BULLETIN *of the*
MALAYSIAN MATHEMATICAL
SCIENCES SOCIETY
http://math.usm.my/bulletin

Bull. Malays. Math. Sci. Soc. (2) **34**(2) (2011), 331–336

On the Poisson Approximation to the Negative Hypergeometric Distribution

K. TEERAPABOLARN
Department of Mathematics, Faculty of Science,
Burapha University, Chonburi 20131, Thailand
kanint@buu.ac.th

Abstract. In this paper, the Stein-Chen method and the w-function associated with the negative hypergeometric random variable are used to give a result of the Poisson approximation to the negative hypergeometric distribution in terms of the total variation distance. Some numerical examples are presented to illustrate the result obtained.

2010 Mathematics Subject Classification: 60F05

Keywords and phrases: Negative hypergeometric distribution, Poisson approximation, Stein-Chen method, w-function.

1. Introduction

Suppose a set or a population of items of size $S + R$ consists of S items of special type and R items of non-special type, respectively. Items are drawn at random, one at a time, without replacement from this population until the number of non-special items reaches a fixed number r. Let X be the number of special items in the sample. Then X has a negative hypergeometric distribution, sometimes called inverse hypergeometric distribution, with parameters R, S and r, denoted by $\mathcal{NH}(R, S, r)$. Its probability function can be expressed as

$$p_X(k) = \frac{\binom{r+k-1}{k}\binom{R-r+S-k}{S-k}}{\binom{R+S}{S}}, \quad k = 0, 1, ..., S, \tag{1.1}$$

where $R, S \in \mathbb{N}$ and $r \in \{1, ..., R\}$. The mean and variance of X are $\mu = \mathbb{E}(X) = \frac{rS}{R+1}$ and $\sigma^2 = \mathbb{V}ar(X) = \frac{rS(R+S+1)(R-r+1)}{(R+1)^2(R+2)}$, respectively. This distribution was used by Kaigh and Lachenbruch [7] in resampling for nonparametric quantile estimation. Its other applications can be found in [5, 6, 9, 13, 14], etc.

Note that this distribution is a finite sample analogy to the negative binomial distribution, which arises in a scheme of sampling with replacement. Moreover, if

Communicated by Anton Abdulbasah Kamil.
Received: Novemeber 12, 2008; *Revised:* January 19, 2010.

$R, S \to \infty$ in such a way $\frac{S}{R+1}$ tends to a constant θ, then the negative hypergeometric distribution converges to the negative binomial with parameters r and $\frac{\theta}{1+\theta}$. Similarly, this distribution may converge to the binomial or Poisson or normal distribution if the conditions on their parameters are appropriate. Let us consider the probability function in (1.1), it can be seen that, for $k \in \{0, 1, ..., S\}$,

$$\begin{aligned} p_X(k) &= \binom{S}{k} \frac{(r+k-1)!}{(r-1)!} \frac{(R-r+S-k)!}{(R-r)!} \frac{R!}{(R+S)!} \\ &= \binom{S}{k} (r+k-1)(r+k-2)\cdots r \\ &\quad \times \frac{(R-r+S-k)(R-r+S-k-1)\cdots(R-r+1)}{(R+S)(R+S-1)\cdots(R+1)} \\ &= \binom{S}{k} \frac{r+k-1}{R+1} \frac{r+k-2}{R+1} \cdots \frac{r}{R+1} \\ &\quad \times \frac{\frac{R-r+S-k}{R+1} \frac{R-r+S-k-1}{R+1} \cdots \frac{R-r+1}{R+1}}{\frac{R+S}{R+1} \frac{R+S-1}{R+1} \cdots \frac{R+1}{R+1}} \\ &= \binom{S}{k} \left(\frac{r}{R+1}\right)^k \left(\frac{R-r+1}{R+1}\right)^{S-k} \left(1 + \frac{k-1}{r}\right)\left(1 + \frac{k-2}{r}\right) \cdots 1 \\ &\quad \times \frac{\left(1 + \frac{S-k-1}{R-r+1}\right)\left(1 + \frac{S-k-2}{R-r+1}\right)\cdots 1}{\left(1 + \frac{S-1}{R+1}\right)\left(1 + \frac{S-2}{R+1}\right)\cdots 1}, \end{aligned} \tag{1.2}$$

and if $R, r \to \infty$ in such a way $\frac{r}{R+1} \to p = 1 - q$, then $p_X(k) \to \binom{S}{k} p^k q^{S-k}$, $k = 0, 1, ..., S$, which is the binomial distribution with parameters S and p, denoted by $\mathcal{B}(S, p)$. It is well known that if p is small and $Sp \to \lambda$ as $S \to \infty$, then $\mathcal{B}(S, p)$ can be approximated by $\mathcal{P}o(\lambda)$, where $\mathcal{P}o(\lambda)$ is the Poisson distribution with mean λ. Therefore, in view of (1.2), $\mathcal{NH}(R, S, r)$ can also be approximated by $\mathcal{P}o(\lambda)$, under appropriate conditions on their parameters. It should be noted that if $\frac{r}{R+1}$ is not small and S is sufficiently large, then $\mathcal{NH}(R, S, r)$ can also be approximated by the normal distribution with mean $\frac{Sr}{R+1}$ and variance $\frac{Sr(R-r+1)}{(R+1)^2}$. In this case, a bound on the normal approximation can be derived by using the same method in Boonta and Neammanee [2].

In this paper, the negative hypergeometric distribution has been approximated by Poisson distribution, and the accuracy of the approximation is measured in terms of an upper bound for the total variation distance. The total variation distance between $\mathcal{NH}(R, S, r)$ and $\mathcal{P}o(\lambda)$ is defined by

$$\begin{aligned} d_{TV}(\mathcal{NH}(R, S, r), \mathcal{P}o(\lambda)) &= \sup_A |\mathcal{NH}(R, S, r)\{A\} - \mathcal{P}o(\lambda)\{A\}| \\ &= \sup_A \left| \sum_{k \in A} p_X(k) - \sum_{k \in A} \frac{e^{-\lambda}\lambda^k}{k!} \right|, \end{aligned} \tag{1.3}$$

where A is a subset of $\mathbb{N} \cup \{0\}$ and $p_X(k)$ is defined in (1.1).

The tools for giving an upper bound for the total variation distance are the Stein-Chen method and the w-function associated with the negative hypergeometric random variable, which are introduced and applied in Section 2. In Section 3, some numerical examples are presented to illustrate the obtained result.

2. Main result

We will prove our main result using the Stein-Chen method together with the w-function associated with the negative hypergeometric random variable.

2.1. The w-function

The w-functions were studied and used by many authors, among others by Cacoullos and Papathanasiou [3], Papathanasiou and Utev [10], and Majsnerowska [8]. In the note by the latter, the following recurrence relation can be found:

$$w(k) = \frac{p_X(k-1)}{p_X(k)} w(k-1) + \frac{\mu - k}{\sigma^2} \geq 0, \ k \in S(x) \setminus \{0\}, \tag{2.1}$$

where $w(0) = \frac{\mu}{\sigma^2}$, $S(x)$ is support of X, $p_X(k) > 0$ for all $k \in S(x)$ and μ and $\sigma^2 \in (0, \infty)$ are mean and variance of X, respectively. Using the relation (2.1), we give the form of the w-function associated with the negative hypergeometric random variable in the following lemma.

Lemma 2.1. *Let $w(X)$ be the w-function associated with the negative hypergeometric random variable X, then*

$$w(k) = \frac{(r+k)(S-k)}{(R+1)\sigma^2}, \ k = 0, 1, ..., S, \tag{2.2}$$

where $\sigma^2 = \frac{rS(R+S+1)(R-r+1)}{(R+1)^2(R+2)}$.

Proof. Following (2.1), the recurrence relation of w-function associated with the random variable X can be written as

$$w(k) = \frac{rS}{(R+1)\sigma^2} + w(k-1)\frac{k(R-r+S-k+1)}{(r+k-1)(S-k+1)} - \frac{k}{\sigma^2}, \ k = 1, ..., S, \tag{2.3}$$

where $w(0) = \frac{rS}{(R+1)\sigma^2}$. In the next step, we shall show that (2.2) holds for every $k \in \{1, ..., S\}$. From (2.3),

$$w(1) = \frac{rS}{(R+1)\sigma^2} + w(0)\frac{R-r+S}{rS} - \frac{1}{\sigma^2} = \frac{(r+1)(S-1)}{(R+1)\sigma^2}.$$

Assuming that $w(i-1) = \frac{(r+i-1)(S-i+1)}{(R+1)\sigma^2}$, for $1 < i-1 < S$, we have

$$w(i) = \frac{rS}{(R+1)\sigma^2} + w(i-1)\frac{i(R-r+S-i+1)}{(r+i-1)(S-i+1)} - \frac{i}{\sigma^2} = \frac{(r+i)(S-i)}{(R+1)\sigma^2}.$$

Therefore, by mathematical induction, (2.2) holds for every $k \in \{1, ..., S\}$. ∎

The next relation stated by Cacoullos and Papathanasiou [3] is crucial for obtaining our main result.

If a non-negative integer-valued random variable X has $p_X(k) > 0$ for every $k \in S(x)$ and $0 < \sigma^2 = \mathbb{V}ar(X) < \infty$, then

$$\mathbb{C}ov(X, g(X)) = \sigma^2 \mathbb{E}[w(X)\Delta g(X)], \tag{2.4}$$

for any function $g : \mathbb{N} \cup \{0\} \to \mathbb{R}$ for which $\mathbb{E}|w(X)\Delta g(X)| < \infty$, where $\Delta g(x) = g(x+1) - g(x)$.

2.2. The Stein-Chen method

The classical Stein's method was first introduced by Stein [11] in 1972. It is a starting tool for approximating the distribution of random elements. His original work was applied to central limit theorem for sums of random variables. The version appropriate for the Poisson case was first developed by Chen [4] in 1975. It is referred to as the Stein-Chen method. The Stein-Chen equation for the Poisson distribution with parameter $\lambda > 0$ is, for given h, of the form

$$h(x) - \mathcal{P}_\lambda(h) = \lambda g(x+1) - xg(x), \tag{2.5}$$

where $\mathcal{P}_\lambda(h) = \sum_{l=0}^{\infty} h(l)\frac{e^{-\lambda}\lambda^l}{l!}$ and g and h are bounded real-valued functions defined on $\mathbb{N} \cup \{0\}$.

For $A \subseteq \mathbb{N} \cup \{0\}$, let $h_A : \mathbb{N} \cup \{0\} \to \mathbb{R}$ be defined by

$$h_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \tag{2.6}$$

Barbour *et al.* showed in [1] that the solution $g = g_A$ of (2.5) can be expressed in the form

$$g(x) = \begin{cases} (x-1)!\lambda^{-x}e^{\lambda}[\mathcal{P}_\lambda(h_{A\cap C_{x-1}}) - \mathcal{P}_\lambda(h_A)\mathcal{P}_\lambda(h_{C_{x-1}})] & \text{if } x \geq 1, \\ 0 & \text{if } x = 0, \end{cases} \tag{2.7}$$

where $C_x = \{0, ..., x\}$, and that for any subset A of $\mathbb{N} \cup \{0\}$ and $k \in \mathbb{N}$, the following bound is valid.

$$\sup_{A,k} |\Delta g(k)| = \sup_{A,k} |g(k+1) - g(k)| \leq \lambda^{-1}(1 - e^{-\lambda}). \tag{2.8}$$

Putting $h = h_A$ and taking expectation in (2.5), we get

$$\mathcal{NH}(R, S, r)\{A\} - \mathcal{P}o(\lambda)\{A\} = \mathbb{E}[\lambda g(X+1) - Xg(X)], \tag{2.9}$$

where g is defined in (2.7).

In the following theorem using the Stein-Chen method, we present an upper bound for the total variation distance between the negative hypergeometric and Poisson distributions.

Theorem 2.1. *Let X be the negative hypergeometric random variable, $\lambda = \frac{rS}{R+1}$ and $r \geq (S-1)$. Then, for $A \subseteq \mathbb{N} \cup \{0\}$,*

$$d_{TV}(\mathcal{NH}(R, S, r), \mathcal{P}o(\lambda)) \leq \left(1 - e^{-\lambda}\right) \frac{(R+1)(r+1) - S(R-r+1)}{(R+1)(R+2)}. \tag{2.10}$$

Proof. From (2.9) and (2.4), it follows that

$$\begin{aligned} |\mathcal{NH}(R, S, r)\{A\} - \mathcal{P}o(\lambda)\{A\}| &= |\lambda\mathbb{E}[g(X+1)] - \mathbb{E}[Xg(X)]| \\ &= |\lambda\mathbb{E}[g(X+1)] - \mathbb{C}ov(X, g(X)) - \mu\mathbb{E}[g(X)]| \\ &= |\lambda\mathbb{E}[\Delta g(X)] - \mathbb{C}ov(X, g(X))| \\ &= |\lambda\mathbb{E}[\Delta g(X)] - \sigma^2\mathbb{E}[w(X)\Delta g(X)]| \end{aligned}$$

$$\le \mathbb{E}|[\lambda - \sigma^2 w(X)]\Delta g(X)|$$
$$\le \sup_{x\ge 1} |\Delta g(x)|\mathbb{E}|\lambda - \sigma^2 w(X)|.$$

Thus, by (2.8),

$$d_{TV}(\mathcal{NH}(R,S,r), \mathcal{P}o(\lambda)) \le \lambda^{-1}(1-e^{-\lambda})\mathbb{E}|\lambda - \sigma^2 w(X)|. \tag{2.11}$$

In view of Lemma 2.1, we have, for $k \in \{0,1,...,S\}$,

$$\begin{aligned}\lambda - \sigma^2 w(k) &= \frac{rS}{R+1} - \frac{(r+k)(S-k)}{R+1}\\ &= \frac{(r-S+k)k}{R+1} \ge 0,\end{aligned}$$

which implies that

$$\begin{aligned}\mathbb{E}|\lambda - \sigma^2 w(X)| &= \lambda - \sigma^2\mathbb{E}[w(X)]\\ &= \lambda - \sigma^2\\ &= \lambda\frac{(R+1)(r+1) - S(R-r+1)}{(R+1)(R+2)}.\end{aligned}$$

Substituting this result into (2.11), finishes the proof of the theorem. ∎

Corollary 2.1. *For $r = S-1$, we have*

$$d_{TV}(\mathcal{NH}(R,S,r), \mathcal{P}o(\lambda)) \le \frac{\left(1-e^{-\lambda}\right)S(S-1)}{(R+1)(R+2)}.$$

Remark 2.1. It is observed from Theorem 2.1 that

$$\begin{aligned}d_{TV}(\mathcal{NH}(R,S,r), \mathcal{P}o(\lambda)) &\le \left(1-e^{-\lambda}\right)\frac{(R+1)(r+1) - S(R-r+1)}{(R+1)(R+2)}\\ &< \frac{r}{R},\end{aligned}$$

that is, if $\frac{r}{R}$ is small, then (2.10) yields a good Poisson approximation, as mentioned previously.

3. Numerical examples

The following numerical examples are given to illustrate how well the Poisson distribution approximates the negative hypergeometric distribution and to see how tight is the upper bound for the total variation distance between two distributions given in Theorem 2.1.

Table 1 provides numerical examples of the total variation distance between negative hypergeometric and Poisson distributions and its upper bound for given $R = 100, 300, 600$, $S = 5, 10, 15, 20, 25, 30, 35, 40, 45$ and $r = 15, 30, 45$. The numerical examples in the table suggest that the Poisson approximation to the negative hypergeometric distribution is quite efficient provided that $\frac{r}{R}$ is small, that is, the estimate of the total variation distance between two distributions is close to the true value of the distance provided that r is small and close to S and R is large.

Table 1. Numerical examples of the total variation distance between negative hypergeometric and Poisson distributions and its upper bound.

R	S	r	λ	$d_{TV}(\mathcal{NH}, \mathcal{P}o)$	upper bound
100	5	15	0.74257	0.03107	0.06034
100	10	15	1.48515	0.01809	0.05676
100	15	15	2.22772	0.00727	0.02823
300	20	30	1.99336	0.00993	0.03716
300	25	30	2.49169	0.00674	0.02579
300	30	30	2.99003	0.00298	0.01255
600	35	45	2.62063	0.00547	0.02098
600	40	45	2.99501	0.00347	0.01419
600	45	45	3.36938	0.00170	0.00701

Acknowledgement. The author would like to thank the referees for their constructive suggestions that helped improve the presentation of the paper.

References

[1] A. D. Barbour, L. Holst and S. Janson, *Poisson Approximation*, Oxford Studies in Probability, 2, Oxford Univ. Press, New York, 1992.

[2] S. Boonta and K. Neammanee, Bounds on random infinite urn model, *Bull. Malays. Math. Sci. Soc. (2)* **30** (2007), no. 2, 121–128.

[3] T. Cacoullos and V. Papathanasiou, Characterizations of distributions by variance bounds, *Statist. Probab. Lett.* **7** (1989), no. 5, 351–356.

[4] L. H. Y. Chen, Poisson approximation for dependent trials, *Ann. Probability* **3** (1975), no. 3, 534–545.

[5] J. Cuzick, Event-based analysis times for randomised clinical trials, *Biometrika* **88** (2001), no. 1, 245–253.

[6] P. G. Graz, On the systematic and system-based study of grapheme frequencies: A re-analysis of German letter frequencies, *Glottometrics*, **15** (2007), 82–91.

[7] W. D. Kaigh and P. A. Lachenbruch, A generalized quantile estimator, *Comm. Statist. A—Theory Methods* **11** (1982), no. 19, 2217–2238.

[8] M. Majsnerowska, A note on Poisson approximation by w-functions, *Appl. Math. (Warsaw)* **25** (1998), no. 3, 387–392.

[9] T. Obilade, A weighted likelihood ratio of two related negative hypergeomeric distributions, *Acta Math. Appl. Sin. Engl. Ser.* **20** (2004), no. 4, 647–654.

[10] V. Papathanasiou and S. A. Utev, Integro-differential inequalities and the Poisson approximation, *Siberian Adv. Math.* **5** (1995), no. 1, 120–132.

[11] C. Stein, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, 583–602, Univ. California Press, Berkeley, CA.

[12] C. M. Stein, *Approximate Computation of Expectations*, IMS, Hayward California, 1986.

[13] X.-G. Wang, Phase properties of hypergeometric states and negative hypergeometric states, *J. Opt. B Quantum Semiclass. Opt.* **2** (2000), no. 1, 29–32.

[14] J. Wesołowski and M. Ahsanullah, Linearity of regression for non-adjacent weak records, *Statist. Sinica* **11** (2001), no. 1, 39–52.