International Journal of Pure and Applied Mathematics
Volume 82 No. 5 2013, 737-745
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: http://dx.doi.org/10.12732/ijpam.v82i5.7
ijpam.eu

THE LEAST UPPER BOUND ON THE POISSON-BINOMIAL RELATIVE ERROR

K. Teerapabolarn[1 §], A. Boondirek

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

Abstract: A simple mathematical method is used to determine the least upper bound on the relative error between the binomial cumulative distribution function with parameters n and p and the Poisson cumulative distribution function with mean $\lambda = \frac{np}{q} = \frac{np}{1-p}$. With this upper bound, the Poisson cumulative distribution function with this mean can be used as an estimate of the binomial cumulative distribution function when p is sufficiently small even if λ is rather large. By numerical comparison, the upper bound obtained in this study is sharper than those reported in Teerapabolarn [4].

AMS Subject Classification: 62E17, 60F05
Key Words: binomial distribution, cumulative distribution function, least upper bound, Poisson distribution, relative error

1. Introduction

Let the distribution of a non-negative integer-valued random variable X be defined as follows:

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

The distribution (1.1) is the well-known *binomial distribution* with parameters

Received: October 25, 2012

[§]Correspondence author

$n \in \mathbb{N}$ and $p \in (0,1)$, where the mean and variance of X are $E(X) = np$ and $Var(X) = npq$, respectively. In addition, by the experimental process of the binomial distribution, the random variable X can be represented as the number of successes in a sequence of n independent Bernoulli trials, where success occurs on each trial with a probability of p and failure occurs on each trial with a probability of $q = 1 - p$. It is well known that if the number of trials $n \to \infty$ and the probability of success $p \to 0$ while the product np remains fixed, then the binomial distribution with parameters n and p converges to the Poisson distribution with mean np, which is due to the *Law of Small Numbers*. Therefore the Poisson distribution with mean $\lambda = np$ can be used as an approximation to the binomial distribution with parameters n and p if n is sufficiently large and p is sufficiently small. Correspondingly, the binomial cumulative distribution function

$$\mathbb{B}_{n,p}(x) = \sum_{j=0}^{x} \binom{n}{j} p^j q^{n-j},\ x = 0, 1, ..., n \tag{1.2}$$

can be also approximated by the Poisson cumulative distribution function

$$\mathbb{P}_\lambda(x) = \sum_{j=0}^{x} \frac{e^{-\lambda}\lambda^j}{j!},\ x = 0, 1, ..., \tag{1.3}$$

for $\lambda = np$. In this case, some authors have tried to give some approximate relation between the binomial and Poisson cumulative distribution functions. For example, Anderson and Samuels [1] gave the error

$$\mathbb{P}_\lambda(x) - \mathbb{B}_{n,p}(x) \begin{cases} > 0 & \text{if } x \leq \frac{\lambda n}{n+1}, \\ < 0 & \text{if } x \geq \lambda, \end{cases} \tag{1.4}$$

for $x \in \{0, 1, ..., n\}$. Ivchenko [2] gave the asymptotic relation of the ratio between the binomial and Poisson cumulative distribution functions in the form of

$$\frac{\mathbb{B}_{n,p}(x)}{\mathbb{P}_\lambda(x)} = 1 + o(1) \tag{1.5}$$

for $x < \lambda$. It is seen that the approximate relations (1.4) and (1.5) do not give any conditions in order to have a good Poisson approximation and the relation (1.5) is obtained on the specific region of values of $x < \lambda$ only. Thus, in order to give a good approximation for $x \in \{0, 1, ..., n\}$, Teerapabolarn [3] used the Stein-Chen method to give a non-uniform bound on the relative error between the binomial and Poisson cumulative distribution functions, which is a criterion for measuring the accuracy of the approximation, as follows.

$$\left|1-\frac{\mathbb{P}_\lambda(x)}{\mathbb{B}_{n,p}(x)}\right| \leq \frac{p(e^\lambda-1)\Delta(x)}{x+1}, \; x=0,1,...,n, \tag{1.6}$$

where

$$\Delta(x)=\begin{cases}\frac{e^{-\lambda}}{q^n} & \text{if } \; x<\lambda,\\ 1 & \text{if } \; x\geq\lambda.\end{cases} \tag{1.7}$$

It is noted that the result (1.6) depends on the value of x and it must always consider the value of x for measuring the accuracy of the approximation where, in some cases, consideration of the value of x is not a convenience. Thus, Teerapabolarn [4] used the Stein-Chen method and the characterization associated with the binomial random variable to give a uniform bound for the approximation in (1.6), that is,

$$\max_{0\leq x\leq n}\left|1-\frac{\mathbb{P}_\lambda(x)}{\mathbb{B}_{n,p}(x)}\right| \leq \frac{(1-e^{-\lambda})(1-q^n)}{nq^n}. \tag{1.8}$$

Let us consider the distribution (1.1), by setting $\lambda=\frac{np}{q}$ or $p=\frac{\lambda}{n+\lambda}$, it follows that

$$\begin{aligned} p_X(x) &= q^n\frac{n!}{x!(n-x)!}\left(\frac{p}{q}\right)^x, \; x=0,1,...,n\\ &= \begin{cases}\left(\frac{1}{1+\frac{\lambda}{n}}\right)^n & \text{if } x=0,\\ \left(\frac{1}{1+\frac{\lambda}{n}}\right)^n \frac{\lambda^x}{x!}\left(\frac{n}{n}\cdots\frac{n-x+1}{n}\right) & \text{if } x=1,...,n.\end{cases}\end{aligned} \tag{1.9}$$

In view of (1.9), it can be observed that if $n\to\infty$ and $p\to 0$ while $\lambda=\frac{np}{q}$ remains fixed, then $p_X(x)=\binom{n}{x}p^xq^{n-x}\to\frac{e^{-\lambda}\lambda^x}{x!}$ for every $x\in\{0,1,...,n\}$, that is, the binomial distribution with parameters n and p also converges to the Poisson distribution with mean $\lambda=\frac{np}{q}$. Therefore the Poisson distribution with mean $\lambda=\frac{np}{q}$ can be used as an approximation of the binomial distribution with parameters n and p when n is sufficiently large and p is sufficiently small. In other words, the binomial cumulative distribution function can be properly approximated by the Poisson cumulative distribution function with mean $\lambda=\frac{np}{q}$. In this article, we are interested to approximate the binomial cumulative distribution function with parameters n and p by the Poisson cumulative distribution function with mean $\lambda = \frac{np}{q}$. For the accuracy of the approximation, it is measured in terms of the least upper bound on the relative error between two such cumulative distribution functions, which is similar to the result in (1.8).

In this study, a simple mathematical method is used to obtain the main result, which is mentioned in Section 2. In Section 3, numerical results are provided to illustrate the obtained result and we also compare the obtained result and the result in (1.8), and the conclusion of this study is presented in the last section.

2. Main Result

The result of this study is the least upper bound for the relative error between the binomial and Poisson cumulative distribution functions. This result can be directly derived by a simple mathematical method, which is the same method mentioned in [5]. The following theorem shows the least upper bound on the relative error between two such cumulative distribution functions.

Theorem 2.1. *Let $\lambda = \frac{np}{q}$, then the following inequality holds:*

$$\max_{0\leq x\leq n}\left|1-\frac{\mathbb{P}_\lambda(x)}{\mathbb{B}_{n,p}(x)}\right| = \max_{0\leq x\leq n}\left\{1-\frac{\mathbb{P}_\lambda(x)}{\mathbb{B}_{n,p}(x)}\right\} = 1-e^{-\lambda}q^{-n}. \tag{2.1}$$

Proof. We shall show that (2.1) holds. Let $\Delta(x) = 1-\dfrac{\mathbb{P}_\lambda(x)}{\mathbb{B}_{n,p}(x)}$. It is observed that if $\Delta(x)$ is decreasing for $x \in \{1,...,n\}$, then $\Delta(0) = \Delta(1) > \Delta(2) > \cdots > \Delta(n) = 1-\mathbb{P}_\lambda(n) > 0$. Note that $\Delta(x)-\Delta(x+1) > 0$ if and only if $\mathbb{P}_\lambda(x+1)\mathbb{B}_{n,p}(x)-\mathbb{P}_\lambda(x)\mathbb{B}_{n,p}(x+1) > 0$. Let $\delta(x) = \mathbb{P}_\lambda(x+1)\mathbb{B}_{n,p}(x)-\mathbb{P}_\lambda(x)\mathbb{B}_{n,p}(x+1)$, then we have

$$\begin{aligned}
\delta(x) = e^{-\lambda}q^n\Bigg\{&\left[\sum_{k=0}^{x}\frac{\lambda^k}{k!}+\frac{\lambda^{x+1}}{(x+1)!}\right]\left[1+\sum_{j=1}^{x}\frac{\lambda^j}{j!}\left(\frac{n}{n}\cdots\frac{n-j+1}{n}\right)\right] \\
&-\sum_{k=0}^{x}\frac{\lambda^k}{k!}\left[1+\sum_{j=1}^{x}\frac{\lambda^j}{j!}\left(\frac{n}{n}\cdots\frac{n-j+1}{n}\right)+\frac{\lambda^{x+1}}{(x+1)!}\left(\frac{n}{n}\cdots\frac{n-x}{n}\right)\right]\Bigg\} \\
= e^{-\lambda}p^n\frac{\lambda^{x+1}}{(x+1)!}&\left\{1+\sum_{j=1}^{x}\frac{\lambda^j}{j!}\left(\frac{n}{n}\cdots\frac{n-j+1}{n}\right)-\sum_{k=0}^{x}\frac{\lambda^k}{k!}\left(\frac{n}{n}\cdots\frac{n-x}{n}\right)\right\}
\end{aligned}$$

$> 0,$

which implies that $\Delta(x) - \Delta(x+1) > 0$, that is, $0 < \Delta(x) < \Delta(1) = \Delta(0) = 1 - e^{-\lambda}q^{-n}$. Hence (2.1) holds. □

Similarly, following the same arguments detailed as in the proof of Theorem 2.1, we can obtain

$$0 < \frac{\mathbb{B}_{n,p}(x)}{\mathbb{P}_{\lambda}(x)} - 1 \leq e^{\lambda}q^{n} - 1$$

for all $x \geq 0$, which gives the following corollary.

Corollary 2.1. *For* $\lambda = \frac{np}{q}$,

$$\max_{0\leq x\leq n}\left|\frac{\mathbb{B}_{n,p}(x)}{\mathbb{P}_{\lambda}(x)} - 1\right| = \max_{0\leq x\leq n}\left\{\frac{\mathbb{B}_{n,p}(x)}{\mathbb{P}_{\lambda}(x)} - 1\right\} = e^{\lambda}q^{n} - 1. \tag{2.2}$$

It is observed that if $e^{\lambda}q^{n}$ tends to 1, then the bound in (2.2) tends to 0, that is, $\mathbb{B}_{n,p}(x)$ is close to $\mathbb{P}_{\lambda}(x)$. Therefore the result in (2.2) yields a good approximation when $e^{\lambda}q^{n}$ tends to 1.

Remark 2.1. 1. Consider the results in Theorem 2.1 and Corollary 2.1. It is found that if $e^{\lambda}q^{n}$ tends to 1, then both least upper bounds in (2.1) and (2.2) tend to 0. This indicates that the binomial cumulative distribution function can be properly approximated by the Poisson cumulative distribution when $e^{\lambda}q^{n}$ is close to 1, that is, p is sufficiently small.

2. Consider the inequality (1.8), because

$$\frac{1}{e^{q\lambda}q^{n}} - 1 = \left|1 - \frac{\mathbb{P}_{q\lambda}(x)}{\mathbb{B}_{n,p}(x)}\right| \leq \max_{0\leq x\leq n}\left|1 - \frac{\mathbb{P}_{q\lambda}(x)}{\mathbb{B}_{n,p}(x)}\right| \leq \frac{(1-e^{-q\lambda})(1-q^{n})}{nq^{n}}$$

for $\lambda = \frac{np}{q}$, we have $\frac{1}{e^{q\lambda}q^{n}} - 1 < \frac{(1-e^{-q\lambda})(1-q^{n})}{nq^{n}}$. It is observed that the least upper in (2.1), $1 - e^{-\lambda}q^{-n}$, is very close to $\frac{1}{e^{q\lambda}q^{n}} - 1$ when p is sufficiently small. Thus, we have $1 - e^{-\lambda}q^{-n} < \frac{(1-e^{-q\lambda})(1-q^{n})}{nq^{n}}$.

3. Numerical Results

First, we provide examples to illustrate how well the Poisson cumulative distribution function with mean $\lambda = \frac{np}{q}$ approximates the binomial cumulative distribution function with parameters n and p using the obtained results. Because it is difficult to compare the upper bounds in (2.1) and (1.8) in theoretical terms, numerical comparison of the upper bounds are presented.

3.1. Examples

3.1.1. Let $n = 10$ and $p = 0.1$, then $\lambda = \frac{10}{9}$ and the numerical results are in the following:

$$\max_{0\le x\le 10}\left|1-\frac{\mathbb{P}_{\frac{10}{9}}(x)}{\mathbb{B}_{10,0.1}(x)}\right| = 1-e^{-\frac{10}{9}}(0.9^{-10}) = 0.055883731$$

and

$$\max_{0\le x\le 10}\left|\frac{\mathbb{B}_{10,0.1}(x)}{\mathbb{P}_{\frac{10}{9}}(x)}-1\right| = e^{\frac{10}{9}}(0.9^{10})-1 = 0.059191578.$$

3.1.2. Let $n = 50$ and $p = 0.05$, then $\lambda = \frac{50}{19}$ and the numerical results are in the following:

$$\max_{0\le x\le 50}\left|1-\frac{\mathbb{P}_{\frac{50}{19}}(x)}{\mathbb{B}_{50,0.05}(x)}\right| = 1-e^{-\frac{50}{19}}(0.95^{-50}) = 0.064724582$$

and

$$\max_{0\le x\le 50}\left|\frac{\mathbb{B}_{50,0.05}(x)}{\mathbb{P}_{\frac{50}{19}}(x)}-1\right| = e^{\frac{50}{19}}(0.95^{50})-1 = 0.069203766.$$

3.1.3. Let $n = 500$ and $p = 0.01$, then $\lambda = \frac{500}{99}$ and the numerical results are in the following:

$$\max_{0\le x\le 500}\left|1-\frac{\mathbb{P}_{\frac{500}{99}}(x)}{\mathbb{B}_{500,0.01}(x)}\right| = 1-e^{-\frac{500}{99}}(0.99^{-500}) = 0.025018833$$

and

$$\max_{0\le x\le 500}\left|\frac{\mathbb{B}_{500,0.01}(x)}{\mathbb{P}_{\frac{500}{99}}(x)}-1\right| = e^{\frac{500}{99}}(0.99^{500})-1 = 0.025660837.$$

3.1.4. Let $n = 1000$ and $p = 0.01$, then $\lambda = \frac{1000}{99}$ and the numerical results are in the following:

$$\max_{0\le x\le 1000}\left|1-\frac{\mathbb{P}_{\frac{1000}{99}}(x)}{\mathbb{B}_{1000,0.01}(x)}\right| = 1-e^{-\frac{1000}{99}}(0.99^{-1000}) = 0.049411723$$

and

$$\max_{0\le x\le 1000}\left|\frac{\mathbb{B}_{1000,0.01}(x)}{\mathbb{P}_{\frac{1000}{99}}(x)}-1\right| = e^{\frac{1000}{99}}(0.99^{1000})-1 = 0.051980152.$$

The numerical results show that the least upper bound of each relative error is small when p is sufficiently small. Moreover, in the case where λ is rather large and p is sufficiently small (Example 3.1.4), each upper bound in this case is also an appropriate criterion for measuring the accuracy of the approximation.

3.2. Numerical Comparison of Two Upper Bounds

As mentioned in Remark 2.1, it is more possibly that $1-e^{-\lambda}q^{-n} < \frac{(1-e^{-q\lambda})(1-q^n)}{nq^n}$ for $\lambda = \frac{np}{q}$. Thus, we shall show numerical comparison of these upper bounds as follows. Let $B_1 = 1 - e^{-\lambda}q^{-n}$ and $B_2 = \frac{(1-e^{-\lambda q})(1-q^n)}{nq^n}$. Some numerical values of the upper bounds B_1 and B_2 are presented in Table 3.1.

Table 3.1 provides numerical values of the upper bounds B_1 and B_2 for given $n = 10, 30, 50, 100, 300, 500$ and $p = 0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1$. It shows that $1 - e^{-\lambda}q^{-n} < \frac{(1-e^{-\lambda q})(1-q^n)}{nq^n}$, which satisfies Remark 2.1. Therefore, the upper bound in this study is sharper than those reported in (1.8).

4. Conclusion

In this study, the least upper bound for the relative error between the binomial cumulative distribution function with parameters n and p and the Poisson cumulative distribution with a different mean $\lambda = \frac{np}{q} = \frac{np}{1-p}$ is obtained by using the same method mentioned in [5]. With this upper bound, the Poisson cumulative distribution function with this mean can be used as an estimate of the binomial cumulative distribution function when p is sufficiently small even if λ is rather large. By numerical comparison between the upper bounds in (2.1) and (1.8), it can be concluded that the upper bound obtained in this study is sharper than the upper bound reported in Teerapabolarn [4].

Acknowledgments

The authors would like to thank Faculty of Science, Burapha University, for financial support to do this research in 2012.

n	p	λ	B_1	B_2
10	0.0001	$\frac{10}{9999}$	0.00000005	0.00000010
	0.0005	$\frac{10}{9995}$	0.00000125	0.00000250
	0.001	$\frac{10}{999}$	0.00000501	0.00001001
	0.005	$\frac{10}{995}$	0.00012583	0.00025070
	0.01	$\frac{10}{99}$	0.00050661	0.00100613
	0.05	$\frac{10}{95}$	0.01329369	0.02636963
	0.1	$\frac{10}{9}$	0.05588373	0.11807835
30	0.0001	$\frac{30}{9999}$	0.00000015	0.00000030
	0.0005	$\frac{30}{9995}$	0.00000375	0.00000750
	0.001	$\frac{30}{999}$	0.00001502	0.00003002
	0.005	$\frac{30}{995}$	0.00037744	0.00075344
	0.01	$\frac{30}{99}$	0.00151907	0.00304019
	0.05	$\frac{30}{95}$	0.03935326	0.09475198
	0.1	$\frac{30}{9}$	0.15845674	0.71550479
50	0.0001	$\frac{50}{9999}$	0.00000025	0.00000050
	0.0005	$\frac{50}{9995}$	0.00000625	0.00001250
	0.001	$\frac{50}{999}$	0.00002503	0.00005004
	0.005	$\frac{50}{995}$	0.00062899	0.00126009
	0.01	$\frac{50}{99}$	0.00253051	0.00513773
	0.05	$\frac{50}{95}$	0.06472458	0.22023168
	0.1	$\frac{50}{9}$	0.24988577	3.83463758
100	0.0001	$\frac{100}{9999}$	0.00000050	0.00000100
	0.0005	$\frac{100}{9995}$	0.00001251	0.00002501
	0.001	$\frac{100}{999}$	0.00005007	0.00010014
	0.005	$\frac{100}{995}$	0.00125759	0.00256066
	0.01	$\frac{100}{99}$	0.00505461	0.01094832
	0.05	$\frac{100}{95}$	0.12525989	1.66772493
	0.1	$\frac{100}{9}$	0.43732864	376.459103
300	0.0001	$\frac{300}{9999}$	0.00000150	0.00000300
	0.0005	$\frac{300}{9995}$	0.00003752	0.00007516
	0.001	$\frac{300}{999}$	0.00015019	0.00030243
	0.005	$\frac{300}{995}$	0.00376802	0.00905981
	0.01	$\frac{300}{99}$	0.01508730	0.06141906
	0.05	$\frac{300}{95}$	0.33067489	16061.9007
	0.1	$\frac{300}{9}$	0.82185877	1.779E+11
500	0.0001	$\frac{500}{9999}$	0.00000250	0.00000500
	0.0005	$\frac{500}{9995}$	0.00006254	0.00012569
	0.001	$\frac{500}{999}$	0.00025030	0.00051083
	0.005	$\frac{500}{995}$	0.00627215	0.02066985
	0.01	$\frac{500}{99}$	0.02501883	0.30035412
	0.05	$\frac{500}{95}$	0.48785233	274933304
	0.1	$\frac{500}{9}$	0.94360065	1.513E+20

Table 3.1: Numerical values of the upper bounds B_1 and B_2

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*, vol. 1, 1965, 1-12.

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

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

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

[5] K. Teerapabolarn, The least upper bound on the Poisson-negative binomial relative error, *Communications in Statistics-Theory and Methods*, **41** (2012), 1833-1838.