International Journal of Pure and Applied Mathematics
Volume 91 No. 1 2014, 83-86
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: http://dx.doi.org/10.12732/ijpam.v91i1.9
ijpam.eu

AN IMPROVED GEOMETRIC DISTRIBUTION TO APPROXIMATE THE YULE DISTRIBUTION

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

Abstract: An improved geometric distribution with parameter p is derived to approximate the Yule distribution with parameter c, where $p = 1-q = \frac{c-1}{c}$. The improved approximation is more accurate than the geometric approximation when c is sufficiently large.

AMS Subject Classification: 62E17, 60F05
Key Words: geometric approximation, geometric probability function, Yule probability function

1. Introduction

Let X be the Yule random variable with parameter $c > 2$, and its probability function is given by

$$\mathbf{y}_c(x) = \frac{cx!\Gamma(c+1)}{\Gamma(c+x+2)}\ ,\ x = 0, 1, ..., \tag{1.1}$$

and its mean and variance are $E(X) = \frac{1}{c-1}$ and $Var(X) = \frac{c^2}{(c-2)(c-1)^2}$, respectively. This distribution was developed by Yule [7] in connection with biological

Received: November 22, 2013

data, and it can be found to provide some useful models in [1, 3, 4, 6]. Let us consider (1.1), it can be written as follows:

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

From (1.2), if c is large then $\mathbf{y}_c(x)$ goes to $\mathbf{g}_p(x) = \frac{c-1}{c}\left(\frac{1}{c}\right)^x$ for every $x \in \mathbb{N}\cup\{0\}$, that is, the Yule probability function with parameter c can be approximated by a geometric probability function with parameter $p = \frac{c-1}{c}$, $c > 2$. In this case, Teerapabolarn [5] gave a bound on $\left|\mathbf{y}_c(x) - \mathbf{g}_p(x)\right|$ for $x \in \mathbb{N} \cup \{0\}$.

In this paper, we give an improved geometric probability function, $\widehat{\mathbf{g}}_p(x)$, for approximating the Yule probability function, and the accuracy of the approximation is measured in the form of $|\mathbf{y}_c(x) - \widehat{\mathbf{g}}_p(x)|$ for $x \in \mathbb{N} \cup \{0\}$. The result of this study is in Section 2. In Section 3, some numerical examples are given to illustrate the improved approximation and the conclusion of this study is presented in the last section.

2. Result

Applying the property in [2], the following lemma is also obtained.

Lemma 2.1. *For $x, N \in \mathbb{N}$, then*

$$\prod_{i=0}^{x-1}\left(q + \frac{i}{N}\right) = q^x\left[1 + \frac{x(x-1)}{2Nq}\right] + O\left(\frac{1}{N^2}\right). \tag{2.1}$$

Theorem 2.1. *Let $x \in \mathbb{N} \cup \{0\}$ and $p = \frac{c-1}{c}$. Then we have the following:*

$$\mathbf{y}_c(x) = \widehat{\mathbf{g}}_p(x) + O\left(\frac{1}{c^2}\right) \tag{2.2}$$

and for large c,

$$\widehat{\mathbf{g}}_p(x) = \mathbf{y}_c(x), \tag{2.3}$$

where $\widehat{\mathbf{g}}_p(x) = \mathbf{g}_p(x)\left\{1 + \frac{x(x-1)}{2}\right\} / \left\{1 + \frac{x(x+3)}{2c}\right\}$.

Proof. For $x = 0$, applying Lemma 2.1, we have that

$$\mathbf{y}_c(0) = \frac{c-1}{c}\left(\frac{c^2}{c^2-1}\right) = p\left(1 + \frac{1}{c^2-1}\right) = \widehat{\mathbf{g}}_p(0) + O\left(\frac{1}{c^2}\right).$$

Next, we have to show that (2.2) holds for $x \in \mathbb{N} \cup \{0\}$. Following (1.2), by using Lemma 2.1, we can obtain

$$\begin{aligned}\mathbf{y}_c(x) &= pq^x \left(\frac{c^2}{c-1}\right) \frac{\prod_{i=0}^{x-1}\left(q+\frac{i}{c}\right)}{\prod_{i=1}^{x+1}\left(1+\frac{i}{c}\right)} \\ &= \frac{\mathbf{g}_p(x)}{1+\frac{x(x+3)}{2c}} \left\{1+\frac{x(x-1)}{2}\right\} + O\left(\frac{1}{c^2}\right) \\ &= \widehat{\mathbf{g}}_p(x) + O\left(\frac{1}{c^2}\right).\end{aligned}$$

As c is large, we get $O\left(\frac{1}{c^2}\right) = 0$. Hence $\widehat{\mathbf{g}}_p(x) = \mathbf{y}_c(x)$. □

3. Numerical examples

The following examples have been given to illustrate how well the improved geometric distribution approximates the Yule distribution (when c is sufficiently large).

3.1. Let $c = 10$, then $p = 0.9$ and the numerical results are as follows:

x	$\mathbf{y}_c(x)$	$\widehat{\mathbf{g}}_p(x)$	$\mathbf{g}_p(x)$	$\lvert\mathbf{y}_c(x)-\widehat{\mathbf{g}}_p(x)\rvert$	$\lvert\mathbf{y}_c(x)-\mathbf{g}_p(x)\rvert$
0	0.83333333	0.80000000	0.80000000	0.03333333	0.03333333
1	0.11904762	0.11428571	0.16000000	0.00476190	0.04095238
2	0.02976190	0.03200000	0.03200000	0.00223810	0.00223810
3	0.00992063	0.00914286	0.00640000	0.00077778	0.00352063
4	0.00396825	0.00235789	0.00128000	0.00161036	0.00268825
5	0.00180375	0.00056320	0.00025600	0.00124055	0.00154775
6	0.00090188	0.00012800	0.00005120	0.00077388	0.00085068
7	0.00048563	0.00002816	0.00001024	0.00045747	0.00047539
8	0.00027750	0.00000606	0.00000205	0.00027144	0.00027545
9	0.00016650	0.00000128	0.00000041	0.00016522	0.00016609
10	0.00010406	0.00000027	0.00000008	0.00010379	0.00010398
11	0.00006733	0.00000006	0.00000002	0.00006728	0.00006732

3.2. Let $c = 15$, then $p = \frac{14}{15}$ and the numerical results are as follows:

x	$\mathbf{y}_c(x)$	$\widehat{\mathbf{g}}_p(x)$	$\mathbf{g}_p(x)$	$\lvert\mathbf{y}_c(x)-\widehat{\mathbf{g}}_p(x)\rvert$	$\lvert\mathbf{y}_c(x)-\mathbf{g}_p(x)\rvert$
0	0.93750000	0.93333333	0.93333333	0.00416667	0.00416667
1	0.05514706	0.05490196	0.06222222	0.00024510	0.00707516
2	0.00612745	0.00622222	0.00414815	0.00009477	0.00197930
3	0.00096749	0.00069136	0.00027654	0.00027613	0.00069095
4	0.00019350	0.00006675	0.00001844	0.00012675	0.00017506
5	0.00004607	0.00000579	0.00000123	0.00004028	0.00004484
6	0.00001256	0.00000047	0.00000008	0.00001210	0.00001248
7	0.00000382	0.00000004	0.00000001	0.00000379	0.00000382

For the approximation of the Yule distribution in the examples 3.1 and 3.2, it can be seen that the improved geometric distribution is more appropriate than the geometric distribution.

4. Conclusion

In this study, an improved geometric distribution with parameter $p = \frac{c-1}{c}$ was derived to approximate the Yule distribution. In addition, the improvement of the geometric approximation is more accurate than the geometric approximation when c is sufficiently large.

References

[1] F.A. Haight, Some statistical problems in connection with word association data, *J. Math. Psychol.*, **3** (1966), 217–233.

[2] D.P. Hu, Y.Q. Cui, A.H. Yin , An improved negative binomial approximation for negative hypergeometric distribution, *Appl. Mech. Mater.*, **427-429** (2013), 2549–2553.

[3] H.A. Simon, On a class of skew distributions, *Biometrika*, **42** (1955), 425–440.

[4] H.A. Simon, Some further notes on a class of skew distribution functions, *Information and Control*, **3** (1960), 80-88.

[5] K. Teerapabolarn, A non-uniform bound on geometric approximation to the Yule distribution, *Int. J. Pure Appl. Math.*, **82** (2013), 731–736.

[6] E. Xekalaki, A property of the Yule distribution and its applications, *Comm. Statist. Ser. A*, **12** (1983), 1181–1189.

[7] 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.