Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 11, Number 4 (2015), pp. 1983-1986

http://www.ripublication.com

Poisson Approximation for A Sum of Independent Beta Binomial Random Variables

K. Teerapabolarn

Department of Mathematics,
Faculty of Science,
Burapha University,
Chonburi 20131, Thailand.
E-mail: kanint@buu.ac.th

Abstract

The Stein-Chen method and the beta binomial w-functions are used to determine a bound for the total variation distance between the distribution of a sum of n independent beta binomial random variables, each with parameters n_i, α_i and β_i, and a Poisson distribution with mean $\sum_{i=1}^{n} \frac{n_i\alpha_i}{\alpha_i + \beta_i}$. The result obtained in this study gives a good approximation when β_i is large with respect to n_i and α_i for every i.

AMS subject classification: 60F05, 60G50.
Keywords: Beta binomial distribution, Poisson approximation, Stein-Chen method, beta binomial w-function.

1. Introduction

Let $X_1, \ldots, X_n$ be independently distributed beta binomial random variables, each with the probability $P(X_i = k) = \binom{n_i}{k} \frac{B(k + \alpha_i, n_i - k + \beta_i)}{B(\alpha_i, \beta_i)}$ for $k = 0, \ldots, n_i$, where B is the complete beta function, and mean $\mu_i = \frac{n_i\alpha_i}{\alpha_i + \beta_i}$ and variance $\sigma_i^2 = \frac{n_i\alpha_i\beta_i(n_i + \alpha_i + \beta_i)}{(\alpha_i + \beta_i)^2(1 + \ alpha_i + \beta_i)}$. Let $\mathbf{S}_n = \sum_{i=1}^{n} X_i$ and $\mathbf{P}_\lambda$ denote the the Poisson random variable with mean λ. Let $\lambda = \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} \mu_i$, if all λ_i are small, the distribution of $\mathbf{S}_n$

can be approximated by the Poisson distribution with mean λ. For $n = 1$, Teerapabolarn [3] gave a bound for the total variation distance between a beta binomial distribution with parameter n, α and β and a Poisson distribution with mean $\lambda = \dfrac{n\alpha}{\alpha+\beta}$ as follows:

$$d(\mathbf{BB}_{n,\alpha,\beta}, \mathbf{P}_\lambda) \le \left(1 - e^{-\lambda}\right) \frac{(\alpha+1)(\alpha+\beta) - n\beta}{(\alpha+\beta)(\alpha+\beta+1)}, \tag{1.1}$$

where $d(\mathbf{BB}_{n,\alpha,\beta}, \mathbf{P}_\lambda) = \sup_{A\subseteq\mathbb{N}\cup\{0\}} |P(\mathbf{BB}_{n,\alpha,\beta} \in A) - P(\mathbf{P}_\lambda \in A)|$ and $\mathbf{BB}_{n,\alpha,\beta}$ is the beta binomial random variable with parameters n, α and β.

In this paper, we are interested to determine a bound for approximating the distribution of a sum of $n(>1)$ independent beta binomial random variables by a Poisson distribution with mean λ, in the form of $d(\mathbf{S}_n, \mathbf{P}_\lambda)$. The tools for giving the desired result are the Stein-Chen method and the beta binomial w-functions, which are in Section 2. In Section 3, our result is derived by these tools and the conclusion of this study is presented in the last section.

2. Method

The following lemma gives the beta binomial w-functions, which are directly obtained from [3].

Lemma 2.1. For $1 \le i \le n$, let w_i be the w-function associated with the beta binomial random variable X_i, then we have the following:

$$w_i(k) = \frac{(n_i - k)(\alpha_i + k)}{(\alpha_i + \beta_i)\sigma_i^2}\ ,\ k = 0, \dots, n_i. \tag{2.1}$$

The following relation is an important property for proving the result, which was stated by [2].

$$\begin{aligned} Cov(\mathbf{S}_n, f(\mathbf{S}_n)) &= \sum_{i=1}^{n} Cov\left(X_i, f\left(X_i + \sum_{j\neq i} X_j\right)\right) \\ &= \sum_{i=1}^{n} \sigma_i^2 E[w_i(X_i)\Delta f(\mathbf{S}_n)], \end{aligned} \tag{2.2}$$

for any function $f : \mathbb{N}\cup\{0\} \to \mathbb{R}$ for which $E|w_i(X_i)\Delta f(\mathbf{S}_n)| < \infty$, where $\Delta f(x) = f(x+1) - f(x)$.

For the Stein-Chen method, following [1], it is applied for every constant $\lambda > 0$, every subset A of $\mathbb{N}\cup\{0\}$ and the bounded real valued function $f = f_A : \mathbb{N}\cup\{0\} \to \mathbb{R}$. Thus, Stein's equation for the Poisson distribution with mean λ is of the form

$$P(\mathbf{S}_n \in A) - P(\mathbf{P}_\lambda \in A) = E[\lambda_n f(\mathbf{S}_n + 1) - \mathbf{S}_n f(\mathbf{S}_n)]. \tag{2.3}$$

For any subset A of $\mathbb{N} \cup \{0\}$ and for every $x \in \mathbb{N}$, [1] showed that

$$\sup_{A,x} |\Delta f(x)| = \sup_{A,x} |f(x+1) - f(x)| \leq \frac{1}{x}. \tag{2.4}$$

3. Result

The following theorem presents the main result of this study.

Theorem 3.1. Let $\lambda = \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} \mu_i$, $\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$ and $\alpha_i \geq n_i - 1$ for every $i \in \{1, \ldots, n\}$. Then we have the following:

$$d(\mathbf{S}_n, \mathbf{P}_\lambda) \leq \frac{1 - e^{-\lambda}}{\lambda} \sum_{i=1}^{n} \frac{n_i \alpha_i [(\alpha_i + 1)(\alpha_i + \beta_i) - n_i \beta_i]}{(\alpha_i + \beta_i)^2 (\alpha_i + \beta_i + 1)}. \tag{3.1}$$

Proof. From (2.3), it follows that

$$\begin{aligned} d(\mathbf{S}_n, \mathbf{P}_\lambda) &= |\lambda E[f(\mathbf{S}_n + 1)] - E[\mathbf{S}_n f(\mathbf{S}_n)]| \\ &= |\lambda E[\Delta f(\mathbf{S}_n)] - Cov(\mathbf{S}_n, f(\mathbf{S}_n))| \\ &= \left| \sum_{i=1}^{n} \mu_i E[\Delta f(\mathbf{S}_n)] - \sum_{i=1}^{n} Cov(X_i, f(\mathbf{S}_n)) \right|. \end{aligned}$$

Using (2.2) and Lemma 1, we have

$$\begin{aligned} d(\mathbf{S}_n, \mathbf{P}_\lambda) &= \left| \sum_{i=1}^{n} \left\{ E\left[\mu_i \Delta f(\mathbf{S}_n)\right] - \sigma_i^2 E[w_i(X_i) \Delta f(\mathbf{S}_n)] \right\} \right| \\ &\leq \sum_{i=1}^{n} E\{|n_i \alpha_i / (\alpha_i + \beta_i) - \sigma_i^2 w_i(X_i)||\Delta f(\mathbf{S}_n)|\} \\ &\leq \sup_{A,x} |\Delta f(x)| \sum_{i=1}^{n} E\, |n_i \alpha_i / (\alpha_i + \beta_i) - (n_i - X_i)(\alpha_i + X_i)/(\alpha_i + \beta_i)| \\ &\leq \frac{1 - e^{-\lambda}}{\lambda} \sum_{i=1}^{n} \frac{n_i \alpha_i [(\alpha_i + 1)(\alpha_i + \beta_i) - n_i \beta_i]}{(\alpha_i + \beta_i)^2 (\alpha_i + \beta_i + 1)}. \end{aligned}$$

Hence, (3.1) is obtained. ■

If X_i's are identically distributed, then the following corollary is an immediately consequence of the Theorem 3.1.

Corollary 3.2. If $n_i = m$, $\alpha_i = \alpha$ and $\beta_i = \beta$ for every i, then $\lambda = \dfrac{nm\alpha}{\alpha + \beta}$ and the following inequality holds:

$$d(\mathbf{S}_n, \mathbf{P}_\lambda) \leq \left(1 - e^{-\lambda}\right) \frac{(\alpha + 1)(\alpha + \beta) - m\beta}{(\alpha + \beta)(\alpha + \beta + 1)}. \tag{3.2}$$

4. Conclusion

In this study, a bound on the error of Poisson approximation to the distribution of a sum of n independent beta binomial random variables was obtained, by using the Stein-Chen method and the beta binomial w-functions. The distribution of this summands can be approximated by a Poisson distribution with mean $\sum_{i=1}^{n} \dfrac{n_i \alpha_i}{\alpha_i + \beta_i}$ when β_i is large with respect to n_i and α_i for every i. In addition, when $n = 1$, the result in Theorem 1 is the same result that reported in [3].

References

[1] Barbour, A. D., Holst, L., Janson, S., 1992, "Poisson approximation", Oxford Studies in Probability 2, Clarendon Press, Oxford.

[2] Cacoullos, T., Papathanasiou, V., 1989, "Characterization of distributions by variance bounds", Statist. Probab. Lett., 7(5), 351–356.

[3] Teerapabolarn, K., 2008, "Poisson approximation to the beta binomial distribution", Int. J. Contemp. Math. Sciences, 3(10), 457–461.