On Pointwise Binomial Approximation for Independent Binomial Random Variables

K. Teerapabolarn

Department of Mathematics, Faculty of Science
Burapha University, Chonburi 20131, Thailand

Copyright © 2014 K. Teerapabolarn. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

We use Stein’s method and the binomial $w$-functions to give a uniform bound for the point metric between the distribution of a sum of $n$ independent binomial random variables, each with parameters $n_i$ and $p_i$, by a binomial distribution with parameters $m = \sum_{i=1}^{n} n_i$ and $p = \frac{1}{m} \sum_{i=1}^{n} n_i p_i$. When all $p_i$ are small or all $p_i$ are close to $p$, the result of the study gives a good binomial approximation.

Mathematics Subject Classification: 60F05, 60G50

Keywords: Binomial approximation, Binomial $w$-function, Stein’s method

1 Introduction

Let $X_1, ..., X_n$ be independently distributed binomial random variables, each with probability function $p_{X_i}(k) = \binom{n_i}{k} p_i^k q_i^{n_i-k}$ for $k \in \{0, ..., n_i; n_i \in \mathbb{N}\}$, and with mean $\mu_i = n_i p_i$ and variance $\sigma_i^2 = n_i p_i q_i$, where $q_i = 1 - p_i$. Let $S_n = \sum_{i=1}^{n} X_i$ and $B_{m,p}$ denote the binomial random variable with parameters $m = \sum_{i=1}^{n} n_i$ and $p = \frac{1}{m} \sum_{i=1}^{n} \mu_i = \frac{1}{m} \sum_{i=1}^{n} n_i p_i$. In this case, for $A \subseteq \{0, ..., m\}$, Teerapabolarn [3] used the Stein-Chen method and the binomial $w$-functions to give a uniform bound for the distance between the distributions of $S_n$ and $B_{m,p}$, $d_A(S_n, B_{m,p}) = |P(S_n \in A) - P(B_{m,p} \in A)|$, as follows:

$$d_A(S_n, B_{m,p}) \leq \frac{1 - p^n + q^n}{(m+1)pq} \sum_{i=1}^{n} |p_i - p| n_i p_i.$$  (1.1)
However, for \( A = \{x_0; x_0 \in \{0, ..., m\} \} \) and \( d_{x_0}(S_n, B_{m,p}) = |P(S_n = x_0) - P(B_{m,p} = x_0)| \), the result in (1.1) becomes
\[
d_{x_0}(S_n, B_{m,p}) \leq \frac{1 - p^{m+1} + q^{m+1}}{(m + 1)pq} \sum_{i=1}^{n} |p_i - p| n_i p_i \tag{1.2}
\]
for every \( x_0 \). It is observed that the bound is a uniform constant for the point metric \( d_{x_0}(S_n, B_{m,p}) \). With this situation, a non-uniform bound with respect to \( x_0 \) is required. In this paper, we focus on deriving a non-uniform bound for the point metric between the distribution of \( S_n \) and the distribution of \( B_{m,p} \), where \( x_0 \in \{0, ..., m\} \).

The tools for giving the desired result consist of Stein’s method and the binomial \( w \)-functions, which are in Section 2. In Section 3, we give a non-uniform bound for \( d_{x_0}(S_n, B_{m,p}) \), and the conclusion of this study is presented in the last section.

2 Method

The following lemma gives the binomial \( w \)-functions, which are directly obtained from [4].

**Lemma 2.1.** For \( 1 \leq i \leq n \), let \( w_i \) be the \( w \)-function associated with the binomial random variable \( X_i \), then we have the following:
\[
w_i(k) = \frac{(n_i - k)p_i}{\sigma_i^2}, \quad k \in \{0, ..., n_i\}. \tag{2.1}
\]

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

\[
Cov(S_n, f(S_n)) = \sum_{i=1}^{n} Cov \left( X_i, f \left( X_i + \sum_{j \neq i} X_j \right) \right)
= \sum_{i=1}^{n} \sigma_i^2 E[w_i(X_i) \Delta f(S_n)], \tag{2.2}
\]
for any function \( f : \mathbb{N} \cup \{0\} \to \mathbb{R} \) for which \( E|w_i(X_i)\Delta f(S_n)| < \infty \), where \( \Delta f(x) = f(x + 1) - f(x) \).

For Stein’s method in the binomial approximation, it can be applied for every \( m \in \mathbb{N} \) and \( 0 < p = 1 - q < 1 \), for every \( x_0 \in \{0, ..., m\} \) and bounded real-valued function \( f = f(x_0) : \mathbb{N} \cup \{0\} \to \mathbb{R} \) defined as in [1], where \( f(0) = f(1) \) and \( f(x) = f(m) \) for \( x \geq m \). So, Stein’s equation for these conditions is as follows:
\[
P(S_n - x_0) - P(B_{m,p} = x_0) = E[(m - S_n)p f(S_n + 1) - qS_n f(S_n)]. \tag{2.3}
\]
For \( x_0, x \in \mathbb{N} \cup \{0\} \), [5] showed that

\[
\sup_{x \geq 0} |\Delta f(x)| \leq \delta(x_0) = \begin{cases} 
\frac{1-p^m}{np} & \text{if } x_0 = 0, \\
\min \left\{ \frac{1-p^m}{x_0q}, \frac{1-p^{m+1} q^{m+1}}{(m+1)pq} \right\} & \text{if } x_0 > 0.
\end{cases}
\] (2.4)

3 Result

The following theorem presents a non-uniform bound on the error of a pointwise binomial approximation to the probability function of \( S_n \).

**Theorem 3.1.** For \( x_0 \in \{0, ..., m\} \), then we have the following:

\[
d_{x_0}(S_n, \mathcal{B}_{m,p}) \leq \delta(x_0) \sum_{i=1}^{n} |p_i - p| n_i p_i,
\] (3.1)

where \( \delta(x_0) \) is defined in (2.4).

**Proof.** From (2.3), it follows that

\[
d_{TV}(S_n, \mathcal{B}_{m,p}) = |E[(m-S_n)p f(S_n+1) - qS_n f(S_n)]|
\]
\[
= |E[m p f(S_n+1) - p S_n \Delta f(S_n) - S_n f(S_n)]|
\]
\[
= |E[m p \Delta f(S_n)] - p E[S_n \Delta f(S_n)] - Cov(S_n, f(S_n))|
\]
\[
= \left| \sum_{i=1}^{n} \{ E[\mu_i \Delta f(S_n)] - p E[X_i \Delta f(S_n)] - Cov(X, f(S_n)) \} \right|.
\]

Using (2.2) and Lemma 2.1, we have

\[
d_{TV}(S_n, \mathcal{B}_{m,p}) = \left| \sum_{i=1}^{n} \{ E[(\mu_i - p X_i) \Delta f(S_n)] - \sigma_i^2 E[w_i(X_i) \Delta f(S_n)] \} \right|
\]
\[
\leq \sum_{i=1}^{n} E|n_i p_i - p X_i - \sigma_i^2 w_i(X_i)||\Delta f(S_n)|
\]
\[
\leq \sup_{x \geq 0} |\Delta f(x)| \sum_{i=1}^{n} E |n_i p_i - p X_i - (n_i - X_i) p_i|
\]
\[
\leq \sup_{x \geq 0} |\Delta f(x)| \sum_{i=1}^{n} |p_i - p| n_i p_i.
\]

Hence, by (2.4), (3.2) is obtained. \( \square \)

If \( n_i = 1 \) for every \( i \in \{1, ..., n\} \), then \( S_n \) has the Poisson binomial distribution with parameter \( p = (p_1, ..., p_n) \). Thus, an immediately consequence of Theorem 3.1, a binomial approximation to the Poisson binomial distribution
is also obtained.

**Corollary 3.1.** For \( x_0 \in \{0, \ldots, m\} \), if \( n_1 = \cdots = n_n = 1 \), then the following inequality holds:

\[
d_{x_0}(S_n, B_{m,p}) \leq \delta(x_0) \sum_{i=1}^{n} |p_i - p| p_i. \tag{3.2}
\]

4 Conclusion

In this study, a non-uniform bound on the error of a pointwise binomial approximation to the probability function of a sum of \( n \) independent binomial random variables was derived by Stein’s method and the binomial \( w \)-functions. It gives a good approximation when all \( p_i \) are small or all \( p_i \) are close to \( p \). In addition, the bound in this study is sharper than that presented in (1.2).

References


**Received: December 3, 2014; Published: January 7, 2015**