Submitted to the Brazilian Journal of Probability and Statistics

Approximations Related to the Sums of $m$-dependent Random Variables

Amit N. Kumar$^a$, Neelshed S. Upadhye$^b$ and P. Vellaisamy$^c$

$^a$Department of Mathematical Sciences,
Indian Institute of Technology (BHU),
Varanasi-221005, India.
$^b$Department of Mathematics,
Indian Institute of Technology Madras,
Chennai-600036, India.
$^c$Department of Mathematics,
Indian Institute of Technology Bombay,
Powai, Mumbai-400076, India.

Abstract. In this paper, we mainly focus on the sums of non-negative integer-valued 1-dependent random variables and its approximation to the power series distribution. We first discuss some relevant results for power series distribution such as the Stein operator, uniform and non-uniform bounds on the solution of the Stein equation. Using Stein’s method, we obtain error bounds for the approximation problem considered. The obtained results can also be applied to the sums of $m$-dependent random variables via appropriate rearrangements of random variables. As special cases, we discuss two applications, namely, 2-runs and $(k_1, k_2)$-runs, and compare our bounds with existing bounds.

1 Introduction and Preliminaries

The sum of $m$-dependent random variables (rvs) has special attention due to its applicability in many real-life applications such as runs and patterns, DNA sequences, and reliability theory, among many others. However, its distribution is difficult or sometimes intractable, especially if the setup is arising from non-identical rvs concentrated on $\mathbb{Z}_+ = \{0, 1, 2, \ldots \}$, the set of non-negative integers. Therefore, there is a need to approximate such a distribution with some known and easy-to-use distributions. In this article, we mainly focus on power series distribution (PSD) approximation to the sums of 1-dependent rvs, however, the approximation results are also useful for the sums of $m$-dependent rvs. We consider the PSD family that satisfies Panjer’s recursive relation with support $\{0, 1, \ldots, N\}$. $N \in \mathbb{Z}_+ \cup \{\infty\}$. Of course, the bound can directly apply for special distributions of the PSD family. An advantage of approximation to the PSD family is that we can obtain the error bounds for approximation to some specific distributions such as Poisson and negative binomial distributions. For some related works, we refer the reader to Lin and Liu (2012), Čekanavičius and Vellaisamy (2015), Barbour and Xia (1999), Brown and Xia (2001), Fu and Johnson (2009), Vellaisamy (2004), Wang and Xia (2008), and Soon (1996), Kumar, Vellaisamy and Viens (2021), and the reference therein.

A sequence of rvs $\{Y_k\}_{k \geq 1}$ is called $m$-dependent if, for $j - i > m$, $\sigma(Y_1, Y_2, \ldots, Y_j)$ and $\sigma(Y_j, Y_{j+1}, \ldots)$ are independent, where $\sigma(Y)$ denotes the sigma-algebra generated by $Y$. Let $S_n = \sum_{i=1}^{n} Y_i$ be the sum of $m$-dependent rvs. Then, grouping the consecutive summations in the following form

$$X_i := \sum_{j=(i-1)m+1}^{\min(im,n)} Y_j, \quad i = 1, 2, \ldots, \left\lceil n/m \right\rceil,$$

Keywords and phrases. Power series distribution, $m$-dependent random variables, Stein’s method, Runs.
where \([x]\) is the smallest integer greater than or equal to \(x\), we can reduce the sum of \(m\)-dependent rvs to the sum of 1-dependent rvs, that is,

\[
S_n = \sum_{i=1}^{n} Y_i = \sum_{i=1}^{\lfloor n/m \rfloor} X_i,
\]

where \(Y_i\)'s are \(m\)-dependent and \(X_i\)'s are 1-dependent rvs. Therefore, the approximation results for the sum of 1-dependent rvs can also be applied to the sum of \(m\)-dependent rvs (see Subsection 4.2 for an example). Note that, for the fixed value of \(m\), an individual study of the random variable (rv) \(X_i\) is needed.

Hereafter, \(1_A\) denotes the indicator function of \(A \subseteq \mathbb{Z}_+\). Let \(X\) be a rv concentrated on \(\mathbb{Z}_+\),

\[
\mathcal{G} = \{ f : \mathbb{Z}_+ \to \mathbb{R} \mid f \text{ is bounded} \},
\]

and

\[
\mathcal{G}_X = \{ g \in \mathcal{G} \mid g(0) = 0 \text{ and } g(x) = 0 \text{ for } x \not\in \mathcal{S}(X) \}, \tag{1.1}
\]

associated with the rv \(X\), where \(\mathcal{S}(X)\) denotes the support of the rv \(X\). We now briefly discuss Stein’s method (Stein (1972)) which we use to derive our approximation results in Section 3. Stein’s method can be discussed in the following three steps.

(a) Identify a Stein operator, denoted by \(A_X\) for a rv \(X\), such that \(\mathbb{E}(A_X g(X)) = 0\), for \(g \in \mathcal{G}_X\).
(b) Solve the Stein equation \(A_X g(k) = f(k) - \mathbb{E}(f(X))\), for \(f \in \mathcal{G}\) and \(g \in \mathcal{G}_X\).
(c) Replace \(k\) by a rv \(Y\) in Stein equation, and taking expectation and supremum to get

\[
d_{TV}(X, Y) := \sup_{f \in \mathcal{H}} |\mathbb{E}(f(X)) - \mathbb{E}(f(Y))| = \sup_{f \in \mathcal{H}} |\mathbb{E}(A_X g_f(Y))|,
\]

where \(g_f\) is the solution of the Stein equation and \(\mathcal{H} = \{ 1_A \mid A \subseteq \mathbb{Z}_+ \} \). For \(\mathbb{Z}_+\)-valued rvs \(X\) and \(Y\), the total variation distance is equivalent to

\[
d_{TV}(X, Y) = \frac{1}{2} \sum_{k=0}^{\infty} |\mathbb{P}(X = k) - \mathbb{P}(Y = k)|. \tag{1.2}
\]

For additional details on Stein’s method, see Barbour, Holst and Janson (1992), Barbour and Chen (2014), Ley, Reinert and Swan (2017), Reinert (2005), Upadhye, Čekanavičius and Vellaisamy (2017), and the references therein.

This article is organized as follows. In Section 2, we discuss the PSD and its related results to Stein’s method. In Section 3, we derive the error bound for PSD approximation to the sums of 1-dependent rvs and discuss some relevant remarks. In Section 4, we discuss two important applications of our results to the sums of 2-runs and \((k_1, k_2)\)-runs.

## 2 Power Series Distribution and Related Results

Let \(Z\) be a rv. We say its distribution belongs to the PSD family, denoted by \(\mathcal{P}\), if \(\mathbb{P}(Z = k) = p_k\) is of the form

\[
p_k = \frac{a_k \theta^k}{\gamma(\theta)}, \quad k \in \mathcal{S}(Z) = \{0, 1, \ldots, N\}, \tag{2.1}
\]

where \(\theta \in (0, r)\), with \(r > 0\), and \(a_k \geq 0\), are called series parameter and coefficient function, respectively, \(N \in \mathbb{Z}_+ \cup \{\infty\}\), and \(\gamma(\theta) = \sum_{k=0}^{N} a_k \theta^k\). Note that \(a_k \neq 0\) for \(k \in \mathcal{S}(Z)\) and \(r\) is the radius of convergence of the series \(\gamma(\theta) = \sum_{k=0}^{N} a_k \theta^k\). Many distributions such
as Poisson, binomial, negative binomial, logarithmic series, and inverse sine distributions, among many others, belong to the PSD family. For more details, we refer the reader to Edwin (2014), Noack (1950), Patil (1962), Kumar, Vellaisamy and Viens (2021), and the references therein.

Next, we give a brief discussion about Stein’s method for PSD, in fact, many results follow from Eichelsbacher and Reinert (2008). The following proposition gives a Stein operator for PSD.

**Proposition 2.1.** Let the rv $Z$ having distribution belonging to the PSD family defined in (2.1) with $E(Z) < \infty$. Then a Stein operator for $Z$ is given by

$$A_Z g(k) = \theta(k + 1) \frac{a_{k+1}}{a_k} g(k + 1) - k g(k), \quad g \in \mathcal{G}_Z, \quad k \in \{0, 1, \ldots, N\}. \quad (2.2)$$

**Proof.** From (2.1), it can be easily verified that

$$\theta(k + 1) \frac{a_{k+1}}{a_k} p_k - (k + 1) p_{k+1} = 0, \quad \text{for} \quad k \in \{0, 1, \ldots, N\}. \quad (2.3)$$

Let $g \in \mathcal{G}_Z$ defined in (1.1), then

$$\sum_{k=0}^{N} g(k + 1) \left[ \theta(k + 1) \frac{a_{k+1}}{a_k} p_k - (k + 1) p_{k+1} \right] = 0.$$ 

The above expression leads to

$$\sum_{k=0}^{N} \left[ \theta(k + 1) \frac{a_{k+1}}{a_k} g(k + 1) - k g(k) \right] p_k = 0.$$ 

This proves the result. \qed

Next, we discuss the solution of the Stein equation

$$\theta(k + 1) \frac{a_{k+1}}{a_k} g(k + 1) - k g(k) = f(k) - E(f(Z)), \quad f \in \mathcal{G}, \quad g \in \mathcal{G}_Z, \quad k \in \{0, 1, \ldots, N\}. \quad (2.4)$$

We first describe discrete Gibbs measure (DGM), a large class of distributions, studied by Eichelsbacher and Reinert (2008). If a rv $U$ has the distribution of the form

$$P(U = k) = \frac{1}{C_w} e^{V(k)} \frac{w^k}{k!}, \quad k \in \{0, 1, \ldots, N\}, \quad (2.5)$$

for some function $V : \mathbb{Z}_+ \to \mathbb{R}$, $w > 0$, and $C_w = \sum_{k=0}^{\infty} e^{V(k)} \frac{w^k}{k!}$, then we say the rv $U$ belongs to the DGM family. Note that if we take $a_k = e^{V(k)} / k! \iff V(k) = \ln(a_k k!)$, $\theta = w$, and $\gamma(\theta) = C_w$, which are valid choices, then the results derived by Eichelsbacher and Reinert (2008) are valid for the PSD family. Therefore, the solution of (2.4) can be directly obtained from (2.5) and (2.6) of Eichelsbacher and Reinert (2008) and is given by

$$g(k) = \frac{1}{k a_k \theta^k} \sum_{j=0}^{k-1} a_j \theta^j [f(j) - E(f(Z))]$$

$$= -\frac{1}{k a_k \theta^k} \sum_{j=k}^{N} a_j \theta^j [f(j) - E(f(Z))].$$

Also, the Lemma 2.1 of Eichelsbacher and Reinert (2008) can be written for the PSD family in the following manner.
Lemma 2.1. Let $G_1 = \{ f : \mathbb{Z}_+ \to [0, 1] \}$, $F(k) = \sum_{i=0}^{k} p_i$ and $\bar{F}(k) = \sum_{i=k}^{N} p_i$. Assume that

$$\frac{kF(k)}{F(k-1)} \geq \theta(k+1) \frac{a_{k+1}}{a_k} \geq k \frac{\bar{F}(k+1)}{\bar{F}(k)}.$$  

Then, for $f \in G_1$ and $g_f$, the solution of (2.4), we have

$$\sup_{f \in G_1} |\Delta g_f(k)| = \frac{a_k}{\theta(k+1)a_{k+1}} \bar{F}(k+1) + \frac{1}{k} F(k-1),$$

where $\Delta g_f(k) = g_f(k+1) - g_f(k)$. Moreover,

$$\sup_{f \in G_1} |\Delta g_f(k)| \leq \frac{1}{k} \wedge \frac{a_k}{\theta(k+1)a_{k+1}},$$

(2.6)

where $x \wedge y$ denotes the minimum of $x$ and $y$, and $k \in \{0,1,\ldots,N\}$.  

Next, it is not easy to use the direct form of the Stein operator (2.2) as $a_k$ is unknown and depends on $k$. So, we consider the PSD family with Panjer’s recursive relation (see Panjer and Wang (1995), Sundt and Jewell (1981), and Hess, Liewald and Schmidt (2002) for details), denoted by $P_1$, which is given by

$$(k+1)\frac{p_{k+1}}{p_k} = a \theta k \Rightarrow \theta(k+1) \frac{a_{k+1}}{a_k} = a \theta k, \text{ for some } a, b \in \mathbb{R} \text{ and } k \in \{0,1,\ldots,N\}.$$  

(2.7)

Therefore, the Stein operator given in (2.2) can be written as

$$A_f g(k) = (a + bk)g(k+1) - kg(k), \quad k \in \{0,1,\ldots,N\}.$$  

(2.8)

Also, the bound given in (2.6) becomes

$$\sup_{f \in G_1} |\Delta g_f(k)| \leq \frac{1}{k} \wedge \frac{1}{a + bk}, \quad k \in \{0,1,\ldots,N\}.$$  

(2.9)

Note that if $a, b \geq 0$ (PSD family satisfies Panjer recursive relation with $a, b \geq 0$, denoted by $P_2$) then the bound given in (2.9) becomes uniform and is given by

$$\sup_{f \in G_1} |\Delta g_f(k)| \leq 1 \wedge \frac{1}{a}, \quad k \in \{1,\ldots,N\}.$$  

(2.10)

Note that $P_2 \subset P_1 \subset P$. Also, observe that $a = \lambda, b = 0$ ($a_k = 1/k!$, $\theta = \lambda$ and $\gamma(\theta) = e^\theta$) and $a = nq, b = q$ ($a_k = \binom{n+k-1}{k}$), $\theta = q$ and $\gamma(\theta) = (1-\theta)^{-n}$ for Poisson (with parameter $\lambda$) and negative binomial (with parameter $n$ and $p = 1 - q$) distributions, respectively, and hence the bounds (from (2.10)) are $1 \wedge \frac{1}{\lambda}$ and $1 \wedge \frac{1}{nq}$, respectively, which are well-known bounds for Poisson and negative binomial distributions. Many distributions satisfy the condition $a, b \geq 0$. However, if the condition is not satisfied, one can still use (2.9) to compute the uniform bound. For example, if $a_k = \binom{n}{k}, \theta = p/q$, and $\gamma(\theta) = (1+\theta)^n$, then $Z \sim Bi(n,p)$, and $\frac{\theta(k+1)a_{k+1}}{a_k} = \frac{q}{p}(n-k)$, and hence $a = np/q$ and $b = -p/q \leq 0$. Therefore, the bound given in (2.9) is

$$\sup_{f \in G_1} |\Delta g_f(k)| \leq \frac{1}{k} \wedge \frac{q}{p(n-k)}$$

$$= \begin{cases} \frac{1}{k} \frac{q}{p(n-k)} & \text{if } k \geq np \\ \frac{q}{(n-k)p} & \text{if } k \leq np \end{cases}$$
\[
  \begin{cases}
  \frac{1}{np} & \text{if } k \geq np \\
  \frac{1}{np} & \text{if } k \leq np
  \end{cases}
\]

which leads to a uniform bound for the binomial distribution. Note here that the Stein operator (from (2.8)) is

\[
  A_Z g(k) = p \frac{n - k}{q} g(k + 1) - kg(k), \quad k \in \{0, 1, \ldots, n\}.
\]

But, the well-known Stein operator for the binomial distribution is

\[
  A_Z g(k) = p \frac{n - k}{q} g(k + 1) - kg(k), \quad k \in \{0, 1, \ldots, n\},
\]

which follows by multiplying \(q\) in (2.12) (see (5) of Upadhye, Čekanavičius and Vellaisamy (2017)). Also, the uniform bound will be changed and is given by \(\frac{1}{npq}\) (that is, divided by \(q\)), which is a well-known bound concerning the Stein operator (2.13) (see (34) of Upadhye, Čekanavičius and Vellaisamy (2017)). Hence, throughout this article, we use \(\|\Delta g\| = \sup_k |\Delta g(k)|\) and the uniform bound for \(\|\Delta g\|\) can be obtained from (2.10) or may be computed explicitly for some applications.

Next, let \(\phi_Z(\cdot)\) be the probability generating function of \(Z\). Then, using (2.7), it can be seen that

\[
  \phi'_Z(t) = a \phi_Z(t) \frac{1}{1 - bt}.
\]

Hence, the mean and variance of the PSD are given by

\[
  \mathbb{E}(Z) = \frac{a}{1 - b} \quad \text{and} \quad \text{Var}(Z) = \frac{a}{(1 - b)^2}, \quad b \neq 1.
\]

For more details, we refer the reader to Edwin (2014), and Panjer and Wang (1995).

### 3 Approximation Results

In this section, we derive an error bound for PSD approximation to the sums of 1-dependent rvs in total variation distance and discuss some relevant remarks. Also, we discuss approximation results for Poisson and negative binomial distribution as special cases. Throughout this section, we assume \(X_1, X_2, \ldots, X_n, n \geq 1\), is a sequence of 1-dependent rvs and

\[
  W_n = \sum_{i=1}^{n} X_i.
\]

For a \(\mathbb{Z}_+\)-valued rv \(Y\), let \(D(Y) := 2d_{TV}(Y, Y + 1)\), where \(d_{TV}(X, Y)\) as defined in (1.2). Also, let

\[
  N_{i, \ell} := \{j : |j - i| \leq \ell\} \cap \{1, 2, \ldots, n\} \quad \text{and} \quad X_{N_{i, \ell}} := \sum_{j \in N_{i, \ell}} X_j, \quad \text{for } \ell = 1, 2.
\]

Note that \(X_{N_{i, 2}} - X_{N_{i, 1}} = X_{N_{i, 2} - N_{i, 1}}\). From (3.1), it can be verified that \(\mathbb{E}(W_n) = \sum_{i=1}^{n} \mathbb{E}(X_i)\) and

\[
  \text{Var}(W_n) = \sum_{i=1}^{n} \sum_{|j - i| \leq 1} \left[\mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)\right] = \sum_{i=1}^{n} \left[\mathbb{E}(X_i X_{N_{i, 1}}) - \mathbb{E}(X_i)\mathbb{E}(X_{N_{i, 1}})\right].
\]

Now, the following theorem gives the error bound for \(Z\)-approximation to \(W_n\).
Theorem 3.1. Let $Z \in \mathcal{P}_1$ and $W_n$ be defined as in (3.1). Assume that $\mathbb{E}(Z) = \mathbb{E}(W_n)$, and $\tau = \text{Var}(W_n) - \text{Var}(Z)$. Then, for $n \geq 6$,

$$d_{TV}(W_n, Z) \leq \|\Delta g\| \left\{ \frac{|1-b|}{2} \left[ \sum_{i=1}^{n} \mathbb{E}(X_i)\mathbb{E}(X_{\tau,i} (2X_{\tau,i} - X_{\tau,i} - 1)D(W_n|X_{\tau,i}, X_{\tau,i}))) 
+ \sum_{i=1}^{n} \mathbb{E}(X_iX_{\tau,i} (2X_{\tau,i} - X_{\tau,i} - 1)D(W_n|X_{\tau,i}, X_{\tau,i})) \right] 
+ \sum_{i=1}^{n} \left[ (1-b) \{ \mathbb{E}(X_i)\mathbb{E}(X_{\tau,i}) - \mathbb{E}(X_iX_{\tau,i}) \} + \mathbb{E}(X_i)\mathbb{E}((X_{\tau,i} - 1)D(W_n|X_{\tau,i})) \right] 
+ \sum_{i=1}^{n} \mathbb{E}(X_i(X_{\tau,i} - 1)D(W_n|N_i, 2)) + |\tau(1-b)| \right\}. \quad (3.3)$$

Proof. Consider the Stein operator given in (2.8) and taking expectation with respect to $W_n$, we have

$$\mathbb{E}(A_Z g(W_n)) = a\mathbb{E}(g(W_n + 1)) + b\mathbb{E}(W_n g(W_n + 1)) - \mathbb{E}(W_n g(W_n)))$$

$$= a\mathbb{E}(g(W_n + 1)) - (1-b)\mathbb{E}(W_n g(W_n + 1)) + \mathbb{E}(W_n \Delta g(W_n))$$

$$= (1-b) \left[ \frac{a}{(1-b)} \mathbb{E}(g(W_n + 1)) - \mathbb{E}(W_n g(W_n + 1)) \right] + \mathbb{E}(W_n \Delta g(W_n)).$$

Applying the first moment matching condition, $\mathbb{E}(Z) = a/(1-b) = \mathbb{E}(W_n)$, we get

$$\mathbb{E}(A_Z g(W_n)) = (1-b) \left[ \mathbb{E}(W_n)\mathbb{E}(g(W_n + 1)) - \mathbb{E}(W_n g(W_n + 1)) \right] + \mathbb{E}(W_n \Delta g(W_n)). \quad (3.4)$$

Let now

$$W_{i,n} := W_n - X_{\tau,i},$$

so that $X_i$ and $W_{i,n}$ are independent. Consider the following expression from (3.4)

$$\mathbb{E}(W_n)\mathbb{E}(g(W_n + 1)) - \mathbb{E}(W_n g(W_n + 1)) = \sum_{i=1}^{n} \mathbb{E}(X_i)\mathbb{E}(g(W_n + 1)) - \sum_{i=1}^{n} \mathbb{E}(X_i g(W_n + 1))$$

$$= \sum_{i=1}^{n} \mathbb{E}(X_i)\mathbb{E}(g(W_n + 1)) - \sum_{i=1}^{n} \mathbb{E}(X_i g(W_n + 1))$$

$$- \sum_{i=1}^{n} \mathbb{E}(X_i g(W_{i,n} + 1)) + \sum_{i=1}^{n} \mathbb{E}(X_i g(W_{i,n} + 1))$$

$$= \sum_{i=1}^{n} \mathbb{E}(X_i)\mathbb{E}(g(W_n + 1) - g(W_{i,n} + 1))$$

$$- \sum_{i=1}^{n} \mathbb{E}(X_i(g(W_n + 1) - g(W_{i,n} + 1))). \quad (3.5)$$
It can be seen that
\[
g(W_n + 1) - g(W_{i,n} + 1) = g(W_{i,n} + X_{N_{i,1}} + 1) - g(W_{i,n} + 1) \\
= g(W_{i,n} + X_{N_{i,1}} + 1) - g(W_{i,n} + X_{N_{i,1}}) \\
+ g(W_{i,n} + X_{N_{i,1}}) - g(W_{i,n} + X_{N_{i,1}} - 1) \\
+ \cdots \\
+ g(W_{i,n} + 2) - g(W_{i,n} + 1) \\
= \sum_{j=1}^{X_{N_{i,1}}} \Delta g(W_{i,n} + j). \tag{3.6}
\]

Using (3.6) in (3.5), we get
\[
E(W_n)E(g(W_n + 1)) - E(W_n g(W_n + 1)) = \sum_{i=1}^{n} \mathbb{E}(X_i)E \left( \sum_{j=1}^{X_{N_{i,1}}} \Delta g(W_{i,n} + j) \right) \\
- \sum_{i=1}^{n} \mathbb{E}(X_i) \left( \sum_{j=1}^{X_{N_{i,1}}} \Delta g(W_{i,n} + j) \right). \tag{3.7}
\]
Substituting (3.7) in (3.4), we have
\[
E(A_Z g(W_n)) = (1-b) \left\{ \sum_{i=1}^{n} \mathbb{E}(X_i)E \left( \sum_{j=1}^{X_{N_{i,1}}} \Delta g(W_{i,n} + j) \right) \\
- \sum_{i=1}^{n} \mathbb{E}(X_i) \left( \sum_{j=1}^{X_{N_{i,1}}} \Delta g(W_{i,n} + j) \right) \right\} \\
+ \sum_{i=1}^{n} \mathbb{E}(X_i) \Delta g(W_n)). \tag{3.8}
\]
Note that \( \mathbb{E}(Z) = a/(1-b) = \mathbb{E}(W_n) = \sum_{i=1}^{n} \mathbb{E}(X_i). \) Therefore, from (2.14),
\[
\text{Var}(Z) = \frac{a}{(1-b)^2} = \frac{1}{(1-b)} \sum_{i=1}^{n} \mathbb{E}(X_i).
\]
Hence,
\[
\tau = \text{Var}(W_n) - \text{Var}(Z) = \sum_{i=1}^{n} \mathbb{E}(X_i X_{N_{i,1}}) - \sum_{i=1}^{n} \mathbb{E}(X_i) \mathbb{E}(X_{N_{i,1}}) - \frac{1}{(1-b)} \sum_{i=1}^{n} \mathbb{E}(X_i).
\]
This implies
\[
(1-b) \left\{ \sum_{i=1}^{n} \mathbb{E}(X_i X_{N_{i,1}}) - \sum_{i=1}^{n} \mathbb{E}(X_i) \mathbb{E}(X_{N_{i,1}}) \right\} - \sum_{i=1}^{n} \mathbb{E}(X_i) - \tau (1-b) = 0. \tag{3.9}
\]
Next, define
\[
V_{i,n} := W_n - X_{N_{i,2}} \tag{3.10}
\]
so that \( X_{N_{i,1}} \) and \( V_{i,n} \) are independent, and \( X_i \) and \( V_{i,n} \) are independent. Using (3.9) and (3.10) in (3.8), we get
\[
\mathbb{E}(A_Z g(W_n)) = (1-b) \left\{ \sum_{i=1}^{n} \mathbb{E}(X_i)E \left( \sum_{j=1}^{X_{N_{i,1}}} \Delta g(W_{i,n} + j) - \Delta g(V_{i,n}) \right) \right\}
\]
Consider first

Substituting (3.12) and (3.13) in (3.11), we have

\[
- \sum_{i=1}^{n} X_{N_{i,1}} \sum_{j=1}^{X_{N_{i,1}}} \left( \Delta g(W_{i,n} + j) - \Delta g(V_{i,n}) \right)
\]

\[
- \sum_{i=1}^{n} \left[ (1-b) \{ \mathbb{E}(X_i) \mathbb{E}(X_{N_{i,1}}) - \mathbb{E}(X_i X_{N_{i,1}}) \} + \mathbb{E}(X_i) \mathbb{E}(\Delta g(W_n) - \Delta g(V_{i,n})) \right]
\]

\[
+ \sum_{i=1}^{n} \mathbb{E}(X_i (\Delta g(W_n) - \Delta g(V_{i,n}))) - \tau (1-b) \mathbb{E}(\Delta g(W_n)) \tag{3.11}
\]

Note also that

\[
\Delta g(W_{i,n} + j) - \Delta g(V_{i,n}) = \Delta g(V_{i,n} + X_{N_{i,2}-N_{i,1}} + j) - \Delta g(V_{i,n})
\]

\[
= \sum_{k=1}^{X_{N_{i,2}-N_{i,1}+j-1}} \Delta^2 g(V_{i,n} + k). \tag{3.12}
\]

and

\[
\Delta g(W_n) - \Delta g(V_{i,n}) = \Delta g(V_{i,n} + X_{N_{i,2}}) - \Delta g(V_{i,n})
\]

\[
= \sum_{k=1}^{X_{N_{i,2}-1}} \Delta^2 g(V_{i,n} + k). \tag{3.13}
\]

Substituting (3.12) and (3.13) in (3.11), we have

\[
\mathbb{E}(A_2 g(W_n)) = (1-b) \left\{ \sum_{i=1}^{n} \mathbb{E}(X_i) \mathbb{E} \left( \sum_{j=1}^{X_{N_{i,1}}} \sum_{k=1}^{X_{N_{i,2}-N_{i,1}+j-1}} \Delta^2 g(V_{i,n} + k) \right) \right\}
\]

\[
- \sum_{i=1}^{n} \mathbb{E} \left( X_i \sum_{j=1}^{X_{N_{i,1}}} \sum_{k=1}^{X_{N_{i,2}-N_{i,1}+j-1}} \Delta^2 g(V_{i,n} + k) \right) \right\}
\]

\[
- \sum_{i=1}^{n} \left[ (1-b) \{ \mathbb{E}(X_i) \mathbb{E}(X_{N_{i,1}}) - \mathbb{E}(X_i X_{N_{i,1}}) \} + \mathbb{E}(X_i) \mathbb{E}(\Delta g(W_n) - \Delta g(V_{i,n})) \right]
\]

\[
\times \mathbb{E} \left( \sum_{j=1}^{X_{N_{i,2}-1}} \Delta^2 g(V_{i,n} + j) \right)
\]

\[
+ \sum_{i=1}^{n} \mathbb{E} \left( X_i \sum_{j=1}^{X_{N_{i,2}-1}} \Delta^2 g(V_{i,n} + j) \right) - \tau (1-b) \mathbb{E}(\Delta g(W_n)). \tag{3.14}
\]

Consider first

\[
\mathbb{E} \left( X_i \sum_{j=1}^{X_{N_{i,2}-1}} \Delta^2 g(V_{i,n} + j) \right) = \mathbb{E} \left( \mathbb{E} \left( X_i \sum_{j=1}^{X_{N_{i,2}-1}} \Delta^2 g(V_{i,n} + j) \middle| X_{N_{i,2}} \right) \right)
\]

\[
= \mathbb{E} \left( \mathbb{E}(X_i | X_{N_{i,2}}) \mathbb{E} \left( \sum_{j=1}^{X_{N_{i,2}-1}} \Delta^2 g(W_n - X_{N_{i,2}} + j) \middle| X_{N_{i,2}} \right) \right), \tag{3.15}
\]
since \( X_i \) and \( V_{i,n} \) are independent given \( X_{N_{i,2}} \). Observe that

\[
\left| E \left( \sum_{j=1}^{X_{N_{i,2}}-1} \Delta^2 g(W_n - X_{N_{i,2}} + j) \bigg| X_{N_{i,2}} = n_{i,2} \right) \right| \leq \| \Delta g \| \| n_{i,2} - 1 \| D(W_n|X_{N_{i,2}} = n_{i,2}).
\]  

(3.16)

Using (3.16) in (3.15), we have

\[
\left| E \left( X_i \sum_{j=1}^{X_{N_{i,2}}-1} \Delta^2 g(V_{i,n} + j) \right) \right| \leq \| \Delta g \| \| E(V_i|X_{N_{i,2}} - 1 \| D(W_n|X_{N_{i,2}}))
\]

\[
= \| \Delta g \| \| E(X_i(X_{N_{i,2}} - 1) D(W_n|X_{N_{i,2}})) \),
\]  

(3.17)

since \( X_i X_{N_{i,2}} \geq X_i \iff X_i(X_{N_{i,2}} - 1) \geq 0 \). Consider next the following expression from (3.14)

\[
E \left( \sum_{j=1}^{X_{N_{i,1}} X_{N_{i,2}-N_{i,1}}+j} \sum_{k=1}^{X_{N_{i,2}}-X_{N_{i,1}}+j-1} \Delta^2 g(V_{i,n} + k) \right)
\]

\[
= E \left( E \left( \sum_{j=1}^{X_{N_{i,1}} X_{N_{i,2}}-X_{N_{i,1}}+j} \sum_{k=1}^{X_{N_{i,2}}-X_{N_{i,1}}+j-1} \Delta^2 g(W_n - X_{N_{i,2}} + k) \bigg| X_{N_{i,2}} = n_{i,2} \right) \right).
\]  

(3.18)

Then

\[
\left| E \left( \sum_{j=1}^{X_{N_{i,1}} X_{N_{i,2}-N_{i,1}}+j} \sum_{k=1}^{X_{N_{i,2}}-X_{N_{i,1}}+j-1} \Delta^2 g(W_n - X_{N_{i,2}} + k) \bigg| X_{N_{i,1}} = n_{i,1}, X_{N_{i,2}} = n_{i,2} \right) \right|
\]

\[
\leq \| \Delta g \| \| n_{i,1} |2n_{i,2} - n_{i,1} - 1 \| D(W_n|X_{N_{i,1}} = n_{i,1}, X_{N_{i,2}} = n_{i,2}).
\]  

(3.19)

Using (3.19) in (3.18), we get

\[
\left| E \left( \sum_{j=1}^{X_{N_{i,1}} X_{N_{i,2}-N_{i,1}}+j} \sum_{k=1}^{X_{N_{i,2}}-X_{N_{i,1}}+j-1} \Delta^2 g(V_{i,n} + k) \right) \right|
\]

\[
\leq \| \Delta g \| \| E(X_{N_{i,1}} |2X_{N_{i,2}} - X_{N_{i,1}} - 1 \| D(W_n|X_{N_{i,1}}, X_{N_{i,2}}))
\]

\[
= \| \Delta g \| \| E(X_{N_{i,1}} |2X_{N_{i,2}} - X_{N_{i,1}} - 1 \| D(W_n|X_{N_{i,1}}, X_{N_{i,2}})),
\]  

(3.20)

since \( X_{N_{i,2}} X_{N_{i,1}} - X_{N_{i,1}}^2 \geq 0 \) and \( X_{N_{i,2}} X_{N_{i,1}} - X_{N_{i,1}} \geq 0 \) which imply \( X_{N_{i,1}} |2X_{N_{i,2}} - X_{N_{i,1}} - 1 \geq 0 \). Similarly,

\[
\left| E \left( X_i \sum_{j=1}^{X_{N_{i,1}} X_{N_{i,2}-N_{i,1}}+j} \sum_{k=1}^{X_{N_{i,2}}-X_{N_{i,1}}+j-1} \Delta^2 g(V_{i,n} + k) \right) \right|
\]

\[
\leq \| \Delta g \| \| E(X_i X_{N_{i,1}} |2X_{N_{i,2}} - X_{N_{i,1}} - 1 \| D(W_n|X_{N_{i,1}}, X_{N_{i,2}}))
\]  

(3.21)
and
\[
\left| \sum_{i=1}^{n} \left( (1-b) \{ \mathbb{E}(X_i) \mathbb{E}(X_{N_{i,1}}) - \mathbb{E}(X_{i}X_{N_{i,1}}) \} + \mathbb{E}(X_i) \right) \right| \\
\leq \| \Delta g \| \sum_{i=1}^{n} \left| \left( (1-b) \{ \mathbb{E}(X_i) \mathbb{E}(X_{N_{i,1}}) - \mathbb{E}(X_iX_{N_{i,1}}) \} + \mathbb{E}(X_i) \right) \mathbb{E}((X_{N_{i,2}} - 1)D(W_{n}X_{N_{i,2}})) \right|.
\]

Finally, using (3.17), (3.20), (3.21) and (3.22) in (3.14), the proof follows.

\[(3.22)\]

**Remark 3.1.** (i) For \( n \geq 1 \), we can use (3.4) to obtain the following crude upper bound for \( d_{TV}(W_{n}, Z) \).

\[
d_{TV}(W_{n}, Z) \leq (2|1-b||g| + \| \Delta g \|) \sum_{i=1}^{n} \mathbb{E}(X_i).
\]

Note however that for \( n \geq 6 \), the bound given in (3.3) would be better than the one given in (3.23) when \( \text{Var}(W_{n}) = \text{Var}(Z) \). Also, observe that the condition \( n \geq 6 \) is needed from the proof point of view (see (3.10)).

(ii) Assume \( D(W_{n}|X_{N_{i,2}}) \leq c_{i}(n) \) so that \( D(W_{n}|X_{N_{i,1}}, X_{N_{i,2}}) \leq c_{i}(n) \). Then the bound given in (3.3) becomes

\[
d_{TV}(W_{n}, Z) \leq \| \Delta g \| \left\{ |\tau(1-b)| + \sum_{i=1}^{n} c_{i}(n) \left[ \frac{1-b}{2} \left[ \mathbb{E}(X_i) \mathbb{E}(X_{N_{i,1}}(2X_{N_{i,2}} - X_{N_{i,1}} - 1)) + \mathbb{E}(X_i(X_{N_{i,2}} - 1)) \right] + \left[ \left( 1-b \right) \mathbb{E}(X_iX_{N_{i,1}}(2X_{N_{i,2}} - X_{N_{i,1}} - 1)) + \mathbb{E}(X_i) \mathbb{E}(X_{N_{i,2}} - 1) \right] \right] \right\}
\]

\[
= d_{1}(n).
\]

Furthermore, let us denote \( \mathcal{L}(W_{i,n}^{*}) = \mathcal{L}(W_{n}|X_{N_{i,2}}) \) and \( Z_{e} = (X_{2}, X_{4}, \ldots, X_{2[\frac{n}{2}]}) = \{X_{2m}| m \in \{1, \ldots, \frac{n}{2}\} \} \), where \( [x] \) is the greatest integer less than or equal to \( x \). Then \( \mathcal{L}(W_{i,n}^{*}|Z_{e} = z_{e}) \) can be written as the sum of independent rvs, say \( X_{j}^{(z_{e})} \), for \( j = 1, 2, \ldots, n_{z_{e}} \). Therefore, using (5.11) of Röllin (2008), we have

\[
D(W_{i,n}^{*}) \leq \mathbb{E}(\mathbb{E}(D(W_{i,n}^{*})|Z_{e})) \leq \mathbb{E} \left( \frac{2}{V_{i,Z_{e}}} \right), \quad (3.25)
\]

where

\[
V_{i,Z_{e}} = \sum_{j=1}^{n_{z_{e}}} \min \left\{ \frac{1}{2}, 1 - \frac{1}{2}D(X_{j}^{(z_{e})}) \right\}.
\]

On the other hand, let

\[
m^{*} = \begin{cases} \left\lfloor \frac{n}{2} \right\rfloor + 1, & \text{if } n \text{ is odd} \\ \frac{n}{2}, & \text{if } n \text{ is even} \end{cases} \quad \text{and} \quad Z_{0} = \{X_{2m-1}| m \in \{1, \ldots, m^{*}\}\}.
\]

(3.27)
Then, applying the similar argument as above, we get

\[ D(W_{i,n}^*) \leq \mathbb{E}(\mathbb{E}(D(W_{i,n}^*)|Z_o)) \leq \mathbb{E}\left( \frac{2}{V_{i,Z_o}^{1/2}} \right), \]  

(3.28)

where \( V_{i,z_o} \) is defined in a similar way to \( V_{i,z} \). Hence, from (3.25) and (3.28), we have

\[ D(W_{i,n}^*) \leq \min \left\{ \mathbb{E}\left( \frac{2}{V_{i,Z_o}^{1/2}} \right), \mathbb{E}\left( \frac{2}{V_{i,Z_o}^{1/2}} \right) \right\} = c_i(n). \]

As \( \mathcal{L}(W_{i,n}^*|Z_e = z_e) \) can be represented as the sum of independent rvs which depends on \( n \), and therefore, from Corollary 1.6 Mattner and Roos (2007), the upper bound for \( \mathcal{L}(W_{i,n}^*) \) is of \( O(n^{-1/2}) \). For more details, we refer the reader to Section 5.3 and Section 5.4 of Röllin (2008), Corollary 1.6 of Mattner and Roos (2007), and Proposition 4.6 of Barbour and Xia (1999).

(iii) The bound given in Theorem 3.1 can also be used for the case of matching the first two moments (i.e., \( \tau = 0 \)), whenever that is possible with the approximating distribution.

(iv) From (3.8), it can be easily verified that in the case of first moment matching, we have

\[ d_{TV}(W_n, Z) \leq \|\Delta g\| \left\{ |1-b| \sum_{i=1}^n [\mathbb{E}(X_i)\mathbb{E}(X_{N,i}) + \mathbb{E}(X_iX_{N,i})] + \sum_{i=1}^n \mathbb{E}(X_i) \right\} 

=: d_2(n), \]

and then we have \( d_{TV}(W_n, Z) \leq \min\{d_1(n), d_2(n)\} \), where \( d_1(n) \) is defined in (3.24).

(v) Observe that if \( \tau = 0 \) then the bound given in (3.3) is of optimal order \( O(n^{-1/2}) \) and is comparable with the existing bounds (Theorems 3.1 3.3, and 3.4 for Poisson, negative binomial, and binomial, respectively) given by Čekanavičius and Vellaisamy (2015) with the relaxation of the conditions (3.1) – (3.3). For example, the bound given in Theorem 4.1 of Čekanavičius and Vellaisamy (2015) is valid when \( p \leq 1/20 \) and \( np^2 \geq 1 \) for negative binomial approximation to 2-runs; however, from (4.9), our bound is valid for all values of \( p \) and sufficiently large values of \( n \).

As discussed in Section 2, the upper bound for \( \|\Delta g\| \) becomes uniform for a rv \( Z \in \mathcal{P}_2 \). Therefore, we restrict ourselves to \( Z \in \mathcal{P}_2 \) in the following corollary and demonstrate the approximation result for the restricted PSD family \( \mathcal{P}_2 \).

**Corollary 3.1.** Assume that the conditions of Theorem 3.1 hold. Then, for any \( X \in \mathcal{P}_2 \) and \( n \geq 6 \),

\[
d_{TV}(W_n, X) \leq \min \left\{ 1, \frac{1}{a} \right\} \left\{ \tau (1-b) + \sum_{i=1}^n c_i(n) \left[ \frac{|1-b|}{2} \left[ \mathbb{E}(X_i)\mathbb{E}(X_{N,i}(2X_{N,i} - X_{N,i} - 1)) + \mathbb{E}(X_iX_{N,i}(2X_{N,i} - X_{N,i} - 1)) \right] + \mathbb{E}(X_i(X_{N,i} - 1)) \right] + \left[ \left( 1 - b \right) \left\{ \mathbb{E}(X_i)\mathbb{E}(X_{N,i}) - \mathbb{E}(X_iX_{N,i}) \right\} + \mathbb{E}(X_i) \right] \mathbb{E}(X_{N,i} - 1) \right\}. \]

(3.29)

Next, using Corollary 3.1, we demonstrate our results for Poisson and negative binomial distributions, as special cases, which are useful for applications in Section 4.
Example. Assume that the conditions of Corollary 3.1 hold. Moreover, let \( Y \sim \text{Poi}(\lambda) \), the Poisson rv, so that \( a = \lambda \) and \( b = 0 \). Then, for \( n \geq 6 \),

\[
d_{TV}(W_n, Y) \leq \min \left\{ 1, \frac{1}{\lambda} \right\} \left\{ |\bar{\tau}_1| + \sum_{i=1}^{n} c_i(n) \left[ \frac{1}{2} \left( \mathbb{E}(X_i)\mathbb{E}(X_{N_{i,1}}(2X_{N_{i,2}} - X_{N_{i,1}} - 1)) + \mathbb{E}(X_i(X_{N_{i,2}} - 1)) + \mathbb{E}(X_i)\mathbb{E}(X_{N_{i,1}}(2X_{N_{i,2}} - X_{N_{i,1}} - 1)) + \mathbb{E}(X_i(X_{N_{i,2}} - 1)) \right] \right\},
\]

where \( \bar{\tau}_1 = \text{Var}(W_n) - \lambda \).

Example. Assume that the conditions of Corollary 3.1 hold. Moreover, let \( U \sim \text{NB}(\alpha, p) \), the negative binomial rv, so that \( a = \alpha(1 - p) \) and \( b = 1 - p \). Then, for \( n \geq 6 \),

\[
d_{TV}(W_n, U) \leq \min \left\{ 1, \frac{1}{\alpha(1 - p)} \right\} \left\{ |\bar{\tau}_2| + \sum_{i=1}^{n} c_i(n) \left[ \frac{p}{2} \left( \mathbb{E}(X_i)\mathbb{E}(X_{N_{i,1}}(2X_{N_{i,2}} - X_{N_{i,1}} - 1)) + \mathbb{E}(X_i(X_{N_{i,2}} - 1)) + \mathbb{E}(X_i)\mathbb{E}(X_{N_{i,1}}(2X_{N_{i,2}} - X_{N_{i,1}} - 1)) + \mathbb{E}(X_i(X_{N_{i,2}} - 1)) \right] \right\},
\]

where \( \bar{\tau}_2 = \text{Var}(W_n) - \alpha(1 - p)/p^2 \).

4 Applications to Runs

The distribution of runs and patterns has been applied successfully in many areas such as reliability theory, machine maintenance, quality control, and statistical testing, among many others. Also, it is not tractable if the underlying setup is arising from non-identical trials. So, the approximation of the runs has been studied by several researchers that include, among others, Fu and Johnson (2009), Godbole and Schaffner (1993), Upadhye and Kumar (2018), Kumar and Upadhye (2020), Vellaisamy (2004), and Wang and Xia (2008). In this section, we mainly focus on 2-runs and \((k_1, k_2)\)-runs, however, the results can also be extended to other types of runs.

4.1 2-runs

We consider here the setup similar to the one discussed in Chapter 5 of Balakrishnan and Koutras (2002), p. 166, for 2-runs. Let \( \eta_1, \eta_2, \ldots, \eta_{n+1} \) be a sequence of independent Bernoulli trials with success probability \( P(\eta_i = 1) = p_i = 1 - P(\eta_i = 0) \), for \( i = 1, 2, \ldots, n+1 \), and

\[
R_n := \sum_{i=1}^{n} X_i,
\]

(4.1)

where \( X_i = \eta_i\eta_{i+1} \), \( 1 \leq i \leq n \), is a sequence of 1-dependent rvs. Observe that \( R_n \) counts the number of overlapping success runs of length 2 in \( n + 1 \) trials. It is easy to see that
\[ E X_i = P(X_i = 1) = p_i p_{i+1} := a_1(p_i). \] Similarly, \[ E(X_i X_{i+1}) = p_i p_{i+1} p_{i+2} := a_2(p_i) \] and \[ E(X_i X_{i+1} X_{i+2}) = p_i p_{i+1} p_{i+2} p_{i+3} := a_3(p_i). \] Now, consider the first term in \((3.3)\). Then

\[
E(X_{N_{i,1}}(2X_{N_{i,2}} - X_{N_{i,1}} - 1)) = E(X_{i-1} + X_i + X_{i+1})^2
+ E((X_{i-1} + X_i + X_{i+1})(2X_{i-2} + 2X_{i+2} - 1))
= 2[a_1(p_{i-1})a_1(p_{i+1}) + a_1(p_{i-2})a_1(p_i) + a_1(p_{i+1})]
+ a_1(p_{i+2})(a_1(p_{i-1}) + a_1(p_i)) + 2 \sum_{j=i-2}^{i+1} a_2(p_j)
:= \bar{a}_1(p_i).
\] (4.2)

Similarly,
\[
E(X_i X_{N_{i,1}}(2X_{N_{i,2}} - X_{N_{i,1}} - 1)) = E(X_i(X_{i-1} + X_i + X_{i+1})^2)
+ E(X_i(X_{i-1} + X_i + X_{i+1})(2X_{i-2} + 2X_{i+2} - 1))
= 2a_1(p_i)(a_1(p_{i-2}) + a_1(p_{i+2})) + a_2(p_{i-1})(1 + a_1(p_{i+2}))
+ 2a_2(p_i)(1 + a_1(p_{i-2})) + 2 \sum_{j=i-2}^{i} a_3(p_j) =: \bar{a}_2(p_i),
\] (4.3)

\[
|[(1 - b)\{E(X_i)E(X_{N_{i,1}}) - E(X_i X_{N_{i,1}})\} + E(X_i)]E(X_{N_{i,2}} - 1)|
= \left|\left(1 - b\right)\left\{\sum_{j=i-1}^{i+1} a_1(p_j) - a_2(p_{i-1}) - a_2(p_i)\right\} + ba_1(p_i)\right| \left(\sum_{j=i-2}^{i+2} a_1(p_j) - 1\right)
=: \bar{a}_3(p_i)
\] (4.4)

and
\[
E(X_i(X_{N_{i,2}} - 1)) = E\left(X_i \sum_{j=i-2}^{i+2} X_j\right) - E(X_i)
= a_1(p_i) \sum_{|j-i|=2} a_1(p_j) + \sum_{j=i-1}^{i} a_2(p_j) =: \bar{a}_4(p_i).
\] (4.5)

Next, recall from Remark 3.1 \((ii)\) with \(W_n = R_0\) and \(W_{i,n} = R_{i,n}, \mathcal{L}(R_{i,n}|Z_e = z_e)\) can be written as the sum of independent rvs, say \(X_j^{(z_e)}\), for \(j \in \{1, 2, \ldots, n_e\} \cap \{\ell : |\ell - i| > 2\} =: \mathcal{C}_i\), for \(i = 1, 2, \ldots, n\). Note that \(n_{z_e} = m^*\) defined in \((3.27)\) and \(X_j^{(z_e)} = X_{2j - 1}\) depends only on \(X_{2j-2}\) \((2j \notin \{2, i + 4\})\) and \(X_{2j}\) \((2j \neq i - 2, 2j \leq n\)\), \(j \in \mathcal{C}_i\), for all values of \(z_e\). So, for simplicity, let us write
\[
X_j^{(z_e)} = X_{2j-1,2j}, \quad j \in \mathcal{C}_i,
\]
where \(x_{2j-2}\) and \(x_{2j}\) are corresponding values of the rvs \(X_{2j-2}\) and \(X_{2j}\), respectively. Therefore, from \((3.26)\), we have
\[
V_{i,z_e} = \sum_{j \in \mathcal{C}_i} \min \left\{ \frac{1}{2}, 1 - \frac{1}{2} D \left(X_{2j-1,2j}^{(x_{2j-2},x_{2j})}\right) \right\}.
\]
Note that
\[
\mathbb{P}
\left( X_{2j-1}^{(1,0)} = 1 \right) = \mathbb{P}(X_{2j-1} = 1 | X_{2j-2} = 1, X_{2j} = 0) = \frac{\mathbb{P}(X_{2j-2} = 1, X_{2j-1} = 1, X_{2j} = 0)}{\mathbb{P}(X_{2j-2} = 1)\mathbb{P}(X_{2j} = 0)}
\]
\[
= \frac{\mathbb{P}(\eta_{2j-2} = 1, \eta_{2j-1} = 1, \eta_{2j} = 1, \eta_{2j+1} = 0)}{\mathbb{P}(\eta_{2j-2} = 1, \eta_{2j-1} = 1)(1 - \mathbb{P}(\eta_{2j} = 1, \eta_{2j+1} = 1))} = \frac{p_{2j}(1 - p_{2j+1})}{1 - p_{2j}p_{2j+1}}.
\]

Using similar steps, we get
\[
\delta_{p_{2j-1}}^{(x_{2j-2}, x_{2j})} = \begin{cases} 
1, & \text{if } x_{2j-2} = x_{2j} = 1; \\
\frac{p_{2j}(1 - p_{2j+1})}{1 - p_{2j}p_{2j+1}}, & \text{if } x_{2j-2} = 1, x_{2j} = 0; \\
\frac{(1 - p_{2j-2})p_{2j-1}}{1 - p_{2j-2}p_{2j-1}}, & \text{if } x_{2j-2} = 0, x_{2j} = 1; \\
\frac{(1 - p_{2j-2})p_{2j-1}p_{2j}(1 - p_{2j+1})}{(1 - p_{2j-2}p_{2j-1})(1 - p_{2j}p_{2j+1})}, & \text{if } x_{2j-2} = 0, x_{2j} = 0; \\
p_{2j-1}, & \text{if } x_{2j} = 1 \text{ and } 2j - 1 \in \{1, i + 3\}; \\
p_{2j-1}p_{2j}(1 - p_{2j+1}){1 - p_{2j}p_{2j+1}}, & \text{if } x_{2j} = 0 \text{ and } 2j - 1 \in \{1, i + 3\}; \\
p_{2j}, & \text{if } x_{2j-2} = 1 \text{ and } 2j - 1 \in \{i - 3, m^*\}; \\
\frac{(1 - p_{2j-2})p_{2j-1}p_{2j}}{1 - p_{2j-2}p_{2j-1}}, & \text{if } x_{2j-2} = 0 \text{ and } 2j - 1 \in \{i - 3, m^*\},
\end{cases}
\]
where \(\delta_{p_{2j-1}}^{(x_{2j-2}, x_{2j})} := \mathbb{P}
\left( X_{2j-1}^{(x_{2j-2}, x_{2j})} = 1 \right)\), for \(x_{2j-2}, x_{2j} \in \{0, 1\}\). Therefore,
\[
\mathcal{D}
\left( X_{2j-1}^{(x_{2j-2}, x_{2j})} \right) = \sum_{k=0}^{1} \left| \mathbb{P}
\left( X_{2j-1}^{(i,j)} = k \right) - \mathbb{P}
\left( X_{2j-1}^{(i,j)} = k \right) \right|
\]
\[
= 1 - \delta_{p_{2j-1}}^{(x_{2j-2}, x_{2j})} + \left| 1 - 2\delta_{p_{2j-1}}^{(x_{2j-2}, x_{2j})} \right|.
\]

Next, let
\[
\delta_{p_{2j-1}} = \max_{0 \leq x_{2j-2} \leq x_{2j} \leq 1} \mathcal{D}
\left( X_{2j-1}^{(x_{2j-2}, x_{2j})} \right).
\]

Hence, for all values of \(z_e\), we have
\[
V_{i,z_e} = \sum_{j \in C_i} \min \left\{ \frac{1}{2}, 1 - \frac{1}{2} \mathcal{D}
\left( X_{2j-1}^{(x_{2j-2}, x_{2j})} \right) \right\} \geq \sum_{j \in C_i} \min \left\{ \frac{1}{2}, 1 - \frac{1}{2} \delta_{p_{2j-1}}^{*} \right\}.
\]

Next, from (3.25), we have
\[
D(R_i^{*}, n) \leq \mathbb{E}
\left( \frac{2}{V_{i,z_e}^{1/2}} \right) \leq 2 \left( \sum_{j \in C_i} \min \left\{ \frac{1}{2}, 1 - \frac{1}{2} \delta_{p_{2j-1}}^{*} \right\} \right)^{-1/2} =: c_{p_i}(n).
\]

Also, note that \(D(R_i^{*}, n) \leq 2\), and therefore, we get
\[
D(R_i^{*}, n) \leq \min \{2, c_{p_i}(n)\} =: c_{p_i}(n) \quad \text{(4.6)}
\]

Hence, using (4.2)-(4.6), Theorem 3.1 and Remark 3.1 (ii), we obtain the following theorem.
Theorem 4.1. Let \( Z \in \mathcal{P}_1 \) and \( R_n \) be defined as in (4.1). Assume that \( \mathbb{E}(Z) = \mathbb{E}(R_n) \), and \( \tau = \text{Var}(R_n) - \text{Var}(Z) \). Then, for \( n \geq 6 \),

\[
d_{TV}(R_n, Z) \leq \|\Delta g\| \left\{ \sum_{i=1}^{n} c_p^*(n) \left[ \frac{1-b}{2} \left[ a_1(p_i)a_1(p_i) + \bar{a}_2(p_i) \right] + \bar{a}_3(p_i) + \bar{a}_4(p_i) \right] + |\tau(1-b)| \right\}. \tag{4.7}
\]

Remark 4.1. Note that if \( \|\Delta g\| \) is of \( O(n^{-1}) \) and \( \text{Var}(Z) = \text{Var}(R_n) \) then the above bound become of order \( O(n^{-1/2}) \) and is comparable with the bounds given by Barbour and Xia (1999), Brown and Xia (2001), Daly, Lefèvre and Utev (2012), and Wang and Xia (2008). For instance, if \( p_i = p \), for all \( 1 \leq i \leq n + 1 \), then \( a_1(p) = p^2 \), \( \bar{a}_1(p) = 8p^3 + 10p^4 \), \( \bar{a}_2(p) = 4p^3 + 10p^4 + 4p^5 \), \( \bar{a}_3(p) = ((1-b)(3p^4 - 2p^3) + bp^2)(5p^2 - 1) \), and \( \bar{a}_4(p) = 2p^3 + 2p^4 \), for all \( 1 \leq i \leq n + 1 \). Hence, from (4.7), we have

\[
d_{TV}(R_n, Z) \leq n\|\Delta g\|c_p^*(n) \left\{ \frac{1-b}{2} \left[ 4p^3 + 10p^4 + 12p^5 + 10p^6 \right] + 2(p^3 + p^4) + \bar{a}_3(p) \right\}. \tag{4.8}
\]

where

\[ c_p^*(n) = 2 \times \min \left\{ 1, \left( \min \left\{ \frac{1}{2}, 1 - \frac{1}{2}\delta_p^* \right\} (m^* - 3) \right) \right\}^{-1/2} \]

with

\[ \delta_p^* = \max \left\{ 1, 1 - p + |1 - 2p|, \frac{2 - p}{2 + 4p - p^2}, \frac{2 + 2p - 3p^2}{1 + p} \right\}. \]

Now, let \( Z \sim \text{NB} (\alpha, \bar{p}) \), the negative binomial distribution with parameters \( \alpha \) and \( \bar{p} \). Then, \( b = 1 - \bar{p} \) with \( \bar{p} = 1/(1 + 2p - 3p^3) \) and \( \|\Delta g\| \leq \frac{1}{\alpha(1-p)} = \frac{1 + 2p - 3p^2}{\alpha p^2} \), where \( \alpha \) and \( \bar{p} \) are obtained from the first two moments matching condition, and hence, we get

\[
d_{TV}(R_n, \text{NB}(\alpha, \bar{p})) \leq c_p^*(n)p[4 + 11p + 10p^2 - p^3 - 6p^4 + (1 - p)((3p - 2)(5p^2 - 1))]. \tag{4.9}
\]

Also, from Theorem 4.2 of Brown and Xia (2001), for \( n \geq 2 \) and \( p < 2/3 \), we have

\[
d_{TV}(R_n, \text{NB}(\alpha, \bar{p})) \leq \frac{32.2p}{\sqrt{(n-1)(1-p)^3}}. \tag{4.10}
\]

Note that the above bound is valid for \( p < 2/3 \), however, our bound is valid for all values of \( p \) and sufficiently large values of \( n \). Also, our bound is comparable with given by Brown and Xia (2001). For example, for various values of \( n \) and \( p \), we compare our bound with the one given in (4.10), due to Brown and Xia (2001). Some numerical computations are given in the following table.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( p )</th>
<th>From (4.9)</th>
<th>From (4.10)</th>
<th>( n )</th>
<th>From (4.9)</th>
<th>From (4.10)</th>
<th>( n )</th>
<th>From (4.9)</th>
<th>From (4.10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.804590</td>
<td>0.579588</td>
<td>0.804590</td>
<td>0.322880</td>
<td>0.804590</td>
<td>0.248395</td>
<td>10</td>
<td>1.129290</td>
<td>0.837338</td>
</tr>
<tr>
<td>0.07</td>
<td>1.129290</td>
<td>0.837338</td>
<td>0.129290</td>
<td>0.46692</td>
<td>1.129290</td>
<td>0.359031</td>
<td>0.09</td>
<td>1.456370</td>
<td>1.112790</td>
</tr>
<tr>
<td></td>
<td>1.456370</td>
<td>1.112790</td>
<td>1.456370</td>
<td>0.619922</td>
<td>1.456370</td>
<td>0.476912</td>
<td>0.50</td>
<td>0.806372</td>
<td>0.91094</td>
</tr>
<tr>
<td>0.53</td>
<td>0.861243</td>
<td>1.05950</td>
<td>0.735966</td>
<td>0.831539</td>
<td>0.735966</td>
<td>0.681275</td>
<td>2500</td>
<td>0.861243</td>
<td>1.05950</td>
</tr>
<tr>
<td>0.56</td>
<td>0.915475</td>
<td>1.23590</td>
<td>0.835543</td>
<td>1.128180</td>
<td>0.773452</td>
<td>1.044460</td>
<td>0.56</td>
<td>0.915475</td>
<td>1.23590</td>
</tr>
</tbody>
</table>
Note that our bound become better for sufficiently large values of $n$ and $p$. However, for small values of $n$ and $p$, the bounds given in (4.10) are better than our bounds. This is because $\delta^*_p$ is large and $c^*_p(n)$ becomes 0 for small values of $p$ and therefore our bound is of constant order. But, if $n$ is sufficiently large then $c^*_p(n)$ becomes of $O(n^{-1/2})$ and so our bounds also decrease as $n$ increases.

### 4.2 $(k_1, k_2)$-runs

In this subsection, we consider the setup similar to Huang and Tsai (1991) and Vellaisamy (2004). Let $I_1, I_2, \ldots$ be a sequence of independent Bernoulli trials. Here, we consider $I_1, I_2, \ldots, I_{(n+1)(k_1+k_2-1)}$ with success probability $\mathbb{P}(I_i = 1) = p_i = 1 - \mathbb{P}(I_i = 0)$, for $i = 1, 2, \ldots, (n+1)(k_1+k_2-1)$. Define $m := k_1 + k_2 - 1$ and

$$Y_j := (1 - I_j) \cdots (1 - I_{j+k_1-1})I_{j+k_1} \cdots I_{j+k_1+k_2-1}, \quad j = 1, 2, \ldots, nm. \quad (4.11)$$

Note that $Y_1, Y_2, \ldots, Y_{nm}$ is a sequence of $m$-dependent rvs. Now, let us also define

$$X_i := \sum_{j=(i-1)m+1}^{im} Y_j, \quad \text{for } i = 1, 2, \ldots, n.$$

Then $X_1, X_2, \ldots, X_n$ become a sequence of $1$-dependent rvs, that is, we reduced $m$-dependent sequence to $1$-dependent sequence of rvs. From (4.11), it is clear that $Y_i$, for $i = 1, 2, \ldots, nm$, are Bernoulli rvs and if $Y_i = 1$ then $Y_j = 0$, for all $j$ such that $|j - i| \leq m$ and $j \neq i$. Therefore, $X_i, i = 1, 2, \ldots, n$, are also Bernoulli rvs. Next, let

$$R'_n = \sum_{i=1}^{nm} Y_i = \sum_{i=1}^{n} X_i, \quad (4.12)$$

the sum of the corresponding $1$-dependent rvs ($X_i$’s). The distribution of $R'_n$ is called the distribution of $(k_1, k_2)$-runs or modified distribution of order $k$ or distribution of order $(k_1, k_2)$. For more details, we refer the reader to Balakrishnan and Koutras (2002), Huang and Tsai (1991), Upadhye and Kumar (2018), Vellaisamy (2004), and reference therein.

Next, note that

$$\mathbb{E}(Y_j) = \mathbb{P}(Y_j = 1) = (1 - p_j) \cdots (1 - p_{j+k_1-1})p_{j+k_1} \cdots p_{j+k_1+k_2-1} =: a(p_j), \quad \text{for } j = 1, 2, \ldots, nm,$$

and hence

$$\mathbb{E}(X_i) = \sum_{j=(i-1)m+1}^{im} \mathbb{E}Y_j = \sum_{j=(i-1)m+1}^{im} a(p_j) =: a^*(p_i), \quad \text{for } i = 1, 2, \ldots, n. \quad (4.13)$$

Also,

$$\mathbb{E}(X_iX_{i+1}) = \sum_{\ell_1=(i-1)m+1}^{i-1} a(p_{\ell_1}) \sum_{\ell_2=\ell_1+m+1}^{(i+1)m} a(p_{\ell_2}) + \sum_{\ell_1=i}^{(i+1)m} a(p_{\ell_1}) \sum_{\ell_2=(i-1)m+1}^{\ell_1-m-1} a(p_{\ell_2}) =: a^*(p_ip_{i+1}). \quad (4.14)$$

and

$$\mathbb{E}(X_iX_{i+1}X_{i+2}) = \sum_{\ell_1=(i-1)m+1}^{im} a(p_{\ell_1}) \sum_{\ell_2=\ell_1+m+1}^{(i+1)m-1} a(p_{\ell_2}) \sum_{\ell_3=\ell_2+m+1}^{(i+2)m} a(p_{\ell_3}).$$
Using the steps similar to (4.2)-(4.5) with (4.13), (4.14) and (4.15), we have

\[
\begin{align*}
\sum_{\ell_1=im+2}^{(i+1)m-1} a(p_{\ell_1}) & \sum_{\ell_2=(i-1)m+1}^{\ell_1-m-1} a(p_{\ell_2}) \sum_{\ell_3=\ell_1+m+1}^{(i+2)m} a(p_{\ell_3}) \\
+ \sum_{\ell_2=(i+1)m+3}^{(i+2)m} a(p_{\ell_2}) \sum_{\ell_3=(i-1)m+1}^{\ell_2-m-1} a(p_{\ell_3}) \sum_{\ell_1=im+2}^{\ell_2-m-1} a(p_{\ell_1}) \\
=: a^*(p_i p_{i+1} p_{i+2}).
\end{align*}
\]

Using the steps similar to (4.2)-(4.5) with (4.13), (4.14) and (4.15), we have

\[
\mathbb{E}(X_{N_i,1}(2X_{N_i,2} - X_{N_i,1} - 1)) \leq 2 \sum_{j=i-2}^{i+1} a^*(p_j p_{j+1}) + 2[a^*(p_{i-2}) a^*(p_{i-1}) + a^*(p_{i+1})] \\
+ a^*(p_{i-1}) a^*(p_{i+1}) + a^*(p_{i+2}) a^*(p_{i-1} + a^*(p_i)) \\
=: a_1^*(p_i),
\]

\[
\mathbb{E}(X_i X_{N_i,1}(2X_{N_i,2} - X_{N_i,1} - 1)) \leq 2a^*(p_i) a^*(p_{i-2}) + a^*(p_{i+1}) + 2a^*(p_{i-1} p_i) (1 + a^*(p_{i+2})) \\
+ 2a^*(p_i p_{i+1}) (1 + a^*(p_{i-2})) + 2 \sum_{j=i-2}^{i} a^*(p_j p_{j+1} p_{j+2}) \\
=: a_2^*(p_i),
\]

\[
\left| (1 - b) \{ \mathbb{E}(X_i) \mathbb{E}(X_{N_i,1}) - \mathbb{E}(X_i X_{N_i,1}) \} + \mathbb{E}(X_i) \mathbb{E}(X_{N_i,2} - 1) \right|
\]

\[
= \left| (1 - b) \left[ a^*(p_i) \sum_{j=i-1}^{i+1} a^*(p_j) - a^*(p_{i-1} p_i) - a^*(p_i p_{i+1}) \right] + b a^*(p_i) \left( \sum_{j=i-2}^{i+2} a^*(p_j) - 1 \right) \right|
\]

\[
=: a_3^*(p_i),
\]

and

\[
\mathbb{E}(X_i (X_{N_i,2} - 1)) \leq a^*(p_i) \sum_{|j-i|=2} a^*(p_j) + \sum_{j=i-1}^{i} a^*(p_j p_{j+1}) =: a_4^*(p_i).
\]

Next, from Subsection 4.1, following the discussion about Remark 3.1 (ii), we have

\[
V_{i,z_i} = \min_{j \in C_i} \left\{ \frac{1}{2} \left[ 1 - \frac{1}{2} D \left( X_{2j-1}^{(x_{2j-2},x_{2j})} \right) \right] \right\}.
\]

Note that

\[
\frac{1}{2} D \left( X_{2j-1}^{(x_{2j-2},x_{2j})} \right) = \frac{1}{2} \left[ \mathbb{P} \left( X_{2j-1}^{(x_{2j-2},x_{2j})} = 0 \right) + \mathbb{P} \left( X_{2j-1}^{(x_{2j-2},x_{2j})} = 1 \right) \right]
\]

\[
\leq \frac{1}{2} \left[ 1 + \mathbb{P} \left( X_{2j-1}^{(x_{2j-2},x_{2j})} = 0 \right) \right]
\]

\[
\leq \frac{1}{2} \left[ 1 + \bar{a}(p_{2j-1}) \right],
\]

where

\[
\bar{a}(p_{2j-1}) = \max_{0 \leq x_{2j-2}, x_{2j} \leq 1} \mathbb{P} \left( X_{2j-1}^{(x_{2j-2},x_{2j})} = 0 \right).
\]
Next, using (4.20), we have

\[
\frac{1}{V_{i,z_e}^{1/2}} \leq \left( \frac{1}{2} \sum_{j \in C_i} \min \{1, 1 - \bar{a}(p_{2j-1})\} \right)^{-1/2} = \left( \frac{1}{2} \sum_{j \in C_i} (1 - \bar{a}(p_{2j-1})) \right)^{-1/2}, \quad \text{for all } z_e.
\]

Therefore, from (3.25), we have

\[
D(R'_{i,n}) \leq \mathbb{E} \left( \frac{2}{V_{i,Z_e}^{1/2}} \right) \leq 2 \left( \frac{1}{2} \sum_{j \in C_i} (1 - \bar{a}(p_{2j-1})) \right)^{-1/2} =: V^*_i.\]

Similarly,

\[
D(R''_{i,n}) \leq \mathbb{E} \left( \frac{2}{V_{i,Z_e}^{1/2}} \right) \leq 2 \left( \frac{1}{2} \sum_{j \in D_i} (1 - \bar{a}(p_{2j-1})) \right)^{-1/2} =: V^*_{i,o},
\]

where \(D_i = \{1, 2, \ldots, [n/2]\} \cap \{\ell : |\ell - i| > 2\}\). Therefore,

\[
c^*_i(n) = \min\{2, V^*_i, V^*_{i,o}\}. \tag{4.22}
\]

Using (4.13), (4.16)-(4.19), (4.22), Theorem 3.1 and Remark 3.1 (ii), the following result is established.

**Theorem 4.2.** Let \(Z \in \mathcal{P}_1\) and \(R'_{n}\) be defined as in (4.12). Assume that \(\mathbb{E}(Z) = \mathbb{E}(R'_{n})\), and \(\text{Var}(R'_{n}) = \text{Var}(Z)\). Then, for \(n \geq 3m\), \(a(p_{2j-1}) < 1\) defined in (4.21),

\[
d_{TV}(R'_{n}, Z) \leq \|\Delta g\| \sum_{i=1}^{n} c^*_i(n) \left[ \frac{1-b|}{2} \left( a^*(p_i) a^*_1(p_i) + a^*_2(p_i) \right) + a^*_3(p_i) + a^*_4(p_i) \right].
\]

**Remark 4.2.** Note that the above bound is of order \(O(n^{-1/2})\) which is comparable with the existing bounds given by Upadhye and Kumar (2018) and an improvement over the bounds given by Barbour, Holst and Janson (1992), Godbole (1993), Godbole and Schaffner (1993) (with \(k_1 = 1\)), and Vellaisamy (2004) which are of \(O(1)\). Also, note that we have used a slightly different form of \((k_1, k_2)\)-runs, that is, we use \(I_1, I_2, \ldots, I_{(n+1)(k_1+k_2-1)}\) instead of \(I_1, I_2, \ldots, I_n\), so that \(X_1, X_2, \ldots, X_n\) become a sequence of 1-dependent rvs and we can directly apply our result. However, we can also use some other forms and derive the corresponding results.

**Acknowledgement**

A part of this work was carried out when the first author was a Post Doctoral Fellow at the Department of Mathematics, IIT Bombay.

**References**


Amit N. Kumar 
Department of Mathematical Sciences, 
Indian Institute of Technology (BHU), 
Varanasi-221005, India. 
E-mail: amit.kumar2703@gmail.com 
P. Vellaisamy 
Department of Mathematics, 
Indian Institute of Technology Bombay, 
Powai, Mumbai-400076, India. 
E-mail: pv@math.iitb.ac.in 

Neelesh S. Upadhye 
Department of Mathematics, 
Indian Institute of Technology Madras, 
Chennai-600036, India. 
E-mail: neelesh@iitm.ac.in