International Journal of Mathematics and
Computer Science, **18**(2023), no. 3, 557–562

A non-uniform bound on Poisson approximation for a sum of independent non-negative integer-valued random variables

Pronsuda Kongkam, Kanint Teerapaolarn

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

email: Pronsuda.K1991@hotmail.com, kanint@buu.ac.th

(Received February 1, 2023, Accepted March 8, 2023,
Published March 31, 2023)

Abstract

In this paper, we use the Stein-Chen method and w-functions to determine a non-uniform bounds on Poisson approximation to the cumulative distribution function of a sum of independent non-negative integer-valued random variables. The obtained bound is more suitable for measuring the accuracy of this approximation.

1 Introduction

In the past few decades, mathematicians and statisticians have studied many research related to the context of Poisson approximation. These studies have yielded useful results that can be applied to probability and statistics. The most useful results are concerned with the Poisson approximation for sums of Bernoulli random variables and for sums of non-negative integer-valued random variables. From which, many authors have tried to obtain some good bounds for this approximation. In particular, many accurate results were created by the well-know Stein-Chen method, which was introduced by Chen[3]. The Stein-Chen method is a powerful and efficient

Key words and phrases: Non-uniform bound, cumulative distribution function, Poisson approximation, Stein-Chen method, w-functions.
AMS (MOS) Subject Classifications: 60F05, 60G50, 62E17.
The corresponding author is Kanint Teerapabolarn.
ISSN 1814-0432, 2023, http://ijmcs.future-in-tech.net

technique for finding uniform and non-uniform bounds on the Poisson approximation. In addition, this study also uses the method to determine a non-uniform bound on Poisson approximation for a sum of non-negative integer-valued random variables. Let $S = \sum_{i=1}^{n} X_i$, where $X_1, X_2..., X_n$ are independent non-negative integer-valued random variables, each with the probability $p_{X_i}(x) = P(X_i = x) > 0$ for every x in the space of X_i, $\mathcal{X}_i$, and have mean μ_i, finite variance $\sigma_i^2 > 0$ and w-function associated with X_i, $w_i(X_i)$, where $\sigma_i^2 w_i(x) p_{X_i}(x) = \sum_{j=0}^{x}(\mu_i - j) p_X(j)$ for every $x \in \mathcal{X}_i$. Let $Z = \sum_{i=1}^{n} Z_i$, where $Z_1, Z_2..., Z_n$ are independent Poisson random variables with mean $\lambda_1, \lambda_2, ..., \lambda_n$, respectively and $\mu = \sum_{i=1}^{n} \mu_i$. In this case, Teerapabolarn[6] used the Stein-Chen method and w-functions to give a uniform bound for approximating the distribution of S, $P(S \in A)$, by a Poisson distribution, $P(Z \in A)$, with mean $\lambda = \sum_{i=1}^{n} \lambda_i$ in the form of

$$d_A(S,Z) \leq \frac{1-e^{-\lambda}}{\lambda} \sum_{i=1}^{n} \mathbb{E}\left|\lambda_i - \sigma_i^2 w_i(X_i)\right| + \sqrt{\frac{2}{\lambda e}} |\lambda - \mu| \left(1 - p_S(0)\right), \quad (1.1)$$

and if $\lambda = \mu$, then

$$d_A(S,Z) \leq \frac{1-e^{-\lambda}}{\lambda} \sum_{i=1}^{n} \mathbb{E}\left|\lambda_i - \sigma_i^2 w_i(X_i)\right| \quad (1.2)$$

for $A \subseteq \mathbb{N} \cup \{0\}$, where $d_A(S,Z) = |P(S \in A) - P(Z \in A)|$. It is noted that the result gives a good approximation when $\lambda = \mu$ and for $A = C_{x_0} = \{0, ..., x_0\}$ as $x_0 \in \mathcal{S}$, where $\mathcal{S}$ is the space of S. Thus, (1.2) becomes

$$d_{C_{x_0}}(S,Z) \leq \frac{1-e^{-\lambda}}{\lambda} \sum_{i=1}^{n} \mathbb{E}\left|\lambda_i - \sigma_i^2 w_i(X_i)\right|, \quad (1.3)$$

where $d_{C_{x_0}}(S,Z) = |P(S \leq x_0) - P(Z \leq x_0)|$. In this paper, we focus on improving the uniform bound in (1.3) to be a non-uniform bound, which is a better bound for measuring the accuracy of this approximation.

2 Method

The method of this study consists of the Stein-Chen method and w-functions. For w-functions that are satisfying to $X_1, X_2, ..., X_n$, it can be applied the result in Cacoullos and Papathanasiou [2] to define a function w_i associated with non-negative integer-valued random variable X_i in the relation

$$\sigma_i^2 w_i(x) p_{X_i}(x) = \sum_{j=0}^{x} (\mu_i - j) p_{X_i}(j), \quad x \in \mathcal{X}_i \tag{2.4}$$

and according to [2], it yields

$$Cov\big(X_i, g(X_i)\big) = \sigma_i^2 \mathbb{E}\big[w_i(X_i) \Delta g(X_i)\big], \tag{2.5}$$

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

For the stein-Chen method, we first need to consider Stein's equation for the Poisson distribution with parameter $\lambda > 0$ is, for given h, of the form

$$h(x) - \mathcal{P}_\lambda(h) = \lambda g(x+1) - x g(x), \tag{2.6}$$

where $\mathcal{P}_\lambda(h) = E[h(X)]$ and g and h are bounded real valued functions defined on $\mathbb{N} \cup \{0\}$. For $A \subseteq \mathbb{N} \cup \{0\}$, let function $h_A : \mathbb{N} \cup \{0\} \to \mathbb{R}$ be defined by

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

Following Barbour et al. [1], the solution g_A of (2.6) is of the form

$$g_A(x) = \begin{cases} (x-1)!\lambda^{-x} e^{\lambda}[\mathcal{P}_\lambda(h_{A \cap C_{x-1}}) - \mathcal{P}_\lambda(h_A)\mathcal{P}_\lambda(h_{C_{x-1}})] & \text{if } x \geq 1, \\ 0 & \text{if } x = 0. \end{cases} \tag{2.7}$$

Similarly, for $A = C_{x_0} = \{0, ..., x_0\}$ where $x_0 \in \mathbb{N} \cup \{0\}$, $g_{C_{x_0}}$ can be expressed as

$$g_{C_{x_0}}(x) = \begin{cases} (x-1)!\lambda^{-x} e^{\lambda}[\mathcal{P}_\lambda(h_{C_{x-1}})\mathcal{P}_\lambda(1 - h_{C_{x_0}})] & \text{if } x \leq x_0, \\ (x-1)!\lambda^{-x} e^{\lambda}[\mathcal{P}_\lambda(h_{C_{x_0}})\mathcal{P}_\lambda(1 - h_{C_{x-1}})] & \text{if } x > x_0. \\ 0 & \text{if } x = 0. \end{cases} \tag{2.8}$$

For $x_0 \in \mathbb{N} \cup \{0\}$, let $\Delta g_{C_{x_0}}(x) = g_{C_{x_0}}(x+1) - g_{C_{x_0}}(w)$. The following lemmas are also need to prove the main result.

Lemma 2.1. *With the above definitions of w-function and for each $i \in \{1, ..., n\}$, we then have the following:*

$$Cov\big[X_i, g(S)\big] = \sigma_i^2 \mathbb{E}\big[w_i(X_i) \Delta g(S)\big]. \tag{2.9}$$

Proof. It suffices to show (2.9) for $n = 2$. Following the fact that

$$\begin{aligned} Cov\big[X_1, g(X_1+X_2)\big] &= \mathbb{E}\big[Cov\big(X_1, g(X_1+X_2)\big)|X_2\big] \\ &= \mathbb{E}\big\{\sigma_1^2\mathbb{E}\big[\big(w_1(X_1)\Delta g(X_1+X_2)\big)|X_2\big]\big\} \text{ (by (2.5))} \\ &= \sigma_1^2\mathbb{E}\big[w_1(X_1)\Delta g(X_1+X_2)\big], \end{aligned}$$

this implies (2.9). □

Because $Cov\big[S, g(S)\big] = \sum_{i=1}^{n} Cov\big[X_i, g(S)\big]$ and by applying (2.9), we have

$$Cov\big[S, g(S)\big] = \sum_{i=1}^{n} \sigma_i^2\mathbb{E}\big[w_i(x_i)\Delta g(S)\big]. \tag{2.10}$$

Lemma 2.2. *For $x_0 \in \mathbb{N}\cup\{0\}$ and $x \in \mathbb{N}$, then the following inequality holds:*

$$\big|\Delta g_{C_{x_0}}(x)\big| \le \begin{cases} \lambda^{-2}(e^{-\lambda} - \lambda - 1) & if \quad x_0 = 0, \\ \min\left\{\lambda^{-1}(1-e^{-\lambda}),\ \dfrac{1}{x_0+1}\right\} & if \quad x_0 > 0. \end{cases}$$

Proof. For $x_0 = 0$, it is directly obtained from Lemma 2.1 in Teerapabolarn and Neammanee [8]. Thus, we have to show that $\big|\Delta g_{C_{x_0}}(x)\big| \le \dfrac{1}{x_0+1}$ for $x_0 > 0$. For $x \le x_0$, Teerapabolarn [7] showed that $0 < \Delta g_{C_{x_0}}(x) \le \Delta g_{C_{x_0}}(x_0) = \Delta g_{C_{x_0-1}}(x_0) + \Delta g_{\{x_0\}}(x_0)$, where $\Delta g_{\{x_0\}}(x) = g_{\{x_0\}}(x+1) - g_{\{x_0\}}(x)$. Using (2.8) and combining Teerapabolarn [7] and Stein [4], we have that

$$
\begin{aligned}
\left|\Delta g_{C_{x_0}}(x)\right| &\leq (x_0-1)!\mathcal{P}_\lambda(h_{C_{x_0-1}})\left[\sum_{k=x_0+1}^{\infty}(x_0-k)\frac{\lambda^{k-(x_0+1)}}{k!}\right] \\
&\quad + e^{-\lambda}\left[\frac{1}{\lambda}\sum_{k=x_0+1}^{\infty}\frac{\lambda^k}{k!}+\frac{1}{x_0}\sum_{k=0}^{x_0-1}\frac{\lambda^k}{k!}\right] \\
&= -\frac{\mathcal{P}_\lambda(h_{C_{x_0-1}})}{x_0}\left[\frac{1}{x_0+1}+\frac{2\lambda}{(x_0+2)(x_0+1)}+\dots\right]+\frac{\mathcal{P}_\lambda(h_{C_{x_0-1}})}{x_0} \\
&\quad + \frac{e^{-\lambda}}{\lambda}\sum_{k=x_0+1}^{\infty}\frac{\lambda^k}{k!} \\
&\leq \left(1-\frac{1}{x_0+1}\right)\frac{\mathcal{P}(h_{C_{x_0-1}})}{x_0}+\frac{e^{-\lambda}}{\lambda}\sum_{k=x_0+1}^{\infty}\frac{\lambda^k}{k!} \\
&\leq \frac{1}{x_0+1}\mathcal{P}(h_{C_{x_0-1}})+\frac{1}{x_0+1}\mathcal{P}(1-h_{C_{x_0-1}}) \\
&= \frac{1}{x_0+1}. \qquad (2.11)
\end{aligned}
$$

For $x > x_0$, Teerapabolarn [7] and Barbour et al. [1] showed that

$$0 > \Delta g_{C_{x_0}}(x) \geqslant -\Delta g_{\{x\}}(x) \geq -\min\left\{\frac{1-e^{-\lambda}}{\lambda}, \frac{1}{x+1}\right\},$$

which yields

$$\left|\Delta g_{C_{x_0}}(x)\right| \leq \min\left\{\frac{1-e^{-\lambda}}{\lambda}, \frac{1}{x_0+1}\right\}. \qquad (2.12)$$

Hence, from (2.11) and (2.12), the inequality in Lemma 2.2 holds. □

3 Main Result

The following theorem presents the our main result of this study.

Theorem 3.1. *For $\lambda = \mu$, the following inequality holds:*

$$d_{C_{x_0}}(S, Z) \leq \delta(x_0)\sum_{i=1}^{n}\mathbb{E}\left|\lambda_i - \sigma_i^2 w_i(X_i)\right|, \qquad (3.13)$$

$$\text{where } \delta(x_0) = \begin{cases} \lambda^{-2}(e^{-\lambda} - \lambda - 1) & if \quad x_0 = 0, \\ \min\left\{\lambda^{-1}(1 - e^{-\lambda}),\ \dfrac{1}{x_0+1}\right\} & if \quad x_0 > 0. \end{cases}$$

Proof. We shall show that (3.13) holds. By following (2.6) and using Teerapabolarn [5] and setting $g = g_{C_{x_0}}$, we have

$$\begin{aligned} d_{C_{x_0}}(S, Z) &= |\lambda \mathbb{E}[g(S+1)] - \mathbb{E}[Sg(S)]| \\ &= |\lambda \mathbb{E}[\Delta g(S)] - \mathbb{C}ov(S, g(S)) + (\lambda - \mu)\mathbb{E}[Sg(S)]| \\ &= \left|\sum_{i=1}^{n} \lambda_i \mathbb{E}[\Delta g(S)] - \sum_{i=1}^{n} \sigma_i^2 \mathbb{E}[w_i(x_i)\Delta g(S)]\right| \text{ (by (2.10))} \\ &\le \sup_{s \in \mathcal{S}} |\Delta g(s)| \sum_{i=1}^{n} \mathbb{E}\left|\lambda_i - \sigma_i^2 w_i(X_i)\right| \\ &\le \delta(x_0) \sum_{i=1}^{n} \mathbb{E}\left|\lambda_i - \sigma_i^2 w_i(X_i)\right| \text{ (by Lemma 2.2).} \end{aligned}$$

Therefore, the inequality (3.13) holds. □

4 Conclusion

In this study, a non-uniform bound for the distance between the cumulative distribution of a sum of independent non-negative integer-valued random variables and a Poisson cumulative distribution function was obtained by using the Stein-Chen method and w-functions. By comparing the two related bounds, the bound of this study is sharper than that reported in (1.3). From which, the obtained bound is more suitable for measuring the accuracy of this approximation.

References

[1] A. D. Barbour, L. Holst, S. Janson, Poisson Approximation, Oxford Studies in Probability 2, 1992.

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

[3] L. H. Y. Chen, Poisson approximation for dependent trials, Ann. Probab., **3,** (1975), 534–545.

[4] C. M. Stein, Approximate Computation of Expectations, Institute of Mathematicl Statistics, Hayward, California, 1986.

[5] K. Teerapabolarn, A non-uniform bound on Poisson approximation by w-functions, Int. J. Pure Appl. Math., **78**, (2012), 1063–1075.

[6] K. Teerapabolarn, An extension of Poisson approximation by w-functions, Int. J. Pure Appl. Math., **87**, (2013), 529–534.

[7] K. Teerapabolarn, An improvement of Poisson approximation for sums of dependent Bernoulli random variables, Commun. Stat. Theory Methods, **43**, (2014), 17581777.

[8] K. Teerapabolarn, K. Neammanee, A non-uniform bound in somatic cell hybrid model, Math. BioSci., **195**, (2005), 56–64.