International Journal of Pure and Applied Mathematics
Volume 87 No. 4 2013, 535-540
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: http://dx.doi.org/10.12732/ijpam.v87i4.4
ijpam.eu

A BOUND ON THE POISSON BINOMIAL-POISSON RELATIVE ERROR

K. Teerapabolarn

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

Abstract: We use the Stein-Chen method to determine a bound on the relative error between the Poisson binomial distribution function with parameter $\mathbf{p} = (p_1, ..., p_n)$ and the Poisson distribution function with mean $\lambda = \sum_{i=1}^{n} \frac{p_i}{1-p_i}$. With this bound, the Poisson distribution function with this mean can be used as an estimate of the Poisson binomial distribution function whenever all p_i are small.

AMS Subject Classification: 62E17, 60F05, 60G50
Key Words: distribution function, Poisson distribution, Poisson binomial distribution, relative error, Stein-Chen method

1. Introduction

Let $Y_1, ..., Y_n$ be n independently distributed Bernoulli random variables, each with probability of success $p_i = P(Y_i = 1) = 1 - P(Y_i = 0)$, and let $X = \sum_{i=1}^{n} Y_i$. Then the distribution of X is called the *Poisson binomial* distribution with parameter $\mathbf{p} = (p_1, ..., p_n)$. This distribution can be thought of as the distribution of the number of successes in a sequence of n independent Bernoulli trials, where success occurs on each ith trial with a probability of p_i and failure occurs on each ith trial with a probability of $q_i = 1 - p_i$. The probability function of X [4] is of the form

Received: May 7, 2013

$$P(X=k)=\sum_{\mathbf{d}\in D^k}\prod_{i=1}^{n}\left(\frac{p_i}{q_i}\right)^{d_i}\prod_{j=1}^{n}q_j,\ \ k=0,1,...,n, \tag{1.1}$$

where $D^k=\{\mathbf{d}=(d_1,...,d_n):d_i=0,1;d_1+\cdots+d_n=k\}$, and the mean and variance of X are $E(X)=\sum_{i=1}^{n}p_i$ and $Var(X)=\sum_{i=1}^{n}p_iq_i$, respectively. It is well known that the Poisson distribution with mean $\sum_{i=1}^{n}p_i$ can be used as an estimate of the Poisson binomial distribution with parameter $\mathbf{p}$ if n is large and all p_i are small. In the past few years, many authors have tried to approximate the Poisson binomial distribution by the Poisson distribution with mean $\lambda=\sum_{i=1}^{n}p_i$. For example, Barbour et al. [1] gave a uniform bound for the difference of the Poisson binomial and Poisson distributions

$$\left|\sum_{k\in A}\sum_{\mathbf{d}\in D^k}\prod_{i=1}^{n}\left(\frac{p_i}{q_i}\right)^{d_i}\prod_{j=1}^{n}q_j-\sum_{k\in A}\frac{e^{-\lambda}\lambda^k}{k!}\right|\le\lambda^{-1}(1-e^{-\lambda})\sum_{i=1}^{n}p_i^2, \tag{1.2}$$

where $A\subseteq\mathbb{N}\cup\{0\}$. In the case of pointwise approximation, Neammanee [5] gave a non-uniform bound for the difference of the Poisson binomial and Poisson probability functions

$$\left|\sum_{\mathbf{d}\in D^{x_0}}\prod_{i=1}^{n}\left(\frac{p_i}{q_i}\right)^{d_i}\prod_{j=1}^{n}q_j-\frac{e^{-\lambda}\lambda^{x_0}}{x_0!}\right|\le\min\left\{\frac{1}{x_0},\lambda^{-1}\right\}\sum_{i=1}^{n}p_i^2, \tag{1.3}$$

where $x_0\in\{1,...,n-1\}$. For cumulative probability approximation, the Poisson binomial distribution function

$$\mathbb{PB}_{\mathbf{p}}(k)=\sum_{j=0}^{k}\sum_{\mathbf{d}\in D^j}\prod_{i=1}^{n}\left(\frac{p_i}{q_i}\right)^{d_i}\prod_{j=1}^{n}q_j,\ \ k=0,1,...,n \tag{1.4}$$

is also approximated by the Poisson distribution function

$$\mathbb{P}_\lambda(k)=\sum_{j=0}^{k}\frac{e^{-\lambda}\lambda^j}{j!},\ \ k=0,1,..., \tag{1.5}$$

for $\lambda=\sum_{i=1}^{n}p_i$. In this case, Teerapabolarn and Neammanee [7] gave a non-uniform bound in the form of

$$|\mathbb{PB}_{\mathbf{p}}(x_0)-\mathbb{P}_\lambda(x_0)|\le\lambda^{-1}(1-e^{-\lambda})\min\left\{1,\frac{e^\lambda}{x_0+1}\right\}\sum_{i=1}^{n}p_i^2, \tag{1.6}$$

where $x_0 \in \{0,1,...,n\}$. Note that, the results in (1.2), (1.3) and (1.6) were created by the Stein-Chen method.

Following the *Law of Small Numbers*, it is observed that the Poisson distribution with mean $\lambda = \sum_{i=1}^{n} \frac{p_i}{q_i}$ can also be used as an estimate of the Poisson binomial distribution with parameter $\mathbf{p}$ when n is large and all p_i are small. Similarly, the Poisson binomial cumulative distribution function can also be properly approximated by the Poisson cumulative distribution function with mean $\lambda = \sum_{i=1}^{n} \frac{p_i}{q_i}$. Let $\mathbb{PB}_{\mathbf{p}}(x_0)$ and $\mathbb{P}_\lambda(x_0)$ be the Poisson binomial and Poisson distribution functions at $x_0 \in \{0,1,...,n\}$. In this paper, we are interested to approximate $\mathbb{PB}_{\mathbf{p}}(x_0)$ by $\mathbb{P}_\lambda(x_0)$ with mean $\lambda = \sum_{i=1}^{n} \frac{p_i}{q_i}$ in terms of the relative error $1 - \frac{\mathbb{P}_\lambda(x_\mathbf{0})}{\mathbb{PB}_{\mathbf{p}}(x_\mathbf{0})}$ and its bound.

The Stein-Chen method is the tool for deriving a bound for the such relative error, which is mentioned in Section 2. In Section 3, the Stein-Chen method is applied to obtain the desired result. Conclusion is presented in the last section.

2. Method

Stein's method was originally formulated for normal approximation [3]. It was adapted and applied to the Poisson case by Chen [2], which is refer to as the Stein-Chen method. Following Teerapabolarn [6], Stein's equation of the Poisson distribution function with parameter $\lambda > 0$ is of the form

$$h_{x_\mathbf{0}}(x) - \mathbb{P}_\lambda(x_0) = \lambda f_{x_\mathbf{0}}(x+1) - x f_{x_\mathbf{0}}(x), \tag{2.1}$$

where $x_0, x \in \mathbb{N} \cup \{0\}$, and for $h_{x_\mathbf{0}}(x) = 1$ if $x \le x_0$ and $h_{x_\mathbf{0}}(x) = 0$ if $x > x_0$. For $x_0 \in \mathbb{N}$, he showed that

$$0 < \sup_{x \ge 2} f_{x_\mathbf{0}}(x) \le \frac{\lambda^{-2}(e^\lambda - \lambda - 1)\mathbb{P}_\lambda(x_0)}{x_0 + 1}. \tag{2.2}$$

3. Result

The following theorem shows a result of the Poisson approximation to the Poisson binomial distribution function in the relative error form and its bound.

Theorem 3.1. *For* $x_0 \in \{1, ..., n\}$ *and* $\lambda = \sum_{i=1}^{n} \frac{p_i}{q_i}$, *then the following inequality holds:*

$$0 \leq 1 - \frac{\mathbb{P}_\lambda(x_0)}{\mathbb{PB}_\mathbf{p}(x_0)} \leq \frac{\lambda^{-2}(e^\lambda - \lambda - 1)}{x_0 + 1} \sum_{i=1}^{n} \frac{p_i^2}{q_i}, \tag{3.1}$$

where $1 - \frac{\mathbb{P}_\lambda(x_\mathbf{0})}{\mathbb{PB}_p(x_\mathbf{0})} = 1 - \frac{e^{-\lambda}}{\Pi_{i=\mathbf{1}}^n q_i}$ *when* $x_0 = 0$.

Proof. Substituting x by X and taking expectation in (2.1), yields

$$\begin{aligned}\mathbb{PB}_\mathbf{p}(x_0) - \mathbb{P}_\lambda(x_0) &= E[\lambda f(X+1) - Xf(X)] \\ &= E\left[\sum_{i=1}^{n} \frac{p_i}{q_i} f(X+1) - \sum_{i=1}^{n} Y_i f(X)\right] \\ &= \sum_{i=1}^{n} E\left[\frac{p_i}{q_i} f(X+1) - Y_i f(X)\right],\end{aligned} \tag{3.2}$$

where $f = f_{x_\mathbf{0}}$ is defined as in (2.1). Let $\delta(f) = E\left[\frac{p_i}{q_i} f(X+1) - Y_i f(X)\right]$ and let $X_i = X - Y_i$. Then, for each i,

$$\begin{aligned}\delta(f) &= E\left[\frac{p_i}{q_i} f(X_i + Y_i + 1) - Y_i f(X_i + Y_i)\right] \\ &= E\left\{E\left[\left(\frac{p_i}{q_i} f(X_i + Y_i + 1) - Y_i f(X_i + Y_i)\right) |Y_i\right]\right\} \\ &= E\left[\left(\frac{p_i}{q_i} f(X_i + Y_i + 1) - Y_i f(X_i + Y_i)\right) |Y_i = 0\right] q_i \\ &\quad + E\left[\left(\frac{p_i}{q_i} f(X_i + Y_i + 1) - Y_i f(X_i + Y_i)\right) |Y_i = 1\right] p_i \\ &= E[p_i f(X_i + 1)] + E\left[\frac{p_i^2}{q_i} f(X_i + 2) - p_i f(X_i + 1)\right] \\ &= \frac{p_i^2}{q_i} E[f(X_i + 2)].\end{aligned} \tag{3.3}$$

Combining (3.2) and (3.3) and using (2.2), it follows that

$$\begin{aligned}0 \leq \mathbb{PB}_\mathbf{p}(x_0) - \mathbb{P}_\lambda(x_0) &= \sum_{i=1}^{n} \frac{p_i^2}{q_i} E[f(X_i + 2)] \\ &\leq \sum_{i=1}^{n} \frac{p_i^2}{q_i} \sup_{x \geq 2} f(x)\end{aligned}$$

$$\leq \frac{\lambda^{-2}(e^{\lambda}-\lambda-1)\mathbb{P}_{\lambda}(x_0)}{x_0+1}\sum_{i=1}^{n}\frac{p_i^2}{q_i}. \qquad (3.4)$$

Dividing (3.4) by $\mathbb{PB}_{\mathbf{p}}(x_0)$ and since $\frac{\mathbb{P}_{\lambda}(x_{\mathbf{0}})}{\mathbb{PB}_{\mathbf{p}}(x_{\mathbf{0}})} \leq 1$, it follows that the inequality (3.1) holds. □

Remark. Consider the result in Theorem 3.1, if all p_i are small, then the bound (3.1) approaches 0, that is, the Poisson binomial distribution function can be properly approximated by the Poisson distribution with this mean.

4. Conclusion

A bound in Theorem 3.1 is an estimate of the relative error between the Poisson binomial distribution function with parameter $\mathbf{p} = (p_1, ..., p_n)$ and the Poisson distribution with mean $\lambda = \sum_{i=1}^{n}\frac{p_i}{q_i} = \sum_{i=1}^{n}\frac{p_i}{1-p_i}$. It is indicated that the Poisson distribution function with this mean can be used as an estimate of the Poisson binomial distribution function whenever all p_i are small.

References

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

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

[3] C.M. Stein, A bound for the error in normal approximation to the distribution of a sum of dependent random variables, *Proc. Sixth Berkeley Sympos. Math. Statist. Probab.*, **3** (1972), 583-602.

[4] S.X. Chen, J.S. Liu, Statistical applications of the Poisson-binomial and condition bernoulli distributions. *Statist. Sinica*, **7** (1997), 875-892.

[5] K. Neammanee, Pointwise Approximation of Poisson binomial by Poisson distribution. *Stochastic Modelling and Applications*, **6** (2003), 20-26.

[6] K. Teerapabolarn, A new bound on Poisson approximation for independent geometric variables, *International Journal of Pure and Applied Mathematics*, **84** (2013), 419-422.

[7] K. Teerapabolarn, K. Neammanee, Poisson approximation for sums of dependent Bernoulli random variables. *ActaMathematica Academiae Paedagogicae Nyiregyhaziensis*, **22** (2006), 87-99.