Approximations for the length of the longest monotone run in a sequence of i.i.d. random variables

Alexandru Amărioarei

Polytech'Lille
Université de Lille 1, INRIA/Modal Team, France

The 18^{th} Conference of the Romanian Society of Statistics and Probability
8 May, 2015, București, România

Outline

1. Introduction
   - Framework
   - Problem
2. Methodology
   - One dimensional discrete scan statistics
   - Longest increasing / non-decreasing run and scan statistics
   - Approximations for scan statistics
3. Comparison study
   - Novak's result
   - Numerical examples
4. References

Outline

1. Introduction
   - Framework
   - Problem
2. Methodology
   - One dimensional discrete scan statistics
   - Longest increasing / non-decreasing run and scan statistics
   - Approximations for scan statistics
3. Comparison study
   - Novak's result
   - Numerical examples
4. References

Definitions and notations

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 : $\underline{1\ \ 3\ \ 5}$ 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

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

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

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

Example ($T_1 = 10$)

$$\begin{array}{rcccccccccc}
X_i: & 1 & 3 & 5 & 2 & 4 & 7 & 1 & 3 & 3 & 8 \\
IR: & \underline{1 \;\; 3 \;\; 5} & & & \underline{2 \;\; 4 \;\; 7} & & & 1 & 3 & 3 & 8 \\
NDR: & 1 & 3 & 5 & 2 & 4 & 7 & 1 & 3 & 3 & 8
\end{array}$$

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

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

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

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

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

Example ($T_1 = 10$)

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

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

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

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

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

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

Example ($T_1 = 10$)

$$X_i : \quad 1 \quad 3 \quad 5 \quad 2 \quad 4 \quad 7 \quad 1 \quad 3 \quad 3 \quad 8$$

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

$$NDR : \quad 1 \quad 3 \quad 5 \quad 2 \quad 4 \quad 7 \quad 1 \quad 3 \quad 3 \quad 8$$

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

NDR : $\underline{1\ 3\ 5}$ 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 : $\underline{1\ 3\ 5}$ $\underline{2\ 4\ 7}$ $\underline{1\ 3}$ $\underline{3\ 8}$ $M^I_{10} = 3$

NDR : $\underline{1\ 3\ 5}$ $\underline{2\ 4\ 7}$ 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 $(T_1 = 10)$

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

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

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

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 ($T_1 = 10$)

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

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

$$NDR:\ \underline{1\ \ 3\ \ 5}\ \ \underline{2\ \ 4\ \ 7}\ \ \underline{1\ \ 3\ \ 3\ \ 8} \qquad M^{ND}_{10} = 4$$

Outline

1. Introduction
   - Framework
   - Problem
2. Methodology
   - One dimensional discrete scan statistics
   - Longest increasing / non-decreasing run and scan statistics
   - Approximations for scan statistics
3. Comparison study
   - Novak's result
   - Numerical examples
4. References

Objective and related work

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]

Outline

1. Introduction
   - Framework
   - Problem
2. Methodology
   - One dimensional discrete scan statistics
   - Longest increasing / non-decreasing run and scan statistics
   - Approximations for scan statistics
3. Comparison study
   - Novak's result
   - Numerical examples
4. References

One dimensional discrete scan statistics

Introducing the Model

Let $1 \leq m_1 \leq T_1$ be positive integers, $X_1, X_2, \dots, X_{T_1}$ a sequence of i.i.d. r.v.'s. and $\mathcal{S} : \mathbb{R}^{m_1} \to \mathbb{R}$ a measurable real valued function.

Then, the one dimensional discrete scan statistics is defined as

$$\mathbf{S}(m_1, T_1, \mathcal{S}) = \max_{1 \leq i_1 \leq T_1 - m_1 + 1} \mathcal{S}\left(X_{i_1}, X_{i_1+1}, \dots, X_{i_1+m_1-1}\right).$$

Remark

If, in particular, we consider $\mathcal{S}(x_1, \dots, x_{m_1}) = x_1 + \dots + x_{m_1}$ then

$$S_{m_1}(T_1) := \mathbf{S}(m_1, T_1, \mathcal{S}) = \max_{1 \leq i_1 \leq T_1 - m_1 + 1} \sum_{i=i_1}^{i_1+m_1-1} X_i$$

is the *classical* one dimensional discrete scan statistics ([Glaz et al., 2001]).

Example ($T_1 = 26,\ m_1 = 6,\ X_i \sim \mathcal{B}(p),\ Y_{i_1} = X_{i_1} + \dots + X_{i_1+m_1-1},\ 1 \leq i_1 \leq 21$)

1	1	0	1	1	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	0	1	1	1	1	0

$Y_1 = 4$

Related statistics: classical model

Let $X_1, \dots, X_{T_1}$ be a sequence of i.i.d. $0-1$ Bernoulli of parameter p

- $W_{m_1,k}$ - the waiting time until we first observe at least k successes in a window of size m_1

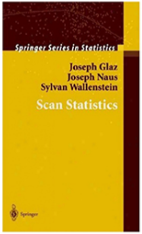
$$\mathbb{P}\left(W_{m_1,k} \leq T_1\right) = \mathbb{P}\left(S_{m_1}(T_1) \geq k\right)$$

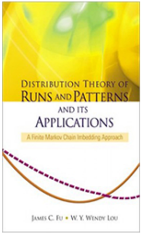
- $D_{T_1}(k)$ - the length of the smallest window that contains at least k successes

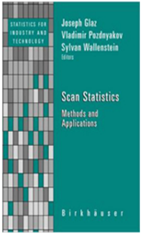
$$\mathbb{P}\left(D_{T_1}(k) \leq m_1\right) = \mathbb{P}\left(S_{m_1}(T_1) \geq k\right)$$

- L_{T_1} - the length of the longest success run

$$\mathbb{P}\left(L_{T_1} \geq m_1\right) = \mathbb{P}\left(S_{m_1}(T_1) \geq m_1\right) = \mathbb{P}\left(S_{m_1}(T_1) = m_1\right)$$







Outline

1. Introduction
   - Framework
   - Problem
2. Methodology
   - One dimensional discrete scan statistics
   - Longest increasing / non-decreasing run and scan statistics
   - Approximations for scan statistics
3. Comparison study
   - Novak's result
   - Numerical examples
4. References

Longest increasing / non-decreasing run and scan statistics

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 $X_1, \dots, X_{T_1}$ 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

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

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

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

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

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

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

We have, for $k \geq 1$

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

Outline

1 Introduction
- Framework
- Problem

2 Methodology
- One dimensional discrete scan statistics
- Longest increasing / non-decreasing run and scan statistics
- Approximations for scan statistics

3 Comparison study
- Novak's result
- Numerical examples

4 References

Scan statistics and 1-dependent sequences

$\mathbf{S}(m_1, T_1, \mathcal{S})$ viewed as maximum of 1-dependent r.v.'s

Let $L_1 = \frac{T_1}{m_1 - 1}$, be a positive integer

- Define for each $k_1 \in \{1, 2, \dots, L_1 - 1\}$ the random variables

$$Z_{k_1} = \max_{(k_1-1)(m_1-1)+1 \leq i_1 \leq k_1(m_1-1)} \mathcal{S}\left(X_{i_1}, \dots, X_{i_1+m_1-1}\right)$$

- $(Z_{k_1})_{k_1}$ is 1-dependent (i.e. $\sigma\left(\{Z_1, \dots, Z_h\}\right) \perp\!\!\!\perp \sigma\left(\{Z_{h+2}, \dots\}\right), \forall h \geq 1$) and stationary
- Observe

$$\mathbf{S}(m_1, T_1, \mathcal{S}) = \max_{1 \leq k_1 \leq L_1 - 1} Z_{k_1}$$

Illustration of 1-dependence

$$X_1, X_2, \dots, X_{m_1-1}, X_{m_1}, \dots, X_{2(m_1-1)}, X_{2m_1-1}, \dots, X_{3(m_1-1)}, X_{3m_1-2}, \dots, X_{4(m_1-1)}$$

$\mathbf{S}(m_1, T_1, \mathcal{S})$ viewed as maximum of 1-dependent r.v.'s

Let $L_1 = \frac{T_1}{m_1 - 1}$, be a positive integer

- Define for each $k_1 \in \{1, 2, \ldots, L_1 - 1\}$ the random variables

$$Z_{k_1} = \max_{(k_1-1)(m_1-1)+1 \leq i_1 \leq k_1(m_1-1)} \mathcal{S}\left(X_{i_1}, \ldots, X_{i_1+m_1-1}\right)$$

- $(Z_{k_1})_{k_1}$ is 1-dependent (i.e. $\sigma\left(\{Z_1, \ldots, Z_h\}\right) \perp\!\!\!\perp \sigma\left(\{Z_{h+2}, \ldots\}\right), \forall h \geq 1$) and stationary
- Observe

$$\mathbf{S}(m_1, T_1, \mathcal{S}) = \max_{1 \leq k_1 \leq L_1 - 1} Z_{k_1}$$

Illustration of 1-dependence

$$\underbrace{X_1, X_2, \ldots, X_{m_1-1}, X_{m_1}, \ldots, X_{2(m_1-1)}}_{Z_1}, X_{2m_1-1}, \ldots, X_{3(m_1-1)}, X_{3m_1-2}, \ldots, X_{4(m_1-1)}$$

$\mathbf{S}(m_1, T_1, \mathcal{S})$ viewed as maximum of 1-dependent r.v.'s

Let $L_1 = \frac{T_1}{m_1 - 1}$, be a positive integer

- Define for each $k_1 \in \{1, 2, \ldots, L_1 - 1\}$ the random variables

$$Z_{k_1} = \max_{(k_1-1)(m_1-1)+1 \leq i_1 \leq k_1(m_1-1)} \mathcal{S}\left(X_{i_1}, \ldots, X_{i_1+m_1-1}\right)$$

- $(Z_{k_1})_{k_1}$ is 1-dependent (i.e. $\sigma\left(\{Z_1, \ldots, Z_h\}\right) \perp\!\!\!\perp \sigma\left(\{Z_{h+2}, \ldots\}\right), \forall h \geq 1$) and stationary
- Observe

$$\mathbf{S}(m_1, T_1, \mathcal{S}) = \max_{1 \leq k_1 \leq L_1 - 1} Z_{k_1}$$

Illustration of 1-dependence

$$\underbrace{X_1, X_2, \ldots, X_{m_1-1}, \overbrace{X_{m_1}, \ldots, X_{2(m_1-1)}}}_{Z_1}\!\!{}^{}\overbrace{, X_{2m_1-1}, \ldots, X_{3(m_1-1)}}^{Z_2}, X_{3m_1-2}, \ldots, X_{4(m_1-1)}$$

$\mathbf{S}(m_1, T_1, \mathcal{S})$ viewed as maximum of 1-dependent r.v.'s

Let $L_1 = \frac{T_1}{m_1 - 1}$, be a positive integer

- Define for each $k_1 \in \{1, 2, \dots, L_1 - 1\}$ the random variables

$$Z_{k_1} = \max_{(k_1-1)(m_1-1)+1 \leq i_1 \leq k_1(m_1-1)} \mathcal{S}\left(X_{i_1}, \dots, X_{i_1+m_1-1}\right)$$

- $(Z_{k_1})_{k_1}$ is 1-dependent (i.e. $\sigma\left(\{Z_1, \dots, Z_h\}\right) \perp\!\!\!\perp \sigma\left(\{Z_{h+2}, \dots\}\right), \forall h \geq 1$) and stationary

- Observe

$$\mathbf{S}(m_1, T_1, \mathcal{S}) = \max_{1 \leq k_1 \leq L_1 - 1} Z_{k_1}$$

Illustration of 1-dependence

$$\underbrace{X_1, X_2, \dots, X_{m_1-1}, \overbrace{X_{m_1}, \dots, X_{2(m_1-1)}}_{Z_1}, X_{2m_1-1}, \dots, X_{3(m_1-1)}}^{Z_2}, X_{3m_1-2}, \dots, X_{4(m_1-1)}$$

$$Z_1 = \{X_1, \dots, X_{2(m_1-1)}\},\quad Z_2 = \{X_{m_1}, \dots, X_{3(m_1-1)}\},\quad Z_3 = \{X_{2m_1-1}, \dots, X_{4(m_1-1)}\}$$

Extremes of 1-dependent stationary sequences

Let $(Z_n)_{n\geq 1}$ be a 1-dependent stationary sequence of r.v.'s.

Notation

For $x < \sup\{u|\mathbb{P}(Z_1 \leq u) < 1\}$,

$$q_n = q_n(x) = \mathbb{P}(\max(Z_1, \dots, Z_n) \leq x)$$

Theorem [Amărioarei, 2012]

For x such that $\mathbb{P}(Z_1 > x) = 1 - q_1 < 0.1$ and $n > 3$ we have

$$\left| q_n - \frac{2q_1 - q_2}{[1 + q_1 - q_2 + 2(q_1 - q_2)^2]^n} \right| \leq nF(q_1, n)(1 - q_1)^2$$

- $F(q_1, n) = 1 + \frac{3}{n} + \left[K(1 - q_1) + \frac{\Gamma(1-q_1)}{n}\right](1 - q_1)$.

▸ Selected values for $K(\cdot)$ and $\Gamma(\cdot)$

Approximation and error bounds

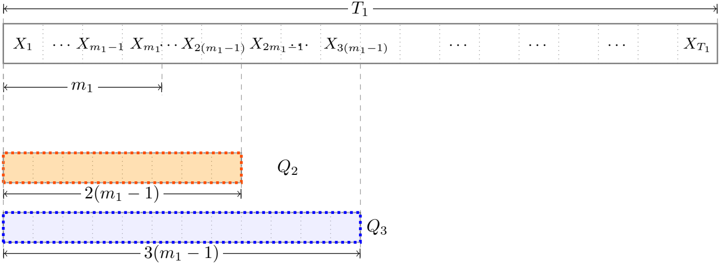
Theorem [Amărioarei, 2014]

Let $t_1 \in \{2,3\}$ and $Q_{t_1} = Q_{t_1}(\tau) = \mathbb{P}\left(\max_{1\le i_1 \le (t_1-1)(m_1-1)} \mathcal{S}\left(X_{i_1},\dots,X_{i_1+m_1-1}\right) \le \tau\right)$.

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

$$\left|\mathbb{P}\left(\mathbf{S}(m_1, T_1, \mathcal{S}) \le \tau\right) - \left(2\hat{Q}_2 - \hat{Q}_3\right)\left[1+\hat{Q}_2-\hat{Q}_3+2(\hat{Q}_2-\hat{Q}_3)^2\right]^{1-L_1}\right| \le E_{total}(1),$$

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



Outline

1 Introduction
- Framework
- Problem

2 Methodology
- One dimensional discrete scan statistics
- Longest increasing / non-decreasing run and scan statistics
- Approximations for scan statistics

3 Comparison study
- Novak's result
- Numerical examples

4 References

Another approach

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

Outline

1 Introduction
- Framework
- Problem

2 Methodology
- One dimensional discrete scan statistics
- Longest increasing / non-decreasing run and scan statistics
- Approximations for scan statistics

3 Comparison study
- Novak's result
- Numerical examples

4 References

Numerical results

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 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 $\tilde{X}_i = \mathbf{1}_{\{X_i \leq X_{i+1}\}}$. 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 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 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

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

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

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

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 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

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

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: $G = \mathcal{U}(\{1, \dots, s\})$

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 $Iter = 10^5$.

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 $Iter = 10^5$.

thank you!

Amărioarei, A. (2012).
Approximation for the distribution of extremes of one dependent stationary sequences of random variables.
arXiv:1211.5456v1, submitted.

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

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.

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

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.

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.

Selected Values for $K(\cdot)$ and $\Gamma(\cdot)$

Table 1 : Selected values for $K(\cdot)$ and $\Gamma(\cdot)$

$1-q_1$	$K(1-q_1)$	$\Gamma(1-q_1)$
0.1	38.63	480.69
0.05	21.28	180.53
0.025	17.56	145.20
0.01	15.92	131.43

Return

Selected Values for $K(\cdot)$ and $\Gamma(\cdot)$

Table 1 : Selected values for $K(\cdot)$ and $\Gamma(\cdot)$

$1 - q_1$	$K(1 - q_1)$	$\Gamma(1 - q_1)$
0.1	38.63	480.69
0.05	21.28	180.53
0.025	17.56	145.20
0.01	15.92	131.43

Return