Songklanakarin J. Sci. Technol.
41 (1), 96-108, Jan. – Feb. 2019

http://www.sjst.psu.ac.th

Original Article

A non-uniform bound on binomial approximation to the beta binomial cumulative distribution function

Kanint Teerapabolarn[1, 2*] and Khunakorn Sae-Jeng[1]

[1] *Department of Mathematics, Faculty of Science, Burapha University, Mueang, Chonburi, 20131 Thailand*

[2] *Centre of Excellence in Mathematics, Commission on Higher Education, Ministry of Education, Ratchathewi, Bangkok, 10400 Thailand*

Received: 29 June 2017; Revised: 31 August 2017; Accepted: 2 October 2017

Abstract

This paper uses Stein's method and the characterization of beta binomial random variable to determine a non-uniform bound for the distance between the beta binomial cumulative distribution function with parameters $n \in \mathbb{N}$, $\alpha > 0$ and $\beta > 0$ and the binomial cumulative distribution function with parameters n and $\frac{\alpha}{\alpha+\beta}$. Some numerical examples are given to illustrate the obtained result.

Keywords: beta binomial cumulative distribution function, binomial approximation, characterization of the beta binomial random variable, non-uniform bound, Stein's method

1. Introduction

Let X be the binomial random variable with parameters $n \in \mathbb{N}$ and $p \in (0,1)$. Its probability mass function is as follows:

$$b(x) = \binom{n}{x} p^x q^{n-x},\ x \in \{0,...,n\}, \tag{1.1}$$

where $q = 1 - p$ and $E(X) = np$ and $Var(X) = npq$ are its mean and variance, respectively. It is well-known that if the

*Corresponding author
Email address: kanint@buu.ac.th

probability of success in the binomial distribution is a random variable and has a beta distribution with parameters $\alpha > 0$ and $\beta > 0,$ then a new resulting distribution is referred to as the beta binomial distribution with parameters $n,$ α and $\beta.$ Let Y be the beta binomial distribution with parameters $n,$ α and $\beta.$ and its probability mass function is of the form

$$bb(y) = \binom{n}{y} \frac{B(\alpha+y, \beta+n-y)}{B(\alpha,\beta)},\ y \in \{0,...,n\}, \tag{1.2}$$

where B is the complete beta function and $\mu = \frac{n\alpha}{\alpha+\beta}$ and $\sigma^2 = \frac{n\alpha\beta(n+\alpha+\beta)}{(\alpha+\beta+1)(\alpha+\beta)^2}$ are the mean and variance of $Y,$ respectively. Some useful applications of this distribution can be found in field such as animal teratology experiments in Gupta and Naradajah (2004), statistical process control in Sant'Anna and Caten (2012), inferential statistics in Salem and Abu El Azm (2013) and balancing revenues and repair costs in Ding, Rusmevichientong, and Topaloglu (2014).

For limiting distribution, it follows from (1.2) that if $\alpha, \beta \to \infty$ in such a way $\frac{\alpha}{\alpha+\beta}$ tends to a constant, then the beta binomial distribution with parameters $n,$ α and β converges to a binomial distribution with parameters $n,$ and $\frac{\alpha}{\alpha+\beta}$. In this case, Teerapabolarn (2008) used Stein's method and the binomial w-function to give a uniform bound on binomial approximation to the beta binomial distribution as follows:

$$d_A(Y, X) \le (1 - p^{n+1} - q^{n+1}) \frac{(n-1)n}{(n+1)(\alpha+\beta+1)} \tag{1.3}$$

for every subset A of $\{0,...,n\},$ where $d_A(Y,X) = |P(Y \in A) - P(X \in A)|$ is the distance between the beta binomial distribution with parameters $n,$ α and β and the binomial distribution with parameters n and $\frac{\alpha}{\alpha+\beta}.$ For $A = C_{x_0} = \{0,...,x_0\},\ x_0 \in \{0,...,n\},$ the result in (1.3) becomes

$$d_{C_{x_0}}(Y, X) \le (1 - p^{n+1} - q^{n+1}) \frac{(n-1)n}{(n+1)(\alpha+\beta+1)}, \tag{1.4}$$

where $d_{C_{x_0}}(Y,X) = |P(Y \le x_0) - P(X \le x_0)|$ is the distance between the beta binomial cumulative distribution function with parameters $n,$ α and β and the binomial cumulative distribution function with parameters n and $\frac{\alpha}{\alpha+\beta}$ at $x_0.$

We observe that the bound in (1.4) is uniform in $x_0 \in \{0,...,n\},$ that is, it does not change along $x_0 \in \{0,...,n\}.$ So, it may be inappropriate for measuring the accuracy of the approximation. In this paper, we are interested to determine a non-uniform bound with respect to the bounds in (1.4) by using Stein's method and the characterization of beta binomial random variable, which are described and determined in Sections 2 and 3, respectively. In Section 4, some numerical examples provided to illustrate the obtained result, and the conclusion of this study is presented in the last Section.

2. Method

Stein's method and the characterization of beta binomial random variable are both tools which can be used to obtain the desired result.

2.1 Stein's method for binomial distribution

Stein (1972) proposed a powerful method for normal approximation, which is called Stein's method. In 1986, he also applied this method to binomial approximation (Stein, 1986). Following Barbour, Holst, and Janson (1992), Stein's equation for binomial distribution with parameters $n \in \mathbb{N}$ and $0 < p = 1 - q < 1$, for given *h*, is of the form

$$h(x) - \mathcal{B}_{n,p}(h) = (n-x)pf(x+1) - qxf(x), \tag{2.1}$$

where $\mathcal{B}_{n,p}(h) = \sum_{k=0}^{n} h(k)\binom{n}{k}p^k q^{n-k}$ and *f* and *h* are bounded real valued functions defined on $\{0,...,n\}$. For $A \subseteq \{0,...,n\}$, let $h_A : \{0,...,n\} \to \mathbb{R}$ be defined by

$$h_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \tag{2.2}$$

Following Barbour et al. (1992), let $f_A : \mathbb{N} \cup \{0\} \to \mathbb{R}$ satisfy (2.1), where $f_A(0) = f_A(1)$ and $f_A(x) = f_A(n)$ for $x \geq n$. Therefore, the solution f_A of (2.1) can be expressed as

$$f_A(x) = \frac{\mathcal{B}_{n,p}(h_{A \cap C_{x-1}}) - \mathcal{B}_{n,p}(h_A)\mathcal{B}_{n,p}(h_{C_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}}, \tag{2.3}$$

where $C_{x-1} = \{0,...,x-1\}$. Similarly, for $A = \{x_0\}$ and $A = C_{x_0}$ when $x_0 \in \{0,...,n\}$ and by setting $h_{x_0} = h_{\{x_0\}}$, thus the solutions $f_{x_0} = f_{\{x_0\}}$ and $f_{C_{x_0}}$ are as follows:

$$f_{x_0}(x) = \begin{cases} -\dfrac{\mathcal{B}_{n,p}(h_{x_0})\mathcal{B}_{n,p}(h_{C_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}} & \text{if } x \leq x_0, \\ \dfrac{\mathcal{B}_{n,p}(h_{x_0})\mathcal{B}_{n,p}(1-h_{C_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}} & \text{if } x > x_0 \end{cases} \tag{2.4}$$

and

$$f_{C_{x_0}}(x) = \begin{cases} \dfrac{\mathcal{B}_{n,p}(h_{C_{x-1}})\mathcal{B}_{n,p}(1-h_{C_{x_0}})}{x\binom{n}{x}p^x q^{n-x+1}} & \text{if } x \leq x_0, \\ \dfrac{\mathcal{B}_{n,p}(h_{C_{x_0}})\mathcal{B}_{n,p}(1-h_{C_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}} & \text{if } x > x_0. \end{cases} \tag{2.5}$$

Let $\Delta f_{x_0}(x) = f_{x_0}(x+1) - f_{x_0}(x)$ and $\Delta f_{C_{x_0}}(x) = f_{C_{x_0}}(x+1) - f_{C_{x_0}}(x)$. It is seen that $\Delta f_{x_0}(x) = 0$ and $\Delta f_{C_{x_0}}(x) = 0$ for $x = 0, n$. Thus, for $x \in \{1,..,n-1\}$ and by (2.4), we obtain

$$\Delta f_{x_0}(x) = \begin{cases} \dfrac{\mathcal{B}_{n,p}(h_{x_0})}{\binom{n}{x}p^x q^{n-x}}\left[-\dfrac{\mathcal{B}_{n,p}(h_{C_x})}{(n-x)p} + \dfrac{\mathcal{B}_{n,p}(h_{C_{x-1}})}{xq}\right] & \text{if } x < x_0, \\ \dfrac{\mathcal{B}_{n,p}(h_{x_0})}{\binom{n}{x}p^x q^{n-x}}\left[\dfrac{\mathcal{B}_{n,p}(1-h_{C_x})}{(n-x)p} + \dfrac{\mathcal{B}_{n,p}(h_{C_{x-1}})}{xq}\right] & \text{if } x = x_0, \\ \dfrac{\mathcal{B}_{n,p}(h_{x_0})}{\binom{n}{x}p^x q^{n-x}}\left[\dfrac{\mathcal{B}_{n,p}(1-h_{C_x})}{(n-x)p} - \dfrac{\mathcal{B}_{n,p}(1-h_{C_{x-1}})}{xq}\right] & \text{if } x > x_0 \end{cases}$$

$$= \begin{cases} \dfrac{\binom{n}{x_0}p^{x_0}q^{n-x_0}}{\binom{n}{x}p^x q^{n-x}}\left[-\dfrac{\sum_{k=0}^{x}\binom{n}{k}p^k q^{n-k}}{(n-x)p} + \dfrac{\sum_{k=0}^{x-1}\binom{n}{k}p^k q^{n-k}}{xq}\right] & \text{if } x < x_0, \\ \dfrac{\binom{n}{x_0}p^{x_0}q^{n-x_0}}{\binom{n}{x}p^x q^{n-x}}\left[\dfrac{\sum_{k=x+1}^{n}\binom{n}{k}p^k q^{n-k}}{(n-x)p} + \dfrac{\sum_{k=0}^{x-1}\binom{n}{k}p^k q^{n-k}}{xq}\right] & \text{if } x = x_0, \\ \dfrac{\binom{n}{x_0}p^{x_0}q^{n-x_0}}{\binom{n}{x}p^x q^{n-x}}\left[\dfrac{\sum_{k=x+1}^{n}\binom{n}{k}p^k q^{n-k}}{(n-x)p} - \dfrac{\sum_{k=x}^{n}\binom{n}{k}p^k q^{n-k}}{xq}\right] & \text{if } x > x_0. \end{cases} \tag{2.6}$$

By (2.5), we also obtain

$$\Delta f_{C_{x_0}}(x) = \begin{cases} \dfrac{\mathcal{B}_{n,p}(1-h_{C_{x_0}})}{\binom{n}{x}p^x q^{n-x}}\left[\dfrac{\mathcal{B}_{n,p}(h_{C_x})}{(n-x)p} - \dfrac{\mathcal{B}_{n,p}(h_{C_{x-1}})}{xq}\right] & \text{if } x \le x_0, \\ \dfrac{\mathcal{B}_{n,p}(h_{C_{x_0}})}{\binom{n}{x}p^x q^{n-x}}\left[\dfrac{\mathcal{B}_{n,p}(1-h_{C_x})}{(n-x)p} - \dfrac{\mathcal{B}_{n,p}(1-h_{C_{x-1}})}{xq}\right] & \text{if } x > x_0 \end{cases}$$

$$= \begin{cases} \dfrac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^x q^{n-x}}\left[\dfrac{\sum_{k=0}^{x}\binom{n}{k}p^k q^{n-k}}{(n-x)p} - \dfrac{\sum_{k=0}^{x-1}\binom{n}{k}p^k q^{n-k}}{xq}\right] & \text{if } x \le x_0, \\ \dfrac{\sum_{j=0}^{x_0}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^x q^{n-x}}\left[\dfrac{\sum_{k=x+1}^{n}\binom{n}{k}p^k q^{n-k}}{(n-x)p} - \dfrac{\sum_{k=x}^{n}\binom{n}{k}p^k q^{n-k}}{xq}\right] & \text{if } x > x_0. \end{cases} \tag{2.7}$$

The following lemmas present some necessarily properties of Δf_{x_0} and $\Delta f_{C_{x_0}}$, which are used to prove the main result.

Lemma 2.1. *Let* $x \in \{1,...,n-1\}$ *and* $x_0 = \{1,...,n\}$, *then the following inequalities hold:*

$$\Delta f_{x_0}(x) \begin{cases} > 0 & if \ x_0 = x, \\ < 0 & if \ x_0 \neq x \end{cases} \tag{2.8}$$

and

$$\Delta f_{C_{x_0}}(x) \begin{cases} > 0 & if \ x \leq x_0, \\ < 0 & if \ x > x_0. \end{cases} \tag{2.9}$$

Proof. Firstly, we have to show that (2.8) holds. For $x_0 = x$, it follows from (2.6) that

$$\Delta f_{x_0}(x) > 0. \tag{2.10}$$

For $x_0 \neq x$, we have to show $\Delta f_{x_0}(x) < 0$ when $x_0 > x$ and $x_0 < x$ as follows. From (2.6), for $x_0 > x$,

$$\begin{aligned}\Delta f_{x_0}(x) &= \frac{\binom{n}{x_0} p^{x_0} q^{n-x_0}}{\binom{n}{x} p^x q^{n-x}} \left[-\frac{\sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k}}{(n-x)p} + \frac{\sum_{k=0}^{x-1} \binom{n}{k} p^k q^{n-k}}{xq} \right] \\ &= \frac{\binom{n}{x_0} p^{x_0} q^{n-x_0}}{x\binom{n}{x} p^{x+1} q^{n-x+1}(n-x)} \left[-x\sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k+1} + \sum_{k=0}^{x-1} (n-x)\binom{n}{k} p^{k+1} q^{n-k} \right].\end{aligned}$$

Let $\xi_1(x) = -x\sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k+1} + \sum_{k=0}^{x-1} (n-x)\binom{n}{k} p^{k+1} q^{n-k}$, then

$$\begin{aligned}\xi_1(x) &= -\sum_{k=0}^{x} \frac{x(n-k+1)}{n+1} \binom{n+1}{k} p^k q^{n-k+1} + \sum_{k=0}^{x-1} \frac{(k+1)(n-x)}{n+1} \binom{n+1}{k+1} p^{k+1} q^{n+1-(k+1)} \\ &= -\sum_{k=0}^{x} \frac{x(n-k+1)}{n+1} \binom{n+1}{k} p^k q^{n-k+1} + \sum_{k=0}^{x} \frac{k(n-x)}{n+1} \binom{n+1}{k} p^k q^{n+1-k} \\ &= \sum_{k=0}^{x} \frac{\binom{n+1}{k} p^k q^{n+1-k}}{n+1} \left[-x(n+1) + kn \right] \\ &< 0.\end{aligned}$$

Thus for $x_0 > x$,

$$\Delta f_{x_0}(x) < 0. \tag{2.11}$$

For $x_0 < x$,

$$\Delta f_{x_0}(x) = \frac{\binom{n}{x_0} p^{x_0} q^{n-x_0}}{\binom{n}{x} p^{x} q^{n-x}} \left[\frac{\sum_{k=x+1}^{n} \binom{n}{k} p^k q^{n-k}}{(n-x)p} - \frac{\sum_{k=x}^{n} \binom{n}{k} p^k q^{n-k}}{xq} \right]$$

$$= \frac{\binom{n}{x_0} p^{x_0} q^{n-x_0}}{x\binom{n}{x} p^{x+1} q^{n-x+1}(n-x)} \left[\sum_{k=x+1}^{n} x\binom{n}{k} p^k q^{n-k+1} - \sum_{k=x}^{n} (n-x)\binom{n}{k} p^{k+1} q^{n-k} \right].$$

Let $\xi_2(x) = \sum_{k=x+1}^{n} x\binom{n}{k} p^k q^{n-k+1} - \sum_{k=x}^{n} (n-x)\binom{n}{k} p^{k+1} q^{n-k}$, then

$$\xi_2(x) = \sum_{k=x+1}^{n} \frac{x(n-k+1)}{n+1}\binom{n+1}{k} p^k q^{n-k+1} - \sum_{k=x}^{n} \frac{(k+1)(n-x)}{n+1}\binom{n+1}{k+1} p^{k+1} q^{n+1-(k+1)}$$

$$= \sum_{k=x+1}^{n+1} \frac{x(n-k+1)}{n+1}\binom{n+1}{k} p^k q^{n-k+1} - \sum_{k=x+1}^{n+1} \frac{k(n-x)}{n+1}\binom{n+1}{k} p^k q^{n+1-k}$$

$$= \sum_{k=x+1}^{n+1} \frac{\binom{n+1}{k} p^k q^{n-k+1}}{n+1}\left[x(n+1)-kn\right]$$

$$< 0.$$

Thus for $x_0 < x$,

$$\Delta f_{x_0}(x) < 0. \tag{2.12}$$

Following (2.11) and (2.12), when $x_0 \neq x$, it yields

$$\Delta f_{x_0}(x) < 0. \tag{2.13}$$

Hence, from (2.10) and (2.13), the inequalities in (2.8) hold.

Next, we shall show that (2.9) holds. For $x_0 > x$, it follows from (2.7) that

$$\Delta f_{C_{x_0}}(x) = \frac{\sum_{j=x_0+1}^{n} \binom{n}{j} p^j q^{n-j}}{x\binom{n}{x} p^{x+1} q^{n-x+1}(n-x)} \left[\sum_{k=0}^{x} x\binom{n}{k} p^k q^{n-k+1} - \sum_{k=0}^{x-1} (n-x)\binom{n}{k} p^{k+1} q^{n-k} \right]$$

$$= \frac{\sum_{j=x_0+1}^{n} \binom{n}{j} p^j q^{n-j}}{x\binom{n}{x} p^{x+1} q^{n-x+1}(n-x)} \sum_{k=0}^{x} \frac{\binom{n+1}{k} p^k q^{n-k+1}}{n+1}\left[x(n+1)-kn\right]$$

$$> 0. \tag{2.14}$$

For $x_0 < x$, by (2.7), we obtain

$$\Delta f_{C_{x_0}}(x) = \frac{\sum_{k=0}^{x_0}\binom{n}{k}p^k q^{n-k}}{\binom{n}{x}p^x q^{n-x}}\left[\frac{\sum_{k=x+1}^{n}\binom{n}{k}p^k q^{n-k}}{(n-x)p} - \frac{\sum_{k=x}^{n}\binom{n}{k}p^k q^{n-k}}{xq}\right]$$

$$= \frac{\sum_{j=0}^{x_0}\binom{n}{j}p^j q^{n-j}}{x\binom{n}{x}p^{x+1}q^{n-x+1}(n-x)}\left[\sum_{k=x+1}^{n} x\binom{n}{k}p^k q^{n-k+1} - \sum_{k=x}^{n}(n-x)\binom{n}{k}p^{k+1}q^{n-k}\right]$$

$$= \frac{\sum_{j=0}^{x_0}\binom{n}{j}p^j q^{n-j}}{x\binom{n}{x}p^{x+1}q^{n-x+1}(n-x)}\sum_{k=x+1}^{n+1}\frac{\binom{n+1}{k}p^k q^{n-k+1}}{n+1}\left[x(n+1)-kn\right]$$

$$< 0. \qquad (2.15)$$

Hence, from (2.14) and (2.15), the inequalities in (2.9) hold.

Lemma 2.2. *For* $x_0 \in \{1,\ldots,n-1\}$, $\Delta f_{C_{x_0}}$ *is an increasing function in* $x \in \{1,\ldots,x_0\}$.

Proof. Let $\Delta^2 f_{C_{x_0}}(x) = \Delta f_{C_{x_0}}(x+1) - \Delta f_{C_{x_0}}(x)$. We shall show that $\Delta^2 f_{C_{x_0}}(x) > 0$ for $x \in \{1,\ldots,x_0-1\}$. It follows from (2.7) that

$$\Delta^2 f_{C_{x_0}}(x) = \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x+1}p^{x+1}q^{n-x-1}}\left[\frac{\sum_{k=0}^{x+1}\binom{n}{k}p^k q^{n-k}}{(n-x-1)p} - \frac{\sum_{k=0}^{x}\binom{n}{k}p^k q^{n-k}}{(x+1)q}\right]$$

$$- \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^x q^{n-x}}\left[\frac{\sum_{k=0}^{x}\binom{n}{k}p^k q^{n-k}}{(n-x)p} - \frac{\sum_{k=0}^{x-1}\binom{n}{k}p^k q^{n-k}}{xq}\right]$$

$$= \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{(x+1)\binom{n}{x+1}p^{x+2}q^{n-x}(n-x-1)}\sum_{k=0}^{x+1}\frac{\binom{n+1}{k}p^k q^{n-k+1}}{n+1}\left[(x+1)(n+1)-kn\right]$$

$$- \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{x\binom{n}{x}p^{x+1}q^{n-x+1}(n-x)}\sum_{k=0}^{x}\frac{\binom{n+1}{k}p^k q^{n-k+1}}{n+1}\left[x(n+1)-kn\right]$$

$$= \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x+2}q^{n-x}(n-x)(n+1)} \sum_{k=0}^{x+1}\binom{n+1}{k}p^k q^{n-k+1}\left[\frac{(x+1)(n+1)-kn}{n-x-1}\right]$$

$$-\frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x+2}q^{n-x}(n-x)(n+1)} \sum_{k=0}^{x}\binom{n+1}{k}p^{k+1} q^{n+1-(k+1)}\left[\frac{x(n+1)-kn}{x}\right]$$

$$= \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x+2}q^{n-x}(n-x)(n+1)} \left\{\sum_{k=0}^{x+1}\binom{n+1}{k}p^k q^{n-k+1}\left[\frac{(x+1)(n+1)-kn}{n-x-1}\right]\right.$$

$$\left. -\sum_{k=0}^{x}\binom{n+1}{k+1}\frac{k+1}{n+1-k}p^{k+1} q^{n+1-(k+1)}\left[\frac{x(n+1)-kn}{x}\right]\right\}$$

$$= \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x+2}q^{n-x}(n-x)(n+1)} \left\{\sum_{k=0}^{x+1}\binom{n+1}{k}p^k q^{n-k+1}\left[\frac{(x+1)(n+1)-kn}{n-x-1}\right]\right.$$

$$\left. -\sum_{k=0}^{x+1}\binom{n+1}{k}\frac{k}{n+2-k}p^{k} q^{n+1-k}\left[\frac{x(n+1)-(k-1)n}{x}\right]\right\}$$

$$= \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x+2}q^{n-x}(n-x)(n+1)} \sum_{k=0}^{x+1}\binom{n+1}{k}p^k q^{n-k+1}\left\{\left[\frac{x(n+1)-kn+n+1}{n-x-1}\right]\right.$$

$$\left. -\left[\frac{x(n+1)-kn+n}{n+2-k}\right]\frac{k}{x}\right\}$$

$$> 0.$$

Therefore $\Delta f_{C_{x_0}}(x)$ is an increasing function in $x \in \{1,\ldots,x_0\}$.

Lemma 2.3. *Let* $x \in \{1,\ldots,n\}$, *then we have the following inequalities hold:*

$$\sup_x \left|\Delta f_{C_0}(x)\right| \le \frac{1-q^n}{np} \quad \text{(\textit{Teerapabolarn and Wongkasem, 2011})} \tag{2.16}$$

and

$$\sup_x \left|\Delta f_{C_{x_0}}(x)\right| \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}, \tag{2.17}$$

where $x_0 \in \{1,\ldots,n\}$.

Proof. We shall show that the inequality (2.17) holds. Teerapabolarn and Wongkasem (2011) showed that

$$\Delta f_{x_0}(x_0) \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}. \tag{2.18}$$

From (2.6), we have

$$\Delta f_x(x) = \frac{\sum_{k=x+1}^{n}\binom{n}{k}p^k q^{n-k}}{(n-x)p} + \frac{\sum_{k=0}^{x-1}\binom{n}{k}p^k q^{n-k}}{xq}.$$

Substituting x_0 by x in the proof of $\Delta f_{x_0}(x_0)$ detailed in Teerapabolarn and Wongkasem (2011), we also obtain

$$\Delta f_x(x) \le \min\left\{\frac{1-p^n}{xq}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}. \tag{2.19}$$

In order to prove that $\sup_x\left|\Delta f_{C_{x_0}}(x)\right| \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)p}\right\}$, it suffices to show $\left|\Delta f_{C_{x_0}}(x)\right| \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)p}\right\}$ for every $x \in \{1,\dots,n-1\}$, because $\Delta f_{C_{x_0}}(x) = 0$ as $x = 0, n.$ For $1 \le x \le x_0$, we have

$$\begin{aligned}
0 &< \Delta f_{C_{x_0}}(x) && \text{(by (2.9))}\\
&\le \Delta f_{C_{x_0}}(x_0) && \text{(by Lemma 2.2)}\\
&= \sum_{k=0}^{x_0}\Delta f_k(x_0)\\
&= \Delta f_0(x_0) + \cdots + \Delta f_{x_0}(x_0)\\
&\le \Delta f_{x_0}(x_0) && \text{(by (2.8))}\\
&\le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}, && \text{(by (2.18))}
\end{aligned}$$

which gives

$$\left|\Delta f_{C_{x_0}}(x)\right| \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}. \tag{2.20}$$

For $x_0 < x \le n-1$, we obtain

$$\begin{aligned}
0 &< -\Delta f_{C_{x_0}}(x) && \text{(by (2.9))}\\
&= -\sum_{k=0}^{x_0}\Delta f_k(x) - \sum_{k=x_0+1}^{x}\Delta f_k(x) + \sum_{k=x_0+1}^{x}\Delta f_k(x)\\
&= -\sum_{k=0}^{x}\Delta f_k(x) + \Delta f_{x_0+1}(x) + \cdots + \Delta f_x(x)\\
&= -\Delta f_{C_x}(x) + \Delta f_{x_0+1}(x) + \cdots + \Delta f_x(x)\\
&\le \Delta f_x(x) && \text{(by (2.8) and (2.9))}
\end{aligned}$$

$$\leq \min\left\{\frac{1-p^n}{xq}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} \qquad \text{(by (2.19))}$$

$$\leq \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\},$$

which gives

$$\left|\Delta f_{C_{x_0}}(x)\right| \leq \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}. \tag{2.21}$$

Hence, by (2.20) and (2.21), the inequality (2.17) holds.

2.2 The characterization of beta binomial random variable

For the characterization associated with the beta binomial random variable $Y,$ by applying Lemma 3.1 in Cacoullos and Papathanasiou (1989), the covariance of Y and $f_{C_{x_0}}(Y)$ can be expressed as

$$Cov\left[Y, f_{C_{x_0}}(Y)\right] = \sum_{y=0}^{n}\left[\Delta f_{C_{x_0}}(y)\sum_{k=0}^{y}(\mu-k)bb(k)\right], \tag{2.22}$$

where $\mu = \frac{n\alpha}{\alpha+\beta}$.

Lemma 2.4. *Let* $\varphi(y) = \frac{\sum_{k=0}^{y}(\mu-k)bb(k)}{bb(y)}$, $y = 0,...,n,$ *be the characterization associated with the beta binomial random variable* $Y,$ *then we have the following.*

$$\varphi(y) = \tfrac{(n-y)(\alpha+y)}{\alpha+\beta}, \; y = 0,...,n, \tag{2.23}$$

and

$$Cov\left[Y, f_{C_{x_0}}(Y)\right] = E\left[\Delta f_{C_{x_0}}(Y)\tfrac{(n-Y)(\alpha+Y)}{\alpha+\beta}\right]. \tag{2.24}$$

Proof. We shall show that $\varphi(y) = \frac{(n-y)(\alpha+y)}{\alpha+\beta}$ by mathematical induction as follows. It can be seen that $\varphi(0) = \mu = \frac{n\alpha}{\alpha+\beta}$ and $\varphi(1) = \frac{\sum_{k=0}^{1}(\mu-k)bb(k)}{bb(1)} = \frac{(n-1)(\alpha+1)}{\alpha+\beta}$. For $2 \leq m < n,$ let

$\varphi(m) = \frac{\sum_{k=0}^{m}(\mu-k)bb(k)}{bb(m)} = \frac{(n-m)(\alpha+m)}{\alpha+\beta}$, we have to show that $\varphi(m+1) = \frac{(n-(m+1))(\alpha+(m+1))}{\alpha+\beta}$. Since

$$\varphi(m+1) = \frac{\sum_{k=0}^{m+1}(\mu-k)bb(k)}{bb(m+1)}$$

$$= \frac{bb(m)}{bb(m+1)}\frac{\sum_{k=0}^{m}(\mu-k)bb(k)}{bb(m)} + \mu-(m+1)$$

$$=\frac{(m+1)(n+\beta-m-1)}{\alpha+\beta}+\mu-(m+1) \text{ (by } \frac{\sum_{k=0}^{m}(\mu-k)bb(k)}{bb(m)}=\frac{(n-m)(\alpha+m)}{\alpha+\beta}\text{)}$$

$$=\frac{(n-(m+1))(\alpha+(m+1))}{\alpha+\beta},$$

thus by mathematical induction, (2.23) is obtained.

Substituting (2.23) into (2.22), it becomes

$$Cov(Y, f_{C_{x_0}}(Y)) = \sum_{y=0}^{n}\left[\Delta f_{C_{x_0}}(y)\tfrac{(n-y)(\alpha+y)}{\alpha+\beta}bb(y)\right]$$

$$= E\left[\Delta f_{C_{x_0}}(Y)\tfrac{(n-Y)(\alpha+Y)}{\alpha+\beta}\right],$$

this yields (2.24).

3. Results

The main point of this study is to determine a non-uniform bound for the distance between the binomial and beta binomial cumulative distribution functions, $d_{C_{x_0}}(Y,X)$ as $x_0 \in \{0,...,n\}$. The following theorem presents this desired result.

Theorem 3.1. *Let* $p=\frac{\alpha}{\alpha+\beta}$ *and* $x_0 \in \{0,...,n\}$, *then we have the following.*

$$d_{C_{x_0}}(Y,X) \le \begin{cases} \frac{(1-q^n)(n-1)q}{\alpha+\beta+1} & if\ x_0=0, \\ \min\left\{\frac{1-p^n}{x_0}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)p}\right\}\frac{(n-1)np}{\alpha+\beta+1} & if\ 1\le x_0\le n. \end{cases} \tag{2.25}$$

Proof. Using the same arguments detailed as in the proof of Theorem 2.1 in Teerapabolarn (2008), it follows that

$$d_{C_{x_0}}(Y,X) \le \left|E\left[\mu\Delta f_{C_{x_0}}(Y)\right]-E\left[pY\Delta f_{C_{x_0}}(Y)\right]-Cov\left[Y, f_{C_{x_0}}(Y)\right]\right|$$

$$=\left|E\left[(\mu-pY)\Delta f_{C_{x_0}}(Y)\right]-E\left[\Delta f_{C_{x_0}}(Y)\tfrac{(n-Y)(\alpha+Y)}{\alpha+\beta}\right]\right| \quad \text{(by (2.24))}$$

$$=\left|E\left\{\left[\mu-pY-\tfrac{(n-Y)(\alpha+Y)}{\alpha+\beta}\right]\Delta f_{C_{x_0}}(Y)\right\}\right|$$

$$\le E\left|\left[\mu-pY-\tfrac{(n-Y)(\alpha+Y)}{\alpha+\beta}\right]\Delta f_{C_{x_0}}(Y)\right|$$

$$=E\left|\tfrac{(Y-n)Y}{\alpha+\beta}\Delta f_{C_{x_0}}(Y)\right|$$

$$=E\left[\tfrac{(n-Y)Y}{\alpha+\beta}\left|\Delta f_{C_{x_0}}(Y)\right|\right]$$

$$\le \begin{cases} \frac{1-q^n}{np}E\left[\frac{(n-Y)Y}{\alpha+\beta}\right] & \text{if } x_0=0, \\ \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}E\left[\frac{(n-Y)Y}{\alpha+\beta}\right] & \text{if } 1\le x_0\le n. \end{cases} \quad \text{(by Lemma 2.3)}$$

Because $E\left[\frac{(n-Y)Y}{\alpha+\beta}\right]=\frac{n\mu-\sigma^2-\mu^2}{\alpha+\beta}=\frac{(n-1)n\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)}$, we have

$$d_{C_{x_0}}(Y,X)\le\begin{cases}\frac{1-q^n}{np}\frac{(n-1)n\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)} & \text{if } x_0=0,\\ \min\left\{\frac{1-p^n}{x_0q},\frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}\frac{(n-1)n\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)} & \text{if } 1\le x_0\le n.\end{cases}$$

Hence, the inequality (2.25) is obtained.

Remark. Since $\frac{1-q^n}{np}<\frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}$ and $\min\left\{\frac{1-p^n}{x_0q},\frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}\le\frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}$ when $x_0\in\{1,\ldots,n\}$, the bound in Theorem 3.1 is better than that mentioned in (1.4).

4. Numerical Examples

Teerapabolarn (2008) suggested the result in (1.4) to give a good approximation when $\frac{\alpha}{\beta}$ and $\frac{n}{\beta}$ are small. So, we provide two examples to illustrate the result in Theorem 3.1 by setting parameters n, α and β to satisfy this suggestion.

Example 4.1. Let $n=30$, $\alpha=10$ and $\beta=500$, then the numerical result in Theorem 3.1 is of the form

$$d_{C_{x_0}}(Y,X)\le\begin{cases}0.024922 & \text{if } x_0=0,\\ 0.025195 & \text{if } x_0=1,\\ \frac{0.033383}{x_0} & \text{if } x_0=2,\ldots,30.\end{cases}$$

It is better than the numerical result in (1.4),

$$d_{C_{x_0}}(Y,X)\le 0.025195,\ x_0=0,1,\ldots,30.$$

Example 4.2. Let $n=50$, $\alpha=100$ and $\beta=1000$, then the numerical result in Theorem 3.1 is of the form

$$d_{C_{x_0}}(Y,X)\le\begin{cases}0.040114 & \text{if } x_0=0,\\ 0.043294 & \text{if } x_0=1,2,3,4,\\ \frac{0.202295}{x_0} & \text{if } x_0=5,\ldots,50.\end{cases}$$

It is better than the numerical result in (1.4),

$$d_{C_{x_0}}(Y,X)\le 0.043294,\ x_0=0,1,\ldots,50.$$

The two examples are indicated that the result in Theorem 3.1 gives a good approximation when $\frac{n}{\beta}$ and $\frac{\alpha}{\beta}$ are small, especially, when $\frac{\alpha}{\beta}$ is small. Furthermore, these examples point out that the bound in Theorem 3.1 is better than that shown in (1.4).

5. Conclusions

The bound in this study, non-uniform bound, was determined by using Stein's method and the characterization of beta binomial random variable. It is appropriate to approximate the distance between the beta binomial cumulative distribution function with parameters $n,\ \alpha > 0$ and $\beta > 0$ and the binomial cumulative distribution function with parameters n and $\frac{\alpha}{\alpha+\beta}$, because it changes along $x_0 \in \{0,...,n\}$. In addition, by theoretical and numerical comparison, the result in this study is better than that mentioned in (1.4), and it gives a good binomial approximation when $\frac{n}{\beta}$ and $\frac{\alpha}{\beta}$ are small.

Acknowledgements

The authors would like to thank the anonymous referees for their useful comments and suggestions.

References

Barbour, A. D., Holst, L., & Janson, S. (1992). *Poisson approximation* (Oxford Studies in Probability 2). Oxford, England: Clarendon Press.

Cacoullos, T., & Papathanasiou, V. (1989). Characterization of distributions by variance
bounds. *Statistics and Probability Letters*, *7*(5), 351-356.

Ding, C., Rusmevichientong, P., & Topaloglu, H. (2014). Balancing revenues and repair Costs under partial information about product reliability. *Production and Operations Management*, *23*(11), 1899-1918.

Gupta, A. K., & Nadarajah, S. (2004). *Handbook of beta distribution and its applications*. New York, NY: Marcel Dekker.

Salem, S. A., & Abu El Azm, W. S. (2013). The sampling distribution of the maximum likelihood estimators for the parameters of beta-binomial distribution. *International Mathematical Forum*, *8*(26), 1263-1277.

Sant'Anna, Â. M. O., & Caten, C. S. (2012). Beta control charts for monitoring fraction data. *Expert Systems with Applications*, *39*(11), 10236-10243.

Stein, C. M. (1972). A bound for the error in normal approximation to the distribution of a sum of dependent random variables. In L. M. Le Cam, J. Neyman, & E. L. Scott (Eds.), *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (pp. 583-602), Berkeley, CA: University of California Press.

Stein, C. M. (1986). *Approximate computation of expectations* (Lecture Notes-Monograph Series 7). Hayward, CA: Institute of Mathematical Statistics.

Teerapabolarn, K. (2008). A bound on the binomial approximation to the beta binomial distribution. *International Mathematical Forum*, *3*(28), 1355-1358.

Teerapabolarn K., & Wongkasem, P. (2011). On pointwise binomial approximation by *w*-functions. *International Journal of Pure and Applied Mathematics*, *71*(1), 57-66.