ON BINOMIAL APPROXIMATION FOR INDEPENDENT BERNOULLI RANDOM SUMS

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

Abstract: A non-uniform bound for the distance between the distribution function of random sums of independent Bernoulli random variables and an appropriate binomial distribution function could be obtained. Two examples have been given to illustrate the result obtained.

AMS Subject Classification: 62E17, 60F05, 60G05
Key Words: Bernoulli random variable, binomial approximation, distribution function, random sums

1. Introduction

Let $X_1, X_2, \ldots$ be a sequence of independent Bernoulli random variables, each with probability of success $p_i = P(X_i = 1) = 1 - P(X_i = 0)$, and let $N$ be a non-negative integer-valued random variable and independent of the $X_i$'s. Thus $S_N = \sum_{i=1}^{N} X_i$ is the random sums of $N$ independent Bernoulli random variables. In the case of $N = n \in \mathbb{N}$ is fixed, there have been some research on topic related to the binomial approximation to the distribution of the sum $S_n$, which can be found in [2], [3] and [5]. Let $\lambda_N = \sum_{i=1}^{N} p_i$, $\hat{n} = E(N) \in \mathbb{N}$ and $\hat{p} = 1 - \hat{q} = \frac{\lambda}{\hat{n}}$, where $\lambda = E(\lambda_N)$. For $x \in \{0, 1, \ldots, n\}$, let $S_n(x) = P(S_n \leq x)$ and $B_{n,p}(x) = P(B_{n,p} \leq x)$, where $B_{n,p}$ is the binomial random variable with parameters $n$ and $p$. In this paper, we are interested to give a non-uniform bound for the distance between $S_N(x)$ and $B_{\hat{n}, \hat{p}}(x)$, $|S_N(x) - B_{\hat{n}, \hat{p}}(x)|$, which
is in Section 2. In Section 3, two examples have been given to illustrate the desired result. The conclusion of this study is presented in the last section.

2. Result

The following theorem presents a non-uniform bound for the distance between the distribution functions of $S_N$ and $B_{\hat{n}, \hat{\rho}}$, which is the desired result.

**Theorem 2.1.** For $\hat{n} \in \mathbb{N}$, $x \in \{1, \ldots, \hat{n}\}$ and $\lambda = E(\lambda_N)$, then

\[
|S_N(x) - B_{\hat{n}, \hat{\rho}}(x)| \leq \min \left\{ E \left( \frac{1 - e^{-\lambda N}}{\lambda N} \sum_{i=1}^{N} p_i^2 \right), \frac{1}{x} E \left( \sum_{i=1}^{N} p_i^2 \right) \right\} + \min \left\{ \frac{1 - e^{-\lambda}}{\lambda}, \frac{1}{x} \right\} \left( \text{Var}(\lambda_N) + \frac{\lambda^2}{\hat{n}} \right),
\]

(1)

where $S_N(0) = E(\prod_{i=1}^{N} q_i)$.

**Proof.** Let $P_\lambda(x)$ be the Poisson distribution function with mean $\lambda$ at $x$. It follows the fact that

\[
|S_N(x) - B_{\hat{n}, \hat{\rho}}(x)| \leq |S_N(x) - P_{\lambda_N}(x)| + |P_{\lambda_N}(x) - P_\lambda(x)| + |P_\lambda(x) - B_{\hat{n}, \hat{\rho}}(x)|.
\]

(2)

Following [6], we have

\[
|S_N(x) - P_{\lambda_N}(x)| \leq \min \left\{ E \left( \frac{1 - e^{-\lambda N}}{\lambda N} \sum_{i=1}^{N} p_i^2 \right), \frac{1}{x} E \left( \sum_{i=1}^{N} p_i^2 \right) \right\}.
\]

(3)

Applying Theorem 1.C in [1] together with Lemma 2.1 in [5] and using Corollary 4.1 in [5], we also obtain

\[
|P_{\lambda_N}(x) - P_\lambda(x)| \leq \min \left\{ \frac{1 - e^{-\lambda}}{\lambda}, \frac{1}{x} \right\} \text{Var}(\lambda_N)
\]

(4)

and

\[
|P_\lambda(x) - B_{\hat{n}, \hat{\rho}}(x)| \leq \min \left\{ \frac{1 - e^{-\lambda}}{\lambda}, \frac{1}{x} \right\} \hat{n} \hat{\rho}^2 = \min \left\{ \frac{1 - e^{-\lambda}}{\lambda}, \frac{1}{x} \right\} \frac{\lambda^2}{\hat{n}},
\]

(5)
respectively. Hence, the inequality (1) is obtained by taking the bounds in (3), (4) and (5) into (2).

If $X_i$'s are identically distributed, then the following corollary is an immediately consequence of the Theorem 2.1.

**Corollary 2.1.** For $\tilde{n} \in \mathbb{N}$ and $x \in \{1, ..., \tilde{n}\}$, if $p_1 = p_2 = \cdots = p$, then we have the following:

$$
|S_N(x) - B_{\tilde{n},p}(x)| \leq \min \left\{ E \left( 1 - e^{-Np} \right), \frac{\tilde{n}p}{x} \right\} p + \min \left\{ \frac{1 - e^{-\tilde{n}p}}{\tilde{n}p}, \frac{1}{x} \right\} (Var(N) + \tilde{n})p^2,
$$

where $S_N(0) = E(q^N)$.

3. Examples

Two examples are given to illustrate the result in the case of $X_i$'s are identically distributed.

**Example 3.1.** For $n$ ($n \in \mathbb{N}$) is fixed, let $N$ be a positive integer-valued random variable with probability function

$$
P(N = k) = \begin{cases} 
\frac{1}{2} & , k = 2n, \\
\frac{1}{2} & , k = 4n, \\
0 & , \text{otherwise}.
\end{cases}
$$

Therefore $\tilde{n} = 3n$, $Var(N) = n^2$ and $E \left( 1 - e^{-Np} \right) = 1 - \frac{e^{-2np} + e^{-4np}}{2}$. Let $p_1 = p_2 = \cdots = p$, then we have

$$
|S_N(x) - B_{3n,p}(x)| \leq \min \left\{ 1 - \frac{e^{-2np} + e^{-4np}}{2}, \frac{3np}{x} \right\} p + \min \left\{ \frac{1 - e^{-3np}}{3np}, \frac{1}{x} \right\} (n + 3)p^2
$$

$x \in \{1, ..., 3n\}$.

**Example 3.2.** Let $N$ be a positive integer-valued random variable with probability function

$$
P(N = n) = \frac{1}{20}, \ n = 1, 2, ..., 21,$$
then we have \( \hat{n} = 11 \) and \( \text{Var}(N) = \frac{891}{20} \). Therefore, if \( p_1 = p_2 = \cdots = p \), then we obtain

\[
\left| S_N(x) - B_{11,p}(x) \right| \leq \min \left\{ 1, \frac{11p}{x} \right\} p + \min \left\{ \frac{1 - e^{-11p}}{11p}, \frac{1}{x} \right\} 55.55p^2,
\]

\( x \in \{1, ..., 11\} \).

4. Conclusion

In this study, a non-uniform bound for the distance between the distribution function of random sums of independent Bernoulli random variables and an appropriate binomial distribution function with parameters \( \hat{n} \) and \( \hat{p} \) could be obtained. The bound points out that the distribution function of random sums of independent Bernoulli random variables can be approximated by the binomial distribution function with parameters \( \hat{n} \) and \( \hat{p} \) when \( \hat{p} \) is small, that is, the desired result gives a good binomial approximation when \( \hat{p} \) is small.

References


