]. Thus, this study will examine the quality of Bayes estimators in the Geometric distribution with prior Beta. The characteristics studied are unbiased, efficient and consistent using MSE.

2. Method

2.1. Research Data
The data used in this are simulation data generated using R-Studio Software version 1.2.1.335. The data generated by scenario \( \theta \sim \text{Beta}(\alpha, \beta) \) and \( y \sim \text{Geometric}(\theta) \) as well as with a varying number of sample sizes is 20, 50, 100, 500 and 1000. Paired values \((\alpha, \beta)\) are set as many as nine pairs is \((2,18), (2,8), (2,5), (2,3), (2,2), (3,2), (5,2), (8,2)\) and \((18,2)\) which results \(\theta\) value is 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9.

2.2. Bayes Method
Bayes method is a method that combines the likelihood function and prior distribution of parameters to get the posterior distribution which is the basis in the estimation parameters [1]. If \( B_1, B_2, \ldots, B_k \) is an exclusive and complete event, then for each occurrence of \( B_i \) conditional \( A \) is:

\[
P(B_i|A) = \frac{P(B_i \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_i)P(A|B_i)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)}
\]

Equation 1 is Bayes rule [14].

2.3. Unbiased
The estimator \( \hat{\theta} \) is said to be unbiased for \( \theta \) if the expected value is the same as the parameter of the probability distribution parameter.

\[
E(\hat{\theta}) = \theta
\]

and it is said to be asymptotically biased if the estimator is biased but when the sample size is greater the expected value is the same as the parameter

\[
\lim_{n \to \infty} E(\hat{\theta}) = \theta
\]

or

\[
\lim_{n \to \infty} \text{Biased}(\hat{\theta}) = 0
\]

with

\[
\text{Biased}(\hat{\theta}) = E(\hat{\theta}) - \theta
\]

and \( n \) is the number of sizes sample [13].

2.4. Efficient
The estimator \( \hat{\theta} \) is said to be efficient for \( \theta \) parameter if the estimator \( \hat{\theta} \) has the smallest variance. If \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) are estimator for \( \theta \), then estimator \( \hat{\theta}_1 \) is said to be more efficient than estimator \( \hat{\theta}_2 \) if:

\[
\text{Var} (\hat{\theta}_1) \leq \text{Var} (\hat{\theta}_2)
\]

for \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) is unbiased estimator for \( \theta \) [14]. If \( \hat{\theta} \) is an unbiased estimator for \( \theta \), then the estimator \( \hat{\theta} \) is said to be efficient if and only if the variance of the estimator \( \hat{\theta} \) reaches the Rao-Cramer lower bound.

\[
\text{Var}(\hat{\theta}) \geq \frac{1}{n I(\theta)}
\]
with $\frac{1}{n} l(\theta)$ is the Rao-Cramer lower bound and $l(\theta)$ Fisher’s information [15]. 

$$
I(\theta) = E \left\{ \left( \frac{\partial \ln f(x; \theta)}{\partial \theta} \right)^2 \right\}
$$

or

$$
I(\theta) = -E \left\{ \left( \frac{\partial^2 \ln f(x; \theta)}{\partial \theta^2} \right) \right\}
$$

2.5. Consistent

The estimator $\hat{\theta}$ is said to be consistent for $\theta$ parameter if the estimator is convergent in probability $\theta$ or if the sample size gets larger then the estimator will approach the parameter and if the sample size becomes infinite then the estimator will be equal to the parameter.

$$
\lim_{n \to \infty} p(\theta_n - \theta) \geq \epsilon = 0
$$

atau

$$
\lim_{n \to \infty} E((\hat{\theta} - \theta)^2) = 0
$$

for every $\epsilon > 0$ [16].

The consistency of an estimator also can be evaluated based on the value of MSE. If $\hat{\theta}$ is an estimator of parameter $\theta$, then the MSE for $\hat{\theta}$ is defined as follows:

$$
MSE(\hat{\theta}) = E(\hat{\theta} - E(\hat{\theta}))^2 = E(\hat{\theta} - \theta)^2
= Var(\hat{\theta}) + [Bias(\hat{\theta})]^2
$$

If $\hat{\theta}$ is the an unbiased estimator then MSE will be equal to the variance $\hat{\theta}$ [13].

3. Result and Discussion

3.1 Bayes Estimator

In the Bayes Estimator, the probability function $f(y_i|\theta)$ is expressed by the conditional probability function $f(y_i|\theta)$, so that for $Y_1, ..., Y_n$ a random sample of a population with a Geometric distribution with parameter $\theta$ can be written as follows:

$$
Y_i \sim Geo(\theta) \iff f(y_i|p) = \begin{cases} 
(\theta - \theta^{-1})^{y_i-1} ; & y_i = 1, 2, 3, ..., i = 1, 2, ..., n; 0 \leq \theta \leq 1 \\
0 ; & y_i \text{ lainnya}
\end{cases} (12)
$$

The prior distribution for $Y_i \sim Geo(\theta)$ is $\theta \sim Beta(\alpha, \beta)$. So the probability function of the prior distribution is

$$
f(\theta) = \begin{cases} 
\frac{1}{B(\alpha, \beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1} ; & 0 < \theta < 1 \\
0 ; & \theta \text{ lainnya}
\end{cases} (13)
$$

The likelihood function of $Y_i \sim Geo(\theta)$:

$$
f(y_1, y_2, ..., y_n | \theta) = \prod_{i=1}^{n} \theta(1-\theta)^{y_i-1} = \theta^n(1-\theta)^{\sum_{i=1}^{n} y_i - n} (14)
$$

Joint probability density function of $Y_1, ..., Y_n$ and $\theta$ is:
\[ f(y_1, y_2, \ldots, y_n; \theta) = f(y_1, y_2, \ldots, y_n|\theta) f(\theta) \]

\[ = \theta^n (1 - \theta)^{\sum_{i=1}^{n} y_i - n} \frac{1}{B(\alpha, \beta)} \theta^{\alpha - 1} (1 - \theta)^{\beta - 1} \]

\[ = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha + n - 1} (1 - \theta)^{\beta + \sum_{i=1}^{n} y_i - n - 1} \quad (15) \]

Marginal function of \( Y_i \) is:

\[ m(y_1, y_2, \ldots, y_n) = \int_{-\infty}^{\infty} f(y_1, y_2, \ldots, y_n; \theta) d\theta \]

\[ = \int_{0}^{\infty} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha + n - 1} (1 - \theta)^{\beta + \sum_{i=1}^{n} y_i - n - 1} \]

\[ = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + n)\Gamma(\beta + \sum_{i=1}^{n} y_i - n)}{\Gamma(\alpha + \beta + \sum_{i=1}^{n} y_i)} \quad (16) \]

The posterior distribution is:

\[ f(\theta|y_1, y_2, \ldots, y_n) = \frac{f(y_1, y_2, \ldots, y_n; \theta)}{m(y_1, y_2, \ldots, y_n)} \]

\[ = \frac{\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha + n - 1} (1 - \theta)^{\beta + \sum_{i=1}^{n} y_i - n - 1}}{\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + n)\Gamma(\beta + \sum_{i=1}^{n} y_i - n)}{\Gamma(\alpha + \beta + \sum_{i=1}^{n} y_i)}} \theta^{\alpha + n - 1} (1 - \theta)^{\beta + \sum_{i=1}^{n} y_i - n - 1} \]

\[ = \frac{\Gamma(\alpha + \beta + \sum_{i=1}^{n} y_i)}{\Gamma(\alpha + n)\Gamma(\beta + \sum_{i=1}^{n} y_i - n)} \theta^{\alpha + n - 1} (1 - \theta)^{\beta + \sum_{i=1}^{n} y_i - n - 1} \quad (17) \]

So, the Bayes estimator for \( \theta \) is:

\[ \hat{\theta}^B = E(\theta|y_1, y_2, \ldots, y_n) = \int_{0}^{1} \theta f(\theta|y_1, y_2, \ldots, y_n) \ d\theta \]

\[ = \int_{0}^{1} \frac{\Gamma(\alpha + \beta + \sum_{i=1}^{n} y_i)}{\Gamma(\alpha + n)\Gamma(\beta + \sum_{i=1}^{n} y_i - n)} \theta^{\alpha + n - 1} (1 - \theta)^{\beta + \sum_{i=1}^{n} y_i - n - 1} \ d\theta \]

\[ = \frac{\alpha + n}{\alpha + \beta + \sum_{i=1}^{n} y_i} \quad (18) \]

### 3.2 Unbiased

The estimator \( \hat{\theta}^B \) is said to be unbiased for \( \theta \) if \( E(\hat{\theta}^B) = \theta \).

\[ E(\hat{\theta}^B) = E\left( \frac{\alpha + n}{\alpha + \beta + \sum_{i=1}^{n} y_i} \right) \]
\[ = \frac{(\alpha + n)\theta}{\alpha \theta + \beta \theta + n} \quad (19) \]

Because \( E(\hat{\theta}^B) \neq \theta \), then \( \hat{\theta}^B \) is biased estimator for \( \theta \). But, \( \hat{\theta}^B \) is asymptotically unbiased because:

\[
\lim_{n \to \infty} E(\hat{\theta}^B) = \lim_{n \to \infty} \frac{\alpha \theta + n\theta}{\alpha \theta + \beta \theta + n} \\
= \lim_{n \to \infty} \frac{\alpha \theta + n\theta}{n} \frac{1}{\frac{1}{n} \alpha \theta + \frac{1}{n} \beta \theta + 1} \\
= \frac{0 + \theta}{0 + 0 + 1} = \theta
\]

3.3 Efficient

The estimator \( \hat{\theta}^B \) is efficient for \( \theta \) parameter if the estimator \( \hat{\theta}^B \) has the smallest variance.

\[
Var(\hat{\theta}^B) = Var \left( \frac{\alpha + n}{\alpha + \beta + \sum_{i=1}^{n} Y_i} \right) \\
= (\alpha + n)^2 Var \left( \frac{1}{\alpha + \beta + \sum_{i=1}^{n} Y_i} \right) \quad (21)
\]

Because \( Var \left( \frac{1}{\alpha + \beta + \sum_{i=1}^{n} Y_i} \right) \) cannot be solved analytically, the value of variance will be shown by simulation using R-Studio software version 1.2.1335.

3.4 Consistent

The estimator \( \hat{\theta}^B \) is consistent for \( \theta \) if \( \lim_{n \to \infty} E((\hat{\theta} - \theta)^2) = \lim_{n \to \infty} MSE(\hat{\theta}^B) = 0. \)

\[
\lim_{n \to \infty} E((\hat{\theta}^B - \theta)^2) = \lim_{n \to \infty} MSE(\hat{\theta}^B) \\
= \lim_{n \to \infty} \left\{ Var(\hat{\theta}^B) + [bias(\hat{\theta}^B)]^2 \right\} \quad (22)
\]

Because \( Var(\hat{\theta}^B) \) cannot be solved analytically, the value of MSE will be shown by simulation using R-Studio software version 1.2.1335.

3.5 Simulation

Data simulation using R-Studio Software version 1.2.1.335 with several pairs of \( (\alpha, \beta) \) and varying number of sample sizes yields the bias, variance and MSE values of the Bayes estimator for the Geometric distribution. Tables 1 to 3 display the simulation results for these values.

**Table 1.** The bias value of the Bayes estimator

| \( \alpha \) | \( \beta \) | \( \theta \) | \( n \) | Bias \( \hat{\theta}^B \) | \( \alpha \) | \( \beta \) | \( \theta \) | \( n \) | Bias \( \hat{\theta}^B \) | \( \alpha \) | \( \beta \) | \( \theta \) | \( n \) | Bias \( \hat{\theta}^B \) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 0.015606 | 20 | 0.270064 | 20 | 1.056814 |
| 50 | 0.013049 | 20 | 0.270064 | 20 | 1.056814 |
| 100 | 0.012367 | 20 | 0.270064 | 20 | 1.056814 |

Because \( Var(\hat{\theta}^B) \) cannot be solved analytically, the value of variance will be shown by simulation using R-Studio software version 1.2.1335.
Based on Table 1 and Figure 1, it can be seen that the bias value of Bayes estimator for $\theta < 0.5$ gets smaller when the number of sample sizes (n) gets bigger and vice versa for $\theta > 0.5$ and fluctuative for $\theta = 0.5$. So it can be concluded that the Bayes estimator is asymptotically unbiased estimator for $\theta < 0.5$ and biased for the others.

**Table 2.** The variance value of the Bayes estimator

| $\alpha$ | $\beta$ | $\theta$ | n   | $\text{Var} \hat{\theta}^B$ | $\alpha$ | $\beta$ | $\theta$ | n   | $\text{Var} \hat{\theta}^B$ | $\alpha$ | $\beta$ | $\theta$ | n   | $\text{Var} \hat{\theta}^B$ |
|----------|---------|----------|-----|-----------------------------|----------|---------|----------|-----|-----------------------------|----------|---------|----------|-----|-----------------------------|
| 2        | 18      | 0.1      | 50  | 0.000281                    | 2        | 3       | 0.4      | 50  | 0.015158                    | 5        | 2       | 0.7      | 50  | 0.234915                   |
| 2        | 18      | 0.1      | 100 | 0.000142                    | 2        | 3       | 0.4      | 50  | 0.015158                    | 5        | 2       | 0.7      | 50  | 0.234915                   |
| 2        | 18      | 0.1      | 50  | 0.000281                    | 2        | 3       | 0.4      | 50  | 0.015158                    | 5        | 2       | 0.7      | 50  | 0.234915                   |
| 2        | 18      | 0.1      | 100 | 0.000142                    | 2        | 3       | 0.4      | 50  | 0.015158                    | 5        | 2       | 0.7      | 50  | 0.234915                   |
Based on Table 2 and Figure 2, it can be seen that the variance value for $\theta \leq 0.6$ gets smaller when the number of sample sizes (n) gets bigger and fluctuative for others. The Bayes estimator is efficient when number of samples sizes large and values $\theta \leq 0.6$ and inefficient for the others.

Table 3. The MSE value of the Bayes estimator

| $\alpha$ | $\beta$ | $\theta$ | n  | MSE $\hat{\theta}^B$ | $\alpha$ | $\beta$ | $\theta$ | n  | MSE $\hat{\theta}^B$ | $\alpha$ | $\beta$ | $\theta$ | n  | MSE $\hat{\theta}^B$ |
|----------|---------|----------|----|----------------------|----------|---------|----------|----|----------------------|----------|---------|----------|----|----------------------|
| 2        | 18      | 0.1      | 20 | 0.000919             | 20       | 0.108219| 20       | 0.1288108             |
|          |         |          |    | 50                   | 0.000451 | 50      | 0.0187891 | 5  | 0.2369129             |
|          |         |          |    | 100                  | 0.000295 | 100     | 0.077295  | 100| 2.769218              |
|          |         |          |    |                      |          |         |          |    |                      |          |         |          |    |                      |
Based on Table 3 and Figure 3, it can be seen that MSE value of the Bayes estimator for $\theta \leq 0.5$ gets smaller when the number of sample sizes (n) gets bigger and vice versa for other. So, the Bayes estimator is consistent when value $\theta \leq 0.5$ and inconsistent for other.

4. Conclusion
Based on the result and discussion it can be concluded that the Bayes estimator on the Geometric distribution with prior Beta are asymptotically unbiased estimator for $\theta < 0.5$ and biased for the others, efficient when number of samples sizes large and values $\theta \leq 0.6$ and inefficient for others and consistent when value $\theta \leq 0.5$ and inconsistent for other.
References
[1] Casella G and Berger R 2002 *Statistical Inference Second Edition* (California: Duxbury)
[2] Widiarti, Adityawati N and Nusyirwan 2019 *Journal of Physics: Conf. Series* 1338012039
[3] Widiarti, D D Oktafiani, M Usman, and D Kurniasari 2020 *J. Physics: Conf. Series* 1567022082
[4] Rizki S W, Mara M N and Sulistianingsih E 2017 *Journal of Physics: Conf. Series* 855012036
[5] Hartono B, Kurnia A and Indahwati 2020 *International Journal of Computer Science and Network* 6(6) 2277-5420
[6] Yanuar F, Putri N C E, and Yozza H 2019 *Journal of Physics: Conf. Series* 1317012006
[7] Fulop A and li J 2019 *Journal of Econometrics* 209 114-138
[8] Kinyanjui J K and Korir B C 2020 *International Journal of Statistics and Probability* 9(2)
[9] Pusponegoro N H and Rachmawati N R 2018 *Procedia Computer Science* 135 712-718
[10] Aw A and Cabral E N 2019 *Laboratory of Mathematics* (Ziguinchor: Assane Seck University of Ziguinchor)
[11] Bolstad W M 2007 *Introduction to Bayesian Statistical Second Edition* (USA: John Willey and Sons)
[12] Spiegel M R, Schiller J J and R A Srinivasan 2009 *Probability and Statistics Third Edition* (New York: The McGraw-Hill Companies Inc)
[13] Montgomery D C and Runger G C 2014 *Applied Statistics and Probability for Engineers sixth Edition* (USA: John Wiley and Sons)
[14] Bain L J and Engelhardt 1992 *Introduction to Probability and Mathematical Statistical Second Edition* (California: Duxbury)
[15] Hogg RV, Mckejan J W, and Craig A T 2005 *Introduction to Mathematical Statistics Sixth Edition* (USA: Pearson Education Inc)
[16] Sahoo P 2008 *Probability and Mathematical Statistics* (Louisville: Department of Mathematics University of Louisville)