Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 11, Number 4 (2015), pp. 1975-1978

http://www.ripublication.com

Pointwise Binomial Approximation for Independent Pólya Random Variables

K. Teerapabolarn

Department of Mathematics,
Faculty of Science,
Burapha University,
Chonburi 20131, Thailand.
E-mail: kanint@buu.ac.th

Abstract

Stein's method and the Pólya w-functions are used to determine a non-uniform bound for the point metric between the probability function of a sum of n independent Pólya random variables, each with parameters N_i, n_i, r_i and c_i, and the binomial probability function with parameters $\sum_{i=1}^{n} n_i$ and $\frac{1}{m}\sum_{i=1}^{n} \frac{n_i r_i}{N_i}$. The result of this study gives a good approximation when all N_i are large with respect to all n_i, r_i and c_i.

AMS subject classification: 60F05, 60G50.
Keywords: Binomial approximation, Pólya distribution, Pólya w-function, Stein's method.

1. Introduction

Let $X_1, \ldots, X_n$ be independently distributed Pólya random variables, each with the probability function

$$p_{X_i}(x) = \frac{\binom{\frac{r_i}{c_i}+x-1}{x}\binom{\frac{N_i-r_i}{c_i}+n_i-x-1}{n_i-x}}{\binom{\frac{N_i}{c_i}+n_i-1}{n_i}}$$

for $x = 0, \ldots, n_i$, and mean $\mu_i = \frac{n_i r_i}{N_i}$ and variance $\sigma_i^2 = \frac{r_i n_i (N_i + c_i n_i)(N_i - r_i)}{N_i^2 (N_i + c_i)}$.

Let $\mathcal{S}_n = \sum_{i=1}^{n} X_i$ and $\mathcal{B}_{m,p}$ denote the binomial random variable with parameters $m = \sum_{i=1}^{n} n_i$ and $p = \frac{1}{m} \sum_{i=1}^{n} \mu_i$. For approximating the distribution of $\mathcal{S}_n$ by a binomial distribution with parameters $m = \sum_{i=1}^{n} n_i$ and $p = \frac{1}{m} \sum_{i=1}^{n} \mu_i$, Teerapabolarn [3] gave a bound for the distance between the distributions of $\mathcal{S}_n$ and $\mathcal{B}_{m,p}$ as follows:

$$d_A(\mathcal{S}_n, \mathcal{B}_{m,p}) \leq \Delta(m, p) \sum_{i=1}^{n} \left\{ |pN_i - r_i| + \frac{c_i (n_i - 1)(N_i - r_i)}{N_i + c_i} \right\} \frac{n_i r_i}{N_i^2}, \tag{1.1}$$

where $d_A(\mathcal{S}_n, \mathcal{B}_{m,p}) = |P(\mathcal{S}_n \in A) - P(\mathcal{B}_{m,p} \in A)|$ for $A \subseteq \{0, \ldots, n\}$ and $\Delta(m, p) = \frac{1 - p^{m+1} + q^{m+1}}{(m+1)pq}$. However, for $A = \{x_0; x_0 \in \{0, \ldots, m\}\}$, (1.1) becomes

$$d_{x_0}(\mathcal{S}_n, \mathcal{B}_{m,p}) \leq \Delta(m, p) \sum_{i=1}^{n} \left\{ |pN_i - r_i| + \frac{c_i (n_i - 1)(N_i - r_i)}{N_i + c_i} \right\} \frac{n_i r_i}{N_i^2}, \tag{1.2}$$

where $d_{x_0}(\mathcal{S}_n, \mathcal{B}_{m,p}) = |P(\mathcal{S}_n = x_0) - P(\mathcal{B}_{m,p} = x_0)|$. It is observed that the bound is a uniform constant for the point metric $d_{x_0}(\mathcal{S}_n, \mathcal{B}_{m,p})$. With this situation, a non-uniform bound with respect to x_0 is required. In this paper, we focus on deriving a non-uniform bound for the point metric between the distribution of $\mathcal{S}_n$ and the distribution of $\mathcal{B}_{m,p}$, where $x_0 \in \{0, \ldots, m\}$.

Stein's method and the Pólya w-functions are important tools, we describe in Section 2. In Section 3, we give a non-uniform bound for $d_{x_0}(\mathcal{S}_n, \mathcal{B}_{m,p})$, and the conclusion of this study is presented in the last section.

2. Method

The following lemma gives the Pólya w-functions [3].

Lemma 2.1. For $1 \leq i \leq n$, let w_i be the w-function associated with the Pólya random variable X_i, then we have the following:

$$w_i(x) = \frac{(n_i - x)(r_i + c_i x)}{N_1 \sigma_i^2}, \quad x = 0, \ldots, n_i. \tag{2.1}$$

The following relation is an important property for proving the result, which was stated by [2].

$$Cov(\mathcal{S}_n, f(\mathcal{S}_n)) = \sum_{i=1}^{n} Cov\left(X_i, f\left(X_i + \sum_{j\neq i} X_j\right)\right)$$
$$= \sum_{i=1}^{n} \sigma_i^2 E[w_i(X_i)\Delta f(\mathcal{S}_n)], \tag{2.2}$$

for any function $f : \mathbb{N} \cup \{0\} \to \mathbb{R}$ for which $E|w_i(X_i)\Delta f(\mathcal{S}_n)| < \infty$, where $\Delta f(x) = f(x+1) - f(x)$.

For Stein's method in the binomial approximation, it can be applied for every $m \in \mathbb{N}$ and $0 < p = 1 - q < 1$, for every $x_0 \in \{0, \dots, m\}$ and bounded real-valued function $f = f_{\{x_0\}} : \mathbb{N} \cup \{0\} \to \mathbb{R}$ defined as in [1], where $f(0) = f(1)$ and $f(x) = f(m)$ for $x \geq m$. So, Stein's equation for these conditions is as follows:

$$P(\mathcal{S}_n - x_0) - P(\mathcal{B}_{m,p} = x_0) = E[(m - \mathcal{S}_n)pf(\mathcal{S}_n + 1) - q\mathcal{S}_n f(\mathcal{S}_n)]. \tag{2.3}$$

For $x_0, x \in \mathbb{N} \cup \{0\}$, [4] showed that

$$\sup_{x\geq 0} |\Delta f(x)| \leq \delta(x_0) = \begin{cases} \dfrac{1-q^m}{np} & \text{if } x_0 = 0, \\ \min\left\{\dfrac{1-p^m}{x_0 q}, \dfrac{1-p^{m+1}q^{m+1}}{(m+1)pq}\right\} & \text{if } x_0 > 0. \end{cases} \tag{2.4}$$

3. Result

The following theorem presents the main result of this study.

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

$$d_{x_0}(\mathcal{S}_n, \mathcal{B}_{m,p}) \leq \delta(x_0) \sum_{i=1}^{n} \left\{|pN_i - r_i| + \frac{c_i(n_i - 1)(N_i - r_i)}{N_i + c_i}\right\} \frac{n_i r_i}{N_i^2}. \tag{3.1}$$

Proof. From (2.3), it follows that

$$\begin{aligned} d_{x_0}(\mathcal{S}_n, \mathcal{B}_{m,p}) &= |E[(m - \mathcal{S}_n)pf(\mathcal{S}_n + 1) - q\mathcal{S}_n f(\mathcal{S}_n)]| \\ &= |E[mpf(\mathcal{S}_n + 1) - p\mathcal{S}_n \Delta f(\mathcal{S}_n) - \mathcal{S}_n f(\mathcal{S}_n)]| \\ &= |E[mp\Delta f(\mathcal{S}_n)] - pE[\mathcal{S}_n \Delta f(\mathcal{S}_n)] - Cov(\mathcal{S}_n, f(\mathcal{S}_n))| \\ &= \left|\sum_{i=1}^{n} \{E[\mu_i \Delta f(\mathcal{S}_n)] - pE[X_i \Delta f(\mathcal{S}_n)] - Cov(X_i, f(\mathcal{S}_n))\}\right|. \end{aligned}$$

Using (2.2) and Lemma 1, we have

$$
\begin{aligned}
d_{x_0}(\mathcal{S}_n, \mathcal{B}_{m,p}) &= \left| \sum_{i=1}^{n} \left\{ E\left[(\mu_i - pX_i)\Delta f(\mathcal{S}_n)\right] - \sigma_i^2 E[w_i(X_i)\Delta f(\mathcal{S}_n)] \right\} \right| \\
&\le \sum_{i=1}^{n} E\{|\mu_i - pX_i - \sigma_i^2 w_i(X_i)||\Delta f(\mathcal{S}_n)|\} \\
&\le \sup_x |\Delta f(x)| \sum_{i=1}^{n} E\left\{ \left| \frac{n_i r_i}{N_i} - pX_i - \frac{(n_i - X_i)(r_i + c_i X_i)}{N_i} \right| \right\} \\
&\le \sup_x |\Delta f(x)| \sum_{i=1}^{n} \left\{ |pN_i - r_i|\, E(X_i) + c_i E(n_i X_i - X_i^2) \right\} \frac{1}{N_i} \\
&= \sup_x |\Delta f(x)| \sum_{i=1}^{n} \left\{ |pN_i - r_i| + \frac{c_i(n_i - 1)(N_i - r_i)}{N_i + c_i} \right\} \frac{n_i r_i}{N_i^2}.
\end{aligned}
$$

Hence, by (2.4), (3.1) is obtained. ■

4. Conclusion

In this study, a non-uniform bound on the error of binomial approximation to the probability function of a sum of n independent Pólya random variables was obtained, by using Stein's method and the Pólya w-functions. The probability function of this summands can be approximated by the binomial probability function with parameters $m = \sum_{i=1}^{n} n_i$ and $p = \frac{1}{m} \sum_{i=1}^{n} \frac{n_i r_i}{N_i}$ when all N_i are large with respect to all n_i, r_i and c_i. In addition, the result in Theorem 3.1 is better than the result in (1.1).

References

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

[2] Cacoullos, T., Papathanasiou, V., 1989, "Characterization of distributions by variance bounds", Statist. Probab. Lett., 7(5), 351–356.

[3] Teerapabolarn, K., 2015, "Binomial approximation for a sum of independent Pólya random variables", Appl. Math. Sci., 9(33), 1637—1640.

[4] Teerapabolarn, K., Wongkasem, P., 2011, "On pointwise binomial approximation by w-functions", Int. J. Pure Appl. Math., 71(1), 57–66.