J. Appl. Prob. **54,** 304–319 (2017)
*doi:*10.1017/jpr.2016.101

CORRECTED DISCRETE APPROXIMATIONS FOR THE CONDITIONAL AND UNCONDITIONAL DISTRIBUTIONS OF THE CONTINUOUS SCAN STATISTIC

YI-CHING YAO,* *Academia Sinica*
DANIEL WEI-CHUNG MIAO,** *National Taiwan University of Science and Technology*
XENOS CHANG-SHUO LIN,*** *Aletheia University*

Abstract

The (conditional or unconditional) distribution of the continuous scan statistic in a one-dimensional Poisson process may be approximated by that of a discrete analogue via time discretization (to be referred to as the discrete approximation). Using a change of measure argument, we derive the first-order term of the discrete approximation which involves some functionals of the Poisson process. Richardson's extrapolation is then applied to yield a corrected (second-order) approximation. Numerical results are presented to compare various approximations.

Keywords: Poisson process; Richardson's extrapolation; Markov chain embedding; change of measure; second-order approximation

2010 Mathematics Subject Classification: Primary 60E05; 62E20
Secondary 60J10; 60D05

1. Introduction

The subject of scan statistics in one dimension as well as in higher dimensions has found a great many applications in diverse areas ranging from astronomy to epidemiology, genetics, and neuroscience. See [8] and [9] for a thorough review and comprehensive discussion of scan distribution theory, methods, and applications. See also [10] for a collection of articles on recent developments.

In the one-dimensional setting, let Π be a (homogeneous) Poisson point process of intensity $\lambda > 0$ on the (normalized) unit interval $(0, 1]$. For a specified window size $0 < w < 1$ and integers $N \geq k \geq 2$, we are interested in finding the conditional and unconditional probabilities

$$P(k; N, w) := \mathbb{P}(S_w \geq k \mid |\Pi| = N) \quad \text{and} \quad P^*(k; \lambda, w) := \mathbb{P}(S_w \geq k),$$

where $|\Pi|$ is the cardinality of the point set Π (i.e. the total number of Poisson points) and

$$S_w = S_w(\Pi) := \max_{0 \leq t \leq 1-w} |\Pi \cap (t, t+w]|,$$

Received 27 November 2015; revision received 1 March 2016.
* Postal address: Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan, ROC.
Email address: yao@stat.sinica.edu.tw
** Postal address: Graduate Institute of Finance, National Taiwan University of Science and Technology, Taipei 106, Taiwan, ROC. Email address: miao@mail.ntust.edu.tw
*** Postal address: Accounting Information Department, Aletheia University, New Taipei City, 25103, Taiwan, ROC.
Email address: xenos.lin@gmail.com

the maximum number of Poisson points within any window of size w. The (continuous) scan statistic S_w arises from the likelihood ratio test for the null hypothesis $\mathcal{H}_0$: the intensity function $\lambda(t) = \lambda$ (constant) against the alternative $\mathcal{H}_a\colon \lambda(t) = \lambda + \Delta \mathbf{1}_{(a,a+w]}(t)$ for (unknown) $0 \le a \le 1 - w$ and $\Delta > 0$, where $\mathbf{1}_{\mathcal{A}}$ denotes the indicator function of a set $\mathcal{A}$.

By applying results on coincidence probabilities and the generalized ballot problem (see [1] and [16]), Huntington and Naus [13] and Hwang [14] derived closed-form expressions for $P(k;\ N, w)$ which involve summing up a large number of determinants of large matrices and, hence, are, in general, not amenable to numerical evaluation. Later, by exploiting the fact that $P(k; N, w)$ is a piecewise polynomial in w with (finitely many) different polynomials of w in different ranges, Neff and Naus [20] developed a more computationally feasible approach and presented extensive tables for the *exact* $P(k;\ N, w)$ for various combinations of (k, N, w) with $N \le 25$. (More precisely, each number in the tables has an error bounded by 10^{-9}.) Noting that $P^*(k; \lambda, w)$ is a weighted average of $P(k; N, w)$ over N (with Poisson probabilities as weights), they also provided tables for $P^*(k; \lambda, w)$ with $\lambda \le 16$, where the error size for each tabulated number varies depending on the combination of (k, λ, w). (The errors tend to be greater for smaller values of w.) Huffer and Lin [11], [12] developed an alternative approach (based on spacings) to computing the exact $P(k; N, w)$.

Instead of finding the exact $P^*(k; \lambda, w)$, Naus [19] proposed an accurate product-type approximation based on a heuristic (approximate) Markov property while Janson [15] derived some sharp bounds. See also [7] for related results in a discrete setting. Treating the problem as a boundary crossing for a two-dimensional random field, Loader [18] obtained effective large deviation approximations for the tail probability of the scan statistic in one and higher dimensions. For more general large deviation approximation results, see [2] and [21].

The continuous scan statistic S_w may be approximated by a discrete analogue via time discretization. Specifically, assuming $w = p/q$ (p, q relatively prime integers), partition the (time) interval $(0, 1]$ into n subintervals of length n^{-1}, n a multiple of q. Each subinterval (independently) either contains no point (with probability $1 - \lambda/n$) or exactly one point (with probability λ/n). Since a window of size w covers nw subintervals, as an approximation to S_w, we define the discrete scan statistic $S_w^{(n)}$ to be the maximum number of points within any nw consecutive subintervals. For large n, $P^*(k; \lambda, w) = \mathbb{P}(S_w \ge k)$ may be approximated by $\mathbb{P}(S_w^{(n)} \ge k)$, which can be readily calculated using the Markov chain embedding method (see [4], [5], [17]). Indeed, it is known that $\mathbb{P}(S_w \ge k) - \mathbb{P}(S_w^{(n)} \ge k) = O(n^{-1})$ (see [6], [22]).

In Section 2, as n (multiple of q) tends to ∞, we derive the limit of $n[\mathbb{P}(S_w \ge k) - \mathbb{P}(S_w^{(n)} \ge k)]$, which involves some functionals of Π. In order to establish this limit result, we find it instructive to introduce a slightly different discrete scan statistic (denoted by $S_w'^{(n)}$) which is stochastically smaller than S_w and $S_w^{(n)}$. With a coupling device, we derive the limits of $n[\mathbb{P}(S_w \ge k) - \mathbb{P}(S_w'^{(n)} \ge k)]$ and $n[\mathbb{P}(S_w^{(n)} \ge k) - \mathbb{P}(S_w'^{(n)} \ge k)]$. In Section 3, using a change of measure argument, a similar result is obtained for the conditional probability $\mathbb{P}(S_w \ge k \mid |\Pi| = N)$. Based on these limit results, Richardson's extrapolation is then applied to yield second-order approximations for the conditional and unconditional distributions of the continuous scan statistic. In Section 4, numerical results comparing the various approximations are presented along with some discussion.

2. The unconditional case

Recall the window size $w = p/q$ with p and q relatively prime integers. For $n = mq$ ($m = 1, 2, \dots$), let H_i^n, $i = 1, \dots, n$, be independent and identically distributed (i.i.d.) with $\mathbb{P}(H_i^n = 0) = 1 - \lambda/n$ and $\mathbb{P}(H_i^n = 1) = \lambda/n$, and let I_i^n, $i = 1, \dots, n$, be i.i.d. with $\mathbb{P}(I_i^n = 0) = \mathrm{e}^{-\lambda/n}$

and $\mathbb{P}(I_i^n = 1) = 1 - \mathrm{e}^{-\lambda/n}$. The i.i.d. Bernoulli sequence $(H_1^n, \dots, H_n^n)$ approximates the Poisson point process Π by matching the expected number of points in each subinterval, i.e.

$$\mathbb{E}(H_i^n) = \mathbb{E}\bigg(\bigg|\Pi \cap \bigg(\frac{i-1}{n}, \frac{i}{n}\bigg]\bigg|\bigg) = \frac{\lambda}{n}.$$

On the other hand, the i.i.d. Bernoulli sequence $(I_1^n, \dots, I_n^n)$ approximates Π by matching the probability of no point in each subinterval, i.e.

$$\mathbb{P}(I_i^n = 0) = \mathbb{P}\bigg(\bigg|\Pi \cap \bigg(\frac{i-1}{n}, \frac{i}{n}\bigg]\bigg| = 0\bigg) = \mathrm{e}^{-\lambda/n}.$$

The two discrete scan statistics $S_w^{(n)}$ and $S_w^{\prime(n)}$ are now defined in terms of the two Bernoulli sequences as

$$S_w^{(n)} = S_{w,H}^{(n)} := \max_{i=1,\dots,n\overline{w}+1} \sum_{r=i}^{i+nw-1} H_r^n, \qquad S_w^{\prime(n)} = S_{w,I}^{(n)} := \max_{i=1,\dots,n\overline{w}+1} \sum_{r=i}^{i+nw-1} I_r^n,$$

where $\overline{w} := 1 - w$. Since I_i^n is stochastically smaller than H_i^n and $|\Pi \cap ((i-1)/n, i/n]|$, it follows that $S_{w,I}^{(n)}$ is stochastically smaller than S_w and $S_{w,H}^{(n)}$. In Sections 2.1 and 2.2 we derive $\lim_{n\to\infty} n[\mathbb{P}(S_w \ge k) - \mathbb{P}(S_{w,I}^{(n)} \ge k)]$ and $\lim_{n\to\infty} n[\mathbb{P}(S_w \ge k) - \mathbb{P}(S_{w,H}^{(n)} \ge k)]$, respectively.

2.1. Matching the probability of no point

Since the Bernoulli sequence $(I_1^n, \dots, I_n^n)$ and Π match in the probability of no point in each subinterval, it is instructive to define $(I_1^n, \dots, I_n^n)$ in terms of Π by

$$I_i^n = \mathbf{1}\bigg\{\Pi \cap \bigg(\frac{i-1}{n}, \frac{i}{n}\bigg] \ne \varnothing\bigg\}, \qquad i = 1, \dots, n.$$

Thus, $(I_1^n, \dots, I_n^n)$ and Π are defined on the same probability space. In particular, $S_w \ge S_{w,I}^{(n)}$ with probability 1. For fixed $w = p/q$ and for each (fixed) $k = 2, 3, \dots$, let $\alpha = \mathbb{P}(\mathcal{A})$ and $\alpha_n = \mathbb{P}(\mathcal{A}_n)$, where $\mathcal{A} = \mathcal{A}_{k,w} := \{S_w \ge k\}$ and $\mathcal{A}_n = \mathcal{A}_{n,k,w} := \{S_{w,I}^{(n)} \ge k\}$.

Note that $\alpha = P^*(k; \lambda, w)$, defined in Section 1. In order to derive the limit of $n(\alpha - \alpha_n)$ as $n \to \infty$, we need to introduce some functionals of Π. Let $M := |\Pi|$, which is a Poisson random variable with mean λ. Writing $\Pi = \{Q_1, \dots, Q_M\}$, assume (with probability 1) that $0 < Q_1 < \cdots < Q_M < 1$. Furthermore, assume (with probability 1) that $w \notin \Pi, \overline{w} = 1 - w \notin \Pi$, and $Q_j \pm w \notin \Pi$ for $j = 1, \dots, M$ (i.e. $Q_j - Q_i \ne w$ for all $1 \le i < j \le M$). Define the functionals $\nu(\Pi) = \nu(\{Q_1, \dots, Q_M\})$ and $\tilde{\nu}(\Pi) = \tilde{\nu}(\{Q_1, \dots, Q_M\})$ as

$$\nu(\Pi) := \sum_{\{\ell\colon Q_\ell < 1-w\}} \mathbf{1}\{S_w < k, |\Pi \cap (Q_\ell, Q_\ell + w]| = k-2,$$
$$|\Pi \cap (t, t+w]| \le k-2 \text{ for all } t \text{ with } Q_\ell \le t \le Q_\ell + w\},$$

$$\tilde{\nu}(\Pi) := \sum_{\ell=1}^{M} \mathbf{1}\Big\{S_w < k, \max_{0 \le t \le 1-w} |(\Pi \cup \{Q_\ell\}) \cap (t, t+w]| = k\Big\},$$

where $\Pi \cup \{Q_\ell\}$ is interpreted as a multiset with Q_ℓ having multiplicity 2.

Theorem 2.1. *For $n = mq\,(m = 1, 2, \dots)$,*

$$\lim_{n\to\infty} n(\alpha - \alpha_n) = \tfrac{1}{2}\lambda\mathbb{E}[\nu(\Pi) + \tilde{\nu}(\Pi)].$$

Proof. Denoting the complement of $\mathcal{A}_n$ by $\mathcal{A}_n^c$ and noting that $\mathcal{A}_n \subset \mathcal{A}$, we have $\alpha - \alpha_n = \mathbb{P}(\mathcal{A}) - \mathbb{P}(\mathcal{A}_n) = \mathbb{P}(\mathcal{A} \cap \mathcal{A}_n^c)$. For $i = 1, \dots, n$, let $\tilde{I}_i^n = |\Pi \cap ((i-1)/n, i/n]|$, the number of Poisson points in the ith subinterval. Then $\tilde{I}_i^n = 0$ implies $I_i^n = 0$ and $\tilde{I}_i^n \geq 1$ implies $I_i^n = 1$. Consider the following disjoint events:

$$\mathcal{G}_1 = \{\tilde{I}_j^n \leq 1,\ j = 1, \dots, n\},$$

$$\mathcal{G}_{2,i} = \{\tilde{I}_i^n = 2,\ \tilde{I}_j^n \leq 1 \text{ for all } j \neq i\}, \qquad i = 1, \dots, n,$$

$$\mathcal{G}_3 = \{\tilde{I}_j^n = \tilde{I}_{j'}^n = 2 \text{ for some } j \neq j'\} \cup \{\tilde{I}_j^n \geq 3 \text{ for some } j\}.$$

We have

$$\alpha - \alpha_n = \mathbb{P}(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1) + \sum_{i=1}^{n} \mathbb{P}(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_{2,i}) + \mathbb{P}(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_3). \tag{2.1}$$

We claim that

$$\mathbb{P}(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1) = \frac{1}{2}\sum_{i=1}^{n\overline{w}} P_i^{(n)} + O(n^{-2}), \tag{2.2}$$

$$\sum_{i=1}^{n} \mathbb{P}(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_{2,i}) = \sum_{i=1}^{n} \widetilde{P}_i^{(n)} + O(n^{-2}), \tag{2.3}$$

$$\mathbb{P}(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_3) = O(n^{-2}), \tag{2.4}$$

where

$$P_i^{(n)} = \mathbb{P}\bigg(\mathcal{A}_n^c, \sum_{r=i+1}^{i+nw-1} I_r^n = k-2, I_i^n = I_{i+nw}^n = 1\bigg), \qquad i = 1, \dots, n\overline{w}, \tag{2.5}$$

$$\widetilde{P}_i^{(n)} = \mathbb{P}\bigg(\mathcal{A}_n^c, \tilde{I}_i^n = 2, \sum_{r=i'}^{i'+nw-1} I_r^n = k-1 \text{ for some } i' \text{ with } 1 \leq i' \leq i \leq i' + nw - 1 \leq n\bigg), \qquad i = 1, \dots, n. \tag{2.6}$$

Since $\mathbb{P}(\mathcal{G}_3) = O(n^{-2})$, (2.4) follows easily. To prove (2.2), note that when $\tilde{I}_i^n \leq 1$ for all i (i.e. on the event $\mathcal{G}_1$), each subinterval $((i-1)/n, i/n]$ contains at most one Poisson point. If $\tilde{I}_i^n = 1$, denote the only Poisson point in $((i-1)/n, i/n]$ by $Q_{(i)}$ whose location is uniformly distributed over $((i-1)/n, i/n]$. When $\tilde{I}_i^n \leq 1$ for all i, in order for $\mathcal{A} \cap \mathcal{A}_n^c$ to occur, there must exist some pair (i, i') with $i' = i + nw$ such that $\sum_{r=i+1}^{i'-1} \tilde{I}_r^n = k-2$, $\tilde{I}_i^n = \tilde{I}_{i'}^n = 1$, and $Q_{(i')} - Q_{(i)} < w$. So we have $\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1 = \bigcup_{i=1}^{n\overline{w}} \mathcal{G}_{1,i}$, where, for $i = 1, \dots, n\overline{w}$,

$$\mathcal{G}_{1,i} = \mathcal{A}_n^c \cap \bigg\{\tilde{I}_j^n \leq 1 \text{ for all } j, \sum_{r=i+1}^{i+nw-1} \tilde{I}_r^n = k-2, \tilde{I}_i^n = \tilde{I}_{i+nw}^n = 1, \text{ and } Q_{(i+nw)} - Q_{(i)} < w\bigg\}.$$

https://doi.org/10.1017/jpr.2016.101 Published online by Cambridge University Press

Since $\sum_{1\le i<j\le n\overline{w}}\mathbb{P}(\tilde{I}_i^n=\tilde{I}_{i+nw}^n=\tilde{I}_j^n=\tilde{I}_{j+nw}^n=1)=O(n^{-2})$, we have

$$\mathbb{P}(\mathcal{A}\cap\mathcal{A}_n^c\cap\mathcal{G}_1)=\sum_{i=1}^{n\overline{w}}\mathbb{P}(\mathcal{G}_{1,i})+O(n^{-2})=\frac{1}{2}\sum_{i=1}^{n\overline{w}}\mathbb{P}(\mathcal{G}'_{1,i})+O(n^{-2}), \tag{2.7}$$

where $\mathcal{G}'_{1,i}=\mathcal{A}_n^c\cap\{\tilde{I}_j^n\le 1 \text{ for all } j, \sum_{r=i+1}^{i+nw-1}\tilde{I}_r^n=k-2, \tilde{I}_i^n=\tilde{I}_{i+nw}^n=1\}$. In (2.7), we have used the facts that $\tilde{I}_1^n,\dots,\tilde{I}_n^n$ are independent and that given $\tilde{I}_i^n=\tilde{I}_{i+nw}^n=1$, $Q_{(i)}$ and $Q_{(i+nw)}$ are (conditionally) independent and uniformly distributed over $((i-1)/n, i/n]$ and $((i+nw-1)/n, (i+nw)/n]$, respectively, so that $Q_{(i+nw)}-Q_{(i)}<w$ with (conditional) probability $\frac{1}{2}$, which implies that $\mathbb{P}(\mathcal{G}_{1,i})=\frac{1}{2}\mathbb{P}(\mathcal{G}'_{1,i})$. For $i=1,\dots,n\overline{w}$, define

$$\mathcal{G}''_{1,i}=\mathcal{A}_n^c\cap\left\{\sum_{r=i+1}^{i+nw-1}I_r^n=k-2, I_i^n=I_{i+nw}^n=1\right\}, \tag{2.8}$$

which is the event inside the parentheses on the right-hand side of (2.5), so that $P_i^{(n)}=\mathbb{P}(\mathcal{G}''_{1,i})$. Note that $\mathcal{G}'_{1,i}\subset\mathcal{G}''_{1,i}$ and that $\mathcal{G}''_{1,i}\setminus\mathcal{G}'_{1,i}$ is contained in $\{I_i^n=I_{i+nw}^n=1,\ \tilde{I}_j^n\ge 2 \text{ for some } j\}$, which has a probability of order n^{-3}. By (2.7),

$$\mathbb{P}(\mathcal{A}\cap\mathcal{A}_n^c\cap\mathcal{G}_1)=\frac{1}{2}\sum_{i=1}^{n\overline{w}}\mathbb{P}(\mathcal{G}''_{1,i})+O(n^{-2})=\frac{1}{2}\sum_{i=1}^{n\overline{w}}P_i^{(n)}+O(n^{-2}),$$

establishing (2.2).

To prove (2.3), let $\mathcal{H}=\{I_j=I_{j+nw}=1 \text{ for some } 1\le j\le n\overline{w}\}$. On $\mathcal{G}_{2,i}\cap\mathcal{H}^c$, in order for $\mathcal{A}\cap\mathcal{A}_n^c$ to occur, there must exist some i' with $1\le i'\le i\le i'+nw-1\le n$ such that $\sum_{r=i'}^{i'+nw-1}I_r^n=k-1$ (implying that $\sum_{r=i'}^{i'+nw-1}\tilde{I}_r^n=k$). It follows that $\mathcal{A}\cap\mathcal{A}_n^c\cap\mathcal{G}_{2,i}\cap\mathcal{H}^c\subset\mathcal{G}'_{2,i}\subset\mathcal{A}\cap\mathcal{A}_n^c\cap\mathcal{G}_{2,i}$, where

$$\mathcal{G}'_{2,i}=\mathcal{A}_n^c\cap\Bigg\{\tilde{I}_i^n=2, \tilde{I}_j^n\le 1 \text{ for all } j\ne i, \sum_{r=i'}^{i'+nw-1}I_r^n=k-1 \text{ for some } i' \text{ with } 1\le i'\le i\le i'+nw-1\le n\Bigg\}.$$

Since $\mathbb{P}(\mathcal{G}_{2,i}\cap\mathcal{H})=O(n^{-3})$, we have $\sum_{i=1}^n\mathbb{P}(\mathcal{G}_{2,i}\cap\mathcal{H})=O(n^{-2})$, implying that

$$\sum_{i=1}^n\mathbb{P}(\mathcal{A}\cap\mathcal{A}_n^c\cap\mathcal{G}_{2,i})=\sum_{i=1}^n\mathbb{P}(\mathcal{G}'_{2,i})+O(n^{-2})=\sum_{i=1}^n\mathbb{P}(\mathcal{G}''_{2,i})+O(n^{-2}),$$

where

$$\mathcal{G}''_{2,i}=\mathcal{A}_n^c\cap\left\{\tilde{I}_i^n=2, \sum_{r=i'}^{i'+nw-1}I_r^n=k-1 \text{ for some } i' \text{ with } 1\le i'\le i\le i'+nw-1\le n\right\}.$$

(Note that $\mathcal{G}'_{2,i}\subset\mathcal{G}''_{2,i}$ and $\mathcal{G}''_{2,i}\setminus\mathcal{G}'_{2,i}$ is contained in the event $\{\tilde{I}_i^n=2, \tilde{I}_j^n\ge 2 \text{ for some } j\ne i\}$, which has a probability of order n^{-3}.) By (2.6), $\widetilde{P}_i^{(n)}=\mathbb{P}(\mathcal{G}''_{2,i})$. This establishes (2.3).

By (2.1)–(2.4), we have

$$\alpha - \alpha_n = \frac{1}{2}\sum_{i=1}^{n\overline{w}} P_i^{(n)} + \sum_{i=1}^{n} \widetilde{P}_i^{(n)} + O(n^{-2}). \tag{2.9}$$

For $i = 1, \dots, n\overline{w}$, let $P_i^{\prime(n)} = \mathbb{P}(\mathcal{F}_i)$, where

$$\mathcal{F}_i := \mathcal{A}_n^c \cap \Bigg\{ \sum_{r=i+1}^{i+nw-1} I_r^n = k-2,\ I_i^n = 1,\ I_{i+nw}^n = 0,$$
$$\text{sum of any } nw \text{ consecutive } I_r^n \text{ including } r = i+nw \text{ is at most } k-2 \Bigg\}.$$

We claim that

$$\frac{P_i^{(n)}}{P_i^{\prime(n)}} = \rho_n \qquad \text{for all } i = 1, \dots, n\overline{w}, \tag{2.10}$$

where $\rho_n := \mathbb{P}(I_{i+nw}^n = 1)/\mathbb{P}(I_{i+nw}^n = 0) = (1 - \mathrm{e}^{-\lambda/n})/\mathrm{e}^{-\lambda/n} = \mathrm{e}^{\lambda/n} - 1$. To establish the claim, recall that $P_i^{(n)} = \mathbb{P}(\mathcal{G}_{1,i}'')$, where $\mathcal{G}_{1,i}''$ (see (2.8)) depends only on $(I_1^n, \dots, I_n^n)$. It is instructive to interpret $\mathcal{G}_{1,i}''$ as a collection of configurations $(I_1^n, \dots, I_n^n) = (h_1, \dots, h_n)$, where $(h_1, \dots, h_n)$ satisfies $h_j = 0$ or 1 for all j, $h_i = h_{i+nw} = 1$, $\max_{j=1,\dots,n\overline{w}+1} \sum_{r=j}^{j+nw-1} h_r < k$, and $\sum_{r=i+1}^{i+nw-1} h_r = k-2$. Likewise, the event $\mathcal{F}_i$ is a collection of configurations $(I_1^n, \dots, I_n^n) = (h_1', \dots, h_n')$, where $(h_1', \dots, h_n')$ satisfies $h_j' = 0$ or 1 for all j, $h_i' = 1, h_{i+nw}' = 0$, $\max_{j=1,\dots,n\overline{w}+1} \sum_{r=j}^{j+nw-1} h_r' < k$, $\sum_{r=i+1}^{i+nw-1} h_r' = k-2$, and the *sum of any nw consecutive h_r' including $r = i + nw$ is at most $k-2$*. It is readily seen that a configuration $(I_1^n, \dots, I_n^n) = (h_1, \dots, h_n)$ is in $\mathcal{G}_{1,i}''$ if and only if the configuration $(I_1^n, \dots, I_n^n) = (h_1', \dots, h_n')$ is in $\mathcal{F}_i$, where $(h_1', \dots, h_n') = (h_1, \dots, h_n) - \boldsymbol{e}_{i+nw}$ with $\boldsymbol{e}_{i+nw}$ being the vector of 0s except for the $(i + nw)$th entry being 1. The claim (2.10) now follows from the independence property of $I_1^n, \dots, I_n^n$. By (2.10),

$$\rho_n^{-1} \sum_{i=1}^{n\overline{w}} P_i^{(n)} = \sum_{i=1}^{n\overline{w}} P_i^{\prime(n)} = \sum_{i=1}^{n\overline{w}} \mathbb{P}(\mathcal{F}_i) = \mathbb{E}[\nu^{(n)}(\Pi)], \tag{2.11}$$

where

$$\nu^{(n)}(\Pi) := \sum_{i=1}^{n\overline{w}} \mathbf{1}\Bigg\{ \mathcal{A}_n^c,\ \sum_{r=i+1}^{i+nw-1} I_r^n = k-2,\ I_i^n = 1,\ I_{i+nw}^n = 0,$$
$$\text{sum of any } nw \text{ consecutive } I_r^n \text{ including } r = i+nw \text{ is at most } k-2 \Bigg\}.$$

To deal with the terms $\widetilde{P}_i^{(n)}$, $i = 1, \dots, n$, on the right-hand side of (2.9), let

$$\widetilde{P}_i^{\prime(n)} := \mathbb{P}\Bigg(\mathcal{A}_n^c,\ I_i^n = 1,\ \sum_{r=i'}^{i'+nw-1} I_r^n = k-1 \text{ for some } i' \text{ with}$$
$$1 \le i' \le i \le i' + nw - 1 \le n \Bigg).$$

https://doi.org/10.1017/jpr.2016.101 Published online by Cambridge University Press

By an argument similar to the proof of (2.10), we have $\widetilde{P}_i^{(n)}/\widetilde{P}_i'^{(n)} = \tilde{\rho}_n$ for all $i = 1, \dots, n$, where $\tilde{\rho}_n = \mathbb{P}(\tilde{I}_i^n = 2)/\mathbb{P}(I_i^n = 1) = \mathrm{e}^{-\lambda/n}(\lambda/n)^2/[2(1-\mathrm{e}^{-\lambda/n})]$. So,

$$\tilde{\rho}_n^{-1}\sum_{i=1}^n \widetilde{P}_i^{(n)} = \sum_{i=1}^n \widetilde{P}_i'^{(n)} = \mathbb{E}[\tilde{\nu}^{(n)}(\Pi)], \tag{2.12}$$

where

$$\tilde{\nu}^{(n)}(\Pi) := \sum_{i=1}^n \mathbf{1}\bigg\{\mathcal{A}_n^c,\, I_i^n = 1,\, \sum_{r=i'}^{i'+nw-1} I_r^n = k-1 \text{ for some } i' \text{ with } 1 \le i' \le i \le i'+nw-1 \le n\bigg\}.$$

Since $\rho_n = \lambda/n + O(n^{-2})$ and $\tilde{\rho}_n = \lambda/(2n) + O(n^{-2})$, it follows from (2.9), (2.11), and (2.12) that

$$n(\alpha - \alpha_n) - \tfrac{1}{2}\lambda\mathbb{E}[\nu^{(n)}(\Pi) + \tilde{\nu}^{(n)}(\Pi)] = O(n^{-1}). \tag{2.13}$$

Note that $\nu^{(n)}(\Pi)$ and $\tilde{\nu}^{(n)}(\Pi)$ converge almost surely to $\nu(\Pi)$ and $\tilde{\nu}(\Pi)$, respectively. Since

$$\max\{\nu^{(n)}(\Pi), \tilde{\nu}^{(n)}(\Pi)\} \le \sum_{i=1}^n \mathbf{1}\{I_i^n = 1\} \le |\Pi|,$$

we have, by the dominated convergence theorem, that $\mathbb{E}[\nu^{(n)}(\Pi) + \tilde{\nu}^{(n)}(\Pi)]$ converges to $\mathbb{E}[\nu(\Pi) + \tilde{\nu}(\Pi)]$, which together with (2.13) completes the proof. □

Remark 2.1. With a little more effort, it can be shown that

$$\mathbb{E}[\nu^{(n)}(\Pi) + \tilde{\nu}^{(n)}(\Pi)] - \mathbb{E}[\nu(\Pi) + \tilde{\nu}(\Pi)] = O(n^{-1}),$$

which together with (2.13) yields $\alpha - \alpha_n = C_\alpha n^{-1} + O(n^{-2})$, where $C_\alpha = \frac{1}{2}\lambda\mathbb{E}[\nu(\Pi) + \tilde{\nu}(\Pi)]$.

2.2. Matching the expected number of points

Recall that H_i^n, $i = 1, \dots, n$, are i.i.d. with $\mathbb{P}(H_i^n = 0) = 1 - \lambda/n$ and $\mathbb{P}(H_i^n = 1) = \lambda/n$. Let $\beta_n = \mathbb{P}(\mathcal{B}_n)$, where $\mathcal{B}_n := \{S_{w,H}^{(n)} \ge k\} = \{\max_{i=1,\dots,n\overline{w}+1} \sum_{r=i}^{i+nw-1} H_r^n \ge k\}$.

Lemma 2.1. *For $n = mq$ $(m = 1, 2, \dots)$,*

$$\lim_{n\to\infty} \frac{2n}{\lambda^2}(\beta_n - \alpha_n) = -\alpha + \int_0^1 \mathbb{P}\Big(\max_{0\le t\le 1-w} |(\Pi \cup \{u\}) \cap (t, t+w]| \ge k\Big)\, \mathrm{d}u.$$

Proof. Let L_i^n, $i = 1, \dots, n$, be i.i.d. and independent of $I_1^n, \dots, I_n^n$ such that $\mathbb{P}(L_i^n = 0) = (1-\lambda/n)\mathrm{e}^{\lambda/n} = 1 - \mathbb{P}(L_i^n = 1)$. Letting $\tilde{L}_i^n = \max\{I_i^n, L_i^n\}$ and noting that $\mathbb{P}(\tilde{L}_i^n = 0) = \mathbb{P}(I_i^n = 0 \text{ and } L_i^n = 0) = 1 - \lambda/n = \mathbb{P}(H_i^n = 0)$, we have $\mathcal{L}(\tilde{L}_1^n, \dots, \tilde{L}_n^n) = \mathcal{L}(H_1^n, \dots, H_n^n)$, where $\mathcal{L}(\boldsymbol{V})$ denotes the law of a random vector $\boldsymbol{V}$, so that $\beta_n = \mathbb{P}(\mathcal{B}_n) = \mathbb{P}(\widetilde{\mathcal{B}}_n)$, where $\widetilde{\mathcal{B}}_n = \{\max_{i=1,\dots,n\overline{w}+1} \sum_{r=i}^{i+nw-1} \tilde{L}_r^n \ge k\}$. Since $I_i^n = 1$ implies $\tilde{L}_i^n = 1$, we have $\mathcal{A}_n \subset \widetilde{\mathcal{B}}_n$. Letting $S_n = \sum_{i=1}^n L_i^n$ and noting that $\widetilde{\mathcal{B}}_n \cap \{S_n = 0\} = \mathcal{A}_n \cap \{S_n = 0\}$, and that

$$\mathbb{P}(S_n = 0) = 1 - \frac{\lambda^2}{2n} + O(n^{-2}), \qquad \mathbb{P}(S_n = 1) = \frac{\lambda^2}{2n} + O(n^{-2}), \qquad \mathbb{P}(S_n \ge 2) = O(n^{-2}),$$

we have

$$\begin{aligned}\beta_n = \mathbb{P}(\widetilde{\mathcal{B}}_n) &= \mathbb{P}(\widetilde{\mathcal{B}}_n \mid S_n = 0)\mathbb{P}(S_n = 0) + \mathbb{P}(\widetilde{\mathcal{B}}_n \mid S_n = 1)\mathbb{P}(S_n = 1) \\ &\quad + \mathbb{P}(\widetilde{\mathcal{B}}_n \mid S_n \geq 2)\mathbb{P}(S_n \geq 2) \\ &= \mathbb{P}(\mathcal{A}_n \mid S_n = 0)\bigg(1 - \frac{\lambda^2}{2n}\bigg) + \mathbb{P}(\widetilde{\mathcal{B}}_n \mid S_n = 1)\frac{\lambda^2}{2n} + O(n^{-2}) \\ &= \alpha_n\bigg(1 - \frac{\lambda^2}{2n}\bigg) + \mathbb{P}(\widetilde{\mathcal{B}}_n \mid S_n = 1)\frac{\lambda^2}{2n} + O(n^{-2}). \end{aligned} \tag{2.14}$$

We claim that

$$\lim_{n\to\infty} \mathbb{P}(\widetilde{\mathcal{B}}_n \mid S_n = 1) = \int_0^1 \mathbb{P}\Big(\max_{0\leq t\leq 1-w} |(\Pi \cup \{u\}) \cap (t, t+w]| \geq k\Big)\, \mathrm{d}u, \tag{2.15}$$

which together with (2.14) yields the desired result.

It remains to establish claim (2.15). Let Q be a random point which is uniformly distributed on $(0, 1]$ and independent of Π. Let

$$\hat{I}_i^n = \mathbf{1}\bigg\{(\Pi \cup \{Q\}) \cap \bigg(\frac{i-1}{n}, \frac{i}{n}\bigg] \neq \varnothing\bigg\}, \qquad i = 1, \dots, n.$$

It is readily seen that $\mathcal{L}(\tilde{L}_1^n, \dots, \tilde{L}_n^n \mid S_n = 1) = \mathcal{L}(\hat{I}_1^n, \dots, \hat{I}_n^n)$, which implies that $\mathbb{P}(\widetilde{\mathcal{B}}_n \mid S_n = 1) = \mathbb{P}(\hat{\mathcal{B}}_n)$, where $\hat{\mathcal{B}}_n = \{\max_{i=1,\dots,n\overline{w}+1} \sum_{r=i}^{i+nw-1} \hat{I}_r^n \geq k\}$. Since $\mathbf{1}_{\hat{\mathcal{B}}_n}$ converges almost surely to $\mathbf{1}\{\max_{0\leq t\leq 1-w} |(\Pi \cup \{Q\}) \cap (t, t+w]| \geq k\}$, we have

$$\lim_{n\to\infty} \mathbb{P}(\widetilde{\mathcal{B}}_n \mid S_n = 1) = \lim_{n\to\infty} \mathbb{P}(\hat{\mathcal{B}}_n) = \mathbb{P}\Big(\max_{0\leq t\leq 1-w} |(\Pi \cup \{Q\}) \cap (t, t+w]| \geq k\Big),$$

from which claim (2.15) follows. This completes the proof of the lemma. □

Theorem 2.2. *For $n = mq\,(m = 1, 2, \dots)$,*

$$\begin{aligned}&\lim_{n\to\infty} \frac{2n}{\lambda^2}(\alpha - \beta_n) \\ &\quad = \frac{1}{\lambda}\mathbb{E}[\nu(\Pi) + \tilde{\nu}(\Pi)] + \alpha - \int_0^1 \mathbb{P}\Big(\max_{0\leq t\leq 1-w} |(\Pi \cup \{u\}) \cap (t, t+w]| \geq k\Big)\, \mathrm{d}u.\end{aligned}$$

Proof. Note that

$$\frac{2n}{\lambda^2}(\alpha - \beta_n) = \frac{2n}{\lambda^2}(\alpha - \alpha_n) - \frac{2n}{\lambda^2}(\beta_n - \alpha_n),$$

which together with Theorem 2.1 and Lemma 2.1 yields the desired result. □

Remark 2.2. It can be shown (see Remark 2.1) that $\alpha - \beta_n = C_\beta n^{-1} + O(n^{-2})$, where

$$C_\beta = C_\alpha + \frac{1}{2}\lambda^2\alpha - \frac{\lambda^2}{2}\int_0^1 \mathbb{P}\Big(\max_{0\leq t\leq 1-w} |(\Pi \cup \{u\}) \cap (t, t+w]| \geq k\Big)\, \mathrm{d}u.$$

3. The conditional case

In this section, for given $N \geq k = 2, 3, \dots$, we are interested in approximating

$$\gamma^{(N)} := P(k; N, w) = \mathbb{P}\Big(\max_{0\leq t\leq 1-w} |\Pi \cap (t, t+w]| \geq k \mid M = N\Big), \qquad M := |\Pi|.$$

Denoting by Π^N a set of N i.i.d. uniform random variables on (0, 1], we have $\mathcal{L}(\Pi^N) = \mathcal{L}(\Pi \mid M = N)$ and $\gamma^{(N)} = \mathbb{P}(\mathcal{E}^N)$, where $\mathcal{E}^N := \{\max_{0\leq t\leq 1-w} |\Pi^N \cap (t, t+w]| \geq k\}$. As in Section 2, with $n = mq$ ($m = 1, 2, \dots$), the interval (0, 1] is partitioned into n subintervals of length n^{-1} so that a window of size $w = p/q$ covers nw subintervals. As an approximation to N points uniformly distributed on (0, 1], we randomly select N of the n subintervals and assign a point to each of them. Let $J_i^n = 1$ or 0 according to whether or not the ith subinterval is selected (so as to contain a point). Then $\sum_{i=1}^n J_i^n = N$. For $h_i = 0$ or 1 with $\sum_{i=1}^n h_i = N$, we have $\mathbb{P}_N(J_i^n = h_i,\ i = 1, \dots, n) = 1/\binom{n}{N}$, where the subscript N in $\mathbb{P}_N$ signifies there are N 1s in $J_1^n, \dots, J_n^n$. Unlike in Section 2 where $(I_1^n, \dots, I_n^n)$ is defined in terms of Π so as to allow for a coupling argument, here there is no natural way to define $(J_1^n, \dots, J_n^n)$ and Π^N on the same probability space. As no danger of confusion may arise, we will abuse notation by using the same probability measure notation $\mathbb{P}_N$ for both the probability space where Π^N is defined, and the probability space where $(J_1^n, \dots, J_n^n)$ is defined. Let

$$\gamma_n^{(N)} = \mathbb{P}_N(\mathcal{E}_n^N), \tag{3.1}$$

where $\mathcal{E}_n^N := \{\max_{i=1,\dots,n\overline{w}+1} \sum_{r=i}^{i+nw-1} J_r^n \geq k\}$.

Theorem 3.1. *For N fixed and $n = mq\,(m = 1, 2, \dots)$,*

$$\lim_{n\to\infty} n(\gamma^{(N)} - \gamma_n^{(N)}) = \tfrac{1}{2}N(N-1)(\gamma^{(N-1)} - \gamma^{(N)}) + \tfrac{1}{2}N\mathbb{E}[\nu(\Pi) + \tilde{\nu}(\Pi) \mid M = N-1].$$

Proof. The proof is similar to (but somewhat more involved than) that of Theorem 2.1. Because of space limitation, we only sketch it here and refer the reader to [23] for further details. For notational simplicity, the superscript N in $\mathcal{E}^N$ and $\mathcal{E}_n^N$ is suppressed. But to avoid possible confusion, $\mathbb{P}_N$ is not abbreviated to $\mathbb{P}$ as later a change of measure argument requires consideration of $\mathbb{P}_{N-1}$. Let $\tilde{J}_i = |\Pi^N \cap ((i-1)/N, i/N]|$, $i = 1, \dots, n$, and define the (disjoint) events

$$U_1 = \{\tilde{J}_j^n \leq 1,\ j = 1, \dots, n\}, \qquad U_2 = \bigcup_{i=1}^n U_{2,i},\ U_{2,i} = \{\tilde{J}_i^n = 2,\ \tilde{J}_j^n \leq 1 \text{ for all } j \neq i\},$$

$$U_3 = \{\tilde{J}_j^n = \tilde{J}_{j'}^n = 2 \text{ for some } j \neq j'\} \cup \{\tilde{J}_j^n \geq 3 \text{ for some } j\}.$$

We have $\mathbb{P}_N(U_1) = 1 - N(N-1)/(2n) + O(n^{-2})$, $\mathbb{P}_N(U_2) = N(N-1)/(2n) + O(n^{-2})$, and $\mathbb{P}_N(U_3) = O(n^{-2})$, so that

$$\gamma^{(N)} = \mathbb{P}_N(\mathcal{E}^N) = \mathbb{P}_N(\mathcal{E}) = \mathbb{P}_N(\mathcal{E} \mid U_1)\mathbb{P}_N(U_1) + \mathbb{P}_N(\mathcal{E} \cap U_2) + O(n^{-2}). \tag{3.2}$$

To deal with $\mathbb{P}_N(\mathcal{E} \mid U_1)$, let $\widetilde{\mathcal{E}}_n := \{\max_{i=1,\dots,n\overline{w}+1} \sum_{r=i}^{i+nw-1} \tilde{J}_r^n \geq k\}$ (which is contained in $\mathcal{E}$). Note that $\mathcal{L}(\tilde{J}_1^n, \dots, \tilde{J}_n^n \mid U_1) = \mathcal{L}(J_1^n, \dots, J_n^n)$ and that $\widetilde{\mathcal{E}}_n$ depends on $(\tilde{J}_1^n, \dots, \tilde{J}_n^n)$ in the same way that $\mathcal{E}_n = \mathcal{E}_n^N$ depends on $(J_1^n, \dots, J_n^n)$ (see (3.1)). So we have $\mathbb{P}_N(\widetilde{\mathcal{E}}_n \mid U_1) = \mathbb{P}_N(\mathcal{E}_n) = \gamma_n^{(N)}$ and

$$\mathbb{P}_N(\mathcal{E} \mid U_1) = \mathbb{P}_N(\widetilde{\mathcal{E}}_n \mid U_1) + \mathbb{P}_N(\mathcal{E} \cap \widetilde{\mathcal{E}}_n^c \mid U_1) = \gamma_n^{(N)} + \mathbb{P}_N(\mathcal{E} \cap \widetilde{\mathcal{E}}_n^c \mid U_1). \tag{3.3}$$

If $\tilde{J}_i^n = 1$, denote the only point of Π^N in $((i-1)/n, i/n]$ by $Q_{(i)}$, whose location is uniformly distributed over $((i-1)/n, i/n]$. When $\tilde{J}_i^n \le 1$ for all i (i.e. on the event U_1), in order for $\mathcal{E} \cap \widetilde{\mathcal{E}}_n^c$ to occur, there must exist some pair (i, i') with $i' = i + nw$ such that $\sum_{r=i+1}^{i'-1} \tilde{J}_r^n = k-2$, $\tilde{J}_i^n = \tilde{J}_{i'}^n = 1$, and $Q_{(i')} - Q_{(i)} < w$. So we have $\mathcal{E} \cap \widetilde{\mathcal{E}}_n^c \cap U_1 = \bigcup_{i=1}^{n\overline{w}} U_{1,i}$ where, for $i = 1, \ldots, n\overline{w}$,

$$U_{1,i} = \widetilde{\mathcal{E}}_n^c \cap \bigg\{ \tilde{J}_j^n \le 1 \text{ for all } j, \sum_{r=i+1}^{i+nw-1} \tilde{J}_r^n = k-2, \tilde{J}_i^n = \tilde{J}_{i+nw}^n = 1, Q_{(i+nw)} - Q_{(i)} < w \bigg\}.$$

Since $\sum_{1 \le i < j \le n\overline{w}} \mathbb{P}_N(U_{1,i} \cap U_{1,j}) = O(n^{-2})$, we have

$$\mathbb{P}_N(\mathcal{E} \cap \widetilde{\mathcal{E}}_n^c \mid U_1) = \sum_{i=1}^{n\overline{w}} \mathbb{P}_N(U_{1,i} \mid U_1) + O(n^{-2}) = \frac{1}{2} \sum_{i=1}^{n\overline{w}} \mathbb{P}_N(U'_{1,i} \mid U_1) + O(n^{-2}), \quad (3.4)$$

where $U'_{1,i} = \widetilde{\mathcal{E}}_n^c \cap \{\tilde{J}_j^n \le 1 \text{ for all } j, \sum_{r=i+1}^{i+nw-1} \tilde{J}_r^n = k-2, \tilde{J}_i^n = \tilde{J}_{i+nw}^n = 1\}, i = 1, \ldots, n\overline{w}$. In (3.4), we have used the fact that for any given $h_j = 0$ or 1 $(j = 1, \ldots, n)$ with $\sum_{j=1}^n h_j = N$ and $h_i = h_{i+nw} = 1$, conditional on $\tilde{J}_j^n = h_j$, $j = 1, \ldots, n$, $Q_{(i)}$, and $Q_{(i+nw)}$ are independent and uniformly distributed over $((i-1)/n, i/n]$ and $((i+nw-1)/n, (i+nw)/n]$, respectively, so that $Q_{(i+nw)} - Q_{(i)} < w$ with probability $\frac{1}{2}$, which implies that $\mathbb{P}_N(U_{1,i} \mid U_1) = \frac{1}{2} \mathbb{P}_N(U'_{1,i} \mid U_1)$.

Note that $U'_{1,i}, i = 1, \ldots, n$, depend only on $\tilde{J}_1^n, \ldots, \tilde{J}_n^n$. Since $\mathcal{L}(\tilde{J}_1^n, \ldots, \tilde{J}_n^n \mid U_1) = \mathcal{L}(J_1^n, \ldots, J_n^n)$, we have

$$\mathbb{P}_N(U'_{1,i} \mid U_1) = \mathbb{P}_N(V_i), \qquad i = 1, \ldots, n\overline{w}, \quad (3.5)$$

where $V_i = \mathcal{E}_n^c \cap \{\sum_{r=i+1}^{i+nw-1} J_r^n = k-2, J_i^n = J_{i+nw}^n = 1\}$. (Note that V_i depends on $(J_1^n, \ldots, J_n^n)$ in the same way that $U'_{1,i}$ depends on $(\tilde{J}_1^n, \ldots, \tilde{J}_n^n)$.)

We will simplify $\sum_{i=1}^{n\overline{w}} \mathbb{P}_N(U'_{1,i} \mid U_1) = \sum_{i=1}^{n\overline{w}} \mathbb{P}_N(V_i)$ via a change of measure argument. It is instructive to interpret the event V_i as a collection of configurations $(J_1^n, \ldots, J_n^n) = (h_1, \ldots, h_n)$, where $(h_1, \ldots, h_n)$ satisfies $h_r = 0$ or 1 for $r = 1, \ldots, n$, $\sum_{r=1}^n h_r = N$, $\max_{j=1,\ldots,n\overline{w}+1} \sum_{r=j}^{j+nw-1} h_r < k$, and $\sum_{r=i+1}^{i+nw-1} h_r = k-2$, $h_i = h_{i+nw} = 1$. Let

$$\begin{aligned} V_i^* = \bigg\{ & \sum_{r=1}^n J_r^n = N-1, \max_{j=1,\ldots,n\overline{w}+1} \sum_{r=j}^{j+nw-1} J_r^n < k, \\ & \sum_{r=i+1}^{i+nw-1} J_r^n = k-2, J_i^n = 1, J_{i+nw}^n = 0, \\ & \quad \text{sum of any } nw \text{ consecutive } J_r^n \text{ including } r = i + nw \text{ is at most } k-2 \bigg\}. \end{aligned}$$

We interpret the V_i^* as a collection of configurations $(J_1^n, \ldots, J_n^n) = (h_1^*, \ldots, h_n^*)$, where $(h_1^*, \ldots, h_n^*)$ satisfies $h_r^* = 0$ or 1 for $r = 1, \ldots, n$,

$$\sum_{r=1}^n h_r^* = N-1, \qquad \sum_{r=j}^{j+nw-1} h_r^* < k \quad \text{for } j = 1, \ldots, n\overline{w}+1,$$

$$\sum_{r=i+1}^{i+nw-1} h_r^* = k-2, \qquad h_i^* = 1, \ h_{i+nw}^* = 0,$$

https://doi.org/10.1017/jpr.2016.101 Published online by Cambridge University Press

and the sum of any nw consecutive h_r^* including $r = i + nw$ is at most $k - 2$. If a configuration $(J_1^n, \ldots, J_n^n) = (h_1, \ldots, h_n)$ is in V_i, then the configuration $(J_1^n, \ldots, J_n^n) = (h_1^*, \ldots, h_n^*)$ is in V_i^* provided $h_r^* = h_r$ for all $r \neq i + nw$ and $h_{i+nw} = 1$, $h_{i+nw}^* = 0$. In other words, a configuration is in V_i if and only if the configuration derived from it by replacing the $(i + nw)$th entry with 0 is in V_i^*. Note that the number of nonzero entries for a configuration in V_i^* is equal to $N - 1$. Recall that the notation $\mathbb{P}_N$ ($\mathbb{P}_{N-1}$, respectively) denotes the probability measure for $(J_1^n, \ldots, J_n^n)$ with $\sum_{r=1}^n J_r^n = N$ ($\sum_{r=1}^n J_r^n = N - 1$, respectively). It follows that $\mathbb{P}_N(V_i)/\mathbb{P}_{N-1}(V_i^*) = \binom{n}{N-1}/\binom{n}{N} = N/(n - N + 1)$. Therefore,

$$\left(\frac{n - N + 1}{N}\right) \sum_{i=1}^{n\overline{w}} \mathbb{P}_N(V_i) = \sum_{i=1}^{n\overline{w}} \mathbb{P}_{N-1}(V_i^*) = \mathbb{E}_{N-1}[\nu_1^{(n)}(J_1^n, \ldots, J_n^n)], \tag{3.6}$$

where

$$\nu_1^{(n)}(J_1^n, \ldots, J_n^n) = \sum_{i=1}^{n\overline{w}} \mathbf{1}\Bigg\{ \max_{j=1,\ldots,n\overline{w}+1} \sum_{r=j}^{j+nw-1} J_r^n < k, \sum_{r=i+1}^{i+nw-1} J_r^n = k - 2, \\ J_i^n = 1, J_{i+nw}^n = 0, \text{ sum of any } nw \text{ consecutive } J_r^n \\ \text{including } r = i + nw \text{ is at most } k - 2 \Bigg\}.$$

By (3.3)–(3.6),

$$\begin{aligned} \mathbb{P}_N(\mathcal{E} \mid U_1) &= \gamma_n^{(N)} + \frac{1}{2} \frac{N}{n - N + 1} \mathbb{E}_{N-1}[\nu_1^{(n)}(J_1^n, \ldots, J_n^n)] + O(n^{-2}) \\ &= \gamma_n^{(N)} + \frac{N}{2n} \mathbb{E}[\nu(\Pi) \mid M = N - 1] + o(n^{-1}), \end{aligned} \tag{3.7}$$

since $\lim_{n\to\infty} \mathbb{E}_{N-1}[\nu_1^{(n)}(J_1^n, \ldots, J_n^n)] = \mathbb{E}[\nu(\Pi) \mid M = N - 1]$.

Another change of measure argument can be used to deal with $\mathbb{P}_N(\mathcal{E} \cap U_2)$ (see [23]), yielding

$$\mathbb{P}_N(\mathcal{E} \cap U_2) = \frac{N}{2n}((N - 1)\gamma^{(N-1)} + \mathbb{E}[\tilde{\nu}(\Pi) \mid M = N - 1]) + o(n^{-1}),$$

which together with (3.2) and (3.7) completes the proof. □

Remark 3.1. It can be shown (see Remarks 2.1 and 2.2) that $\gamma^{(N)} - \gamma_n^{(N)} = C_\gamma n^{-1} + O(n^{-2})$, where $C_\gamma = \frac{1}{2}N(N - 1)(\gamma^{(N-1)} - \gamma^{(N)}) + \frac{1}{2}N\mathbb{E}[\nu(\Pi) + \tilde{\nu}(\Pi) \mid M = N - 1]$.

Remark 3.2. Note that α_n and β_n are weighted averages of $\gamma_n^{(N)}$ over N with binomial probabilities $\binom{n}{N} p_n^N (1 - p_n)^{n-N}$ as weights, where $p_n = 1 - \mathrm{e}^{-\lambda/n}$ for α_n and $p_n = \lambda/n$ for β_n. The limits $\lim_{n\to\infty} n(\alpha - \alpha_n)$ and $\lim_{n\to\infty} n(\alpha - \beta_n)$ in Theorems 2.1 and 2.2 can be *formally* derived from $\lim_{n\to\infty} n(\gamma^{(N)} - \gamma_n^{(N)})$ by interchanging $\lim_n$ and Σ_N.

4. Numerical results and discussion

Using the Markov chain embedding method (see [4], [6], [17]), we compute the discrete approximations α_n, β_n, and $\gamma_n^{(N)}$ for various combinations of parameter values (k, w, λ) (the unconditional case) and (k, w, N) (the conditional case). In Figure 1 we plot $n(\alpha - \alpha_n)$, $n(\alpha - \beta_n)$, and $n(\gamma^{(N)} - \gamma_n^{(N)})$ for $n = 25(5)600$ with $k = 5$, $w = 0.4$, $\lambda = 8$, and

https://doi.org/10.1017/jpr.2016.101 Published online by Cambridge University Press

$N = 8$, where the superscript (N) in $\gamma^{(N)}$ and $\gamma_n^{(N)}$ is suppressed for ease of notation.The exact probabilities $\alpha = P^*(k; \lambda, w) = P^*(5; 8, 0.4) = 0.628\,144\,085$ and $\gamma^{(8)} = P(k; N, w) = P(5; 8, 0.4) = 0.780\,861\,440$ are taken from [20]. By Theorems 2.1, 2.2, and 3.1, it follows that $n(\alpha - \alpha_n)$, $n(\alpha - \beta_n)$, and $n(\gamma^{(N)} - \gamma_n^{(N)})$ converge, respectively, to the limits C_α, C_β, and C_γ (see Remarks 2.1, 2.2, and 3.1). These limits were estimated by Monte Carlo simulation with 10^6 replications, resulting in $C_\alpha = 4.6322 \pm 0.0096$ (standard error), $C_\beta = 0.8297 \pm 0.0167$, and $C_\gamma = 2.7279 \pm 0.0114$. In view of Remarks 2.1, 2.2, and 3.1, the rate of convergence, n^{-1}, for α_n, β_n, and $\gamma_n^{(N)}$ can be improved to n^{-2} by using Richardson's extrapolation. Specifically, for $w = p/q$, suppose that n is even such that $n/2$ is a multiple of q. Letting $\tilde{\alpha}_n := 2\alpha_n - \alpha_{n/2}$, $\tilde{\beta}_n := 2\beta_n - \beta_{n/2}$, and $\tilde{\gamma}_n^{(N)} := 2\gamma_n^{(N)} - \gamma_{n/2}^{(N)}$, we have $\alpha - \tilde{\alpha}_n = O(n^{-2})$, $\alpha - \tilde{\beta}_n = O(n^{-2})$, and $\gamma^{(N)} - \tilde{\gamma}_n^{(N)} = O(n^{-2})$. In Table 1 we present numerical results comparing α_n, $\tilde{\alpha}_n$, β_n, and $\tilde{\beta}_n$ for the unconditional case. In Table 2 we compare $\gamma_n^{(N)}$ and $\tilde{\gamma}_n^{(N)}$ for the conditional case.

Remark 4.1. In Tables 1 and 2, we have taken relatively large values of $w = 0.2$ and 0.4 since the *exact* unconditional probabilities reported in [20] are less accurate for $w < 0.2$. It is shown in Figure 1 that $n(\alpha - \alpha_n)$, $n(\alpha - \beta_n)$, and $n(\gamma^{(N)} - \gamma_n^{(N)})$ monotonically approach C_α, C_β, and C_γ, respectively. In Table 1, β_n is consistently more accurate than α_n, which is not surprising since $\alpha_n < \min\{\alpha, \beta_n\}$. According to Tables 1 and 2, when n doubles, the errors of α_n, β_n, and $\gamma_n^{(N)}$ decrease by roughly a factor of two while the errors of the corrected approximations $\tilde{\alpha}_n$, $\tilde{\beta}_n$, and $\tilde{\gamma}_n^{(N)}$ decrease by (very) roughly a factor of four. Our limited numerical studies

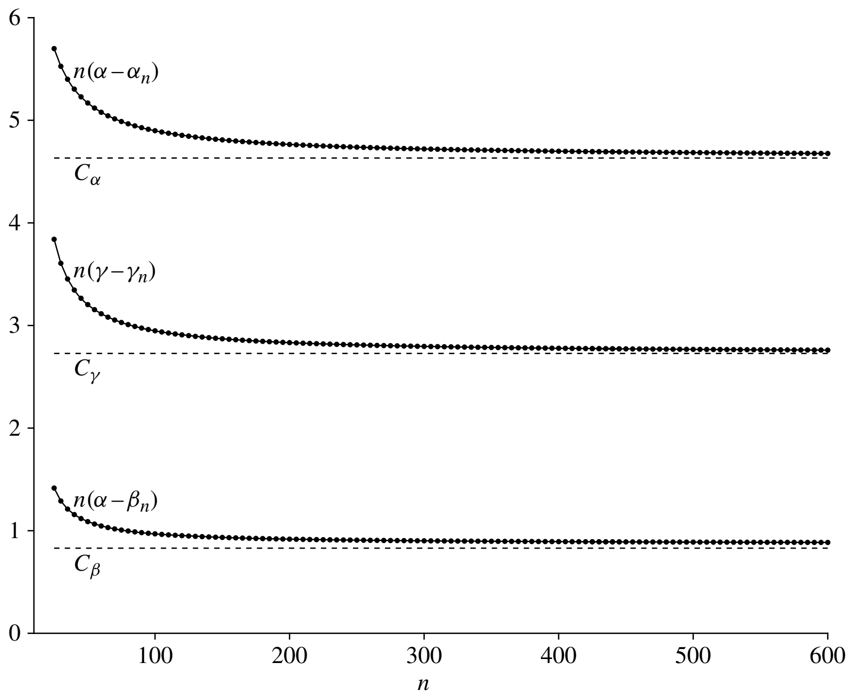

FIGURE 1: Plot of $n(\alpha - \alpha_n)$, $n(\alpha - \beta_n)$, and $n(\gamma - \gamma_n)$ for $n = 25(5)600$ with parameters $w = 0.4$, $k = 5$, $\lambda = 8$, and $N = 8$.

https://doi.org/10.1017/jpr.2016.101 Published online by Cambridge University Press

TABLE 1: The unconditional case.

Parameters				n					Exact
λ	w	k		25	50	100	200	400	α
4	0.2	3	α_n	0.226 474 137	0.297 081 029	0.330 413 369	0.346 549 002	0.354 481 473	0.362322986
			$\alpha-\alpha_n$	0.135 848 849	0.065 241 957	0.031 909 617	0.015 773 984	0.007 841 513	
			$\widetilde{\alpha}_n$		0.367 687 921	0.363 745 709	0.362 684 635	0.362 413 943	
			$\alpha-\widetilde{\alpha}_n$		−0.005 364 935	−0.001 422 723	−0.000 361 649	−0.000 090 957	
			β_n	0.269 466 265	0.321 289 109	0.342 964 036	0.352 912 780	0.357 682 871	
			$\alpha-\beta_n$	0.092 856 721	0.041 033 877	0.019 358 950	0.009 410 206	0.004 640 115	
			$\widetilde{\beta}_n$		0.373 111 952	0.364 638 964	0.362 861 523	0.362 452 963	
			$\alpha-\widetilde{\beta}_n$		−0.010 788 966	−0.002 315 978	−0.000 538 537	−0.000 129 977	
4	0.2	4	α_n	0.028 528 199	0.063 252 847	0.083 861 016	0.094 813 938	0.100 432 989	0.106 139 839
			$\alpha-\alpha_n$	0.077 611 640	0.042 886 992	0.022 278 823	0.011 325 901	0.005 706 850	
			$\widetilde{\alpha}_n$		0.097 977 495	0.104 469 184	0.105 766 860	0.106 052 039	
			$\alpha-\widetilde{\alpha}_n$		0.008 162 344	0.001 670 655	0.000 372 979	0.000 087 800	
			β_n	0.037 826 080	0.071 921 990	0.089 167 692	0.097 701 122	0.101 933 685	
			$\alpha-\beta_n$	0.068 313 759	0.034 217 849	0.016 972 147	0.008 438 717	0.004 206 154	
			$\widetilde{\beta}_n$		0.106 017 899	0.106 413 395	0.106 234 551	0.106 166 248	
			$\alpha-\widetilde{\beta}_n$		0.000 121 940	−0.000 273 556	−0.000 094 712	−0.000 026 409	
8	0.4	5	α_n	0.400 190 890	0.524 770 327	0.579 159 623	0.604 320 002	0.616 397 532	0.628 144 085
			$\alpha-\alpha_n$	0.227 953 195	0.103 373 758	0.048 984 462	0.023 824 083	0.011 746 553	
			$\widetilde{\alpha}_n$		0.649 349 765	0.633 548 918	0.629 480 382	0.628 475 061	
			$\alpha-\widetilde{\alpha}_n$		−0.021 205 680	−0.005 404 833	−0.001 336 297	−0.000 330 976	
			β_n	0.571 524 668	0.606 381 317	0.618 451 977	0.623 556 407	0.625 910 702	
			$\alpha-\beta_n$	0.056 619 417	0.021 762 768	0.009 692 108	0.004 587 678	0.002 233 383	
			$\widetilde{\beta}_n$		0.641 237 966	0.630 522 637	0.628 660 836	0.628 264 997	
			$\alpha-\widetilde{\beta}_n$		−0.013 093 881	−0.002 378 552	−0.000 516 751	−0.000 120 912	
8	0.4	6	α_n	0.156 407 681	0.278 520 053	0.341 202 440	0.372 097 133	0.387 351 968	0.402 452 588
			$\alpha-\alpha_n$	0.246 044 907	0.123 932 535	0.061 250 148	0.030 355 455	0.015 100 620	
			$\widetilde{\alpha}_n$		0.400 632 426	0.403 884 826	0.402 991 826	0.402 606 803	
			$\alpha-\widetilde{\alpha}_n$		0.001 820 162	−0.001 432 238	−0.000 539 238	−0.000 154 215	
			β_n	0.278 663 391	0.351 874 806	0.379 351 117	0.391 387 631	0.397 034 846	
			$\alpha-\beta_n$	0.123 789 197	0.050 577 782	0.023 101 471	0.011 064 957	0.005 417 742	
			$\widetilde{\beta}_n$		0.425 086 221	0.406 827 428	0.403 424 144	0.402 682 062	
			$\alpha-\widetilde{\beta}_n$		−0.022 633 633	−0.004 374 840	−0.000 971 556	−0.000 229 474	

indicate that the corrected approximations are more accurate than the uncorrected ones for $n \geq 50$. Also in both tables, $\tilde{\beta}_{100}$ ($\tilde{\gamma}_{100}^{(N)}$, respectively) is about as accurate as/or more accurate than β_{400} (respectively $\gamma_{400}^{(N)}$).

Remark 4.2. The anonymous referee of this paper raised an important question on the relationship among $w, \lambda(N)$ and the convergence rate. While the convergence rate for the uncorrected (respectively corrected) approximations is n^{-1} (respectively n^{-2}), the error size for the approximations α_n, β_n, $\gamma_n^{(N)}$, $\tilde{\alpha}$, $\tilde{\beta}_n$, and $\tilde{\gamma}_n^{(N)}$ depends on w and $\lambda(N)$ as well as on k in a very complicated way. While addressing this issue in full detail would require extensive analytical and numerical studies, we briefly present in Table 3 the values of $n(\alpha - \beta_n)$ and $n^2(\alpha - \tilde{\beta}_n)$ for $n = 400$, $w \in \{0.1, 0.2, 0.3, 0.4\}$, $\lambda \in \{1, 2, 4, 8\}$, and $k \in \{3, 5\}$. The (absolute) value of $n^2(\alpha - \tilde{\beta}_n)$ is noticeably large for $w = 0.1$ and $\lambda = 8$, indicating that the approximation $\tilde{\beta}_n$ is less accurate when w is small and λ is large.

Remark 4.3. The discrete approximations are usually computed using the Markov chain embedding method. A drawback of this method is the requirement of a very large state space (corresponding to a large computer memory space) for some practical applications. Indeed, it

TABLE 2: The conditional case.

Parameters				*n*					Exact
w	k	N		25	50	100	200	400	γ
0.2	4	6	γ_n	0.080 688 876	0.155 913 836	0.194 799 457	0.214 242 757	0.223 935 622	0.233 600 000
			$\gamma - \gamma_n$	0.152 911 124	0.077 686 164	0.038 800 543	0.019 357 243	0.009 664 378	
			$\widetilde{\gamma}_n$		0.231 138 796	0.233 685 077	0.233 686 058	0.233 628 487	
			$\gamma - \widetilde{\gamma}_n$		0.002 461 204	−0.000 085 077	−0.000 086 058	−0.000 028 487	
		7	γ_n	0.166 798 419	0.294 914 521	0.354 660 825	0.383 179 030	0.397 084 766	0.410 752 000
			$\gamma - \gamma_n$	0.243 953 581	0.115 837 479	0.056 091 175	0.027 572 970	0.013 667 234	
			$\widetilde{\gamma}_n$		0.423 030 623	0.414 407 129	0.411 697 234	0.410 990 502	
			$\gamma - \widetilde{\gamma}_n$		−0.012 278 623	−0.003 655 129	−0.000 945 234	−0.000 238 502	
		8	γ_n	0.291 588 655	0.469 102 180	0.542 215 920	0.575 193 126	0.590 840 920	0.605 949 440
			$\gamma - \gamma_n$	0.314 360 785	0.136 847 260	0.063 733 520	0.030 756 314	0.015 108 520	
			$\widetilde{\gamma}_n$		0.646 615 704	0.615 329 660	0.608 170 332	0.606 488 715	
			$\gamma - \widetilde{\gamma}_n$		−0.040 666 264	−0.009 380 220	−0.002 220 892	−0.000 539 275	
		9	γ_n	0.448 718 168	0.651 907 101	0.723 414 803	0.753 448 974	0.767 220 941	0.780 225 536
			$\gamma - \gamma_n$	0.331 507 368	0.128 318 435	0.056 810 733	0.026 776 562	0.013 004 595	
			$\widetilde{\gamma}_n$		0.855 096 034	0.794 922 506	0.783 483 145	0.780 992 907	
			$\gamma - \widetilde{\gamma}_n$		−0.074 870 498	−0.014 696 970	−0.003 257 609	−0.000 767 371	
0.4	5	6	γ_n	0.162 450 593	0.224 402 953	0.254 838 093	0.269 838 436	0.277 276 942	0.284 672 000
			$\gamma - \gamma_n$	0.122 221 407	0.060 269 047	0.029 833 907	0.014 833 564	0.007 395 058	
			$\widetilde{\gamma}_n$		0.286 355 312	0.285 273 233	0.284 838 778	0.284 715 449	
			$\gamma - \widetilde{\gamma}_n$		−0.001 683 312	−0.000 601 233	−0.000 166 778	−0.000 043 449	
		7	γ_n	0.371 395 881	0.463 718 058	0.504 028 994	0.522 865 538	0.531 971 853	0.540 876 800
			$\gamma - \gamma_n$	0.169 480 919	0.077 158 742	0.036 847 806	0.018 011 262	0.008 904 947	
			$\widetilde{\gamma}_n$		0.556 040 235	0.544 339 929	0.541 702 083	0.541 078 168	
			$\gamma - \widetilde{\gamma}_n$		−0.015 163 435	−0.003 463 129	−0.000 825 283	−0.000 201 368	
		8	γ_n	0.627 251 924	0.716 788 906	0.751 379 277	0.766 696 715	0.773 916 208	0.780 861 440
			$\gamma - \gamma_n$	0.153 609 516	0.064 072 534	0.029 482 163	0.014 164 725	0.006 945 232	
			$\widetilde{\gamma}_n$		0.806 325 887	0.785 969 648	0.782 014 154	0.781 135 700	
			$\gamma - \widetilde{\gamma}_n$		−0.025 464 447	−0.005 108 208	−0.001 152 714	−0.000 274 260	
		9	γ_n	0.864 220 071	0.918 826 852	0.936 992 617	0.944 511 915	0.947 942 424	0.951 173 120
			$\gamma - \gamma_n$	0.086 953 049	0.032 346 268	0.014 180 503	0.006 661 205	0.003 230 696	
			$\widetilde{\gamma}_n$		0.973 433 633	0.955 158 381	0.952 031 214	0.951 372 933	
			$\gamma - \widetilde{\gamma}_n$		−0.022 260 513	−0.003 985 261	−0.000 858 094	−0.000 199 813	

was shown in [3] that to compute α_n, β_n, and $\gamma_n^{(N)}$ using the Markov chain embedding method, the minimum number of states required is $\binom{nw}{k-1} + 1$, which is enormous when nw is large and k is not small. (It should be remarked that [3] was concerned with computation of the reliability for the so-called d-within-consecutive-k-out-of-n system, which is equivalent to the discrete scan statistic.) The corrected discrete approximations partially alleviate the requirement of large memory space since a reasonable accuracy can be achieved with relatively small n.

Remark 4.4. Since the assumption of constant intensity plays a relatively minor role in the proofs of Theorems 2.1, 2.2, and 3.1, the proof method can be extended to the setting of nonhomogeneous Poisson point processes, which is relevant to computation of the power of the continuous scan statistic. In the literature, there appears to be no general method available for computing the exact power under general nonhomogeneous Poisson point processes. The method of corrected discrete approximation may prove to be useful in such a setting as well as in a multiple-window setting (see [22]).

TABLE 3: Values for $n(\alpha - \beta_n)$ and $n^2(\alpha - \tilde{\beta}_n)$ with fixed $n = 400$.

		$k = 3$			
		$w = 0.1$	$w = 0.2$	$w = 0.3$	$w = 0.4$
$n(\alpha - \beta_n)$	$\lambda = 1$	0.109 466	0.162 936	0.182 359	0.183 382
	$\lambda = 2$	0.681 630	0.782 077	0.688 400	0.554 156
	$\lambda = 4$	3.007 367	1.856 046	0.889 771	0.318 677
	$\lambda = 8$	4.757 807	0.419 619	−0.284 002	−0.385 443
$n^2(\alpha - \tilde{\beta}_n)$	$\lambda = 1$	0.988 731	0.370 714	0.037 438	−0.129 136
	$\lambda = 2$	2.851 580	−1.776 830	−2.686 059	−2.468 150
	$\lambda = 4$	−20.107 224	−20.796 382	−10.384 765	−4.388 858
	$\lambda = 8$	−144.200 096	−13.815 051	0.373 053	0.747 002
		$k = 5$			
$n(\alpha - \beta_n)$	$\lambda = 1$	0.000 299	0.001 929	0.005 073	0.009 257
	$\lambda = 2$	0.008 114	0.044 062	0.097 558	0.150 669
	$\lambda = 4$	0.185 741	0.694 806	1.052 768	1.126 564
	$\lambda = 8$	2.891 219	4.290 922	2.499 797	0.893 353
$n^2(\alpha - \tilde{\beta}_n)$	$\lambda = 1$	0.018 808	0.055 607	0.087 602	0.107 105
	$\lambda = 2$	0.468 006	1.013 606	1.124 645	0.906 081
	$\lambda = 4$	8.653 572	7.869 467	0.434 585	−4.666 294
	$\lambda = 8$	70.164 359	−54.042 093	−47.939 426	−19.345 980

References

[1] BARTON, D. E. AND MALLOWS, C. L. (1965). Some aspects of the random sequence. *Ann. Math. Statist.* **36,** 236–260.

[2] CHAN, H. P. AND ZHANG, N. R. (2007). Scan statistics with weighted observations. *J. Amer. Statist. Assoc.* **102,** 595–602.

[3] CHANG, J. C., CHEN, R. J. AND HWANG, F. K. (2001). A minimal-automaton-based algorithm for the reliability of Con(d, k, n) system. *Methodol. Comput. Appl. Prob.* **3,** 379–386.

[4] FU, J. C. (2001). Distribution of the scan statistic for a sequence of bistate trials. *J. Appl. Prob.* **38,** 908–916.

[5] FU, J. C. AND KOUTRAS, M. V. (1994). Distribution theory of runs: a Markov chain approach. *J. Amer. Statist. Assoc.* **89,** 1050–1058.

[6] FU, J. C., WU, T. L. AND LOU, W. Y. W. (2012). Continuous, discrete, and conditional scan statistics. *J. Appl. Prob.* **49,** 199–209.

[7] GLAZ, J. AND NAUS, J. I. (1991). Tight bounds and approximations for scan statistic probabilities for discrete data. *Ann. Appl. Prob.* **1,** 306–318.

[8] GLAZ, J. AND NAUS, J. I. (2010). Scan statistics. In *Methods and Applications of Statistics in the Life and Health Sciences*, ed. N. Balakrishnan. John Wiley, New Jersey, pp. 733–747.

[9] GLAZ, J., NAUS, J. AND WALLENSTEIN, S. (2001). *Scan Statistics*. Springer, New York.

[10] GLAZ, J., POZDNYAKOV, V. AND WALLENSTEIN, S. (eds) (2009). *Scan Statistics*. Birkhäuser, Boston.

[11] HUFFER, F. W. AND LIN, C.-T. (1997). Computing the exact distribution of the extremes of sums of consecutive spacings. *Comput. Statist. Data Anal.* **26,** 117–132.

[12] HUFFER, F. W. AND LIN, C.-T. (1999). An approach to computations involving spacings with applications to the scan statistic. In *Scan Statistics and Applications*. eds J. Glaz and N. Balakrishnan, Birkhäuser, Boston, pp. 141–163.

[13] HUNTINGTON, R. J. AND NAUS, J. I. (1975). A simpler expression for Kth nearest neighbor coincidence probabilities. *Ann. Prob.* **3,** 894–896.

[14] HWANG, F. K. (1977). A generalization of the Karlin–McGregor theorem on coincidence probabilities and an application to clustering. *Ann. Prob.* **5,** 814–817.

[15] JANSON, S. (1984). Bounds on the distributions of extremal values of a scanning process. *Stoch. Process. Appl.* **18,** 313–328.

[16] KARLIN, S. AND MCGREGOR, G. (1959). Coincidence probabilities. *Pacific J. Math.* **9,** 1141–1164.

https://doi.org/10.1017/jpr.2016.101 Published online by Cambridge University Press

[17] Koutras, M. V. and Alexandrou, V. A. (1995). Runs, scans, and urn model distributions: a unified Markov chain approach. *Ann. Inst. Statist. Math.* **47,** 743–776.

[18] Loader, C. (1991). Large deviation approximations to the distribution of scan statistics. *Adv. Appl. Prob.* **23,** 751–771.

[19] Naus, J. I. (1982). Approximations for distributions of scan statistics. *J. Amer. Statist. Assoc.* **77,** 177–183.

[20] Neff, N. D. and Naus, J. I. (1980). *Selected Tables in Mathematical Statistics*, Vol. VI. American Mathematical Society, Providence.

[21] Siegmund, D. and Yakir, B. (2000). Tail probabilities for the null distribution of scanning statistics. *Bernoulli* **6,** 191–213.

[22] Wu, T. L., Glaz, J. and Fu, J. C. (2013). Discrete, continuous and conditional multiple window scan statistics. *J. Appl. Prob.* **50,** 1089–1101.

[23] Yao, Y.-C., Miao, D. W.-C. and Lin, X. C.-S. (2017). Corrected discrete approximations for the conditional and unconditional distributions of the continuous scan statistic. Preprint. Available at http://arxiv.org/abs/1602.02597.