

# Compound Poisson process approximation for locally dependent real valued random variables via a new coupling inequality

Michael V. Boutsikas

*Department of Statistics and Insurance Science,*

University of Piraeus, Karaoli & Dimitriou str. 18534, Piraeus, Greece

September 29, 2005

## Abstract

In this work we present a general and quite simple upper bound for the total variation distance  $d_{TV}$  between any stochastic process  $(X_i)_{i \in \Gamma}$  defined over a countable space  $\Gamma$ , and a compound Poisson process on  $\Gamma$ . This result is sufficient for proving weak convergence for any functional of the process  $(X_i)_{i \in \Gamma}$  when the real valued  $X_i$ 's are rarely nonzero and locally dependent. Our result is being established after introducing and employing a generalization of the basic coupling inequality. Finally, two simple examples of application are presented in order to illustrate the applicability of our results.

**Abbreviated Title:** Compound Poisson process approximation.

**Key words and phrases:** law of small numbers, compound Poisson process approximation, total variation distance, rate of convergence, coupling inequality, locally dependent variables, success runs, moving sums.

**AMS 2000 subject classification:** Primary 60E15, Secondary: 62E17, 60G99

## 1 Introduction

Let  $(X_i)_{i \in \Gamma}$  be a stochastic process with state space  $R$ , where  $\Gamma$  is a countable index set. The main aim of the present work is to provide simple and effective tools for approximating the distribution of any functional of  $(X_i)_{i \in \Gamma}$  when the real valued r.v.'s  $X_i, i \in \Gamma$  are locally dependent and rarely differ from zero. This situation appears in numerous applications involving rare and locally dependent events, e.g. in risk theory, graph theory, extreme value theory, reliability theory, run and scan statistics, biomolecular sequence analysis etc.

In the simplest case when  $X_i$ 's are independent, identically distributed (i.i.d.) binary (0-1) random variables (r.v.'s),  $\sum_i X_i$  follows a binomial distribution, which can be approximated by a Poisson distribution (when  $\mathbb{P}(X_i \neq 0) \approx 0$ ). In the case of dependent  $X_i$ 's, two methods have been mainly used for obtaining Poisson approximation results. The first, initiated by Freedman (1974) and Serfling (1975), concerns sums of dependent indicators  $X_1, X_2, \dots, X_n$  (cf. also Serfling (1978), Serfozo (1986)). Typically, this approach offers bounds for the total variation distance between

the distribution of the sum of indicators  $\sum_i X_i$  and an appropriate Poisson distribution. These bounds are expressed in terms of conditional probabilities of the form  $\mathbb{P}(X_i|\mathcal{F}_{i-1})$  assuming that  $\{X_i\}$  are adapted to a filtration  $\{\mathcal{F}_i\}$ . The method exploits coupling techniques, but is also related to martingale theory. For recent developments of this approach we refer to the work of Vellaisamy and Chaudhuri (1999).

The second and most important method for Poisson approximation (for dependent r.v.'s) is based on an adaptation (by Chen(1975)) for the Poisson distribution of Stein's technique for normal approximation. This much acclaimed method (referred to as Stein's method for Poisson approximation or Stein-Chen method) is based on the solution of a difference equation (Stein's equation) but it also exploits coupling techniques. This method was refined and extended by many authors to various directions and applied to a series of problems and models in diverse research areas. For a complete list of the relevant articles we refer to the monograph of Barbour, Holst and Janson (1992) and the recent review article of Barbour and Chryssaphinou (2001). In the last years, substantial attention has been drawn to results concerning Poisson process approximation through Stein's method. Refer to Arratia, Goldstein and Gordon (1989) for countable carrier spaces, and to Barbour and Mansson (2002), Chen and Xia (2004) and the references therein for more general carrier spaces.

An alternative third way for obtaining compound Poisson approximation results similar to the ones offered by Stein's method has been recently proposed by Boutsikas and Koutras (2000) for sums of integer-valued associated r.v.'s, Boutsikas and Vaggelatou (2002) for sums of real valued associated r.v.'s and Boutsikas and Koutras (2001) for sums of locally dependent r.v.'s. These approaches were couched on direct probabilistic methods, namely, specific dependence concepts, stochastic orders or coupling techniques.

In this paper we present a compound Poisson process approximation result for locally dependent real valued random variables. More specifically we present an upper bound for the total variation distance  $d_{TV}$  between the law of any stochastic process  $(X_i)_{i \in \Gamma}$  defined over a finite or more generally countable space  $\Gamma$ , and an appropriate compound Poisson process on  $\Gamma$ . This bound is small when  $X_i, i \in \Gamma$  are locally dependent and  $\mathbb{P}(X_i \neq 0) \approx 0, i \in \Gamma$ . This result was proved by introducing and exploiting *a new coupling inequality* that can be considered as a generalization of the well known *basic coupling inequality*.

It is remarkable that the form of the bound we provide is similar to the bounds offered by Arratia, Goldstein and Gordon (1989) using Stein's method. Their bounds concern the distance between the law of a sequence of indicator  $X_i$ 's and an appropriate Poisson process and in that sense our result (which concerns real valued  $X_i$ 's and compound Poisson process) can be considered as an extension. It is also worth mentioning that, the so called "magic factor" or "Stein factor" (a factor that decreases as  $\lambda$ , the mean of the approximating Poisson distribution, increases) that appears in the upper bound of many Poisson approximation results through Stein's method, cannot be present in our bounds since we use the total variation distance and a compound Poisson process (cf. Barbour, Holst and Janson (1992), page 203).

## 2 Preliminaries

A *random element*  $\Xi$  is a measurable mapping from a probability space  $(\Omega, \mathfrak{F}, \mathbb{P})$  to a measurable space  $(E, \mathcal{A})$ . A *coupling* of two random elements  $\Xi, \Psi$  from  $(\Omega_1, \mathfrak{F}_1, \mathbb{P}_1), (\Omega_2, \mathfrak{F}_2, \mathbb{P}_2)$  respectively to  $(E, \mathcal{A})$  is any random element  $(\Xi', \Psi')$  from  $(\Omega, \mathfrak{F}, \mathbb{P})$  to  $(E \times E, \mathcal{A} \otimes \mathcal{A})$  such that  $\mathcal{L}(\Xi) = \mathcal{L}(\Xi')$  and  $\mathcal{L}(\Psi) = \mathcal{L}(\Psi')$  where, as usual,  $\mathcal{L}(\Xi)$  denotes the law of  $\Xi$ . Loosely speaking, a coupling of  $\Xi, \Psi$  is any "definition" of  $\Xi, \Psi$  in the same probability space. In order to check how "close" are the laws  $\mathcal{L}(\Xi), \mathcal{L}(\Psi)$  of two random elements  $\Xi, \Psi$  we shall be using the well-known total variation distance

$$d_{TV}(\mathcal{L}(\Xi), \mathcal{L}(\Psi)) = \sup_{A \in \mathcal{A}} |\mathbb{P}_1(\Xi \in A) - \mathbb{P}_2(\Psi \in A)| = \sup_{A \in \mathcal{A}} |\mathbb{P}(\Xi' \in A) - \mathbb{P}(\Psi' \in A)|,$$

which may sometimes be too strong for proving convergence of probability measures (requiring "similarity" of the two measures in *every* event whereas e.g. vague convergence requires "similarity" in events with nonzero measure boundaries) but on the other hand it possesses the following useful property

$$d_{TV}(\mathcal{L}(f(\Xi)), \mathcal{L}(f(\Psi))) \leq d_{TV}(\mathcal{L}(\Xi), \mathcal{L}(\Psi)), \quad (1)$$

for *every* measurable  $f$ . Therefore, if  $\mathcal{L}(\Xi)$  approximates  $\mathcal{L}(\Psi)$  w.r.t.  $d_{TV}$  then, with the same accuracy, the law of any functional of  $\Xi$  approximates the law of the same functional of  $\Psi$ .

A well known result concerning the  $d_{TV}$  is the so called *basic coupling inequality*: for any coupling  $(\Xi', \Psi')$  of two random elements  $\Xi, \Psi$ ,

$$d_{TV}(\mathcal{L}(\Xi), \mathcal{L}(\Psi)) \leq \mathbb{P}(\Xi' \neq \Psi'). \quad (2)$$

The standard way to assure that the event  $[\Xi' = \Psi']$  belongs to the  $\sigma$ -algebra  $\mathfrak{F}$  is to restrict ourselves to state spaces  $(E, \mathfrak{B}(E))$  that are Polish (i.e. complete and separable metric spaces) where  $\mathfrak{B}(E)$  denotes the usual Borel  $\sigma$ -algebra generated by the open sets in  $E$ .

Next, we state two well known preliminary results (see e.g. Serfling (1978), Wang (1986), Wang (1989) or Barbour, Holst and Janson (1992)) that will be used in the sequel. In what follows, whenever dependency or independency of some random elements is mentioned, this will immediately imply that these are defined over the same probability space. The following lemma is easily proved by resorting to the basic coupling inequality and the triangle inequality for  $d_{TV}$ .

**Lemma 1** *Let  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  and  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$  be two collections of random vectors ( $\mathbf{X}_i, \mathbf{Y}_i \in \mathbb{R}^{k_i}, i = 1, 2, \dots, n$ ). If the couples  $(\mathbf{X}_1, \mathbf{Y}_1), (\mathbf{X}_2, \mathbf{Y}_2), \dots, (\mathbf{X}_n, \mathbf{Y}_n)$  are independent then*

$$d_{TV}(\mathcal{L}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n), \mathcal{L}(\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n)) \leq \sum_{i=1}^n d_{TV}(\mathcal{L}(\mathbf{X}_i), \mathcal{L}(\mathbf{Y}_i)).$$

As usual,  $CP(\lambda, F)$  denotes the distribution of the random sum  $\sum_{i=1}^N X_i$  where  $X_1, X_2, \dots$  is a sequence of i.i.d. r.v.'s with common d.f.  $F$  and  $N$  is a r.v. independent of the  $X_i$ 's following Poisson distribution with mean  $\lambda$ .

We can now use the above lemma to derive a simple bound for the total variation distance between the joint distribution of a random vector with independent components and a compound Poisson product measure. For the proof of this bound we also use the inequality  $d_{TV}(\mathcal{L}(X), CP(\lambda, F)) \leq \mathbb{P}(X \neq 0)^2$ ,  $\lambda = \mathbb{P}(X \neq 0)$ ,  $F(x) = \mathbb{P}(X \leq x | X \neq 0)$  which holds for any real valued r.v.  $X$ .

**Proposition 2** *If  $X_1, X_2, \dots, X_n$  are independent real valued r.v.'s, then,*

$$d_{TV}(\mathcal{L}(\mathbf{X}), \prod_{i=1}^n CP(\lambda_i, F_i)) \leq \sum_{i=1}^n \mathbb{P}(X_i \neq 0)^2,$$

where  $\mathbf{X} = (X_1, X_2, \dots, X_n)$ ,  $\lambda_i = \mathbb{P}(X_i \neq 0)$  and  $F_i(x) = \mathbb{P}(X_i \leq x | X_i \neq 0)$ ,  $i = 1, 2, \dots, n$ .

The product measure  $\prod_{i=1}^n CP(\lambda_i, F_i)$  coincides with the distribution  $\mathcal{L}(\mathbf{Y})$  of a random vector  $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n) \in R^n$  with independent components where each  $Y_i$  follows a  $CP(\lambda_i, F_i)$  distribution.

Proposition's 2 inequality, which can be considered a by-product of the basic coupling inequality, can be used to establish compound Poisson process approximation results for sequences of independent r.v.'s. Unfortunately, when we look at cases where the  $X_i$ 's may possibly be dependent, the basic coupling inequality cannot help. In order to obtain similar results for dependent r.v.'s using coupling, it seems reasonable to try first to find an appropriate generalization of the basic coupling inequality. This is the aim of the next section.

### 3 The generalized coupling inequality

The basic coupling inequality (2) offers a bound for the distance between the laws of two random elements. It would be more flexible though, to possess a result concerning the change in  $d_{TV}$  between the laws of two random elements which occurs when we modify these two elements (e.g. change some of their coordinates). Such a result is offered by the next lemma which, apart from its independent interest, is the basic ingredient for the establishment of our main result.

**Lemma 3** *If  $\Xi_1, \Xi_2, \Psi_1, \Psi_2$  are four random elements taking values in a Polish space  $E$ , then*

$$\begin{aligned} & |d_{TV}(\mathcal{L}(\Xi_1), \mathcal{L}(\Xi_2)) - d_{TV}(\mathcal{L}(\Psi_1), \mathcal{L}(\Psi_2))| \\ & \leq \mathbb{P}(\Xi'_1 \neq \Xi'_2, (\Xi'_1, \Xi'_2) \neq (\Psi'_1, \Psi'_2)) + \mathbb{P}(\Psi'_1 \neq \Psi'_2, (\Xi'_1, \Xi'_2) \neq (\Psi'_1, \Psi'_2)), \end{aligned} \quad (3)$$

for any coupling  $(\Xi'_1, \Xi'_2, \Psi'_1, \Psi'_2)$  of  $\Xi_1, \Xi_2, \Psi_1, \Psi_2$ .

**Proof.** Let  $(\Xi'_1, \Xi'_2, \Psi'_1, \Psi'_2)$  be a coupling of  $\Xi_1, \Xi_2, \Psi_1, \Psi_2$ , defined over some probability space  $(\Omega, \mathfrak{F}, \mathbb{P})$  and taking values in  $(E^4, \mathfrak{B}(E^4))$ . For a fixed  $B \in \mathfrak{B}(E)$  define the events of  $\mathfrak{F}$ ,

$$A_1 = [\Xi'_1 \in B], \quad A_2 = [\Xi'_2 \in B], \quad A_3 = [\Psi'_1 \in B], \quad A_4 = [\Psi'_2 \in B].$$

It is easy to see that (as usual,  $A^c$  denotes the complementary set of  $A$ ),

$$\begin{aligned} & \{\mathbb{P}(A_1) - \mathbb{P}(A_2)\} - \{\mathbb{P}(A_3) - \mathbb{P}(A_4)\} = \\ & = \{\mathbb{P}(A_1 A_2^c) - \mathbb{P}(A_1^c A_2)\} - \{\mathbb{P}(A_3 A_4^c) - \mathbb{P}(A_3^c A_4)\} \\ & = \{\mathbb{P}(A_1 A_2^c) - \mathbb{P}(A_3 A_4^c)\} - \{\mathbb{P}(A_1^c A_2) - \mathbb{P}(A_3^c A_4)\} \\ & = \{\mathbb{P}(A_1 A_2^c (A_3 A_4^c)^c) - \mathbb{P}((A_1 A_2^c)^c A_3 A_4^c)\} - \{\mathbb{P}(A_1^c A_2 (A_3^c A_4)^c) - \mathbb{P}((A_1^c A_2)^c A_3^c A_4)\} \\ & \leq \mathbb{P}(A_1 A_2^c (A_3^c \cup A_4)) + \mathbb{P}((A_1 \cup A_2^c) A_3^c A_4) \\ & = \mathbb{P}(\Xi'_1 \in B, \Xi'_2 \notin B, (\Psi'_1 \notin B \text{ or } \Psi'_2 \in B)) + \mathbb{P}((\Xi'_1 \in B \text{ or } \Xi'_2 \notin B), \Psi'_1 \notin B, \Psi'_2 \in B), \end{aligned}$$

which is upper bounded by  $c_1 + c_2$  where

$$\begin{aligned} c_1 &= \mathbb{P}(\Xi'_1 \neq \Xi'_2, \Xi'_1 \neq \Psi'_1 \text{ or } \Xi'_1 \neq \Xi'_2, \Xi'_2 \neq \Psi'_2) = \mathbb{P}(\Xi'_1 \neq \Xi'_2, (\Xi'_1, \Xi'_2) \neq (\Psi'_1, \Psi'_2)), \\ c_2 &= \mathbb{P}(\Psi'_1 \neq \Psi'_2, \Psi'_1 \neq \Xi'_1 \text{ or } \Psi'_1 \neq \Psi'_2, \Psi'_2 \neq \Xi'_2) = \mathbb{P}(\Psi'_1 \neq \Psi'_2, (\Xi'_1, \Xi'_2) \neq (\Psi'_1, \Psi'_2)). \end{aligned}$$

Interchanging  $\Xi_1$  with  $\Psi_1$  and  $\Xi_2$  with  $\Psi_2$  we get

$$\{\mathbb{P}(A_3) - \mathbb{P}(A_4)\} - \{\mathbb{P}(A_1) - \mathbb{P}(A_2)\} \leq c_2 + c_1,$$

and thus,  $|\{\mathbb{P}(A_1) - \mathbb{P}(A_2)\} - \{\mathbb{P}(A_3) - \mathbb{P}(A_4)\}| \leq c_1 + c_2$ . Since  $\|a\| - \|b\| \leq \|a + b\|$ , for every  $a, b \in R$ , we conclude that

$$\|\mathbb{P}(A_1) - \mathbb{P}(A_2)\| - \|\mathbb{P}(A_3) - \mathbb{P}(A_4)\| \leq c_1 + c_2.$$

Hence, for any  $B \in \mathfrak{B}(E)$ ,

$$|\mathbb{P}(\Xi'_1 \in B) - \mathbb{P}(\Xi'_2 \in B)| \leq |\mathbb{P}(\Psi'_1 \in B) - \mathbb{P}(\Psi'_2 \in B)| + c_1 + c_2,$$

and

$$|\mathbb{P}(\Psi'_1 \in B) - \mathbb{P}(\Psi'_2 \in B)| \leq |\mathbb{P}(\Xi'_1 \in B) - \mathbb{P}(\Xi'_2 \in B)| + c_1 + c_2.$$

Considering the supremum with respect to  $B$  at both sides of the above inequalities we deduce

$$d_{TV}(\mathcal{L}(\Xi_1), \mathcal{L}(\Xi_2)) \leq d_{TV}(\mathcal{L}(\Psi_1), \mathcal{L}(\Psi_2)) + c_1 + c_2,$$

and

$$d_{TV}(\mathcal{L}(\Psi_1), \mathcal{L}(\Psi_2)) \leq d_{TV}(\mathcal{L}(\Xi_1), \mathcal{L}(\Xi_2)) + c_1 + c_2,$$

which completes the proof. ■

It is easy to see that Lemma 3 can be considered as a generalization of the basic coupling inequality (2). Indeed, if  $(\Xi'_1, \Xi'_2)$  is a coupling of some random elements  $\Xi_1, \Xi_2$ , then Lemma's 3 inequality (considering the coupling  $(\Xi'_1, \Xi'_2, \Xi'_1, \Xi'_1)$  of  $\Xi_1, \Xi_2, \Xi_1, \Xi_1$ ) leads to

$$d_{TV}(\mathcal{L}(\Xi_1), \mathcal{L}(\Xi_2)) = |d_{TV}(\mathcal{L}(\Xi_1), \mathcal{L}(\Xi_2)) - d_{TV}(\mathcal{L}(\Xi_1), \mathcal{L}(\Xi_1))| \leq \mathbb{P}(\Xi'_1 \neq \Xi'_2).$$

Similar to the basic coupling inequality, (3) can also be used for bounding the  $d_{TV}$  between the laws of  $\Xi_1$  and  $\Xi_2$ . This can be accomplished by choosing appropriate auxiliary elements  $\Psi_1, \Psi_2$  (more precisely an appropriate coupling of  $\Xi_1, \Xi_2, \Psi_1, \Psi_2$ ) so that the upper bound of (3) is small and  $d_{TV}(\mathcal{L}(\Psi_1), \mathcal{L}(\Psi_2))$  can be easily calculated or upper bounded. This approach offers increased flexibility in the bounding procedure, due to the presence of two additional random elements  $\Psi_1, \Psi_2$  to play with.

The proof of Lemma 1 was based on the triangle inequality and the relation,

$$d_{TV}(\mathcal{L}(\mathbf{X}, \mathbf{Z}), \mathcal{L}(\mathbf{Y}, \mathbf{Z})) = d_{TV}(\mathcal{L}(\mathbf{X}), \mathcal{L}(\mathbf{Y})), \quad \mathbf{X}, \mathbf{Y} \in R^k, \mathbf{Z} \in R^r \quad (4)$$

that holds true when the random vector  $\mathbf{Z}$  is independent of  $\mathbf{X}$  and  $\mathbf{Y}$  (actually only the  $\leq$ -part of the above relation is needed), which in turn can be proved using the basic coupling inequality (2). Therefore, we cannot use this relation when the involved r.v.'s are possibly dependent. It would thus be very convenient to possess an analogous result for  $d_{TV}(\mathcal{L}(\mathbf{X}, \mathbf{Z}), \mathcal{L}(\mathbf{Y}, \mathbf{Z}))$  that holds true even when  $\mathbf{Z}$  is dependent on  $\mathbf{X}, \mathbf{Y}$ . The next corollary of Lemma 3 offers such a result which is quite flexible since it involves an arbitrarily chosen random vector  $\mathbf{Z}'$ .

**Corollary 4** For any random vectors  $\mathbf{X}, \mathbf{Y} \in R^k$  and  $\mathbf{Z}, \mathbf{Z}' \in R^r$  defined on the same probability space we have that

$$|d_{TV}(\mathcal{L}(\mathbf{Z}, \mathbf{X}), \mathcal{L}(\mathbf{Z}, \mathbf{Y})) - d_{TV}(\mathcal{L}(\mathbf{Z}', \mathbf{X}), \mathcal{L}(\mathbf{Z}', \mathbf{Y}))| \leq 2\mathbb{P}(\mathbf{X} \neq \mathbf{Y}, \mathbf{Z} \neq \mathbf{Z}'). \quad (5)$$

**Proof.** A direct application of (3) reveals that

$$\begin{aligned} & |d_{TV}(\mathcal{L}(\mathbf{Z}, \mathbf{X}), \mathcal{L}(\mathbf{Z}, \mathbf{Y})) - d_{TV}(\mathcal{L}(\mathbf{Z}', \mathbf{X}), \mathcal{L}(\mathbf{Z}', \mathbf{Y}))| \\ & \leq \mathbb{P}((\mathbf{Z}, \mathbf{X}) \neq (\mathbf{Z}, \mathbf{Y}), ((\mathbf{Z}, \mathbf{X}), (\mathbf{Z}, \mathbf{Y})) \neq ((\mathbf{Z}', \mathbf{X}), (\mathbf{Z}', \mathbf{Y}))) \\ & \quad + \mathbb{P}((\mathbf{Z}', \mathbf{X}) \neq (\mathbf{Z}', \mathbf{Y}), ((\mathbf{Z}, \mathbf{X}), (\mathbf{Z}, \mathbf{Y})) \neq ((\mathbf{Z}', \mathbf{X}), (\mathbf{Z}', \mathbf{Y}))) \\ & = \mathbb{P}(\mathbf{X} \neq \mathbf{Y}, \mathbf{Z} \neq \mathbf{Z}') + \mathbb{P}(\mathbf{X} \neq \mathbf{Y}, \mathbf{Z} \neq \mathbf{Z}'). \end{aligned}$$

■

In the following section we shall exploit the above inequality in order to obtain a compound Poisson process result for locally dependent sequences.

## 4 Compound Poisson Process approximation using the generalized coupling inequality

Consider a collection of real valued r.v.'s  $X_i, i \in \Gamma_n = \{1, 2, \dots, n\}$  and assume that for every  $X_i$  there exist a set of indices  $B_i \subset \Gamma_n - \{i\}$  so that  $X_i$  is independent of or weakly dependent on  $X_j, j \in B_i^c \equiv (\Gamma_n - \{i\}) - B_i$ . Assume also that the sets  $B_i, i \in \Gamma_n$  satisfy the reflexivity condition  $j \in B_i \Leftrightarrow i \in B_j$ . The set  $B_i$  can be considered as the *neighborhood of strong dependence of  $X_i$*  and therefore the set  $B_i \cap \Gamma_i$  is the *left neighborhood of strong dependence of  $X_i$* .

Note that the sets  $B_i, i \in \Gamma_n$  can be chosen arbitrarily but the next theorem offers better (smaller) bounds when  $B_i$ 's are chosen so that every  $X_i$  is independent of or weakly dependent on all  $X_j$ 's outside its neighborhood.

Let also  $X_i^\perp, i \in \Gamma_n$  be a sequence of independent r.v.'s (also independent of  $X_i, i \in \Gamma_n$ ) with the same marginal distributions as  $X_i, i \in \Gamma_n$  (i.e.  $\mathcal{L}(X_i) = \mathcal{L}(X_i^\perp), i \in \Gamma_n$ ).

**Theorem 5** If  $X_i, i \in \Gamma_n$  is a collection of real valued r.v.'s and  $B_i, i \in \Gamma_n$  denote their neighborhoods of strong dependence, then

$$d_{TV}(\mathcal{L}(X_1, \dots, X_n), \mathcal{L}(X_1^\perp, \dots, X_n^\perp)) \leq 2 \sum_{i=2}^n \mathbb{P}((X_b)_{b \in B_i \cap \Gamma_i} \neq \mathbf{0}, X_i \neq X_i^\perp) + c_{\mathbf{X}}(B), \quad (6)$$

where

$$c_{\mathbf{X}}(B) = \sum_{i=2}^n d_{TV}(\mathcal{L}((X_b)_{b \in B_i^c \cap \Gamma_i}, X_i), \mathcal{L}((X_b)_{b \in B_i^c \cap \Gamma_i}, X_i^\perp)). \quad (7)$$

**Proof.** Define the random vectors ( $I(A) = 1$  if relation  $A$  holds true and  $I(A) = 0$  otherwise),

$$\mathbf{Z}_i = (X_1, \dots, X_{i-1}), \quad \mathbf{Z}'_i = (X_1 \cdot I(1 \notin B_i), X_2 \cdot I(2 \notin B_i), \dots, X_{i-1} \cdot I(i-1 \notin B_i)),$$

i.e.  $\mathbf{Z}'_i$  emerges by replacing the "neighbors" of  $X_i$  in  $\mathbf{Z}_i$  with zeros. Applying Corollary 4 we get

$$\begin{aligned} \left| d_{TV}(\mathcal{L}(\mathbf{Z}_i, X_i), \mathcal{L}(\mathbf{Z}_i, X_i^\perp)) - d_{TV}(\mathcal{L}(\mathbf{Z}'_i, X_i), \mathcal{L}(\mathbf{Z}'_i, X_i^\perp)) \right| &\leq 2\mathbb{P}\left(\mathbf{Z}_i \neq \mathbf{Z}'_i, X_i \neq X_i^\perp\right) \\ &= 2\mathbb{P}\left((X_b)_{b \in B_i \cap \Gamma_i} \neq \mathbf{0}, X_i \neq X_i^\perp\right), \end{aligned}$$

(where  $\mathbf{0} = (0, 0, \dots, 0)$ ). Note now that

$$d_{TV}(\mathcal{L}(\mathbf{Z}'_i, X_i), \mathcal{L}(\mathbf{Z}'_i, X_i^\perp)) = d_{TV}(\mathcal{L}((X_b)_{b \in B_i^c \cap \Gamma_i}, X_i), \mathcal{L}((X_b)_{b \in B_i^c \cap \Gamma_i}, X_i^\perp)),$$

where the right side of the above inequality is just the left side without the coordinates of  $\mathbf{Z}'_i$  that are equal to 0 (i.e. all the coordinates with index  $j \in B_i \cap \Gamma_i$ ). This follows from the general equality,  $d_{TV}(\mathcal{L}(\mathbf{X}, \mathbf{0}), \mathcal{L}(\mathbf{Y}, \mathbf{0})) = d_{TV}(\mathcal{L}(\mathbf{X}), \mathcal{L}(\mathbf{Y}))$  that holds for every  $\mathbf{X}, \mathbf{Y}$ , which can be considered as a special case of (4).

Moreover, using again (4) we get,

$$\begin{aligned} d_{TV}(\mathcal{L}(\mathbf{Z}_i, X_i), \mathcal{L}(\mathbf{Z}_i, X_i^\perp)) &= d_{TV}(\mathcal{L}(X_1, \dots, X_i), \mathcal{L}(X_1, \dots, X_{i-1}, X_i^\perp)) \\ &= d_{TV}(\mathcal{L}(X_1, \dots, X_i, X_{i+1}^\perp, \dots, X_n^\perp), \mathcal{L}(X_1, \dots, X_{i-1}, X_i^\perp, \dots, X_n^\perp)). \end{aligned}$$

Therefore, for  $i = 2, \dots, n$ ,

$$\begin{aligned} &d_{TV}(\mathcal{L}(X_1, \dots, X_i, X_{i+1}^\perp, \dots, X_n^\perp), \mathcal{L}(X_1, \dots, X_{i-1}, X_i^\perp, \dots, X_n^\perp)) \\ &\leq 2\mathbb{P}\left((X_b)_{b \in B_i \cap \Gamma_i} \neq \mathbf{0}, X_i \neq X_i^\perp\right) + d_{TV}(\mathcal{L}((X_b)_{b \in B_i^c \cap \Gamma_i}, X_i), \mathcal{L}((X_b)_{b \in B_i^c \cap \Gamma_i}, X_i^\perp)). \quad (8) \end{aligned}$$

Finally, from the triangle inequality we conclude that

$$d_{TV}(\mathcal{L}(\mathbf{X}), \mathcal{L}(\mathbf{X}^\perp)) \leq \sum_{i=1}^n d_{TV}(\mathcal{L}(X_1, \dots, X_i, X_{i+1}^\perp, \dots, X_n^\perp), \mathcal{L}(X_1, \dots, X_{i-1}, X_i^\perp, \dots, X_n^\perp)),$$

which combined with (8) leads to (6) (from Lemma 1, the first term of the above sum equals to 0).

■

From the above proof we understand that  $\mathbb{P}\left((X_b)_{b \in B_i \cap \Gamma_i} \neq \mathbf{0}, X_i \neq X_i^\perp\right) = 0$  when some  $B_i \cap \Gamma_i = \emptyset$ . Now, in order to minimize the upper bound, it is essential that  $X_i$ 's are rarely nonzero. This situation immediately calls for a compound Poisson approximation result. Specifically, we state the following theorem.

**Theorem 6** *If  $X_i, i \in \Gamma_n$  is a collection of real valued r.v.'s and  $B_i, i \in \Gamma_n$  denote their neighborhoods of strong dependence, then*

$$d_{TV}(\mathcal{L}(\mathbf{X}), \prod_{i=1}^n CP(\lambda_i, F_i)) \leq 2 \sum_{i=2}^n \mathbb{P}\left((X_b)_{b \in B_i \cap \Gamma_i} \neq \mathbf{0}, X_i \neq X_i^\perp\right) + \sum_{i=1}^n \mathbb{P}(X_i \neq 0)^2 + c_{\mathbf{X}}(B),$$

where  $\mathbf{X} = (X_1, X_2, \dots, X_n)$ ,  $\lambda_i = \mathbb{P}(X_i \neq 0)$ ,  $F_i(x) = \mathbb{P}(X_i \leq x | X_i \neq 0)$  and  $c_{\mathbf{X}}(B)$  is given by (7).

**Proof.** From Proposition 2 we get that,

$$d_{TV}(\mathcal{L}(\mathbf{X}^\perp), \prod_{i=1}^n CP(\lambda_i, F_i)) \leq \sum_{i=1}^n \mathbb{P}(X_i \neq 0)^2,$$

where  $\mathbf{X}^\perp = (X_1^\perp, X_2^\perp, \dots, X_n^\perp)$ ,  $\lambda_i = \mathbb{P}(X_i \neq 0)$  and  $F_i(x) = \mathbb{P}(X_i \leq x | X_i \neq 0)$ . The proof is now easily completed by the use of the triangle inequality and Theorem 5. ■

Considering that  $[(X_b)_{b \in B_i \cap \Gamma_i} \neq \mathbf{0}] \cap [X_i \neq X_i^\perp] \subset \cup_{b \in B_i \cap \Gamma_i} [X_b \neq 0] \cap ([X_i \neq 0] \cup [X_i^\perp \neq 0])$ , we get the next corollary which offers a slightly worse but computationally more convenient bound.

**Corollary 7** *If  $X_i, i \in \Gamma_n$  is a collection of real valued r.v.'s and  $B_i, i \in \Gamma_n$  denote their neighborhoods of strong dependence, then*

$$\begin{aligned} d_{TV}(\mathcal{L}(\mathbf{X}), \prod_{i=1}^n CP(\lambda_i, F_i)) \leq UB &\equiv \sum_{i=1}^n \sum_{b \in B_i} (\mathbb{P}(X_b \neq 0, X_i \neq 0) + \mathbb{P}(X_b \neq 0)\mathbb{P}(X_i \neq 0)) \\ &+ \sum_{i=1}^n \mathbb{P}(X_i \neq 0)^2 + c_{\mathbf{X}}(B), \end{aligned} \quad (9)$$

where  $\mathbf{X} = (X_1, X_2, \dots, X_n)$ ,  $\lambda_i = \mathbb{P}(X_i \neq 0)$ ,  $F_i(x) = \mathbb{P}(X_i \leq x | X_i \neq 0)$  and  $c_{\mathbf{X}}(B)$  is given by (7).

**Remark 1.** By employing arguments similar to the ones used in Arratia, Goldstein and Gordon (1989), page 22, the above inequality could be extended from the finite carrier space  $\Gamma_n$  to an infinite countable carrier space  $\Gamma$  for the process  $(X_i)_{i \in \Gamma}$ .

**Remark 2.** It is worth stressing that  $c_{\mathbf{X}}(B) = 0$  when each  $X_i$  is independent of  $X_b, b \in B_i^c$  ( $B_i$  is the *neighborhood of dependence* of  $X_i$ ). However, if each  $X_i$  is "weakly" dependent on  $X_b, b \in B_i^c$  ( $B_i$  is the *neighborhood of strong dependence* of  $X_i$ ) then  $c_{\mathbf{X}}(B)$  could be upper bounded. For example, if  $X_1, X_2, \dots$  is a  $\phi$ -*mixing* sequence of integer-valued r.v.'s, then on choosing  $B_i \cap \Gamma_i = \{i - s + 1, \dots, i - 1\}$  for some  $s > 1$  and denoting  $\mathbf{Y} = (X_b)_{b \in B_i^c \cap \Gamma_i} = (X_1, \dots, X_{i-s})$ , we deduce that

$$d_{TV}(\mathcal{L}(\mathbf{Y}, X_i), \mathcal{L}(\mathbf{Y}, X_i^\perp)) = \sup_A \left| \sum_j \{ \mathbb{P}((\mathbf{Y}, j) \in A | X_i = j) - \mathbb{P}((\mathbf{Y}, j) \in A) \} \mathbb{P}(X_i = j) \right| \leq \phi(s),$$

where  $\phi(s) = \sup_k \{ \sup \{ |\mathbb{P}(B|C) - \mathbb{P}(B)|, C \in \sigma(X_i; i \leq k), B \in \sigma(X_i; i \geq k + s) \} \} \rightarrow_{s \rightarrow \infty} 0$  (e.g. for a Doeblin irreducible Markov chain,  $\phi(s) \leq ab^s$  for some  $a > 0, 0 \leq b < 1$ ). In a similar way we can treat  $\alpha$ -*mixing* (strongly mixing) or other types of weakly dependent sequences.

**Remark 3.** Theorem 6 or Corollary 7 can be used for proving weak convergence for any function of the process  $\mathbf{X}$ . More specifically, from (1) and (9) we conclude that

$$\sup_{A \in \mathcal{B}(R^k)} |\mathbb{P}(f(\mathbf{X}) \in A) - \mathbb{P}(f(\mathbf{Y}) \in A)| = d_{TV}(\mathcal{L}(f(\mathbf{X})), \mathcal{L}(f(\mathbf{Y}))) \leq UB,$$

for any measurable function  $f : R^{|\Gamma|} \rightarrow R^k$  where  $\mathbf{Y} \sim \prod_{i \in \Gamma} CP(\lambda_i, F_i)$ . If, for example, we choose  $f(\mathbf{x}) = \sum_i x_i$ , then it readily follows that  $(\lambda = \sum_i \lambda_i)$

$$d_{TV}(\mathcal{L}(\sum X_i), CP(\lambda, \sum \frac{\lambda_i}{\lambda} F_i)) \leq UB,$$

( $UB$  is given by (9)) a generalization of the *Khintchine-Doeblin* inequality. Other choices of  $f$  could for example be  $f(\mathbf{x}) = (\sum_{i \in C_1} x_i, \sum_{i \in C_2} x_i)$  with  $C_1, C_2 \subset \Gamma$ , or  $f(\mathbf{x}) = (\max_i x_i, \sum_i x_i)$  etc.

## 5 Applications

As already mentioned in the introduction, the bound of Corollary 7 has almost the same form as the bounds developed by the aid of the Stein-Chen method (cf. Arratia, Goldstein and Gordon (1989), (1990), or Barbour, Holst and Janson (1992)) for Poisson approximation. Consequently, (9) can almost directly be applied to many of the problems where Stein-Chen method has been applied in the past. These models include problems from graph theory, extreme value theory, run and scan statistics, biomolecular sequence analysis, risk theory, reliability theory etc. Moreover, bound (9) is almost identical to the bounds offered by Boutsikas and Koutras (2001) for compound Poisson approximation (via the Kolmogorov distance). The same authors applied these bounds to scan statistics - related problems in risk theory (cf. Boutsikas and Koutras (2002)). Next, we present only two simple applications elucidating the techniques employed when applying Corollary 7.

**Example 1.** (*Compound Poisson process approximation for overlapping success runs in i.i.d. trials*) Let  $\{Z_i\}_{i \in \mathbb{Z}}$  be a sequence of i.i.d. binary trials with  $\mathbb{P}(Z_i = 1) = p$ ,  $\mathbb{P}(Z_i = 0) = q$ ,  $p + q = 1$ . We are interested in the appearances of overlapping success runs (runs of "1"s) of length  $k$  in trials  $1, 2, \dots, n$ . This model has been studied by many authors in the past, see e.g. Barbour, Holst and Janson (1992) and the relevant references therein.

Define  $X_i = \prod_{j=i}^{i+k-1} Z_j$ ,  $i = 1, 2, \dots, n - k + 1$ . The random vector  $\mathbf{X} = (X_1, X_2, \dots, X_{n-k+1}) \in \{0, 1\}^{n-k+1}$  indicates the starting points of the observed overlapping success runs. In this case  $X_i$  is dependent only on  $X_{i-k+1}, \dots, X_{i+k-1}$ , and therefore we can conveniently choose

$$B_i = \{\max\{1, i - k + 1\}, \dots, i - 1, i + 1, \dots, \min\{n - k + 1, i + k - 1\}\}.$$

With the above choice,  $c_{\mathbf{X}}(B) = 0$  and since  $F_i(x) = I(x \geq 1)$  we see that  $CP(\lambda_i, F_i) \equiv Po(\lambda_i) = Po(p^k)$  (Poisson distribution with parameter  $p^k$ ). From Corollary 7 we readily get that

$$\begin{aligned} d_{TV}(\mathcal{L}(\mathbf{X}), \prod_{i=1}^{n-k+1} Po(p^k)) &\leq 2 \sum_{i=2}^{n-k+1} \sum_{b=\max\{1, i-k+1\}}^{i-1} \left( \mathbb{P}(X_b = 1, X_i = 1) + p^{2k} \right) + \sum_{i=1}^{n-k+1} p^{2k} \\ &\leq (n - k + 1)p^k \left( 2p \frac{1 - p^{k-1}}{1 - p} + (2k - 1)p^k \right), \end{aligned}$$

and therefore  $\mathcal{L}(\mathbf{X})$  can be approximated by a Poisson Process with intensity  $p^k$  on the carrier space  $\Gamma_n$  when  $n$  is large,  $p$  is small ( $k$  is fixed) and  $np^k \rightarrow \lambda$  (the upper bound is of order  $O(p)$ ).

The above bound cannot be used when we assume that  $n \rightarrow \infty$ ,  $k \rightarrow \infty$  and  $p$  is fixed. Under these conditions, the success runs tend to occur in "clumps" (clusters of adjacent success runs). The occurrences of these clumps are rare and asymptotically independent while each clump consists of a random number of overlapping success runs. This situation readily calls for a compound Poisson approximation result. To achieve this, let  $Y_1, Y_2, \dots, Y_{n-k+1}$  represent the sizes of the clumps started at trials  $1, 2, \dots, n - k + 1$  respectively, i.e.

$$Y_i = (1 - Z_{i-1}) \sum_{r=0}^{n-i-k+1} \prod_{j=i}^{i+k+r-1} Z_j, \quad i = 2, 3, \dots, n - k + 1, \quad \text{and} \quad Y_1 = \sum_{r=0}^{n-k} \prod_{j=1}^{k+r} Z_j.$$

Obviously,  $\sum_{i=1}^{n-k+1} Y_i$  is equal to  $\sum_{i=1}^{n-k+1} X_i$ , the total number of overlapping success runs within trials  $1, 2, \dots, n$ . Instead of  $Y_i$ 's, it is more convenient to use the following r.v.'s

$$Y'_i = (1 - Z_{i-1}) \sum_{r=0}^{k-1} \prod_{j=i}^{i+k+r-1} Z_j, \quad i = 1, 2, \dots, n - k + 1,$$

which represent the *truncated* sizes of clumps ( $Y'_i \leq k$ ) starting at positions  $1, 2, \dots, n - k + 1$  (to obtain stationarity, we have also allowed the last clumps to extend further than trial  $n$ ). When  $k, n$  increase while the expected number of runs  $(n - k + 1)p^k$  remains bounded, the processes  $\mathbf{Y} = (Y_1, \dots, Y_{n-k+1})$ ,  $\mathbf{Y}' = (Y'_1, \dots, Y'_{n-k+1})$  rarely differ. This is expressed by the next inequality

$$\begin{aligned} d_{TV}(\mathcal{L}(\mathbf{Y}), \mathcal{L}(\mathbf{Y}')) &\leq \mathbb{P}(\mathbf{Y} \neq \mathbf{Y}') \leq \mathbb{P}\left(\bigcup_{i=1}^{n-2k+1} [Z_{i-1} = 0, Z_i = \dots = Z_{i+2k-1} = 1]\right) \\ &\quad + \mathbb{P}(Z_0 = \dots = Z_k = 1) + \mathbb{P}\left(\bigcup_{i=n-2k+3}^{n-k+1} [Z_{i-1} = 0, Z_i = \dots = Z_{n+1} = 1]\right) \\ &\leq (n - 2k + 1)qp^{2k} + p^{k+1} + p^{k+1}(1 - p^{k-1}) \leq (n - 2k + 1)qp^{2k} + 2p^{k+1}. \end{aligned}$$

Now, consider  $B_i \cap \Gamma_i = \{\max\{1, i - 2k + 1\}, \dots, i - 1\}$  and apply Corollary 7 to gain the inequality

$$\begin{aligned} d_{TV}\left(\mathcal{L}(\mathbf{Y}'), \prod_{i=1}^{n-k+1} CP(\lambda_i, F)\right) &\leq 2 \sum_{i=2}^{n-k+1} \sum_{b=\max\{1, i-2k+1\}}^{i-1} (\mathbb{P}(Y'_i \neq 0, Y'_b \neq 0) + \mathbb{P}(Y'_i \neq 0)\mathbb{P}(Y'_b \neq 0)) \\ &\quad + \sum_{i=1}^{n-k+1} \mathbb{P}(Y'_i \neq 0)^2, \end{aligned}$$

where  $\lambda_i = \mathbb{P}(Y'_i \neq 0)$ , and  $F(x) = \mathbb{P}(Y'_i \leq x | Y'_i \neq 0)$ ,  $x \in \mathbf{R}$  ( $c_{\mathbf{X}}(B) = 0$ ). It is easy to check that

$$\begin{aligned} \mathbb{P}(Y'_i \neq 0) &= qp^k, \quad i = 1, 2, \dots, n - k + 1, \quad \mathbb{P}(Y'_i \neq 0, Y'_b \neq 0) = 0, \quad b = i - k, \dots, i - 1, \\ \mathbb{P}(Y'_i \neq 0, Y'_b \neq 0) &\leq q^2p^{2k}, \quad b = i - 2k + 1, \dots, i - k - 1, \end{aligned}$$

and  $\lambda_i = qp^k$ ,  $F(x) = 1 - p^x$ ,  $x = 1, 2, \dots, k$ . Using the above and the triangle inequality we get

$$\begin{aligned} d_{TV}\left(\mathcal{L}(\mathbf{Y}), \prod_{i=1}^{n-k+1} CP(\lambda_i, F)\right) &\leq 2(n - k)(k - 1)q^2p^{2k} + 2(n - k)(2k - 1)q^2p^{2k} \\ &\quad + (n - k + 1)q^2p^{2k} + (n - 2k + 1)qp^{2k} + 2p^{k+1} \\ &\leq \lambda p^k(1 + (6k - 3)q) + 2p^{k+1}, \end{aligned}$$

where  $\lambda = (n - k + 1)qp^k$ . Obviously, if  $n, k \rightarrow \infty$  so that  $(n - k + 1)qp^k \rightarrow \lambda_0 \in (0, \infty)$  then the upper bound vanishes and the law of the clump process  $\mathbf{Y}$  can be approximated by a compound Poisson process (the convergence rate being of order  $O(kp^k)$ ). Therefore, according to Remark

3 above, we can obtain weak convergence results (as  $n, k \rightarrow \infty$ ) for any functional  $f(\mathbf{Y})$  of the process  $\mathbf{Y}$ . For example, choosing  $f(\mathbf{y}) = \sum y_i$ , the distance  $d_{TV}(\mathcal{L}(\sum Y_i), CP(\lambda, F))$  is bounded above by the same quantity  $\lambda p^k(1 + (6k - 3)q) + 2p^{k+1}$  and therefore,  $\sum Y_i$  (the total number of overlapping success runs) follows asymptotically (as  $n, k \rightarrow \infty$ ) a compound Poisson with geometric compounding distribution (Pólya-Aeppli distribution). Note though that for this special case better bounds can be obtained via the Stein-Chen method that include the so called "magic factor" (cf. e.g. Barbour, Chryssaphinou, and Vaggelatou (2001)).

**Example 2.** (*Compound Poisson process approximation for the total excess amount above a high threshold for moving sums of i.i.d. r.v.'s*). In this example we are interested in the exceedances of the moving sum ( $r$ -scan process)

$$S_i = \sum_{j=i}^{i+r-1} X_j, i = 1, 2, \dots, n - r + 1,$$

above a threshold  $b$  where  $X_1, X_2, \dots, X_n$  is a sequence of i.i.d. non-negative unbounded random variables with a common distribution function  $F$ . More specifically we are interested in the process of excess values ("peaks over threshold  $b$ ")

$$Y_i = \max\{S_i - b, 0\}, i = 1, 2, \dots, n - r + 1.$$

Obviously,  $Y_i$ 's are locally dependent and if we choose  $b$  to be "high" (so that  $Y_i$ 's are rarely nonzero) then it is clear that the process of excess values  $\mathbf{Y}$  can be approximated by an appropriate compound Poisson process. Dembo and Karlin (1992) studied the number of exceedances  $\sum_{i=1}^{n-r+1} I(S_i > b)$  and proved (using the Stein-Chen method) that, under appropriate conditions, the number of exceedances converges to a Poisson distribution.

In a more general setup, Rootzen, Leadbetter and De Haan (1998) considered strongly mixing stationary sequences  $\{X_i\}$  and offered results pertaining to the asymptotic distribution of tail array sums of the general form  $\sum \psi(X_i - b)$  for a class of real functions  $\psi$  (which includes the case  $\psi(x) = \max\{0, x\}$  considered above). They proved that, under appropriate conditions, tail array sums converge to a compound Poisson distribution (for very high levels of  $b$ ). Note though, that their approach does not provide any bounds or convergence rates while the parameters of the limiting  $CP(\lambda, G)$  were not explicitly described.

Boutsikas and Koutras (2001) proved that the sum of excess values converges to a compound Poisson distribution. Here, following essentially the same steps, we employ Corollary 7 to obtain a compound Poisson process approximation for the law of the process of excess values  $\mathbf{Y} = (Y_i)_{i \in \Gamma_{n-r+1}}$ . As in Boutsikas and Koutras (2001), we first choose the left neighborhoods of dependence  $B_i \cap \Gamma_i = \{\max\{i - r + 1, 1\}, \dots, i - 1\}$  and then apply Corollary 7 to get

$$d_{TV} \left( \mathcal{L}(\mathbf{Y}), \prod_{i=1}^{n-r+1} CP(\lambda_i, G_b) \right) \leq 2\lambda \sum_{m=2}^r \mathbb{P}(S_m > b | S_1 > b) + 2\lambda r(1 - F^{(r)}(b)) := \lambda \epsilon(r, b), \quad (10)$$

where  $\lambda_i = \mathbb{P}(S_i > b) = (1 - F^{(r)}(b))$ ,  $\lambda = (n - r + 1)(1 - F^{(r)}(b))$  and  $F^{(m)}$  denotes the  $m$ -fold convolution of  $F$ . The compounding distribution  $G_b$  is given by the expression

$$G_b(x) = \mathbb{P}(S_i \leq x + b | S_i > b) = 1 - \frac{1 - F^{(r)}(b + x)}{1 - F^{(r)}(b)}, x \geq 0.$$

It can be shown (cf. Theorem 3 of Dembo and Karlin (1992)) that  $\epsilon(r, b) \rightarrow 0$  for any fixed  $r$ , provided that for each constant  $K > 0$ ,

$$\frac{1 - F(b - K)}{1 - F^{(2)}(b)} \rightarrow 0 \quad \text{as } b \rightarrow \infty. \quad (11)$$

According to Dembo and Karlin (1992), condition (11) holds true for any d.f.  $F$  which is a finite or infinite convolution of exponentials of any scale parameters or has a log concave density. Therefore, if the common d.f.  $F$  of  $X_i$ 's satisfies (11) then the law of the process  $\mathbf{Y}$  of excess values can be approximated by a compound Poisson process and, as in the previous cases, we may establish weak convergence for any functional of the process  $\mathbf{Y}$ . For example, we can get that the sum of excess values  $\sum_{i=1}^{n-r+1} Y_i$  converges to a compound Poisson distribution  $CP(\lambda, G)$  with a convergence rate given by (10) provided that  $n, b \rightarrow \infty$  ( $r$  is fixed) so that  $n(1 - F^{(r)}(b)) \rightarrow \lambda \in (0, \infty)$  and  $G_b(x) \rightarrow_{b \rightarrow \infty} G(x)$ .

## References

- [1] Arratia, R., Goldstein, L. and Gordon, L. (1989) Two moments suffice for Poisson approximations: the Chen-Stein method. *The Annals of Probability* **17**, 9-25.
- [2] Arratia, R., Goldstein, L. and Gordon, L. (1990) Poisson Approximation and the Chen-Stein method. *Statistical Science* **5**, 403-434.
- [3] Barbour A.D. and Chryssaphinou O. (2001) Compound Poisson approximation: a user's guide. *The Annals of Applied Probability* **11**, No 3, 964 -1002.
- [4] Barbour, A.D., Chryssaphinou, O. and E. Vaggelatou, E. (2001) Applications of compound Poisson approximation. In: *Probability and Statistical Models with Applications*, Eds. Ch. A. Charalambides, M. V. Koutras and N. Balakrishnan, Chapman and Hall, Boca Raton, 41-62.
- [5] Barbour A.D., Mansson M. (2002) Compound Poisson process approximation. *The Annals of Probability* **30**, 1492 -1537.
- [6] Barbour, A.D., Holst, L. and Janson, S. (1992) *Poisson Approximation*. Clarendon Press, Oxford.
- [7] Boutsikas, M.V. and Koutras, M.V. (2000) A bound for the distribution of the sum of discrete associated or negatively associated random variables. *The Annals of Applied Probability* **10**, 1137-1150.
- [8] Boutsikas, M.V. and Koutras, M.V. (2001) Compound Poisson approximation for sums of dependent random variables. In: *Probability and Statistical Models with Applications*, Eds. Ch.A. Charalambides, M.V. Koutras, N. Balakrishnan, Chapman and Hall, Boca Raton, 63-86.

- [9] Boutsikas, M.V. and Koutras, M.V. (2002) Modelling claim exceedances over thresholds. *Insurance: Mathematics and Economics* **30**, 67-83.
- [10] Boutsikas, M.V. and Vaggelatou, E. (2002) On the distance between convex-ordered random variables. *Advances in Applied Probability* **34**, 349-374.
- [11] Chen, L.H.Y., and Xia, A. (2004) Stein's method, Palm theory and Poisson process approximation. *The Annals of Probability* **32**, 2545-2569.
- [12] Chen, L.H.Y. (1975). Poisson approximation for dependent trials. *The Annals of Probability* **3**, 534-545.
- [13] Dembo, A. and Karlin, S. (1992) Poisson approximations for r-scan processes. *The Annals of Applied Probability* **2**, 329-357.
- [14] Rootzen, H. Leadbetter, M.R. and De Haan, L. (1998) On the distribution of tail array sums for strongly mixing stationary sequences. *The Annals of Applied Probability* **8**, 868-885.
- [15] Serfling, R.J. (1975) A general Poisson approximation theorem. *The Annals of Probability* **3**, 726-731.
- [16] Serfling, R.J. (1978) Some elementary results on Poisson approximations in a sequence of Bernoulli trials. *SIAM Review* **20**, 567-579.
- [17] Serfozo, R. (1986) Compound Poisson approximation for sums of random variables *The Annals of Probability* **14**, 1391-1398.
- [18] Vellaisamy, P. and Chaudhuri, B. (1999) On compound Poisson approximation for sums of random variables. *Statistics & Probability Letters* **41**, 179-189.
- [19] Wang, Y. H. (1986) Coupling methods in approximations. *The Canadian Journal of Statistics* **14**, 69-74.
- [20] Wang, Y. H. (1989) From Poisson to compound Poisson approximations. *The Mathematical Scientist* **14**, 38-49.

Michael V. Boutsikas,  
 Department of Statistics and Insurance Science,  
 University of Piraeus,  
 80, Karaoli & Dimitriou str.  
 18534, Piraeus  
 Greece

Tel 030 210 4142143  
 email: mbouts@unipi.gr  
 web page: [http://www.unipi.gr/faculty/dep\\_en.php?dep=mbouts](http://www.unipi.gr/faculty/dep_en.php?dep=mbouts)