

Beta-Burr Type X Distribution with Application to Luteinizing Hormone Data

A A D Ikram¹, S Nurrohmah^{2*}, I Fithriani³.

^{1,2,3}*Department of Mathematics, Universitas Indonesia, Depok, 16424, Indonesia.*

^{*}) Corresponding author: snurrohmah@sci.ui.ac.id

Abstract

Beta-Burr Type X distribution is a three parameter distribution and can model right skewed, left skewed, and symmetric data. Beta-Burr Type X distribution is the result of composition distribution functions of beta distribution and Burr Type X distribution. In this study, the characteristics such as probability density function (PDF), cumulative distribution function (CDF), the r -th moment, mean, and variance are presented. The maximum likelihood method is used to estimate the parameters of Beta-Burr Type X distribution, and the solution is obtained using a numerical method. As an illustration, Beta-Burr Type X distribution is used to model the data of luteinizing hormone in female blood samples.

Keywords: beta Distribution, Burr Type X Distribution, maximum likelihood

Introduction

Probability distribution has an important role in data modeling. From the various distribution of opportunities that have been found, the normal distribution is the most commonly used distribution, because it has two parameters that directly state the mean and variance. Besides, the normal distribution has a symmetrical probability density function (Walck, 2007). The properties of the normal distribution are considered ideal. So, the normal distribution is

often used in modeling data. However, in fact, not all data has a symmetrical distribution. Therefore, a suitable distribution is needed to model the data that has an asymmetric distribution shape, right-skewed, or left-skewed.

One of the distributions that suits for modeling data with asymmetrical distribution shape is Burr Distribution. Irving Wingate Burr (1942), introduced Burr distribution. Burr distribution is introduced with twelve forms of distribution functions to model data. Among the twelve distribution functions, one that gets special attention is the Burr Type X distribution, because it can model reliability, health, agriculture, and biological studies. Burr Type X distribution can model data with a form of distribution that is right-skewed and light-tailed. However, the Burr Type X distribution cannot model data that has a symmetric distribution shape or a left-skewed. In real-world problems, not all data has a form of distribution that has a right-skewed.

One of the distributions that can model data with a form of symmetrical distribution, left-skewed, or right-skewed is beta distribution. The beta distribution is very flexible because it can model data with symmetrical distribution shape, left-skewed, or right-skewed. Therefore, beta distribution will have an important role in overcoming the problem of flexibility of a distribution.

Eugene *et al.* (2002) introduced the beta-Generated distribution class. This distribution class composes the distribution functions of beta distribution and the other continuous distribution functions. Then, Khaleel *et al.* (2017) expand the Burr Type X distribution by composing distribution functions of beta distribution and Burr Type X distribution and obtained a distribution named the Beta-Burr Type X distribution. Beta-Burr Type X distribution is more flexible than distribution Burr Type X because there are additional two shape parameters from the beta distribution, so the Beta-Burr Type X distribution can be more flexible in modeling data.

In this paper, beta-Generated distribution will be explained. Then, it will be explained the process of composing distribution functions of beta distribution and Burr Type X distribution. Furthermore, the characteristics such as probability density function (PDF), cumulative distribution function (CDF), the r -th moment, mean, and variance are presented. The maximum likelihood method is used to estimate the parameters of Beta-Burr Type X distribution, and the solution is obtained using a numerical method. As an illustration, Beta-Burr Type X distribution is used to model the data of luteinizing hormone in female blood samples.

Methods

Beta-Generated Distribution

Eugene *et al.* (2002) proposed a new form for building a new distribution. It is known as the beta-Generated class of distribution and it has two shape parameters. If $F(x)$ is the CDF of any random variable, beta-Generated distribution is defined as:

$$F(x) = \frac{1}{B(\alpha, \beta)} \int_0^{G(x)} t^{\alpha-1} (1-t)^{\beta-1} dt, \quad \alpha, \beta > 0, \quad G(x) \in [0,1] \quad (1)$$

where α, β are the two extra shape parameters for the baseline distribution. The PDF of beta-Generated is denoted by equation (1) is

$$f(x) = \frac{1}{B(\alpha, \beta)} [G(x)]^{\alpha-1} [1-G(x)]^{\beta-1} g(x), \quad x > 0, G(x) \in [0,1] \quad (2)$$

whereby

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}. \quad (3)$$

Maximum Likelihood Method

Let X_1, X_2, \dots, X_n be random sample size n from a certain distribution with PDF of X is $f(x; \theta)$ that depends on $\theta \in \Omega$, where Ω is a space of parameters. Then likelihood function can be obtained as joint PDF of X_1, X_2, \dots, X_n , denoted by $L(\theta; x_1, x_2, \dots, x_n)$ or $L(\theta)$, as follows:

$$L(\theta) = f(x_1; \theta)f(x_2; \theta) \dots f(x_n; \theta) = \prod_{i=1}^n f(x_i; \theta); \theta \in \Omega \quad (4)$$

Let $u(x_1, x_2, \dots, x_n)$ be a function from x_1, x_2, \dots, x_n so that if θ replaced by $u(x_1, x_2, \dots, x_n)$, the likelihood function $L(\theta; x_1, x_2, \dots, x_n)$ reach maximum value. So, $u(x_1, x_2, \dots, x_n)$ be the maximum likelihood estimator (MLE) for θ which denoted by $\hat{\theta}$.

Results and Discussion

Beta-Burr Type X Distribution

Let X be a random variable of Burr Type X distribution with shape parameter k , then X has the following CDF,

$$G(x) = \Pr(X \leq x) = (1 - e^{-x^2})^k, \quad x, k > 0 \quad (5)$$

and the PDF of Burr Type X distribution is denoted by equation (4) is

$$g(x) = 2kxe^{-x^2}(1 - e^{-x^2})^{k-1}, \quad x, k > 0 \quad (6)$$

then inserting (5) and (6) into (1), the PDF for the Beta-Burr Type X distribution obtained as follows

$$f(x) = \frac{1}{B(\alpha, \beta)} \left[(1 - e^{-x^2})^k \right]^{\alpha-1} \left[1 - G(1 - e^{-x^2})^k \right]^{\beta-1} 2kxe^{-x^2} \times (1 - e^{-x^2})^{k-1} \quad (7)$$

that can be reduced to

$$f(x) = \frac{2kxe^{-x^2}(1 - e^{-x^2})^{k\alpha-1}}{B(\alpha, \beta)} \left[1 - (1 - e^{-x^2})^k \right]^{\beta-1}, \quad \alpha, \beta, k, x > 0. \quad (8)$$

Furthermore, CDF for the Beta-Burr Type X distribution can be obtained by inserting (5) into (1),

$$F(x) = \frac{1}{B(\alpha, \beta)} \int_0^{(1-e^{-x^2})^k} t^{\alpha-1} (1-t)^{\beta-1} dt, \quad \alpha, \beta > 0, \quad G(x) \in [0,1]. \quad (9)$$

Figure 1 display a variety of possible shapes of the PDF with the various value of the parameters.

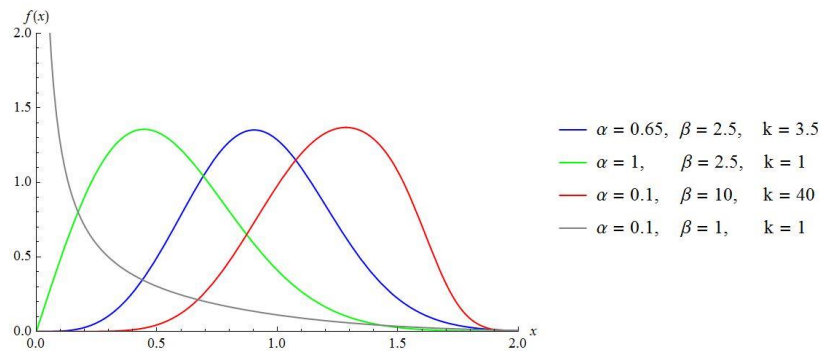


Figure 1. The PDF plots with various values of parameters.

Figure 1 shows that PDF from Beta-Burr Type X distribution can be L-shaped, left-skewed, right-skewed, or symmetrical.

Furthermore, with (8), the r -th moment of Beta-Burr Type X distribution can be defined as

$$E(X^r) = \frac{k\Gamma\left(\frac{r}{2} + 1\right)}{B(\alpha, \beta)} \sum_{i,j=0}^{\infty} \frac{(-1)^{j+i} \Gamma(\beta) \Gamma(k(\alpha + j))}{\Gamma(\beta - j) \Gamma(k(\alpha + j) - i) i! j! (1+i)^{\frac{r}{2}+1}}. \quad (10)$$

Based on the definition of the first, second moment, and (10), mean and variance of Beta-Burr Type X distribution are respectively given by

$$E(X) = \frac{k\sqrt{\pi}}{2B(\alpha, \beta)} \sum_{i,j=0}^{\infty} \frac{(-1)^{j+i} \Gamma(\beta) \Gamma(k(\alpha + j))}{\Gamma(\beta - j) \Gamma(k(\alpha + j) - i) i! j! (1+i)^{\frac{3}{2}}} \quad (11)$$

$$Var(X) = \frac{k}{B(\alpha, \beta)} \sum_{i,j=0}^{\infty} \frac{(-1)^{j+i} \Gamma(\beta) \Gamma(k(\alpha + j))}{\Gamma(\beta - j) \Gamma(k(\alpha + j) - i) i! j! (1+i)^2} \quad (12)$$

$$-\left(\frac{k\sqrt{\pi}}{2B(\alpha, \beta)} \sum_{i,j=0}^{\infty} \frac{(-1)^{j+i} \Gamma(\beta) \Gamma(k(\alpha + j))}{\Gamma(\beta - j) \Gamma(k(\alpha + j) - i) i! j! (1 + i)^{\frac{3}{2}}}\right)^2$$

Parameters Estimation of Beta-Burr Type X Distribution

To use the maximum likelihood method for parameter estimation purpose, the likelihood function or log-likelihood function are needed. Let X_1, X_2, \dots, X_n be independent and identic random variables of size n with PDF (8), then the likelihood function with observed sample $x = (x_1, x_2, \dots, x_n)$

$$L(\alpha, \beta, k|x) = \frac{2^n k^n x^n [\Gamma(\alpha + \beta)]^n}{[\Gamma(\alpha)]^n [\Gamma(\beta)]^n} e^{-\sum_{i=1}^n x_i^2} \prod_{i=1}^n (1 - e^{-x_i^2})^{k\alpha-1} \times \prod_{i=1}^n [1 - (1 - e^{-x_i^2})^k]^{\beta-1} \quad (13)$$

and its log-likelihood function

$$\ln(\alpha, \beta, k|x) = n[\ln 2 + \ln k + \ln x + \ln \Gamma(\alpha + \beta) - \ln \Gamma(\alpha) - \ln \Gamma(\beta)] - \sum_{i=1}^n x_i^2 + (k\alpha - 1) \sum_{i=1}^n \ln(1 - e^{-x_i^2}) + (\beta - 1) \times \sum_{i=1}^n \ln [1 - (1 - e^{-x_i^2})^k] \quad (14)$$

The maximum likelihood estimators of α , β , and k can be obtained by finding the solution of these three equations

$$\frac{\partial \ln l(\alpha, \beta, k|x)}{\partial \alpha} = \psi(\alpha + \beta) - n \psi(\alpha) + k \sum_{i=1}^n \ln(1 - e^{-x_i^2}) = 0 \quad (15)$$

$$\frac{\partial \ln l(\alpha, \beta, k|x)}{\partial \beta} = n \psi(\alpha + \beta) - n \psi(\beta) + \sum_{i=1}^n \ln [1 - (1 - e^{-x_i^2})^k] = 0 \quad (16)$$

$$\frac{\partial \ln l(\alpha, \beta, k|x)}{\partial k} = \frac{n}{k} + k \sum_{i=1}^n \ln(1 - e^{-x_i^2}) - (\beta - 1) \sum_{i=1}^n \frac{(1 - e^{-x_i^2})^k \ln(1 - e^{-x_i^2})}{[1 - (1 - e^{-x_i^2})^k]} = 0 \quad (17)$$

The above equation cannot be solved analytically. Therefore, estimating parameters α , β , and k can be obtained numerically by using the library in the RStudio version 1.1.463 with the R programming language where the library will find the optimum value in each parameter.

Application

The application of the Beta-Burr Type X distribution in modeling data will be illustrated. The data used in this illustration are data about the luteinizing hormone in female blood sample quoted from the book A Biostatistical Introduction (Diggle, 1990). The data consists of 48 observations and has a nanogram per deciliter unit (*ng/dL*).

Table 1. Luteinizing hormone data in female blood samples (*ng/dL*).

1.4	1.5	1.5	1.7	1.7	1.8	1.8	1.8	1.9	1.9	1.9	2
2	2	2.1	2.1	2.1	2.2	2.2	2.2	2.2	2.3	2.3	2.3
2.3	2.4	2.4	2.4	2.4	2.5	2.6	2.6	2.7	2.7	2.7	2.7
2.9	2.9	2.9	3	3	3.1	3.2	3.2	3.3	3.4	3.5	3.5

The estimation parameters obtained by using software RStudio version 1.1.463 with the maximum likelihood method.

Table 2. Estimated parameters.

No.	Distribution	Parameter
1	Burr Type X	$\hat{k} = 56.0822$
2	Beta-Burr Type X	$\hat{\alpha} = 23.4572361, \hat{\beta} = 0.3349163,$ $\hat{k} = 0.812042$

Figure 2 and 3 shows the PDF and CDF of Burr Type X distribution and Beta-Burr Type X distribution based on estimated parameters Table 2.

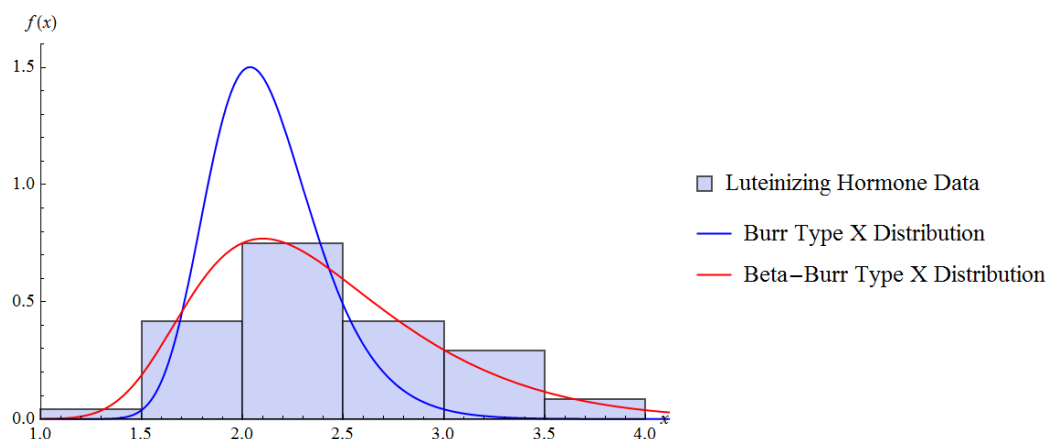


Figure 2. Histogram of luteinizing hormone data, PDF of Burr Type X distribution, and Beta-Burr Type X distribution.

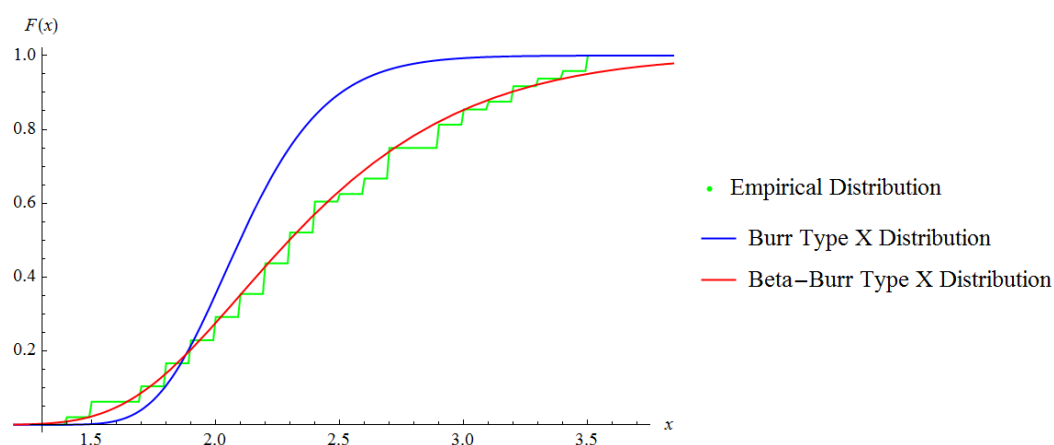


Figure 3. Empirical distribution, CDF of Burr Type X and Beta-Burr Type X distribution.

Table 3. Kolmogorov-Smirnov test statistic.

No.	Distribution	Kolmogorov-Smirnov
1	Burr Type X	0.296113
2	Beta-Burr Type X	0.054108

Figure 2 and 3 have shown that the Beta-Burr Type X distribution provides a better fit than the Burr Type X distribution for the luteinizing hormone data. Furthermore, for $1 - \alpha = 0.95$ and $n = 48$ the value of $w_{0,95} = 0,19629$. Then, with $w_{0,95} = 0,19629$ and at a significance level of $\alpha = 0.05$, Burr Type X distribution has a statistical value greater than $w_{0,95}$, so it is not

suitable for modeling luteinizing hormone data, whereas Beta-Burr Type X distribution has a statistical value smaller than $w_{0,95}$, so Beta-Burr Type X distribution is suitable for modeling data luteinizing hormone.

Conclusion

Based on all the results, it can be concluded that Beta-Burr Type X distribution is the result of composition distribution functions of beta distribution and Burr Type X distribution that has three parameters shape, which is α , β , and k . Then, Beta-Burr Type X distribution can model data with symmetrical distribution shape, left-skewed, or right-skewed. Furthermore, several statistical properties like the r -th moment, mean, and variance are presented. Finally, the luteinizing hormone in female blood samples data is illustrated. The Beta-Burr Type X distribution provides a better fit than Burr Type X distribution.

References

- Burr, I. W. (1942). Cumulative frequency functions. *Annals of Mathematical Statistics*, 13, 215–232
- Diggle, P. (1990). *Time Series: A Biostatistical Introduction*. Oxford.
- Eugene, N., Lee, C., and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and methods*, 31(4):497–512.
- Khaleel, M. A., Ibrahim, N. A., Shitan, M., Merovci, F., & Rehman, E. (2017). *Beta Burr Type X with Application to Rainfall Data*. *Journal of Mathematical Sciences*, 73–86.
- Walck, C. (2007). *Hand-book on statistical distributions for experimentalists*. Sweden: University of Stockholm.