New non-uniform bounds on Poisson approximation for dependent Bernoulli trials

K. Teerapabolarn
Department of Mathematics, Faculty of Science
Burapha University, Chonburi 20131, Thailand
Centre of Excellence in Mathematics, CHE
Sri Ayutthaya Road, Bangkok 10400, Thailand
E-mail : kanint@buu.ac.th

Abstract

The aim of this article is a use of the Stein-Chen method to obtain new non-uniform bounds on the error of the distribution of sums of dependent Bernoulli random variables and the Poisson distribution. The bounds obtained in this study are improved to be more appropriate for measuring the accuracy of Poisson approximation. Examples are provided to illustrate applications of the obtained results.

Keywords: Bernoulli random variable, Poisson approximation, non-uniform bound, Stein-Chen method.

Mathematics Subject Classification: 60F05, 60G05.

1 Introduction

It is well-known that much methodological research on topics related to the Poisson approximation have yielded useful results in applied probability and statistics, and the most valuable findings have concerned the Poisson approximation for sums of independent and dependent Bernoulli random variables. For the independent case, the distribution of sums of n independent Bernoulli random variables is usually referred to as the distribution of the number of successes in a sequence of n independent Bernoulli trials, where success occurs on the i^{th} trial with a probability of $p_i \in (0, 1)$, and failure occurs on the i^{th} trial with a probability of $q_i = 1 - p_i$. This distribution is always called the Poisson binomial distribution with parameter $\mathbf{p} = (p_1, ..., p_n)$. When all p_i are identical and equal to p, the distribution reduces to the binomial distribution with parameters n and p. Similarly, the distribution of a sum of n Bernoulli random variables can also be considered as the distribution of the number of successes in a sequence of n dependent Bernoulli trials for the dependent case. In the past few years, some mathematicians and statisticians have developed a powerful technique known as the Stein-Chen method for approximating the distribution of a sum of Bernoulli random variables, such as Chen [7], Stein [15], Arratia et al. [1, 2], Barbour et al. [5], Neammanee [13], Teerapabolarn and Neammanee [17], Teerapabolarn and Neammanee [19], Teerapabolarn and Santiwipanont [20] and for approximating the specific distribution appeared in Teerapabolarn [21]. In contrast to many asymptotic methods, this approximation carries with it explicit error bounds as follows.

Suppose Γ is an arbitrary finite index set of size $|\Gamma|$. For each $\alpha \in \Gamma$, let X_α be a Bernoulli random variables with success probability $P(X_\alpha = 1) = 1 - P(X_\alpha = 0) = p_\alpha$, and let $W = \sum_{\alpha\in\Gamma} X_\alpha$ and $\lambda = E(W) = \sum_{\alpha\in\Gamma} p_\alpha$. It is well-known that the distribution of W can be approximated by the Poisson distribution with mean λ when the probabilities p_α's are sufficiently small. In recent years, numerous authors have sought to propose a good error bound for measuring the accuracy of this approximation. Many accurate results are derived from the well-known Stein-Chen method as proposed by Chen [7]. For example, when all X_α are independent and $\lambda = \sum_{\alpha\in\Gamma} p_\alpha$, Stein [15] gave an explicit uniform bound for the difference of the distribution of W and the Poisson distribution with mean λ as follows:

$$\left| P(W \in A) - \sum_{k\in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \lambda^{-1}(1 - e^{-\lambda}) \sum_{\alpha\in\Gamma} p_\alpha^2, \tag{1.1}$$

where $A \subseteq \mathbb{N} \cup \{0\}$. For $A = \{w_0\},\ w_0 \in \{1, ..., |\Gamma| - 1\}$, Neammanee [13] gave a non-uniform bound

$$\left| P(W = w_0) - \frac{\lambda^{w_0} e^{-\lambda}}{w_0!} \right| \leq \min\left\{ \frac{1}{w_0}, \lambda^{-1} \right\} \sum_{\alpha\in\Gamma} p_\alpha^2 \tag{1.2}$$

for the point metric between the probability function of W and the Poisson probability function with mean λ. For $A = \{0, ..., w_0\},\ w_0 \in \{0, 1, ..., |\Gamma|\}$, Teerapabolarn and Neammanee [19] gave a non-uniform bound

$$\left| P(W \leq w_0) - \sum_{k=0}^{w_0} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \lambda^{-1}(1 - e^{-\lambda}) \min\left\{ 1, \frac{e^\lambda}{w_0 + 1} \right\} \sum_{\alpha\in\Gamma} p_\alpha^2 \tag{1.3}$$

for approximating the cumulative distribution function of W by the Poisson cumulative distribution function with the same mean. For $A \subseteq \{0, ..., |\Gamma|\}$, Teerapabolarn and Santiwipanont [20] gave a non-uniform bound for the distance between the distribution of W and the Poisson distribution with this mean as follows:

$$\left| P(W \in A) - \sum_{k\in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \lambda^{-1} \min\left\{ 1, \lambda, \frac{\Delta(\lambda)}{M_A + 1} \right\} \sum_{\alpha\in\Gamma} p_\alpha^2, \tag{1.4}$$

where

$$\Delta(\lambda) = \begin{cases} e^\lambda + \lambda - 1 & \text{if } \lambda^{-1}(e^\lambda - 1) \leq M_A, \\ 2(e^\lambda - 1) & \text{if } \lambda^{-1}(e^\lambda - 1) > M_A, \end{cases}$$

and for $C_w = \{0, ..., w\}$,

$$M_A = \begin{cases} \max\{w | C_w \subseteq A\} & \text{if } 0 \in A, \\ \min\{w | w \in A\} & \text{if } 0 \notin A. \end{cases}$$

In the case of dependent Bernoulli summands, we first suppose that, for each $\alpha \in \Gamma$, a neighborhood $B_\alpha \subsetneq \Gamma$ of α can be chosen so that X_α is independent of X_β with $\beta \notin B_\alpha$. Let

$$b_1 = \sum_{\alpha\in\Gamma}\sum_{\beta\in B_\alpha} p_\alpha p_\beta \tag{1.5}$$

and

$$b_2 = \sum_{\alpha\in\Gamma}\sum_{\beta\in B_\alpha\setminus\{\alpha\}} E(X_\alpha X_\beta). \tag{1.6}$$

Barbour et al. [5] gave a uniform bound in the form of

$$\left| P(W \in A) - \sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \lambda^{-1}(1-e^{-\lambda})(b_1+b_2) \tag{1.7}$$

and Janson [9] used the coupling method to determine a uniform bound in the form of

$$\left| P(W \in A) - \sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \lambda^{-1}(1-e^{-\lambda})\sum_{\alpha\in\Gamma} p_\alpha E|W-W_\alpha^*|, \tag{1.8}$$

where W_α^* is a random variable that has the same distribution as $W - X_\alpha$ conditional on $X_\alpha = 1$.

For non-uniform bounds, Teerapabolarn and Neammanee [17] gave two pointwise bounds, that is,

$$\left| P(W = w_0) - \frac{\lambda^{w_0} e^{-\lambda}}{w_0!}\right| \leq \min\left\{\frac{1}{w_0}, \lambda^{-1}\right\}(b_1+b_2) \tag{1.9}$$

and

$$\left| P(W = w_0) - \frac{\lambda^{w_0} e^{-\lambda}}{w_0!}\right| \leq \min\{\frac{1}{w_0}, \lambda^{-1}\}\sum_{\alpha\in\Gamma} p_\alpha E|W-W_\alpha^*|, \tag{1.10}$$

where $w_0 \in \{1, 2, ..., |\Gamma|\}$. They later discovered two non-uniform bounds for $A = \{0, ..., w_0\}$, $w_0 \in \{0, ..., |\Gamma|\}$, in [19], which say that

$$\left| P(W \leq w_0) - \sum_{k=0}^{w_0}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \lambda^{-1}(1-e^{-\lambda})\min\left\{1, \frac{e^\lambda}{w_0+1}\right\}(b_1+b_2) \tag{1.11}$$

and

$$\left| P(W \leq w_0) - \sum_{k=0}^{w_0}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \lambda^{-1}(1-e^{-\lambda})\min\left\{1, \frac{e^\lambda}{w_0+1}\right\}\sum_{\alpha\in\Gamma} p_\alpha E|W-W_\alpha^*|. \tag{1.12}$$

After that, Teerapabolarn and Santiwipanont [20] determined general results of two non-uniform bounds for $A \subseteq \{0, ..., |\Gamma|\}$, that is,

$$\left| P(W \in A) - \sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \lambda^{-1}\min\left\{1, \lambda, \frac{\Delta(\lambda)}{M_A+1}\right\}(b_1+b_2) \tag{1.13}$$

and

$$\left| P(W \in A) - \sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \lambda^{-1}\min\left\{1, \lambda, \frac{\Delta(\lambda)}{M_A+1}\right\}\sum_{\alpha\in\Gamma} p_\alpha E|W-W_\alpha^*|. \tag{1.14}$$

It is observed that each result in (1.13) and (1.14) gives a good Poisson approximation when $\Delta(\lambda)$ is small, that is, e^{λ} is small: however, when e^{λ} is rather large, these results may be inappropriate for approximating the distribution of W. In this article, our goal is to improve the results with respect to the bounds in (1.13) and (1.14) by eliminating the influence of the factor e^{λ}.

The Stein-Chen method is utilized to provide all results in the present study as mentioned in Section 2. In Section 3, we use the Stein-Chen method to yield new results of the approximation and we also compare the obtained results and the results in (1.13) and (1.14). In Section 4, we give some examples to illustrate applications of these results. Concluding remarks are presented in the last section.

2 Method

In 1972, Stein [15] introduced a powerful and general method for bounding the error in the normal approximation. This method was first developed and applied to the Poisson case by Chen [7] which is refer to as the Stein-Chen method mentioned above. Stein's equation for Poisson distribution with mean $\lambda > 0$, for given h, is of the form

$$h(w) - \mathcal{P}_{\lambda}(h) = \lambda f(w+1) - wf(w), \tag{2.1}$$

where $\mathcal{P}_{\lambda}(h) = e^{-\lambda}\sum_{l=0}^{\infty} h(l)\frac{\lambda^l}{l!}$ and f 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

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

Following Barbour et al. [5], the solution f_A of (2.1) is of the form

$$f_A(w) = \begin{cases} (w-1)!\lambda^{-w}e^{\lambda}[\mathcal{P}_{\lambda}(h_{A\cap C_{w-1}}) - \mathcal{P}_{\lambda}(h_A)\mathcal{P}_{\lambda}(h_{C_{w-1}})] & \text{if } w \geq 1, \\ 0 & \text{if } w = 0, \end{cases} \tag{2.2}$$

For $k, w \in \mathbb{N}$, let $\Delta f_{\{k\}}(w) = f_{\{k\}}(w+1) - f_{\{k\}}(w)$ and $\Delta f_{C_k}(w) = f_{C_k}(w+1) - f_{C_k}(w)$. It follows from [15] that

$$\Delta f_{\{k\}}(w) \begin{cases} < 0 & \text{if } w \neq k, \\ > 0 & \text{if } w = k, \end{cases} \tag{2.3}$$

while Barbour et al. [5] showed that

$$\Delta f_{\{w\}}(w) \leq \frac{1}{w}. \tag{2.4}$$

Also, when $w \leq k$, it follows from [16] that

$$0 < \Delta f_{C_k}(w) \leq \Delta f_{C_k}(k). \tag{2.5}$$

The following lemma gives a non-uniform bound for $f_A(w+1) - f_A(w)$ that are used to determine the main results.

Lemma 2.1. *For $A \subseteq \mathbb{N} \cup \{0\}$ and $w \in \mathbb{N}$, let $\Delta f_A(w) = f_A(w+1) - f_A(w)$, $w_A^{\divideontimes} = \min\{w | w \in A\}$ and $w_A^{\bigstar} = \max\{w | C_w \subseteq A\}$, then we have the following:*

$$|\Delta f_A(w)| \leq \min\left\{\lambda^{-1}(1 - e^{-\lambda}), \frac{1}{w_A}\right\}, \tag{2.6}$$

where $\frac{1}{w_A}$ is taken to be 1 when $w_A = 0$ ($w_A^{\bigstar} = 0$ or $w_A^{\divideontimes} = 1$) and for $w_A > 0$, it is given by

$$\frac{1}{w_A} = \begin{cases} \frac{1}{w_A^{\bigstar}} & \text{if } 0 \in A, \\ \frac{1}{w_A^{\divideontimes} - 1} & \text{if } 0 \notin A. \end{cases}$$

Proof. The first bound of $|\Delta f_A(w)|$ follows directly from Barbour et al. [5]. For $w_A = 0$, $\min\left\{\lambda^{-1}(1 - e^{-\lambda}), \frac{1}{w_A}\right\} = \lambda^{-1}(1 - e^{-\lambda})$ because $\lambda^{-1}(1 - e^{-\lambda}) < 1$. The next step, for $w_A > 0$, we shall show that $|\Delta f_A(w)| \leq \frac{1}{w_A}$ as follows.

Case 1. $w > w_A$.

Because $\Delta f_A(w) = \sum_{k \in A} \Delta f_{\{k\}}(w)$ and $f_{A^c}(w) = -f_A(w)$, it follows from (2.3) and (2.4) that

$$\frac{1}{w} \geq \Delta f_{\{w\}}(w) \geq \Delta f_A(w) \geq \Delta f_{\{w\}^c}(w) = -\Delta f_{\{w\}}(w) \geq -\frac{1}{w},$$

this gives

$$|\Delta f_A(w)| \leq \frac{1}{w} \leq \frac{1}{w_A + 1}.$$

Case 2. $w \leq w_A^{\bigstar}$ ($0 \in A$).

Let $\widehat{w} = \max\{w | w \in A\}$. Following (2.5), we obtain

$$0 < \Delta f_{C_{\widehat{w}}}(w) \leq \Delta f_A(w).$$

Thus

$$0 < \Delta f_A(w) \leq \Delta f_{C_{w_A^{\bigstar}}}(w) \leq \Delta f_{C_{w_A^{\bigstar}}}(w_A^{\bigstar}) \leq \Delta f_{\{w_A^{\bigstar}\}}(w_A^{\bigstar}) \leq \frac{1}{w_A^{\bigstar}} = \frac{1}{w_A}.$$

Case 3. $w \leq w_A^{\divideontimes} - 1$ ($0 \notin A$).

It is observed that $\Delta f_A(w) < 0$. Therefore

$$\begin{aligned} 0 < -\Delta f_A(w) &\leq -\Delta f_{C^c_{w_A^{\divideontimes} - 1}}(w) \\ &= \Delta f_{C_{w_A^{\divideontimes} - 1}}(w) \\ &\leq \Delta f_{C_{w_A^{\divideontimes} - 1}}(w_A^{\divideontimes} - 1) \\ &\leq \Delta f_{\{w_A^{\divideontimes} - 1\}}(w_A^{\divideontimes} - 1) \\ &\leq \frac{1}{w_A^{\divideontimes} - 1} \\ &= \frac{1}{w_A}. \end{aligned}$$

Hence, following three cases, (2.6) is obtained. $\square$

Lemma 2.2. *Let* $Z_\alpha = \sum_{\beta \in B_\alpha \setminus \{\alpha\}} X_\beta, Y_\alpha = W - X_\alpha - Z_\alpha = \sum_{\beta \notin B_\alpha} X_\beta$ *and* $f = f_A$ *be defined as above. Then we have the following:*

1. $|E[p_\alpha(f(W+1) - f(Y_\alpha+1))]| \le \lambda^{-1} \min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\} \left(p_\alpha^2 + p_\alpha E(Z_\alpha)\right).$
2. $|E[X_\alpha(f(Y_\alpha + Z_\alpha + 1) - f(Y_\alpha+1))]| \le \lambda^{-1} \min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\} E(X_\alpha Z_\alpha).$

Proof. The inequalities in 1 and 2 follow from the same argument detailed in the proof of Lemma 2.2 in [20] combined with the bound in Lemma 2.1. $\square$

3 Results

The main results of this study are new non-uniform bounds for approximating the distribution of sums of dependent Bernoulli random variables using the Poisson distribution. These results can be obtained with the Stein-Chen method and related properties in Section 2 to improve the results of Teerapabolarn and Santiwipanont [20] in the following theorems.

Theorem 3.1. *With the above definition, for* $A \subseteq \{0, ..., |\Gamma|\}$*, we have the following:*

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \le \lambda^{-1} \min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\} (b_1 + b_2) \tag{3.1}$$

and for $A = \{0\}$,

$$\left|P(W = 0) - e^{-\lambda}\right| \le \lambda^{-2}(\lambda + e^{-\lambda} - 1) \max\{b_1, b_2\}. \tag{3.2}$$

Proof. The inequality (3.2) follows the result in [18]. Now, we have to verify the general result in (3.1).

Let $Z_\alpha = \sum_{\beta \in B_\alpha \setminus \{\alpha\}} X_\beta, Y_\alpha = W - X_\alpha - Z_\alpha = \sum_{\beta \notin B_\alpha} X_\beta$, $W_\alpha = W - X_\alpha$ and $f = f_A$ be defined as in (2.2). Teerapabolarn and Santiwipanont [20] showed that

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \le \sum_{\alpha \in \Gamma} \{|E[p_\alpha(f(W+1) - f(Y_\alpha+1))]| + |E[X_\alpha(f(Y_\alpha + Z_\alpha + 1) - f(Y_\alpha + 1))]|.$$

With Lemma 2.1 and 2.2, we obtain

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \le \lambda^{-1} \min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\} (b_1 + b_2). \qquad \square$$

If it is possible to construct, for each $\alpha \in \Gamma$, a random variable W_α^* on a common probability space with W such that W_α^* has the same distribution as the $W - X_\alpha$ conditional on $X_\alpha = 1$, then the following theorem provides a result along these lines.

Theorem 3.2. *For $A \subseteq \{0, ..., |\Gamma|\}$, we have the following:*

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \lambda^{-1} \min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\} \sum_{\alpha \in \Gamma} p_\alpha E|W - W_\alpha^*| \tag{3.3}$$

and for $A = \{0\}$,

$$\left| P(W = 0) - e^{-\lambda} \right| \leq \lambda^{-2}(\lambda + e^{-\lambda} - 1) \sum_{\alpha \in \Gamma} p_\alpha E|W - W_\alpha^*|. \tag{3.4}$$

Proof. The second inequality follows from the Theorem 2.2 in [20]. In the next step, we shall show that (3.3) holds. Teerapabolarn and santiwipanont [20] showed that

$$\begin{aligned} \left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| &\leq \sum_{\alpha \in \Gamma} p_\alpha E|f(W+1) - f(W_\alpha^* + 1)| \\ &\leq \sup_{w \geq 1} |\Delta f(w)| \sum_{\alpha \in \Gamma} p_\alpha E|W - W_\alpha^*|, \end{aligned}$$

where $f = f_A$ is defined in (2.2). Following Lemma 2.1, we have

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \lambda^{-1} \min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\} \sum_{\alpha \in \Gamma} p_\alpha E|W - W_\alpha^*|,$$

which holds for (3.3). □

If all X_α are independent, then a non-uniform bound of a Poisson approximation to the Poisson binomial distribution can be obtained from the following result.

Corollary 3.1. *If $\{X_\alpha, \alpha \in \Gamma\}$ are independent Bernoulli random variables, then for $A \subseteq \{0, ..., |\Gamma|\}$,*

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \lambda^{-1} \min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\} \sum_{\alpha \in \Gamma} p_\alpha^2. \tag{3.5}$$

Consider the result in Theorem 3.2, if $W \geq W_\alpha^*$ or $W - X_\alpha \leq W_\alpha^*$ for every $\alpha \in \Gamma$, then we have more convenient forms in the following corollaries.

Corollary 3.2. *If $W \geq W_\alpha^*$ for every $\alpha \in \Gamma$, then we have the following:*

$$\left| P(W = 0) - e^{-\lambda} \right| \leq \lambda^{-2}(\lambda + e^{-\lambda} - 1)\{\lambda - Var(W)\} \tag{3.6}$$

and for $A \subseteq \{0, ..., |\Gamma|\}$,

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \lambda^{-1} \min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\} \{\lambda - Var(W)\}. \tag{3.7}$$

Corollary 3.3. *If $W - X_\alpha \leq W^*_\alpha$ for every $\alpha \in \Gamma$, then we have the following:*

$$\left|P(W=0) - e^{-\lambda}\right| \leq \lambda^{-2}(\lambda + e^{-\lambda} - 1)\left\{Var(W) - \lambda + 2\sum_{\alpha\in\Gamma} p_\alpha^2\right\} \tag{3.8}$$

and for $A \subseteq \{0, ..., |\Gamma|\}$,

$$\left|P(W \in A) - \sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \lambda^{-1}\min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\}\left\{Var(W) - \lambda + 2\sum_{\alpha\in\Gamma} p_\alpha^2\right\}. \tag{3.9}$$

Remark. Let us consider the bound of $|\Delta f_A(w)|$ in (2.6) and the bound in Teerapabolarn and Santiwipanont [20], that is, $\lambda^{-1}\min\left\{1-e^{-\lambda}, \frac{\lambda}{w_A}\right\}$ and $\lambda^{-1}\min\left\{1, \lambda, \frac{\Delta(\lambda)}{M_A+1}\right\}$, where $M_A = w_A$ as $w_A = w_A^\bigstar$ and $M_A = w_A + 1$ as $w_A = w_A^{\divideontimes} - 1$. It follows that

1. $1 - e^{-\lambda} < \min\{1, \lambda\}$.
2. For $M_A \leq 2$, $1 - e^{-\lambda} < \frac{\Delta(\lambda)}{M_A+1}$.
3. $\frac{\lambda}{w_A} < \frac{\Delta(\lambda)}{M_A+1}$ when $w_A = w_A^\bigstar > 0$, because $\frac{\lambda}{w_A^\bigstar} \leq \frac{2\lambda}{w_A^\bigstar+1} < \frac{e^\lambda+\lambda-1}{M_A+1} < \frac{2(e^\lambda-1)}{M_A+1}$.
4. $\frac{\lambda}{w_A} < \frac{\Delta(\lambda)}{M_A+1}$ when $w_A = w_A^{\divideontimes} - 1 > 1$, because $\frac{\lambda}{w_A^{\divideontimes}-1} \leq \frac{2\lambda}{w_A^{\divideontimes}+1} < \frac{e^\lambda+\lambda-1}{M_A+1} < \frac{2(e^\lambda-1)}{M_A+1}$.

Following these comparisons, the bounds (3.1) and (3.3) are sharper than the bounds (1.13) and (1.14). Therefore, our results in this study are superior to all results of Teerapabolarn and Santiwipanont [20].

4 Applications

Many applications of the Poisson estimate for dependent Bernoulli trials have been proposed by various authors in recent years. These include the birthday problem and the longest head run in Arratia et al. [1, 2], applications to the theory of random graphs in Barbour et al. [5], the problem of estimating statistical significance in sequence comparison in Goldstein and Waterman [8], sequence comparison significance in Waterman and Vingron [22], applications to time series analysis in Kim [10] and the somatic cell hybrid model in Lange [11], all of which are applications of the result in Theorem 3.1. Some applications of the result in Theorem 3.2 include random graph problems in Barbour [3, 4], Barbour et al. [5] and Janson [9], the random allocation problem in Mikhailov [12], occupancy and urn models in Barbour et al. [5], the empty urn model in Boonyued and Tangkanchanawong [6] and the ménage, birthday and biggest random graph problems in Lange [11]. In this section, we present some results that are applications of Theorems 3.1 and 3.2 and Corollaries 3.2 and 3.3, which are the same applications of the results in Teerapabolarn and Santiwipanont [20].

Example 4.1. (A birthday problem)
Suppose n balls (people) are uniformly and independently distributed into d boxes (days of the year). The birthday problem involves determining an approximate distribution of the number of boxes that receive k or more balls for some fixed positive integer k. Let Γ be the collection of all sets of trials $\alpha \subset \{1, 2, ..., n\}$ having $|\alpha| = k$ elements, where $\{1, 2, ..., n\}$ is a set of n balls.

Let X_α be the indicator of the event that the balls indexed by α all fall into the same box with small probability $p_\alpha = P(X_\alpha = 1) = d^{1-k}$. The number of sets of k balls that fall into the same box is given by $W = \sum_{\alpha\in\Gamma} X_\alpha$. It seems reasonable to approximate W as a Poisson random variable with mean $\lambda = E(W)$ when p_α is small. Because all p_α are identical, we have

$$\lambda = |\Gamma| p_\alpha = \binom{n}{k} d^{1-k}.$$

To bound the error of the difference of the distribution of W and the Poisson distribution, following Arratia et al. [1], we first take $B_\alpha = \{\beta \in \Gamma : \alpha \cap \beta \neq \emptyset\}$ as the neighborhood dependence set for α. It is observed that X_α and X_β are independent when $\alpha\cap\beta = \emptyset$. Because the size of B_α is $|B_\alpha| = \binom{n}{k} - \binom{n-k}{k}$, we have

$$\begin{aligned} b_1 &= |\Gamma||B_\alpha|p_\alpha^2 \\ &= \lambda |B_\alpha| d^{1-k}. \end{aligned}$$

For a given α, we have $1 \le |\alpha\cap\beta| \le k-1$ for $\beta \in B_\alpha \setminus \{\alpha\}$ and

$$\begin{aligned} b_2 &= \binom{n}{k} \sum_{j=1}^{k-1} \binom{k}{j}\binom{n-k}{k-j} d^{1+j-2k} \\ &= \lambda b, \end{aligned}$$

where $b = \sum_{j=1}^{k-1} \binom{k}{j}\binom{n-k}{k-j} d^{j-k}$. By applying Theorem 3.1, we obtain

$$\left| P(W\in A) - \sum_{k\in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \le \min\left\{1-e^{-\lambda}, \frac{\lambda}{w_A}\right\} \left(|B_\alpha| d^{1-k} + b\right),$$

where $A \subseteq \{0, ..., |\Gamma|\}$ and

$$\left| P(W=0) - e^{-\lambda} \right| \le \lambda^{-1}(\lambda + e^{-\lambda} - 1) \max\left\{|B_\alpha| d^{1-k}, b\right\}.$$

Numerical examples:

1. For $n = 5,\ k = 2$ and $d = 30$, we have $\lambda = \frac{1}{3},\ |B_\alpha| = 7$ and $b = 0.2$. Thus for $A \subseteq \{0, ..., 10\}$, an approximation of the distribution of the number of sets of two balls that fall into the same box is

$$\left| P(W\in A) - \sum_{k\in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \le \begin{cases} 0.12283643 & \text{if } w_A \le 1, \\ \frac{0.14444444}{w_A} & \text{if } w_A \ge 2, \end{cases}$$

which is better than the numerical result obtained from (1.13),

$$\left| P(W\in A) - \sum_{k\in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \le \begin{cases} 0.14444444 & \text{if } M_A \le 1, \\ \frac{0.31587650}{M_A+1} & \text{if } M_A \ge 2. \end{cases}$$

2. For $n = 50,\ k = 3$ and $d = 365$, we have $\lambda = \binom{50}{3}(365)^{-2} = 0.14711953,\ |B_\alpha| = \binom{50}{3} - \binom{47}{3} = 3385$ and $b = 3\binom{47}{2}(365)^{-2} + 3(47)(365)^{-1} = 0.41064365$. Thus for $A \subseteq \{0, ..., 19600\}$,

an approximation of the distribution of the number of sets of two balls that fall into the same box is

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \le \begin{cases} 0.05965590 & \text{if } w_A \le 1, \\ \frac{0.06415174}{w_A} & \text{if } w_A \ge 2, \end{cases}$$

which is also better than the numerical result obtained from (1.13),

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \le \begin{cases} 0.06415174 & \text{if } M_A \le 1, \\ \frac{0.13326265}{M_A+1} & \text{if } M_A \ge 2. \end{cases}$$

Example 4.2. (A random graph problem)
Consider the n-dimensional unit cube $[0,1]^n$ random graph with 2^n vertices, each of degree n, with an edge joining pairs of vertices that differ in exactly one coordinate. Suppose that each of the $n2^{n-1}$ edges is independently assigned to one of two equally likely orientations. Let Γ be the set of all 2^n vertices, and for each $\alpha \in \Gamma$, let X_α be the indicator that vertex α has all of its edges directed inward with the probability $p_\alpha = P(X_\alpha = 1) = 2^{-n}$. Let $W = \sum_{\alpha \in \Gamma} X_\alpha$ be the number of vertices at which all n edges point inward. Its distribution can be approximated by a Poisson distribution with mean $\lambda = E(W) = 1$ when n is large.

We follow Arratia et al. [1] by taking $B_\alpha = \{\beta \in \Gamma : |\alpha - \beta| = 1\}$ as the neighborhood of α such that X_α and X_β are independent for every $\beta \notin B_\alpha$. X_α is independent of X_β with $|\alpha - \beta| > 1$ and $E(X_\alpha X_\beta) = 0$ for $|\alpha - \beta| = 1$; hence $b_2 = 0$. Because $|B_\alpha| = n$, we have

$$\begin{aligned} b_1 &= |\Gamma||B_\alpha|p_\alpha^2 \\ &= n2^{-n}. \end{aligned}$$

By applying Theorem 3.1, it follows that

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \le n2^{-n} \min\left\{1 - e^{-1}, \frac{1}{w_A}\right\},$$

where $A \subseteq \{0, ..., 2^{n-1}\}$ and

$$\left|P(W = 0) - e^{-1}\right| \le ne^{-1}2^{-n}.$$

Numerical examples:

1. For $n = 5$ and $A \subseteq \{0, ..., 16\}$, an approximation of the distribution of the number of vertices at which all 5 edges point inward is of the form

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \le \begin{cases} 0.09876884 & \text{if } w_A \le 1, \\ \frac{0.15625000}{w_A} & \text{if } w_A \ge 2, \end{cases}$$

which is better than the numerical result obtained from (1.13),

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \le \begin{cases} 0.15625000 & \text{if } M_A \le 1, \\ \frac{0.42473154}{M_A+1} & \text{if } M_A \ge 2. \end{cases}$$

2. For $n = 10$ and $A \subseteq \{0, ..., 512\}$, an approximation of the distribution of the number of vertices at which all 10 edges point inward is of the form

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \begin{cases} 0.00617305 & \text{if } w_A \leq 1, \\ \frac{0.00976563}{w_A} & \text{if } w_A \geq 2, \end{cases}$$

which is also better than the numerical result obtained from (1.13),

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \begin{cases} 0.00976563 & \text{if } M_A \leq 1, \\ \frac{0.02654572}{M_A + 1} & \text{if } M_A \geq 2. \end{cases}$$

Example 4.3. (The longest perfect head run)
Consider an infinite sequence $Y_1, Y_2, ...$ of independent random indicators with success probability p. For $\Gamma = \{1, ..., n\}$ and a fixed positive integer value of length t, let X_α be the indicator of the event that a successful run of length t or longer begins at position α. Note that $X_1 = \prod_{k=1}^{t} Y_k$ and for $\alpha \in \{2, ..., n\}$,

$$X_\alpha = (1 - Y_{\alpha-1}) \prod_{k=\alpha}^{\alpha+t-1} Y_k.$$

Let $W = \sum_{\alpha \in \Gamma} X_\alpha$ be the number of such successful runs starting in the first n positions. The Poisson heuristic suggests that W is approximately Poisson with mean $\lambda = E(W) = p^t[(n-1)(1-p)+1]$.

Following Arratia et al. [1], we take $B_\alpha = \{\beta \in \Gamma : |\beta - \alpha| \leq t\}$ as the neighborhood of α. It is observed that X_α is independent of X_β for $\beta \notin B_\alpha$ and $E(X_\alpha X_\beta) = 0$; hence $b_2 = 0$ and

$$\begin{aligned} b_1 &= \sum_{\alpha \in \Gamma} \sum_{\beta \in B_\alpha} p_\alpha p_\beta \\ &= p^{2t} + 2tp^{2t}(1-p) + [2nt - t^2 + n - 3t - 1]p^{2t}(1-p)^2 \\ &\leq \frac{\lambda^2(2t+1)}{n} + 2\lambda p^t. \end{aligned}$$

By applying Theorem 3.1, an approximation of the distribution of the number of successful runs starting in the first n positions is of the form

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\} \left[\frac{\lambda(2t+1)}{n} + 2p^t\right],$$

where $A \subseteq \{0, ..., n\}$ and

$$\left| P(W = 0) - e^{-\lambda} \right| \leq \lambda^{-1}(\lambda + e^{-\lambda} - 1) \left[\frac{\lambda(2t+1)}{n} + 2p^t\right].$$

Numerical examples:

1. For $n = 200$, $p = 0.3$ and $t = 4$, we have $\lambda = 1.13643$ and for $A \subseteq \{0, ..., 200\}$, a non-uniform bound for this approximation is of the form

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \begin{cases} 0.04572592 & \text{if } w_A \leq 1, \\ \frac{0.07652646}{w_A} & \text{if } w_A \geq 2, \end{cases}$$

which is better than the numerical result obtained from (1.13),

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \begin{cases} 0.06733935 & \text{if } M_A \leq 2, \\ \frac{0.21899132}{M_A+1} & \text{if } M_A \geq 3. \end{cases}$$

2. For $n = 500$, $p = 0.5$ and $t = 7$, we have $\lambda = 1.95703125$ and for $A \subseteq \{0, ..., 500\}$, a non-uniform bound for this approximation is of the form

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \begin{cases} 0.06383396 & \text{if } w_A \leq 2, \\ \frac{0.14547775}{w_A} & \text{if } w_A \geq 3, \end{cases}$$

which is also better than the numerical result obtained from (1.13),

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \begin{cases} 0.07433594 & \text{if } M_A \leq 7, \\ \frac{0.59731256}{M_A+1} & \text{if } M_A \geq 8. \end{cases}$$

Example 4.4. (A hypergeometric distribution)
Suppose a random sample of size n is chosen without replacement from a finite population containing N elements of two types of which m are of type A and $N - m$ are of type B. For each $\alpha \in \Gamma = \{1, ..., n\}$, let $X_\alpha = 1$ if the α^{th} element in the sample is of type A and $X_\alpha = 0$ otherwise. Then the probability $P(X_\alpha = 1) = \frac{m}{N}$. Let $W = \sum_{\alpha=1}^{n} X_\alpha$, thus W is the number of type A elements in the sample that have the hypergeometric distribution with parameters N, m and n, and its the mean and variance are $E(W) = \frac{nm}{N}$ and $Var(W) = \frac{N-n}{N-1}\frac{nm}{N}\left(1 - \frac{m}{N}\right)$, respectively. If $\dfrac{m}{N}$ and $\dfrac{n}{N}$ are small then it seems reasonable to approximate the distribution of W by a Poisson distribution with mean $\lambda = E(W) = \frac{nm}{N}$.

Consider the coupled random variable W_α^* which has the same distribution as the $W - X_\alpha$ conditional on $X_\alpha = 1$. It is the number of type A elements in the sample other than the α^{th} element conditional on $X_\alpha = 1$ and is obtained by swapping out the α^{th} element chosen if it is of type B, for a randomly chosen an element of type A. Following Barbour [4], we take

$$W_\alpha^* = W - X_\alpha - \sum_{\beta=1, \beta \neq \alpha}^{n} X_\beta I_\beta,$$

where I_β is the indicator of the event that the β^{th} element in the sample is chosen to be swapped with the α^{th}. It is observed that $W \geq W_\alpha^*$ for every $\alpha \in \{1, ..., n\}$. Thus, by Corollary 3.2, we have

$$\left| P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq \min\left\{1 - e^{-\lambda}, \frac{\lambda}{w_A}\right\} \left(\frac{n + m - 1}{N - 1}\right),$$

where $A \subseteq \{0, ..., n\}$ and

$$\left|P(W=0) - e^{-\lambda}\right| \leq \lambda^{-1}(\lambda + e^{-\lambda} - 1)\left(\frac{n+m-1}{N-1}\right).$$

Numerical examples:

1. For $N = 500$, $m = 25$ and $n = 20$, we have $\lambda = 1$ and for $A \subseteq \{0, ..., 20\}$, a Poisson approximation to the hypergeometric distribution is of the form

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases} 0.05573809 & \text{if } w_A \leq 1, \\ \frac{0.08817635}{w_A} & \text{if } w_A \geq 2, \end{cases}$$

which is better than the numerical result obtained from (1.14),

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases} 0.08817635 & \text{if } M_A \leq 1, \\ \frac{0.23968818}{M_A+1} & \text{if } M_A \geq 2. \end{cases}$$

2. For $N = 1000$, $m = 70$ and $n = 30$, we have $\lambda = 2.1$ and for $A \subseteq \{0, ..., 30\}$, a Poisson approximation to the hypergeometric distribution is of the form

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases} 0.08696378 & \text{if } w_A \leq 2, \\ \frac{0.20810811}{w_A} & \text{if } w_A \geq 3, \end{cases}$$

which is also better than the numerical result obtained from (1.14),

$$\left|P(W \in A) - \sum_{k \in A} \frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases} 0.09909910 & \text{if } M_A \leq 8, \\ \frac{0.91826911}{M_A+1} & \text{if } M_A \geq 9. \end{cases}$$

Example 4.5. (A random graph problem)
A random graph $G(n,p)$ is a graph on n labeled vertices $\{1, 2, \ldots, n\}$ where each possible edge $\{\alpha, \beta\}$ is present randomly and independently with probability p, $0 < p < 1$. If we let $E_{\alpha\beta}$ be the independent edge indicator of the event at edge $\{\alpha, \beta\} \in G(n,p)$, then $P(E_{\alpha\beta} = 1) = p$. For each $\alpha \in \Gamma = \{1, ..., n\}$, let $X_\alpha = 1$ if vertex α is an isolated vertex in $G(n,p)$ and $X_\alpha = 0$ otherwise. Then $W = \sum_{\alpha=1}^{n} X_\alpha$ is the number of isolated vertices in $G(n,p)$. We now have the probability $p_\alpha = P(X_\alpha = 1) = (1-p)^{n-1}$, $\lambda = E(W) = n(1-p)^{n-1}$ and $Var(W) = \lambda + n(n-1)(1-p)^{2n-3} - \lambda^2$. Because $E(X_\alpha X_\beta) \neq E(X_\alpha)E(X_\beta)$ for $\alpha \neq \beta$, it indicates that X_α's are not independent.

Consider the number of isolated vertices in $G(n,p)$ other than the α^{th} vertex conditional on $X_\alpha = 1$, which is obtained by deleting all the edges $\{\alpha, \beta\}$ $(1 \leq \beta \leq n,\ \beta \neq \alpha)$ in $G(n,p)$. Following Barbour [4], we take

$$W_\alpha^* = W - X_\alpha + \sum_{\beta=1, \beta \neq \alpha}^{n} E_{\alpha\beta} \prod_{\gamma \neq \alpha, \beta} (1 - E_{\beta\gamma}),$$

where $\sum_{\beta=1,\beta\neq\alpha}^{n} E_{\alpha\beta}\prod_{\gamma\neq\alpha,\beta}(1-E_{\beta\gamma})$ is the number of isolated vertices that are connected to the vertex α. Then W_α^* has the same distribution as $W-X_\alpha$ conditional on $X_\alpha=1$, and we observe that $W_\alpha^* \geq W-X_\alpha$ for every $\alpha\in\{1,...,n\}$. Thus, by Corollary 3.3, an approximation of the distribution of the number of isolated vertices in $G(n,p)$ by a Poisson distribution is of the form

$$\left|P(W\in A)-\sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \min\left\{1-e^{-\lambda},\frac{\lambda}{w_A}\right\}[(n-2)p+1]e^{-(n-2)p},$$

where $A\subseteq\{0,...,n\}$ and

$$\left|P(W=0)-e^{-\lambda}\right| \leq \lambda^{-1}(\lambda+e^{-\lambda}-1)[(n-2)p+1]e^{-(n-2)p}.$$

Numerical examples:

1. For $n=15$ and $p=0.2$, we have $\lambda=0.65970698$ and for $A\subseteq\{0,...,15\}$, a non-uniform bound of the error of this approximation is of the form

$$\left|P(W\in A)-\sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases} 0.12914615 & \text{if } w_A\leq 1,\\ \frac{0.17639567}{w_A} & \text{if } w_A\geq 2,\end{cases}$$

which is better than the numerical result obtained from (1.14),

$$\left|P(W\in A)-\sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases} 0.17639567 & \text{if } M_A\leq 1,\\ \frac{0.42619344}{M_A+1} & \text{if } M_A\geq 2.\end{cases}$$

2. For $n=30$ and $p=0.1$, we have $\lambda=1.41303861$ and for $A\subseteq\{0,...,30\}$, a non-uniform bound of the error of this approximation is of the form

$$\left|P(W\in A)-\sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases} 0.17483320 & \text{if } w_A\leq 1,\\ \frac{0.32652247}{w_A} & \text{if } w_A\geq 2,\end{cases}$$

which is also better than the numerical result obtained from (1.14),

$$\left|P(W\in A)-\sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases} 0.23107824 & \text{if } M_A\leq 3,\\ \frac{1.04481077}{M_A+1} & \text{if } M_A\geq 4.\end{cases}$$

Example 4.6. (The ménage problem)
The classical ménage problem asks for the number of seatings of n married couples at a round table, with men and women alternating such that no one sits next to his or her partner. More generally, we may ask for the probability that a random seating produces exactly k couples sitting together. We number the seats around the table from 1 to $2n$, that is, for $\alpha\in\Gamma=\{1,...,2n\}$, let $X_\alpha=1$ if a couple occupies seats α and $\alpha+1$ and $X_\alpha=0$ otherwise. Then, W, the number of couples sitting next to each other, can be represented by $W=\sum_{\alpha=1}^{2n}X_\alpha$, where $X_{2n+1}=X_1$ and, by symmetry, $p_\alpha=P(X_\alpha=1)=\frac{1}{n}$ and $\lambda=E(W)=2$.

The coupled random variable W_α^* is constructed by exchanging the person in seat $\alpha+1$ with the spouse of the person in seat α; then, we count the number of adjacent spouse pairs, excluding the pair now occupying seats α and $\alpha+1$. From Lange [11], the term $E|W-W_\alpha^*|$ is bound by $\frac{6(n-2)}{n(n-1)}$, that is, $E|W-W_\alpha^*| \leq \frac{6(n-2)}{n(n-1)}$. By applying Theorem 3.2, a result in approximating the distribution of the number of couples sitting next to each other can be approximated as

$$\left|P(W\in A)-\sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \min\left\{1-e^{-2},\frac{2}{w_A}\right\}\frac{6(n-2)}{n(n-1)},$$

where $A\subseteq\{0,...,2n\}$ and

$$\left|P(W=0)-e^{-\lambda}\right| \leq \frac{3(1+e^{-2})(n-2)}{n(n-1)}.$$

Numerical examples:

1. For $n=100$ and $A\subseteq\{0,...,200\}$, a result of this approximation is

$$\left|P(W\in A)-\sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases}0.05135584 & \text{if } w_A\leq 2,\\ \frac{0.11878788}{w_A} & \text{if } w_A\geq 3,\end{cases}$$

which is better than the numerical result obtained from (1.14),

$$\left|P(W\in A)-\sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases}0.05939394 & \text{if } M_A\leq 7,\\ \frac{0.49825909}{M_A+1} & \text{if } M_A\geq 8.\end{cases}$$

2. For $n=200$ and $A\subseteq\{0,...,400\}$, a result of this approximation is

$$\left|P(W\in A)-\sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases}0.02580959 & \text{if } w_A\leq 2,\\ \frac{0.05969849}{w_A} & \text{if } w_A\geq 3,\end{cases}$$

which is also better than the numerical result obtained from (1.14),

$$\left|P(W\in A)-\sum_{k\in A}\frac{\lambda^k e^{-\lambda}}{k!}\right| \leq \begin{cases}0.02984925 & \text{if } M_A\leq 7,\\ \frac{0.25040700}{M_A+1} & \text{if } M_A\geq 8.\end{cases}$$

5 Conclusion

The bounds in Theorems 3.1 and 3.2 and Corollaries 3.1−3.3, which were improved by the Stein-Chen method, provide new general estimates of the error between the distribution of W and the Poisson distribution with mean $\lambda=E(W)$, where W is a sums of dependent Bernoulli random variables. All bounds reported in this study are sharper than the bounds in Teerapabolarn and Santiwipanont [20], including both theoretical and numerical results. In addition, the influence of factor e^λ in the old bounds is eliminated from or reduced in the new bounds. Accordingly, these bounds provide appropriate criteria for measuring the accuracy of approximating the distribution of W by the Poisson distribution with this mean. In short, if each corresponding bound is small, a good Poisson approximation is obtained. When each corresponding bound is not small, however, it is not appropriate to approximate the distribution of W with the Poisson distribution.

Acknowledgement. This research is supported by the Centre of Excellence in Mathematics, the Commission on Higher Education, Thailand. The author would like to thank the referees for their useful comments and suggestions.

References

[1] Arratia R., Goldstein L., Gordon L. 1989. Two moments suffice for Poisson approximations:the Chen-Stein method *Annals of probability* 17:9–25.

[2] Arratia R., Goldstein L., Gordon L. 1990. Poisson approximations and the Chen-Stein method. *Statistical Science* 5:403–434.

[3] Barbour, A. D. 1982. Poisson convergence and random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society* 92:349–359.

[4] Barbour A. D. 2001. Stochastic Processes:Theory and Methods. *Handbook of Statistics.* Elsevier Science 19:79–115.

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

[6] Boonyued, A., Tangkanchanawong, C. 2012. A non-uniform bound on Poisson approximation of empty urn-model via Stein-Chen method. *International Mathematical Forum* 7:2037–2044.

[7] Chen, L. H. Y. 1975. Poisson approximation for dependent trials. *Annals of Probability* 3:534–545.

[8] Goldstein, L., Waterman, M. 1992. Poisson, compound Poisson and process approximations for testing statistical significance in sequence Comparisons. *Bulletin of Mathematical Biology* 54:785–812.

[9] Janson, S. 1994. Coupling and Poisson approximation. *Acta Applicandae Mathematicae* 34:7–15.

[10] Kim, S. T. 2000. A use of the Stein-Chen method in time series analysis. *Journal of Applied Probability* 37:1129–1136.

[11] Lange, K. 2003. *Applied Probability.* Springer-Verlag, New York.

[12] Mikhailov, V. G. 1997. On a Poisson approximation for the distribution of the number of empty cells in a nonhomogeneuos allocation scheme. *Theory of Probability and Its Applications* 42:184–189.

[13] Neammanee, K. 2003. Pointwise approximation of Poisson binomial by Poisson distribution. *Stochastic Modelling and Applications* 6:20–26.

[14] Stein, C. M. 1972. A bound for the error in normal approximation to the distribution of a sum of dependent random variables. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* 3:583–602.

[15] Stein, C. M. 1986. *Approximate Computation of Expectations.* IMS, Hayward California.

[16] Teerapabolarn, K. 2006. A non-uniform bound on Poisson approximation for sums of Bernoulli random variables with small mean. *Thai Journal of Mathematics* 4:179–196.

[17] Teerapabolarn, K., Neammanee, K. 2005. A non-uniform bound on Poisson approximation for dependent trials. *Stochastic Modelling and Applications* 8:17–31.

[18] Teerapabolarn, K., Neammanee, K. 2005. A non-uniform bound in somatic cell hybrid model. *Mathematical BioScience* 195:56–64.

[19] Teerapabolarn, K., Neammanee, K. 2006. Poisson approximation for sums of dependent Bernoulli random variables. *ActaMathematica Academiae Paedagogicae Nyiregyhaziensis* 22:87–99.

[20] Teerapabolarn, K., Santiwipanont, T. 2007. Two non-uniform bounds in the Poisson approximation of sums of dependent indicators. *Thai Journal of Mathematics* 5:15–39.

[21] Teerapabolarn, K. 2011. On the Poisson approximation to the negative hypergeometric distribution. *Bulletin of the Malaysian Mathematical Sciences Society* 34:331–336.

[22] Waterman, S. M., Vingron, M. 1994. Sequence comparison significance and Poisson approximation. *Statistical Science* 9:367–381.