

A NON-UNIFORM BOUND ON GEOMETRIC APPROXIMATION TO THE YULE DISTRIBUTION

K. Teerapabolarn

Department of Mathematics

Faculty of Science

Burapha University

Chonburi, 20131, THAILAND

Abstract: The aim of this paper, we use Stein's method and the w -function associated with the Yule random variable to obtain a non-uniform bound for the point metric of the Yule and geometric distributions. Three numerical examples are provided to illustrate the obtained result.

AMS Subject Classification: 62E17, 60F05

Key Words: geometric approximation, point metric, Stein's method, w -function associated with the Yule random variable, Yule distribution

1. Introduction

A non-negative integer-valued random variable X is said to have the Yule distribution with parameter $c > 0$ if its probability function, $\mathbb{Y}_c(x)$, is given by

$$\mathbb{Y}_c(x) = \frac{cx! \Gamma(c+1)}{\Gamma(c+x+2)}, \quad x = 0, 1, \dots, \quad (1.1)$$

where Γ is the gamma function, and its mean and variance are $\mu = \frac{1}{c-1}$ and $\sigma^2 = \frac{c^2}{(c-2)(c-1)^2}$, respectively. This distribution was developed by Yule [11] in connection with biological data and can be found to provide some useful models

in Mitzutani [4], Simon [5, 6], Kendall [2], Haight [1] and Xekalaki [9, 10]. Let us consider (1.1), it can be written as follows:

$$\mathbb{Y}_c(x) = \begin{cases} \frac{c-1}{c} \left(\frac{c^2}{c^2-1} \right) & \text{if } x = 0, \\ \frac{c-1}{c} \left(\frac{1}{c} \right)^x \left[\frac{c^2}{c-1} \frac{c \cdots cx}{(c+x+1) \cdots (c+1)} \right] & \text{if } x \geq 1. \end{cases} \quad (1.2)$$

From (1.2), if c is large then $\mathbb{Y}_c(x)$ tends to $\frac{c-1}{c} \left(\frac{1}{c} \right)^x$ for every $x \in \mathbb{N} \cup \{0\}$, that is, the Yule distribution with parameter c converges to the geometric distribution with parameter $p = \frac{c-1}{c}$, $c > 2$. Thus, if we set the parameter of the geometric distribution to correspond the parameter of the Yule distribution, then the Yule distribution can also be approximated by the geometric distribution. In this case, Teerapabolarn [7] gave a uniform bound for this approximation, in the point metric form, as follows:

$$|\mathbb{Y}_c(x_0) - \mathbb{G}_p(x_0)| \leq \frac{3c-1}{(c^2-1)(c-1)}, \quad (1.3)$$

where $x_0 \in \mathbb{N} \cup \{0\}$ and $\mathbb{G}_p(x_0) = pq^{x_0} = p(1-p)^{x_0} = \frac{c-1}{c} \left(\frac{1}{c} \right)^{x_0}$. The bound is a good criteria for measuring the accuracy of the approximation; however, it is uniform bound. So, our interesting is to determine a good bound for each $x_0 \in \mathbb{N} \cup \{0\}$. In this study, we focus on determining a non-uniform bound for the point metric in (1.3) by using Stein's method and the w -function associated with the Yule random variable, and we also provide three numerical examples to illustrate the obtained result.

2. Preliminaries

The following lemmas present important properties related to Stein's method and the w -function associated with the Yule random variable.

Lemma 2.1. (Teerapabolarn [7]) *Let $w(X)$ be the w -function associated with the Yule random variable X and $c > 2$. Then, we have the following:*

$$w(x) = \frac{(x+1)^2}{(c-1)\sigma^2}, \quad x = 0, 1, \dots, \quad (2.1)$$

where $\sigma^2 = \frac{c^2}{(c-2)(c-1)^2}$.

Lemma 2.2. For $x_0 \in \mathbb{N} \cup \{0\}$ and $x \in \mathbb{N}$, let $\Delta f_{x_0}(x) = f_{x_0}(x+1) - f_{x_0}(x)$, then the following inequality holds:

$$|\Delta f_{x_0}(x)| \leq \begin{cases} \frac{1}{x(x+1)} & \text{if } x_0 = 0, \\ \min \left\{ \frac{1}{x}, \frac{1}{x_0} \right\} & \text{if } x_0 > 0, \end{cases} \tag{2.2}$$

where f_{x_0} is defined as in [3] when $r = 1$.

Proof. It follows from [8] and [3] for $x_0 = 0$ and $x_0 > 0$, respectively. \square

3. Result

The following Theorem presents our result of this study in terms of the point metric of the Yule and geometric distributions together with its non-uniform bound.

Theorem 3.1. For $x_0 \in \mathbb{N} \cup \{0\}$ and $c > 2$, if $p = \frac{c-1}{c}$, then we have the following:

$$|\mathbb{Y}_c(x_0) - \mathbb{G}_p(x_0)| \leq \begin{cases} \frac{1}{c(c+1)} & \text{if } x_0 = 0, \\ \frac{2}{c-1} \min \left\{ \frac{1}{c+1}, \frac{1}{(c-2)x_0} \right\} & \text{if } x_0 > 0, \end{cases} \tag{3.1}$$

Proof. We apply the proof of Theorem 3.1 in [3], with $r = 1$, to obtain

$$\begin{aligned} |\mathbb{Y}_c(x_0) - \mathbb{G}_p(x_0)| &\leq \mathbb{E}\{|(1 + X)q - \sigma^2 w(X)p|\} |\Delta f_{x_0}(X)| \\ &= \mathbb{E} \left\{ \left| \frac{1 + X}{c} - \frac{\sigma^2 w(X)(c - 1)}{c} \right| |\Delta f_{x_0}(X)| \right\} \\ &= \sum_{x=0}^{\infty} \left| \frac{1 + x}{c} - \frac{\sigma^2 w(x)(c - 1)}{c} \right| |\Delta f_{x_0}(x)| \mathbb{Y}_c(x) \\ &= \sum_{x=1}^{\infty} \frac{x(x + 1)}{c} |\Delta f_{x_0}(x)| \mathbb{Y}_c(x) \quad (\text{by Lemma 2.1}). \end{aligned} \tag{3.2}$$

For $x_0 = 0$, using Lemma 2.2, (3.2) becomes

$$|\mathbb{Y}_c(0) - \mathbb{G}_p(0)| \leq \frac{1}{c(c + 1)}, \tag{3.3}$$

and for $x_0 > 0$, we have

$$|\mathbb{Y}_c(x_0) - \mathbb{G}_p(x_0)| \leq \sum_{x=1}^{\infty} \frac{x(x + 1)}{c} \min \left\{ \frac{1}{x}, \frac{1}{x_0} \right\} \mathbb{Y}_c(x)$$

$$= \frac{2}{c-1} \min \left\{ \frac{1}{c+1}, \frac{1}{(c-2)x_0} \right\}. \quad (3.4)$$

Hence, from (3.3) and (3.4), (3.1) is obtained. \square

Remark. By comparing the two bounds, it is seen that the bound in (3.1) is sharper than the bounds in (1.3).

4. Numerical Examples

This section we give three numerical examples of the result to illustrate how well geometric distribution approximate the Yule distribution, and we also compare the bounds in (1.3) and (3.1).

4.1 For $c = 5$, a non-uniform bound for this approximation is of the form

$$|\mathbb{Y}_c(x_0) - \mathbb{G}_p(x_0)| \leq \begin{cases} 0.03333333 & \text{if } x_0 = 0, \\ 0.08333333 & \text{if } x_0 = 1, \\ \frac{0.16666667}{x_0} & \text{if } x_0 \geq 2, \end{cases}$$

which is sharper than the uniform bound in (1.3),

$$|\mathbb{Y}_c(x_0) - \mathbb{G}_p(x_0)| \leq 0.08333333, \quad x_0 \geq 0.$$

4.2 For $c = 10$, a non-uniform bound for this approximation is of the form

$$|\mathbb{Y}_c(x_0) - \mathbb{G}_p(x_0)| \leq \begin{cases} 0.00909091 & \text{if } x_0 = 0, \\ 0.02020202 & \text{if } x_0 = 1, \\ \frac{0.02777778}{x_0} & \text{if } x_0 \geq 2, \end{cases}$$

which is sharper than the uniform bound in (1.3),

$$|\mathbb{Y}_c(x_0) - \mathbb{G}_p(x_0)| \leq 0.02020202, \quad x_0 \geq 0.$$

4.3 For $c = 20$, a non-uniform bound for this approximation is of the form

$$|\mathbb{Y}_c(x_0) - \mathbb{G}_p(x_0)| \leq \begin{cases} 0.00238095 & \text{if } x_0 = 0, \\ 0.00501253 & \text{if } x_0 = 1, \\ \frac{0.00584795}{x_0} & \text{if } x_0 \geq 2, \end{cases}$$

which is sharper than the uniform bound in (1.3),

$$|\mathbb{Y}_c(x_0) - \mathbb{G}_p(x_0)| \leq 0.00501253, \quad x_0 \geq 0.$$

The three numerical examples suggest that a pointwise geometric approximation to the Yule distribution is quite efficient when c is sufficiently large, that is, the result of Theorem 3.1 gives a good approximation when c is sufficiently large. By numerical comparison, it is observed that the bound in (3.1) is sharper than the bound in (1.3), which satisfies the remark mentioned above.

References

- [1] F.A. Haight, Some statistical problems in connection with word association data, *J. Math. Psychol.*, **3** (1966), 217-233.
- [2] M.G. Kendall, Natural law in the social sciences, *J. Roy. Statist. Soc. Ser. A*, **124** (1961), 1-16.
- [3] P. Malingam, K. Teerapabolarn, A pointwise negative binomial approximation by w -functions, *Int. J. Pure Appl. Math.*, **69** (2011), 453-467.
- [4] S. Mitzutani, *Vocabulary in women's magazines*, Kokken-Hokoku, **4**, National Research Institute, Tokyo, 1953.
- [5] H.A. Simon, On a class of skew distributions, *Biometrika*, **42** (1955), 425-440.
- [6] H.A. Simon, Some further notes on a class of skew distribution functions, *Information and Control*, **3** (1960), 80-88.
- [7] K. Teerapabolarn, Bounds on approximating the Yule distribution by the Poisson and geometric distributions, *Int. J. Appl. Math. Stat.*, **14** (2009), 86-93.
- [8] K. Teerapabolarn, Non uniform bounds on geometric approximation via Stein's method and w -functions, *Comm. Statist. Theory Meth.*, **40** (2011), 145-158.
- [9] E. Xekalaki, Linear regression and the Yule distribution, *J. Econ.*, **24** (1984), 397-403.
- [10] E. Xekalaki, A property of the Yule distribution and its applications, *Comm. Statist. Ser. A*, **12** (1983), 1181-1189.

- [11] G.U. Yule, A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis F.R.S., *Phil. Trans. Roy. Soc. London. Ser. B*, **213** (1925), 21-87.