Journal of Mathematics Research; Vol. 4, No. 3; 2012
ISSN 1916-9795 E-ISSN 1916-9809
Published by Canadian Center of Science and Education

Bounds in Poisson Approximation for Random Sums of Bernoulli Random Variables

Sasithorn Kongudomthrap
Department of Mathematics and Computer Science, Faculty of Science
Chulalongkorn University, Bangkok, Thailand

Nattakarn Chaidee (Corresponding author)
Department of Mathematics and Computer Science, Faculty of Science
Chulalongkorn University, Bangkok, Thailand
&
Centre of Excellence in Mathematics, CHE, Si Ayutthaya Road, Bangkok 10400, Thailand
E-mail: nattakarn.c@chula.ac.th

Received: April 5, 2012 Accepted: April 18, 2012 Online Published: May 28, 2012
doi:10.5539/jmr.v4n3p29 URL: http://dx.doi.org/10.5539/jmr.v4n3p29

Abstract

Let (X_n) be a sequence of Bernoulli random variables and N a positive integer value random variable. Assume that $N, X_1, X_2, \ldots$ are independent. In this paper, we investigate uniform and non-uniform bounds in Poisson approximation for random sums $X_1 + X_2 + \cdots + X_N$.

Keywords: Poisson approximation, Random sums, Bernoulli random variable

1. Introduction and Main Results

Let $X_1, X_2, \ldots, X_n$ be independent Bernoulli random variables with

$$P(X_i = 1) = p_i = 1 - P(X_i = 0)$$

and

$$S_n = X_1 + X_2 + \cdots + X_n.$$

Let U_λ be a Poisson random variable with parameter λ, i.e., $P(U_\lambda = k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$ for $k = 0, 1, 2, \ldots$. Set $\lambda_n = \sum_{i=1}^{n} p_i$. It has long been known that the distribution of S_n can be approximated by the distribution of U_{λ_n} if p_i's are small, that is Poisson approximations are essential in the case of event has small probability of occurring. Many authors investigated the approximations between S_n and U_{λ_n}. For examples, Le Cam in 1960 gave a uniform bound

$$\sup_{x\in\mathbb{Z}_0^+} |P(S_n \le x) - P(U_{\lambda_n} \le x)| \le \sum_{i=1}^{n} p_i^2$$

and Kerstan in 1964 obtained this result in the form of

$$\sup_{x\in\mathbb{Z}_0^+} |P(S_n \le x) - P(U_{\lambda_n} \le x)| \le 1.05\lambda_n^{-1} \sum_{i=1}^{n} p_i^2 \text{ if } \max_{1\le i\le n} p_i \le 1/4.$$

In 1974, Chen used the Stein's method to gave the bound

$$\sup_{x\in\mathbb{Z}_0^+} |P(S_n \le x) - P(U_{\lambda_n} \le x)| \le 5\lambda_n^{-1} \sum_{i=1}^{n} p_i^2$$

and then Barbour and Hall improved the result of Chen (1974) as follows.

Theorem 1.1 (Barbour & Hall, 1984) *We have*

$$\sup_{x\in\mathbb{Z}_0^+} |P(S_n \le x) - P(U_{\lambda_n} \le x)| \le \lambda_n^{-1}(1 - e^{-\lambda_n}) \sum_{i=1}^{n} p_i^2.$$

In 2003, Neammanee gave a non-uniform pointwise bound when $\lambda_n \in (0, 1]$ and $x = 1, 2, \ldots, n-1$

$$|P(S_n = x) - P(U_{\lambda_n} = x)| \le \frac{1}{x} \sum_{i=1}^{n} p_i^2.$$

In the same year, he generalized his result to the case of any positive λ_n.

Theorem 1.2 (Neammanee, 2003) *For $\lambda_n > 0$ and $x = 1, 2, \ldots, n-1$, then*

$$|P(S_n = x) - P(U_{\lambda_n} = x)| \le \min\{\frac{1}{x}, \lambda_n^{-1}\} \sum_{i=1}^{n} p_i^2.$$

In 2005, Teerapabolarn and Neammanee also gave a non-uniform bound as follow.

Theorem 1.3 (Teerapabolarn & Neammanee, 2005) *We have*

$$|P(S_n \le x) - P(U_{\lambda_n} \le x)| \le \lambda_n^{-1}(1 - e^{-\lambda_n}) \min\left\{1, \frac{e^{\lambda_n}}{x+1}\right\} \sum_{i=1}^{n} p_i^2$$

where $x \in \{0, 1, \ldots, n\}$.

Let $X_1, X_2, \ldots$ be a sequence of independent Bernoulli random variables and N a positive integer-value random variable. Assume that $N, X_1, X_2, \ldots$ are independent. Define $S_N = X_1 + X_2 + \cdots + X_N$ which is called *random sums* and $\lambda_N = \sum_{i=1}^{N} p_i$ and $\lambda = E\lambda_N$. In 1991, Yannaros gave uniform bounds between the distribution of S_N and U_λ. The following is his result.

Theorem 1.4 (Yannaros, 1991) *We have*

1. $\sup_{x\in\mathbb{Z}_0^+} |P(S_N \le x) - P(U_\lambda \le x)| \le E|\lambda_N - E\lambda_N| + E\Big(\frac{1 - e^{-\lambda_N}}{\lambda_N} \sum_{i=1}^{N} p_i^2\Big).$

2. *If $p_1 = p_2 = \cdots = p$, then*

$$\sup_{x\in\mathbb{Z}_0^+} |P(S_N \le x) - P(U_{pEN} \le x)| \le \min\left\{\frac{p}{2\sqrt{1-p}}, pE(1 - e^{-pN})\right\} + \frac{1}{2}\sqrt{\frac{p\mathrm{Var}(N)}{EN}} \min\left\{1, 2\sqrt{pEN}\right\}.$$

In this paper, we will find uniform and non-uniform bounds for random sums by Poisson approximation. The followings are our results.

Theorem 1.5 *We have*

1. $|P(S_N = x) - P(U_\lambda = x)| \le \frac{7E\lambda_N}{2x}$ *where $x \in \{1, 2, \ldots\}$,*

2. $\sup_{x\in\mathbb{Z}^+} |P(S_N = x) - P(U_\lambda = x)| \le \frac{3}{2}E\lambda_N + 2\min\{E\lambda_N, E|\lambda - \lambda_N|\}.$

Note that, when $x = 0$ we can compute the exact probability, that is,

$$P(S_N = 0) = \sum_{n=1}^{\infty} P(N = n) \prod_{i=1}^{n} (1 - p_i) = E \prod_{i=1}^{N} (1 - p_i).$$

Theorem 1.6 *For $x \in \{1, 2, \ldots\}$, we have*

$$|P(S_N \le x) - P(U_\lambda \le x)| \le \frac{3E\lambda_N}{x} + E\Big[\lambda_N^{-1}(1 - e^{-\lambda_N}) \min\left\{1, \frac{e^{\lambda_N}}{x+1}\right\} \sum_{i=1}^{N} p_i^2\Big].$$

If X_i's are identically distributed, we have the following result.

Corollary 1.7 *If* $p_1 = p_2 = \cdots = p$, *then*

1. $|P(S_N = x) - P(U_\lambda = x)| \leq \frac{7pEN}{2x}$ *where* $x \in \{1, 2, \ldots\}$,

2. $\sup_{x\in\mathbb{Z}^+} |P(S_N = x) - P(U_\lambda = x)| \leq \frac{3pEN}{2} + 2p\min\{EN, E|N - EN|\}$,

3. $|P(S_N \leq x) - P(U_\lambda \leq x)| \leq \frac{3pEN}{x} + pE\Big[(1 - e^{-pN})\min\Big\{1, \frac{e^{pN}}{x+1}\Big\}\Big]$ *where* $x \in \{1, 2, \ldots\}$.

2. Proof of Main Results

Proof of Theorem 1.5

1. Let $x \in \{1, 2, \ldots\}$. Note that

$$\begin{aligned}
|P(S_N = x) - P(U_\lambda = x)| &= \Big|\sum_{n=1}^{\infty} P(N = n)P(S_n = x) - P(U_\lambda = x)\Big| \\
&= \sum_{n=1}^{\infty} P(N = n)|P(S_n = x) - P(U_\lambda = x)| \\
&\leq \sum_{n=1}^{\infty} P(N = n)|P(U_{\lambda_n} = x) - P(U_\lambda = x)| + \sum_{n=1}^{\infty} P(N = n)|P(S_n = x) - P(U_{\lambda_n} = x)| \\
&=: A_1 + A_2. \qquad (2.1)
\end{aligned}$$

By Chebyshev's inequality, we obtain

$$|P(U_{\lambda_n} = x) - P(U_\lambda = x)| \leq P(U_{\lambda_n} \geq x) + P(U_\lambda \geq x) \leq \frac{EU_{\lambda_n}}{x} + \frac{EU_\lambda}{x} = \frac{\lambda_n + \lambda}{x},$$

and then

$$A_1 \leq \frac{2E\lambda_N}{x}. \qquad (2.2)$$

To bound A_2, we note that

$$A_2 = \sum_{\substack{n=1\\ n\neq x}}^{\infty} P(N = n)|P(S_n = x) - P(U_{\lambda_n} = x)| + P(N = x)|P(S_x = x) - P(U_{\lambda_x} = x)| =: A_{21} + A_{22}. \qquad (2.3)$$

By Theorem 1.2 and the fact that $P(S_n = x) = 0$ for $n = 1, 2, \ldots, x - 1$, we have

$$\begin{aligned}
A_{21} &= \sum_{n=1}^{x-1} P(N = n)|P(S_n = x) - P(U_{\lambda_n} = x)| + \sum_{n=x+1}^{\infty} P(N = n)|P(S_n = x) - P(U_{\lambda_n} = x)| \\
&\leq \sum_{n=1}^{x-1} P(N = n)P(U_{\lambda_n} = x) + \frac{1}{x}\sum_{n=x+1}^{\infty} P(N = n)\sum_{i=1}^{n} p_i^2 \\
&\leq \frac{1}{x}\sum_{n=1}^{x-1} P(N = n)EU_{\lambda_n} + \frac{1}{x}\sum_{n=x+1}^{\infty} P(N = n)\lambda_n \\
&= \frac{1}{x}\sum_{\substack{n=1\\ n\neq x}}^{\infty} P(N = n)\lambda_n. \qquad (2.4)
\end{aligned}$$

To bound A_{22}, we note that

$$P(S_x = x) = \prod_{i=1}^{x} p_i \leq (\prod_{i=1}^{x} p_i)^{\frac{1}{x}} \leq \frac{p_1 + p_2 + \cdots + p_x}{x} = \frac{\lambda_x}{x} \tag{2.5}$$

by applying AM-GM inequality.

Next, we will show that

$$P(U_{\lambda_x} = x) \leq \frac{\lambda_x}{2x} \text{ for } x = 2, 3, \ldots. \tag{2.6}$$

Assume that $x \geq 2$. If $\lambda_x \leq x - 1$, then

$$e^{\lambda_x} \geq \frac{\lambda_x^{x-2}}{(x-2)!} + \frac{\lambda_x^{x-1}}{(x-1)!} = \frac{\lambda_x^{x-1}}{(x-1)!}(\frac{x-1}{\lambda_x} + 1) \geq \frac{2\lambda_x^{x-1}}{(x-1)!}$$

this implies that

$$P(U_{\lambda_x} = x) = \frac{e^{-\lambda_x}\lambda_x^x}{x!} \leq \frac{\lambda_x}{2x}. \tag{2.7}$$

If $\lambda_x = x$, we have

$$e^{\lambda_x} \geq \frac{\lambda_x^{x-1}}{(x-1)!} + \frac{\lambda_x^x}{x!} = \frac{\lambda_x^x}{x!}(\frac{x}{\lambda} + 1) \geq \frac{2\lambda_x^x}{x!}.$$

Hence

$$P(U_{\lambda_x} = x) = \frac{e^{-\lambda_x}\lambda_x^x}{x!} \leq \frac{1}{2}. \tag{2.8}$$

Combine (2.7) and (2.8), we obtain (2.6).

Hence, by (2.5) and (2.6), for $x = 2, 3, \ldots$

$$|P(S_x = x) - P(U_{\lambda_x} = x)| \leq \prod_{i=1}^{x} p_i + \frac{e^{-\lambda_x}\lambda_x^x}{x!} \leq \frac{\lambda_x}{x} + \frac{\lambda_x}{2x} = \frac{3\lambda_x}{2x}. \tag{2.9}$$

Observe that if $x = 1$, then

$$|P(S_x = x) - P(U_{\lambda_x} = x)| = |p_1 - e^{-1}p_1| = p_1|1 - e^{-p_1}| \leq p_1 \leq \frac{3\lambda_1}{2}. \tag{2.10}$$

By (2.9) and (2.10),

$$|P(S_x = x) - P(U_{\lambda_x} = x)| \leq \frac{3\lambda_x}{2x} \text{ for } x = 1, 2, \ldots. \tag{2.11}$$

By (2.3), (2.4) and (2.11), we have

$$A_2 \leq \frac{1}{x}\sum_{\substack{n=1\\ n\neq x}}^{\infty} P(N = n)\lambda_n + \frac{3}{2x}P(N = x)\lambda_x \leq \frac{3}{2x}E\lambda_N. \tag{2.12}$$

Hence, by (2.1), (2.2) and (2.11), we obtain (1) as desire.

2. Freedman (1974, pp. 260) showed that for all $\mu_1, \mu_2 > 0$

$$\sup_{x\in\mathbb{Z}_0^+} |P(U_{\mu_1} \leq x) - P(U_{\mu_2} \leq x)| \leq |\mu_1 - \mu_2|. \tag{2.13}$$

Then

$$\begin{aligned}
A_1 &= \sum_{n=1}^{\infty} P(N=n)|P(U_{\lambda_n}=x)-P(U_\lambda=x)| \\
&\leq \sum_{n=1}^{\infty} P(N=n)\{|P(U_{\lambda_n}\leq x)-P(U_\lambda\leq x)|+|P(U_\lambda\leq x-1)-P(U_{\lambda_n}\leq x-1)|\} \\
&\leq 2\sum_{n=1}^{\infty} P(N=n)|\lambda-\lambda_n| \\
&= 2E|\lambda-\lambda_N|. \qquad (2.14)
\end{aligned}$$

By (2.1), (2.2), (2.12) and (2.14), we conclude that

$$\sup_{x\in\mathbb{Z}^+}|P(S_N=x)-P(U_\lambda=x)|\leq \frac{3}{2}E\lambda_N+2\min\{E\lambda_N, E|\lambda-\lambda_N|\}.$$

Proof of Theorem 1.6

By the same argument of (2.2), we have

$$B_1 := \sum_{n=1}^{\infty} P(N=n)|P(U_{\lambda_n}\leq x)-P(U_\lambda\leq x)|\leq \sum_{n=1}^{\infty} P(N=n)|P(U_{\lambda_n}\geq x)+P(U_\lambda\geq x)|\leq \frac{2E\lambda_N}{x}. \qquad (2.15)$$

Using the fact that $P(S_n\leq x)=1$ for $n=1,2,\ldots,x$, we obtain

$$\begin{aligned}
B_2 &:= \sum_{n=1}^{x} P(N=n)|P(S_n\leq x)-P(U_{\lambda_n}\leq x)| \\
&\leq \sum_{n=1}^{x} P(N=n)P(U_{\lambda_n}\geq x) \\
&\leq \frac{1}{x}\sum_{n=1}^{\infty} P(N=n)\lambda_n \\
&= \frac{1}{x}E\lambda_N. \qquad (2.16)
\end{aligned}$$

By Theorem 1.3, we get

$$\begin{aligned}
B_3 &:= \sum_{n=x+1}^{\infty} P(N=n)|P(S_n\leq x)-P(U_{\lambda_n}\leq x)| \\
&\leq \sum_{n=x+1}^{\infty} P(N=n)\lambda_n^{-1}(1-e^{-\lambda_n})\min\left\{1,\frac{e^{\lambda_n}}{x+1}\right\}\sum_{i=1}^{n}p_i^2 \\
&\leq E[\lambda_N^{-1}(1-e^{-\lambda_N})\min\left\{1,\frac{e^{\lambda_N}}{x+1}\right\}\sum_{i=1}^{N}p_i^2]. \qquad (2.17)
\end{aligned}$$

From the fact that

$$|P(S_n=x)-P(U_\lambda=x)|\leq B_1+B_2+B_3$$

and (2.15) – (2.17), we complete the proof.

3. Examples

Applying our main results together with the facts that

1. $E\lambda_N=\frac{1}{2}(\lambda_n+\lambda_{2n})$ and $E|\lambda_N-E\lambda_N|=\frac{1}{2}(\lambda_{2n}-\lambda_n)$,
2. $EN=2$ and $E|N-EN|=1$ and

3. $EN = \mu$ and $E|N - EN| = 2\mu e^{-\mu}$

in Example 3.1–Example 3.3, respectively, we conclude the following bounds.

Example 3.1 Fix $n \in \mathbb{N}$, let N be random variable defined by

$$P(N = n) = \frac{1}{2} \quad \text{and} \quad P(N = 2n) = \frac{1}{2}.$$

Then

1. $|P(S_N = x) - P(U_\lambda = x)| \leq \dfrac{7(\lambda_n + \lambda_{2n})}{4x}$ for $x = 1, 2, \ldots,$

2. $\sup\limits_{x\in\mathbb{Z}^+} |P(S_N = x) - P(U_\lambda = x)| \leq \dfrac{1}{4}(7\lambda_{2n} - \lambda_n)$ and

3. $|P(S_N \leq x) - P(U_\lambda \leq x)| \leq \dfrac{3(\lambda_n + \lambda_{2n})}{2x} + \dfrac{1}{2(x+1)}\Big\{\dfrac{e^{\lambda_n} - 1}{\lambda_n}\sum\limits_{i=1}^{n} p_i^2 + \dfrac{e^{\lambda_{2n}} - 1}{\lambda_{2n}}\sum\limits_{i=1}^{2n} p_i^2\Big\}$ for $x = 1, 2, \ldots$

where $\lambda = \frac{1}{2}(\lambda_n + \lambda_{2n})$. Furthermore if $p_1 = p_2 = \cdots = p$, then

1. $|P(S_N = x) - P(U_{3np/2} = x)| \leq \dfrac{21np}{4x}$ for $x = 1, 2, ...,$

2. $\sup\limits_{x\in\mathbb{Z}^+} |P(S_N = x) - P(U_{3np/2} = x)| \leq \dfrac{13np}{4}$ and

3. $|P(S_N \leq x) - P(U_{3np/2} \leq x)| \leq \dfrac{9np}{2x} + \dfrac{p(e^{np} + e^{2np} - 2)}{2(x+1)}$ for $x = 1, 2, \ldots.$ (3.1)

Example 3.2 Let N be random variable defined by

$$P(N = n) = \frac{1}{2^n}$$

for all $n \in \{1, 2, \ldots\}$. Assume that $p_1 = p_2 = \cdots = p$. Then

1. $|P(S_N = x) - P(U_{2p} = x)| \leq \dfrac{7p}{x}$ for $x = 1, 2, \ldots,$

2. $\sup\limits_{x\in\mathbb{Z}^+} |P(S_N = x) - P(U_{2p} = x) \leq 5p$ and

3. if $e^p < 2$, then $|P(S_N \leq x) - P(U_{2p} \leq x)| \leq \dfrac{6p}{x} + \dfrac{2p}{x+1}\Big(\dfrac{e^p - 1}{2 - e^p}\Big)$ for $x = 1, 2, \ldots.$ (3.2)

Example 3.3 Let $0 < \mu \leq 1$ and let N be a random variable defined by

$$P(N = n) = \frac{e^{-\mu}\mu^n}{n!} \quad \text{for } n = 0, 1, 2, \ldots.$$

Assume that $p_1 = p_2 = \cdots = p$. Then

1. $|P(S_N = x) - P(U_{\mu p} = x)| \leq \dfrac{7\mu p}{2x}$ for $x = 1, 2, \ldots,$

2. $\sup\limits_{x\in\mathbb{Z}^+} |P(S_N = x) - P(U_{\mu p} = x)| \leq \dfrac{7\mu p}{2}$ and

3. $|P(S_N \le x) - P(U_{\mu p} \le x)| \le \dfrac{3\mu p}{x} + \dfrac{p(e^{\mu(e^p-1)} - 1)}{x+1}$ for $x = 1, 2, \ldots$. (3.3)

Remark 3.4 In the case of i.i.d., the uniform bounds of Yannaros (Theorem 1.4 (2)) in Example 3.1 – Example 3.3 are

(2.1) $\min\Big\{\dfrac{p}{2\sqrt{1-p}}, p\Big(1 - \dfrac{e^{-np} + e^{-2np}}{2}\Big)\Big\} + \dfrac{\sqrt{np}}{2}$,

(2.2) $\min\Big\{\dfrac{p}{2\sqrt{1-p}}, p\Big(1 - \dfrac{1}{2e^p - 1}\Big)\Big\} + \sqrt{2}p$ and

(2.3) $\min\Big\{\dfrac{p}{2\sqrt{1-p}}, p\Big(1 - e^{\mu(e^{-p}-1)}\Big)\Big\} + \sqrt{\mu}p$,

respectively. We observe that the bounds of Yannaros ((2.1)–(2.3)) and our non-uniform bounds ((3.1)–(3.3)) have the same order but our bounds are better if x is large enough.

Acknowledgments

The authors would like to thank K. Neammanee for his valuable comments and suggestions which have helped improving our work.

References

Barbour, A. D., & Hall, P. (1984). On the rate of Poisson convergence. *Math. Proc. Cambridge Philos Soc., 95*, 473-480. http://dx.doi.org/10.1017/S0305004100061806

Chen, L. H. Y. (1974). On the convergence of Poisson binomial to Poisson distributions. *Ann. Probab., 2*, 178-180. http://dx.doi.org/10.1214/aop/1176996766

Freedman, D. (1974). The Poisson approximation for dependent events. *Ann. Probab., 2*, 256-269. http://dx.doi.org/10.1214/aop/1176996707

Kerstan, J. (1964). Verallgemeinerung eines Satzes von Procharov und Le Cam. *Z. Wahrsch. Verw. Febiete, 2*, 173-179. http://dx.doi.org/10.1007/BF00533378

Le Cam, L. (1960). An approximation theorem for the Poisson binomial distribution. *Pacific J. Math., 10*, 1181-1197.

Neammanee, K. (2003). A non-uniform bound for the approximation of Poisson binomial by Poisson distribution. *Int. J. Math. Math. Sci., 48*, 3041-3046.

Neammanee, K. (2003). Pointwise approximation of Poisson binomial by Poisson distribution. *Int. J. Stochastic Model. Appl., 6*, 20-26.

Teerapabolarn, K., & Neammanee, K. (2005). Poisson approximation for sums of dependent Bernoulli random variables. *Acta Math. Acad. Paedagog. Nyhazi. (N. S.), 22*, 87-99.

Yannaros, N. (1991). Poisson approximation for random sums of Bernoulli random variables. *Statist. Probab. Lett., 11*, 161-165. http://dx.doi.org/10.1016/0167-7152(91)90135-E