Extension of the classical scan statistics framework with applications

Alexandru Amărioarei

Faculty of Mathematics and Computer Science
University of Bucharest, Romania

The 6^{th} IMS-China International Conference on Statistics and Probability

June 28 - July 2, 2017
Nanning, Guangxi, China

Outline

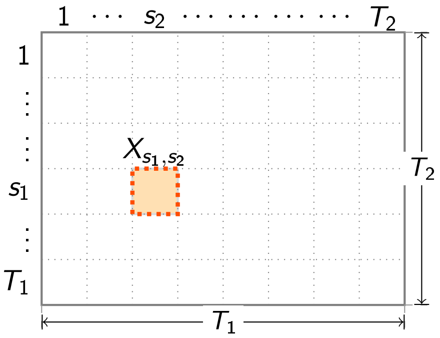
1. Introduction
   - Framework
   - Problem
2. Methodology
   - Approximation
3. Applications
   - Application: Length of the Longest increasing/non-decreasing run
   - Application: Scanning with windows of arbitrary shape
   - Application: Scanning the surface of a cylinder

Outline

Definitions and notations

Preliminary notations

Let T_1, T_2 be positive integers



- Rectangular region $\mathcal{R}_2 = [0, T_1] \times [0, T_2]$
- $(X_{s_1,s_2})_{\substack{1\leq s_1\leq T_1\\ 1\leq s_2\leq T_2}}$ i.i.d. r.v.'s
  - Bernoulli($\mathcal{B}(1, p)$)
  - Binomial($\mathcal{B}(n, p)$)
  - Poisson($\mathcal{P}(\lambda)$)
  - Normal($\mathcal{N}(\mu, \sigma^2)$)
- X_{s_1,s_2} number of observed events in the elementary subregion $r_{s_1,s_2} = [s_1 - 1, s_1] \times [s_2 - 1, s_2]$

Two dimensional scan statistic

Let $2 \leq m_s \leq T_s$, $s \in \{1, 2\}$ be positive integers

- Define for $1 \leq i_s \leq T_s - m_s + 1$ and $1 \leq j_s \leq m_s$ the 2-way tensor $\mathcal{X}_{i_1,i_2} \in \mathbb{R}^{m_1 \times m_2}$,

$$\mathcal{X}_{i_1,i_2}(j_1, j_2) = X_{i_1+j_1-1, i_2+j_2-1}$$

- Take $\mathcal{S} : \mathbb{R}^{m_1 \times m_2} \to \mathbb{R}$ to be a measurable real valued function (*score function*) and define

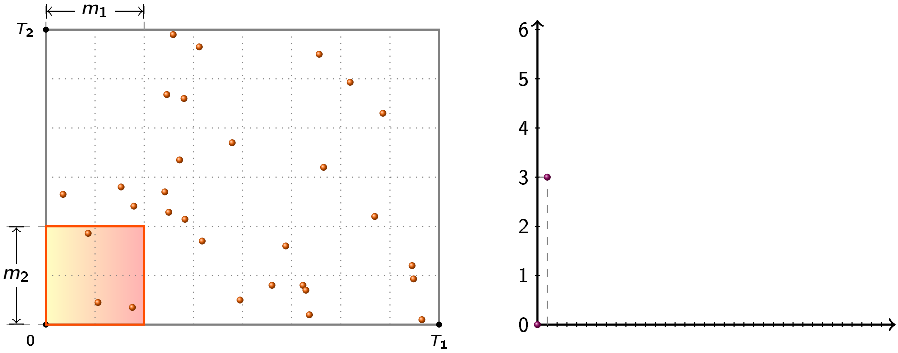
$$Y_{i_1,i_2}(\mathcal{S}) = \mathcal{S}\left(\mathcal{X}_{i_1,i_2}\right)$$

Definition

The two dimensional scan statistic with score function $\mathcal{S}$ is defined by

$$S_{m_1,m_2}(T_1, T_2; \mathcal{S}) = \max_{\substack{1 \leq i_1 \leq T_1 - m_1 + 1 \\ 1 \leq i_2 \leq T_2 - m_2 + 1}} Y_{i_1,i_2}(\mathcal{S})$$

Animation for 2 dimensional scan statistics



Number of points: 3

Outline

Problem and related work

Objective

Find a good estimate for the distribution of the two dimensional discrete scan statistic with score function $\mathcal{S}$

$$Q_{\mathbf{m}}(\mathbf{T};\mathcal{S}) = \mathbb{P}\left(S_{\mathbf{m}}(\mathbf{T};\mathcal{S}) \leq \tau\right)$$

with $\mathbf{m} = (m_1, m_2)$ and $\mathbf{T} = (T_1, T_2)$

Previous work:

- Continuous scan statistics
  - Rectangles: [Loader, 1991], [Glaz et al., 2001], [Glaz et al., 2009]
  - Circles: [Anderson and Titterington, 1997]
  - Triangles, ellipses and other convex shapes: [Alm, 1983, Alm, 1997, Alm, 1998], [Tango and Takahashi, 2005], [Assunção et al., 2006]
- Discrete scan statistics
  - **No results !**

Outline

Approximation methodology for the scan statistic

Approximation and error bounds

Theorem (Generalization of [Amărioarei, 2014])

Let $t_1, t_2 \in \{2,3\}$, $Q_{t_1,t_2} = \mathbb{P}\left(S_{\mathbf{m}}\left(t_1(m_1-1), t_2(m_2-1); \mathcal{S}\right) \leq \tau\right)$ and $L_s = \left\lfloor \frac{T_s}{m_s - 1} \right\rfloor$, $s \in \{1,2\}$. If $\hat{Q}_{t_1,t_2}$ is an estimate of Q_{t_1,t_2}, $\left|\hat{Q}_{t_1,t_2} - Q_{t_1,t_2}\right| \leq \beta_{t_1,t_2}$ and τ is such that $1 - \hat{Q}_{2,2}(\tau) \leq 0.1$ then

$$\left|\mathbb{P}\left(S_{\mathbf{m}}(\mathbf{T};\mathcal{S}) \leq \tau\right) - \left(2\hat{Q}_2 - \hat{Q}_3\right)\left[1 + \hat{Q}_2 - \hat{Q}_3 + 2(\hat{Q}_2 - \hat{Q}_3)^2\right]^{1-L_1}\right| \leq E_{sf} + E_{sapp},$$

where, for $t \in \{2,3\}$

$$\hat{Q}_t = \left(2\hat{Q}_{t,2} - \hat{Q}_{t,3}\right)\left[1 + \hat{Q}_{t,2} - \hat{Q}_{t,3} + 2(\hat{Q}_{t,2} - \hat{Q}_{t,3})^2\right]^{1-L_2}$$

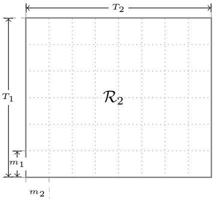
$$E_{sf} = (L_1 - 1)(L_2 - 1)\left(\beta_{2,2} + \beta_{2,3} + \beta_{3,2} + \beta_{3,3}\right)$$

$$E_{sapp} = (L_1 - 1)\left[F_1\left(1 - \hat{Q}_2 + A_2 + C_2\right)^2 + (L_2 - 1)(F_2 C_2 + F_3 C_3)\right]$$

$$A_2 = (L_2 - 1)\left(\beta_{2,2} + \beta_{2,3}\right)$$

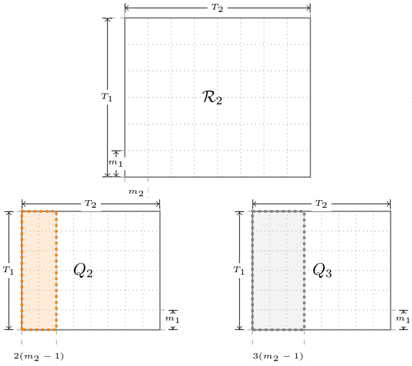
$$C_t = (L_2 - 1)F_t\left(1 - \hat{Q}_{t,2} + \beta_{t,2}\right)^2.$$

Illustration of the approximation process



Find Approximation

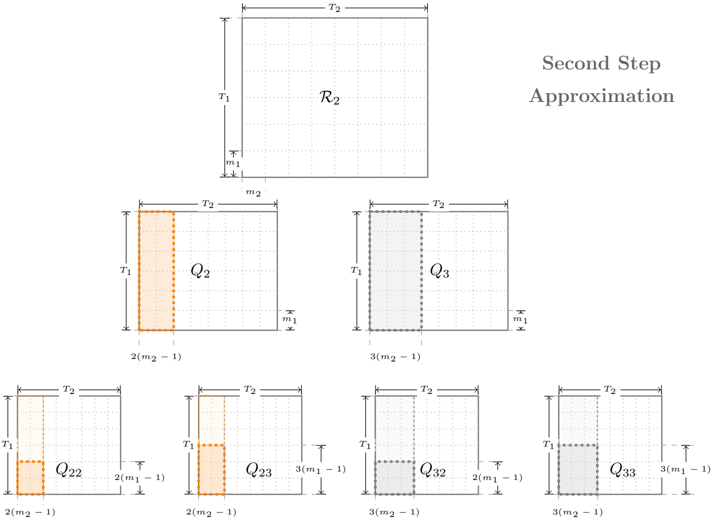
Illustration of the approximation process



First Step

Approximation

Illustration of the approximation process



Second Step

Approximation

Outline

Application 1:

Longest increasing/non-decreasing run

Longest increasing/non-decreasing run

Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. r.v.'s with the common distribution G.

Increasing run

A subsequence $(X_k, \dots, X_{k+l-1})$ forms an *increasing run* of length $l \geq 1$, starting at position $k \geq 1$, if

$$X_{k-1} > X_k < X_{k+1} < \cdots < X_{k+l-1} > X_{k+l}$$

Non-decreasing run

A subsequence $(X_k, \dots, X_{k+l-1})$ forms an *non-decreasing run* of length $l \geq 1$, starting at position $k \geq 1$, if

$$X_{k-1} > X_k \leq X_{k+1} \leq \cdots \leq X_{k+l-1} > X_{k+l}$$

Longest increasing/non-decreasing run

Notations

- $M^{I}_{T_1}$ = the length of the longest increasing run among the first T_1 r.v.'s

$$M^{I}_{T_1} = \max\{l \,|\, X_k < \cdots < X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

- $M^{ND}_{T_1}$ = the length of the longest non-decreasing run among the first T_1 r.v.'s

$$M^{ND}_{T_1} = \max\{l \,|\, X_k \leq \cdots \leq X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

Example $(T_1 = 10)$

X_i :	1	3	5	2	4	7	1	3	3	8
IR :	1	3	5	2	4	7	1	3	3	8
NDR :	1	3	5	2	4	7	1	3	3	8

Longest increasing/non-decreasing run

Notations

- $M^{I}_{T_1}$ = the length of the longest increasing run among the first T_1 r.v.'s

$$M^{I}_{T_1} = \max\{l \,|\, X_k < \cdots < X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

- $M^{ND}_{T_1}$ = the length of the longest non-decreasing run among the first T_1 r.v.'s

$$M^{ND}_{T_1} = \max\{l \,|\, X_k \leq \cdots \leq X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

Example ($T_1 = 10$)

X_i : 1 3 5 2 4 7 1 3 3 8

IR : <u>1 3 5</u> 2 4 7 1 3 3 8

NDR : 1 3 5 2 4 7 1 3 3 8

Longest increasing/non-decreasing run

Notations

- $M^{I}_{T_1}$ = the length of the longest increasing run among the first T_1 r.v.'s

$$M^{I}_{T_1} = \max\{l \,|\, X_k < \cdots < X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

- $M^{ND}_{T_1}$ = the length of the longest non-decreasing run among the first T_1 r.v.'s

$$M^{ND}_{T_1} = \max\{l \,|\, X_k \leq \cdots \leq X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

Example ($T_1 = 10$)

X_i : 1 3 5 2 4 7 1 3 3 8

IR : $\underline{1\ 3\ 5}$ $\underline{2\ 4\ 7}$ 1 3 3 8

NDR : 1 3 5 2 4 7 1 3 3 8

Longest increasing/non-decreasing run

Notations

- $M^{I}_{T_1}$ = the length of the longest increasing run among the first T_1 r.v.'s

$$M^{I}_{T_1} = \max\{l \mid X_k < \cdots < X_{k+l-1} \text{ for some } k,\ 1 \le k \le T_1 - l + 1\}$$

- $M^{ND}_{T_1}$ = the length of the longest non-decreasing run among the first T_1 r.v.'s

$$M^{ND}_{T_1} = \max\{l \mid X_k \le \cdots \le X_{k+l-1} \text{ for some } k,\ 1 \le k \le T_1 - l + 1\}$$

Example ($T_1 = 10$)

X_i : 1 3 5 2 4 7 1 3 3 8

IR : <u>1 3 5</u> <u>2 4 7</u> <u>1 3</u> 3 8

NDR : 1 3 5 2 4 7 1 3 3 8

Longest increasing/non-decreasing run

Notations

- $M^{I}_{T_1}$ = the length of the longest increasing run among the first T_1 r.v.'s

$$M^{I}_{T_1} = \max\{l \,|\, X_k < \cdots < X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

- $M^{ND}_{T_1}$ = the length of the longest non-decreasing run among the first T_1 r.v.'s

$$M^{ND}_{T_1} = \max\{l \,|\, X_k \leq \cdots \leq X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

Example ($T_1 = 10$)

X_i : 1 3 5 2 4 7 1 3 3 8

IR : 1 3 5 | 2 4 7 | 1 3 | 3 8

NDR : 1 3 5 2 4 7 1 3 3 8

Longest increasing/non-decreasing run

Notations

- $M^{I}_{T_1}$ = the length of the longest increasing run among the first T_1 r.v.'s

$$M^{I}_{T_1} = \max\{l \,|\, X_k < \cdots < X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

- $M^{ND}_{T_1}$ = the length of the longest non-decreasing run among the first T_1 r.v.'s

$$M^{ND}_{T_1} = \max\{l \,|\, X_k \leq \cdots \leq X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

Example ($T_1 = 10$)

X_i : 1 3 5 2 4 7 1 3 3 8

IR : 1 3 5 | 2 4 7 | 1 3 | 3 8 $\quad M^{I}_{10} = 3$

NDR : 1 3 5 2 4 7 1 3 3 8

Longest increasing/non-decreasing run

Notations

- $M^{I}_{T_1}$ = the length of the longest increasing run among the first T_1 r.v.'s

$$M^{I}_{T_1} = \max\{l \,|\, X_k < \cdots < X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

- $M^{ND}_{T_1}$ = the length of the longest non-decreasing run among the first T_1 r.v.'s

$$M^{ND}_{T_1} = \max\{l \,|\, X_k \leq \cdots \leq X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

Example ($T_1 = 10$)

X_i : 1 3 5 2 4 7 1 3 3 8

IR : 1 3 5 | 2 4 7 | 1 3 | 3 8 $\qquad M^{I}_{10} = 3$

NDR : 1 3 5 2 4 7 1 3 3 8

Longest increasing/non-decreasing run

Notations

- $M^{I}_{T_1}$ = the length of the longest increasing run among the first T_1 r.v.'s

$$M^{I}_{T_1} = \max\{l \mid X_k < \cdots < X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

- $M^{ND}_{T_1}$ = the length of the longest non-decreasing run among the first T_1 r.v.'s

$$M^{ND}_{T_1} = \max\{l \mid X_k \leq \cdots \leq X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

Example $(T_1 = 10)$

X_i : 1 3 5 2 4 7 1 3 3 8

IR : <u>1 3 5</u> <u>2 4 7</u> <u>1 3</u> <u>3 8</u> $\quad M^{I}_{10} = 3$

NDR : <u>1 3 5</u> <u>2 4 7</u> 1 3 3 8

Longest increasing/non-decreasing run

Notations

- $M^{I}_{T_1}$ = the length of the longest increasing run among the first T_1 r.v.'s

$$M^{I}_{T_1} = \max\{l \mid X_k < \cdots < X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

- $M^{ND}_{T_1}$ = the length of the longest non-decreasing run among the first T_1 r.v.'s

$$M^{ND}_{T_1} = \max\{l \mid X_k \leq \cdots \leq X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

Example ($T_1 = 10$)

X_i : 1 3 5 2 4 7 1 3 3 8

IR : 1 3 5 | 2 4 7 | 1 3 | 3 8 — $M^{I}_{10} = 3$

NDR : 1 3 5 | 2 4 7 | 1 3 3 8

Longest increasing/non-decreasing run

Notations

- $M_{T_1}^{I}$ = the length of the longest increasing run among the first T_1 r.v.'s

$$M_{T_1}^{I} = \max\{l \,|\, X_k < \cdots < X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

- $M_{T_1}^{ND}$ = the length of the longest non-decreasing run among the first T_1 r.v.'s

$$M_{T_1}^{ND} = \max\{l \,|\, X_k \leq \cdots \leq X_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

Example ($T_1 = 10$)

X_i : 1 3 5 2 4 7 1 3 3 8

IR : <u>1 3 5</u> <u>2 4 7</u> <u>1 3</u> <u>3 8</u> $\qquad M_{10}^{I} = 3$

NDR : <u>1 3 5</u> <u>2 4 7</u> <u>1 3 3 8</u> $\qquad M_{10}^{ND} = 4$

Problem

Goal

Find a good estimate for the distribution of the longest *increasing* or *non-decreasing* run in the sequence $(X_n)_{n\geq 1}$ of i.i.d. r.v.'s

$$\mathbb{P}\left(M^{I}_{T_1} \leq k\right) \text{ and } \mathbb{P}\left(M^{ND}_{T_1} \leq k\right)$$

The asymptotic distribution was studied

- G continuous distribution: [Pittel, 1981], [Révész, 1983], [Grill, 1987], [Novak, 1992]

$$\mathbb{P}\left(M^{I}_{T_1} = M^{ND}_{T_1}\right) = 1$$

- G discrete distribution:
  - IR: geometric [Grabner et al., 2003], [Louchard and Prodinger, 2003]
  - NDR: geometric [Csaki and Foldes, 1996], [Eryilmaz, 2006]
  - NDR: Poisson [Csaki and Foldes, 1996]
  - NDR: uniform [Louchard, 2005]

Relation: IR / NDR - scan statistics

Let $1 \leq m_1 \leq T_1$ be positive integers and $X_1, \ldots, X_{T_1}$ a sequence of i.i.d. r.v.'s. Define $\mathcal{S}_1, \mathcal{S}_2 : \mathbb{R}^{m_1} \to \mathbb{R}$ by

$$\mathcal{S}_1(x_1, \ldots, x_{m_1}) = \sum_{i=1}^{m_1-1} \mathbf{1}_{\{x_i < x_{i+1}\}}, \quad \mathcal{S}_2(x_1, \ldots, x_{m_1}) = \sum_{i=1}^{m_1-1} \mathbf{1}_{\{x_i \leq x_{i+1}\}}$$

Example $(X_i \sim \mathcal{U}(0,1),\ \tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}},\ T_1 = 10)$

X_i : 0.79 0.31 0.52 0.16 0.60 0.26 0.65 0.68 0.74 0.45

$\tilde{X}_i$:

Relation: IR / NDR - scan statistics

Let $1 \leq m_1 \leq T_1$ be positive integers and $X_1, \ldots, X_{T_1}$ a sequence of i.i.d. r.v.'s. Define $\mathcal{S}_1, \mathcal{S}_2 : \mathbb{R}^{m_1} \to \mathbb{R}$ by

$$\mathcal{S}_1(x_1, \ldots, x_{m_1}) = \sum_{i=1}^{m_1 - 1} \mathbf{1}_{\{x_i < x_{i+1}\}}, \quad \mathcal{S}_2(x_1, \ldots, x_{m_1}) = \sum_{i=1}^{m_1 - 1} \mathbf{1}_{\{x_i \leq x_{i+1}\}}$$

Example $(X_i \sim \mathcal{U}(0,1),\ \tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}},\ T_1 = 10)$

X_i : 0.79 0.31 0.52 0.16 0.60 0.26 0.65 0.68 0.74 0.45

$\tilde{X}_i$: 0

Relation: IR / NDR - scan statistics

Let $1 \leq m_1 \leq T_1$ be positive integers and $X_1, \dots, X_{T_1}$ a sequence of i.i.d. r.v.'s. Define $\mathcal{S}_1, \mathcal{S}_2 : \mathbb{R}^{m_1} \to \mathbb{R}$ by

$$\mathcal{S}_1(x_1, \dots, x_{m_1}) = \sum_{i=1}^{m_1 - 1} \mathbf{1}_{\{x_i < x_{i+1}\}}, \quad \mathcal{S}_2(x_1, \dots, x_{m_1}) = \sum_{i=1}^{m_1 - 1} \mathbf{1}_{\{x_i \leq x_{i+1}\}}$$

Example $(X_i \sim \mathcal{U}(0,1),\ \tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}},\ T_1 = 10)$

X_i : 0.79 0.31 0.52 0.16 0.60 0.26 0.65 0.68 0.74 0.45

↘ ↙ ↘ ↙

$\tilde{X}_i$: 0 1

Relation: IR / NDR - scan statistics

Let $1 \leq m_1 \leq T_1$ be positive integers and $X_1, \ldots, X_{T_1}$ a sequence of i.i.d. r.v.'s. Define $\mathcal{S}_1, \mathcal{S}_2 : \mathbb{R}^{m_1} \to \mathbb{R}$ by

$$\mathcal{S}_1(x_1, \ldots, x_{m_1}) = \sum_{i=1}^{m_1-1} \mathbf{1}_{\{x_i < x_{i+1}\}}, \quad \mathcal{S}_2(x_1, \ldots, x_{m_1}) = \sum_{i=1}^{m_1-1} \mathbf{1}_{\{x_i \leq x_{i+1}\}}$$

Example $(X_i \sim \mathcal{U}(0,1),\ \tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}},\ T_1 = 10)$

X_i : 0.79 0.31 0.52 0.16 0.60 0.26 0.65 0.68 0.74 0.45

$\tilde{X}_i$: 0 1 0

Relation: IR / NDR - scan statistics

Let $1 \leq m_1 \leq T_1$ be positive integers and $X_1, \dots, X_{T_1}$ a sequence of i.i.d. r.v.'s. Define $\mathcal{S}_1, \mathcal{S}_2 : \mathbb{R}^{m_1} \to \mathbb{R}$ by

$$\mathcal{S}_1(x_1, \dots, x_{m_1}) = \sum_{i=1}^{m_1-1} \mathbf{1}_{\{x_i < x_{i+1}\}}, \quad \mathcal{S}_2(x_1, \dots, x_{m_1}) = \sum_{i=1}^{m_1-1} \mathbf{1}_{\{x_i \leq x_{i+1}\}}$$

Example ($X_i \sim \mathcal{U}(0,1)$, $\tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}}$, $T_1 = 10$)

$$\begin{array}{lcccccccccccccccccccc}
X_i : & 0.79 & & 0.31 & & 0.52 & & 0.16 & & 0.60 & & 0.26 & & 0.65 & & 0.68 & & 0.74 & & 0.45 \\
& & \searrow\ \swarrow & & \searrow\ \swarrow & & \searrow\ \swarrow & & \searrow\ \swarrow & & \searrow\ \swarrow & & \searrow\ \swarrow & & \searrow\ \swarrow & & \searrow\ \swarrow & & \searrow\ \swarrow & \\
\tilde{X}_i : & & 0 & & 1 & & 0 & & 1 & & 0 & & 1 & & 1 & & 1 & & 0 &
\end{array}$$

Relation: IR / NDR - scan statistics

Let $1 \leq m_1 \leq T_1$ be positive integers and $X_1, \ldots, X_{T_1}$ a sequence of i.i.d. r.v.'s. Define $\mathcal{S}_1, \mathcal{S}_2 : \mathbb{R}^{m_1} \to \mathbb{R}$ by

$$\mathcal{S}_1(x_1, \ldots, x_{m_1}) = \sum_{i=1}^{m_1 - 1} \mathbf{1}_{\{x_i < x_{i+1}\}}, \quad \mathcal{S}_2(x_1, \ldots, x_{m_1}) = \sum_{i=1}^{m_1 - 1} \mathbf{1}_{\{x_i \leq x_{i+1}\}}$$

Example $(X_i \sim \mathcal{U}(0,1),\ \tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}},\ T_1 = 10)$

X_i : 0.79 0.31 0.52 0.16 0.60 0.26 0.65 0.68 0.74 0.45

$\tilde{X}_i$: 0 1 0 1 0 1 1 1 0

Relation: IR / NDR - scan statistics

Let $1 \leq m_1 \leq T_1$ be positive integers and $X_1, \ldots, X_{T_1}$ a sequence of i.i.d. r.v.'s. Define $\mathcal{S}_1, \mathcal{S}_2 : \mathbb{R}^{m_1} \to \mathbb{R}$ by

$$\mathcal{S}_1(x_1, \ldots, x_{m_1}) = \sum_{i=1}^{m_1-1} \mathbf{1}_{\{x_i < x_{i+1}\}}, \quad \mathcal{S}_2(x_1, \ldots, x_{m_1}) = \sum_{i=1}^{m_1-1} \mathbf{1}_{\{x_i \leq x_{i+1}\}}$$

Example $(X_i \sim \mathcal{U}(0,1),\ \tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}},\ T_1 = 10)$

X_i :	0.79		0.31		0.52		0.16		0.60		0.26		0.65		0.68		0.74		0.45
$\tilde{X}_i$:		0		1		0		1		0		1		1		1		0	

We have, for $k \geq 1$

$$\mathbb{P}\left(M^{I}_{T_1} \leq k\right) = \mathbb{P}\left(L_{T_1-1} < k\right) = \mathbb{P}\left(\mathbf{S}(k+1, T_1, \mathcal{S}_1) < k\right),$$
$$\mathbb{P}\left(M^{ND}_{T_1} \leq k\right) = \mathbb{P}\left(L_{T_1-1} < k\right) = \mathbb{P}\left(\mathbf{S}(k+1, T_1, \mathcal{S}_2) < k\right).$$

Novak's result

Let $\left(\tilde{X}_n\right)_{n\geq 1}$ be a 1 - dependent stationary sequence of r.v.'s with $\tilde{X}_n \in \{0,1\}$,

$$s(k) = \mathbb{P}\left(\tilde{X}_1 = \cdots = \tilde{X}_k = 1\right),$$
$$r(k) = s(k+1) - s(k),$$

and let L_{T_1} be the length of the longest success run among the first T_1 trials

$$L_{T_1} = \max\{l \,|\, \tilde{X}_k = \cdots = \tilde{X}_{k+l-1} \text{ for some } k,\ 1 \leq k \leq T_1 - l + 1\}$$

Theorem ([Novak, 1992])

If there exists positive constants $t, C < \infty$ such that

$$\frac{s(k+1)}{s(k)} \geq \frac{1}{Ck^t} \quad \text{for all } k \geq C,$$

then, as $T_1 \to \infty$

$$\max_{1\leq k\leq T_1} \left|\mathbb{P}\left(L_{T_1} < k\right) - e^{T_1 r(k)}\right| = \mathcal{O}\left(\frac{(\log(T_1))^d}{T_1}\right)$$

where $d = \max\{t, 1\}$.

Longest increasing run: $G = \mathcal{U}([0,1])$

Let $X_1, \dots, X_{T_1}$ be a sequence of i.i.d. r.v.'s with the common distribution $G = \mathcal{U}([0,1])$ and $\tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}}$. In the view of [Novak, 1992] result we have

$$s(k) = \frac{1}{(k+1)!}, \quad r(k) = \frac{k+1}{(k+2)!}, \quad C = 2, \quad t = 1, \quad d = 1$$

and since $\mathbb{P}\left(M_{T_1}^I \le k\right) = \mathbb{P}\left(L_{T_1 - 1} < k\right) = \mathbb{P}\left(\mathbf{S}(k+1, T_1, \mathcal{S}_1) < k\right)$,

$$\max_{1 \le k \le T_1} \left| \mathbb{P}\left(M_{T_1}^I \le k\right) - e^{-(T_1 - 1)\frac{k+1}{(k+2)!}} \right| = \mathcal{O}\left(\frac{\ln T_1}{T_1}\right)$$

k	Sim	AppH	$E_{total}(1)$	LimApp
5	0.00000700	0.00000733	0.14860299	0.00000676
6	0.17567262	0.17937645	0.01089628	0.17620431
7	0.80257424	0.80362353	0.00110990	0.80215088
8	0.97548510	0.97566460	0.00011579	0.97550345
9	0.99749821	0.99751049	0.00001114	0.99749792
10	0.99977074	0.99977183	0.00000098	0.99977038
11	0.99998075	0.99998083	0.00000008	0.99998073
12	0.99999851	0.99999851	0.00000001	0.99999851
13	0.99999989	0.99999989	0.00000000	0.99999989
14	0.99999999	0.99999999	0.00000000	0.99999999
15	1.00000000	1.00000000	0.00000000	1.00000000

We used $T_1 = 10001$ and $Iter = 10^5$.

Longest increasing run: $G = \mathcal{U}([0,1])$

Let $X_1, \dots, X_{T_1}$ be a sequence of i.i.d. r.v.'s with the common distribution $G = \mathcal{U}([0,1])$ and $\tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}}$. In the view of [Novak, 1992] result we have

$$s(k) = \frac{1}{(k+1)!}, \quad r(k) = \frac{k+1}{(k+2)!}, \quad C = 2, \quad t = 1, \quad d = 1$$

and since $\mathbb{P}\left(M^{I}_{T_1} \leq k\right) = \mathbb{P}\left(L_{T_1 - 1} < k\right) = \mathbb{P}\left(\mathbf{S}(k+1, T_1, \mathcal{S}_1) < k\right)$,

$$\max_{1 \leq k \leq T_1} \left| \mathbb{P}\left(M^{I}_{T_1} \leq k\right) - e^{-(T_1 - 1)\frac{k+1}{(k+2)!}} \right| = \mathcal{O}\left(\frac{\ln T_1}{T_1}\right)$$

k	Sim	AppH	$E_{total}(1)$	LimApp
5	0.00000700	0.00000733	0.14860299	0.00000676
6	0.17567262	0.17937645	0.01089628	0.17620431
7	0.80257424	0.80362353	0.00110990	0.80215088
8	0.97548510	0.97566460	0.00011579	0.97550345
9	0.99749821	0.99751049	0.00001114	0.99749792
10	0.99977074	0.99977183	0.00000098	0.99977038
11	0.99998075	0.99998083	0.00000008	0.99998073
12	0.99999851	0.99999851	0.00000001	0.99999851
13	0.99999989	0.99999989	0.00000000	0.99999989
14	0.99999999	0.99999999	0.00000000	0.99999999
15	1.00000000	1.00000000	0.00000000	1.00000000

We used $T_1 = 10001$ and $Iter = 10^5$.

Longest non-decreasing run: $G = Geom(p)$

Let $X_1, \dots, X_{T_1}$ be a sequence of i.i.d. r.v.'s with the common distribution $G = Geom(p)$ and $\tilde{X}_i = \mathbf{1}_{\{X_i \le X_{i+1}\}}$. In the view of [Novak, 1992] result we have ([Eryilmaz, 2006])

$$s(k) = \frac{p^{k+1}}{\prod_{l=1}^{k+1}\left[1-(1-p)^l\right]}, \quad r(k) = \frac{(1-p)p^{k+1}}{\prod_{l=1}^{k}\left[1-(1-p)^l\right]\left[1-(1-p)^{k+2}\right]}, \quad C=2, \quad t=1, \quad d=1$$

and since $\mathbb{P}\left(M_{T_1}^{ND} \le k\right) = \mathbb{P}\left(L_{T_1-1} < k\right) = \mathbb{P}\left(\mathbf{S}(k+1, T_1, \mathcal{S}_2) < k\right)$,

$$\max_{1 \le k \le T_1}\left|\mathbb{P}\left(M_{T_1}^{ND} \le k\right) - e^{-(T_1-1)r(k)}\right| = \mathcal{O}\left(\frac{\ln T_1}{T_1}\right)$$

k	Sim	AppH	$E_{total}(1)$	LimApp
6	0.00910000	0.00881996	0.04299442	0.00955270
7	0.41785119	0.43020013	0.00530043	0.43655368
8	0.86812059	0.86944409	0.00077029	0.87208008
9	0.97847345	0.97856327	0.00011366	0.97901482
10	0.99681593	0.99681619	0.00001621	0.99689102
11	0.99955034	0.99955248	0.00000222	0.99956349
12	0.99993975	0.99993967	0.00000029	0.99994116
13	0.99999211	0.99999214	0.00000004	0.99999234
14	0.99999900	0.99999900	0.00000000	0.99999903
15	0.99999988	0.99999988	0.00000000	0.99999988

We used $T_1 = 10001$, $p = 0.1$ and $Iter = 10^5$.

Longest non-decreasing run: $G = Geom(p)$

Let $X_1, \dots, X_{T_1}$ be a sequence of i.i.d. r.v.'s with the common distribution $G = Geom(p)$ and $\tilde{X}_i = \mathbf{1}_{\{X_i \leq X_{i+1}\}}$. In the view of [Novak, 1992] result we have ([Eryilmaz, 2006])

$$s(k) = \frac{p^{k+1}}{\prod_{l=1}^{k+1}\left[1-(1-p)^l\right]}, \quad r(k) = \frac{(1-p)p^{k+1}}{\prod_{l=1}^{k}\left[1-(1-p)^l\right]\left[1-(1-p)^{k+2}\right]}, \quad C=2, \quad t=1, \quad d=1$$

and since $\mathbb{P}\left(M_{T_1}^{ND} \leq k\right) = \mathbb{P}\left(L_{T_1-1} < k\right) = \mathbb{P}\left(\mathbf{S}(k+1, T_1, \mathcal{S}_2) < k\right)$,

$$\max_{1 \leq k \leq T_1}\left|\mathbb{P}\left(M_{T_1}^{ND} \leq k\right) - e^{-(T_1-1)r(k)}\right| = \mathcal{O}\left(\frac{\ln T_1}{T_1}\right)$$

k	Sim	AppH	$E_{total}(1)$	LimApp
6	0.00910000	0.00881996	0.04299442	0.00955270
7	0.41785119	0.43020013	0.00530043	0.43655368
8	0.86812059	0.86944409	0.00077029	0.87208008
9	0.97847345	0.97856327	0.00011366	0.97901482
10	0.99681593	0.99681619	0.00001621	0.99689102
11	0.99955034	0.99955248	0.00000222	0.99956349
12	0.99993975	0.99993967	0.00000029	0.99994116
13	0.99999211	0.99999214	0.00000004	0.99999234
14	0.99999900	0.99999900	0.00000000	0.99999903
15	0.99999988	0.99999988	0.00000000	0.99999988

We used $T_1 = 10001$, $p = 0.1$ and $Iter = 10^5$.

Longest increasing run: $G = Geom(p)$

Let $X_1, \ldots, X_{T_1}$ be a sequence of i.i.d. r.v.'s with the common distribution $G = Geom(p)$ and $\tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}}$. The result of [Novak, 1992] cannot be applied since

$$s(k) = \frac{p^{k+1}}{\prod_{l=1}^{k+1}\left[1-(1-p)^l\right]}(1-p)^{\frac{(k+1)(k+2)}{2}}, \qquad \frac{s(k+1)}{s(k)} = \frac{p(1-p)^{k+1}}{1-(1-p)^{k+2}}.$$

For this case, [Louchard and Prodinger, 2003] showed that

$$\begin{aligned}
\mathbb{P}\left(M^{I}_{T_1} \leq k\right) &\sim \exp\left(-\exp \eta\right),\\
\eta &= \frac{k(k+1)}{2}\log\frac{1}{1-p} + k\log\frac{1}{p} - \log T_1 - \log p + \log D(k),\\
D(k) &= \prod_{l=1}^{k}\left[1-(1-p)^l\right]\left[1-(1-p)^{k+2}\right]
\end{aligned}$$

k	Sim	AppH	$E_{total}(1)$	LimApp
6	0.56445934	0.56997462	0.00255592	0.56810748
7	0.95295406	0.95325180	0.00018554	0.95294598
8	0.99658057	0.99659071	0.00001214	0.99657969
9	0.99979460	0.99979550	0.00000068	0.99979435
10	0.99998950	0.99998950	0.00000003	0.99998947

We used $T_1 = 10001$, $p = 0.1$ and $Iter = 10^5$.

Longest increasing run: $G = Geom(p)$

Let $X_1, \dots, X_{T_1}$ be a sequence of i.i.d. r.v.'s with the common distribution $G = Geom(p)$ and $\tilde{X}_i = \mathbf{1}_{\{X_i < X_{i+1}\}}$. The result of [Novak, 1992] cannot be applied since

$$s(k) = \frac{p^{k+1}}{\prod_{l=1}^{k+1}\left[1-(1-p)^l\right]}(1-p)^{\frac{(k+1)(k+2)}{2}}, \qquad \frac{s(k+1)}{s(k)} = \frac{p(1-p)^{k+1}}{1-(1-p)^{k+2}}.$$

For this case, [Louchard and Prodinger, 2003] showed that

$$\mathbb{P}\left(M^{I}_{T_1} \leq k\right) \sim \exp\left(-\exp \eta\right),$$
$$\eta = \frac{k(k+1)}{2}\log\frac{1}{1-p} + k\log\frac{1}{p} - \log T_1 - \log p + \log D(k),$$
$$D(k) = \prod_{l=1}^{k}\left[1-(1-p)^l\right]\left[1-(1-p)^{k+2}\right]$$

k	Sim	AppH	$E_{total}(1)$	LimApp
6	0.56445934	0.56997462	0.00255592	0.56810748
7	0.95295406	0.95325180	0.00018554	0.95294598
8	0.99658057	0.99659071	0.00001214	0.99657969
9	0.99979460	0.99979550	0.00000068	0.99979435
10	0.99998950	0.99998950	0.00000003	0.99998947

We used $T_1 = 10001$, $p = 0.1$ and $Iter = 10^5$.

Longest non-decreasing run: $G = \mathcal{U}(\{1,\ldots,s\})$

Let $X_1,\ldots,X_{T_1}$ be a sequence of i.i.d. r.v.'s with the common distribution $G = \mathcal{U}(\{1,\ldots,s\})$ and $\tilde{X}_i = \mathbf{1}_{\{X_i \leq X_{i+1}\}}$. By [Novak, 1992] result ([Louchard, 2005]) we have for $k \geq s$

$$s(k) = \binom{k+s}{s-1}\left(\frac{1}{s}\right)^{k+1}, \quad r(k) = (k+1)\binom{k+s}{s-2}\left(\frac{1}{s}\right)^{k+2}, \quad C = s, \quad t = 0, \quad d = 1$$

and since $\mathbb{P}\left(M_{T_1}^{ND} \leq k\right) = \mathbb{P}\left(L_{T_1-1} < k\right) = \mathbb{P}\left(\mathbf{S}(k+1, T_1, \mathcal{S}_2) < k\right)$,

$$\max_{1\leq k\leq T_1}\left|\mathbb{P}\left(M_{T_1}^{ND} \leq k\right) - e^{-(T_1-1)r(k)}\right| = \mathcal{O}\left(\frac{\ln T_1}{T_1}\right)$$

k	Sim	AppH	$E_{total}(1)$	LimApp
6	0.00011600	0.00009250	0.12199130	0.00012230
7	0.12501359	0.13542539	0.01560743	0.14301582
8	0.66274522	0.66691156	0.00260740	0.67447410
9	0.92424548	0.92504454	0.00046466	0.92720370
10	0.98565802	0.98582491	0.00008240	0.98623886
11	0.99748606	0.99747899	0.00001420	0.99756110
12	0.99956827	0.99957165	0.00000238	0.99958439
13	0.99992879	0.99992933	0.00000039	0.99993136
14	00.99998862	0.99998861	0.00000006	0.99998897

We used $T_1 = 10001$, $s = 10$ and $Iter = 10^5$.

Longest non-decreasing run: $G = \mathcal{U}(\{1,\dots,s\})$

Let $X_1,\dots,X_{T_1}$ be a sequence of i.i.d. r.v.'s with the common distribution $G = \mathcal{U}(\{1,\dots,s\})$ and $\tilde{X}_i = \mathbf{1}_{\{X_i \leq X_{i+1}\}}$. By [Novak, 1992] result ([Louchard, 2005]) we have for $k \geq s$

$$s(k) = \binom{k+s}{s-1}\left(\frac{1}{s}\right)^{k+1}, \quad r(k) = (k+1)\binom{k+s}{s-2}\left(\frac{1}{s}\right)^{k+2}, \quad C = s, \quad t = 0, \quad d = 1$$

and since $\mathbb{P}\left(M_{T_1}^{ND} \leq k\right) = \mathbb{P}\left(L_{T_1-1} < k\right) = \mathbb{P}\left(\mathbf{S}(k+1, T_1, \mathcal{S}_2) < k\right)$,

$$\max_{1\leq k\leq T_1}\left|\mathbb{P}\left(M_{T_1}^{ND} \leq k\right) - e^{-(T_1-1)r(k)}\right| = \mathcal{O}\left(\frac{\ln T_1}{T_1}\right)$$

k	Sim	AppH	$E_{total}(1)$	LimApp
6	0.00011600	0.00009250	0.12199130	0.00012230
7	0.12501359	0.13542539	0.01560743	0.14301582
8	0.66274522	0.66691156	0.00260740	0.67447410
9	0.92424548	0.92504454	0.00046466	0.92720370
10	0.98565802	0.98582491	0.00008240	0.98623886
11	0.99748606	0.99747899	0.00001420	0.99756110
12	0.99956827	0.99957165	0.00000238	0.99958439
13	0.99992879	0.99992933	0.00000039	0.99993136
14	00.99998862	0.99998861	0.00000006	0.99998897

We used $T_1 = 10001$, $s = 10$ and $Iter = 10^5$.

Outline

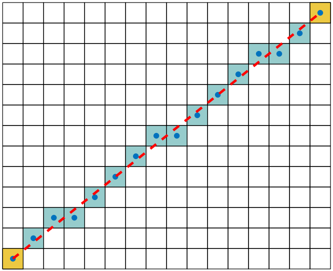
Application 3:

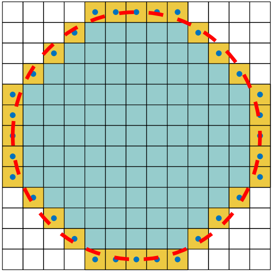
Scanning with windows of arbitrary shape

Shape of the scanning window

Let G be the geometrical shape of the scanning window (rectangle, quadrilateral, ellipse, etc.) and $\tilde{G}$ be its corresponding discrete form.

- Rasterization algorithms (computer vision): continuous shape $\rightarrow$ discrete shape
    - Line - Bresenham line algorithm ([Bresenham, 1965])
    - Circle - Bresenham circle algorithm ([Bresenham, 1977])
    - Bezier curves - [Foley, 1995]

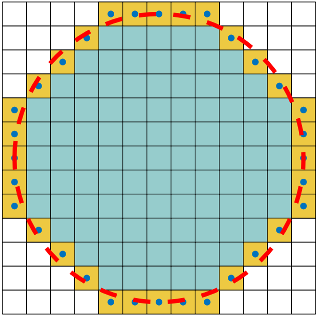




Shape of the scanning window

To each discrete shape $\tilde{G}$ it corresponds an unique matrix (2-way tensor) $A(G) = A\left(\tilde{G}\right)$ (of smallest size) with 0 and 1 entries (1 if there is a point and 0 otherwise):

$\tilde{G}$: **Circle**



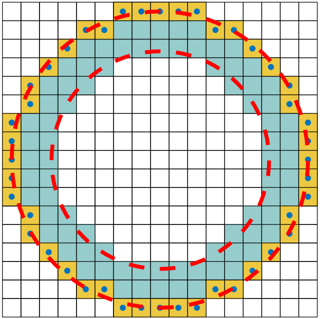
$A\left(\tilde{G}\right)$: **Circle**

				1	1	1	1	1				
			1	1	1	1	1	1	1			
		1	1	1	1	1	1	1	1	1		
	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	
		1	1	1	1	1	1	1	1	1		
			1	1	1	1	1	1	1			
				1	1	1	1	1				

Shape of the scanning window

To each discrete shape $\tilde{G}$ it corresponds an unique matrix (2-way tensor) $A(G) = A\left(\tilde{G}\right)$ (of smallest size) with 0 and 1 entries (1 if there is a point and 0 otherwise):

$\tilde{G}$: Annulus



$A\left(\tilde{G}\right)$: Annulus

						1	1	1	1	1						
				1	1	1	1	1	1	1	1	1				
			1	1	1	1	1	1	1	1	1	1	1			
		1	1	1	1						1	1	1	1		
	1	1	1	1								1	1	1	1	
	1	1	1										1	1	1	
1	1	1												1	1	1
1	1	1												1	1	1
1	1	1												1	1	1
1	1	1												1	1	1
1	1	1												1	1	1
	1	1	1										1	1	1	
	1	1	1	1								1	1	1	1	
		1	1	1	1						1	1	1	1		
			1	1	1	1	1	1	1	1	1	1	1			
				1	1	1	1	1	1	1	1	1				
						1	1	1	1	1						

Arbitrary window scan statistic

Let G be a geometric shape and $A = A(G)$ its corresponding $\{0,1\}$ matrix of size $m_1 \times m_2$.

- Define the score function $\mathcal{S}$ associated to the shape G by

$$\mathcal{S}\left(\mathcal{X}_{i_1,i_2}\right) = A \circ \mathcal{X}_{i_1,i_2} = \sum_{s_1=i_1}^{i_1+m_1-1} \sum_{s_2=i_2}^{i_2+m_2-1} A(s_1 - i_1 + 1, s_2 - i_2 + 1) X_{s_1,s_2}$$

Remark

If, in particular, the shape G is a rectangle of size $m_1 \times m_2$ than its corresponding $\{0,1\}$ matrix of the same size has all the entries equal to 1 so the score function

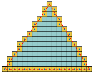
$$\mathcal{S}\left(\mathcal{X}_{i_1,i_2}\right) = \sum_{s_1=i_1}^{i_1+m_1-1} \sum_{s_2=i_2}^{i_2+m_2-1} X_{s_1,s_2}$$

is the *classical* rectangular window of the two dimensional scan statistics.

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 1: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Triangle

Window's shape	Triangle ($m_1 = 14,\ m_2 = 18,\ Nt = 133,\ IS = 1e5,\ IA = 1e6$)



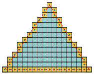
$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.916397	0.918667	0.004333	59	0.863336	0.897101	0.004902
4	0.997488	0.997483	0.000082	60	0.936684	0.948160	0.002010
5	0.999947	0.999947	0.000002	61	0.971529	0.974938	0.000894
6	0.999999	0.999999	0	62	0.987094	0.988279	0.000412
7	0.999999	0.999999	0	63	0.994372	0.994664	0.000192
8	1.000000	1.000000	0	64	0.997563	0.997643	0.000089
9	1.000000	1.000000	0	65	0.998946	0.998971	0.000041
10	1.000000	1.000000	0	66	0.999564	0.999567	0.000018
11	1.000000	1.000000	0	67	0.999817	0.999820	0.000008

$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.739866	0.820277	0.011813	50	0.911170	0.934999	0.002737
60	0.871829	0.902957	0.004483	51	0.945219	0.957777	0.001655
61	0.939577	0.950977	0.001911	52	0.966494	0.972997	0.001026
62	0.971673	0.975649	0.000890	53	0.979944	0.982929	0.000644
63	0.987075	0.988206	0.000425	54	0.987840	0.989469	0.000406
64	0.994104	0.994492	0.000204	55	0.992768	0.993477	0.000257
65	0.997384	0.997452	0.000098	56	0.995667	0.996022	0.000162
66	0.998821	0.998855	0.000046	57	0.997412	0.997574	0.000102
67	0.999489	0.999490	0.000022	58	0.998509	0.998563	0.000063

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 1: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Triangle

Window's shape — Triangle ($m_1 = 14,\ m_2 = 18,\ Nt = 133,\ IS = 1e5,\ IA = 1e6$)



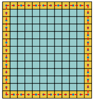
$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.916397	0.918667	0.004333	59	0.863336	0.897101	0.004902
4	0.997488	0.997483	0.000082	60	0.936684	0.948160	0.002010
5	0.999947	0.999947	0.000002	61	0.971529	0.974938	0.000894
6	0.999999	0.999999	0	62	0.987094	0.988279	0.000412
7	0.999999	0.999999	0	63	0.994372	0.994664	0.000192
8	1.000000	1.000000	0	64	0.997563	0.997643	0.000089
9	1.000000	1.000000	0	65	0.998946	0.998971	0.000041
10	1.000000	1.000000	0	66	0.999564	0.999567	0.000018
11	1.000000	1.000000	0	67	0.999817	0.999820	0.000008

$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.739866	0.820277	0.011813	50	0.911170	0.934999	0.002737
60	0.871829	0.902957	0.004483	51	0.945219	0.957777	0.001655
61	0.939577	0.950977	0.001911	52	0.966494	0.972997	0.001026
62	0.971673	0.975649	0.000890	53	0.979944	0.982929	0.000644
63	0.987075	0.988206	0.000425	54	0.987840	0.989469	0.000406
64	0.994104	0.994492	0.000204	55	0.992768	0.993477	0.000257
65	0.997384	0.997452	0.000098	56	0.995667	0.996022	0.000162
66	0.998821	0.998855	0.000046	57	0.997412	0.997574	0.000102
67	0.999489	0.999490	0.000022	58	0.998509	0.998563	0.000063

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 2: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Rectangle

Window's shape — **Rectangle ($m_1 = 11$, $m_2 = 12$, $Nt = 132$, $IS = 1e5$, $IA = 1e6$)**



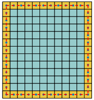
	$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$		
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.947896	0.948226	0.002110	59	0.856993	0.857094	0.005485
4	0.997983	0.997986	0.000058	60	0.919416	0.919589	0.002300
5	0.999943	0.999943	0.000001	61	0.956420	0.956569	0.001024
6	0.999999	0.999999	0	62	0.977142	0.977065	0.000471
7	0.999999	0.999999	0	63	0.988228	0.988208	0.000220
8	1.000000	1.000000	0	64	0.994130	0.994095	0.000103
9	1.000000	1.000000	0	65	0.997121	0.997107	0.000048
10	1.000000	1.000000	0	66	0.998610	0.998607	0.000022
11	1.000000	1.000000	0	67	0.999342	0.999342	0.000010

	$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$		
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.764372	0.764495	0.013285	50	0.865858	0.865484	0.004572
60	0.857872	0.859289	0.005320	51	0.904555	0.904805	0.002691
61	0.918972	0.918732	0.002307	52	0.933323	0.933206	0.001620
62	0.954682	0.954579	0.001059	53	0.953950	0.953807	0.000993
63	0.975271	0.975391	0.000502	54	0.968340	0.968483	0.000617
64	0.986996	0.986966	0.000242	55	0.978632	0.978687	0.000386
65	0.993240	0.993261	0.000117	56	0.985791	0.985752	0.000242
66	0.996557	0.996551	0.000056	57	0.990621	0.990585	0.000152
67	0.998290	0.998283	0.000027	58	0.993837	0.993815	0.000096

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 2: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Rectangle

Window's shape	Rectangle ($m_1 = 11,\ m_2 = 12,\ Nt = 132,\ IS = 1e5,\ IA = 1e6$)



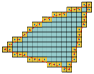
$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.947896	0.948226	0.002110	59	0.856993	0.857094	0.005485
4	0.997983	0.997986	0.000058	60	0.919416	0.919589	0.002300
5	0.999943	0.999943	0.000001	61	0.956420	0.956569	0.001024
6	0.999999	0.999999	0	62	0.977142	0.977065	0.000471
7	0.999999	0.999999	0	63	0.988228	0.988208	0.000220
8	1.000000	1.000000	0	64	0.994130	0.994095	0.000103
9	1.000000	1.000000	0	65	0.997121	0.997107	0.000048
10	1.000000	1.000000	0	66	0.998610	0.998607	0.000022
11	1.000000	1.000000	0	67	0.999342	0.999342	0.000010

$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.764372	0.764495	0.013285	50	0.865858	0.865484	0.004572
60	0.857872	0.859289	0.005320	51	0.904555	0.904805	0.002691
61	0.918972	0.918732	0.002307	52	0.933323	0.933206	0.001620
62	0.954682	0.954579	0.001059	53	0.953950	0.953807	0.000993
63	0.975271	0.975391	0.000502	54	0.968340	0.968483	0.000617
64	0.986996	0.986966	0.000242	55	0.978632	0.978687	0.000386
65	0.993240	0.993261	0.000117	56	0.985791	0.985752	0.000242
66	0.996557	0.996551	0.000056	57	0.990621	0.990585	0.000152
67	0.998290	0.998283	0.000027	58	0.993837	0.993815	0.000096

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 3: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Quadrilateral

Window's shape — **Quadrilateral** ($m_1 = 14,\ m_2 = 18,\ Nt = 131,\ IS = 1e5,\ IA = 1e6$)



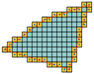
	$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$		
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.926068	0.927398	0.003806	59	0.914546	0.927613	0.002942
4	0.997622	0.997627	0.000075	60	0.959599	0.963873	0.001255
5	0.999946	0.999946	0.000002	61	0.981235	0.982506	0.000571
6	0.999999	0.999999	0	62	0.991423	0.991796	0.000266
7	0.999999	0.999999	0	63	0.996113	0.996233	0.000124
8	1.000000	1.000000	0	64	0.998283	0.998337	0.000057
9	1.000000	1.000000	0	65	0.999266	0.999266	0.000026
10	1.000000	1.000000	0	66	0.999684	0.999684	0.000012
11	1.000000	1.000000	0	67	0.999868	0.999869	0.000005

	$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$		
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.835054	0.870351	0.006852	50	0.920004	0.935266	0.002571
60	0.917972	0.931040	0.002768	51	0.950232	0.957711	0.001556
61	0.960397	0.964711	0.001237	52	0.968755	0.972594	0.000964
62	0.981228	0.982451	0.000585	53	0.980695	0.982566	0.000606
63	0.991142	0.991510	0.000281	54	0.988110	0.989060	0.000383
64	0.995855	0.995971	0.000136	55	0.992626	0.993110	0.000242
65	0.998108	0.998124	0.000065	56	0.995569	0.995771	0.000153
66	0.999135	0.999153	0.000031	57	0.997361	0.997394	0.000096
67	0.999620	0.999622	0.000014	58	0.998379	0.998435	0.000060

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 3: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Quadrilateral

Window's shape — Quadrilateral ($m_1 = 14$, $m_2 = 18$, $Nt = 131$, $IS = 1e5$, $IA = 1e6$)



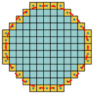
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
3	0.926068	0.927398	0.003806	59	0.914546	0.927613	0.002942
4	0.997622	0.997627	0.000075	60	0.959599	0.963873	0.001255
5	0.999946	0.999946	0.000002	61	0.981235	0.982506	0.000571
6	0.999999	0.999999	0	62	0.991423	0.991796	0.000266
7	0.999999	0.999999	0	63	0.996113	0.996233	0.000124
8	1.000000	1.000000	0	64	0.998283	0.998337	0.000057
9	1.000000	1.000000	0	65	0.999266	0.999266	0.000026
10	1.000000	1.000000	0	66	0.999684	0.999684	0.000012
11	1.000000	1.000000	0	67	0.999868	0.999869	0.000005

τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$			
59	0.835054	0.870351	0.006852	50	0.920004	0.935266	0.002571
60	0.917972	0.931040	0.002768	51	0.950232	0.957711	0.001556
61	0.960397	0.964711	0.001237	52	0.968755	0.972594	0.000964
62	0.981228	0.982451	0.000585	53	0.980695	0.982566	0.000606
63	0.991142	0.991510	0.000281	54	0.988110	0.989060	0.000383
64	0.995855	0.995971	0.000136	55	0.992626	0.993110	0.000242
65	0.998108	0.998124	0.000065	56	0.995569	0.995771	0.000153
66	0.999135	0.999153	0.000031	57	0.997361	0.997394	0.000096
67	0.999620	0.999622	0.000014	58	0.998379	0.998435	0.000060

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 4: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Circle

Window's shape | **Circle** $(m_1 = 13,\ m_2 = 13,\ Nt = 129,\ IS = 1e54,\ IA = 1e6)$



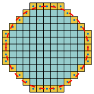
	$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$		
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.950311	0.950461	0.002195	59	0.920229	0.920388	0.002318
4	0.998118	0.998114	0.000059	60	0.956814	0.957143	0.001016
5	0.999947	0.999947	0.000001	61	0.977460	0.977614	0.000462
6	0.999999	0.999999	0	62	0.988568	0.988567	0.000214
7	0.999999	0.999999	0	63	0.994312	0.994309	0.000099
8	1.000000	1.000000	0	64	0.997229	0.997228	0.000046
9	1.000000	1.000000	0	65	0.998678	0.998679	0.000021
10	1.000000	1.000000	0	66	0.999380	0.999381	0.000009
11	1.000000	1.000000	0	67	0.999715	0.999715	0.000004

	$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$		
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.858454	0.859178	0.005391	50	0.888137	0.887891	0.003485
60	0.919182	0.919586	0.002310	51	0.921173	0.921549	0.002058
61	0.955229	0.955388	0.001047	52	0.945761	0.945644	0.001243
62	0.976023	0.975987	0.000491	53	0.962790	0.962825	0.000760
63	0.987414	0.987344	0.000234	54	0.974848	0.974878	0.000470
64	0.993477	0.993502	0.000112	55	0.983235	0.983263	0.000293
65	0.996706	0.996703	0.000054	56	0.988885	0.988907	0.000182
66	0.998372	0.998365	0.000025	57	0.992730	0.992734	0.000114
67	0.999207	0.999203	0.000012	58	0.995269	0.995287	0.000071

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 4: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Circle

Window's shape	Circle ($m_1 = 13,\ m_2 = 13,\ Nt = 129,\ IS = 1e54,\ IA = 1e6$)



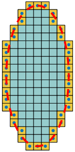
$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.950311	0.950461	0.002195	59	0.920229	0.920388	0.002318
4	0.998118	0.998114	0.000059	60	0.956814	0.957143	0.001016
5	0.999947	0.999947	0.000001	61	0.977460	0.977614	0.000462
6	0.999999	0.999999	0	62	0.988568	0.988567	0.000214
7	0.999999	0.999999	0	63	0.994312	0.994309	0.000099
8	1.000000	1.000000	0	64	0.997229	0.997228	0.000046
9	1.000000	1.000000	0	65	0.998678	0.998679	0.000021
10	1.000000	1.000000	0	66	0.999380	0.999381	0.000009
11	1.000000	1.000000	0	67	0.999715	0.999715	0.000004

$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.858454	0.859178	0.005391	50	0.888137	0.887891	0.003485
60	0.919182	0.919586	0.002310	51	0.921173	0.921549	0.002058
61	0.955229	0.955388	0.001047	52	0.945761	0.945644	0.001243
62	0.976023	0.975987	0.000491	53	0.962790	0.962825	0.000760
63	0.987414	0.987344	0.000234	54	0.974848	0.974878	0.000470
64	0.993477	0.993502	0.000112	55	0.983235	0.983263	0.000293
65	0.996706	0.996703	0.000054	56	0.988885	0.988907	0.000182
66	0.998372	0.998365	0.000025	57	0.992730	0.992734	0.000114
67	0.999207	0.9992032	0.000012	58	0.995269	0.995287	0.000071

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 5: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Ellipse

Window's shape — **Ellipse** $(m_1 = 19,\ m_2 = 9,\ Nt = 135,\ IS = 1e5,\ IA = 1e6)$



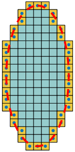
$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.944001	0.944211	0.002297	59	0.764871	0.763482	0.009128
4	0.997757	0.997758	0.000069	60	0.860903	0.860548	0.004127
5	0.999935	0.999935	0.000002	61	0.921089	0.920882	0.001941
6	0.999998	0.999998	0	62	0.956735	0.956866	0.000934
7	1.000000	1.000000	0	63	0.977118	0.977094	0.000452
8	1.000000	1.000000	0	64	0.988182	0.988152	0.000218
9	1.000000	1.000000	0	65	0.994044	0.994012	0.000104
10	1.000000	1.000000	0	66	0.997037	0.997037	0.000049
11	1.000000	1.000000	0	67	0.998554	0.998558	0.000023

$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.638962	0.639102	0.019442	50	0.843156	0.844113	0.004369
60	0.768283	0.769666	0.008610	51	0.887477	0.887692	0.002755
61	0.861614	0.860885	0.004012	52	0.920601	0.920385	0.001757
62	0.919144	0.919301	0.001948	53	0.944398	0.944328	0.001127
63	0.954941	0.954864	0.000965	54	0.961682	0.961667	0.000725
64	0.975255	0.975369	0.000481	55	0.973797	0.973864	0.000468
65	0.986869	0.986900	0.000240	56	0.982232	0.982330	0.000301
66	0.993115	0.993127	0.000119	57	0.988183	0.988132	0.000193
67	0.996485	0.996472	0.000058	58	0.992138	0.992134	0.000123

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 5: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Ellipse

Window's shape — Ellipse ($m_1 = 19$, $m_2 = 9$, $Nt = 135$, $IS = 1e5$, $IA = 1e6$)



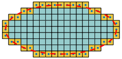
$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.944001	0.944211	0.002297	59	0.764871	0.763482	0.009128
4	0.997757	0.997758	0.000069	60	0.860903	0.860548	0.004127
5	0.999935	0.999935	0.000002	61	0.921089	0.920882	0.001941
6	0.999998	0.999998	0	62	0.956735	0.956866	0.000934
7	1.000000	1.000000	0	63	0.977118	0.977094	0.000452
8	1.000000	1.000000	0	64	0.988182	0.988152	0.000218
9	1.000000	1.000000	0	65	0.994044	0.994012	0.000104
10	1.000000	1.000000	0	66	0.997037	0.997037	0.000049
11	1.000000	1.000000	0	67	0.998554	0.998558	0.000023

$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.638962	0.639102	0.019442	50	0.843156	0.844113	0.004369
60	0.768283	0.769666	0.008610	51	0.887477	0.887692	0.002755
61	0.861614	0.860885	0.004012	52	0.920601	0.920385	0.001757
62	0.919144	0.919301	0.001948	53	0.944398	0.944328	0.001127
63	0.954941	0.954864	0.000965	54	0.961682	0.961667	0.000725
64	0.975255	0.975369	0.000481	55	0.973797	0.973864	0.000468
65	0.986869	0.986900	0.000240	56	0.982232	0.982330	0.000301
66	0.993115	0.993127	0.000119	57	0.988183	0.988132	0.000193
67	0.996485	0.996472	0.000058	58	0.992138	0.992134	0.000123

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 6: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Ellipse2

Window's shape	Ellipse2 $(m_1 = 9,\ m_2 = 19,\ Nt = 135,\ IS = 1e5,\ IA = 1e6)$



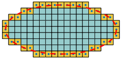
$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.944180	0.943779	0.003420	59	0.764090	0.763753	0.036701
4	0.997761	0.997767	0.000071	60	0.859511	0.860602	0.011219
5	0.999934	0.999935	0.000002	61	0.921173	0.920669	0.003816
6	0.999998	0.999998	0	62	0.956920	0.956693	0.001440
7	0.999999	0.999999	0	63	0.977198	0.977129	0.000586
8	1.000000	1.000000	0	64	0.988162	0.988177	0.000253
9	1.000000	1.000000	0	65	0.993979	0.994008	0.000113
10	1.000000	1.000000	0	66	0.997032	0.997036	0.000051
11	1.000000	1.000000	0	67	0.998558	0.998562	0.000023

$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.638389	0.640805	0.117483	50	0.843792	0.844342	0.013346
60	0.770899	0.769795	0.034030	51	0.888070	0.887823	0.006857
61	0.860847	0.861092	0.010890	52	0.920460	0.920691	0.003626
62	0.919522	0.919537	0.003909	53	0.944514	0.944368	0.001974
63	0.954873	0.954742	0.001516	54	0.961591	0.961748	0.001109
64	0.975326	0.975379	0.000635	55	0.973861	0.973830	0.000640
65	0.986856	0.986843	0.000282	56	0.982294	0.982269	0.000377
66	0.993154	0.993119	0.000130	57	0.988180	0.988143	0.000226
67	0.996482	0.996478	0.000061	58	0.992154	0.992123	0.000138

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 6: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Ellipse2

Window's shape	Ellipse2 ($m_1 = 9$, $m_2 = 19$, $Nt = 135$, $IS = 1e5$, $IA = 1e6$)



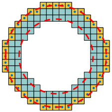
$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.944180	0.943779	0.003420	59	0.764090	0.763753	0.036701
4	0.997761	0.997767	0.000071	60	0.859511	0.860602	0.011219
5	0.999934	0.999935	0.000002	61	0.921173	0.920669	0.003816
6	0.999998	0.999998	0	62	0.956920	0.956693	0.001440
7	0.999999	0.999999	0	63	0.977198	0.977129	0.000586
8	1.000000	1.000000	0	64	0.988162	0.988177	0.000253
9	1.000000	1.000000	0	65	0.993979	0.994008	0.000113
10	1.000000	1.000000	0	66	0.997032	0.997036	0.000051
11	1.000000	1.000000	0	67	0.998558	0.998562	0.000023

$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.638389	0.640805	0.117483	50	0.843792	0.844342	0.013346
60	0.770899	0.769795	0.034030	51	0.888070	0.887823	0.006857
61	0.860847	0.861092	0.010890	52	0.920460	0.920691	0.003626
62	0.919522	0.919537	0.003909	53	0.944514	0.944368	0.001974
63	0.954873	0.954742	0.001516	54	0.961591	0.961748	0.001109
64	0.975326	0.975379	0.000635	55	0.973861	0.973830	0.000640
65	0.986856	0.986843	0.000282	56	0.982294	0.982269	0.000377
66	0.993154	0.993119	0.000130	57	0.988180	0.988143	0.000226
67	0.996482	0.996478	0.000061	58	0.992154	0.992123	0.000138

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 7: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Annulus

Window's shape — **Annulus** ($m_1 = 17$, $m_2 = 17$, $Nt = 124$, $IS = 1e5$, $IA = 1e6$)



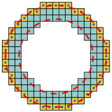
	$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$		
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
3	0.881798	0.882489	0.004812	59	0.951170	0.951245	0.000699
4	0.995434	0.995465	0.000069	60	0.975772	0.975727	0.000255
5	0.999883	0.999883	0.000001	61	0.988275	0.988270	0.000099
6	0.999998	0.999998	0	62	0.994460	0.994466	0.000041
7	0.999999	0.999999	0	63	0.997439	0.997440	0.000017
8	1.000000	1.000000	0	64	0.998839	0.998840	0.000007
9	1.000000	1.000000	0	65	0.999484	0.999484	0.000003
10	1.000000	1.000000	0	66	0.999775	0.999775	0.000001
11	1.000000	1.000000	0	67	0.999903	0.999903	0.000000

	$X_{s_1,s_2} \sim \mathcal{P}(0.25)$				$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$		
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
59	0.903956	0.903852	0.002128	50	0.860708	0.860416	0.004097
60	0.949083	0.949059	0.000735	51	0.904651	0.904644	0.001977
61	0.973814	0.973844	0.000277	52	0.936077	0.935938	0.000987
62	0.986909	0.986876	0.000111	53	0.957564	0.957650	0.000508
63	0.993576	0.993570	0.000047	54	0.972378	0.972311	0.000270
64	0.996910	0.996907	0.000020	55	0.982139	0.982136	0.000148
65	0.998539	0.998539	0.000009	56	0.988564	0.988587	0.000082
66	0.999320	0.999320	0.000004	57	0.992769	0.992769	0.000047
67	0.999689	0.999689	0.000002	58	0.995471	0.995466	0.000027

Scanning a region of size $T_1 \times T_2 = 250 \times 250$

Table 7: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Annulus

Window's shape — **Annulus** $(m_1 = 17,\ m_2 = 17,\ Nt = 124,\ IS = 1e5,\ IA = 1e6)$



$X_{s_1,s_2} \sim \mathcal{B}(1, 0.01)$

τ	Sim	AppH	ETotal
3	0.881798	0.882489	0.004812
4	0.995434	0.995465	0.000069
5	0.999883	0.999883	0.000001
6	0.999998	0.999998	0
7	0.999999	0.999999	0
8	1.000000	1.000000	0
9	1.000000	1.000000	0
10	1.000000	1.000000	0
11	1.000000	1.000000	0

$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$

τ	Sim	AppH	ETotal
59	0.951170	0.951245	0.000699
60	0.975772	0.975727	0.000255
61	0.988275	0.988270	0.000099
62	0.994460	0.994466	0.000041
63	0.997439	0.997440	0.000017
64	0.998839	0.998840	0.000007
65	0.999484	0.999484	0.000003
66	0.999775	0.999775	0.000001
67	0.999903	0.999903	0.000000

$X_{s_1,s_2} \sim \mathcal{P}(0.25)$

τ	Sim	AppH	ETotal
59	0.903956	0.903852	0.002128
60	0.949083	0.949059	0.000735
61	0.973814	0.973844	0.000277
62	0.986909	0.986876	0.000111
63	0.993576	0.993570	0.000047
64	0.996910	0.996907	0.000020
65	0.998539	0.998539	0.000009
66	0.999320	0.999320	0.000004
67	0.999689	0.999689	0.000002

$X_{s_1,s_2} \sim \mathcal{N}(0, 1)$

τ	Sim	AppH	ETotal
50	0.860708	0.860416	0.004097
51	0.904651	0.904644	0.001977
52	0.936077	0.935938	0.000987
53	0.957564	0.957650	0.000508
54	0.972378	0.972311	0.000270
55	0.982139	0.982136	0.000148
56	0.988564	0.988587	0.000082
57	0.992769	0.992769	0.000047
58	0.995471	0.995466	0.000027

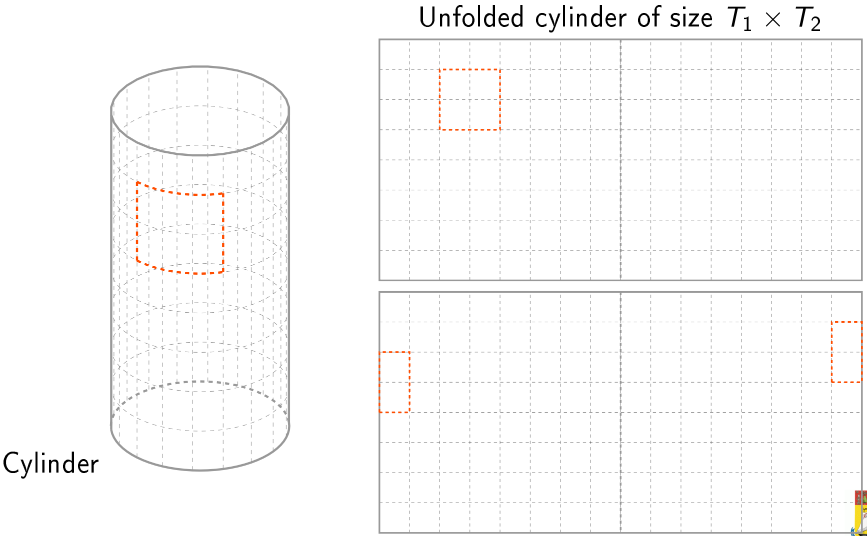
Outline

1. Introduction
   - Framework
   - Problem
2. Methodology
   - Approximation
3. Applications
   - Application: Length of the Longest increasing/non-decreasing run
   - Application: Scanning with windows of arbitrary shape
   - Application: Scanning the surface of a cylinder

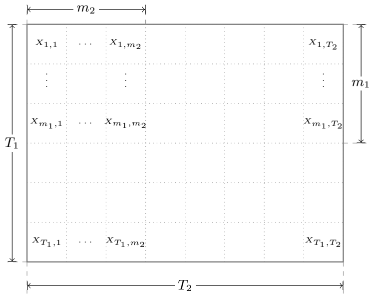
Application 3:

Scanning the surface of a cylinder

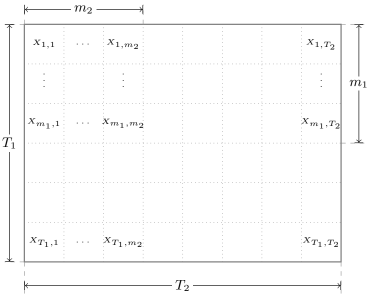
Scanning the surface of a cylinder



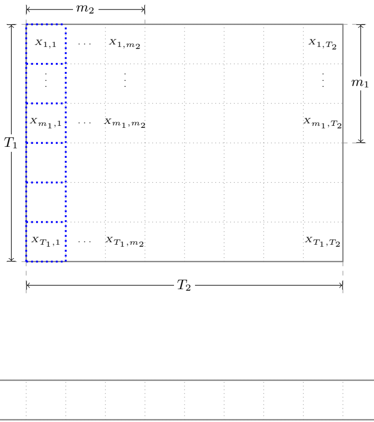
Transformation of the unfolded cylinder



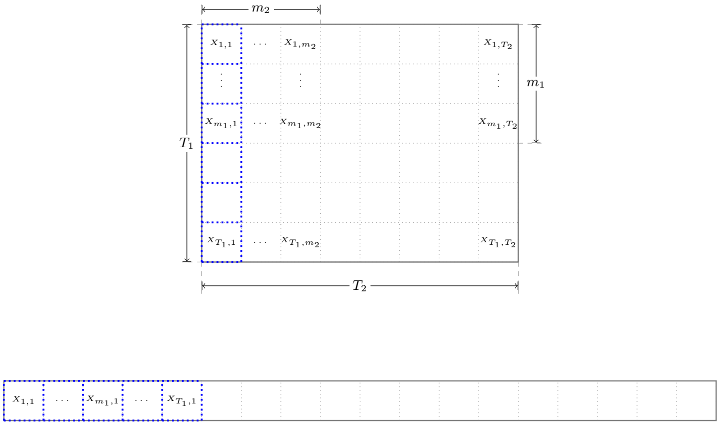
Transformation of the unfolded cylinder



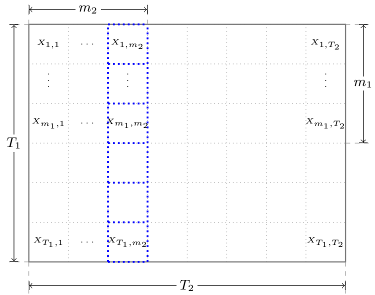
Transformation of the unfolded cylinder



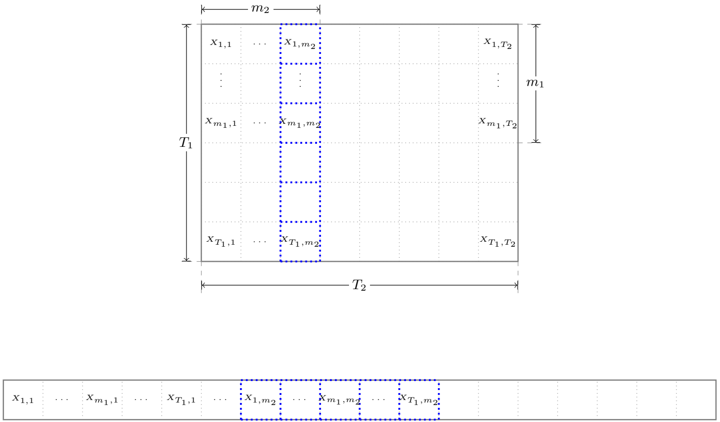
Transformation of the unfolded cylinder



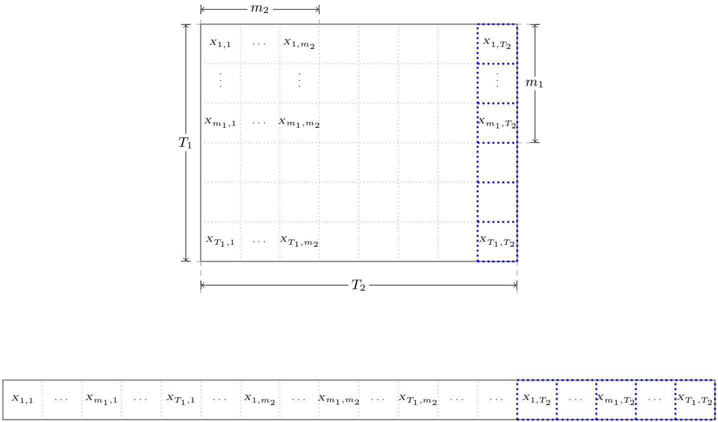
Transformation of the unfolded cylinder



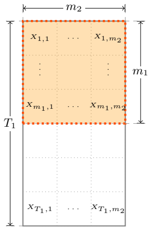
Transformation of the unfolded cylinder



Transformation of the unfolded cylinder



Defining the score function

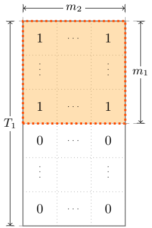


- The size of the scanning window is $\tilde{m}_1 = T_1 m_2 - (T_1 - m_1)$
- The score function is defined by

$$\mathcal{S}\left(x_1, \ldots, x_{\tilde{m}_1}\right) = \sum_{i_1=1}^{\tilde{m}_1} A(i_1) X_{i_1}$$

where A is the corresponding $\{0,1\}$ vector

Defining the score function

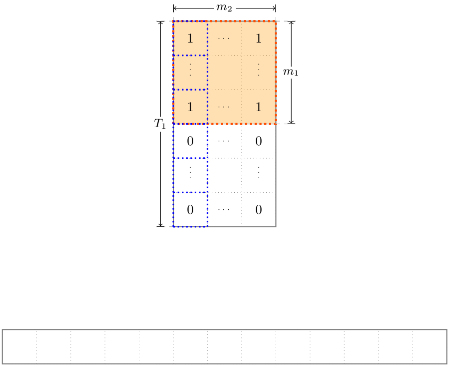


- The size of the scanning window is $\tilde{m}_1 = T_1 m_2 - (T_1 - m_1)$
- The score function is defined by

$$\mathcal{S}(x_1, \ldots, x_{\tilde{m}_1}) = \sum_{i_1=1}^{\tilde{m}_1} A(i_1) X_{i_1}$$

where A is the corresponding $\{0, 1\}$ vector

Defining the score function

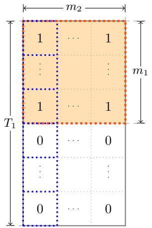


- The size of the scanning window is $\tilde{m}_1 = T_1 m_2 - (T_1 - m_1)$
- The score function is defined by

$$\mathcal{S}(x_1, \ldots, x_{\tilde{m}_1}) = \sum_{i_1=1}^{\tilde{m}_1} A(i_1) X_{i_1}$$

where A is the corresponding $\{0, 1\}$ vector

Defining the score function

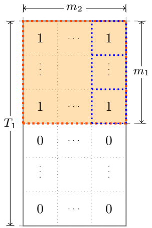


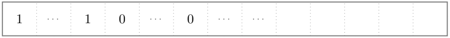
- The size of the scanning window is $\tilde{m}_1 = T_1 m_2 - (T_1 - m_1)$
- The score function is defined by

$$\mathcal{S}(x_1, \dots, x_{\tilde{m}_1}) = \sum_{i_1=1}^{\tilde{m}_1} A(i_1) X_{i_1}$$

where A is the corresponding $\{0, 1\}$ vector

Defining the score function



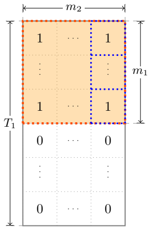


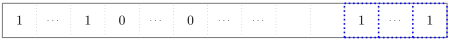
- The size of the scanning window is $\tilde{m}_1 = T_1 m_2 - (T_1 - m_1)$
- The score function is defined by

$$\mathcal{S}(x_1, \dots, x_{\tilde{m}_1}) = \sum_{i_1=1}^{\tilde{m}_1} A(i_1) X_{i_1}$$

where A is the corresponding $\{0, 1\}$ vector

Defining the score function





- The size of the scanning window is $\tilde{m}_1 = T_1 m_2 - (T_1 - m_1)$
- The score function is defined by

$$\mathcal{S}(x_1, \dots, x_{\tilde{m}_1}) = \sum_{i_1=1}^{\tilde{m}_1} A(i_1) X_{i_1}$$

where A is the corresponding $\{0, 1\}$ vector

Approximation and error bounds

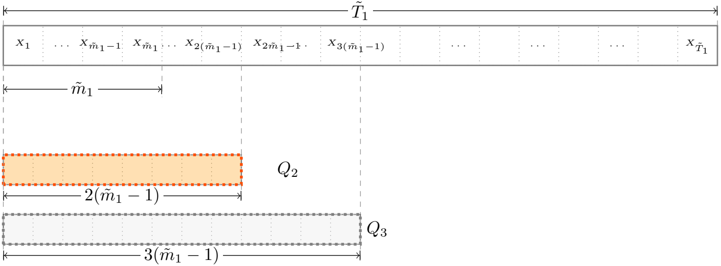
Theorem [Amărioarei, 2014]

Let $t_1 \in \{2,3\}$ and $Q_{t_1} = Q_{t_1}(\tau) = \mathbb{P}\left(S_{\tilde{m}_1}\left(t_1(\tilde{m}_1 - 1); \mathcal{S}\right) \leq \tau\right)$ and $L_1 = \left\lfloor \frac{\tilde{T}_1}{\tilde{m}_1 - 1} \right\rfloor$

If $\hat{Q}_{t_1}$ is an estimate of Q_{t_1} with $\left|\hat{Q}_{t_1} - Q_{t_1}\right| \leq \beta_{t_1}$ and τ is such that $1 - \hat{Q}_2(\tau) \leq 0.1$ then

$$\left|\mathbb{P}\left(S_{\tilde{m}_1}(\tilde{T}_1, \mathcal{S}) \leq \tau\right) - \left(2\hat{Q}_2 - \hat{Q}_3\right)\left[1 + \hat{Q}_2 - \hat{Q}_3 + 2(\hat{Q}_2 - \hat{Q}_3)^2\right]^{1-L_1}\right| \leq E_{total}(1),$$

$$E_{total}(1) = (L_1 - 1)\left[\beta_2 + \beta_3 + F\left(\hat{Q}_2, L_1 - 1\right)\left(1 - \hat{Q}_2 + \beta_2\right)^2\right].$$



Scanning a region of size $T_1 \times T_2 = 300 \times 350$

Table 8: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Cylinder

Cylinder: $(m_1 = 10,\ m_2 = 15,\ T_1 = 300,\ T_2 = 350,\ IS = 1e4,\ IA = 1e5)$

$X_{s_1,s_2} \sim \mathcal{B}(1, 0.1)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
33	0.871559	0.870200	0.003674	68	0.955593	0.955671	0.000938
34	0.946216	0.946527	0.001177	69	0.976348	0.976285	0.000461
35	0.979458	0.979381	0.000393	70	0.987406	0.987574	0.000227
36	0.992379	0.992465	0.000131	71	0.993526	0.993593	0.000111
37	0.997384	0.997367	0.000043	72	0.996772	0.996751	0.000054
38	0.999116	0.999111	0.000014	73	0.998401	0.998391	0.000026
39	0.999713	0.999714	0.000004	74	0.999214	0.999212	0.000012
40	0.999911	0.999911	0.000001	75	0.999623	0.999626	0.000006

Scanning a region of size $T_1 \times T_2 = 300 \times 350$

Table 8: Numerical results for $\mathbb{P}(S \leqslant \tau)$: Cylinder

Cylinder: $(m_1 = 10,\ m_2 = 15,\ T_1 = 300,\ T_2 = 350,\ IS = 1e4,\ IA = 1e5)$

$X_{s_1,s_2} \sim \mathcal{B}(1, 0.1)$				$X_{s_1,s_2} \sim \mathcal{B}(5, 0.05)$			
τ	Sim	AppH	ETotal	τ	Sim	AppH	ETotal
33	0.871559	0.870200	0.003674	68	0.955593	0.955671	0.000938
34	0.946216	0.946527	0.001177	69	0.976348	0.976285	0.000461
35	0.979458	0.979381	0.000393	70	0.987406	0.987574	0.000227
36	0.992379	0.992465	0.000131	71	0.993526	0.993593	0.000111
37	0.997384	0.997367	0.000043	72	0.996772	0.996751	0.000054
38	0.999116	0.999111	0.000014	73	0.998401	0.998391	0.000026
39	0.999713	0.999714	0.000004	74	0.999214	0.999212	0.000012
40	0.999911	0.999911	0.000001	75	0.999623	0.999626	0.000006

thank you!

Error bounds: approximation error

Approximation error

$$E_{app}(d) = \sum_{s=1}^{d} (L_1 - 1)\cdots(L_s - 1) \sum_{t_1,\ldots,t_{s-1}\in\{2,3\}} F_{t_1,\ldots,t_{s-1}} \left(1 - \gamma_{t_1,\ldots,t_{s-1},2} + B_{t_1,\ldots,t_{s-1},2}\right)^2,$$

where for $2 \leq s \leq d$

$$F_{t_1,\ldots,t_{s-1}} = F\left(Q_{t_1,\ldots,t_{s-1},2}, L_s - 1\right), \; F = F\left(Q_2, L_1 - 1\right),$$

$$B_{t_1\ldots,t_{s-1}} = (L_s - 1)\left[F_{t_1,\ldots,t_{s-1}}\left(1 - \gamma_{t_1,\ldots,t_{s-1},2} + B_{t_1\ldots,t_{s-1},2}\right)^2 + \sum_{t_s\in\{2,3\}} B_{t_1\ldots,t_s}\right],$$

$$B_{t_1\ldots,t_{d-1}} = (L_d - 1)F_{t_1,\ldots,t_{d-1}}\left(1 - \gamma_{t_1,\ldots,t_{d-1},2} + B_{t_1\ldots,t_{d-1},2}\right)^2, \; B_{t_1\ldots,t_d} = 0,$$

and for $s = 1$: $\sum_{t_1,t_0\in\{2,3\}} x = x$, $F_{t_1,t_0} = F$, $\gamma_{t_1,t_0,2} = \gamma_2$ and $B_{t_1,t_0,2} = B_2$.

Return

Error bounds: simulation errors

Simulation errors

$$E_{sf}(d) = (L_1 - 1)\dots(L_d - 1) \sum_{t_1,\dots,t_d \in \{2,3\}} \beta_{t_1,\dots,t_d}$$

$$E_{sapp}(d) = \sum_{s=1}^{d} (L_1 - 1)\cdots(L_s - 1) \sum_{t_1,\dots,t_{s-1} \in \{2,3\}} F_{t_1,\dots,t_{s-1}} \Big(1 - \hat{Q}_{t_1,\dots,t_{s-1},2} + A_{t_1,\dots,t_{s-1},2} + C_{t_1,\dots,t_{s-1},2}\Big)^2$$

where for $2 \leq s \leq d$

$$A_{t_1,\dots,t_{s-1}} = (L_s - 1)\dots(L_d - 1) \sum_{t_s,\dots,t_d \in \{2,3\}} \beta_{t_1,\dots,t_d},\ A_{t_1,\dots,t_d} = \beta_{t_1,\dots,t_d}$$

$$C_{t_1\dots,t_{s-1}} = (L_s - 1)\left[F_{t_1,\dots,t_{s-1}} \left(1 - \hat{Q}_{t_1,\dots,t_{s-1},2} + A_{t_1\dots,t_{s-1},2} + C_{t_1\dots,t_{s-1},2}\right)^2 + \sum_{t_s \in \{2,3\}} C_{t_1\dots,t_s}\right]$$

Return

Simulation methods for Normal data

Importance sampling algorithm

Test the null hypothesis of randomness against an alternative of clustering

H_0: The r.v.'s X_{s_1,s_2} are i.i.d. $\mathcal{N}(\mu,\sigma^2)$

H_1: There exists $\mathcal{R}(i_1,i_2)=[i_1-1,i_1+m_1-1]\times[i_2-1,i_2+m_2-1]\subset\mathcal{R}_2$ where the r.v.'s $X_{s_1,s_2}\sim\mathcal{N}(\mu_1,\sigma^2)$, $\mu_1>\mu$ and $X_{s_1,s_2}\sim\mathcal{N}(\mu,\sigma^2)$ outside $\mathcal{R}(i_1,i_2)$

Objective

Find a good estimate for $\mathbb{P}_{H_0}\left(S_{\mathbf{m}}(\mathbf{T};\mathcal{S})\geq\tau\right)$.

We are interested in evaluating the probability

$$\mathbb{P}_{H_0}\left(S_{\mathbf{m}}(\mathbf{T};\mathcal{S})\geq\tau\right)=\mathbb{P}\left(\bigcup_{i_1=1}^{T_1-m_1+1}\bigcup_{i_2=1}^{T_2-m_2+1}E_{i_1,i_2}(\mathcal{S})\right)$$

where $E_{i_1,i_2}(\mathcal{S})=\{Y_{i_1,i_2}(\mathcal{S})\geq\tau\}$.

Importance sampling algorithm

Algorithm 1 Importance Sampling Algorithm for Scan Statistics

Begin

Repeat for each k from 1 to $ITER$ (iterations number)

1: Generate uniformly the couple $(i_1^{(k)}, i_2^{(k)})$ from the set $\{1,\dots,T_1-m_1+1\}\times\{1,\dots,T_2-m_2+1\}$.

2: Given the couple $(i_1^{(k)}, i_2^{(k)})$, generate a sample of the random field $\tilde{\mathbf{X}}^{(\mathbf{k})}=\left\{\tilde{X}_{s_1,s_2}^{(k)}\right\}$, with $s_j\in\{1,\dots,T_j\}$ and $j\in\{1,2\}$, from the conditional distribution of $\mathbf{X}$ given $\left\{Y_{i_1^{(k)},i_2^{(k)}}(\mathcal{S})\geq\tau\right\}$.

3: Take $c_k=C(\tilde{\mathbf{X}}^{(\mathbf{k})})$ the number of all couples (i_1,i_2) for which $\tilde{Y}_{i_1,i_2}(\mathcal{S})\geq\tau$ and put $\widehat{\rho}_k(2)=\frac{1}{c_k}$.

End Repeat

Return

$$\widehat{\rho}(2)=\frac{1}{ITER}\sum_{k=1}^{ITER}\widehat{\rho}_k(2),\quad Var\left[\widehat{\rho}(2)\right]\approx\frac{1}{ITER-1}\sum_{k=1}^{ITER}\left(\widehat{\rho}_k(2)-\frac{1}{ITER}\sum_{k=1}^{ITER}\widehat{\rho}_k(2)\right)^2$$

End

Importance sampling algorithm: $\mathcal{N}(\mu,\sigma^2)$

Step 2 requires to sample:

- $Y_{i_1^{(k)},i_2^{(k)}}(\mathcal{S})$ from the tail distribution $\mathbb{P}\left(Y_{i_1^{(k)},i_2^{(k)}}(\mathcal{S})\geq\tau\right)$ ([?])
- for the other indices, from the conditional distribution given $\left\{Y_{i_1^{(k)},i_2^{(k)}}(\mathcal{S})\geq\tau\right\}$

Lemma (Generalization of [?, Lemma 3.4.4])

Let N be a positive integer, $\mathbf{X}=(X_1,X_2,\dots,X_N)$ be a vector of i.i.d. $\mathcal{N}(\mu,\sigma^2)$ and $\mathbf{a}=(a_1,\dots,a_N)\in\mathbb{R}^N$ a non zero constant vector ($a_j\neq 0$ for some particular j). Then conditionally given $\langle\mathbf{a},\mathbf{X}\rangle=t$, the r.v.'s X_s with $s\neq j$ are jointly distributed as

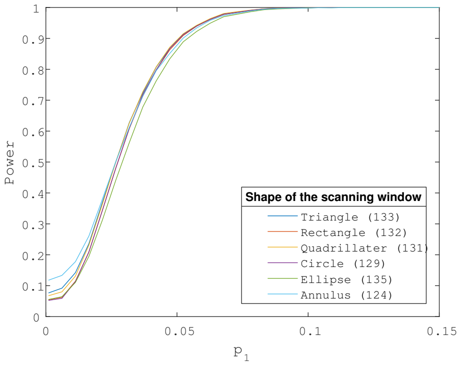
$$\tilde{X}_s=\frac{a_s}{\|a\|}\left[\frac{t-\mu a_j}{\|a\|}-\frac{1}{\|a\|-|a_j|}\sum_{i\neq j}a_i\left(Z_i-\frac{\mu|a_j|}{\|a\|}\right)\right]+Z_s$$

where Z_s are i.i.d. $\mathcal{N}(\mu,\sigma^2)$ r.v.s.

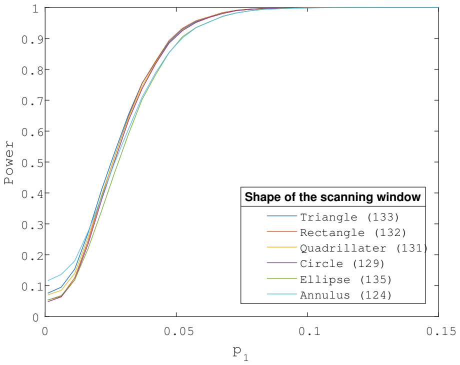
Power of the scan statistic test

Power evaluation for $\mathcal{B}(1, 0.001)$ model

Triangular simulated cluster

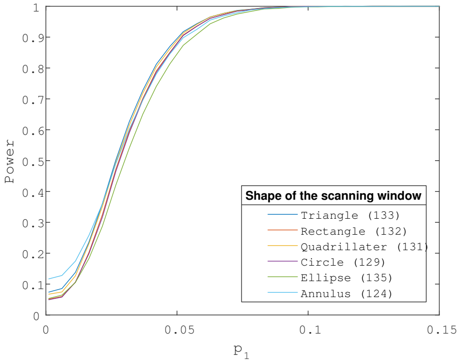


Rectangular simulated cluster

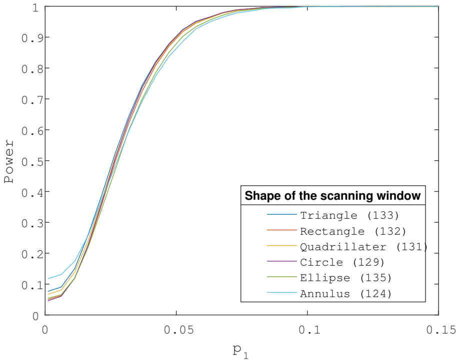


Power evaluation for $\mathcal{B}(1, 0.001)$ model

Quadrilateral simulated cluster

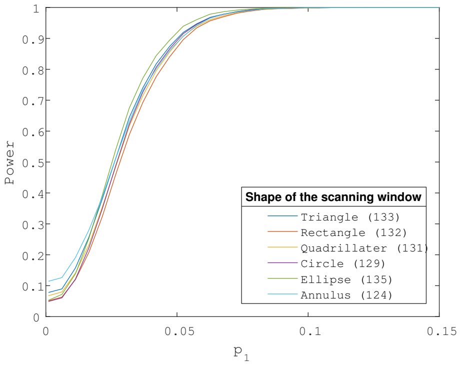


Circular simulated cluster

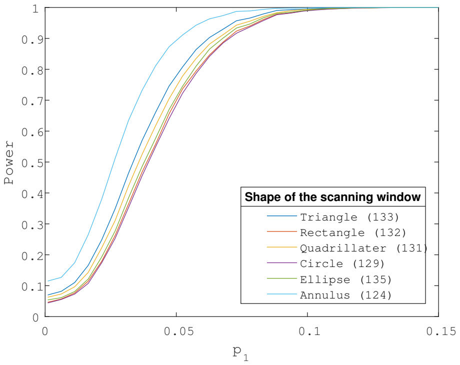


Power evaluation for $\mathcal{B}(1, 0.001)$ model

Ellipsoidal simulated cluster

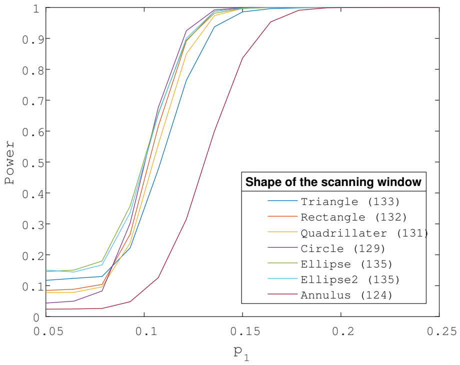


Annular simulated cluster

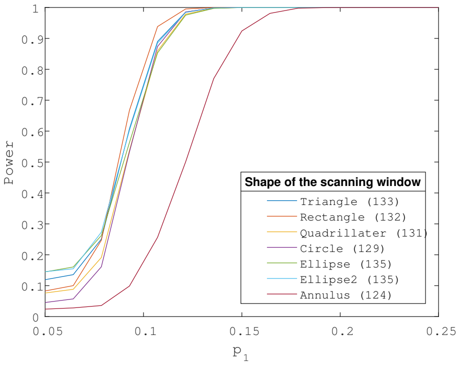


Power evaluation for $\mathcal{B}(5, 0.05)$ model

Triangular simulated cluster

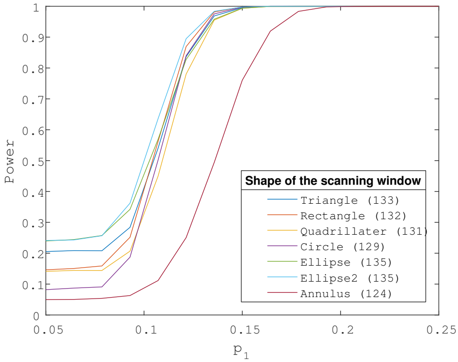


Rectangular simulated cluster

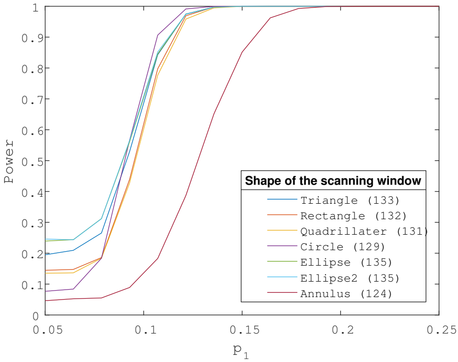


Power evaluation for $\mathcal{B}(5, 0.05)$ model

Quadrilateral simulated cluster

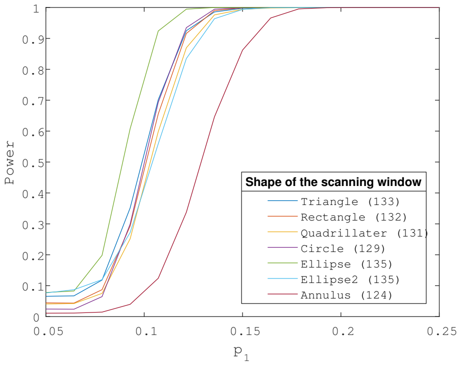


Circular simulated cluster

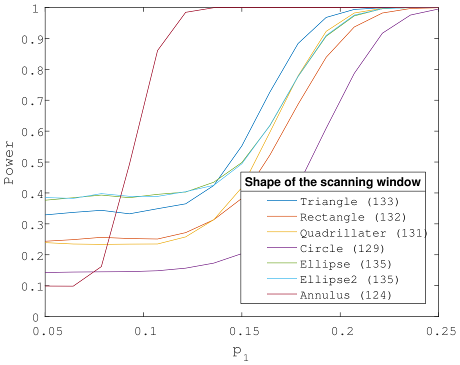


Power evaluation for $\mathcal{B}(5, 0.05)$ model

Ellipsoidal simulated cluster



Annular simulated cluster



Alm, S. E. (1983).
On the distribution of scan statistic of poisson process.
In Gut, A. and Helst, L., editors, *Probability and Mathematical Statistics*, pages 1–10. Upsalla University Press.

Alm, S. E. (1997).
On the distributions of scan statistics of a two-dimensional poisson process.
Advances in Applied Probability, pages 1–18.

Alm, S. E. (1998).
Approximation and simulation of the distributions of scan statistics for poisson processes in higher dimensions.
Extremes, 1(1):111–126.

Amărioarei, A. (2014).
Approximations for the multidimensional discrete scan statistics.
PhD thesis, University of Lille 1.

Anderson, N. H. and Titterington, D. M. (1997).
Some methods for investigating spatial clustering, with epidemiological applications.
Journal of the Royal Statistical Society: Series A (Statistics in Society), 160(1):87–105.

Assunção, R., Costa, M., Tavares, A., and Ferreira, S. (2006).
Fast detection of arbitrarily shaped disease clusters.
Statistics in Medicine, 25(5):723–742.

Bresenham, J. (1965).
Algorithm for computer control of a digital printer.
IBM Systems Journal, 4(1).

Bresenham, J. (1977).
A linear algorithm for incremental digital display of circular arcs.
Communications of the ACM, 20(2):100–106.

Csaki, E. and Foldes, A. (1996).
On the length of theh longest monnotone block.
Studio Scientiarum Mathematicarum Hungarica, 31:35–46.

Eryilmaz, S. (2006).
A note on runs of geometrically distributed random variables.
Discrete Mathematics, 306:1765–1770.

Foley, J. (1995).
Computer Graphics, Principles and Practice in C.

Addison-Wesley Professional.

Glaz, J., Naus, J., and Wallenstein, S. (2001).
Scan Statistics.
Springer.

Glaz, J., Pozdnyakov, V., and Wallenstein, S. (2009).
Scan Statistics: Methods and Applications.
Birkhaüser Boston.

Grabner, P., Knopfmacher, A., and Prodinger, H. (2003).
Combinatorics of geometrically distributed random variables: run statistics.
Theoret. Comput. Sci., 297:261–270.

Grill, K. (1987).
Erdos-Révész type bounds for the length of the longest run from a stationary mixing sequence.
Probab. Theory Relat. Fields, 75:169–179.

Loader, C. R. (1991).
Large-deviation approximations to the distribution of scan statistics.
Advances in Applied Probability, pages 751–771.

Louchard, G. (2005).
Monotone runs of uniformly distributed integer random variables: a probabilistic analysis.
Theoret. Comput. Sci., 346:358–387.

Louchard, G. and Prodinger, H. (2003).
Ascending runs of sequences of geometrically distributed random variables: a probabilistic analysis.
Theoretical Computer Science, 304:59–86.

Novak, S. (1992).
Longest runs in a sequence of m-dependent random variables.
Probab. Theory Relat. Fields, 91:269–281.

Pittel, B. (1981).
Limiting behavior of a process of runs.
Ann. Probab., 9:119–129.

Révész, P. (1983).
Three problems on the llength of increasing runs.
Stochastic Process. Appl., 5:169–179.

Tango, T. and Takahashi, K. (2005).
A flexibly shaped spatial scan statistic for detecting clusters.
Int J Health Geogr, 4:11.
Tango, Toshiro Takahashi, Kunihiko England Int J Health Geogr. 2005 May 18;4:11.