POINTWISE BINOMIAL APPROXIMATION TO
THE GENERALIZED HYPERGEOMETRIC DISTRIBUTION

K. Teerapabolarn
Department of Mathematics
Faculty of Science
Burapha University
Chonburi, 20131, THAILAND

Abstract: In this paper, we use the result in [2] and the $w$-function associated
with the generalized hypergeometric random variable to give a pointwise bound
for the point metric between the generalized hypergeometric distribution with
parameters $\alpha$, $\beta$ and $N$ and the binomial distribution with parameters $n = N - 1$
and $p = 1 - q = \frac{\beta + 1}{\alpha + \beta + 2}$. With this bound, it is observed that the desired result
gives a good binomial approximation when $\alpha$ is sufficiently large.

AMS Subject Classification: 62E17, 60F05
Key Words: binomial approximation, generalized hypergeometric distribution, point metric

1. Introduction

Let a non-negative integer-valued random variable $X$ have the generalized hy-
pergeometric distribution with parameters $\alpha$, $\beta$ and $N$. Its probability mass
function is given by

$$GH_{\alpha, \beta, N}(x) = \frac{(N-1)! \Gamma(N+\alpha-x)\Gamma(\beta+1+x)\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)\Gamma(\alpha+\beta+N+1)}, \quad x = 0, ..., N - 1,$$

where $N \in \mathbb{N} \setminus \{1\}$, $\alpha \geq 0$ and $\beta > -1$ and the mean and variance of $X$ are
\( \mu = \frac{(N-1)(\beta+1)}{\alpha+\beta+2} \) and \( \sigma^2 = \frac{(N-1)(\beta+1)(\alpha+\beta+N+1)}{(\alpha+\beta+2)^2(\alpha+\beta+3)} \), respectively [1]. We know that this distribution can be approximated by some appropriate discrete distributions if some conditions of their parameters are satisfied. In this case, Crosu [1] used Stein’s method and the \( w \)-function associated with the generalized hypergeometric random variable to obtain a bound for the total variation distance between the generalized hypergeometric distribution and a Poisson distribution with mean \( \mu = \frac{(N-1)(\beta+1)}{\alpha+\beta+2} \), where \( \beta + 2 \geq N \). After that, Teerapabolarn [3] used the same tools to give a bound for the total variation distance between the generalized hypergeometric distribution and a binomial distribution with parameters \( n = N - 1 \) and \( p = \frac{\beta+1}{\alpha+\beta+2} \). In this paper, we extend the result in [3] to approximate the generalized hypergeometric probability function by determining an appropriate pointwise bound for the point metric \( |GH_{\alpha,\beta,N}(x) - B_{n,p}(x)| \) when \( x \in \{0, ..., n\} \), where \( B_{n,p}(x) \) is the binomial probability function with parameters \( n \) and \( p \), which is in Section 2. The conclusion of this study is presented in the last section.

2. Result

The following lemma presents the \( w \)-function associated with the generalized hypergeometric random variable, which obtained from [1].

**Lemma 2.1.** Let \( w(X) \) be the \( w \)-function associated with the generalized hypergeometric random variable \( X \). Then, we have the following:

\[
w(x) = \frac{(\beta + x + 1)(N - x - 1)}{(\alpha + \beta + 2)\sigma^2}, \quad x = 0, ..., N - 1.
\]

(2.1)

The desired result of this study is a pointwise bound for the point metric between \( GH_{\alpha,\beta,N}(x) \) and \( B_{n,p}(x) \), which presents in the following theorem.

**Theorem 2.1.** Let \( n = N - 1 \) and \( p = 1 - q = \frac{\beta+1}{\alpha+\beta+2} \), for \( x \in \{0, ..., N-1\} \), then we have

\[
|GH_{\alpha,\beta,N}(x) - B_{n,p}(x)| \leq \begin{cases} (1-q^{N-1})(N-2)(\alpha+1) \\ (\alpha+\beta+2)(\alpha+\beta+3) \end{cases} \min \{ \frac{1-p^{N-1}}{x}, \frac{1-p^n-Nq^N}{Np} \} \quad \text{if } x = 0,
\]

\[
(N-1)(N-2)(\beta+1) \quad \text{if } x > 0.
\]

(2.2)

**Proof.** Because \( (n-x)p - \sigma^2 w(x) = \frac{(\beta+1)(N-x-1)}{\alpha+\beta+2} - \frac{(\beta+x+1)(N-x-1)}{\alpha+\beta+2} \leq 0 \) for
every \(0 \leq x \leq N - 1\) and by following Corollary 3.1 in [2], we have that

\[
|GH_{\alpha,\beta,N}(x) - B_{n,p}(x)| \leq \begin{cases} \frac{1}{np} \left| \mu q - \sigma^2 \right| & \text{if } x = 0, \\ \min \left\{ \frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq} \right\} \left| \mu q - \sigma^2 \right| & \text{if } x > 0. \end{cases}
\]

Hence, by substituting these parameters, the inequality (2.2) is obtained. □

**Corollary 2.1.** If \(\beta = 0\), then, for \(x \in \{0, ..., N - 1\}\), we have

\[
|GH_{\alpha,N}(x) - B_{n,p}(x)| \leq \begin{cases} \frac{(1-q^{N-1})(N-2)(\alpha+1)}{(\alpha+2)(\alpha+3)} & \text{if } x = 0, \\ \min \left\{ \frac{1-p^{N-1}}{x}, \frac{1-p^{N}-q^{N}}{Np} \right\} \frac{(N-1)(N-2)}{(\alpha+2)(\alpha+3)} & \text{if } x > 0. \end{cases}
\]  

(2.3)

**Remark.** It can be seen that the result gives a good binomial approximation when \(p\) is small, or \(\beta\) is small and \(\alpha\) is large, rather than \(p\) is large. Especially, in the case of \(\beta = 0\), the bound is a good measurement of the accuracy of this approximation when \(\frac{N}{\alpha}\) is small, or \(\alpha\) is large.

### 3. Conclusion

In this study, a pointwise bound for the point metric between the generalized hypergeometric with parameters \(\alpha, \beta\) and \(N\) and an appropriate binomial distribution with parameters \(n = N - 1\) and \(p = \frac{\beta+1}{\alpha+\beta+2}\) was obtained. With this bound, it is observed that the binomial probability function can be used as an estimate of the generalized hypergeometric probability function when \(\frac{N}{\alpha}\) is small, or \(\alpha\) is sufficiently large.

### References


