

Total Variation Distance Between Poisson Distribution and Polya Distribution and It's Non-uniform Upper Bound

Bright Okore Osu^{1, *}, Samson Ogu-Egege², Emmanuel Inyang²

¹Department of Mathematics, Michael Okpara University of Agriculture, Umudike, Nigeria

²Department of Mathematics, Abia State University, Uturu, Nigeria

Abstract

The approximation of distribution due to Egege et al [1] is further extended to a wider approximation called total variation distances between Poisson and Pólya distribution, using Stein's Chen method and ω -function to give a non-uniform bound. In this work it was found that the results obtain for non-uniform are better than the results obtained for uniform bound in literature. Polya distribution approximate Poisson sufficiently enough provided r is close to n and N is large. If the upper bound is very small then a good approximation is obtained.

Keywords

Variation Distance, Poisson Distribution, Pólya Distribution, Stein Method, Non-uniform Upper Bound

Received: April 2, 2018 / Accepted: May 7, 2018 / Published online: June 7, 2018

© 2018 The Authors. Published by American Institute of Science. This Open Access article is under the CC BY license.

<http://creativecommons.org/licenses/by/4.0/>

1. Introduction

Pólya distribution was first studied and presented by G. Pólya 1923. Polyadistribution can be explained in a better way with the use of urn scheme. Suppose that we have an urn that initially contains a red and green balls denoted by r and b respectively, where r and b are positive integers. At each discrete time (*trial*), a ball is selected from the urn and then return the ballot the urn along with c , new balls of the same color. It is a discrete distribution that depend on four parameters (r, n, p, c) . Anon $-$ negative integer $-$ valued random variable X taking non-negative integer values k , $0 \leq k \leq n$. It's probability distribution function is of the form

$$P_X(k) = \binom{n}{k} \frac{(p;\gamma)_{k-1}(q;\gamma)_{n-k-1}}{(1;\gamma)_{n-1}}, \quad (1)$$

Where $n > 0, 0 < p < 1, q = 1 - p$ and $\gamma > 0$.

The form (1) can also be expressed as

$$\frac{\binom{p}{\gamma}^{k-1} \binom{q}{\gamma}^{n-k-1}}{\binom{1}{\gamma}^{n-1}}, k = 0, 1, 2 \dots n. \quad (2)$$

The mean and variance are given as np and $\frac{np(1+\gamma n)}{(1+\gamma)}$ respectively.

Equation (2) is equivalent to Egege, et al [1], if for a specialcases $c = 1, p = \frac{r}{r+b}$ and $\gamma = \frac{c}{r+b}$. From Equation (2) one gets;

$$P_X(k) = \binom{n}{k} \frac{[\frac{p}{\gamma} \dots \frac{p}{\gamma} + (k-1)] [\frac{q}{\gamma} \dots \frac{q}{\gamma} + (n-k-1)]}{[\frac{N}{\gamma} \dots \frac{N}{\gamma} + (n-1)]} = \binom{n}{k} \frac{[\frac{p}{\gamma} \dots \frac{p}{\gamma} + \tau] [\frac{q}{\gamma} \dots \frac{q}{\gamma} + \delta]}{[\frac{N}{\gamma} \dots \frac{N}{\gamma} + (n-1)]},$$

where $\tau = \begin{cases} 0 & \text{if } k = 0 \\ k - 1 & \text{if } k = 1, 2, \dots n \end{cases}$ and

* Corresponding author

E-mail address: osu.bright@mouau.edu.ng (B.O. Osu)

$$\delta = \begin{cases} 0 & \text{if } k = 0, 1, \dots, n \\ n - k - 1 & \text{if } k = 0 \end{cases}$$

$$= \binom{n}{k} \frac{N \left(\frac{r}{N} \dots \frac{r}{N} + \frac{\tau}{N^2\gamma} \right) \left[N \left(\frac{b}{N} \dots \frac{b}{N} + \frac{\delta}{N^2\gamma} \right) \right]}{N \left[\frac{1}{N^2\gamma} \dots 1 + \frac{(n-1)}{N^2\gamma} \right]}$$

$$= \binom{n}{k} \frac{N^k \left(\frac{r}{N} \dots \frac{r}{N} + \frac{\tau}{N^2\gamma} \right) \left[N^{n-k} \left(\frac{b}{N} \dots \frac{b}{N} + \frac{\delta}{N^2\gamma} \right) \right]}{N^n \left[\frac{1}{N^2\gamma} \dots 1 + \frac{n-1}{N^2\gamma} \right]}$$

Note $\frac{N^k \cdot N^{n-k}}{N^n} = N^{n-n} = 1$ and

$$P_X(k) = \binom{n}{k} \frac{\left[\frac{r}{N} \dots \frac{r}{N} + \frac{\tau}{N^2\gamma} \right] \left[\frac{b}{N} \dots \frac{b}{N} + \frac{\delta}{N^2\gamma} \right]}{\left[1 \dots 1 + \frac{n-1}{N^2\gamma} \right]} \tag{3}$$

If $b \rightarrow \infty, r \rightarrow \infty$ and $\gamma \rightarrow 0$ while $\frac{r}{N}$ remain fixed implies $\frac{b}{N}$ is also fixed then

$$P_X(k) = \binom{n}{k} \frac{\left[\frac{r}{N} \dots \frac{r}{N} + \frac{\tau}{N^2\gamma} \right] \left[\frac{b}{N} \dots \frac{b}{N} + \frac{\delta}{N^2\gamma} \right]}{\left[1 \dots 1 + \frac{n-1}{N^2\gamma} \right]} \rightarrow \binom{n}{k} \left(\frac{r}{N} \right)^n \left(\frac{b}{N} \right)^{n-k} \tag{4}$$

In view of (4) Teerapabolarn [2] used Stein’s Chen method [3] and ω -function associated with Polya random variable to give a bound for total variation distance between Binomial

$$d_{tv}(P_X(A), \wp_\lambda(A)) \leq \lambda^{-1}(\lambda + e^{-\lambda} - 1) \frac{r(N+N\gamma) - b(n-1)N\gamma}{N(N+N\gamma)} \text{ for } x_0 = 0, \tag{7}$$

and

$$d_{tv}(P_X(A), \wp_\lambda(A)) \leq \min \left\{ 1 - e^{-\lambda}, \frac{\lambda}{x_0} \right\} \frac{r(N+N\gamma) - b(n-1)N\gamma}{N(N+N\gamma)} \text{ for } x_0 > 0. \tag{8}$$

2. Method

The tools for giving the desired result consist of Stein’s Chen for Poisson distribution method and ω -function associated with Pólya random variable.

2.1. Stein-Chen Method

The Stein- Chen method is a general method in probability theory to obtain bounds on distance between two probability distributions with respect to a probability metric. This method was first introduced in [4] to obtain a bound between the distributions, a sum of m – dependent sequence of random variable and standard normal distribution in the Kolmogorov (uniform) metric. Stein’s Chen gave an equation for Poisson distribution with parameter $\lambda > 0$ and for a given h then is given as

$$h(x) + \wp_\lambda(x) = \lambda g(x + 1) - xg(x) \tag{9}$$

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

and Pólya of the form

$$d \left(B(n, p), P_\gamma(N, n, r, c) \right) \leq \frac{(1-p)^{n+1} - q^{n+1})c(n-1)n}{(n+1)(N+c)}, \tag{5}$$

Where N, n, r, c are non-negative valued integers, $B(n, p)$ and $P_\gamma(N, n, r, c)$ are binomial and Polya distribution respectively.

Also in view of (4) Egege, et al [1] use Poisson to approximate Polya distribution in terms of point metric to obtain a uniform upper bound given as

$$|P_X(x) - \wp_\lambda(x)| \leq (1 - e^{-\lambda}) \frac{r(r+b+c) - b(n-1)c}{(r+b)(r+b+c)}, \tag{6}$$

where $x \in \mathbb{N}, \wp_\lambda(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \lambda = \frac{nr}{r+b}$ and $(n - 1)c \leq r$.

It is observed that the upper bound in (6) can be seen to uniform bound with respect to x_0 and do not depend on x_0 . Hence the uniform upper bound in(6) may not be sufficiently good for measuring the accuracy of the approximation. To overcome these limitations, we proposed a non-uniform upper bound that depends on x_0 of the form.

For $A \subseteq \mathbb{N} \cup \{0\}$, define an indicator function $h_A(x) = \mathbb{N} \cup \{0\} \rightarrow \mathbb{R}$ such that

$$h_A(x) = \begin{cases} 1 & \text{if } x \in \Omega \\ 0 & \text{if } x \notin \Omega \end{cases} \tag{10}$$

for $A = \{x_0\}$ with $x_0 \in \mathbb{N} \cup \{0\}$. Letting $g = g_A$ and $h = h_A$. Barbour, et al [5] obtained a solution of (9) expressed in the form.

$$f_A(x) = \begin{cases} (x - 1)! \lambda^{-x} e^{\lambda} [\wp_\lambda(h_\Omega) \wp_\lambda(1 - h_{C_{x-1}})] & \text{if } x > A, \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Where $C_x = (0, 1, \dots, n)$.

Lemma 2.1: let $A \in \mathbb{N} \cup \{0\}, x \in \mathbb{N}$ defining a non-increasing function as

$\Delta g_A(x) = g_A(x + 1) - g_A(x) \forall x \in \mathbb{N}$ then following holds;

$$|\Delta g_A(x)| \leq \lambda^{-1} (1 - e^{-\lambda}). \tag{12}$$

Proof

For $A \subseteq \mathbb{N} \cup \{0\}$ and x_0 then $x > x_0$.

By definition of a function $\Delta g_A(x) = g_A(x+1) - g_A(x) \forall x \in \mathbb{N}$

$$|\Delta g_A(x)| = |g_A(x+1) - g_A(x)| \leq |g_A(x)|$$

And lemma(2.1) gives $|g_A(x)| = \lambda^{-1}(1 - e^{-\lambda})$. Thus

$$|\Delta g_A(x)| \leq \lambda^{-1}(1 - e^{-\lambda}). \quad (13)$$

Lemma 2.2: let $x_0 \in \mathbb{N} \cup \{0\}$ and $x \in \mathbb{N}$ then the following hold;

1. For $x_0 = 0$

$$|\Delta g_A(x)| \leq \lambda^{-2}(\lambda + e^{-\lambda} - 1). \quad (14)$$

2. For $x_0 > 0$

$$|\Delta g_A(x)| \leq \min\left\{\lambda^{-1}(1 - e^{-\lambda}), \frac{1}{x_0}\right\}. \quad (15)$$

Proof(15)

$$\begin{aligned} g_A(1) &= (x-1)! \lambda^{-x} e^{\lambda} [\wp_{\lambda} h_{(0)} \wp_{\lambda}(1 - h_{c_0})] \\ &= (x-1)! \lambda^{-x} e^{\lambda} [e^{-\lambda}(1 - e^{-\lambda})] \\ &\leq (x-1)! \lambda^{-x} e^{\lambda} [e^{-\lambda} - e^{-2\lambda}] \\ &\leq (x-1)! \lambda^{-1}(1 - e^{-\lambda}) \\ &= (x-1)! \lambda^{-1} \left[1 - 1 - \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} - \frac{\lambda^3}{3!} + \dots\right] \\ &= \frac{(x-1)!}{1!} \lambda^1 \left[\frac{-\lambda^1}{1!} + \frac{\lambda^2}{2!} - \frac{\lambda^3}{3!} + \dots\right]. \end{aligned}$$

If $x_0 = x > 0$, then

$$g_A(1) \leq \frac{\lambda^{-1}}{x_0}(1 - e^{-\lambda}).$$

Thus

$$|\Delta g_A(x)| = |g_A(x+1) - g_A(x)| \leq |g_A(1)| \forall x \in \mathbb{N}$$

$$|\Delta g_A(x)| \leq |g_A(1)| \leq \frac{\lambda^{-1}}{x_0}(1 - e^{-\lambda})$$

$$|\Delta g_A(x)| \leq \frac{\lambda^{-1}}{x_0}(1 - e^{-\lambda}). \quad (16)$$

Combining (13) and (16) gives

$$|\Delta g_A(x)| \leq \min\left\{\lambda^{-1}(1 - e^{-\lambda}), \frac{1}{x_0}\right\}. \quad (17)$$

2.2. ω - Function

ω - function had been studied and used by many authors. Teerepabolan [7, 8] and Wongkasem and Teerepabolan [9]

$$d_{tv}(P_X(A), \wp_{\lambda}(A)) \leq \lambda^{-2}(\lambda + e^{-\lambda} - 1)E|\lambda - \sigma^2\omega(x)| + |\lambda - \mu|E|g_A(X)|$$

2. For $x_0 > 0$

among others defined a ω -function associated with non-negative integer valued random variable in the relation

$$\omega(x) = \frac{1}{p(x)\sigma^2} \sum_{k=0}^x [\mu - h(x)]p(x), x \in S(X). \quad (18)$$

Lemma 2.3: let $P_X(k) > 0, \forall k \in \mathbb{N}$ there exist $w(x)$ such that

$$\frac{P(x-1)}{P(x)} = \frac{x(b+(n-x)N\gamma)}{(n-x+1)(r+(x-1)N\gamma)}. \quad (19)$$

Lemma 2.4: For $P_X(k) > 0$ there exist

$$\omega(x) = \omega(x-1) \frac{P(x-1)}{P(x)} + \frac{\mu - h(x)}{\sigma^2} \forall x \in \mathbb{N}. \quad (20)$$

Lemma 2.5: Let $\omega(x)$ be the ω - function associated with Pólya random variable X and $P_X(k) > 0$, for $0 \leq k \leq n$ then

$$\omega(x) = \frac{(n-x)(r+N\gamma x)}{N\sigma^2}, \quad (21)$$

where $\sigma^2 = \frac{npq(1+\gamma n)}{(1+\gamma)}$.

Proof

$$\omega(x) = \frac{(n-x)(r+N\gamma x)}{N\sigma^2}$$

and

$$\omega(x-1) = \frac{(n-x+1)(r+(x-1)N\gamma)}{N\sigma^2}. \quad (22)$$

Substituting (22) in Lemma 2.4 we obtained

$$\omega(x) = \frac{(r+N\gamma x)(n-x)}{N\sigma^2}.$$

Lemma 2.6: If a non-negative valued random variable X have $P_X(k) > 0$, for every x in the support of X , and finite variance $0 < \sigma^2 < \infty$ then

$$E[(X - \mu)g(x)] = \sigma^2 E[\omega(x)\Delta g(x)] < \infty. \quad (24)$$

And for any function $f: \mathbb{N} \cup \{0\} \rightarrow R$ for which $|\omega(x)\Delta g(x)| < \infty$, where $\Delta g(x) = g(x+1) - g(x)$ then $E[\omega(x)] = 1$

3. Main Result

Theorem 3.1: Let X be a Polya random variable with $\lambda = np$ and $(n-1)N\gamma \leq r$ and $c > 0 \forall x_0 \in \mathbb{N} \cup (0)$ then following holds

1. For $x_0 = 0$

$$d_{tv}(P_X(A), \wp_\lambda(A)) \leq \min \left\{ \lambda^{-1}(1 - e^{-\lambda}), \frac{1}{x_0} \right\} E|\lambda - \sigma^2 \omega(x)| + |\lambda - \mu| E|g_A(X)|$$

Proof

Substituting x by X from the right hand side (9) and taking expectation gives as in [10];

$$\begin{aligned} E[h_A(x) - \wp_\lambda(h)] &= E[\lambda g(X + 1) - Xg(X)] \\ d_{tv}(P_X(A), \wp_\lambda(A)) &= E\lambda[g(X + 1)] - E[Xg(X)] \\ &= \lambda E[g(X + 1) - g(X) + g(X)] - E[Xg(X)] \\ &= E[\Delta g(X) + g(X)] - E[Xg(X)] \\ &= \lambda E[\Delta g(X) + g(X)] - E[(X - \mu)g(X)] - E[\mu g(X)] \\ &= E[\lambda \Delta g(X) + \lambda g(X)] - E[(\lambda - \mu)g(X)] - E[(X - \mu)g(X)] \end{aligned} \tag{24}$$

From lemma 2.6

$$\begin{aligned} E[(X - \mu)g(X)] &= \sigma^2 E[\omega(X)\Delta g(X)] \\ &= \lambda E[\Delta g(X)] + E[\lambda - \mu]g(X) - \sigma^2 E[\omega(X)\Delta g(X)] \\ &= E[(\lambda - \sigma^2 \omega(X))\Delta g(X)] + E[(\lambda - \mu)]g(X) \\ |p_X(A) - \wp_\lambda(A)| &= E[(\lambda - \sigma^2 \omega(X))\Delta f(X)] + E[(\lambda - \mu)]g(X) \\ &\leq [|\lambda - \sigma^2 \omega(X)|\Delta g(X)] + |\lambda - \mu|E|g(X)| \\ &\leq \sup_{x \geq 1} |\Delta g(X)|E|\lambda - \sigma^2 \omega(X)| + |\lambda - \mu|E|g(X)| \end{aligned}$$

By lemma 2.2 and [10] the theorem is proved

Corollary 3.1

If $E|\lambda - \sigma^2 \omega(x)| \geq 0, \lambda - \omega(x) = np - \frac{np(1+\gamma n)}{(1+\gamma)}$ and $\lambda = \mu$

1. For $x_0 = 0$

$$|P_X(A) - \wp_\lambda(A)| \leq \lambda^{-1}(\lambda + e^{-\lambda} - 1) \frac{r(N+N\gamma) - b(n-1)N\gamma}{N(N+N\gamma)} \tag{25}$$

2. For $x_0 > 0$

$$|P_X(A) - \wp_\lambda(A)| \leq \min \left\{ \frac{\lambda}{x_0}, (1 - e^{-\lambda}) \right\} \frac{r(N+N\gamma) - b(n-1)N\gamma}{N(N+N\gamma)} \tag{26}$$

4. Numerical Result

Using the same numerical examples in Egege, et al [1] to illustrate how well Pólya can be approximated by Poisson in terms of total variation distance and its non-uniform upper bound with $c = 1$.

Table 1. Numerical examples of the total distances between Pólya distribution and Poisson distribution and its upper bound.

$x_0, r, n, N, \lambda,$	$P_X(A)$	$\wp_\lambda(A)$	d_{tv}	B_1 $A = \{x_0\} \in \{0\}$	B_2 $A = x_0 \in \mathbb{N}$	Uniformupper
0, 10, 5, 1000, 0.05	0.951085826	0.95122945	0.000143624	0.000148612	$\min \left\{ \begin{matrix} 0.000294767, \\ \frac{0.05}{x_0} (0.00643956) \end{matrix} \right\}$	0.000294767
1, 10, 5, 1000, 0.05	0.047841339	0.047561471	0.000279868		$\min \left\{ \begin{matrix} 0.000294767, \\ 0.000321978 \end{matrix} \right\}$	0.000294767

$x_0, r, n, N, \lambda,$	$P_X(A)$	$\varphi_\lambda(A)$	d_{tv}	B_1 $A = \{x_0\} \in \{0\}$	B_2 $A = x_0 \in \mathbb{N}$	Uniformupper
0, 10, 10, 1000, 0.1	0.904790652	0.904837418	0.000046766	0.000053123	$\min \left\{ \begin{matrix} 0.0000104574, \\ \frac{0.1}{x_0} (0.0001098901) \end{matrix} \right\}$	0.000104574
2, 10, 10, 1000, 0.1	0.004492181	0.004524187	0.000397623		$\min \left\{ \begin{matrix} 0.000104574, \\ 0.000054945 \end{matrix} \right\}$	0.000104574
0, 25, 20, 1000, 0.5	0.605592734	0.606530660	0.000937926	0.001378258	$\min \left\{ \begin{matrix} 0.002545169, \\ \frac{0.5}{x_0} (0.006493506) \end{matrix} \right\}$	0.002545169
1, 25, 20, 1000, 0.5	0.304624112	0.303265330	0.001358782		$\min \left\{ \begin{matrix} 0.002545169, \\ 0.0016233776 \end{matrix} \right\}$	0.002545169
2, 25, 20, 1000, 0.5	0.075772564	0.075816332	0.000043768		$\min \left\{ \begin{matrix} 0.002545169, \\ 0.0016233776 \end{matrix} \right\}$	0.002545169
0, 25, 25, 1000, 0.625	0.535059233	0.535261429	0.000202196	0.0004162264	$\min \left\{ \begin{matrix} 0.000754446, \\ \frac{0.626}{x_0} (0.001623377) \end{matrix} \right\}$	0.000754446

Where $B_1 = \lambda^{-1}(\lambda + e^{-\lambda} - 1) \frac{r(N+N\gamma) - b(n-1)N\gamma}{N(N+N\gamma)}$ for $x_0 = 0$

$B_2 = \min \left\{ 1 - e^{-\lambda}, \frac{\lambda}{x_0} \right\} \frac{r(N+N\gamma) - b(n-1)N\gamma}{N(N+N\gamma)}$ for $x_0 \in \mathbb{N}$

Corollary 4.1

1. $\min \left\{ (1 - e^{-\lambda}), \frac{\lambda}{x_0} \right\} < \lambda^{-1}(\lambda + e^{-\lambda} - 1) < (1 - e^{-\lambda})$ for $x_0 = 0$
2. $\lambda^{-2}(\lambda + e^{-\lambda} - 1) < \min \left\{ \lambda^{-1}(1 - e^{-\lambda}), \frac{1}{x_0} \right\} \leq \lambda^{-1}(1 - e^{-\lambda})$ for $x_0 > 0$.

5. Discussion

In view of table 1, when n is close to r and very small a good approximation is obtained. The non-uniform upper bound used in this work is a criterion use for measuring the accuracy of approximation of Pólya by Poisson. It was also observed that the result obtained for non-uniform bound is better than uniform bound in [1], which is in agreement with corollary 4.1. If the upper bound is very small then a good approximation is obtained. And the total variation distance will be close the upper bounds. This is in agreement with [1]. To obtain a small upper bounds and a total distance variation close to the bounds the following must exist

1. $n \leq r$
2. $\frac{r}{N}$ small and $N\gamma = 1$
3. N Sufficiently large.

6. Conclusion

In this work, a wider approximation called total variation distances between Poisson and Pólya distribution was considered. Stein’s Chen method and ω – function were applied to give a non-uniform bound. It was found that non-uniform upper bound approximation of Pólya distribution by

Poisson distribution is better than uniform upper bound provided r is close to n and N is large.

References

- [1] Samson O Egege, Bright O. Osu and Emmanuel J. Ekepenyoung (2017) A Stein –Chen upper bound approximation of Pólya distribution Via Poisson Distribution with the ω -function. Journal of the Nigerian association of mathematical Physics. vol 42 pp 41-46.
- [2] Teerapabolan K. (2012) Binomial approximation to the Polya distribution, International journal of pure and applied mathematics pp 635-640.
- [3] Stein C. M (1972) A bound for the error in normal approximation to the distribution of a sum of dependent random variable, Proc, sixth Berkeley Sympos. maths statist probvol 3, 583-602.
- [4] Muhammad I., Syed. I A. S and Muhammad A (2011), Polya, Binomial and Hypergeometric distribution comparison regarding skewness and kurtosis, Middle East journal of scientific research No 8 vol 6 pp 1027-1028.
- [5] Barbour A. B, Host. I and Janson, S. (1992) Poisson approximation, Oxford Studies in Probability 2, Clarendon press Oxford.
- [6] Cacoullos T. and Papathanasious (1989) Characterization of distribution by variance bounds, Statistics probability Lett 7 pp 351-356.
- [7] Teerapabolan K (2012) A point wise approximation of generalized Binomial by Poisson distribution, Applied mathematics science vol 6, no 22, pp 1095-1104.
- [8] Teerapabolan K (2013) Stein’s identity for discrete distribution, International Journal of pure and applied mathematics 83 (4): 565-573.
- [9] Wongkasem P. and Teerapabolan K. (2008) On approximating a generalized Binomial by Binomial and Poisson distribution, international journal of statistics and system, 3 (2): 113-124.
- [10] Feller, W. (1968) An introduction to probability theory and its application, 3rd edition, Wiley, New York, pp 34-40.