BOUNDS FOR THE DISTRIBUTION OF TWO-DIMENSIONAL BINARY SCAN STATISTICS

MICHAEL V. BOUTSIKAS AND MARKOS V. KOUTRAS

Department of Statistics and Insurance Science
University of Piraeus
Piraeus, Greece
E-mail: mbouts@unipi.gr, mkoutras@unipi.gr

In the present article, we develop some efficient bounds for the distribution function of a two-dimensional scan statistic defined on a (double) sequence of independent and identically distributed (i.i.d.) binary trials. The methodology employed here takes advantage of the connection between the scan statistic problem and an equivalent reliability structure and exploits appropriate techniques of reliability theory to establish tractable bounds for the distribution of the statistic of interest. An asymptotic result is established and a numerical study is carried out to investigate the efficiency of the suggested bounds.

1. INTRODUCTION

Assume that a two-dimensional rectangular region $R = [0, L_1] \times [0, L_2]$ is observed and our interest is focused on the patterns in which a certain event $\mathcal{E}$ occurs in R. Let n_1 and n_2 be two positive integers and define $h_i = L_i/n_i$, $i = 1,2$. Furthermore, assume that n_1 and n_2 are large enough so that in each of the $n_1 \times n_2$ rectangular subregions

$$R_{ij} = [(i-1)h_1, ih_1] \times [(j-1)h_2, ih_2], \qquad i = 1,2,\ldots,n_1, j = 1,2,\ldots,n_2, \quad (1)$$

the event $\mathcal{E}$ may occur only once or does not occur at all. A question that comes up naturally within this setup is whether reasonable criteria providing evidence of clustering of the occurrences of $\mathcal{E}$ over R could be established.

In order to establish a probabilistic model, let us introduce the Bernoulli random variables (r.v.'s) X_{ij}, $i = 1, 2, \ldots, n_1$, $j = 1, 2, \ldots, n_2$ indicating the occurrence ($X_{ij} = 1$) or nonoccurrence ($X_{ij} = 0$) of the event $\mathcal{E}$ in subregion R_{ij}. We assume that X_{ij} are independent and identically distributed (i.i.d.) r.v.'s with success probabilities

$$q = P(X_{ij} = 1) = 1 - P(X_{ij} = 0) = 1 - p. \tag{2}$$

For $1 \le a \le n_1 - k_1 + 1$ and $1 \le b \le n_2 - k_2 + 1$, we define

$$S(a, b) = \sum_{i=a}^{a+k_1-1} \sum_{j=b}^{b+k_2-1} X_{ij} \tag{3}$$

to be the number of events in a rectangular region comprised of $k_1 \times k_2$ adjacent subregions R_{ij}, $a \le i \le a + k_1 - 1$, $b \le j \le b + k_2 - 1$ (k_1 and k_2 are positive integers). If $S(a, b)$ exceeds a preassigned value r, we will say that (at least) r events have been clustered within the inspected region. The two-dimensional (binary) discrete scan statistic is then defined as the largest number of events in any of the $(n_1 - k_1 + 1)(n_2 - k_2 + 1)$ observed regions; that is,

$$S = S_{k_1, k_2, n_1, n_2}$$
$$= \max\{S(a, b) : a = 1, 2, \ldots, n_1 - k_1 + 1, b = 1, 2, \ldots, n_2 - k_2 + 1\}. \tag{4}$$

It is clear that such a statistic can be exploited for testing the null hypothesis of uniformity against an alternative hypothesis of two-dimensional clustering that results in an increased occurrence of events in square subgrids of the rectangular arrangement of the observations.

Astronomy is probably one of the earliest sciences in which the scan statistic was fruitfully used. The search of the heavens for clustered galaxies, particles, signals, or radiation sources attracted the interest of astronomers from the very beginning. Orford [10] reviewed approaches for analyzing cosmic ray data in one and two dimensions, with an extensive discussion of the scan statistic. Two-dimensional scans play a critical role in several other scientific areas as well, for example, ecology, computer science, mechanical engineering, biosciences, image analysis, minefield detection, and pattern recognition; for a detailed review, see Glaz, Naus, and Wallenstein [8] and the references therein.

In addition to such diverse and fascinating applications, the study of scan statistics involves many sophisticated and intricate theoretical methods (e.g., Poisson and compound Poisson approximations, large deviation theory, Bonferroni-type inequalities, order statistics and spacings, Monte Carlo simulation, etc.).

The distribution of the scan statistic S can alternatively be expressed as the reliability of a two-dimensional r-within-consecutive-$k_1 \times k_2$-out-of-$n_1 \times n_2$ system. This system consists of $n_1 n_2$ independent components arranged in a square grid of size $n_1 \times n_2$ and fails if and only if there exists at least one square block of size $k_1 \times k_2$ ($1 < k_1 \le n_1$, $1 < k_2 \le n_2$) which contains at least $r \ge 2$ failed components. A typical application of this system in electrical engineering comes from the TFT

liquid crystal display technology where an increased number of nonfunctional pixels in fixed-size rectangular subgrids of the screen is considered as a criterion for rejecting the produced item.

If $R = R_{r,k_1,k_2,n_1,n_2}$ denotes the reliability of a two-dimensional r-within-consecutive-$k_1 \times k_2$-out-of-$n_1 \times n_2$ system with component failure probabilities q_{ij}, $i = 1,2,\ldots,n_1, j = 1,2,\ldots,n_2$, it is clear that $P(S_{k_1,k_2,n_1,n_2} \leq r - 1) = R_{r,k_1,k_2,n_1,n_2}$. Hence, the determination of the distribution of S is equivalent to the assessment of the reliability of a two-dimensional r-within-consecutive-$k_1 \times k_2$-out-of-$n_1 \times n_2$ system. Unfortunately, the exact reliability evaluation of this system is a formidable (or nonfeasible) task even for moderate values of k_1 and k_2 (cf. Akiba and Yamamoto [1]). Therefore, it seems reasonable to resort to approximations or bounds. Chen and Glaz [5] and Makri and Psillakis [9] used Bonferroni inequalities to establish reliability bounds for this system; Chen and Glaz [5] investigated also several product-type approximations for the distribution of the two-dimensional scan statistic S. Recently, Boutsikas and Koutras [4] employed a finite Markov chain imbedding method to derive accurate approximations for R_{r,k_1,k_2,n_1,n_2}. For a detailed review of these approaches, the interested reader may consult Glaz et al. [8].

The purpose of this article is to develop improved (upper and lower) bounds for the distribution of the discrete scan statistic S in the case of independent and identical X_{ij}'s. These bounds are extremely useful, especially for the cases where the lower and upper bounds are close to each other, therefore giving rise to narrow interval estimates for the unknown exact distribution. The consideration here of the special case of i.i.d. Bernoulli trials was for typographical convenience only. It is straightforward to verify that many of the results developed in this article could be easily adjusted to the case where the trials are independent but not necessarily identical. This task is left to the interested reader as an exercise.

The derivation of the new bounds is chiefly couched in the general results provided in a recent publication of the authors of the present article. More specifically, Boutsikas and Koutras [3] introduced several generalizations of well-known reliability bounds based on arbitrary partitions of the family of minimal path or cut sets of a coherent structure. After the introduction of the necessary definitions and notations (Section 2), we proceed to the establishment of product-type upper and lower bounds for the cumulative distribution function of the scan statistic (Section 3). In Section 4, we provide some additional bounds of a different nature (nonproduct-type), and in Section 5, a detailed numerical experimentation is carried out, which elucidates the quality of the suggested bounds. Finally, in Section 6 the asymptotic behavior of S is investigated by the aid of the results developed in the previous sections.

2. DEFINITIONS AND NOTATIONS

As already mentioned, the development of bounds for the distribution function of the scan statistic $S = S_{k_1,k_2,n_1,n_2}$ will be carried out by considering the equivalent two-dimensional r-within-consecutive-$k_1 \times k_2$-out-of-$n_1 \times n_2$ system, bounding its

reliability $R = R_{r,k_1,k_2,n_1,n_2}$ from above and below and subsequently using the transition formula $P(S < r) = R$. We will introduce now the necessary notations for the presentation that will follow. Let $I = \{(i,j) : i = 1,2,\ldots,n_1, j = 1,2,\ldots,n_2\}$ be the collection of the system's components (equipped with the lexicographic order) and denote by $X'_{ij} = 1 - X_{ij}$ the indicator r.v.'s describing the components' state [i.e., $X'_{ij} = 1$ if component (i,j) is functioning and $X'_{ij} = 0$ if component (i,j) has failed]. Then, the failure probabilities q of system's components can be expressed as

$$q = P(X'_{ij} = 0) = P(X_{ij} = 1). \tag{5}$$

In order to describe the minimal cut sets of the system, we set $N_1 = n_1 - k_1 + 1$, $N_2 = n_2 - k_2 + 1$ and introduce the sets

$$A_{ij} = \{(a,b), a = i, i+1, \ldots, i+k_1-1, b = j, j+1, \ldots, j+k_2-1\}, \qquad i,j \geq 1. \tag{6}$$

By the definition of the system, its minimal cut sets C are all the subsets of $A_{ij}, i = 1,2,\ldots,N_1, j = 1,2,\ldots,N_2$ with cardinality $|C| = r$. Therefore, the family $\mathbb{C}$ of all minimal cut sets can be expressed as

$$\mathbb{C} = \bigcup_{i=1}^{N_1}\bigcup_{j=1}^{N_2} \mathbb{A}_{ij}, \tag{7}$$

where

$$\mathbb{A}_{ij} = \{C : C \subseteq A_{ij} \text{ and } |C| = r\}.$$

A more convenient description of $\mathbb{C}$ is achieved by considering the disjoint families of cut sets

$$\begin{aligned}
\mathbb{C}_{ij} &= \mathbb{A}_{ij}\backslash(\mathbb{A}_{i+1,j} \cup \mathbb{A}_{i,j+1}), & i < N_1 \text{ and } j < N_2,\\
\mathbb{C}_{i,N_2} &= \mathbb{A}_{i,N_2}\backslash\mathbb{A}_{i+1,N_2}, & i < N_1,\\
\mathbb{C}_{N_1,j} &= \mathbb{A}_{N_1,j}\backslash\mathbb{A}_{N_1,j+1}, & j < N_2,\\
\mathbb{C}_{N_1,N_2} &= \mathbb{A}_{N_1,N_2}
\end{aligned} \tag{8}$$

and observing that the next equality is true as well:

$$\mathbb{C} = \bigcup_{i=1}^{N_1}\bigcup_{j=1}^{N_2} \mathbb{C}_{ij}. \tag{9}$$

If $\mathbb{A}$ is a collection of subsets of I, the symbol $R(\mathbb{A})$ will denote the reliability of a system which has as cut sets all the elements of $\mathbb{A}$; that is,

$$R(\mathbb{A}) = E\left(\prod_{A \in \mathbb{A}}\left(1 - \prod_{(i,j) \in A}(1 - X'_{ij})\right)\right). \tag{10}$$

Finally, we use the notation

$$F_{i,m}^c = \sum_{x=\max\{i,0\}}^{m} \binom{m}{x} q^x p^{m-x} \tag{11}$$

for the tail probabilities of the binomial distribution.

3. PRODUCT-TYPE BOUNDS

One of the earliest approaches to the problem of developing product-type reliability bounds may be attributed to Esary and Proschan [6], who suggested the simple lower bound

$$L_{\mathrm{EP}} = \prod_{C \in \mathbb{C}} \left(1 - \prod_{a \in C} q_a \right) \tag{12}$$

(see also Barlow and Proschan [2]). A direct application of this result in the special case of a two-dimensional-r-within-consecutive-$k_1 \times k_2$-out-of-$n_1 \times n_2$ system yields Proposition 1.

PROPOSITION 1: *The cumulative probability $P(S \leq r-1) = R_{r,k_1,k_2,n_1,n_2}$ of the scan statistic S is bounded below by*

$$L_{\mathrm{EP}} = (1 - q^r)^{|\mathbb{C}|}, \tag{13}$$

where

$$
\begin{aligned}
|\mathbb{C}| = {}&(n_1 - k_1)(n_2 - k_2) \\
&\times \left(\binom{k_1 k_2}{r} - \binom{(k_1 - 1)k_2}{r} - \binom{k_1(k_2 - 1)}{r} + \binom{(k_1 - 1)(k_2 - 1)}{r} \right) \\
&+ (n_1 - k_1) \left(\binom{k_1 k_2}{r} - \binom{(k_1 - 1)k_2}{r} \right) \\
&+ (n_2 - k_2) \left(\binom{k_1 k_2}{r} - \binom{k_1(k_2 - 1)}{r} \right) + \binom{k_1 k_2}{r}.
\end{aligned}
\tag{14}
$$

PROOF: Observe first that for each cut set $C \in A_{ij}$, we have

$$\prod_{(i,j) \in C} q_{ij} = \prod_{(i,j) \in C} q = q^{|C|} = q^r \tag{15}$$

and, therefore,

$$L_{\mathrm{EP}} = \prod_{C \in \mathbb{C}} (1 - q^r) = (1 - q^r)^{|\mathbb{C}|}. \tag{16}$$

The cardinality of $\mathbb{C}$ can be easily evaluated if we note that, in view of (9),

$$|\mathbb{C}| = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} |\mathbb{C}_{ij}| \tag{17}$$

and then proceed to a careful enumeration of the elements of each family $\mathbb{C}_{ij}$. ∎

The lower bound of Proposition 1 is pleasing in its simplicity; however, it succeeds in approximating satisfactorily the exact cumulative probability only for low values of q. A better approximation may be established by using Proposition 2.

PROPOSITION 2: *The cumulative probability* $P(S \le r-1) = R_{r,k_1,k_2,n_1,n_2}$ *of the scan statistic S is bounded below by*

$$L_{EP}^{(G)} = (1-Q_1)^{(n_1-k_1)(n_2-k_2)}(1-Q_2)^{n_1-k_1}(1-Q_3)^{n_2-k_2}(1-Q_4), \tag{18}$$

where

$$Q_1 = F_{r,k_1k_2}^c - p^{k_2}F_{r,(k_1-1)k_2}^c - p^{k_1}F_{r,k_1(k_2-1)}^c + p^{k_1+k_2-1}F_{r,(k_1-1)(k_2-1)}^c, \tag{19}$$

$$Q_2 = F_{r,k_1k_2}^c - p^{k_2}F_{r,(k_1-1)k_2}^c, \quad Q_3 = F_{r,k_1k_2}^c - p^{k_1}F_{r,k_1(k_2-1)}^c, \quad Q_4 = F_{r,k_1k_2}^c. \tag{20}$$

PROOF: Since $\mathbb{C}_{ij}, i = 1,2,\ldots,N_1, j = 1,2,\ldots,N_2$, are disjoint and satisfy (9), they form a partition of the family $\mathbb{C}$ of system's minimal cut sets. As Boutsikas and Koutras [3] have proven (cf. their Theorem 2), the quantity

$$L_{EP}^{(G)} = \prod_{i=1}^{N_1} \prod_{j=1}^{N_2} R(\mathbb{C}_{ij}) \tag{21}$$

will consist a lower bound of system's reliability. We observe next that each $R(\mathbb{C}_{ij})$ can be expressed in terms of the tail probabilities of appropriate binomial distributions. More specifically, on introducing the events

$$W_{ij} = \left[\sum_{a=i}^{i+k_1-1} \sum_{b=j}^{j+k_2-1} X_{ab} \ge r \right],$$

$$H_{ij} = \left[\sum_{b=j}^{j+k_2-1} X_{ib} = 0 \right], \quad V_{ij} = \left[\sum_{a=i}^{i+k_1-1} X_{aj} = 0 \right] \tag{22}$$

and the index sets

$$\begin{aligned}
T_1 &= \{(i,j), & i &= 1,2,\ldots,N_1-1, \ j = 1,2,\ldots,N_2-1\}, \\
T_2 &= \{(i,N_2), & i &= 1,2,\ldots,N_1-1\}, \\
T_3 &= \{(N_1,j), & j &= 1,2,\ldots,N_2-1\}, \\
T_4 &= \{(N_1,N_2)\},
\end{aligned} \tag{23}$$

we can write

$$
\begin{aligned}
1 - R(\mathbb{C}_{ij}) &= P(W_{ij} H_{ij}^c V_{ij}^c) = P(W_{ij}) - P(W_{ij} H_{ij}) \\
&\quad - P(W_{ij} V_{ij}) + P(W_{ij} H_{ij} V_{ij}) = Q_1, &&(i,j) \in T_1, \\
1 - R(\mathbb{C}_{ij}) &= P(W_{ij} H_{ij}^c) = P(W_{ij}) - P(W_{ij} H_{ij}) = Q_2, &&(i,j) \in T_2, \\
1 - R(\mathbb{C}_{ij}) &= P(W_{ij} V_{ij}^c) = P(W_{ij}) - P(W_{ij} V_{ij}) = Q_3, &&(i,j) \in T_3, \\
1 - R(\mathbb{C}_{ij}) &= P(W_{ij}) = Q_4, &&(i,j) \in T_4.
\end{aligned}
\tag{24}
$$

The desired result can now be easily verified by substituting all $R(\mathbb{C}_{ij})$ in (21). ∎

So far, the cumulative probabilities of the scan statistic S were lower bounded by computationally tractable bounds. A natural question at this point is whether upper bounds could also be established by exploiting the connection between the scan statistic distribution and the reliability of a two-dimensional r-within-consecutive-$k_1 \times k_2$-out-of-$n_1 \times n_2$ system. Unfortunately, should we wish to calculate the Esary and Proschan [6] upper reliability bounds, we need to establish first all of the minimal path sets of the system. The computational complexity of the last task is huge and, therefore, such an approach seems quite intricate. Fu and Koutras [7] developed a set of upper reliability bounds that are based on the minimal cut sets of a coherent structure and offer a satisfactory counterpart for the Esary and Proschan lower bounds. In order to apply these, one has to consider a certain ordering (arrangement) of the system's minimal cut sets and introduce appropriate index sets. If L_C denotes the index set associated with the minimal cut set $C \in \mathbb{C}$ (for the details on the construction of it, see Fu and Koutras [7]), the upper bounds take on the form

$$
L_{\mathrm{FK}} = \prod_{C \in \mathbb{C}} \left(1 - \left(\prod_{a \in L_C} p_a \right) \left(\prod_{b \in C} q_b \right) \right).
\tag{25}
$$

In the case of a two-dimensional r-within-consecutive-$k_1 \times k_2$-out-of-$n_1 \times n_2$ system, a simple arrangement (linearization) of the family $\mathbb{C}$ of minimal cut sets is established as follows: First place (in any order) the $|\mathbb{C}_{11}|$ cut sets of the family $\mathbb{C}_{11}$; next, place (in any order) the $|\mathbb{C}_{12}|$ cut sets of the family $\mathbb{C}_{12}$; and so on (use lexicographic ordering for the indices (i,j) in $\mathbb{C}_{ij}$ …; place at the end (in any order) the $|\mathbb{C}_{N_1 N_2}|$ cut sets of the family $\mathbb{C}_{N_1 N_2}$. It can be easily ascertained that the family of index sets

$$
\begin{aligned}
L_{C_i} = (\{(u,v) : u &= \max\{a - k_1 + 1, 1\}, \ldots, a + k_1 - 2, \\
v &= \max\{b - k_2 + 1, 1\}, \ldots, \min\{b + 2k_2 - 2, n_2\}\} \\
&\cup \{(a + k_1 - 1, v) : v = \max\{b - k_2 + 1, 1\}, \ldots, b + k_2 - 1\}) \backslash C_i \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad (a, b : C_i \in \mathbb{C}_{ab}),
\end{aligned}
\tag{26}
$$

$i = 2, 3, \ldots, |\mathbb{C}|, L_{C_1} = \emptyset$, satisfies the conditions required in the Fu and Koutras [7] approach. Thus, in the i.i.d. case, we get the upper bound

$$U_{\text{FK}} = \prod_{j=1}^{|\mathbb{C}|} (1 - p^{|L_{C_j}|} q^r), \tag{27}$$

and taking into account that

$$|L_{C_j}| \leq (2k_1 - 2)(3k_2 - 2) + 2k_2 - 1 - r, \tag{28}$$

we easily arrive at the following result.

PROPOSITION 3: *The cumulative probability* $P(S \leq r - 1) = R_{r, k_1, k_2, n_1, n_2}$ *of the scan statistic S is bounded above by*

$$U_{\text{FK}} = (1 - p^{(2k_1 - 2)(3k_2 - 2) + 2k_2 - 1 - r} q^r)^{|\mathbb{C}|}. \tag{29}$$

A substantially better upper bound is provided by the next proposition, which makes use of the generalized Fu and Koutras upper bound given in Boutsikas and Koutras [3].

PROPOSITION 4: *The cumulative probability* $P(S \leq r - 1) = R_{r, k_1, k_2, n_1, n_2}$ *of the scan statistic S is bounded above by*

$$U_{\text{FK}}^{(G)} = \prod_{i=1}^{N_1-1} \prod_{j=1}^{N_2-1} (1 - P_{ij} Q_1) \prod_{i=1}^{N_1-1} (1 - P_{i, N_2} Q_2) \prod_{j=1}^{N_2-1} (1 - P_{N_1, j} Q_3)(1 - P_{N_1, N_2} Q_4), \tag{30}$$

where $Q_i, i = 1, 2, 3, 4$ *are given as in Proposition 2, and*

$$P_{ij} = R(\mathbb{C}_{ij}^{**}) = E\left(\prod_{C \in \mathbb{C}_{ij}^{**}} \left(1 - \prod_{(a,b) \in C} X_{ab} \right) \right). \tag{31}$$

The quantities $\mathbb{C}_{ij}^{**}$ *are defined as*

$$\mathbb{C}_{ij}^{**} = \text{minimal}\{C \backslash A_{ij} : C \cap A_{ij} \neq \emptyset, C \in \cup_{(a,b)<(i,j)} \mathbb{C}_{ab}\}, \tag{32}$$

where A_{ij} *are defined in (6) and* $\text{minimal}(\mathbb{A})$ *denotes the set of all minimal sets of* $\mathbb{A}$ *with respect to the partial ordering induced by the inclusion operator* $\subset$.

PROOF: Recalling once more the partition described in the proof of Proposition 2 and applying Corollary 1 of Boutsikas and Koutras [3] (the optimum choice of the required binary functions a_s should be invoked), we readily obtain the upper bound

$$U_{\text{FK}}^{(G)} = \prod_{i=1}^{N_1} \prod_{j=1}^{N_2} (1 - R(\mathbb{C}_{ij}^{**})(1 - R(\mathbb{C}_{ij}))). \tag{33}$$

The evaluation of the quantities $R(\mathbb{C}_{ij})$ has already been discussed in Proposition 2. ∎

Because the evaluation of P_{ij} is, in general, tedious, a convenient lower bound for these quantities would be quite helpful. Such a bound can be easily constructed if we observe that, on introducing the index sets

$$
\begin{aligned}
K_{ij} = \{&(u,v): u = \max\{i - k_1 + 1, 1\}, \ldots, i - 1, \\
&v = \max\{j - k_2 + 1, 1\}, \ldots, \min\{j + 2k_2 - 2, n_2\}\} \\
\cup \{&(u,v): u = i, \ldots, i + k_1 - 1, v = \max\{j - k_2 + 1, 1\}, \ldots, j - 1\} \\
\cup \{&(u,v): u = i, \ldots, i + k_1 - 2, v = j + k_2, \ldots, \min\{j + 2k_2 - 2, n_2\}\},
\end{aligned} \quad \textbf{(34)}
$$

we can write the inequality

$$
\prod_{C \in \mathbb{C}_{ij}^{**}} \left(1 - \prod_{(u,v) \in C} (1 - X'_{uv})\right) \geq \prod_{(u,v) \in K_{ij}} X'_{uv}, \quad \textbf{(35)}
$$

which guarantees the validity of the inequality

$$
P_{ij} \geq \prod_{(u,v) \in K_{ij}} p = p^{|K_{ij}|}. \quad \textbf{(36)}
$$

Therefore, it can be easily verified that $P_{ij} \geq p^{(k_1-1)(3k_2-2)+k_1(k_2-1)+(k_1-1)(k_2-1)}$ and a more detailed investigation reveals Proposition 5.

PROPOSITION 5: *The cumulative probability* $P(S \leq r - 1) = R_{r,k_1,k_2,n_1,n_2}$ *of the scan statistic S is bounded above by*

$$
\begin{aligned}
U_{\text{FK}}^{(G)} \leq &(1 - Q_1)(1 - p^{(k_1-1)(3k_2-2)+(2k_1-1)(k_2-1)}Q_1)^{(n_1-k_1-1)(n_2-k_2-1)} \\
&\times (1 - p^{k_1(k_2-1)}Q_1)^{n_2-k_2-1}(1 - p^{(k_1-1)(2k_2-1)+(k_1-1)(k_2-1)}Q_1)^{n_1-k_1-1} \\
&\times (1 - p^{(k_1-1)(2k_2-1)+k_1(k_2-1)}Q_2)^{n_1-k_1} \\
&\times (1 - p^{(k_1-1)(3k_2-2)+k_1(k_2-1)+(k_1-1)(k_2-1)}Q_3)^{n_2-k_2} \\
&\times (1 - p^{(k_1-1)(2k_2-1)+k_1(k_2-1)}Q_4),
\end{aligned} \quad \textbf{(37)}
$$

where $Q_i, i = 1,2,3,4$ *are given in Proposition 2.*

For small r's, the right-hand side and the left-hand side of formula (36) are close together, whereas for large r's the actual value of P_{ij} exhibits strong deviations from the right-hand-side term. Consequently, the upper bound in (37) is sharp only when r is relatively small. A better bound could be established though by conducting a careful inspection of the set $\mathbb{C}_{ij}^{**}$ for specific values of k_1, k_2, and r. For example, for $r = 2$, we may derive the following slightly better bound:

$$
\begin{aligned}
U_{\text{FK}}^{(G)} \leq &(1 - Q_1)(1 - p^{(k_1-1)(3k_2-2)+k_1(k_2-1)}Q_1)^{(n_1-k_1-1)(n_2-k_2-1)} \\
&\times (1 - p^{k_1(k_2-1)}Q_1)^{n_2-k_2-1}(1 - p^{(k_1-1)(2k_2-1)}Q_1)^{n_1-k_1-1} \\
&\times (1 - p^{(k_1-1)(2k_2-1)+k_1(k_2-1)}Q_2)^{n_1-k_1}(1 - p^{(k_1-1)(3k_2-2)+k_1(k_2-1)}Q_3)^{n_2-k_2} \\
&\times (1 - p^{(k_1-1)(2k_2-1)+k_1(k_2-1)}Q_4).
\end{aligned} \quad \textbf{(38)}
$$

4. COVARIANCE BOUNDS

In this section, we develop an upper bound different in nature than the bounds developed so far. The key tool for the construction of this bound is a general result given by Boutsikas and Koutras [3]; since the latter involves covariances of appropriate discrete random variables, the resulting upper bound will be referred to as the covariance bound. To start with, let us denote by B_{ab} the set of all components (i,j) such that $(i,j) < (a,b)$ (use lexicographic ordering) and $A_{ij} \cap A_{ab} \neq \varnothing$ (so that $R(\mathbb{C}_{ij} \cup \mathbb{C}_{ab}) = R(\mathbb{C}_{ij})R(\mathbb{C}_{ab})$ for $(i,j) < (a,b), (i,j) \notin B_{ab}$). It can be easily verified that B_{ab} can be decomposed as $B_{ab} = \cup_{i=1}^{4} B_{ab}^{i}$, where

$$B_{ab}^{1} = \{(i,j) : i = \max\{1, a - k_1 + 1\}, \ldots, a - 1, \ j = \max\{1, b - k_2 + 1\}, \ldots, b - 1\},$$

$$B_{ab}^{2} = \{(a,j) : j = \max\{1, b - k_2 + 1\}, \ldots, b - 1\},$$

$$B_{ab}^{3} = \{(i,b) : i = \max\{1, a - k_1 + 1\}, \ldots, a - 1\},$$

$$B_{ab}^{4} = \{(i,j) : i = \max\{1, a - k_1 + 1\}, \ldots, a - 1, \ j = b + 1, \ldots, \min\{N_1, b + k_2 - 1\}\},$$

$$(39)$$

$a = 1, 2, \ldots, N_1$ and $b = 1, 2, \ldots, N_2$ (note that B_{ab}^{i} may occasionally be empty). Finally, let us denote by $F(x, y, z)$ the sum

$$F(x, y, z) = \sum_{w=0}^{z} F_{r-w, x-z}^{c} F_{r-w, y-z}^{c} \binom{z}{w} q^{w} p^{z-w}, \tag{40}$$

and by **i** the quadruple (i_1, i_2, i_3, i_4).

PROPOSITION 6: *The cumulative probability* $P(S \leq r - 1) = R_{r, k_1, k_2, n_1, n_2}$ *of the scan statistic S is bounded above by*

$$U_{CB}^{\prime(G)} = (1 - Q_1)^{N_1 N_2} + \sum_{m=1}^{4} S_m, \tag{41}$$

where Q_1 *is given as in Proposition 2,*

$$S_j = \sum_{a=a_0}^{k_2 - 1} \sum_{b=b_0}^{k_1 - 1} (N_1 - k_1 + b)(N_2 - k_2 + a) \left(\sum_{\mathbf{i} \in \{0,1\}^4} (-1)^{\sum_{k=1}^{4} i_k} F_{\mathbf{i}}^{(j)}(a, b) - Q_1^2 \right),$$

$$j = 1, 4, \tag{42}$$

$$S_2 = \sum_{a=a_0}^{k_2 - 1} N_1(N_2 - k_2 + a) \left(\sum_{\mathbf{i} \in \{0,1\}^4} (-1)^{\sum_{k=1}^{4} i_k} F_{\mathbf{i}}^{(2)}(a) - Q_1^2 \right), \tag{43}$$

$$S_3 = \sum_{b=b_0}^{k_1 - 1} N_2(N_1 - k_1 + b) \left(\sum_{\mathbf{i} \in \{0,1\}^4} (-1)^{\sum_{k=1}^{4} i_k} F_{\mathbf{i}}^{(3)}(b) - Q_1^2 \right), \tag{44}$$

with

$$a_0 = \max\{2k_2 - n_2, 1\}, \qquad b_0 = \max\{2k_1 - n_1, 1\}, \tag{45}$$

and

$$F_{\mathbf{i}}^{(1)}(a, b) = F((k_1 - i_2)(k_2 - i_1) - i_3 b - i_4 a$$
$$+ i_3 i_4, (k_1 - i_4)(k_2 - i_3), (a - i_3)(b - i_4))$$
$$\times p^{i_1 k_1 + i_2 k_2 - i_1 i_2 + i_3 k_1 + i_4 k_2 - i_3 i_4}, \tag{46}$$

$$F_{\mathbf{i}}^{(2)}(a) = F((k_1 - i_2)(k_2 - i_1 - i_3) - i_4(1 - i_2)(a - i_3), (k_1 - i_4)(k_2 - i_3)$$
$$- i_2(1 - i_4)(a - i_3), (k_1 - i_2)(a - i_3) - i_4(1 - i_2)(a - i_3))$$
$$\times p^{i_1 k_1 + i_2 k_2 - i_1 i_2 + i_3 k_1 + i_4 k_2 - i_3 i_4 - i_2 i_3 - i_2 i_4 a + i_2 i_3 i_4}, \tag{47}$$

$$F_{\mathbf{i}}^{(3)}(b) = F((k_2 - i_2)(k_1 - i_1 - i_3) - i_4(1 - i_2)(b - i_3), (k_2 - i_4)(k_1 - i_3)$$
$$- i_2(1 - i_4)(b - i_3), (k_2 - i_2)(b - i_3) - i_4(1 - i_2)(b - i_3))$$
$$\times p^{i_1 k_2 + i_2 k_1 - i_1 i_2 + i_3 k_2 + i_4 k_1 - i_3 i_4 - i_2 i_3 - i_2 i_4 b + i_2 i_3 i_4}, \tag{48}$$

$$F_{\mathbf{i}}^{(4)}(a, b) = F((k_1 - i_2)(k_2 - i_1) - i_3(b - i_2), (k_1 - i_4)(k_2 - i_3)$$
$$- i_2(a - i_3), (a - i_3)(b - i_2))$$
$$\times p^{i_1 k_1 + i_2 k_2 - i_1 i_2 + i_3 k_1 + i_4 k_2 - i_3 i_4 - i_2 i_3}. \tag{49}$$

∎

PROOF: Applying Theorem 3 of Boutsikas and Koutras [3] for the partition $\mathcal{C}_1 = \{\mathbb{C}_{ij}, i = 1, 2, \ldots, N_1, j = 1, 2, \ldots, N_2\}$ of $\mathbb{C}$, we deduce an upper bound of the form

$$U_{CB}^{(G)} = L_{EP}^{(G)} + \sum_{a=1}^{N_1} \sum_{b=1}^{N_2} \sum_{(i,j) \in B_{ab}} (R(\mathbb{C}_{ij} \cup \mathbb{C}_{ab}) - R(\mathbb{C}_{ij}) R(\mathbb{C}_{ab})). \tag{50}$$

Because $\mathcal{C}_1$ consists of four different types of sets with similar structure (cf. (8)), the final form of the bound (50) will be rather cumbersome. In order to gain a simpler (and slightly worse of course) upper bound, we consider the system whose family of minimal cut sets is given by $\mathbb{C}' = \bigcup_{i=1}^{N_1} \bigcup_{j=1}^{N_2} \mathbb{C}'_{ij}$, where

$$\mathbb{C}'_{ij} = \mathbb{A}_{ij} \backslash (\mathbb{A}_{i+1,j} \cup \mathbb{A}_{i,j+1}), \ i = 1, 2, \ldots, N_1 \text{ and } j = 1, 2, \ldots, N_2. \tag{51}$$

Let us also denote by $R'_{r, k_1, k_2, n_1, n_2}$ the reliability of this last system. Since $\mathbb{C}'_{ij} = \mathbb{C}_{ij}$ for $i < N_1, j < N_2$ and $\mathbb{C}'_{ij} \subseteq \mathbb{C}_{ij}$ for $i = N_1$ or $j = N_2$, it is obvious that $R_{r, k_1, k_2, n_1, n_2} \leq R'_{r, k_1, k_2, n_1, n_2}$. Applying once more Theorem 3 of Boutsikas and Koutras [3] for the partition $\mathcal{C}_2 = \{\mathbb{C}'_{ij}, i = 1, 2, \ldots, N_1, j = 1, 2, \ldots, N_2\}$, we easily arrive at the upper bound for $R'_{r, k_1, k_2, n_1, n_2}$ (which is also an upper bound for $R_{r, k_1, k_2, n_1, n_2}$):

$$U_{CB}^{\prime(G)} = \prod_{a=1}^{N_1} \prod_{b=1}^{N_2} R(\mathbb{C}'_{ab}) + \sum_{a=1}^{N_1} \sum_{b=1}^{N_2} \sum_{(i,j) \in B_{ab}} (R(\mathbb{C}'_{ij} \cup \mathbb{C}'_{ab}) - R(\mathbb{C}'_{ij}) R(\mathbb{C}'_{ab})), \tag{52}$$

where, in this case, $R(\mathbb{C}'_{ab}) = 1 - Q_1$. On setting $\delta(\mathbb{A}) = \Pi_{A \in \mathbb{A}}(1 - \Pi_{i \in A}(1 - X'_i))$, the triple sum appearing in $U'^{(G)}_{\mathrm{CB}}$ takes on the form

$$
\sum_{a=1}^{N_1} \sum_{b=1}^{N_2} \sum_{(i,j) \in B_{ab}} \mathrm{Cov}(\delta(\mathbb{C}'_{ij}), \delta(\mathbb{C}'_{ab}))
$$

$$
= \sum_{a=1}^{N_1} \sum_{b=1}^{N_2} \sum_{(i,j) \in B_{ab}} \mathrm{Cov}(1 - \delta(\mathbb{C}'_{ij}), 1 - \delta(\mathbb{C}'_{ab}))
$$

$$
= \sum_{a=1}^{N_1} \sum_{b=1}^{N_2} \sum_{(i,j) \in B_{ab}} (P(\delta(\mathbb{C}'_{ij}) = 0, \delta(\mathbb{C}'_{ab}) = 0) - Q_1^2)
$$

$$
= \sum_{m=1}^{4} S_m, \tag{53}
$$

where

$$
S_m = \sum_{a=1}^{N_1} \sum_{b=1}^{N_2} \sum_{(i,j) \in B^m_{ab}} (P(\delta(\mathbb{C}'_{ab}) = 0, \delta(\mathbb{C}'_{ij}) = 0) - Q_1^2). \tag{54}
$$

For the evaluation of S_1, we observe that

$$
S_1 = \sum_{a=a_0}^{k_2-1} \sum_{b=b_0}^{k_1-1} (N_1 - k_1 + b)(N_2 - k_2 + a)
$$

$$
\times (P(\delta(\mathbb{C}'_{11}) = 0, \delta(\mathbb{C}'_{k_1-b+1, k_2-a+1}) = 0) - Q_1^2); \tag{55}
$$

the computation of $P(\delta(\mathbb{C}'_{11}) = 0, \delta(\mathbb{C}'_{k_1-b+1, k_2-a+1}) = 0)$ can be carried out by expressing the event of interest as an intersection of events $AB_1^c B_2^c B_3^c B_4^c$, where (cf. (22))

$$
A = W_{11} W_{k_1-b+1, k_2-a+1}, \qquad B_1 = V_{11}, \qquad B_2 = H_{11},
$$

$$
B_3 = V_{k_1-b+1, k_2-a+1}, \qquad B_4 = H_{k_1-b+1, k_2-a+1}. \tag{56}
$$

More specifically, by virtue of the inclusion–exclusion theorem, we can write

$$
P\left(A \cap \left(\bigcap_{i=1}^{4} B_i^c\right)\right) = P(A) - P\left(\bigcup_{i=1}^{4} AB_i\right)
$$

$$
= P(A) - \sum_{s=1}^{4} \sum_{D \subseteq \{1,2,3,4\}: |D|=s} (-1)^{s-1} P\left(\bigcap_{i \in D} AB_i\right), \tag{57}
$$

and the evaluation of the probabilities $P(\cap_{i \in D} AB_i)$ can be accomplished if we recognize that the events of interest can be expressed via i.i.d. binary r.v.'s Z_i such that $P(Z_i = 1) = q$. In particular, it is not difficult to check that, on setting

$$F(x, y, z) = P\left(\sum_{i=1}^{x} Z_i \geq r, \sum_{i=x-z+1}^{x-z+y} Z_i \geq r\right)$$

$$= \sum_{w=0}^{z} P\left(\sum_{i=1}^{x-z} Z_i \geq r - w, \sum_{i=x+1}^{x-z+y} Z_i \geq r - w\right) P\left(\sum_{i=x-z+1}^{x} Z_i = w\right) \quad \textbf{(58)}$$

(a convenient expression for that is given in (40)), we can write $P(\cap_{i \in D} A B_i) = F_{\mathbf{i}}^{(1)}(a, b)$, $\mathbf{i} = (i_1, i_2, i_3, i_4)$ with the indices i_j, $j = 1, 2, 3, 4$ being 1 or 0 according to whether $j \in D$ or $j \notin D$, respectively, and

$$F_{\mathbf{i}}^{(1)}(a, b) = F_{\mathbf{i}}^{(1)}(a, b; k_1, k_2) = p^{i_1 k_1 + i_2 k_2 - i_1 i_2 + i_3 k_1 + i_4 k_2 - i_3 i_4}$$

$$\times F((k_1 - i_2)(k_2 - i_1) - i_3 b - i_4 a + i_3 i_4,$$

$$(k_1 - i_4)(k_2 - i_3), (a - i_3)(b - i_4)). \quad \textbf{(59)}$$

Combining all the above, we deduce

$$S_1 = \sum_{a=a_0}^{k_2-1} \sum_{b=b_0}^{k_1-1} (N_1 - k_1 + b)(N_2 - k_2 + a) \left(\sum_{\mathbf{i} \in \{0,1\}^4} (-1)^{\Sigma_{k=1}^{4} i_k} F_{\mathbf{i}}^{(1)}(a, b) - Q_1^2\right).$$

$$\textbf{(60)}$$

Arguing in exactly the same way, we may also compute S_2, S_3, and S_4 (the technical details are lengthy but straightforward). ∎

5. NUMERICAL COMPARISONS

In Table 1 we provide, for selected values of r, k_1, k_2, n_1, and p, the values of the bounds described in formulas (13), (18), (29), (37) (or (38) for $r = 2$), and (41) (they have been labeled L_{EP}, L_{EP}^G, U_{FK}, U_{FK}^G, and U_{CB}^G, respectively) along with the Bonferroni-type bounds developed by Makri and Psillakis [9] (for the case $n_1 = n_2$ and $k_1 = k_2$). These bounds have been labeled as L_{Bonf} and U_{Bonf}. We also provide a simulated value (R_{simul}) and an approximate value (R_{approx}) based on the results of Boutsikas and Koutras [4].

It is worth stressing that for large systems (e.g., $n_1, n_2 > 30$), the computer times for assessing the simulated values of R are extremely large, especially when a large number of iterations are performed. This is why, in most cases, we were forced to reduce the number of iterations to 10^4 or 10^5 (only in Table 1 do we employ 10^6 iterations). This fact may have resulted in quite crude approximations of the exact reliability. However, the bounds are easily evaluated no matter how large the values of n and k are.

A careful inspection of Table 1 reveals that L_{EP}^G is the best lower bound of the three for each of the cases we consider, whereas, for very high component survival probabilities, U_{CB}^G is the best of the four upper bounds. For lower-component survival probabilities, the leading bounds are U_{FK}^G (for small r) or U_{Bonf}.

TABLE 1. Values of Bounds for Selected Values
of r, k_1, k_2, n_1, n_2, and p

p	L_{EP}	L_{EP}^{G}	L_{Bonf}	R_{approx}	R_{simul}	U_{Bonf}	U_{FK}	U_{FK}^{G}	U_{CB}^{G}
$r=2$, $k_1=3$, $k_2=3$, $n_1=30$, $n_2=30$
.982	0.0402	**0.0476**	−7.2781	0.0777	*0.0838*	0.3536	0.1604	**0.1154**	0.7995
.986	0.1431	**0.1551**	−4.1247	0.1980	*0.2071*	0.4603	0.2849	**0.2400**	0.5339
.99	0.3709	**0.3821**	−1.6731	0.4199	*0.4284*	0.6127	0.4837	**0.4514**	0.5387
.994	0.6997	**0.7043**	0.0171	0.7197	*0.7237*	0.8044	0.7436	**0.7314**	0.7478
.998	0.9611	**0.9613**	0.8886	0.9622	*0.9624*	0.9717	0.9634	**0.9628**	0.9645
.999	0.9901	**0.9902**	0.9720	0.9903	*0.9902*	0.9926	0.9904	**0.9903**	0.9908
$r=3$, $k_1=4$, $k_2=4$, $n_1=100$, $n_2=100$
.99	0.1429	**0.1669**	−3.7749	0.2080	*0.2207*	0.5003	0.3596	**0.3383**	0.5403
.992	0.3692	**0.3937**	−1.4936	0.4321	*0.4464*	0.6486	0.5511	**0.5359**	0.5497
.994	0.6568	**0.6704**	−0.0729	0.6910	*0.6959*	0.8033	0.7513	0.7439	**0.7217**
.996	0.8829	**0.8865**	0.6758	0.8920	*0.8916*	0.9281	0.9081	0.9061	**0.8973**
.998	0.9846	**0.9848**	0.9587	0.9852	*0.9850*	0.9899	0.9864	0.9862	**0.9856**
$r=6$, $k_1=3$, $k_2=3$, $n_1=1000$, $n_2=1000$
.93	0.0001	**0.0003**	−7.1775	0.0008	*0.0009*	**0.1346**	0.2581	0.2451	1.0089
.94	0.0221	**0.0386**	−2.3322	0.0522	*0.0518*	**0.2645**	0.4883	0.4772	0.3769
.95	0.2791	**0.3264**	−0.1464	0.3554	*0.3614*	0.4953	0.7265	0.7205	**0.4196**
.96	0.7157	**0.7397**	0.6913	0.7532	*0.7545*	0.7658	0.8948	0.8928	**0.7591**
.97	0.9422	**0.9464**	0.9436	0.9487	*0.9470*	0.9491	0.9742	0.9738	**0.9490**
.98	0.9948	**0.9950**	0.9949	0.9952	*0.9951*	0.9952	0.9970	0.9969	**0.9952**
$r=12$, $k_1=4$, $k_2=4$, $n_1=100$, $n_2=100$
.7	0.0001	**0.0822**	−1.5003	0.1777	*0.1862*	**0.4462**	1.	0.9994	1.4161
.72	0.0188	**0.2992**	−0.2075	0.4164	*0.4265*	**0.6067**	1.	0.9996	0.8312
.74	0.1955	**0.5789**	0.4529	0.6603	*0.6615*	**0.7515**	1.	0.9998	0.7769
.76	0.5354	**0.7944**	0.7697	0.8336	*0.8352*	0.8647	1.	0.9999	**0.8625**
.78	0.8026	**0.9149**	0.911	0.9297	*0.9292*	0.9364	1.	0.9999	**0.9363**
.8	0.9323	**0.9694**	0.9689	0.9740	*0.9737*	0.9755	1.	1.	**0.9754**
.82	0.9804	**0.9905**	0.9904	0.9917	*0.9916*	0.9919	1.	1.	**0.9919**
$r=5$, $k_1=5$, $k_2=5$, $n_1=100$, $n_2=100$
.97	0.0025	**0.0246**	−6.1884	0.0755	*0.0967*	**0.4292**	0.8001	0.7533	2.4923
.975	0.0903	**0.1993**	−2.1440	0.3018	*0.3331*	**0.6088**	0.8554	0.8288	1.0415
.98	0.4547	**0.5641**	−0.1208	0.6361	*0.6543*	0.7956	0.9149	0.9030	**0.7895**
.985	0.8294	**0.8631**	0.7107	0.8838	*0.8888*	0.9313	0.9641	0.9605	**0.9042**
.99	0.9757	**0.9792**	0.9586	0.9815	*0.9821*	0.9882	0.9917	0.9912	**0.9829**
.995	0.9992	**0.9993**	0.9986	0.9993	*0.9993*	0.9996	0.9996	0.9995	**0.9994**
$r=20$, $k_1=5$, $k_2=5$, $n_1=1000$, $n_2=1000$
.64	0.0000	**0.0002**	−7.6897	0.0014	*0.0013*	**0.1673**	1.	1.	3.3426
.66	0.0000	**0.0411**	−2.1925	0.0809	*0.0793*	**0.3328**	1.	1.	1.0427
.68	0.0013	**0.3362**	−0.0902	0.4106	*0.4066*	**0.5641**	1.	1.	0.6145
.7	0.1592	**0.7097**	0.6570	0.7489	*0.7481*	0.7869	1.	1.	**0.7807**
.72	0.6298	**0.9063**	0.9016	0.9182	*0.9120*	0.9229	1.	1.	**0.9227**
.74	0.9003	**0.9749**	0.9746	0.9777	*0.9780*	0.9783	1.	1.	**0.9783**
.76	0.9790	**0.9942**	0.9942	0.9947	*0.9945*	0.9948	1.	1.	**0.9948**

(continued)

TABLE 1. *continued*

p	L_{EP}	L^G_{EP}	L_{Bonf}	R_{approx}	R_{simul}	U_{Bonf}	U_{FK}	U^G_{FK}	U^G_{CB}
$r = 2, k_1 = 4, k_2 = 4, n_1 = 500, n_2 = 500$									
.9992	0.0221	**0.0225**	−17.827	0.0239	*0.0250*	0.2750	0.0268	**0.0254**	0.1312
.9993	0.0540	**0.0546**	−13.428	0.0570	*0.0588*	0.3301	0.0614	**0.0594**	0.1281
.9994	0.1171	**0.1179**	−9.6105	0.1211	*0.1223*	0.4005	0.1271	**0.1243**	0.1650
.9995	0.2255	**0.2264**	−6.3754	0.2300	*0.2334*	0.4892	0.2365	**0.2335**	0.2547
.9996	0.3855	**0.3863**	−3.7247	0.3894	*0.3900*	0.5985	0.3950	**0.3924**	0.4018
.9997	0.5850	**0.5855**	−1.6601	0.5875	*0.5893*	0.7251	0.5910	**0.5894**	0.5928
.9998	0.7879	**0.7882**	−0.1834	0.7890	*0.7883*	0.8554	0.7904	**0.7897**	0.7908
.9999	0.9422	**0.9422**	0.7039	0.9423	*0.9424*	0.9592	0.9425	**0.9424**	0.9427
.99995	0.9852	**0.9852**	0.9259	0.9852	*0.9852*	0.9895	0.9853	**0.9852**	0.9853
$r = 3, k_1 = 2, k_2 = 5, n_1 = 100, n_2 = 20$									
.965	0.0155	**0.0292**		0.0608	*0.0678*		0.2640	**0.2106**	1.1422
.97	0.0726	**0.1026**		0.1549	*0.1635*		0.3718	**0.3232**	0.7218
.975	0.2193	**0.2595**		0.3192	*0.3282*		0.5092	**0.4712**	0.5743
.98	0.4598	**0.4931**		0.5386	*0.5476*		0.6656	0.6418	**0.6340**
.985	0.7205	**0.7368**		0.7585	*0.7640*		0.8170	0.8063	**0.7877**
.99	0.9074	**0.9115**		0.9170	*0.9175*		0.9300	0.9292	**0.9237**
.995	0.9879	**0.9882**		0.9886	*0.9888*		0.9897	0.9895	**0.9893**
$r = 15, k_1 = 6, k_2 = 3, n_1 = 1000, n_2 = 100$									
.7	0.3197	**0.6543**		0.7071	*0.7090*		1.	1.	**0.7581**
.72	0.6669	**0.8498**		0.8723	*0.8718*		1.	1.	**0.8832**
.74	0.8752	**0.9439**		0.9515	*0.9513*		1.	1.	**0.9538**
.76	0.9607	**0.9815**		0.9837	*0.9830*		1.	1.	**0.9842**
.78	0.9892	**0.9946**		0.9951	*0.9949*		1.	1.	**0.9952**

6. ASYMPTOTIC RESULTS

Many real-life applications call for the study of the two-dimensional scan statistic S in a large square grid of size $n_1 \times n_2$ (i.e., for large n_1 and n_2) or, equivalently, for the study of huge two-dimensional r-within-consecutive-$k_1 \times k_2$-out-of-$n_1 \times n_2$ reliability structures with high-quality components. In these cases, the evaluation of lower and upper bounds, although feasible, quickly become tedious. An asymptotic estimate of the tail probability or system's reliability, can, therefore, be useful. In this section, we aim at an asymptotic result concerning the case when n_1 and n_2 converge to infinity and, at the same time, $q \to 0$ so that $n_1 n_2 q^r$ converges to a positive number. To achieve this, we may use either the bounds developed in the previous sections (e.g., L_{EP} and U_{FK}) or the general results for coherent structures presented by Fu and Koutras [7].

PROPOSITION 7: *If r, k_1, and k_2 are fixed and $n_1, n_2 \to \infty$ and $q \to 0$ such that*

$$n_1 n_2 \left(\binom{k_1 k_2}{r} - \binom{(k_1 - 1)k_2}{r} - \binom{k_1(k_2 - 1)}{r} \right.$$

$$\left. + \binom{(k_1 - 1)(k_2 - 1)}{r} \right) q^r \to \lambda \in (0, \infty), \tag{61}$$

then

$$R_{r, k_1, k_2, n_1, n_2} \to e^{-\lambda}. \tag{62}$$

PROOF: It suffices to observe that $\max |L_{C_j}|$ remains bounded as $n_1, n_2 \to \infty$ (cf. (28)) and then apply Corollary 4.2. of Fu and Koutras [7]. The desired result follows immediately, using expression (14) for $|\mathbb{C}|$. An alternative proof of the same result can also be established by using the simpler bounds L_{EP} and U_{FK} developed in the previous sections. ∎

We now present an interesting consequence of Proposition 7 relating to the asymptotic distribution of the time to failure T of a two-dimensional r-within-consecutive-$k_1 \times k_2$-out-of-$n_1 \times n_2$ reliability system. More specifically, suppose that this system starts functioning at time 0 equipped with $n_1 n_2$ new (unused) components and denote by T the time until the system's failure and by T_{ij} the lifetime of the (i, j) component. Then we have the next result.

COROLLARY 1: *Let the common lifetime distributions of the components be of the form*

$$P(T_{ij} \le t) = q(t) = \theta^a t^a + o(t^a), \qquad t \ge 0, \tag{63}$$

where a and θ are positive constants. Then, the asymptotic lifetime T of the system (properly normalized) follows a Weibull distribution with parameter

$$\theta' = \theta \left(\binom{k_1 k_2}{r} - \binom{(k_1 - 1)k_2}{r} - \binom{k_1(k_2 - 1)}{r} + \binom{(k_1 - 1)(k_2 - 1)}{r} \right)^{1/ar};$$

$$\tag{64}$$

more specifically,

$$P((n_1 n_2)^{1/ra} T > t) = R((n_1 n_2)^{-1/ra} t) \to e^{-(\theta' t)^{ar}} \tag{65}$$

as $n_1, n_2 \to \infty$.

PROOF: It is an immediate consequence of Proposition 7 because $q = q((n_1 n_2)^{-1/ra} t) \to 0$ and

$$
\begin{aligned}
n_1 n_2 q^r &= n_1 n_2 q((n_1 n_2)^{-1/ra} t)^r \\
&= n_1 n_2 (\theta^a (n_1 n_2)^{-1/r} t^a + o((n_1 n_2)^{-1/r}))^r \\
&= \left(\theta^a t^a + \frac{o((n_1 n_2)^{-1/r})}{(n_1 n_2)^{-1/r}} \right)^r \to (\theta t)^{ar}
\end{aligned}
\tag{66}
$$

as $n_1, n_2 \to \infty$. ∎

Note that the assumption made above that the components' failure distribution satisfies $q(t) = \theta^a t^a + o(t^a), t \geq 0$, is a rather mild one (it is equivalent to $\lim_{t \to 0^+} t^{-a} q(t) \in (0, \infty)$ for some $a > 0$). There are quite a few lifetime distributions which satisfy this condition, including exponential, Weibull, gamma, uniform, and so forth.

References

1. Akiba, T. & Yamamoto, H. (2001). Reliability of a two-dimensional *k*-within-consecutive-*r* × *s*-out-of-*m* × *n*:F system. *Naval Research Logistics* 48: 625–637.
2. Barlow, R.E. & Proschan, F. (1981). *Statistical theory of reliability and life testing*. Silver Spring, MD: To Begin With.
3. Boutsikas, M.V. & Koutras, M.V. (2000). Generalized reliability bounds for coherent structures. *Journal of Applied Probability* 37: 778–794.
4. Boutsikas, M.V. & Koutras, M.V. (2000). Reliability approximation for Markov chain imbeddable systems. *Methodology and Computing in Applied Probability* 2(4): 393–411.
5. Chen, J. & Glaz, J. (1996). Two-dimensional discrete scan statistics. *Statistics and Probability Letters* 31: 59–68.
6. Esary, J.D. & Proschan, F. (1963). Coherent structures of non-identical components. *Technometrics* 5: 191–209.
7. Fu, J.C. & Koutras, M.V. (1995). Reliability bounds for coherent structures with independent components. *Statistics and Probability Letters* 22: 137–148.
8. Glaz, J., Naus, J., & Wallenstein, S. (2001). *Scan statistics*. New York: Springer-Verlag.
9. Makri, F.S. & Psillakis, Z.M. (1997). Bounds for reliability of *k*-within 2-dimensional consecutive-*r*-out-of-*n* failure systems. *Microelectronics and Reliability* 37: 1217–1224.
10. Orford, K.J. (2000). The analysis of cosmic ray data. *Journal of Physics G: Nuclear and Particle Physics* 26(4): R1–R26.