

A NON-UNIFORM BOUND ON THE POINT METRIC OF POISSON BINOMIAL AND BINOMIAL DISTRIBUTIONS

K. Teerapabolarn

Department of Mathematics

Faculty of Science

Burapha University

Chonburi, 20131, THAILAND

Abstract: We use Stein's method to determine a non-uniform bound on the point metric of the Poisson binomial distribution with parameter $\mathbf{p} = (p_1, \dots, p_n)$ and the binomial distribution with parameters n' and p' , where n' is the integral part of $\left\{ \frac{(\sum_{i=1}^n p_i)^2}{\sum_{i=1}^n p_i^2} + \frac{1}{2} \right\}$ and $p' = \frac{1}{n'} \sum_{i=1}^n p_i$. The bound in this study is more appropriate for measuring the accuracy of the approximation.

AMS Subject Classification: 62E17, 60F05, 60G50

Key Words: binomial distribution, non-uniform bound, Poisson binomial distribution, point metric, Stein's method

1. Introduction and Theorem

Let Y_1, \dots, 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 usually called the *Poisson binomial* distribution with parameter $\mathbf{p} = (p_1, \dots, p_n)$. The probability function of X is of the form

$$\text{pb}_{\mathbf{p}}(k) = \sum_{\mathbf{d} \in D^k} \prod_{i=1}^n \left(\frac{p_i}{q_i} \right)^{d_i} \prod_{j=1}^n q_j, \quad k = 0, 1, \dots, n, \quad (1.1)$$

where $q_i = 1 - p_i$, $D^k = \{\mathbf{d} = (d_1, \dots, d_n) : d_i = 0, 1; d_1 + \dots + d_n = k\}$, and

the mean and variance of X are $E(X) = \sum_{i=1}^n p_i$ and $Var(X) = \sum_{i=1}^n p_i q_i$, 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 $\lambda = \sum_{i=1}^n p_i$. Furthermore, there have been some authors have also tried to approximate the Poisson binomial distribution by the binomial distribution. For example, Ehm [2] gave a uniform bound for the difference of the Poisson binomial and binomial distributions

$$|\mathcal{PB}_{\mathbf{p}}(A) - \mathcal{B}_{n,p}(A)| \leq \frac{1 - p^{n+1} - q^{n+1}}{(n + 1)pq} \sum_{i=1}^n (p_i - p)^2, \tag{1.2}$$

for $A \subseteq \{0, 1, \dots, n\}$, $\mathcal{PB}_{\mathbf{p}}(A) = \sum_{k \in A} \text{pb}_{\mathbf{p}}(k)$, $\mathcal{B}_{n,p}(A) = \sum_{k \in A} \binom{n}{k} p^k q^{n-k}$ and $p = 1 - q = \frac{1}{n} \sum_{i=1}^n p_i$. Later, Soon [3] improved the above bound to be more appropriate in the form of

$$|\mathcal{PB}_{\mathbf{p}}(A) - \mathcal{B}_{n',p'}(A)| \leq C_{n',p'} \left\{ 2 \left(\lambda_3 - \frac{\lambda_2^2}{\lambda} \right) + \lambda |p' - p^*| \right\}, \tag{1.3}$$

where $A \subseteq \{0, 1, \dots, n'\}$, $C_{n',p'} = \frac{1 - (p')^{n'+1} - (1-p')^{n'+1}}{(n'+1)p'(1-p')}$, n' is the integral part of $\left\{ \frac{(\sum_{i=1}^n p_i)^2}{\sum_{i=1}^n p_i^2} + \frac{1}{2} \right\}$, $p' = \frac{1}{n'} \sum_{i=1}^n p_i$, $\lambda_i = \sum_{j=1}^n p_j^i$, $\lambda = \lambda_1$ and $p^* = \frac{\lambda_2}{\lambda}$. For $A = \{x_0\}$ as $x_0 \in \{0, 1, \dots, n'\}$, let $\text{b}_{n',p'}(x_0) = \binom{n}{x_0} p^{x_0} q^{n-x_0}$ then (1.3) becomes the point metric

$$|\text{pb}_{\mathbf{p}}(x_0) - \text{b}_{n',p'}(x_0)| \leq C_{n',p'} \left\{ 2 \left(\lambda_3 - \frac{\lambda_2^2}{\lambda} \right) + \lambda |p' - p^*| \right\}. \tag{1.4}$$

It is observed that the bound in (1.4) is a uniform bound for $x_0 \in \{0, 1, \dots, n'\}$, which is inappropriate for measuring the accuracy of the approximation. Therefore, a non-uniform bound, depends on x_0 , is more appropriate for this case. In this study, we focus on determining a non-uniform bound for the point metric $|\text{pb}_{\mathbf{p}}(x_0) - \text{b}_{n',p'}(x_0)|$ by using the Stein's method.

2. Method

We will prove our main result by Stein's method, which was first developed and applied in the setting of the binomial distribution by Stein [4]. The Stein's

equation for the binomial distribution with parameters $m \geq 1$ and $\wp \in (0, 1)$ is, for given h , defined by

$$(m - x)\wp g(x + 1) - (1 - \wp)xg(x) = h(x) - \mathcal{B}_{m,\wp}(h), \tag{2.1}$$

where $\mathcal{B}_{m,\wp}(h) = \sum_{k=0}^n h(k) \binom{n}{k} p^k q^{n-k}$ and g and h are bounded real-valued functions defined on $\{0, 1, \dots, m\}$.

For $A \subseteq \{0, 1, \dots, m\}$, let $h_A : \{0, 1, \dots, m\} \rightarrow \mathbb{R}$ be defined by

$$h_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \tag{2.2}$$

Following [1], let $g_A : \mathbb{N} \cup \{0\} \rightarrow \mathbb{R}$ satisfy (2.1), where $g_A(0) = g_A(1)$ and $g_A(x) = g_A(m)$ for $x \geq m$. For $A = \{x_0\}$, $x_0 \in \{0, 1, \dots, m\}$, the solution $g_{x_0} = g_{\{x_0\}}$ of (2.1) can be written as

$$g_{x_0}(x) = \begin{cases} -\frac{\binom{m}{x_0} \wp^{x_0-x} \mathcal{B}_{m,\wp}(h_{C_{x-1}})}{x \binom{m}{x} (1-\wp)^{x_0-(x-1)}} & \text{if } x_0 \geq x \geq 1, \\ \frac{\binom{m}{x_0} \wp^{x_0-x} \mathcal{B}_{m,\wp}(1-h_{C_{x-1}})}{x \binom{m}{x} (1-\wp)^{x_0-(x-1)}} & \text{if } x_0 < x, \end{cases} \tag{2.3}$$

where $C_x = \{0, \dots, x\}$.

Before proving our main result, we give the following lemma.

Lemma 2.1. *For $x_0 \in \{1, \dots, m\}$, let $\Delta g_{x_0}(x) = g_{x_0}(x + 1) - g_{x_0}(x)$. Then the following inequalities hold:*

$$\sup_{x \geq 1} |\Delta g_{x_0}(x)| \leq \min \left\{ \frac{1 - \wp^m}{(1 - \wp)x_0}, \frac{1 - \wp^{m+1} - (1 - \wp)^{m+1}}{(m + 1)\wp(1 - \wp)} \right\} \quad ([5]) \tag{2.4}$$

3. Result

The following theorem presents the point metric of Poisson binomial and binomial distributions, which can be derived by Stein’s method.

Theorem 3.1. *Let $m = n'$ and $\wp = p'$ and for $x_0 \in \{1, \dots, n'\}$, then we have the following:*

$$|\text{pb}_{\mathbf{p}}(x_0) - \text{b}_{n',p'}(x_0)| \leq \min \left\{ \frac{1 - (p')^n}{(1 - p')x_0}, C_{n',p'} \right\} \lambda |p' - p^*|, \tag{3.1}$$

where $pb_{\mathbf{p}}(0) = \prod_{i=1}^n (1 - p_i)$.

Proof. Substituting x by X and putting $h = h_{x_0}$ and then taking expectation in (2.1), we have

$$\begin{aligned} pb_{\mathbf{p}}(x_0) - b_{n',p'}(x_0) &= E[(n' - X)p'g(X + 1) - (1 - p')Xg(X)] \\ &= E[\lambda g(X + 1) - Xg(X)] - p'E[X\Delta g(X)], \end{aligned} \tag{3.2}$$

where $g = g_{x_0}$ is defined in (2.3).

Let $X_i = \sum_{j \neq i} Y_j$, then we have $p'E[X\Delta g(X)] = \sum_{i=1}^n p'p_i E[\Delta g(X_i + 1)|Y_i = 1] = \sum_{i=1}^n p'p_i E[\Delta g(X_i + 1)]$ and

$$\begin{aligned} E[\lambda g(X + 1) - Xg(X)] &= \sum_{i=1}^n p_i E[g(X + 1)] + \sum_{i=1}^n E[Y_i g(X)] \\ &= \sum_{i=1}^n p_i^2 E[\Delta g(X_i + 1)]. \end{aligned}$$

Thus, the right hand side of (3.2) can be expressed as

$$\sum_{i=1}^n p_i(p^* - p') E[\Delta g(X_i + 1)]. \tag{3.3}$$

Therefore, from (3.2) and (3.3), we obtain

$$\begin{aligned} |pb_{\mathbf{p}}(x_0) - b_{n',p'}(x_0)| &\leq \sum_{i=1}^n p_i |p^* - p'| E|\Delta g(X_i + 1)| \\ &\leq \sup_{x \geq 1} |\Delta g(x)| \lambda |p' - p^*| \\ &\leq \min \left\{ \frac{1 - (p')^n}{(1 - p')x_0}, C_{n',p'} \right\} \lambda |p' - p^*| \text{ (by (2.4))}, \end{aligned}$$

which gives the result in (3.1). □

Remark. It is clear that the bound in (3.1) is sharper than the bound in (1.4), because $\min \left\{ \frac{1 - (p')^n}{(1 - p')x_0}, C_{n',p'} \right\} \leq C_{n',p'}$ and $\lambda |p' - p^*| \leq 2 \left(\lambda_3 - \frac{\lambda_2^2}{\lambda} \right) + \lambda |p' - p^*|$.

4. Conclusion

A non-uniform bound in Theorem 3.1 is an estimate of the point metric of the Poisson binomial distribution with parameter $\mathbf{p} = (p_1, \dots, p_n)$ and the binomial distribution with parameters n' and p' . By comparing the bounds in (1.4) and (3.1), it is clear that the bound in (3.1) is sharper than the bound in (1.4). Therefore, the bound in this study is an appropriate criteria for measuring the accuracy of this approximation.

References

- [1] A.D. Barbour, L. Holst and S. Janson, *Poisson Approximation*, Oxford Studies in probability 2, Clarendon Press, Oxford, 1992.
- [2] W. Ehm, Binomial approximation to the Poisson binomial distribution, *Statist. Probab. Lett.*, **11** (1991), 7-16.
- [3] Y.T. Soon Spario, Binomial approximation for dependent indicators, *Statist. Sinica*, **6** (1996), 703-714.
- [4] C.M. Stein, *Approximate Computation of Expectations*, IMS, Hayward California, 1986.
- [5] K. Teerapabolarn, P. Wongkasem, On pointwise binomial approximation by w -functions, *Int. J. Pure Appl. Math.*, **71** (2011), 57-66.

