Applied Probability Trust (14 March 2018)

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

* 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

epidemiology, genetics and neuroscience. See Glaz, Naus and Wallenstein [10] and Glaz and Naus [8] for a thorough review and comprehensive discussion of scan distribution theory, methods and applications. See also Glaz, Pozdnyakov and Wallenstein [9] 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

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

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

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

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 $\mathcal{H}_a : \lambda(t) = \lambda + \Delta \mathbf{1}_{(a,a+w]}(t)$ for (unknown) $0 \leq a \leq 1 - w$ 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 (*cf.* Karlin and McGregor [16] and Barton and Mallows [1]), Huntington and Naus [11] and Hwang [14] derived closed-form expressions for $P(k; N, w)$ which require to sum 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 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 $N \leq 25$. (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 $\lambda \leq 16$ 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 [12, 13] 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 Glaz and Naus [7] for related results in a discrete setting. Treating the problem as 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 Siegmund and Yakir [21] and Chan and Zhang [2].

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) contains either 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, $P^*(k;\lambda,w) = \mathbb{P}(S_w \geq k)$ may be approximated by $\mathbb{P}(S_w^{(n)} \geq k)$, which can be readily calculated using the Markov chain embedding method (*cf.* [4, 5, 17]). Indeed, it is known that $\mathbb{P}(S_w \geq k) - \mathbb{P}(S_w^{(n)} \geq k) = O(n^{-1})$ (*cf.* [6, 22]).

In Section 2, as n (multiple of q) tends to infinity, we derive the limit of $n[\mathbb{P}(S_w \geq k) - \mathbb{P}(S_w^{(n)} \geq k)]$, 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 $S_w'^{(n)}$) which is stochastically smaller than S_w and $S_w^{(n)}$. With a coupling device, we derive the limits of $n[\mathbb{P}(S_w \geq k) - \mathbb{P}(S_w'^{(n)} \geq k)]$ and $n[\mathbb{P}(S_w^{(n)} \geq k) - \mathbb{P}(S_w'^{(n)} \geq k)]$. In Section 3, using a change-of-measure argument, a similar result is obtained for the conditional probability $\mathbb{P}(S_w \geq k \mid |\Pi| = N)$. 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 *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 $\mathbb{P}(I_i^n = 0) = e^{-\lambda/n}$ and $\mathbb{P}(I_i^n = 1) = 1 - e^{-\lambda/n}$. 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. $\mathbb{E}(H_i^n) = \mathbb{E}\Big(\Big|\Pi \cap \Big(\frac{i-1}{n}, \frac{i}{n}\Big]\Big|\Big) = \frac{\lambda}{n}$. 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. $\mathbb{P}(I_i^n = 0) = \mathbb{P}\Big(\Big|\Pi \cap \Big(\frac{i-1}{n}, \frac{i}{n}\Big]\Big| = 0\Big) = e^{-\lambda/n}$. The two discrete scan statistics $S_w^{(n)}$ and $S_w'^{(n)}$ are now defined in terms of the two Bernoulli sequences as follows:

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

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 $\lim_{n\to\infty} n[\mathbb{P}(S_w \geq k) - \mathbb{P}(S_{w,I}^{(n)} \geq k)]$ and $\lim_{n\to\infty} n[\mathbb{P}(S_w \geq k) - \mathbb{P}(S_{w,H}^{(n)} \geq k)]$, 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 $I_i^n = \mathbf{1}\Big\{\Pi \cap \Big(\frac{i-1}{n}, \frac{i}{n}\Big] \neq \emptyset\Big\}$, $i = 1, \dots, n$. Thus, $(I_1^n, \dots, I_n^n)$ and Π are defined on the same probability space. In particular, $S_w \geq S_{w,I}^{(n)}$ 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 $\mathcal{A} = \mathcal{A}_{k,w} := \{S_w \geq k\}$ and $\mathcal{A}_n = \mathcal{A}_{n,k,w} := \{S_{w,I}^{(n)} \geq k\}$.

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 $0 < Q_1 < \cdots < Q_M < 1$. Further 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. $Q_j - Q_i \neq w$ for all $1 \leq i < j \leq M$). Define the functionals $\nu(\Pi) = \nu(\{Q_1, \dots, Q_M\})$ and $\tilde{\nu}(\Pi) =$

$\tilde{\nu}(\{Q_1,\dots,Q_M\})$ as follows:

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

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

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

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

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

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 $\tilde{I}_i^n = \left|\Pi \cap \left(\frac{i-1}{n}, \frac{i}{n}\right]\right|$, the number of Poisson points in the i-th subinterval. Then $\tilde{I}_i^n = 0$ implies $I_i^n = 0$ and $\tilde{I}_i^n \ge 1$ implies $I_i^n = 1$. Consider the following disjoint events

$$\begin{aligned}
\mathcal{G}_1 &= \{\tilde{I}_j^n \le 1, j = 1, \dots, n\}, \\
\mathcal{G}_{2,i} &= \{\tilde{I}_i^n = 2, \tilde{I}_j^n \le 1 \text{ for all } j \ne i\},\ i = 1, \dots, n, \\
\mathcal{G}_3 &= \{\tilde{I}_j^n = \tilde{I}_{j'}^n = 2 \text{ for some } j \ne j'\} \cup \{\tilde{I}_j^n \ge 3 \text{ for some } j\}.
\end{aligned}$$

We have

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

Claim that

$$\begin{aligned}
\mathbb{P}(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1) &= \frac{1}{2}\sum_{i=1}^{n\overline{w}} P_i^{(n)} + O(n^{-2}), & (2.2)\\
\sum_{i=1}^{n} \mathbb{P}(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_{2,i}) &= \sum_{i=1}^{n} \widetilde{P}_i^{(n)} + O(n^{-2}), & (2.3)\\
\mathbb{P}(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_3) &= O(n^{-2}), & (2.4)
\end{aligned}$$

where

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

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

Since $\mathbb{P}(\mathcal{G}_3) = O(n^{-2})$, (2.4) follows easily. To prove (2.2), note that when $\tilde{I}_i^n \le 1$ 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 $\tilde{I}_i^n \le 1$ 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 $\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1 = \cup_{i=1}^{n\overline{w}} \mathcal{G}_{1,i}$ where for $i = 1, \ldots, n\overline{w}$,

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

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 $\tilde{I}_1^n, \ldots, \tilde{I}_n^n$ 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 1/2, which implies $\mathbb{P}(\mathcal{G}_{1,i}) = \frac{1}{2}\mathbb{P}(\mathcal{G}'_{1,i})$. For $i = 1, \ldots, n\overline{w}$, 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 $\mathcal{G}''_{1,i} \backslash \mathcal{G}'_{1,i}$ is contained in $\{I_i^n = I_{i+nw}^n = 1,\ \tilde{I}_j^n \ge$

2 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 $\mathcal{G}''_{2,i} \backslash \mathcal{G}'_{2,i}$ 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 $i = 1, \ldots, n\overline{w}$, let $P_i'^{(n)} = \mathbb{P}(\mathcal{F}_i)$ where

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

Claim that

$$P_i^{(n)}/P_i'^{(n)} = \rho_n \quad \text{for all} \quad i = 1, \ldots, n\overline{w}, \tag{2.10}$$

where $\rho_n := \mathbb{P}(I_{i+nw}^n = 1)/\mathbb{P}(I_{i+nw}^n = 0) = (1 - e^{-\lambda/n})/e^{-\lambda/n} = e^{\lambda/n} - 1$. To establish the claim, recall that $P_i^{(n)} = \mathbb{P}(\mathcal{G}''_{1,i})$ where $\mathcal{G}''_{1,i}$ (*cf.* (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 that $h'_j = 0$ 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 *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 $(h'_1, \dots, h'_n) = (h_1, \dots, h_n) - \mathbf{e}_{i+nw}$ with $\mathbf{e}_{i+nw}$ being the vector of zeroes 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),

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

where

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

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

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

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 $\tilde{\rho}_n = \mathbb{P}(\tilde{I}_i^n = 2)/\mathbb{P}(I_i^n = 1) = e^{-\lambda/n}(\lambda/n)^2/[2(1-e^{-\lambda/n})]$. So,

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

where

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

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

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

Note that $\nu^{(n)}(\Pi)$ and $\tilde{\nu}^{(n)}(\Pi)$ converge a.s. 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

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

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

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 $\mathcal{B}_n := \{S_{w,H}^{(n)} \ge k\} = \left\{\max_{i=1,\dots,n\overline{w}+1} \sum_{r=i}^{i+nw-1} H_r^n \ge k\right\}$.

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

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

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 $\mathbb{P}(L_i^n = 0) = (1 - \lambda/n)e^{\lambda/n} = 1 - \mathbb{P}(L_i^n = 1)$. 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 $\mathcal{L}(\mathbf{V})$ denotes the law of a random vector $\mathbf{V}$, so that $\beta_n = \mathbb{P}(\mathcal{B}_n) = \mathbb{P}(\widetilde{\mathcal{B}}_n)$ where $\widetilde{\mathcal{B}}_n = \left\{\max_{i=1,\dots,n\overline{w}+1} \sum_{r=i}^{i+nw-1} \tilde{L}_r^n \ge k\right\}$. 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

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

we have

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

Claim that

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

which together with (2.14) yields the desired result.

It remains to establish the claim (2.15). Let Q be a random point which is uniformly distributed on $(0,1]$ and independent of Π. Let $\hat{I}_i^n = \mathbf{1}\Big\{(\Pi\cup\{Q\})\cap\Big(\frac{i-1}{n},\frac{i}{n}\Big]\neq\emptyset\Big\}$, $i=1,\ldots,n$. It is readily seen that $\mathcal{L}(\tilde{L}_1^n,\ldots,\tilde{L}_n^n\mid S_n=1)=\mathcal{L}(\hat{I}_1^n,\ldots,\hat{I}_n^n)$, which implies $\mathbb{P}(\widetilde{\mathcal{B}}_n|S_n=1)=\mathbb{P}(\hat{\mathcal{B}}_n)$, where $\hat{\mathcal{B}}_n=\Big\{\max_{i=1,\ldots,n\overline{w}+1}\sum_{r=i}^{i+nw-1}\hat{I}_r^n\geq k\Big\}$. Since $\mathbf{1}_{\hat{\mathcal{B}}_n}$ converges a.s. to $\mathbf{1}\{\max_{0\leq t\leq 1-w}|(\Pi\cup\{Q\})\cap(t,t+w]|\geq k\}$, we have

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

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

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

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

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 (*cf.* Remark 2.1) that $\alpha-\beta_n=C_\beta n^{-1}+O(n^{-2})$, where

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

3. The conditional case

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

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

Denoting by Π^N a set of N *i.i.d.* uniform random variables on $(0, 1]$, we have $\mathcal{L}(\Pi^N) = \mathcal{L}(\Pi | M = N)$ and $\gamma^{(N)} = \mathbb{P}(\mathcal{E}^N)$ where $\mathcal{E}^N := \left\{ \max_{0 \leq t \leq 1-w} \left| \Pi^N \cap (t, t+w] \right| \geq k \right\}$. 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 i-th 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 $\sum_{i=1}^n h_i = N$, $\mathbb{P}_N(J_i^n = h_i, i = 1, \dots, n) = 1/\binom{n}{N}$ where the subscript N in $\mathbb{P}_N$ signifies that there are N 1's in $J_1^n, \dots, J_n^n$. While in Section 2, $(I_1^n, \dots, I_n^n)$ is defined in terms of Π in order to make use of a coupling argument, 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 use 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

$$\gamma_n^{(N)} = \mathbb{P}_N(\mathcal{E}_n^N), \quad \text{where } \mathcal{E}_n^N := \left\{ \max_{i=1,\dots,n\overline{w}+1} \sum_{r=i}^{i+nw-1} J_r^n \;\geq\; k \right\}. \tag{3.1}$$

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

$$\lim_{n \to \infty} n(\gamma^{(N)} - \gamma_n^{(N)}) = \frac{1}{2} N(N-1)(\gamma^{(N-1)} - \gamma^{(N)}) + \frac{1}{2} N \mathbb{E}\left[\nu(\Pi) + \tilde{\nu}(\Pi) \,\middle|\, M = N-1 \right].$$

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 details. For notational simplicity, the superscript N in $\mathcal{E}^N$ and $\mathcal{E}_n^N$ is suppressed while 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 $\tilde{J}_i = \left| \Pi^N \cap ((i-1)/N, i/N] \right|$, $i = 1, \dots, n$, and define the (disjoint) events

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

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

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

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

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

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 $\tilde{J}_i^n \leq 1$ 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 $i = 1, \ldots, n\overline{w}$,

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

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

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

where $U'_{1,i} = \widetilde{\mathcal{E}}_n^c \cap \left\{\tilde{J}_j^n \leq 1 \text{ for all } j,\ \sum_{r=i+1}^{i+nw-1} \tilde{J}_r^n = k-2,\ \tilde{J}_i^n = \tilde{J}_{i+nw}^n = 1\right\}$, $i = 1, \ldots, n\overline{w}$. 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 $1/2$, which implies $\mathbb{P}_N(U_{1,i}|U_1) = \frac{1}{2}\mathbb{P}_N(U'_{1,i}|U_1)$.

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

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

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

We will simplify $\sum_{i=1}^{n\overline{w}} \mathbb{P}_N(U'_{1,i}|U_1) = \sum_{i=1}^{n\overline{w}} \mathbb{P}_N(V_i)$ 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 $r = 1, \ldots, n$, $\sum_{r=1}^n h_r = N, \max_{j=1,\ldots,n\overline{w}+1} \sum_{r=j}^{j+nw-1} h_r < k, \sum_{r=i+1}^{i+nw-1} h_r = k-2, h_i = h_{i+nw} = 1$. Let $V_i^* = \{\sum_{r=1}^n J_r^n = N-1, \max_{j=1,\ldots,n\overline{w}+1} \sum_{r=j}^{j+nw-1} J_r^n < k, \sum_{r=i+1}^{i+nw-1} J_r^n = k-2, J_i^n = 1, J_{i+nw}^n = 0,$ sum of any nw consecutive J_r^n including $r = i+nw$ is at most $k-2\}$. We interpret 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 that $h_r^* = 0$ or 1 for $r = 1, \ldots, n$, $\sum_{r=1}^n h_r^* = N-1, \sum_{r=j}^{j+nw-1} h_r^* < k$ for $j = 1, \ldots, n\overline{w}+1$, $\sum_{r=i+1}^{i+nw-1} h_r^* = k-2$, $h_i^* = 1, h_{i+nw}^* = 0$, and 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 with the $(i+nw)$-th entry replaced by 0, it is in V_i^*. Note that the number of nonzero entries for a configuration in V_i^* equals $N-1$. Recall that the notation $\mathbb{P}_N$ ($\mathbb{P}_{N-1}$, *resp.*) denotes the probability measure for $(J_1^n, \ldots, J_n^n)$ with $\sum_{r=1}^n J_r^n = N$ ($\sum_{r=1}^n J_r^n = N-1$, *resp.*). 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,

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

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

By (3.3)–(3.6),

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

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

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

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

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

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

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 $p_n = 1-e^{-\lambda/n}$ 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 (*cf.* [4, 6, 17]), we computed 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). Figure 1 plots $n(\alpha - \alpha_n), n(\alpha-\beta_n)$ and $n(\gamma^{(N)} - \gamma_n^{(N)})$ for $n = 25(5)600$ with $k = 5, w = 0.4, \lambda = 8$ and $N = 8$, where the superscript (N) in $\gamma^{(N)}$ and $\gamma_n^{(N)}$ is suppressed for ease of notation. The exact probabilities $\alpha = P^*(k;\lambda,w) = P^*(5;8,0.4) = 0.628144085$ and $\gamma^{(8)} = P(k;N,w) = P(5;8,0.4) = 0.780861440$ are taken from [20]. By Theorems 2.1, 2.2 and 3.1, $n(\alpha-\alpha_n), n(\alpha-\beta_n)$ and $n(\gamma^{(N)}-\gamma_n^{(N)})$ converge, respectively, to the limits C_α, C_β and C_γ (*cf.* Remarks 2.1, 2.2 and 3.1). These limits were estimated by Monte Carlo simulation with 10^6 replications, resulting in $C_\alpha = 4.6322 \pm 0.0096$(Std. Err.), $C_\beta = 0.8297 \pm 0.0167, C_\gamma = 2.7279 \pm 0.0114$. 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 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 $\gamma^{(N)} - \tilde{\gamma}_n^{(N)} = O(n^{-2})$. Table 1 presents numerical results comparing $\alpha_n, \tilde{\alpha}_n, \beta_n$ and $\tilde{\beta}_n$ for the unconditional case. Table 2 compares $\gamma_n^{(N)}$ and $\tilde{\gamma}_n^{(N)}$ for the conditional case.

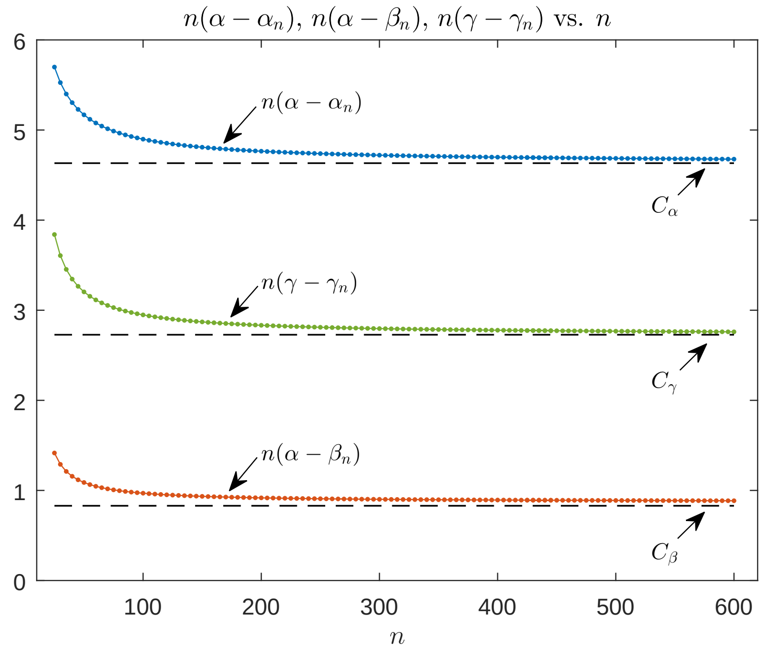

FIGURE 1: Plot of $n(\alpha-\alpha_n)$, $n(\alpha-\beta_n)$, and $n(\gamma-\gamma_n)$ for $n = 25(5)600$ with parameters $w = 0.4, k = 5, \lambda = 8, N = 8$.

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 $w < 0.2$. Figure 1 shows 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 2 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 4. Our limited numerical studies indicate that the corrected approximations are more accurate than the uncorrected ones for $n \geq 50$. Also in the two tables, $\tilde{\beta}_{100}$ ($\tilde{\gamma}_{100}^{(N)}$, *resp.*) is about as accurate as or more accurate than β_{400} ($\gamma_{400}^{(N)}$, *resp.*).

Remark 4.2. The referee of this paper raises an important question on the relationship among w, λ (N) and the convergence rate. While the convergence rate for the uncorrected (corrected, *resp.*) approximations is n^{-1} (n^{-2}, *resp.*), the error size for the approximations $\alpha_n, \beta_n, \gamma_n^{(N)}, \tilde{\alpha}, \tilde{\beta}_n, \tilde{\gamma}_n^{(N)}$ depends on w and λ (N) as well as on k in a

Table 1: The unconditional case

Parameters				n					Exact
λ	w	k		25	50	100	200	400	α
4	0.2	3	α_n	0.226474137	0.297081029	0.330413369	0.346549002	0.354481473	0.362322986
			$\alpha - \alpha_n$	0.135848849	0.065241957	0.031909617	0.015773984	0.007841513	
			$\tilde{\alpha}_n$		0.367687921	0.363745709	0.362684635	0.362413943	
			$\alpha - \tilde{\alpha}_n$		−0.005364935	−0.001422723	−0.000361649	−0.000090957	
			β_n	0.269466265	0.321289109	0.342964036	0.352912780	0.357682871	
			$\alpha - \beta_n$	0.092856721	0.041033877	0.019358950	0.009410206	0.004640115	
			$\tilde{\beta}_n$		0.373111952	0.364638964	0.362861523	0.362452963	
			$\alpha - \tilde{\beta}_n$		−0.010788966	−0.002315978	−0.000538537	−0.000129977	
4	0.2	4	α_n	0.028528199	0.063252847	0.083861016	0.094813938	0.100432989	0.106139839
			$\alpha - \alpha_n$	0.077611640	0.042886992	0.022278823	0.011325901	0.005706850	
			$\tilde{\alpha}_n$		0.097977495	0.104469184	0.105766860	0.106052039	
			$\alpha - \tilde{\alpha}_n$		0.008162344	0.001670655	0.000372979	0.000087800	
			β_n	0.037826080	0.071921990	0.089167692	0.097701122	0.101933685	
			$\alpha - \beta_n$	0.068313759	0.034217849	0.016972147	0.008438717	0.004206154	
			$\tilde{\beta}_n$		0.106017899	0.106413395	0.106234551	0.106166248	
			$\alpha - \tilde{\beta}_n$		0.000121940	−0.000273556	−0.000094712	−0.000026409	
8	0.4	5	α_n	0.400190890	0.524770327	0.579159623	0.604320002	0.616397532	0.628144085
			$\alpha - \alpha_n$	0.227953195	0.103373758	0.048984462	0.023824083	0.011746553	
			$\tilde{\alpha}_n$		0.649349765	0.633548918	0.629480382	0.628475061	
			$\alpha - \tilde{\alpha}_n$		−0.021205680	−0.005404833	−0.001336297	−0.000330976	
			β_n	0.571524668	0.606381317	0.618451977	0.623556407	0.625910702	
			$\alpha - \beta_n$	0.056619417	0.021762768	0.009692108	0.004587678	0.002233383	
			$\tilde{\beta}_n$		0.641237966	0.630522637	0.628660836	0.628264997	
			$\alpha - \tilde{\beta}_n$		−0.013093881	−0.002378552	−0.000516751	−0.000120912	
8	0.4	6	α_n	0.156407681	0.278520053	0.341202440	0.372097133	0.387351968	0.402452588
			$\alpha - \alpha_n$	0.246044907	0.123932535	0.061250148	0.030355455	0.015100620	
			$\tilde{\alpha}_n$		0.400632426	0.403884826	0.402991826	0.402606803	
			$\alpha - \tilde{\alpha}_n$		0.001820162	−0.001432238	−0.000539238	−0.000154215	
			β_n	0.278663391	0.351874806	0.379351117	0.391387631	0.397034846	
			$\alpha - \beta_n$	0.123789197	0.050577782	0.023101471	0.011064957	0.005417742	
			$\tilde{\beta}_n$		0.425086221	0.406827428	0.403424144	0.402682062	
			$\alpha - \tilde{\beta}_n$		−0.022633633	−0.004374840	−0.000971556	−0.000229474	

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 is 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 $\binom{nw}{k-1} + 1$,

TABLE 2: The conditional case

Parameters				n					Exact
w	k	N		25	50	100	200	400	γ
0.2	4	6	γ_n	0.080688876	0.155913836	0.194799457	0.214242757	0.223935622	0.233600000
			$\gamma-\gamma_n$	0.152911124	0.077686164	0.038800543	0.019357243	0.009664378	
			$\widetilde{\gamma}_n$		0.231138796	0.233685077	0.233686058	0.233628487	
			$\gamma-\widetilde{\gamma}_n$		0.002461204	−0.000085077	−0.000086058	−0.000028487	
		7	γ_n	0.166798419	0.294914521	0.354660825	0.383179030	0.397084766	0.410752000
			$\gamma-\gamma_n$	0.243953581	0.115837479	0.056091175	0.027572970	0.013667234	
			$\widetilde{\gamma}_n$		0.423030623	0.414407129	0.411697234	0.410990502	
			$\gamma-\widetilde{\gamma}_n$		−0.012278623	−0.003655129	−0.000945234	−0.000238502	
		8	γ_n	0.291588655	0.469102180	0.542215920	0.575193126	0.590840920	0.605949440
			$\gamma-\gamma_n$	0.314360785	0.136847260	0.063733520	0.030756314	0.015108520	
			$\widetilde{\gamma}_n$		0.646615704	0.615329660	0.608170332	0.606488715	
			$\gamma-\widetilde{\gamma}_n$		−0.040666264	−0.009380220	−0.002220892	−0.000539275	
		9	γ_n	0.448718168	0.651907101	0.723414803	0.753448974	0.767220941	0.780225536
			$\gamma-\gamma_n$	0.331507368	0.128318435	0.056810733	0.026776562	0.013004595	
			$\widetilde{\gamma}_n$		0.855096034	0.794922506	0.783483145	0.780992907	
			$\gamma-\widetilde{\gamma}_n$		−0.074870498	−0.014696970	−0.003257609	−0.000767371	
0.4	5	6	γ_n	0.162450593	0.224402953	0.254838093	0.269838436	0.277276942	0.284672000
			$\gamma-\gamma_n$	0.122221407	0.060269047	0.029833907	0.014833564	0.007395058	
			$\widetilde{\gamma}_n$		0.286355312	0.285273233	0.284838778	0.284715449	
			$\gamma-\widetilde{\gamma}_n$		−0.001683312	−0.000601233	−0.000166778	−0.000043449	
		7	γ_n	0.371395881	0.463718058	0.504028994	0.522865538	0.531971853	0.540876800
			$\gamma-\gamma_n$	0.169480919	0.077158742	0.036847806	0.018011262	0.008904947	
			$\widetilde{\gamma}_n$		0.556040235	0.544339929	0.541702083	0.541078168	
			$\gamma-\widetilde{\gamma}_n$		−0.015163435	−0.003463129	−0.000825283	−0.000201368	
		8	γ_n	0.627251924	0.716788906	0.751379277	0.766696715	0.773916208	0.780861440
			$\gamma-\gamma_n$	0.153609516	0.064072534	0.029482163	0.014164725	0.006945232	
			$\widetilde{\gamma}_n$		0.806325887	0.785969648	0.782014154	0.781135700	
			$\gamma-\widetilde{\gamma}_n$		−0.025464447	−0.005108208	−0.001152714	−0.000274260	
		9	γ_n	0.864220071	0.918826852	0.936992617	0.944511915	0.947942424	0.951173120
			$\gamma-\gamma_n$	0.086953049	0.032346268	0.014180503	0.006661205	0.003230696	
			$\widetilde{\gamma}_n$		0.973433633	0.955158381	0.952031214	0.951372933	
			$\gamma-\widetilde{\gamma}_n$		−0.022260513	−0.003985261	−0.000858094	−0.000199813	

which is enormous when nw is large and k is not small. (It should be remarked that [3] is 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 method of proof 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 nonhomoge-

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

		$k=3$				$k=5$			
		$w=0.1$	$w=0.2$	$w=0.3$	$w=0.4$	$w=0.1$	$w=0.2$	$w=0.3$	$w=0.4$
$n(\alpha-\beta_n)$	$\lambda=1$	0.109466	0.162936	0.182359	0.183382	0.000299	0.001929	0.005073	0.009257
	$\lambda=2$	0.681630	0.782077	0.688400	0.554156	0.008114	0.044062	0.097558	0.150669
	$\lambda=4$	3.007367	1.856046	0.889771	0.318677	0.185741	0.694806	1.052768	1.126564
	$\lambda=8$	4.757807	0.419619	−0.284002	−0.385443	2.891219	4.290922	2.499797	0.893353
$n^2(\alpha-\tilde{\beta}_n)$	$\lambda=1$	0.988731	0.370714	0.037438	−0.129136	0.018808	0.055607	0.087602	0.107105
	$\lambda=2$	2.851580	−1.776830	−2.686059	−2.468150	0.468006	1.013606	1.124645	0.906081
	$\lambda=4$	−20.107224	−20.796382	−10.384765	−4.388858	8.653572	7.869467	0.434585	−4.666294
	$\lambda=8$	−144.200096	−13.815051	0.373053	0.747002	70.164359	−54.042093	−47.939426	−19.345980

neous 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 (*cf.* [22]).

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. Probab.* **3**, 379–386.

[4] Fu, J.C. (2001). Distribution of the scan statistic for a sequence of bistate trials. *J. Appl. Probab.* **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. and Wu, T.L. and Lou, W.Y.W. (2012). Continuous, discrete, and conditional scan statistics. *J. Appl. Probab.* **49**, 199–209.

[7] Glaz, J. and Naus, J.I. (1991). Tight bounds and approximations for scan statistic probabilities for discrete data. *Ann. Appl. Probab.* **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* (N. Balakrishnan ed.), 733–747. John Wiley & Sons, New Jersey.

[9] Glaz, J. and Pozdnyakov, V. and Wallenstein, S. (2009). *Scan Statistics: Methods and Applications.* Birkhauser, Boston.

[10] Glaz, J. and Naus, J. and Wallenstein, S. (2001). *Scan Statistics.* Springer, New York.

[11] Huntington, R.J. and Naus, J.I. (1975). A simpler expression for Kth nearest neighbor coincidence probabilities. *Ann. Probab.* **3**, 894–896.

[12] 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.

[13] 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* (J. Glaz and N. Balakrishnan eds.), 141–163. Birkhauser, Boston.

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

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

[16] Karlin, S. and McGregor, G. (1959). Coincidence probabilities. *Pacific J. Math.* **9**, 1141–1164.

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

[18] Loader, C. (1991). Large deviation approximations to the distribution of scan statistics. *Adv. Appl. Probab.* **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, Rhode Island.

[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. Probab.* **50**, 1089–1101.

[23] Yao, Y.-C., Miao, D. W.-C. and Lin, X. C.-S. (2016). Corrected discrete approximations for the conditional and unconditional distributions of the continuous scan statistic. http://arxiv.org/abs/1602.02597