Improved Negative Binomial Approximation to Negative Pólya Distribution

Kanint Teerapabolarn [#1]

[#] *Department of Mathematics, Faculty of Science, Burapha University,*
Chonburi, 20131, Thailand

Abstract : *We determine an improved negative binomial distribution with parameters $n \in \mathbb{N}$ and $0 < p < 1$ from a negative Pólya distribution with parameters N, n and r, where $p = \frac{n-r-1}{N-1}$. Following improved negative binomial and negative binomial approximations to the negative Pólya distribution, the improved negative binomial approximation is more accurate than the negative binomial approximation..*

Keywords *Improved negative binomial distribution, negative binomial distribution, negative Pólya distribution.*

I. INTRODUCTION

Let X be a non-negative integer-valued random variable that has a negative Pólya distribution with parameters N, n and r. By following [1], its probability function can be expressed as

$$np_{N,n,r}(x) = \binom{n+x-1}{x} \frac{(r+x-1)!(N-r+n-1)!(N-1)!}{(r-1)!(N-r-1)!(N+n+x-1)!}, \quad x = 0,1,..., \tag{1}$$

where the mean and variance of X are $\mu = E(X) = \frac{nr}{N-r-1}$ and $\sigma^2 = Var(X) = \frac{nr(n+N-r-1)(N-1)}{(N-r-2))(N-r-1)^2}$, respectively.

Let us consider the probability function in (1), it is seen that

$$np_{N,n,r}(x) = \begin{cases} \frac{(N-r+n-1)\cdots(N-r)}{(N+n-1)\cdots N} & , x = 0, \\ \binom{n+x-1}{x} \frac{[(r+x-1)\cdots r][(N-r+n-1)\cdots(N-r)]}{(N+n+x-1)\cdots N}, & x = 1,2,... \end{cases}$$

$$= \begin{cases} \frac{\prod_{i=1}^{n}\left(\frac{N-r-1}{N-1}+\frac{i}{N-1}\right)}{\prod_{i=1}^{n}\left(1+\frac{i}{N-1}\right)} & , x = 0, \\ \binom{n+x-1}{x} \frac{\prod_{i=0}^{x-1}\left(\frac{r}{N-1}+\frac{i}{N-1}\right)\prod_{i=1}^{n}\left(\frac{N-r-1}{N-1}+\frac{i}{N-1}\right)}{\prod_{i=1}^{n+x}\left(1+\frac{i}{N-1}\right)}, & x = 1,2,... \end{cases} \tag{2}$$

From which, it can be observed that if $N, r \to \infty$ in such a way that $0 < \frac{r}{N-1} < 1$ remains a constant, then $np_{N,n,r}(x) \to nb_{n,p}(x)$ for every $x \in \mathbb{N} \cup \{0\}$, where $nb_{n,p}(x) = \binom{n+x-1}{x} q^x p^n$ and $q = \frac{r}{N-1}$ and $p = \frac{N-r-1}{N-1}$. Therefore, the negative Pólya distribution with parameters N, n and r can be approximated by a negative binomial distribution with parameters n and $p = \frac{N-r-1}{N-1}$. However, it may be improved the negative binomial approximation to the Pólya distribution by using the methodology of [2]. Thus, in this study, we are interested to determine an improved negative binomial distribution by improving from the Pólya distribution form in (2).

II. RESULT

This section, we use the methodology of [2] and the Pólya distribution form of (2) to derive an improved negative binomial distribution. Before giving the result, we also need the following lemma, which can be directly applied from [2].

Lemma 1. *For $x, M \in \mathbb{N}$ and $0 < q < 1$, the following equalities hold:*

$$\prod_{i=1}^{x}\left(1+\frac{i}{M}\right) = 1 + \frac{x(x+1)}{2M} + O\left(\frac{1}{M^2}\right). \tag{3}$$

and

$$\prod_{i=0}^{x-1}\left(q+\frac{i}{M}\right) = q^x\left[1+\frac{x(x-1)}{2Mq}\right] + O\left(\frac{1}{M^2}\right). \tag{4}$$

Proof. The results in (3) and (4) directly follow mathematical induction. $\square$

The following theorem presents the desired result.

Theorem 1. *Let* $x \in ¥ \cup \{0\}$, $p = \frac{N-r-1}{N-1}$ *and* $q = \frac{r}{N-1}$, *we then have the following:*

$$np_{N,n,r}(x) = \tilde{n}b_{n,p}(x) + O\left(\frac{1}{(N-1)^2}\right) \tag{5}$$

and if N *is large, then*

$$np_{N,n,r}(x) \approx \tilde{n}b_{n,p}(x), \tag{6}$$

where $\tilde{n}b_{n,p}(x) = \frac{nb_{n,p}(x)}{1+\frac{(n+x)(n+x+1)}{2(N-1)}}\left\{1 + \frac{x(x-1)}{2r} + \frac{n(n+1)}{2(N-r-1)}\right\}$.

Proof. We shall show the result in (5) by following (2). For $x = 0$, we obtain

$$\begin{aligned}
np_{N,n,r}(0) &= \frac{\prod_{i=1}^{n}\left(\frac{N-r-1}{N-1} + \frac{i}{N-1}\right)}{\prod_{i=1}^{n}\left(1 + \frac{i}{N-1}\right)} \\
&= \frac{\prod_{i=1}^{n}\left(p + \frac{i}{N-1}\right)}{\prod_{i=1}^{n}\left(1 + \frac{i}{N-1}\right)} \\
&= \frac{p^n\left[1 + \frac{n(n+1)}{2(N-1)p} + O\left(\frac{1}{(N-1)^2}\right)\right]}{1 + \frac{n(n+1)}{2(N-1)} + O\left(\frac{1}{(N-1)^2}\right)} \qquad \text{(by Lemma 1)} \\
&= \frac{nb_{n,p}(0)}{1 + \frac{n(n+1)}{2(N-1)}}\left\{1 + \frac{n(n+1)}{2(N-1)p}\right\} + O\left(\frac{1}{(N-1)^2}\right) \\
&= \tilde{n}b_{n,p}(0) + O\left(\frac{1}{(N-1)^2}\right).
\end{aligned} \tag{7}$$

For $x > 0$, we have

$$\begin{aligned}
np_{N,n,r}(x) &= \binom{n+x-1}{x}\frac{\prod_{i=0}^{x-1}\left(\frac{r}{N-1} + \frac{i}{N-1}\right)\prod_{i=1}^{n}\left(\frac{N-r-1}{N-1} + \frac{i}{N-1}\right)}{\prod_{i=1}^{n+x}\left(1 + \frac{i}{N-1}\right)} \\
&= \binom{n+x-1}{x}\frac{\prod_{i=0}^{x-1}\left(q + \frac{i}{N-1}\right)\prod_{i=1}^{n}\left(p + \frac{i}{N-1}\right)}{\prod_{i=1}^{n+x}\left(1 + \frac{i}{N-1}\right)} \\
&= \frac{\binom{n+x-1}{x}q^x p^n}{1 + \frac{(n+x)(n+x+1)}{2(N-1)}}\left\{1 + \frac{x(x-1)}{2(N-1)q} + \frac{n(n+1)}{2(N-1)p}\right\} + O\left(\frac{1}{(N-1)^2}\right) \\
&= \frac{nb_{n,p}(x)}{1 + \frac{(n+x)(n+x+1)}{2(N-1)}}\left\{1 + \frac{x(x-1)}{2r} + \frac{n(n+1)}{2(N-r-1)}\right\} + O\left(\frac{1}{(N-1)^2}\right) \\
&= \tilde{n}b_{n,p}(x) + O\left(\frac{1}{(N-1)^2}\right).
\end{aligned} \tag{8}$$

Hence, by (7) and (8), the result in (5) is obtained. $\square$

III. NUMERICAL RESULTS

We give some numerical examples to illustrate how well the improved negative binomial distribution approximates the negative Pólya distribution. In addition, we also give some numerical comparisons of improved negative binomial and negative binomial approximations to the negative Pólya distribution.

Example 1. Let $N = 20, n = 5, r = 5$ and $p = \frac{N-r-1}{N-1} = 0.73684211$ then the numerical results are as follows:

x	$np_{N,n,r}(x)$	$nb_{n,p}(x)$	$\tilde{n}b_{n,p}(x)$	$\lvert np_{N,n,r}(x) - nb_{n,p}(x)\rvert$	$\lvert np_{N,n,r}(x) - \tilde{n}b_{n,p}(x)\rvert$
0	0.09791588	0.04717852	0.05970178	0.05073735	0.03821410
1	0.16319313	0.12415401	0.13677807	0.03903912	0.02641506
2	0.17372172	0.17969659	0.18051700	0.00597488	0.00679529
3	0.15200650	0.18915431	0.18063008	0.03714781	0.02862358
4	0.11976270	0.16177671	0.15191728	0.04201401	0.03215458
5	0.08876529	0.11920389	0.11289454	0.03043860	0.02412924
6	0.06340378	0.07842361	0.07621908	0.01501983	0.01281530
7	0.04428201	0.04717210	0.04756781	0.00289009	0.00328580
8	0.03051868	0.02637913	0.02777346	0.00413955	0.00274522
9	0.02088120	0.01388375	0.01530819	0.00699745	0.00557301
10	0.01424205	0.00694188	0.00802235	0.00730017	0.00621970
11	0.00971049	0.00332147	0.00402088	0.00638902	0.00568961
12	0.00663155	0.00152963	0.00193702	0.00510193	0.00469454
13	0.00454249	0.00068121	0.00090068	0.00386128	0.00364181
14	0.00312391	0.00029451	0.00040570	0.00282940	0.00271820
15	0.00215834	0.00012400	0.00017759	0.00203433	0.00198075
16	0.00149884	0.00005099	0.00007574	0.00144786	0.00142310
17	0.00104651	0.00002052	0.00003155	0.00102599	0.00101495
18	0.00073478	0.00000810	0.00001287	0.00072668	0.00072192
19	0.00051886	0.00000314	0.00000514	0.00051572	0.00051371
20	0.00036850	0.00000120	0.00000202	0.00036730	0.00036648
21	0.00026321	0.00000045	0.00000078	0.00026276	0.00026243
22	0.00018908	0.00000017	0.00000030	0.00018891	0.00018878
23	0.00013659	0.00000006	0.00000011	0.00013653	0.00013648
24	0.00009922	0.00000002	0.00000004	0.00009920	0.00009918

Example 2. Let $N = 50,\ n = 5,\ r = 10$ and $p = \frac{N-r-1}{N-1} = 0.79591837$ then the numerical results are as follows:

x	$np_{N,n,r}(x)$	$nb_{n,p}(x)$	$\tilde{n}b_{n,p}(x)$	$\lvert np_{N,n,r}(x) - nb_{n,p}(x)\rvert$	$\lvert np_{N,n,r}(x) - \tilde{n}b_{n,p}(x)\rvert$
0	0.34340065	0.31940568	0.33860073	0.02399497	0.00479992
1	0.31218241	0.32592416	0.31589573	0.01374175	0.00371332
2	0.18396464	0.19954541	0.18852157	0.01558077	0.00455693
3	0.09036859	0.09502162	0.09227846	0.00465303	0.00190987
4	0.04051006	0.03878434	0.04012370	0.00172572	0.00038636
5	0.01730260	0.01424731	0.01600714	0.00305529	0.00129546
6	0.00720942	0.00484602	0.00595623	0.00236339	0.00125319
7	0.00297156	0.00155412	0.00208945	0.00141745	0.00088212
8	0.00122218	0.00047575	0.00069679	0.00074642	0.00052538
9	0.00050439	0.00014024	0.00022243	0.00036415	0.00028196
10	0.00020964	0.00004007	0.00006837	0.00016957	0.00014127
11	0.00008796	0.00001115	0.00002033	0.00007681	0.00006763
12	0.00003732	0.00000303	0.00000588	0.00003428	0.00003144
13	0.00001602	0.00000081	0.00000166	0.00001521	0.00001437
14	0.00000697	0.00000021	0.00000046	0.00000676	0.00000651
15	0.00000307	0.00000005	0.00000012	0.00000302	0.00000295

Example 3. Let $N = 80,\ n = 10,\ r = 5$ and $p = \frac{N-r-1}{N-1} = 0.93670886$ then the numerical results are as follows:

x	$np_{N,n,r}(x)$	$nb_{n,p}(x)$	$\tilde{nb}_{n,p}(x)$	$\lvert np_{N,n,r}(x) - nb_{n,p}(x)\rvert$	$\lvert np_{N,n,r}(x) - \tilde{nb}_{n,p}(x)\rvert$
0	0.54297267	0.52005142	0.53447398	0.02292125	0.00849868
1	0.30165148	0.32914647	0.31261246	0.02749499	0.01096097
2	0.10939010	0.11457630	0.11203389	0.00518620	0.00264379
3	0.03329264	0.02900666	0.03158590	0.00428598	0.00170674
4	0.00930762	0.00596656	0.00753979	0.00334106	0.00176783
5	0.00249523	0.00105736	0.00157126	0.00143787	0.00092398
6	0.00065664	0.00016730	0.00029159	0.00048934	0.00036505
7	0.00017198	0.00002420	0.00004898	0.00014777	0.00012300
8	0.00004521	0.00000326	0.00000755	0.00004196	0.00003766
9	0.00001199	0.00000041	0.00000108	0.00001158	0.00001091
10	0.00000322	0.00000005	0.00000015	0.00000317	0.00000308

The numerical results in Examples 3.1, 3.2 and 3.3 indicate that (1) the theorem gives a good improved negative binomial approximation to the negative Pólya distribution if N is large (2) the improved negative binomial approximation is more accurate than the negative binomial approximation.

IV. CONCLUSION

In this study, an improvement of the negative binomial distribution for approximating the negative Pólya distribution was obtained. The improved negative binomial distribution can be used as a good approximation of the negative Pólya distribution when N is large. Moreover, by comparing the approximation, the improved negative binomial distribution also gives more accurate than the negative binomial distribution.

REFERENCES

[1] A. Hassan and S. Bilal, "On estimation of negative Polya-Eggenberger distribution and its applications", *Journal of the Korea Society for Industrial and Applied Mathematics*, vol. 12, pp. 81-59, 2008.

[2] D. P. Hu, Y. Q. Cui and A. H. Yin, An improved negative binomial approximation for negative hypergeometric distribution", *Applied Mechanics and Materials*, Vols. 427-429, pp. 2549-2553, 2013.