A Bound on Poisson Approximation for Independent Binomial Random Variables

K. Teerapabolarn

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

Abstract

This paper uses the Stein-Chen method and the binomial $w$-functions to derive a non-uniform bound for approximating the cumulative distribution function of a sum of independent binomial random variables, each with parameters $m_i$ and $p_i$, by a cumulative Poisson distribution function with mean $\sum_{i=1}^{n} m_i p_i$. With this bound, the cumulative Poisson distribution function can be used as an approximation of the cumulative distribution function of the summands when all $p_i$ are small.

Mathematics Subject Classification: 60F05, 60G50

Keywords: Binomial distribution, Binomial $w$-function, Poisson approximation, Stein-Chen method

1 Introduction

Let $X_1, ..., X_n$ be independently distributed binomial random variables, each with probability $P(X_i = k) = \binom{m}{k} p_i^k q_i^{m_i-k}$ for $k \in \{0, 1, ..., m_i\}$, and mean $\mu_i = m_i p_i$ and variance $\sigma_i^2 = m_i p_i q_i$, where $0 < p_i = 1 - q_i < 1$ and $m_i \in \mathbb{N}$. Let $S_n = \sum_{i=1}^{n} X_i$ and $Z_\lambda$ denote the the Poisson random variable with mean $\lambda = \sum_{i=1}^{n} \mu_i$. If all $p_i$ are small, then the distribution of $S_n$ can be approximated by the Poisson distribution with mean $\lambda$. In this case, for $A \subseteq \mathbb{N} \cup \{0\}$, Teerapabolarn [5] used the Stein-Chen method and the binomial $w$-functions
to give a bound for the distance between the distributions of $S_n$ and $Z_\lambda$ as follows:

$$d_A(S_n, Z_\lambda) = |P(S_n \in A) - P(Z_\lambda \in A)| \leq \frac{1 - e^{-\lambda}}{\lambda} \sum_{i=1}^{n} m_i p_i^2.$$  \hspace{1cm} (1.1)

However, when $A = C_{x_0} = \{0, ..., x_0\}$, $x_0 \in \mathbb{N} \cup \{0\}$, the result in (1.1) becomes

$$d_{C_{x_0}}(S_n, Z_\lambda) = |P(S_n \leq x_0) - P(Z_\lambda \leq x_0)| \leq \frac{1 - e^{-\lambda}}{\lambda} \sum_{i=1}^{n} m_i p_i^2$$  \hspace{1cm} (1.2)

for every $x_0$. It can be seen that the bound in (1.2) is a uniform constant with respect to $x_0$. So, a non-uniform bound with respect to $x_0$ is required for this situation. In this paper, we are interested to derived a non-uniform bound for $d_{C_{x_0}}(S_n, Z_\lambda)$ by using the same tools in [5], which are described in Section 2. In Section 3, we use these tools to derive the desired result and the conclusion in this study is presented in the last section.

### 2 Method

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

**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:

$$w_i(k) = \frac{(m_i - k)p_i}{\sigma_i^2}, \quad k \in \{0, ..., m_i\}. \hspace{1cm} (2.1)$$

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

$$Cov(S_n, f(S_n)) = \sum_{i=1}^{n} Cov(X_i, f(X_i + \sum_{j \neq i} X_j)) = \sum_{i=1}^{n} \sigma_i^2 E[w_i(X_i)\Delta f(S_n)], \hspace{1cm} (2.2)$$

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 the Stein-Chen method, following [1], it can be applied for every constant $\lambda > 0$, for every $x_0 \in \mathbb{N} \cup \{0\}$ and the bounded real valued function $f = f_{C_{x_0}} : \mathbb{N} \cup \{0\} \to \mathbb{R}$. Thus, Stein’s equation for the Poisson distribution with these conditions is of the form

$$P(S_n \leq x_0) - P(Z_{\lambda n} \leq x_0) = E[\lambda f(S_n + 1) - S_n f(S_n)]. \hspace{1cm} (2.3)$$
For \( x_0 \in \mathbb{N} \cup \{0\} \) and \( x \in \mathbb{N} \), \([4]\) showed that
\[
\sup_{x \geq 1} |\Delta f(x)| \leq \min \left\{ \frac{1 - e^{-\lambda}}{\lambda}, \frac{1}{x_0} \right\}.
\] (2.4)

3 Result

The following theorem gives a non-uniform bound for the distance \( d_{C_{x_0}}(S_n, Z_\lambda) \).

**Theorem 3.1.** For \( x_0 \in \mathbb{N} \cup \{0\} \), then the following inequality holds:
\[
d_{C_{x_0}}(S_n, Z_\lambda) \leq \min \left\{ \frac{1 - e^{-\lambda}}{\lambda}, \frac{1}{x_0} \right\} \sum_{i=1}^{n} m_i p_i^2.
\] (3.1)

**Proof.** From (2.3), it follows that
\[
d_{C_{x_0}}(S_n, Z_\lambda) = |\lambda E[f(S_n + 1)] - E[S_n f(S_n)]| = |\lambda E[\Delta f(S_n)] - Cov(S_n, f(S_n))|
\]
\[
= \left| \sum_{i=1}^{n} \mu_i E[\Delta f(S_n)] - \sum_{i=1}^{n} Cov(X_i, f(S_n)) \right|.
\]

Using (2.2) and Lemma 2.1, we have
\[
d_{C_{x_0}}(S_n, Z_\lambda) = \left| \sum_{i=1}^{n} \{ E[\mu_i \Delta f(S_n)] - \sigma_i^2 E[w_i(X_i) \Delta f(S_n)] \} \right|
\]
\[
\leq \sum_{i=1}^{n} E \{|m_i p_i - \sigma_i^2 w_i(X_i)| \Delta f(S_n)|\}
\]
\[
\leq \sup_{x \geq 1} |\Delta f(x)| \sum_{i=1}^{n} E|m_i p_i - (m_i - X_i)p_i|
\]
\[
\leq \sup_{x \geq 1} |\Delta f(x)| \sum_{i=1}^{n} m_i p_i^2.
\]

Hence, by (2.4), (3.1) is obtained. \(\square\)

**Corollary 3.1.** For \( m_1 = \cdots = m_n = 1 \), then \( \lambda = \sum_{i=1}^{n} p_i \) and
\[
d_{C_{x_0}}(S_n, Z_\lambda) \leq \min \left\{ \frac{1 - e^{-\lambda}}{\lambda}, \frac{1}{x_0} \right\} \sum_{i=1}^{n} p_i^2.
\] (3.2)
The result (3.2) is a result in Poisson approximation to the distribution of a sum of independent Bernoulli random variables. When all $X_i$ are identically distributed random variables, thus immediately from the Theorem 3.1, we have the following Corollary.

**Corollary 3.2.** If $m_1 = \cdots = m_n = m$ and $p_1 = \cdots = p_n = p$, then $\lambda = nmp$ and the following inequality holds:

$$d_{C_{x_0}}(S_n, Z_\lambda) \leq \min \left\{ \left(1 - e^{-\lambda}\right), \frac{\lambda}{x_0} \right\} p.$$  \hfill (3.3)

### 4 Conclusion

In this study, a non-uniform bound for the distance between the cumulative distribution function of a sum of independent binomial random variables and a cumulative Poisson distribution function was derived by the Stein-Chen method and the binomial $w$-functions. In view of this bound, it can be seen that the cumulative distribution function of that summands can be approximated by a cumulative Poisson distribution function with mean $\lambda = \sum_{i=1}^{n} m_i p_i$ when all $p_i$ are small. In addition, the bound in this study is sharper than that presented in (1.2).

### References


**Received:** December 1, 2014; **Published:** January 3, 2015