Statistica Sinica **23** (2013), 1541-1551
doi:http://dx.doi.org/10.5705/ss.2012.062s

THE MAXIMUM OF A RATCHET SCANNING PROCESS OVER A POISSON RANDOM FIELD

I-Ping Tu

Academia Sinica

Abstract: Scan statistic is a popular method in searching non-random clusters over some random field. Motivated by a high energy particle detection problem, we are interested in a ratchet scan statistic whose scanning window is a grid set, instead of a rectangular box. We further generalize it to an r-dimensional problem for any $r \in \mathbb{Z}^+$ and provide a tail probability approximation for a ratchet scan statistic over an r-dimensional homogeneous Poisson process. We show that the ratchet effect can be factored out as an overshoot function ν in each dimension.

Key words and phrases: Cosmic ray, high dimension, Poisson process, ratchet, scan statistic, tail probability approximation.

1. Introduction

We are interested in the tail distribution of a rachet scan statistic over an r-dimensional Poisson random field. The ratchet scan statistic $M_{\mathbf{L},\boldsymbol{\Delta},\mathbf{w}}$ is defined as follows:

$$M_{\mathbf{L},\boldsymbol{\Delta},\mathbf{w}} = \max_{\mathbf{t}\in T_{\mathbf{L},\boldsymbol{\Delta}}} X_{\mathbf{w}}(\mathbf{t}), \tag{1.1}$$

where $X_{\mathbf{w}}(\mathbf{t}) \equiv N(\mathbf{A}_{\mathbf{w}}(\mathbf{t}))$ counts the events occurred in the scanning window $\mathbf{A}_{\mathbf{w}}(\mathbf{t}) = \prod_{1\le i\le r}[t_i, t_i + w_i)$, and we model N as a counting process over an r-dimensional homogeneous Poisson process with rate λ_0 under the null hypothesis. $T_{\mathbf{L},\boldsymbol{\Delta}} = \prod_{i=1}^{r}([0, L_i] \bigcap \Delta_i\mathbb{Z}_i)$ is a grid set and $\mathbb{Z} = (\mathbb{Z}_1, \ldots, \mathbb{Z}_r)$ is the r-dimensional integer set.

Scan statistics over a Poisson random field have been applied in molecular biology (Chan and Zhang (2007)), epidemiology, and geostatistics (Kulldorff (1997); Glaz, Naus, and Wallenstein (2001)) to search non-random clusters. The tail distributions of these statistics are of interest and their approximations have been studied in the literature (Chan (2009)); Chan and Zhang (2007); Siegmund and Yakir (2000); Alm (1997); Rabinowitz and Siegmund (1997); Loader (1991); Naus (1965)). While interest may be in various forms of the statistic $X_{\mathbf{w}}(\mathbf{t})$, the common point is that their scanning windows are over a rectangular box. Akin to (1.1), the scan statistic with index over a rectangular box is

$$M_{\mathbf{L},\mathbf{w}} = \max_{\mathbf{t}\in T_{\mathbf{L}}} X_{\mathbf{w}}(\mathbf{t}), \tag{1.2}$$

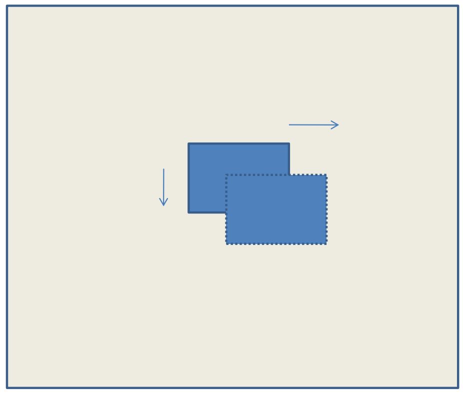

Figure 1. The detector scans the whole sky by shifting the window by size Δ at each step in either the longitudinal or latitudinal directions.

where $T_{\mathbf{L}} = \prod_{i=1}^{r}([0, L_i])$.

Motivated by a high energy particle detection problem, we study the tail probability of the ratchet scan statistic (1.1) that captures the characteristic of a step-shifting scan of a detector (Glaz, Naus, and Wallenstein (2001, p.68-71)). As high energy cosmic rays travel into our atmosphere, pion particles are generated through nuclear reactions. Pion, a short-lived heavy particle, further decays into muons, which are charged leptons that can be detected relatively easily. By scanning muon particles across the whole sky, showers of cosmic rays coming from a specific direction can be detected. For instance, the Soudan-2 detector, an underground muon telescope, was designed to search for cosmic ray muon sources or hot spots in the sky. The search covers the sky region from $-m_1$ to m_2 declination degrees and all right accessions. Each observation counts the total number of particles within a window size $w_1 \times w_2$ in a fixed length of time. The detector scans the whole sky through shifting the window by size Δ at each step in either the longitudinal or latitudinal directions, see Figure 1.

In this paper, we provide a p-value approximation for the ratchet scan statistic (1.1). We show that the ratchet effect in the approximation can be factored out as a function ν (Tu (2009)) in each dimension. When the shift size $\Delta \to 0$, the overshoot correction function $\nu \to 1$, and the tail probability approximation for (1.1) converges to that for (1.2), which is consistent with the result in Chan (2009), Chan and Zhang (2007), Alm (1997), and Aldous (1989). We take the last time conditioning approach developed by Woodroofe (1976) to do the approximation. In developing the overshoot correction term for higher dimensions, we generalize the two-dimensional conditional distribution approximation in Siegmund (1988) (Theorem 1). We first show the approximation for $r = 1$, then apply the induction rule to get the results for any positive integer r.

This paper is organized as follows. In Section 2, we state the main results. In Section 3, we give a sketch of the proof. In Section 4, we report on a simulation study. The paper ends with a brief discussion. We provide a computing formula for the overshoot correction function ν in the Appendix.

2. Results

Our results give the asymptotic distributions for $M_{\mathbf{L},\boldsymbol{\Delta},\mathbf{w}}$. To simplify the notation, we assume torus boundary conditions to avoid boundary effects, and let $w_i = w$, $L_i = L$, and $\Delta_i = \Delta$.

Theorem 1. *Suppose $b \to \infty$, $\lambda_1 := \lambda_1(b, \mathbf{w}) = b/w^r \in (\lambda_0 + \delta, 1/\delta)$ for some small δ, and $1 \gg \frac{w}{L} > \epsilon$ for some $\epsilon > 0$ and $w \gg \Delta$. Then, for $r \geq 1$, we have*

$$P\{M_{\mathbf{L},\boldsymbol{\Delta},\mathbf{w}} \geq b\} \sim (1 - \frac{\lambda_0}{\lambda_1})^{2r-1} \cdot b^r \frac{\exp[-b\theta + w^r \cdot (\lambda_1 - \lambda_0)]}{\sqrt{2\pi b}} \left((\frac{L}{w}) \nu(\theta, \alpha) \right)^r, \tag{2.1}$$

where $\nu(\theta, \alpha) = (1 - E_\theta \exp(-\theta S_{\tau_+}))/((1 - e^{-\theta}) E_\theta S_{\tau_+})$, $\theta = \log(\lambda_1/\lambda_0)$, and $\alpha = (\lambda_0, b, w, \Delta)$ parametrizes the local random walk $\{S_n, n > 1\}$, with S_n's i.i.d. sums of the difference of independent Poissons rates $(\lambda_1 w^{r-1}\Delta)$ and $(\lambda_0 w^{r-1}\Delta)$, and $\tau_+ = \inf\{n : S_n > 0\}$.

Remark 1. $\nu(\theta, \alpha)$ is a correction term for the overshoot of the process S_n to exceed b. S_n's are i.i.d. partial sum of two independent Poissons with mean greater than 0, which means that $P(S_{\tau_+} \geq 1) = 1$ by the definition of τ_+. In the case of no overshoot $P(S_{\tau_+} = 1) = 1$, $\nu(\theta, \alpha) = 1$. It can be shown that $0 < \nu(\theta, \alpha) \leq 1$ for non-negative θ. The computation formula can be found in Woodroofe (1979) for non-arithmetic iid random variables and in Tu (2009) for arithmetic iid random variables.

Remark 2. The condition on the lower bound of w/L ensures that the expression (2.1) is a small number, but it is not a crucial assumption in this approximation. If w/L is so small that (2.1) is greater than .1, we can rewrite it as $1 - \exp(-\text{expression (2.1)})$.

Theorem 2. *If $\Delta \to 0$,*

$$P\{M_{\mathbf{L},\mathbf{w}} \geq b\} \sim (1 - \frac{\lambda_0}{\lambda_1})^{2r-1} \cdot b^r \frac{\exp[-b\theta + w^r \cdot (\lambda_1 - \lambda_0)]}{\sqrt{2\pi b}} \left(\frac{L}{w} \right)^r. \tag{2.2}$$

This result is consistent with those in Chan (2009) Chan and Zhang (2007), Alm (1997), and Aldous (1989).

Remark 3. The difference between (2.1) and (2.2) is the overshoot correction function $\nu(\theta, \alpha)$. When the ratchet effect is ignored, (2.2) is usually employed to do the approximation for $P\{M_{\mathbf{L},\boldsymbol{\Delta},\mathbf{w}} \geq b\}$.

Theorem 3. *Let $\lambda_0 = n/L^r$ and $B(n,b,p) = \binom{n}{b} p^b(1-p)^{n-b}$. If the total number of events n is known, then*

$$P\{M_{\mathbf{L},\mathbf{\Delta},\mathbf{w}} \geq b \mid N(T_{\mathbf{L},\mathbf{\Delta}}) = n\} \sim (1-\frac{\lambda_0}{\lambda_1})^{2r-1} \cdot b^r B(n,b,\frac{w^r}{L^r}) \left((\frac{L}{w})\nu(\theta,\alpha)\right)^r, \quad (2.3)$$

$$P\{M_{\mathbf{L},\mathbf{w}} \geq b \mid N(T_{\mathbf{L}}) = n\} \sim (1-\frac{\lambda_0}{\lambda_1})^{2r-1} \cdot b^r B(n,b,\frac{w^r}{L^r}) \left(\frac{L}{w}\right)^r. \quad (2.4)$$

Remark 4. The conditional distribution of $\{X_{\mathbf{w}}(t) \mid N(L)=n\}$ is binomial(n,p), where $p = w^r/L^r$. By Lemma 1 in Section 3, it can be observed that $P_0\{X_{\mathbf{w}}(t) = b+k\} \sim P_0\{X_{\mathbf{w}}(t) = b\} \cdot e^{-\theta k}$. This is true as long as $X_{\mathbf{w}}(t)$ belongs to an exponential family. Then $(\exp[-b\theta + w^r \cdot (\lambda_1 - \lambda_0)])/(\sqrt{2\pi b})$ in (2.1) and (2.2) is replaced by $B(n,b,p)$, which leads to (2.3) and (2.4).

3. Sketch of the Proof for Theorem 1

Our approximation applies the last time conditioning approach developed by Woodroofe (1976). The tail probability can be written as a product of the probability of a fixed time event and a conditional probability of a maximum of the process. Let $\mathbf{i} = (i_1,\ldots,i_r)$ be the r-dimensional index. We write $\mathbf{i} > \mathbf{j}$ if and only if $\{i_r > j_r\}$ or $\{i_{r-1} > j_{r-1}\} \cap \{i_r = j_r\}$, or $\ldots$ or $\{i_1 > j_1\} \cap \{i_2 = j_2\} \cap \cdots \cap \{i_r = j_r\}$. Let $D(\mathbf{i_0}) = \{\mathbf{i} \mid \mathbf{i} > \mathbf{i_0}\}$.

Then

$$P\{\max_{\mathbf{t}\in T_{\mathbf{L},\mathbf{\Delta}}} X_{\mathbf{w}}(\mathbf{t}) \geq b\}$$
$$\sim \sum_{\mathbf{t_0}\in T_{\mathbf{L},\mathbf{\Delta}}} \sum_{k=0}^{\infty} P\{X_{\mathbf{w}}(\mathbf{t_0}) = b+k\} P\{\max_{\mathbf{t}\in D(\mathbf{t_0})} X_{\mathbf{w}}(\mathbf{t}) - X_{\mathbf{w}}(\mathbf{t_0}) < -k \mid X_{\mathbf{w}}(\mathbf{t_0}) = b+k\}. \quad (3.1)$$

We need some lemmas to complete the proof.

Lemma 1. *Let Y be a Poisson random variable with mean $w^r\lambda_0$. Let $\theta = \log(\lambda_1/\lambda_0)$, $\lambda_1 = b/w^r$, and $A_r = \exp[-b\theta + w^r \cdot (\lambda_1 - \lambda_0)]/\sqrt{2\pi b}$. Then*

$$P\{Y = b+k\} \sim A_r e^{-\theta k}. \quad (3.2)$$

Proof of Lemma 1. This approximation uses Stirling's formula.

$$P\{Y = b+k\} = \frac{(w^r\lambda_0)^{b+k}}{(b+k)!} e^{-w^r\lambda_0}$$
$$\sim \frac{(w^r\lambda_0)^{b+k}}{\sqrt{2\pi(b+k)}\,(b+k)^{b+k}\,e^{-(b+k)}} e^{-w^r\lambda_0}$$

$$\sim \frac{(w^r\lambda_0)^{b+k}}{\sqrt{2\pi b}\, b^{b+k}\, e^k\, e^{-(b+k)}} e^{-w^r\lambda_0}$$
$$= \frac{\exp[-b\theta + w^r \cdot (\lambda_1 - \lambda_0)]}{\sqrt{2\pi b}} e^{-\theta k}$$

Lemma 2. *Let $x_i^{(m)}$ be the difference of independent Poissons with parameters $w^{r-1}\Delta\lambda_0$ and $w^{r-1}\Delta\lambda_1$, let $S_{i_m}^{(m,+)}$ be the $i_m{}^{th}$ partial sum of i.i.d. $x_{i_m}^{(m)}$'s, and let $S_{i_m}^{(m,-)}$'s be i.i.d. copies of $S_{i_m}^{(m,+)}$ for all m and i_m. If $B_0(k)$ is a complete set, $B_1(k) = \{\max_{i_1>0} S_{i_1}^{(1,+)} < -k\}$, and*

$$B_r(k) = B_{r-1}(k) \cap \{ \max_{\substack{i_1,\ldots,i_{r-1}\geq 0 \\ i_r>0}} S_{i_1}^{(1,-)} + S_{i_2}^{(2,-)} + \cdots + S_{i_{r-1}}^{(r-1,-)} + S_{i_r}^{(r,+)} < -k \}, \quad (3.3)$$

then the distribution of

$$\{ \max_{\mathbf{t}\in D(\mathbf{t_0})} X_{\mathbf{w}}(\mathbf{t}) - X_{\mathbf{w}}(\mathbf{t_0}) < -k \mid X_{\mathbf{w}}(\mathbf{t_0}) = b+k \} \quad (3.4)$$

can be approximated by the distribution of the set $B_r(k)$.

Proof of Lemma 2. The proof of Lemma 2 for any r comes by extending the method to prove the case for $r = 2$ in Siegmund (1988) and $r = 3$ in Loader (1991).

By our definition, $D(\mathbf{t_0}) = \{\mathbf{t} \in T_{\mathbf{L},\mathbf{\Delta}} \mid \{t_r > t_{0r}\}$ or $\{t_{r-1} > t_{0r-1}\} \cap \{t_r = t_{0r}\}$ or $\ldots$ or $\{t_1 > t_{01}\} \cap \{t_2 = t_{02}\} \cap \cdots \cap \{t_r = t_{0r}\}\}$, and each index $\mathbf{t} \in D(\mathbf{t_0})$ has to satisfy (3.4). $\{\mathbf{t} \in T_{\mathbf{L},\mathbf{\Delta}} \mid t_r > t_{0r}\}$ means that $\{\mathbf{t} \in T_{\mathbf{L},\mathbf{\Delta}} \mid t_r > t_{0r}, \forall t_1, \ldots, t_{r-1}\}$. By Siegmund (1988), we can approximate this conditional set for the index set $\{\mathbf{t} \in T_{\mathbf{L},\mathbf{\Delta}} \mid t_r > t_{0r}\}$ as

$$\{ \max_{\substack{i_1,\ldots,i_{r-1}\geq 0 \\ p\in\{+,-\} \\ i_r>0}} S_{i_1}^{(1,p)} + S_{i_2}^{(2,p)} + \cdots + S_{i_{r-1}}^{(r-1,p)} + S_{i_r}^{(r,+)} < -k \}.$$

Using the same argument, we can approximate this conditional distribution, with index $\{\mathbf{t} \in T_{\mathbf{L},\mathbf{\Delta}} \mid \{t_{r_0} > t_{0r_0}\} \cap \{t_{r_0+1} = t_{0r_0+1}\} \cap \cdots \cap \{t_r = t_{0r}\}\}$, as

$$\{ \max_{\substack{i_1,\ldots,i_{r_0-1}\geq 0 \\ p\in\{+,-\} \\ i_{r_0}>0}} S_{i_1}^{(1,p)} + S_{i_2}^{(2,p)} + \cdots + S_{i_{r_0-1}}^{(r_0-1,p)} + S_{i_{r_0}}^{(r_0,+)} < -k \}.$$

Thus, we have

$$B_r(k) = B_{r-1}(k) \cap \{ \max_{\substack{i_1,\ldots,i_{r-1}\geq 0 \\ p\in\{+,-\} \\ i_r>0}} S_{i_1}^{(1,p)} + S_{i_2}^{(2,p)} + \cdots + S_{i_{r-1}}^{(r-1,p)} + S_{i_r}^{(r,+)} < -k \}. \quad (3.5)$$

From (3.5), it can be observed that

$$\{\max_{\substack{i_1,\dots,i_{r-1}\geq 0\\ i_r>0}} S_{i_1}^{(1,+)} + S_{i_2}^{(2,+)} + \cdots + S_{i_{r-1}}^{(r-1,+)} + S_{i_r}^{(r,+)} < -k\}$$

is implied by

$$\{\max_{i_1>0} S_{i_1}^{(1,+)} < -k\} \bigcap \ \cdots \ \bigcap \{\max_{i_r>0} S_{i_r}^{(r,+)} < -k\}.$$

So, we can simplify (3.5) as

$$B_r(k) = B_{r-1}(k) \cap \{\max_{\substack{i_1,\dots,i_{r-1}\geq 0\\ i_r>0}} S_{i_1}^{(1,-)} + S_{i_2}^{(2,-)} + \cdots + S_{i_{r-1}}^{(r-1,-)} + S_{i_r}^{(r,+)} < -k\}.$$

Lemma 3. *Let S_n be the partial sum of i.i.d. x_i's, with x_i the difference of Poissons with parameters $w^{r-1}\Delta\lambda_0$ and $w^{r-1}\Delta\lambda_1$. Let P_θ be the exponential tilted measure of the original measure P_0, with $\theta = \log(\lambda_1/\lambda_0)$. Then*

$$P_\theta\{S_n = l\} = P_0\{S_n = l\} \cdot e^{\theta l} \ \ \textit{and} \ \ P_0\{x_i = l\} = P_\theta\{x_i = -l\}. \tag{3.6}$$

Proof of Lemma 3. The proof of Lemma 3 uses a change of measure method. Let $x_i = a_i - b_i$, where a_i and b_i are i.i.d. Poissons with means $w^{r-1}\Delta\lambda_0$ and $w^{r-1}\Delta\lambda_1$. Notice that $\varphi(\theta) = \log(Ee^{S_1\theta}) = \log(Ee^{a_1\theta}) + \log(Ee^{b_1(-\theta)}) = 0$. Thus, we find $P_\theta\{S_n = l\} = P_0\{S_n = l\} \cdot e^{\theta l}$.

$$\begin{aligned}
P_0\{x_i = l\} &= \sum_{k=0}^{\infty} P_{\lambda_0}\{a_i = k+l\} P_{\lambda_1}\{b_i = k\} \\
&= \sum_{k=0}^{\infty} P_{\lambda_1}\{a_i = k+l\} e^{-\theta_0(k+l)+w^{r-1}\Delta(\lambda_1-\lambda_0)} \\
&\quad \cdot P_{\lambda_0}\{b_i = k\} e^{\theta_0 k + w^{r-1}\Delta(\lambda_0-\lambda_1)} \\
&= \sum_{k=0}^{\infty} P_{\lambda_0}\{a_i = k\} P_{\lambda_1}\{b_i = k+l\} e^{-\theta_0 l} \\
&= P_0\{x_i = -l\} e^{-\theta_0 l} && (3.7)\\
&= P_\theta\{x_i = -l\}. && (3.8)
\end{aligned}$$

From (3.7) to (3.8), we use (3.6) with $n = 1$. It can be observed that, by changing the measure to P_θ, a_i and b_i exchange their parameters.

Lemma 4. *Let $\tau_+ = \inf\{n > 0 \mid S_n > 0\}$, $\tau_- = \inf\{n > 0 \mid S_n < 0\}$, and $\mu = w^{r-1}\Delta(\lambda_1 - \lambda_0) = b(\Delta/w)(1 - e^{-\theta})$. In the notation of Lemma 3,*

$$\sum_{k=0}^{\infty} e^{-\theta k} P\{\max_{j>0} S_j < -k\} = \mu \frac{(1 - E_\theta e^{-\theta S_{\tau_+}})}{(1 - e^{-\theta}) E_\theta S_{\tau_+}} = \mu\nu(\theta, \alpha). \tag{3.9}$$

Remark of Lemma 4. Lemma 4 uses the result of Lemma 3, $P\{\max_{j>0} S_j < -k\} = P_\theta\{\min_{j>0} S_j > k\}$, and applies the Siegmund (1985) reuslt that $P\{\min_{j>0} S_j > k\} = \mu P(S_{\tau_+} > k)/ES_{\tau_+}$.

Lemma 5. *With the notation of Lemma* 3*, let* η *be a positive integer. Then*

$$\sum_{k=0}^{\eta} e^{\theta(\eta-k)} P\{\max_{i>0} S_i < -k\}\, P\{\max_{i\geq 0} S_i = \eta - k\} = \mu(1-e^{-\theta})\nu(\theta,\alpha). \qquad (3.10)$$

Proof of Lemma 5. By Lemma 3, we have $P\{\max_{i>0} S_i < -k\} = P_\theta\{\min_{i>0} S_i > k\} = \mu P_\theta(S_{\tau_+} > k)/E_\theta S_{\tau_+}$. Let $Z(k) = e^{\theta k} P\{\max_{i\geq 0} S_i = k\}$, $F(k) = P_\theta(S_{\tau_+} > k)$, and $f(k) = P_\theta(S_{\tau_+} = k)$. Then, Lemma 5 is equivalent to $F * Z(\eta) \equiv \sum_{k=0}^{\eta} F(k) Z(\eta - k) = 1 - E_\theta e^{-S_{\tau_+}} = P(\tau_+ = \infty)$ for any non-negative integer η. Notice that $Z(0) = P\{\max_{i\geq 0} S_i = 0\} = P(\tau_+ = \infty)$. When $k \geq 1$, we have

$$\begin{aligned} Z(k) &= e^{\theta k} \sum_{k'=1}^{k} P(S_{\tau_+} = k') P\{\max_{i\geq 0} S_i = k - k'\} \\ &= \sum_{k'=1}^{k} e^{\theta k'} P(S_{\tau_+} = k') Z(k-k') \\ &= \sum_{k'=1}^{k} f(k') Z(k-k') \\ &= \sum_{k'=1}^{k} \big(F(k'-1) - F(k')\big) Z(k-k') \\ &= F * Z(k-1) - F * Z(k) + F(0) Z(k). \end{aligned} \qquad (3.11)$$

From (3.11), $F * Z(k-1) = F * Z(k)$ for any positive integer k, because $F(0) = 1$. Thus, $F * Z(\eta) = F * Z(0) = P(\tau_+ = \infty)$, and the proof is done.

Now, we continue with (3.1). For $r = 1$, we get the last time conditional set as $\{\max_{j>0} S_j < -k\}$ by Lemma 2. Thus the approximation at (2.1) can be derived by using Lemma 1 and Lemma 4. (3.1) is approximated by

$$\begin{aligned} &\frac{L}{\Delta} \frac{\exp[-b\theta + w\cdot(\lambda_1 - \lambda_0)]}{\sqrt{2\pi b}} \sum_{k=0}^{\infty} e^{-\theta k} P_\theta\{\max_{j>0} S_j < -k\} \\ &\quad = (1 - \frac{\lambda_0}{\lambda_1}) \cdot b \frac{\exp[-b\theta + w\cdot(\lambda_1 - \lambda_0)]}{\sqrt{2\pi b}} \left(\frac{L}{w}\nu(\theta,\alpha)\right). \end{aligned} \qquad (3.12)$$

Note that (3.12) gives (2.1) when $r = 1$.

For $r \geq 2$, we rewrite the conditional set first. We condition on $\{\max_{i_1\geq 0} S_{i_1}^{1,-} = \eta - k\}$, then take this set and the set $\{\max_{i_1>0} S_{i_1}^{1,+} < -k\}$ out by independencies. The random variable $\max_{i_1\geq 0} S_{i_1}^{1,-}$ must be non-negative because the index includes 0. We sum over η from 0 to ∞ and thus force k to be summed from 0 to η. Let $h(r)=\sum_{k=0}^{\infty} e^{-\theta k}P(B_r(k))$, where $B_r(k)$ is defined in Lemma 2, then (3.1) can be approximated as $(L/w)^r A_r h(r)$ by Lemma 2, where A_r is defined in Lemma 1.

$$
\begin{aligned}
h(r) &= \sum_{k=0}^{\infty} e^{-\theta k} P\{\max_{i_1>0} S_{i_1}^{1,+} < -k; \max_{i_1\ge 0; i_2>0} S_{i_1}^{1,-} + S_{i_2}^{2,+} < -k; \ldots; \\
&\qquad \max_{\substack{i_1,\ldots,i_{r-1}\ge 0\\ i_r>0}} S_{i_1}^{1,-} + \cdots + S_{i_{r-1}}^{r-1,-} + S_{i_r}^{r,+} < -k\} \\
&= \sum_{\eta=0}^{\infty} e^{-\theta\eta} P\{\max_{i_2>0} S_{i_2}^{2,+} < -\eta; \ldots; \max_{\substack{i_2,\ldots,i_{r-1}\ge 0\\ i_r>0}} S_{i_2}^{2,-} + \ldots + S_{i_{r-1}}^{r-1,-} + S_{i_r}^{r,+} < -\eta\} \\
&\qquad \sum_{k=0}^{\eta} e^{\theta(\eta-k)} P\{\max_{i_1>0} S_{i_1}^{1,+} < -k\} P\{\max_{i_1\ge 0} S_{i_1}^{1,-} = \eta - k\} \qquad (3.13) \\
&= h(r-1)\cdot \mu(1-e^{-\theta})\nu(\theta,\alpha). \qquad (3.14)
\end{aligned}
$$

From (3.13) and (3.14), we use Lemma 5 and observe that the first line in (3.13) is $h(r-1)$. By Lemma 1 and Lemma 4, we know that $h(1)=\mu\nu(\theta,\alpha)$, also shown in deriving (3.12). By (3.14), we get $h(r)=(1-e^{-\theta})^{r-1}\mu^r\nu^r(\theta,\alpha)= b^r(1-e^{-\theta})^{2r-1}(w/\Delta)^r\nu^r(\theta,\alpha)$. Thus, (2.1) is proved for any $r\ge 2$.

4. Simulation Examples

Applying Theorem 1, the thresholds for the Soudan-2 detector are provided for various background intensities λ_0 (see Table 1). A simulation experiment was set to check the accuracies of the approximation (2.1) with $r=2$ in Table 2. The approximations are compared with Normal approximation, an importance sampling simulation with 1,000 repetitions, and the direct Monte Carlo simulation with 10,000 repetitions. To highlight the impact of the ratchet factor, an approximation ignoring the ratchet effect is also presented.

Starting from the direct Monte Carlo procedure, we simulated Poissons on each unit square $\Delta\times\Delta$, where Δ is the shifting size. With $\{Q_{ij}^{(k)} \mid 1\le i\le L_1, 1\le j\le L_2\}$, $1\le k\le B$, a B independent realization of the Poisson random field with rate λ_0. $X^{(k)}(i,j)=\sum_{m=i}^{i+w_1-1}\sum_{n=j}^{j+w_2-1} Q_{ij}^{(k)}$, $\hat P_D = B^{-1}\sum_{k=1}^{B} I_{[\max_{i,j} X^{(k)}(i,j)\ge b]}$, and $\hat{s.e.}_D=[\hat P_D(1-\hat P_D)/B]^{1/2}$ are the unbiased direct Monte Carlo estimate of $p := P(\max_{i,j} X^{(k)}(i,j)\ge b)$ and its consistent standard error estimate.

Following Chan and Zhang (2007), we used a change of measure importance sampling procedure for two-dimensional ratchet scan statistics that can be modified easily to higher dimensional cases. We took $(u^{(k)}, v^{(k)})$ randomly generated from $([1, L_1] \cap \mathbb{Z}) \times ([1, L_2] \cap \mathbb{Z})$, and generated $X^{(k)}(i, j)$ as an alternative measure P^* with $Q_{ij}^{(k)}$ generated as independent Poissons with rate λ_1 for $(i, j) \in [u^{(k)}, u^{(k)} + w_1 - 1] \times [v^{(k)}, v^{(k)} + w_2 - 1]$, and with rate λ_0 elsewhere. Under this setting,

$$L_k = \frac{dP^{*(k)}}{dP_0} = \sum_{1 \le i \le L_1, 1 \le j \le L_2} e^{(\lambda_0 - \lambda_1) w_1 w_2} \frac{(\lambda_1/\lambda_0)^{X^{(k)}(i,j)}}{L_1 L_2}.$$

Then

$$\hat{P}_I = B^{-1} \sum_{k=1}^{B} L_k^{-1} I_{[\max_{i,j} X^{(k)}(i,j) \ge b]}$$

and

$$s.\hat{e}._I = B^{-1} (\sum_{k=1}^{B} L_k^{-2} I_{[\max_{i,j} X^{(k)}(i,j) \ge b]} - B \hat{P}_I^2)^{1/2}$$

are unbiased estimate of p and its standard error estimate.

We give an example on deriving the threshold in detecting the muon source. We use the window shifting size Δ as the unit length. The window size is 29×56, and the whole sky space is 1160×4480. Table 1 lists the scaled thresholds for .05 significance level for both $M_{\mathbf{L},\mathbf{\Delta},\mathbf{w}}$ and $M_{\mathbf{L},\mathbf{w}}$, applying Theorem 1 and Theorem 2. We also show the scaled threshold calculated by Normal approximation for both with and without ratchet effect. The results show that the tests are too conservative if we ignore the ratchet effect. The threshold values in Table 1 have been scaled as the numbers of standard deviation. We also show the p-value for fixed b in Table 2, and we increase the ratio between window shifting size and window size to magnify the ratchet effect. The window size is 9×9 and the whole sky space is 180×180. We also present the normal approximations in the tables as a comparison. The important resampling method and the direct Monte Carlo simulation are both quite consistent with Theorem 1.

It is a reasonable practice to approximate a Poisson random variable by a Gaussian when the mean of $X(\mathbf{t})$ is greater than five. However, in this case, we see that even when the mean is above 30, discrepancies in threshold approximations still exist between these two, as skewness plays an important role in large deviation approximations (Tu and Siegmund (1999); Tang and Siegmund (2001)). For a scan statistic, the threshold is usually four or five standard deviations above the mean, which is really a large deviation problem.

5. Discussion

Ratchet scan statistics are of practical interest. When ignoring the ratchet effect, the critical value for a hypothesis test tends to be too conservative and

Table 1. 2-dim Poisson Random Field.
Window Size = 29×56, Whole Space = 1160×4480
Thresholds for Type I Error = 0.05

		scaled b (SD)	
λ_0	$E_0 X_{\mathbf{w}}(\mathbf{t})$	$M_{\mathbf{L},\mathbf{\Delta},\mathbf{w}}$	$M_{\mathbf{L},\mathbf{w}}$
0.002	3.248	7.34	7.46
0.02	32.48	5.99	6.18
0.2	324.8	5.48	5.73
2.	3248	5.32	5.58
Norm.	App.	5.23	5.52

Table 2. 2-dim Poisson Random Field.
Window Size: $w_1 = 9$, $w_2 = 9$, Whole Space: $L_1 = 180$, $L_2 = 180$
$P(M_{\mathbf{L},\mathbf{\Delta},\mathbf{w}} \geq b)$.

λ_0	b	Normal Approx	Thm. 1 with ratchet effect	Importance Sampling 1,000 repetitions	Direct Monte Carlo 10,000 repetitions	Thm. 2 ignore ratchet effect
0.1	24	0.0002	0.0559	0.0535	0.0517	0.2093
	25	0.00002	0.0194	0.0200	0.0174	0.0781
0.3	49	0.0045	0.0893	0.0833	0.0810	0.4446
	50	0.0016	0.0453	0.0454	0.0424	0.2385
0.5	72	0.0060	0.0690	0.0736	0.0616	0.3990
	73	0.0028	0.0395	0.0394	0.0341	0.2398
1.0	125	0.0080	0.0506	0.0491	0.0475	0.3498
	126	0.0047	0.0307	0.0320	0.0300	0.3320

the test loses its power. We develop an approximation method to factor out the ratchet effect on the tail probabilities as an overshoot function ν in each dimension for any r-dimensional Poisson process. Although we use $w_1 = \cdots = w_r$, $\Delta_1 = \cdots = \Delta_r$, and $L_1 = \cdots = L_r$ for notation simplicity, it should be very easy to generalize to any set of parameters $\{(L_i, w_i, \Delta_i), 1 \leq i \leq r\}$ given a reasonable region. This method can also be extended to do the tail probability approximation for a ratchet scan statistic over a compound Poisson process.

Acknowledgement

This paper is orignated from my Ph.D. work at Stanford University. I would like to thank my thesis advisor, Professor David Siegmund, for his guidance and support. I am very grateful to a referee of Advances in Applied Probability for his critical suggestions.

Appendix

A computing formula for ν (Tu (2009)) Let x_1, x_2, $\ldots$ be arithmetic i.i.d.

random variables with span h. If $\mu = Ex_1 > 0$, $Ex_1^2 < \infty$, $\tau_+ = \inf\{n : S_n > 0\}$, $\phi(t) = E\exp(itx_1)$, and $\xi(t) = \sum_{n=1}^{\infty} \phi^n(t)/n = -\log(1-\phi(t))$, then

$$\begin{aligned}\nu(\theta) &= h\frac{1 - E\exp(-\theta S_{\tau_+})}{(1-\exp(-\theta h))ES_{\tau_+}} \\ &= \exp\Big\{\frac{-h}{2\pi}\int_{-\pi/h}^{\pi/h} dt\Big[\Big(\xi(t) + \log(\frac{\mu(1-e^{iht})}{h})\Big)(\frac{e^{-\theta h - iht}}{1-e^{-\theta h - iht}} + \frac{1}{1-e^{iht}})\Big]\Big\}.\end{aligned}$$

References

Aldous, D. (1989). *Probability Approximations via the Poisson Clumping Heuristic*. Springer-Verlag, New York.

Alm, S. E. (1997). On the distribution of scan statistics of a two dimensional Poisson process. *Adv. Appl. Probab.* **29**, 1-18.

Chan, H. P. (2009). Maxima of moving sums in a Poisson random field. *Adv. Appl. Probab.* **41**, 647-663.

Chan, H. and Zhang, N. (2007). Scan statistics with weighted observations. *J. Amer. Statist. Assoc.* **102**, 595-602.

Glaz, J., Naus, J. and Wallenstein, S. (2001). *Scan Statistics*. Springer-Verlag, New York.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics - Theory and Methods* **26**, 1481-1496.

Loader, C. R. (1991). Large-deviation approximations to the distribution of scan statistics. *Adv. Appl. Probab.* **23**, 751-771.

Naus, J. (1965). Clustering of random points in two dimensions. *Biometrika* **52**, 263-267.

Rabinowitz, D. and Siegmund, D. (1997). The appoximate distribution of the maximum of a smoothed Poisson random field. *Statist. Sinica* **7**, 167-180.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*. Springer-Verlag, New York.

Siegmund, D. (1988). Approximate tail probabilities for the maximum of some random fields. *Ann. Probab.* **16**, 487-501.

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

Tang, H. K. and Siegmund, D. (2001). Mapping quantitative trait loci in oligogenic models. *Biostatistics* **2**, 147-162.

Tu, I. (2009). Asymptotic overshoot for arithmetic iid random variables. *Statist. Sinica* **19**, 315-323.

Tu, I. and Siegmund, D. (1999). The maximum of a function of a Markov Chain and application to linkage analysis. *Adv. Appl. Probab.* **31**, 510-531.

Woodroofe, M. (1976). Frequentist properties of Bayesian sequential tests. *Biometrika* **63**, 101-110.

Woodroofe, M. (1979). Repeated likelihood ratio tests. *Biometrika* **66**, 454-463.

Institute of Statistical Science, Academia Sinica, 128, Academia Rd. Sec. 2, Taipei 115, Taiwan, R.O.C.

E-mail: iping@stat.sinica.edu.tw

(Received February 2012; accepted September 2012)