International Journal of Pure and Applied Mathematics
Volume 80 No. 5 2012, 711-719
ISSN: 1311-8080 (printed version)
url: http://www.ijpam.eu

ijpam.eu

AN IMPROVEMENT OF BOUND ON THE POISSON-BINOMIAL RELATIVE ERROR

K. Teerapabolarn

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

Abstract: We use the Stein-Chen method to obtain new uniform bounds for the relative error of the binomial cumulative distribution function with parameters n and p and the Poisson cumulative distribution function with mean $\lambda = np$. The results obtained in this study are better than those reported in [7].

AMS Subject Classification: 62E17, 60F05
Key Words: cumulative distribution function, Poisson approximation, relative error, Stein-Chen method

1. Introduction

Let X be a discrete random variable with the probability mass function

$$p_X(x) = \binom{n}{x} p^x q^{n-x},\ x = 0, 1, ..., n, \tag{1.1}$$

for $n \in \mathbb{N}$, $p \in (0, 1)$ and $q = 1 - p$. With the probability function, the distribution of X is called the binomial distribution with parameters n and p, where np and npq are its mean and variance. Consider the probability function (1.1), if $n \to \infty$ and $p \to 0$ while $\lambda = np$ remains fixed then $\binom{n}{x} p^x q^{n-x} \to \frac{e^{-\lambda}\lambda^x}{x!}$ for every $x = 0, 1, ..., n$. Therefore, the binomial distribution with parameters n and p can be approximated by the Poisson distribution with mean $\lambda = np$

Received: August 21, 2012

when n is sufficiently large and p is sufficiently small. Correspondingly, the Poisson cumulative distribution function can also be used as an approximation of the binomial cumulative distribution function when n is sufficiently large and p is sufficiently small. In the past, there has been some research related to approximation relations between the Poisson and binomial cumulative distribution functions. For example, Anderson and Samuels [1] gave an inequality of the error

$$\mathbb{P}_\lambda(x_0) - \mathbb{B}_{n,p}(x_0) \begin{cases} > 0 & \text{if } x_0 \leq \dfrac{\lambda n}{n+1}, \\ < 0 & \text{if } x_0 \geq \lambda, \end{cases} \tag{1.2}$$

where $\mathbb{P}_\lambda(x_0) = \sum_{k=0}^{x_0} \frac{e^{-\lambda}\lambda^k}{k!}$ and $\mathbb{B}_{n,p}(x_0) = \sum_{k=0}^{x_0} \binom{n}{k} p^k q^{n-k}$ are the Poisson and binomial cumulative distribution functions at $x_0 \in \{0, ..., n\}$. Ivchenko [3] gave the asymptotic relation on the ratio of these distribution functions

$$\frac{\mathbb{B}_{n,p}(x_0)}{\mathbb{P}_\lambda(x_0)} = 1 + o(1), \tag{1.3}$$

which is uniformly in $x_0 < \lambda$. Because the relation (1.2) does not give any conditions for a good approximation and the relation (1.3) is obtained on the specific region of x_0. Afterward, Teerapabolarn [6] used the Stein-chen method to give a non-uniform bound for each relative error of two such cumulative distributions functions as follows:

$$\left| \frac{\mathbb{P}_\lambda(x_0)}{\mathbb{B}_{n,p}(x_0)} - 1 \right| \leq \frac{(e^\lambda - 1)p\Delta(x_0)}{x_0 + 1}, \; x_0 = 0, 1, ..., n, \tag{1.4}$$

where $\Delta(x_0) = e^{-\lambda} q^{-n}$ if $x_0 < \lambda$ and $\Delta(x_0) = 1$ if $x_0 \geq \lambda$ and

$$\left| \frac{\mathbb{B}_{n,p}(x_0)}{\mathbb{P}_\lambda(x_0)} - 1 \right| \leq \frac{(e^\lambda - 1)p}{x_0 + 1}, \; x_0 = 0, 1, ..., n. \tag{1.5}$$

The bounds in (1.4) and (1.5) depend on the value of x_0 and we must always consider the value of x_0 where, in some cases, consideration the value of x_0 is not a convenience. So, Teerapabolarn [7] used the Stein-Chen method and the characterization associated with the binomial random variable to give uniform bounds for the relative error in (1.4) and (1.5) as follows:

$$\sup_{0 \leq x_0 \leq n} \left| \frac{\mathbb{P}_\lambda(x_0)}{\mathbb{B}_{n,p}(x_0)} - 1 \right| \leq \frac{(1 - e^{-\lambda})(1 - q^n)}{nq^n} \tag{1.6}$$

and

$$\sup_{0\le x_0\le n}\left|\frac{\mathbb{B}_{n,p}(x_0)}{\mathbb{P}_\lambda(x_0)}-1\right|\le\frac{(e^\lambda-1)(1-q^n)}{n}. \tag{1.7}$$

In this paper, we use the Stein-Chen method to improve uniform bounds in (1.6) and (1.7) sharper than ever, which are described in Section 2 and 3, respectively. In Section 4, numerical examples are provided to show applications of the results. Conclusion of this study is presented in the last section.

2. Method

Stein's method was first introduced by Stein [4]. The version for the Poisson case was first developed by Chen [2], which is referred to as the Stein-Chen method. This method is an important tool for giving the results.

Following Teerapabolarn [6], Stein's equation of the Poisson cumulative distribution function with parameter $\lambda > 0$ is of the form

$$h_{x_0}(x)-\mathbb{P}_\lambda(x_0)=\lambda f_{x_0}(x+1)-xf_{x_0}(x) \tag{2.1}$$

for $h_{x_0}(x)=1$ if $x\le x_0$ and $h_{x_0}(x)=0$ if $x>x_0$ and

$$f_{x_0}(x)=\begin{cases}(x-1)!\lambda^{-x}e^\lambda[\mathbb{P}_\lambda(x-1)[1-\mathbb{P}_\lambda(x_0)]] & \text{if } x\le x_0,\\ (x-1)!\lambda^{-x}e^\lambda[\mathbb{P}_\lambda(x_0)[1-\mathbb{P}_\lambda(x-1)]] & \text{if } x> x_0,\\ 0 & \text{if } x=0,\end{cases} \tag{2.2}$$

where $x_0, x\in\mathbb{N}\cup\{0\}$.

Lemma 2.1. *For $x_0, x\in\mathbb{N}$, let $\Delta f_{x_0}(x)=f_{x_0}(x+1)-f_{x_0}(x)$. Then the following inequality holds:*

$$\sup_{x\ge 1}|\Delta f_{x_0}(x)|\le\lambda^{-2}(e^\lambda-\lambda-1)\mathbb{P}_\lambda(x_0)\min\left\{1,\frac{2}{x}\right\}. \tag{2.3}$$

Proof. For $x\le x_0$, because Δf_{x_0} is an increasing function for $x\le x_0$ (Lemma 2.1 in [5]), we obtain

$$\begin{aligned}\Delta f_{x_0}(x)&\le\Delta f_{x_0}(x_0)\\ &=(x_0-1)!e^{-\lambda}\sum_{k=0}^{x_0}(x_0-k)\frac{\lambda^k}{k!}\sum_{j=x_0+1}^{\infty}\frac{\lambda^{j-(x_0+1)}}{j!}\quad\text{(by (2.4) in [5])}\end{aligned}$$

$$\leq \mathbb{P}_\lambda(x_0)x_0! \sum_{j=x_0+1}^{\infty} \frac{\lambda^{j-(x_0+1)}}{j!}$$
$$= \frac{\mathbb{P}_\lambda(x_0)}{x_0+1}\left\{1+\frac{\lambda}{x_0+2}+\frac{\lambda^2}{(x_0+2)(x_0+3)}+\cdots\right\}$$

$$\leq \frac{\mathbb{P}_\lambda(x_0)\lambda^{-2}}{x_0+1}\left\{\lambda^2+\frac{\lambda^3}{3}+\frac{\lambda^4}{12}+\cdots\right\}$$
$$\leq \lambda^{-2}(e^\lambda-\lambda-1)\mathbb{P}_\lambda(x_0)\min\left\{1,\frac{2}{x}\right\}.$$

For $x > x_0$, because Δf_{x_0} is a negative function for this case (Lemma 2.1 in [6]), we have

$$\begin{aligned}
0 &< -\Delta f_{x_0}(x) \\
&= (x-1)!e^{-\lambda}\sum_{j=0}^{x_0}\frac{\lambda^j}{j!}\sum_{k=x+1}^{\infty}(k-x)\frac{\lambda^{k-(x+1)}}{k!} \quad \text{(by (2.4) in [5])} \\
&= \mathbb{P}_\lambda(x_0)\frac{(x-1)!}{x!}\left\{\frac{1}{x+1}+\frac{2\lambda}{(x+1)(x+2)}+\cdots\right\} \\
&\leq \frac{\mathbb{P}_\lambda(x_0)}{x}\lambda^{-2}\left\{\frac{\lambda^2}{3}+\frac{\lambda^3}{6}+\frac{\lambda^4}{20}+\cdots\right\} \\
&\leq \lambda^{-2}(e^\lambda-\lambda-1)\mathbb{P}_\lambda(x_0)\min\left\{1,\frac{2}{x}\right\}.
\end{aligned}$$

Hence, from two cases, (2.3) holds. □

Lemma 2.2. *For $x_0 \in \{0,1,...,n\}$ and $\lambda = np$, then*

$$\frac{\mathbb{P}_\lambda(x_0)}{\mathbb{B}_{n,p}(x_0)} \leq e^{-\lambda}q^{-n} \quad \textit{(Teerapabolarn [7])}. \tag{2.4}$$

3. Main Results

The main results of this study are new uniform bounds on two forms of the relative error of the binomial and Poisson cumulative distribution functions.

Theorem 3.1. *With the above definitions, we have the following:*

$$\sup_{0\le x_0\le n}\left|\frac{\mathbb{P}_\lambda(x_0)}{\mathbb{B}_{n,p}(x_0)}-1\right| \le \max\left\{e^{-\lambda}q^{-n}-1, \frac{1-(1+\lambda)e^{-\lambda}}{nq^n}\min\left(1,\frac{2(1-q^n)}{\lambda}\right)\right\}. \quad (3.1)$$

Proof. The first bound follows from $\left|\frac{\mathbb{P}_\lambda(0)}{\mathbb{B}_{n,p}(0)}-1\right|$. Next, we shall derive the second bound. Teerapabolarn [7] showed that

$$\begin{aligned}|\mathbb{B}_{n,p}(x_0)-\mathbb{P}_\lambda(x_0)| &\le p\sum_{x=1}^{n}|\Delta f_{x_0}(x)|\, xp_X(x)\\ &\le p\sum_{x=1}^{n}\sup_{x\ge 1}|\Delta f_{x_0}(x)|\, xp_X(x)\\ &\le p\sum_{x=1}^{n}\frac{(e^\lambda-\lambda-1)\mathbb{P}_\lambda(x_0)}{\lambda^2}\min\left\{1,\frac{2}{x}\right\}xp_X(x) \text{ (by (2.3))}\\ &= \frac{(e^\lambda-\lambda-1)\mathbb{P}_\lambda(x_0)}{n}\min\left\{1,\frac{2(1-q^n)}{\lambda}\right\},\end{aligned}$$

and dividing by $\mathbb{B}_{n,p}(x_0)$ and using the fact that $e^{-\lambda}q^{-n}\ge 1$, we obtain

$$\begin{aligned}\left|\frac{\mathbb{P}_\lambda(x_0)}{\mathbb{B}_{n,p}(x_0)}-1\right| &\le \frac{(e^\lambda-\lambda-1)\mathbb{P}_\lambda(x_0)}{n\mathbb{B}_{n,p}(x_0)}\min\left\{1,\frac{2(1-q^n)}{\lambda}\right\}\\ &\le \frac{1-(1+\lambda)e^{-\lambda}}{nq^n}\min\left\{1,\frac{2(1-q^n)}{\lambda}\right\}\end{aligned}$$

for every $1\le x_0\le n$. Hence, (3.1) is obtained. □

Corollary 3.1. *We have the following inequality*

$$\sup_{0\le x_0\le n}\left|\frac{\mathbb{B}_{n,p}(x_0)}{\mathbb{P}_\lambda(x_0)}-1\right| \le \max\left\{1-e^\lambda q^n, \frac{e^\lambda-\lambda-1}{n}\min\left(1,\frac{2(1-q^n)}{\lambda}\right)\right\}. \quad (3.2)$$

Proof. The proof is similar to that of Theorem 3.1. □

Remark. 1. Consider the results in Theorem 3.1 and Corollary 3.1, if p or λ is small and n is large, then the bounds in (3.1) and (3.2) approach 0. It indicates that the Poisson cumulative distribution function can be used as an approximation of the binomial cumulative distribution function when p or λ is sufficiently small and n is sufficiently large.

2. By comparing between the bounds in (1.6) and (1.7) and the bounds in Theorem 3.1 and Corollary 3.1, because $e^{-\lambda}q^{-n} - 1 \leq \sup_{0\leq x_0\leq n}\left|\frac{\mathbb{P}_\lambda(x_0)}{\mathbb{B}_{n,p}(x_0)} - 1\right| \leq \frac{(1-e^{-\lambda})(1-q^n)}{nq^n}$ and $1-e^{\lambda}q^n \leq \sup_{0\leq x_0\leq n}\left|\frac{\mathbb{B}_{n,p}(x_0)}{\mathbb{P}_\lambda(x_0)} - 1\right| \leq \frac{(e^{\lambda}-1)(1-q^n)}{n}$ and $2\lambda^{-1}(e^{\lambda}-\lambda-1) < e^{\lambda}-1$, the bounds in Theorem 3.1 and Corollary 3.1 are sharper than the bounds in (1.6) and (1.7), respectively.

4. Numerical Examples

4.1. Let $n = 100$ and $p = 0.005$, then $\lambda = 0.5$ and the numerical result of Theorem 3.1 is of the form

$$\sup_{0\leq x_0\leq 100}\left|\frac{\mathbb{P}_{0.5}(x_0)}{\mathbb{B}_{100,0.005}(x_0)} - 1\right| \leq 0.00148908,$$

which is better than the numerical result obtained from (1.6),

$$\sup_{0\leq x_0\leq 100}\left|\frac{\mathbb{P}_{0.5}(x_0)}{\mathbb{B}_{100,0.005}(x_0)} - 1\right| \leq 0.00256066.$$

The numerical result of Corollary 3.1 is of the form

$$\sup_{0\leq x_0\leq 100}\left|\frac{\mathbb{B}_{100,0.005}(x_0)}{\mathbb{P}_{0.5}(x_0)} - 1\right| \leq 0.00148721,$$

which is also better than the numerical result obtained from (1.7),

$$\sup_{0\leq x_0\leq 100}\left|\frac{\mathbb{B}_{100,0.005}(x_0)}{\mathbb{P}_{0.5}(x_0)} - 1\right| \leq 0.00255745.$$

4.2. Let $n = 100$ and $p = 0.01$, then $\lambda = 1.0$ and the numerical result of Theorem 3.1 is of the form

$$\sup_{0\leq x_0\leq 100}\left|\frac{\mathbb{P}_{1.0}(x_0)}{\mathbb{B}_{100,0.01}(x_0)} - 1\right| \leq 0.00721906,$$

which is better than the numerical result obtained from (1.6),

$$\sup_{0\leq x_0\leq 100}\left|\frac{\mathbb{P}_{1.0}(x_0)}{\mathbb{B}_{100,0.01}(x_0)} - 1\right| \leq 0.01094832.$$

The numerical result of Corollary 3.1 is of the form

$$\sup_{0\le x_0\le 100}\left|\frac{\mathbb{B}_{100,0.01}(x_0)}{\mathbb{P}_{1.0}(x_0)}-1\right|\le 0.00718282,$$

which is also better than the numerical result obtained from (1.7),

$$\sup_{0\le x_0\le 100}\left|\frac{\mathbb{B}_{100,0.01}(x_0)}{\mathbb{P}_{1.0}(x_0)}-1\right|\le 0.01089335.$$

4.3. Let $n = 250$ and $p = 0.01$, then $\lambda = 2.5$ and the numerical result of Theorem 3.1 is of the form

$$\sup_{0\le x_0\le 250}\left|\frac{\mathbb{P}_{2.5}(x_0)}{\mathbb{B}_{250,0.01}(x_0)}-1\right|\le 0.02585517,$$

which is better than the numerical result obtained from (1.6),

$$\sup_{0\le x_0\le 250}\left|\frac{\mathbb{P}_{2.5}(x_0)}{\mathbb{B}_{250,0.01}(x_0)}-1\right|\le 0.04162475.$$

The numerical result of Corollary 3.1 is of the form

$$\sup_{0\le x_0\le 250}\left|\frac{\mathbb{B}_{250,0.01}(x_0)}{\mathbb{P}_{2.5}(x_0)}-1\right|\le 0.02553185,$$

which is also better than the numerical result obtained from (1.7),

$$\sup_{0\le x_0\le 250}\left|\frac{\mathbb{B}_{250,0.01}(x_0)}{\mathbb{P}_{2.5}(x_0)}-1\right|\le 0.04110423.$$

4.4. Let $n = 1000$ and $p = 0.005$, then $\lambda = 5.0$ and the numerical result of Theorem 3.1 is of the form

$$\sup_{0\le x_0\le 1000}\left|\frac{\mathbb{P}_{5.0}(x_0)}{\mathbb{B}_{1000,0.005}(x_0)}-1\right|\le 0.05730038,$$

which is better than the numerical result obtained from (1.6),

$$\sup_{0\le x_0\le 1000}\left|\frac{\mathbb{P}_{5.0}(x_0)}{\mathbb{B}_{1000,0.005}(x_0)}-1\right|\le 0.14828037.$$

The numerical result of Corollary 3.1 is of the form

$$\sup_{0\le x_0\le 1000}\left|\frac{\mathbb{B}_{1000,0.005}(x_0)}{\mathbb{P}_{5.0}(x_0)}-1\right|\le 0.05658622,$$

which is also better than the numerical result obtained from (1.7),

$$\sup_{0\le x_0\le 1000}\left|\frac{\mathbb{B}_{1000,0.005}(x_0)}{\mathbb{P}_{5.0}(x_0)}-1\right|\le 0.14643228.$$

From 4.1−4.4, we see that the numerical results in Poisson approximation to binomial cumulative distribution are more accurate as p or λ is small and n is large. In addition, by numerical comparison, the bounds in Theorem 3.1, Corollary 3.1 are sharper than the corresponding bounds in (1.6) and (1.7).

5. Conclusion

In this study, the uniform bounds in Theorems 3.1 and Corollary 3.1, which were improved by the Stein-Chen method, provide new general criteria for measuring the accuracy in approximating the binomial cumulative distribution with parameter n and p by the Poisson cumulative distribution with mean $\lambda = np$. With these bounds, it is indicated that each result gives a good approximation when p or λ is sufficiently small and n is sufficiently large. Moreover, the results reported in this study are better than the bounds in [7], including both theoretical and numerical results.

References

[1] T.W. Anderson, S.M. Samuels, Some inequalities among binomial and Poisson probabilities, In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, **1** (1965), 1-12.

[2] L.H.Y. Chen, Poisson approximation for dependent trials, *Annals of Probability*, **3** (1975), 534-545.

[3] G.I. Ivchenko, On comparison of binomial and Poisson distributions, *Theory of Probability and its Applications*, **19** (1974), 584-587.

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

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

[6] K. Teerapabolarn, A bound on the Poisson-binomial relative error, *Statistical Methodology*, **4** (2007), 407-415.

[7] K. Teerapabolarn, A Poisson-binomial relative error uniform bound, *Statistical Methodology*, **7** (2010), 69-76.