Scan Statistics: Theory and Applications

Alexandru Amărioarei

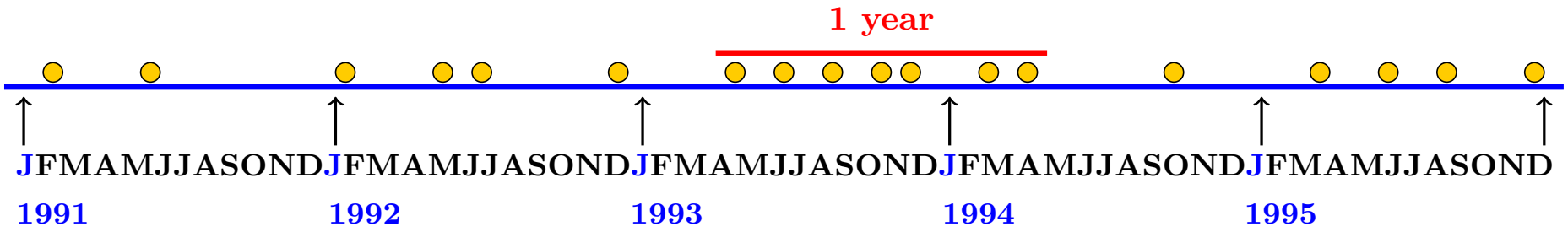
Laboratoire de Mathématiques Paul Painlevé
Département de Probabilités et Statistique
Université de Lille 1, INRIA Modal Team

Seminaire de Probabilité et Statistique
Laboratoire de Mathématiques Paul Painlevé
5 March, 2014, Lille

Outline

Outline

Epidemiology example



Observation of disease cases over time:
N = 19 cases over a period of $T = 5$ years

Observation

The epidemiologist notes a **one year period** (from April 93 - through April 94) with **8** cases: 42%!

Question

Given 19 cases over 5 years, how unusual is it to have a 1 year period containing as many as 8 cases?

The answer: A First approach

A First approach:

X = the number of cases falling in [April 93, April 94]

$X \sim Bin(19, 0.2)$

$$\mathbb{P}(X \geq 8) = 0.023$$

Conclusion: **atypical situation !**

But: it is **not** the answer to our question: the one year period is **not fixed** but identified after the scanning process !

The answer: Correct approach

The scan statistics:

$S =$ the maximum number of cases over **any continuous** one year period in $[0, T]$

Thus,

$$\mathbb{P}(S \geq 8) = 0.379$$

gives the answer to the epidemiologist question.

Conclusion: **no unusual situation !**

Outline

Introducing the Scan Statistics Model

Let T_1 be a positive integer and $X_1, X_2, \ldots, X_{T_1}$ a sequence of i.i.d. r.v.'s. For $m_1 \leq T_1$ integer, consider the moving sums

$$Y_{i_1} = \sum_{j=i_1}^{i_1+m_1-1} X_j$$

The discrete one dimensional scan statistics

$$S_{m_1}(T_1) = \max_{1 \leq i_1 \leq T_1 - m_1 + 1} Y_{i_1}.$$

Problem

Approximate the distribution of one dimensional scan statistic

$$\mathbb{P}\left(S_{m_1}(T_1) \leq k\right).$$

Used for testing the null hypotheses of randomness against the alternative hypothesis of clustering.

Related Statistics

Let $X_1, \ldots, X_{T_1}$ 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

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

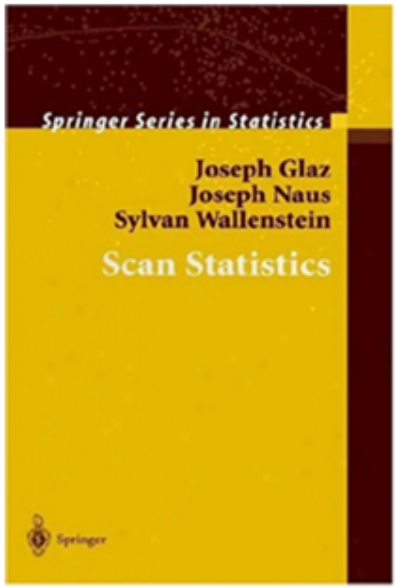
- L_{T_1} - the length of the longest success run

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

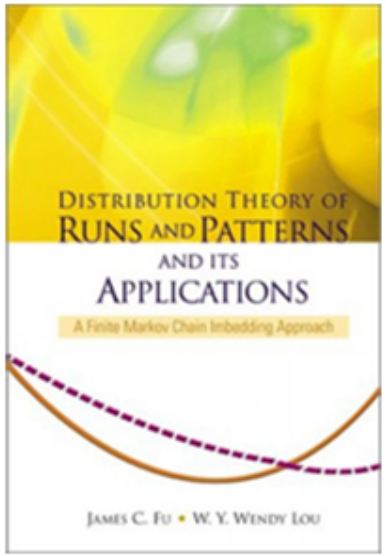
Literature


Joseph Glaz
N. Balakrishnan
Scan Statistics
and Applications
Birkhäuser


Runs and Scans
with Applications
N. Balakrishnan
Markos V. Koutras
WILEY SERIES IN
PROBABILITY AND STATISTICS


Springer Series in Statistics
Joseph Glaz
Joseph Naus
Sylvan Wallenstein
Scan Statistics
Springer


Joseph Glaz
Vladimir Pozdnyakov
Sylvan Wallenstein
Scan Statistics
Methods and
Applications
Birkhäuser


DISTRIBUTION THEORY OF
RUNS AND PATTERNS
AND ITS
APPLICATIONS
A Finite Markov Chain Imbedding Approach
JAMES C. FU • W. Y. WENDY LOU

Outline

Approach

Fu (2001) applied the Markov Chain Imbedding Technique to find the distribution of binary scan statistics.

Main Idea

Express the distribution of the $S_{m_1}(T_1)$ in terms of the waiting time distribution of a special compound pattern

- define for $0 \leq k \leq m_1$

$$\mathcal{F}_{m_1,k} = \{\Lambda_i | \Lambda_1 = \underbrace{1\ldots1}_{k}, \Lambda_2 = 10\underbrace{1\ldots1}_{k-1}, \ldots, \Lambda_l = \underbrace{1\ldots1}_{k-1}\overbrace{0\ldots01}^{m_1}\}$$

$$|\mathcal{F}_{m_1,k}| = \sum_{j=0}^{m_1-k_1} \binom{k-2+j}{j}$$

- the compound pattern $\Lambda = \cup_{i=1}^{l} \Lambda_i,\ \Lambda_i \in \mathcal{F}_{m_1,k}$
  - $\mathbb{P}(S_{m_1}(T_1) < k) = \mathbb{P}(W(\Lambda) \geq T_1 + 1).$
  - $\mathbb{P}(S_{m_1}(T_1) < k) = \boldsymbol{\xi} \boldsymbol{N}^{T_1} \mathbf{1}^{\intercal}$, where $\boldsymbol{\xi} = (1, 0, \ldots, 0)$

Example

Consider the i.i.d. two-state sequence $(X_i)_{i\in\{1,2,\dots,T_1\}}$ with $p = \mathbb{P}(X_1 = 1)$ and $q = \mathbb{P}(X_1 = 0)$.

- A realisation for $T_1 = 20$

  00101011101101010110

- For $k = 3$ and $m_1 = 4$

  $\mathcal{F}_{4,3} = \{\Lambda_1 = 111, \Lambda_2 = 1011, \Lambda_3 = 1101\}$

- The state space

  $\Omega = \{\emptyset, 0, 1, 10, 11, 101, 110, \alpha_1, \alpha_2, \alpha_3\}$

- the principal matrix:

$$N = \begin{pmatrix} 0 & q & p & 0 & 0 & 0 & 0 \\ 0 & q & p & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q & p & 0 & 0 \\ 0 & q & 0 & 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & q \\ 0 & 0 & 0 & q & 0 & 0 & 0 \\ 0 & q & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Outline

Naus Product Type Approximation

Considering that $T_1 = L_1 m_1$, Naus (1982) gave the following product type approximation

$$\mathbb{P}(S_{m_1}(T_1) \leq k) \approx Q_{2m_1} \left(\frac{Q_{3m_1}}{Q_{2m_1}} \right)^{\frac{T_1}{m_1}-2},$$

where $Q_{2m_1} = \mathbb{P}(S_{m_1}(2m_1) \leq k)$ and $Q_{3m_1} = \mathbb{P}(S_{m_1}(3m_1) \leq k)$.
Idea of proof:

$$\mathbb{P}(S_{m_1}(T_1) \leq k) = \mathbb{P}(E_1)\mathbb{P}(E_2|E_1)\cdots\mathbb{P}\left(E_{L-1}|\bigcap_{i=1}^{L-2} E_i\right),$$

where for $j \in \{1, 2, \ldots, L-1\}$,

$$E_j = \left\{ \max_{(j-1)m_1+1 \leq i_1 \leq jm_1} Y_{i_1} \leq k \right\},$$

and use the Markov-like approximation (exchangeability)

$$\mathbb{P}\left(E_l|\bigcap_{i=1}^{l-1} E_i\right) \approx \mathbb{P}(E_l|E_{l-1}) \approx \frac{\mathbb{P}(E_2 \cap E_1)}{\mathbb{P}(E_1)} = \frac{Q_{3m_1}}{Q_{2m_1}}.$$

Product Type Approximation for Bernoulli

For Bernoulli r.v.'s, Naus (1982) derived exact formulas for Q_{2m_1} and Q_{3m_1}:

$$Q_{2m_1} = F^2(k; m_1, p) - kb(k+1; m_1, p)F(k-1; m_1, p) + m_1 pb(k+1; m_1, p)F(k-2, m_1-1),$$
$$Q_{3m_1} = F^3(k; m_1, p) - A_1 + A_2 + A_3 - A_4,$$

where

$$A_1 = 2b(k+1; m_1, p)F(k; m_1, p)[kF(k-1; m_1, p) - m_1 pF(k-2; m_1-1, p)],$$
$$A_2 = 0.5b^2(k+1; m_1, p)\left[k(k-1)F(k-2; m_1, p) - 2(k-1)m_1 F(k-3; m_1-1, p) + m_1(m_1-1)p^2 F(k-4; m_1-2, p)\right],$$
$$A_3 = \sum_{r=1}^{k} b(2(k+1)-r; m_1, p)F^2(r-1; m_1, p),$$
$$A_4 = \sum_{r=2}^{k} b(2(k+1)-r; m_1, p)b(r+1; m_1, p)[rF(r-1; m_1, p) - m_1 pF(r-2; m_1-1, p)].$$

Product Type Approximation for Binomial and Poisson

If $X_i \sim Bin(n,p)$ or $X_i \sim Pois(\lambda)$, we have the approximation

$$\mathbb{P}(S_{m_1}(T_1) \leq k) \approx Q_{2m_1} \left(\frac{Q_{3m_1}}{Q_{2m_1}}\right)^{\frac{T_1}{m_1}-2}, \quad T_1 \geq 3m_1$$

$$\approx Q_{2m_1-1} \left(\frac{Q_{2m_1}}{Q_{2m_1-1}}\right)^{T_1-2m_1+1}, \quad T_1 \geq 2m_1$$

where Q_{2m_1-1}, Q_{2m_1} and Q_{3m_1} are computed by Karwe and Naus (1996) recurrence.

▸ Karwe Naus algorithm for Q_{2m_1-1} and Q_{2m_1}

Bounds

Glaz and Naus (1991) developed a variety of tight bounds:

- Lower Bounds

$$\begin{aligned}\mathbb{P}(S_{m_1}(T_1) \leq k) &\leq \frac{Q_{2m_1}}{\left[1 + \frac{Q_{2m_1-1} - Q_{2m_1}}{Q_{2m_1-1}\, Q_{2m_1}}\right]^{T_1 - 2m_1}}, \quad T_1 \geq 2m_1 \\ &\leq \frac{Q_{3m_1}}{\left[1 + \frac{Q_{2m_1-1} - Q_{2m_1}}{Q_{3m_1-1}}\right]^{T_1 - 3m_1}}, \quad T_1 \geq 3m_1\end{aligned}$$

- Upper Bounds

$$\begin{aligned}\mathbb{P}(S_{m_1}(T_1) \leq k) &\leq Q_{2m_1}\left[1 - Q_{2m_1-1} + Q_{2m_1}\right]^{T_1 - 2m_1}, \quad T_1 \geq 2m_1 \\ &\leq Q_{3m_1}\left[1 - Q_{2m_1-1} + Q_{2m_1}\right]^{T_1 - 3m_1}, \quad T_1 \geq 3m_1\end{aligned}$$

The values $Q_{2m_1-1}, Q_{2m_1}, Q_{3m_1-1}, Q_{3m_1}$ are computed using Karwe and Naus (1996) algorithm.

Haiman Type Approximation: Main Observation

Haiman proposed in 2000 a different approach

Main Observation

The scan statistic r.v. can be viewed as a maximum of a sequence of 1-dependent stationary r.v..

- The idea:
  - discrete and continuous one dimensional scan statistic: Haiman (2000,2007)
  - discrete and continuous two dimensional scan statistic: Haiman and Preda (2002,2006)
  - discrete three dimensional scan statistic: Amărioarei and Preda (2013)

Haiman Type Approximation: Writing the Scan as an Extreme of 1-Dependent R.V.'s

Let $T_1 = L_1 m_1$ positive integer

- Define for $j \in \{1, 2, \ldots, L_1 - 1\}$

$$Z_j = \max_{(j-1)m_1+1 \le i_1 \le jm_1+1} Y_{i_1}$$

- $(Z_j)_j$ is 1-dependent and stationary

$$\underbrace{X_1, X_2, \ldots, X_{m_1}}_{Z_1}, \overbrace{\underbrace{X_{m_1+1}, \ldots, X_{2m_1}}_{Z_2}, \underbrace{X_{2m_1+1}, \ldots, X_{3m_1}}_{Z_3}}, \underbrace{X_{3m_1+1}, \ldots, X_{4m_1}}_{Z_4}$$

- Observe

$$S_{m_1}(T_1) = \max_{1 \le j \le L_1 - 1} Z_j$$

Haiman Type Approximation: Main Tool

Let $(Z_j)_{j\geq 1}$ be a strictly stationary 1-dependent sequence of r.v.'s and let $q_m = q_m(x) = \mathbb{P}(\max(Z_1, \ldots, Z_m) \leq x)$, with $x < \sup\{u | \mathbb{P}(Z_1 \leq u) < 1\}$.

Theorem (Haiman 1999)

For any x such that $\mathbb{P}(Z_1 > x) = 1 - q_1 \leq 0.025$ and any integer $m > 3$,

$$\left| q_m - \frac{6(q_1 - q_2)^2 + 4q_3 - 3q_4}{(1 + q_1 - q_2 + q_3 - q_4 + 2q_1^2 + 3q_2^2 - 5q_1 q_2)^m} \right| \leq \Delta_1^H (1 - q_1)^3,$$

$$\left| q_m - \frac{2q_1 - q_2}{[1 + q_1 - q_2 + 2(q_1 - q_2)^2]^m} \right| \leq \Delta_2^H (1 - q_1)^2,$$

- $\Delta_1^H = 561 + 88m[1 + 124m(1 - q_1)^3]$
- $\Delta_1^H = 9 + 561(1 - q_1) + 3.3m[1 + 4.7m(1 - q_1)^2]$.

Haiman Type Approximation: Improvement

Main Theorem (Amarioarei 2012)

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

$$\left| q_m - \frac{6(q_1 - q_2)^2 + 4q_3 - 3q_4}{(1 + q_1 - q_2 + q_3 - q_4 + 2q_1^2 + 3q_2^2 - 5q_1 q_2)^m} \right| \leq \Delta_1 (1 - q_1)^3,$$

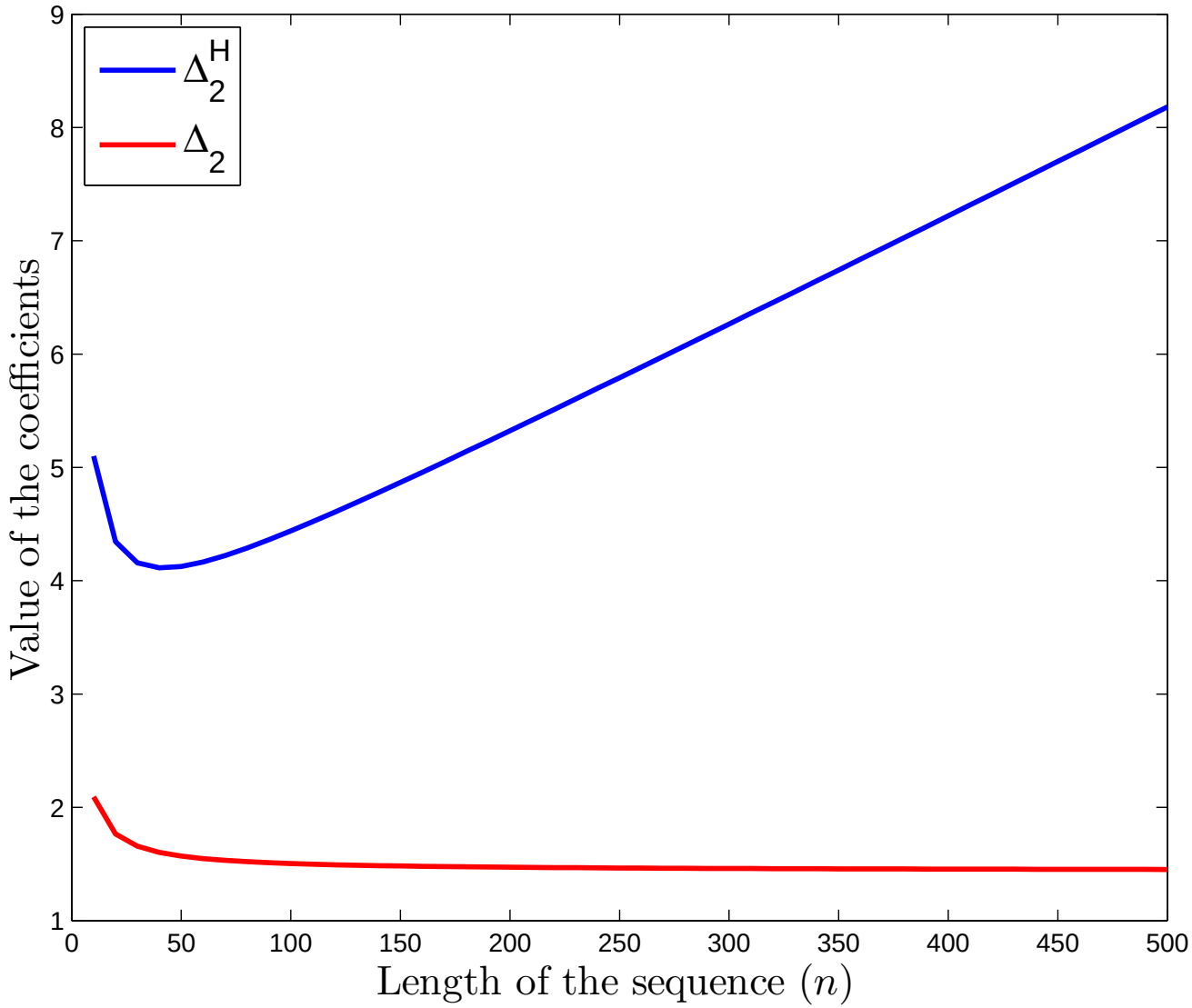
$$\left| q_m - \frac{2q_1 - q_2}{[1 + q_1 - q_2 + 2(q_1 - q_2)^2]^m} \right| \leq \Delta_2 (1 - q_1)^2,$$

- $\Delta_1 = \Delta_1(\alpha, q_1, m) = \Gamma(\alpha) + mK(\alpha)$
- $\Delta_2 = mE(\alpha, q_1, m) = m\left[1 + \frac{3}{m} + K(\alpha)(1 - q_1) + \frac{\Gamma(\alpha)(1-q_1)}{m}\right].$

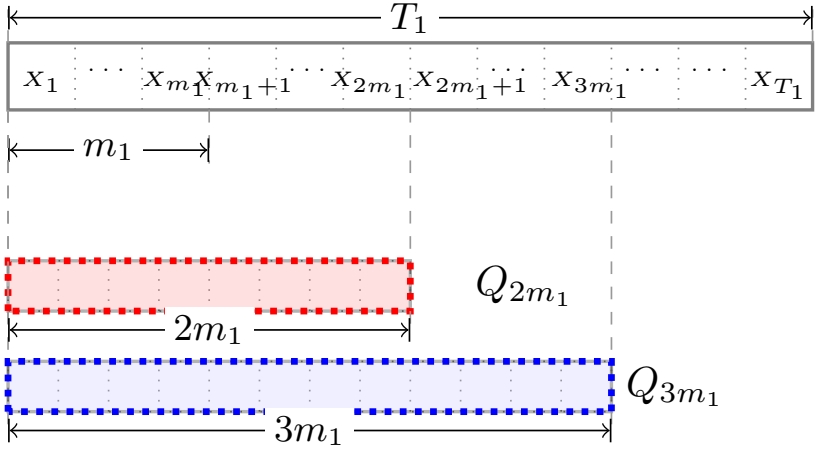
- Increased range of applicability
- Sharp bounds values (ex. $\alpha = 0.025$: $561 \to 145$ and $88 \to 17.5$)

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

Difference between the results: $1 - q_1 = 0.025$



Haiman Type Approximation: Result



- Observe that
$$Q_{2m_1} = \mathbb{P}(Z_1 \leq k)$$
$$Q_{3m_1} = \mathbb{P}(Z_1 \leq k, Z_2 \leq k)$$

- If $1 - Q_{2m_1} \leq 0.1$ the approximation

$$\mathbb{P}(S \leq k) \approx \frac{2Q_{2m_1} - Q_{3m_1}}{\left[1 + Q_{2m_1} - Q_{3m_1} + 2\left(Q_{2m_1} - Q_{3m_1}\right)^2\right]^{L_1 - 1}}$$

where $S = S_{m_1}(T_1)$

- Approximation error, about

$$(L_1 - 1)E(\alpha, L_1 - 1)\,(1 - Q_{2m_1})^2$$

with $E(\alpha, \alpha, m) = E(\alpha, m)$.

- Q_{2m_1} and Q_{3m_1} are evaluated using the previous approaches

Outline

Comparison between methods

Table 1 : $n = 1, p = 0.005, m_1 = 10, T_1 = 1000, It_{App} = 10^4$

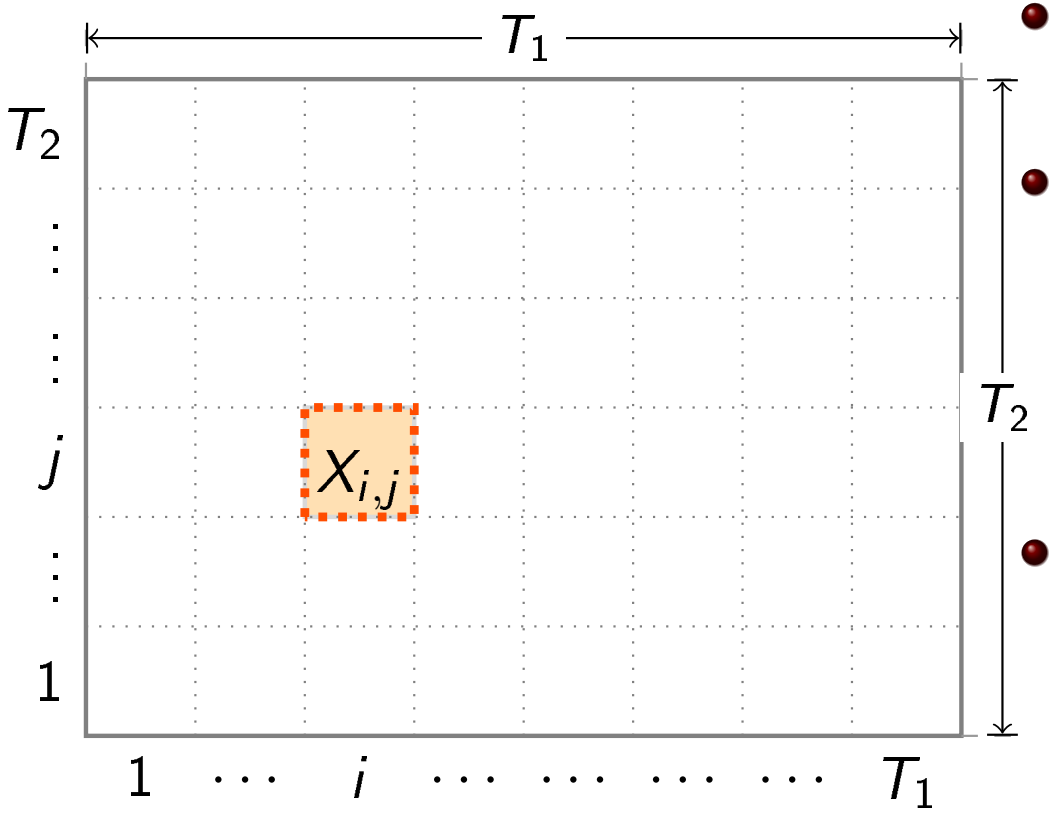
k	Exact	Glaz and Naus Product type	Haiman Approximation	Approximation Error	Lower Bound	Upper Bound
1	0.810209	0.810216	0.810404	0.001111	0.809903	0.810439
2	0.995764	0.995764	0.995764	3×10^{-7}	0.995764	0.995764
3	0.999950	0.999950	0.999950	4×10^{-11}	0.999950	0.999950

Table 2 : $n = 5, p = 0.05, m_1 = 25, T_1 = 500, It_{App} = 10^4, It_{Sim} = 10^3$

k	$\hat{\mathbb{P}}(S \leq k)$	Glaz and Naus Product type	Haiman Approximation	Total Error	Lower Bound	Upper Bound
13	0.712750	0.705787	0.714699	0.039308	0.697431	0.706948
14	0.867498	0.862184	0.865029	0.012502	0.859543	0.862407
15	0.946912	0.943329	0.946177	0.004169	0.942552	0.943362
16	0.980230	0.978959	0.979822	0.001354	0.978733	0.978963
17	0.993486	0.992821	0.993134	0.000433	0.992756	0.992822
18	0.997802	0.997726	0.997849	0.000127	0.997708	0.997726
19	0.999362	0.999327	0.999358	3×10^{-5}	0.999322	0.999327
20	0.999819	0.999813	0.999825	9×10^{-6}	0.999812	0.999813
21	0.999954	0.999951	0.999953	2×10^{-6}	0.999951	0.999951

Outline

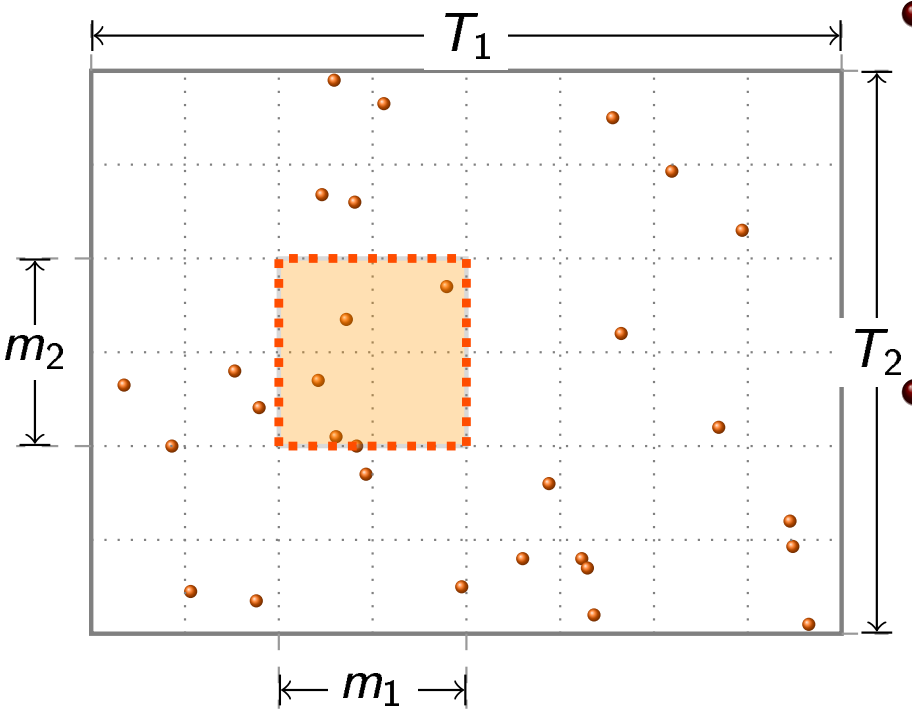
Introducing the Model



Let T_1, T_2 be positive integers

- Rectangular region $\mathcal{R} = [0, T_1] \times [0, T_2]$
- $(X_{ij})_{\substack{1 \le i \le T_1 \\ 1 \le j \le T_2}}$ i.i.d. integer r.v.'s
  - Bernoulli($\mathcal{B}(1, p)$)
  - Binomial($\mathcal{B}(n, p)$)
  - Poisson($\mathcal{P}(\lambda)$)
- X_{ij} number of observed events in the elementary subregion $r_{ij} = [i-1, i] \times [j-1, j]$

Defining the Scan Statistic



Let m_1, m_2 be positive integers

- Define for $1 \leq i_j \leq T_j - m_j + 1$,

$$Y_{i_1 i_2} = \sum_{i=i_1}^{i_1+m_1-1} \sum_{j=i_2}^{i_2+m_2-1} X_{ij}$$

- The two dimensional scan statistic,

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

- Used for testing the null hypotheses of randomness against the alternative hypothesis of clustering

Outline

Product Type Approximation Bernoulli Case

Boutsikas and Koutras (2000) using Markov Chain Imbedding approach proposed the approximation

$$\mathbb{P}\left(S_{m_1,m_2}(T_1,T_2)\leq k\right)\approx \frac{Q(m_1,m_2)^{(T_1-m_1-1)(T_2-m_2-1)}\,Q(m_1+1,m_2+1)^{(T_1-m_1)(T_2-m_2)}}{Q(m_1,m_2+1)^{(T_1-m_1-1)(T_2-m_2)}\,Q(m_1+1,m_2)^{(T_1-m_1)(T_2-m_2-1)}}$$

Where,

$$Q(m_1,m_2)=F\left(k;m_1m_2,p\right)$$

$$Q(m_1+1,m_2)=\sum_{s=0}^{k}F^2\left(k-s;m_2,p\right)b\left(s;(m_1-1)m_2,p\right)$$

$$Q(m_1+1,m_2+1)=\sum_{s_1,s_2=0}^{k}\sum_{t_1,t_2=0}^{k}\sum_{i_1,i_2,i_3,i_4=0}^{1} b(s_1;m_1-1,p)b(s_2;m_1-1,p)b(t_1;m_2-1,p)\times$$

$$b(t_2;m_2-1,p)p^{\sum i_j}(1-p)^{4-\sum i_j}F(u;(m_1-1)(m_2-1),p)$$

$$u=\min\left\{k-s_1-t_1-i_1,k-s_2-t_1-i_2,k-s_1-t_2-i_3,k-s_2-t_2-i_4\right\}$$

$$b(s;n,p)=\binom{n}{s}p^s(1-p)^{n-s}$$

$$F(s;n,p)=\sum_{i=0}^{s}b(i;n,p)$$

Bounds for the Bernoulli Case

The following bounds were established by Boutsikas and Koutras (2003)

- Lower Bound

$$LB = (1-Q_1)^{(T_1-m_1)(T_2-m_2)}\,(1-Q_2)^{T_1-m_1}\,(1-Q_3)^{T_2-m_2}\,(1-Q_4)$$

- Upper Bound

$$\begin{aligned} UB = {} & (1-Q_1)\left(1-q^{(m_1-1)(3m_2-2)+(2m_1-1)(m_2-1)}Q_1\right)^{(T_1-m_1-1)(T_2-m_2-1)}\left(1-q^{m_1(m_2-1)}Q_1\right)^{T_2-m_2-1} \\ & \times\left(1-q^{(m_1-1)(2m_2-1)+(m_1-1)(m_2-1)}Q_1\right)^{T_1-m_1-1}\left(1-q^{(m_1-1)(2m_2-1)+m_1(m_2-1)}Q_2\right)^{T_1-m_1} \\ & \times\left(1-q^{(m_1-1)(3m_2-2)+m_1(m_2-1)+(m_1-1)(m_2-1)}Q_3\right)^{T_2-m_2}\left(1-q^{(m_1-1)(2m_2-1)+m_1(m_2-1)}Q_4\right). \end{aligned}$$

Where $X_{ij} \sim B(p)$, $q = 1-p$ and

$$\begin{aligned} Q_1 &= F^c_{k+1,m_1m_2} - q^{m_2}F^c_{k+1,(m_1-1)m_2} - q^{m_1}F^c_{k+1,m_1(m_2-1)} + q^{m_1+m_2-1}F^c_{k+1,(m_1-1)(m_2-1)}, \\ Q_2 &= F^c_{k+1,m_1m_2} - q^{m_2}F^c_{k+1,(m_1-1)m_2}, \quad Q_3 = F^c_{k+1,m_1m_2} - q^{m_1}F^c_{k+1,m_1(m_2-1)}, \quad Q_4 = F^c_{k+1,m_1m_2} \\ F^c_{i,m} &= 1 - F(i-1;m,p). \end{aligned}$$

Product Type Approximation Binomial and Poisson

For $X_{ij} \sim Bin(n, p)$ or $X_{ij} \sim Pois(\lambda)$, Chen and Glaz (2009) proposed the product type approximation

$$\mathbb{P}(S_{m_1,m_2}(T_1, T_2) \leq k) \approx \frac{Q(m_1+1,m_2+1)^{(T_1-m_1)(T_2-m_2)}}{Q(m_1+1,m_2)^{(T_1-m_1)(T_2-m_2-1)}} \times \frac{Q(m_1,2m_2-1)^{(T_1-m_1-1)(T_2-2m_2)}}{Q(m_1,2m_2)^{(T_1-m_1-1)(T_2-2m_2+1)}}$$

Where,

$$Q(m_1, 2m_2 - 1) = \mathbb{P}(S_{m_1,m_2}(m_1, 2m_2 - 1) \leq k)$$
$$Q(m_1, 2m_2) = \mathbb{P}(S_{m_1,m_2}(m_1, 2m_2) \leq k)$$

To compute the unknown variables we use

- $Q(m_1, 2m_2 - 1)$ and $Q(m_1, 2m_2)$ - adaptation of Karwe and Naus algorithm
- $Q(m_1 + 1, m_2)$ and $Q(m_1 + 1, m_2 + 1)$ - conditioning

▸ Formulas for $Q(m_1 + 1, m_2)$ and $Q(m_1 + 1, m_2 + 1)$

Lower Bound for Binomial and Poisson

Glaz and Chen (1996) gave a lower bound applying Hoover Bonferroni type inequality of order $r \geq 3$,

$$\mathbb{P}\left(S_{m_1,m_2}(T_1,T_2) \geq k\right) = \mathbb{P}\left(\bigcup_{i_1=1}^{T_1-m_1+1}\bigcup_{i_2=1}^{T_2-m_2+1} A_{i_1,i_2}\right) \leq \sum_{i_1=1}^{T_1-m_1+1}\sum_{i_2=1}^{T_2-m_2+1}\mathbb{P}\left(A_{i_1,i_2}\right)$$
$$- \sum_{i_1=1}^{T_1-m_1+1}\sum_{i_2=1}^{T_2-m_2}\mathbb{P}\left(A_{i_1,i_2}\cap A_{i_1,i_2+1}\right) - \sum_{i_1=1}^{T_1-m_1}\mathbb{P}\left(A_{i_1,1}\cap A_{i_1+1,1}\right)$$
$$- \sum_{i_1=1}^{T_1-m_1+1}\sum_{l=2}^{r-1}\sum_{i_2=1}^{T_2-m_2+1-l}\mathbb{P}\left(A_{i_1,i_2}\cap A^c_{i_1,i_2+1}\cdots A^c_{i_1,i_2+l-1}\cap A_{i_1,i_2+l}\right)$$

Where $A_{i_1,i_2} = \{Y_{i_1,i_2} \geq k\}$ and for $r = 4$,

- Lower Bound

$$\mathbb{P}\left(S_{m_1,m_2}(T_1,T_2) \leq k\right) \geq (T_1-m_1)(Q(m_1+1,m_2)-2Q(m_1,m_2)) - (T_1-m_1+1)(T_2-m_2-3)$$
$$\times Q(m_1,m_2+2) + (T_1-m_1+1)(T_2-m_2-2)Q(m_1,m_2+3).$$

- $Q(m_1+1,m_2)$, $Q(m_1,m_2)$, $Q(m_1,m_2+2)$, $Q(m_1,m_2+3)$ - Karwe and Naus algorithm (variant)

Upper Bound for Binomial and Poisson

For the upper bound we adapt the inequality of Kuai et al. (2000) to the two dimensional framework:

$$\mathbb{P}\left(S_{m_1,m_2}\ (T_1,T_2) \le k\right) = 1 - \mathbb{P}\left(\bigcup_{i_1=1}^{T_1-m_1+1}\ \bigcup_{i_2=1}^{T_2-m_2+1} A_{i_1,i_2}\right)$$

$$\le 1 - \sum_{i_1=1}^{T_1-m_1+1}\ \sum_{i_2=1}^{T_2-m_2+1}\left[\frac{\theta_{i_1,i_2}\,\mathbb{P}(A_{i_1,i_2})^2}{\Sigma(i_1,i_2)+(1-\theta_{i_1,i_2})\mathbb{P}(A_{i_1,i_2})} + \frac{(1-\theta_{i_1,i_2})\mathbb{P}(A_{i_1,i_2})^2}{\Sigma(i_1,i_2)-\theta_{i_1,i_2}\,\mathbb{P}(A_{i_1,i_2})}\right]$$

where

$$\Sigma(i_1,i_2) = \sum_{j_1=1}^{T_1-m_1+1}\ \sum_{j_2=1}^{T_2-m_2+1} \mathbb{P}\left(A_{i_1,i_2}\cap A_{j_1,j_2}\right) \text{ and } \theta_{i_1,i_2} = \frac{\Sigma(i_1,i_2)}{\mathbb{P}(A_{i_1,i_2})} - \left\lfloor \frac{\Sigma(i_1,i_2)}{\mathbb{P}(A_{i_1,i_2})}\right\rfloor.$$

We have

$$\mathbb{P}\left(A_{i_1,i_2}\cap A_{j_1,j_2}\right) = \begin{cases} [1-Q(m_1,m_2)]^2\text{, if } |i_1-j_1|\ge m_1 \text{ or } |i_2-j_2|\ge m_2, \\ 1-2Q(m_1,m_2)+\mathbb{P}\left(Y_{i_1,i_2}\le n, Y_{j_1,j_2}\le n\right)\text{, otherwise}\end{cases}$$

$$\mathbb{P}\left(Y_{i_1,i_2}\le n, Y_{j_1,j_2}\le n\right) = \sum_{k=0}^{n}\mathbb{P}(Z=k)\mathbb{P}(Y_{i_1,i_2}-Z\le n-k)^2,$$

$$Z = \sum_{s=i_1\vee j_1}^{(i_1+m_1-1)\wedge(j_1+m_1-1)}\ \sum_{t=i_2\vee j_2}^{(i_2+m_2-1)\wedge(j_2+m_2-1)} X_{s,t}.$$

Haiman Type Approximation and Error Bounds: Writing the Scan as an Extreme of 1-Dependent R.V.'s

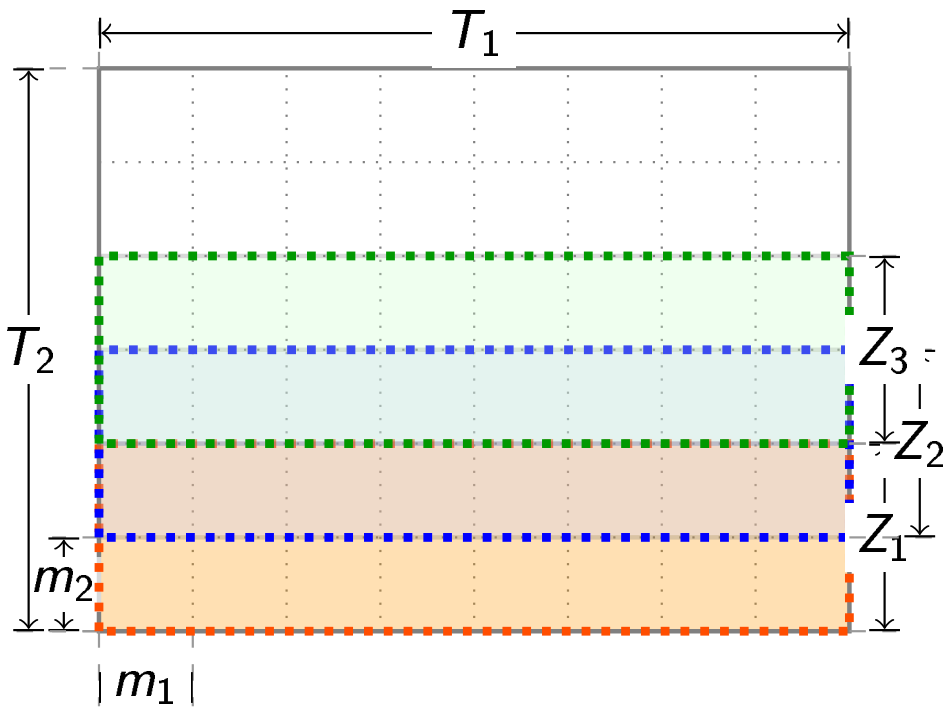
Let $L_j = \frac{T_j}{m_j}$, $j \in \{1, 2\}$ positive integers

- Define for $l \in \{1, 2, \ldots, L_2 - 1\}$

$$Z_l = \max_{\substack{1 \le i_1 \le (L_1-1)m_1+1 \\ (l-1)m_2+1 \le i_2 \le lm_2+1}} Y_{i_1 i_2}$$

- $(Z_l)_l$ is 1-dependent and stationary
- Observe

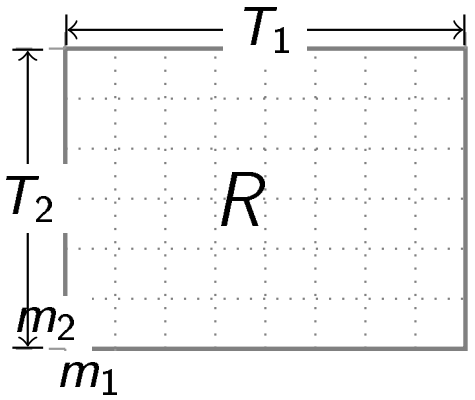
$$S_{m_1,m_2}(T_1, T_2) = \max_{1 \le k \le L_2-1} Z_k$$



Haiman Type Approximation and Error Bounds: First Step Approximation

Using Main Theorem we obtain

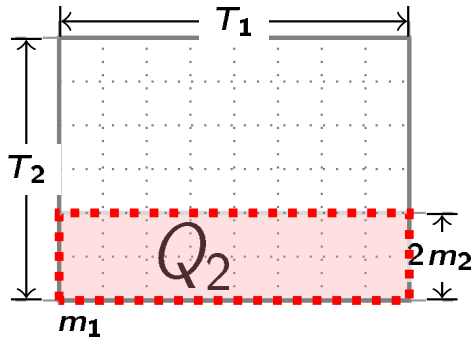
- Define
$$Q_2 = \mathbb{P}(Z_1 \leq k)$$
$$Q_3 = \mathbb{P}(Z_1 \leq k, Z_2 \leq k)$$

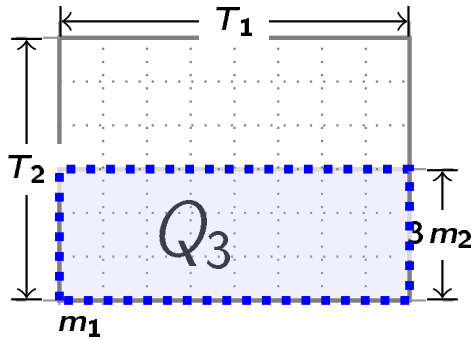


- If $1 - Q_2 \leq \alpha_1 < 0.1$ the (first) approximation

$$\mathbb{P}(S \leq k) \approx \frac{2Q_2 - Q_3}{[1 + Q_2 - Q_3 + 2(Q_2 - Q_3)^2]^{L_2 - 1}}$$

where $S = S_{m_1, m_2}(T_1, T_2)$





- Approximation error

$$(L_2 - 1)E(\alpha_1, L_2 - 1)(1 - Q_2)^2$$

Haiman Type Approximation and Error Bounds: Second Step Approximation

Q_2 :

- For $s \in \{1, 2, \ldots, L_1 - 1\}$

$$Z_s^{(2)} = \max_{\substack{(s-1)m_1+1 \le i_1 \le sm_1+1 \\ 1 \le i_2 \le m_2+1}} Y_{i_1 i_2}$$

- $$Q_2 = \mathbb{P}\left(\max_{1 \le s \le L_1 - 1} Z_s^{(2)} \le k \right)$$

- Define

$$Q_{22} = \mathbb{P}(Z_1^{(2)} \le k)$$

$$Q_{32} = \mathbb{P}(Z_1^{(2)} \le k, Z_2^{(2)} \le k)$$

- Approximation $(1 - Q_{22} \le \alpha_2)$

$$Q_2 \approx \frac{2Q_{22} - Q_{32}}{[1 + Q_{22} - Q_{32} + 2(Q_{22} - Q_{32})^2]^{L_1 - 1}}$$

- $$(L_1 - 1)E(\alpha_2, L_1 - 1)(1 - Q_{22})^2$$

Q_3 :

- For $s \in \{1, 2, \ldots, L_1 - 1\}$

$$Z_s^{(3)} = \max_{\substack{(s-1)m_1+1 \le i_1 \le sm_1+1 \\ 1 \le i_2 \le 2m_2+1}} Y_{i_1 i_2}$$

- $$Q_3 = \mathbb{P}\left(\max_{1 \le l \le L_1 - 1} Z_s^{(3)} \le k \right)$$

- Define

$$Q_{23} = \mathbb{P}(Z_1^{(3)} \le k)$$

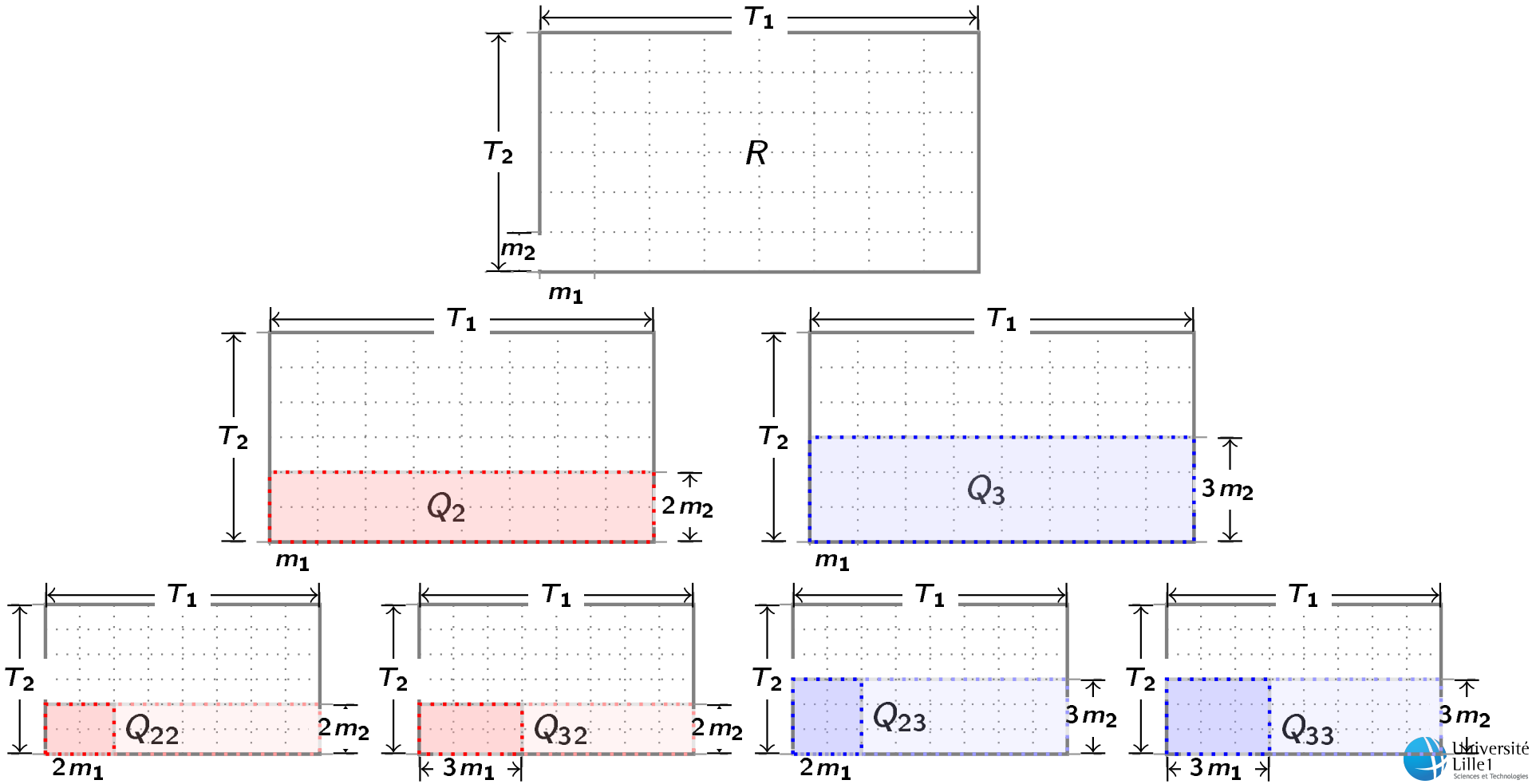
$$Q_{33} = \mathbb{P}(Z_1^{(3)} \le k, Z_2^{(3)} \le k)$$

- Approximation $(1 - Q_{23} \le \alpha_2)$

$$Q_3 \approx \frac{2Q_{23} - Q_{33}}{[1 + Q_{23} - Q_{33} + 2(Q_{23} - Q_{33})^2]^{L_1 - 1}}$$

- $$(L_1 - 1)E(\alpha_2, L_1 - 1)(1 - Q_{23})^2$$

Haiman Type Approximation and Error Bounds: Illustration of the Approximation Process



Outline

Comparison between methods

Table 3 : $n = 1, p = 0.005, m_1 = m_2 = 6, T_1 = T_2 = 30, It_{App} = 10^3, It_{Sim} = 10^3$

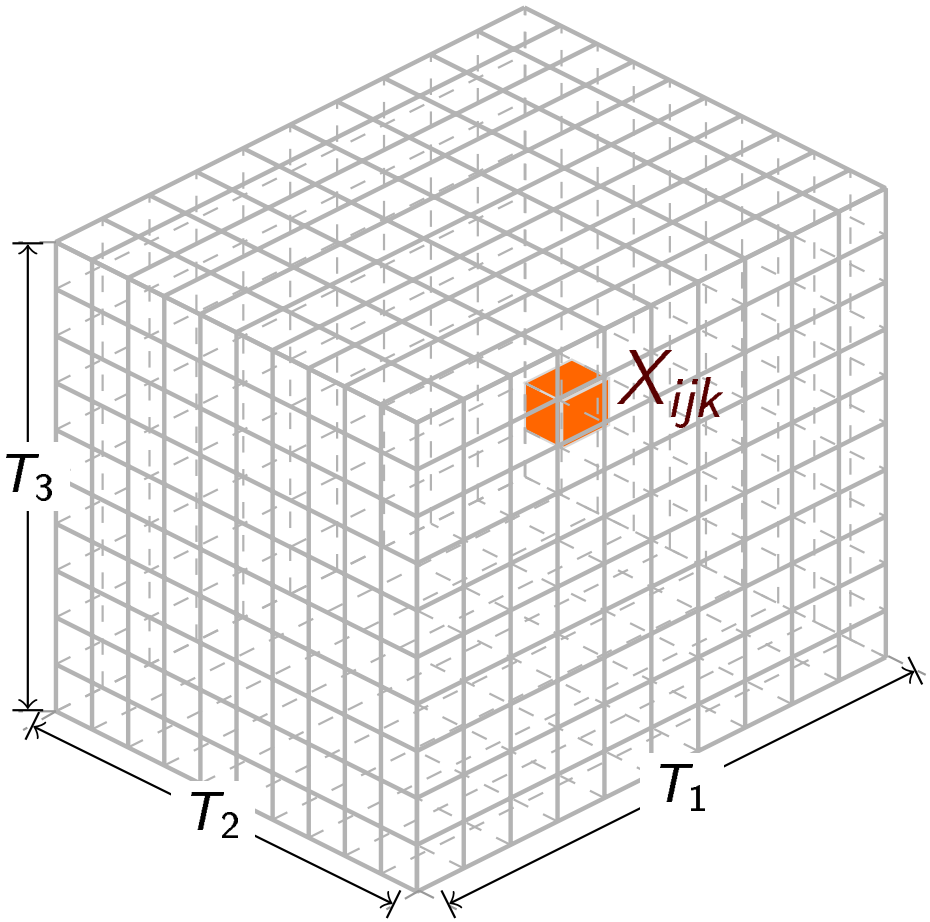
k	$\hat{\mathbb{P}}(S \leq k)$	Glaz and Naus Product type	Haiman Approximation	Total Error(App+Sim)	Lower Bound	Upper Bound
2	0.915903	0.914013	0.920211	0.041483	0.901935	0.945623
3	0.994292	0.994395	0.994578	0.000803	0.993785	0.996638
4	0.999747	0.999757	0.999760	2×10^{-5}	0.999737	0.999858
5	0.999992	0.999992	0.999992	7×10^{-7}	0.999992	0.999995

Table 4 : $n = 5, p = 0.002, m_1 = 5, m_2 = 10, T_1 = 50, T_2 = 80, It_{App} = 10^4, It_{Sim} = 10^3$

k	$\hat{\mathbb{P}}(S \leq k)$	Glaz and Naus Product type	Haiman Approximation	Total Error(App+Sim)	Lower Bound	Upper Bound
4	0.894654	0.873256	0.893724	0.037136	0.803422	0.944318
5	0.988003	0.986249	0.988144	0.002125	0.981418	0.993451
6	0.998963	0.998847	0.998963	0.000152	0.998543	0.999401
7	0.999926	0.999919	0.999925	9×10^{-6}	0.999903	0.999955
8	0.999995	0.999995	0.999995	5×10^{-7}	0.999994	0.999997

Outline

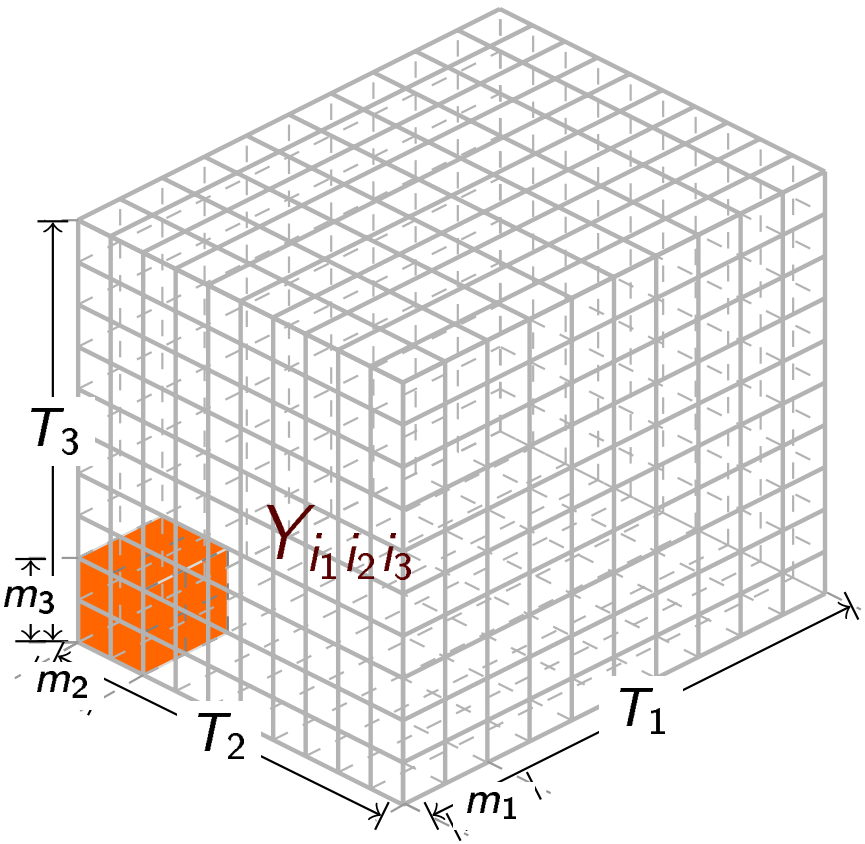
Introducing the Model



Let T_1, T_2, T_3 be positive integers

- Rectangular region
  $\mathcal{R} = [0, T_1] \times [0, T_2] \times [0, T_3]$
- $(X_{ijk})_{\substack{1 \le i \le T_1 \\ 1 \le j \le T_2 \\ 1 \le k \le T_3}}$ i.i.d. integer r.v.'s
  - Bernoulli($\mathcal{B}(1, p)$)
  - Binomial($\mathcal{B}(n, p)$)
  - Poisson($\mathcal{P}(\lambda)$)
- X_{ijk} number of observed events in the elementary subregion
  $r_{ijk} = [i-1, i] \times [j-1, j] \times [k-1, k]$

Defining the Scan Statistic



Let m_1, m_2, m_3 be positive integers

- Define for $1 \leq i_j \leq T_j - m_j + 1$,

$$Y_{i_1 i_2 i_3} = \sum_{i=i_1}^{i_1+m_1-1} \sum_{j=i_2}^{i_2+m_2-1} \sum_{k=i_3}^{i_3+m_3-1} X_{ijk}$$

- The three dimensional scan statistic,

$$S_{m_1,m_2,m_3} = \max_{\substack{1 \leq i_1 \leq T_1-m_1+1 \\ 1 \leq i_2 \leq T_2-m_2+1 \\ 1 \leq i_3 \leq T_3-m_3+1}} Y_{i_1 i_2 i_3}.$$

- Used for testing the null hypotheses of randomness against the alternative hypothesis of clustering

Product Type Approximation

Glaz et al. proposed in the product type approximation

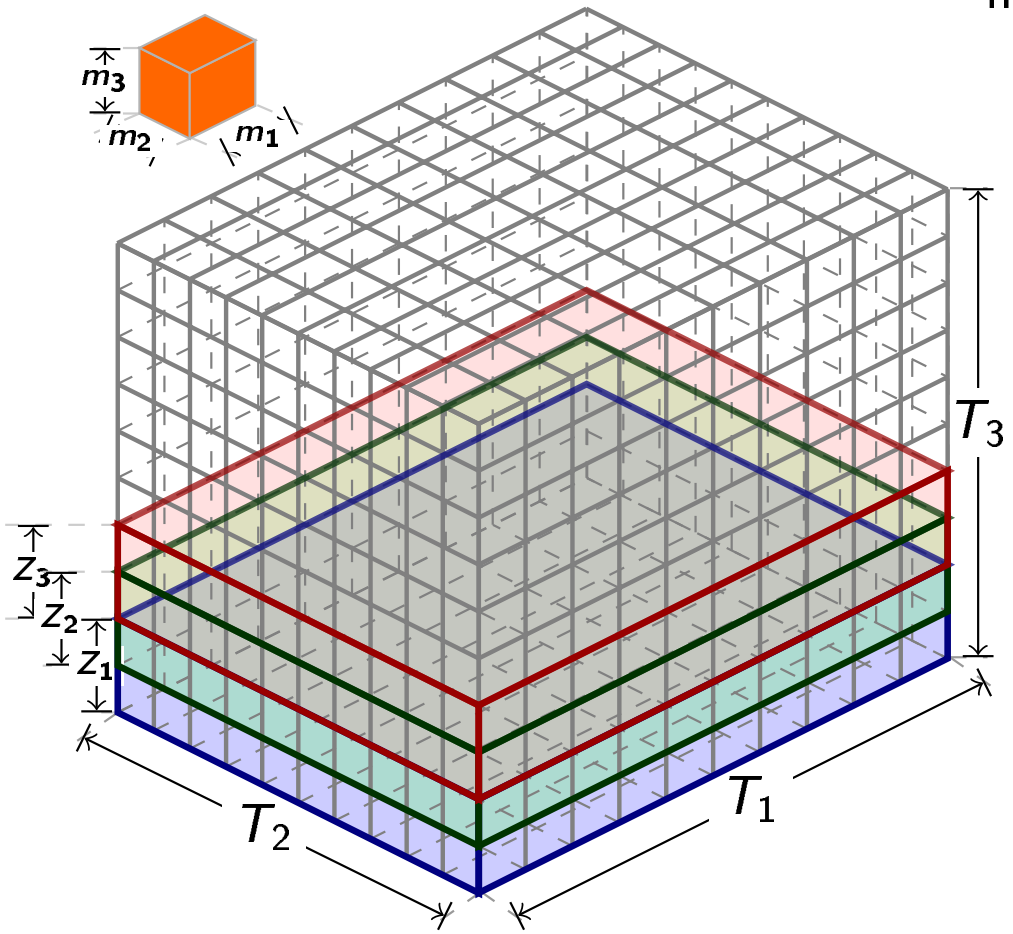
$$\mathbb{P}\left(S_{m_1,m_2,m_3} \leq k\right) \approx$$
$$\frac{Q(m_1+1,m_2+1,m_3+1)^{(T_1-m_1)(T_2-m_2)(T_3-m_3)}\, Q(m_1+1,m_2,m_3)^{(T_1-m_1)(T_2-m_2-1)(T_3-m_3-1)}}{Q(m_1,m_2,m_3)^{(T_1-m_1-1)(T_2-m_2-1)(T_3-m_3-1)}\, Q(m_1+1,m_2+1,m_3)^{(T_1-m_1)(T_2-m_2)(T_3-m_3-1)}} \times$$
$$\frac{Q(m_1,m_2+1,m_3)^{(T_1-m-1-1)(T_2-m_2)(T_3-m_3-1)}\, Q(m_1,m_2,m_3+1)^{(T_1-m_1-1)(T_2-m-2-1)(T_3-m_3-1)}}{Q(m_1+1,m_2,m_3+1)^{(T_1-m-1)(T-2-m_2-1)(T-3-m-3)}\, Q(m_1,m_2+1,m_3+1)^{(T_1-m_1-1)(T_2-m_2)(T_3-m-3)}}$$

Where,

$$Q(N_1,N_2,N_3) = \mathbb{P}\left(S_{m_1,m_2,m_3}(N_1,N_2,N_3) \leq k\right)$$

- The approximation also works for binomial and Poisson distribution
- Three Poisson Type Approximation

Haiman Type Approximation: Writing the Scan as an Extreme of 1-Dependent R.V.'s



Let $L_j = \frac{T_j}{m_j}$, $j \in \{1,2,3\}$ positive integers

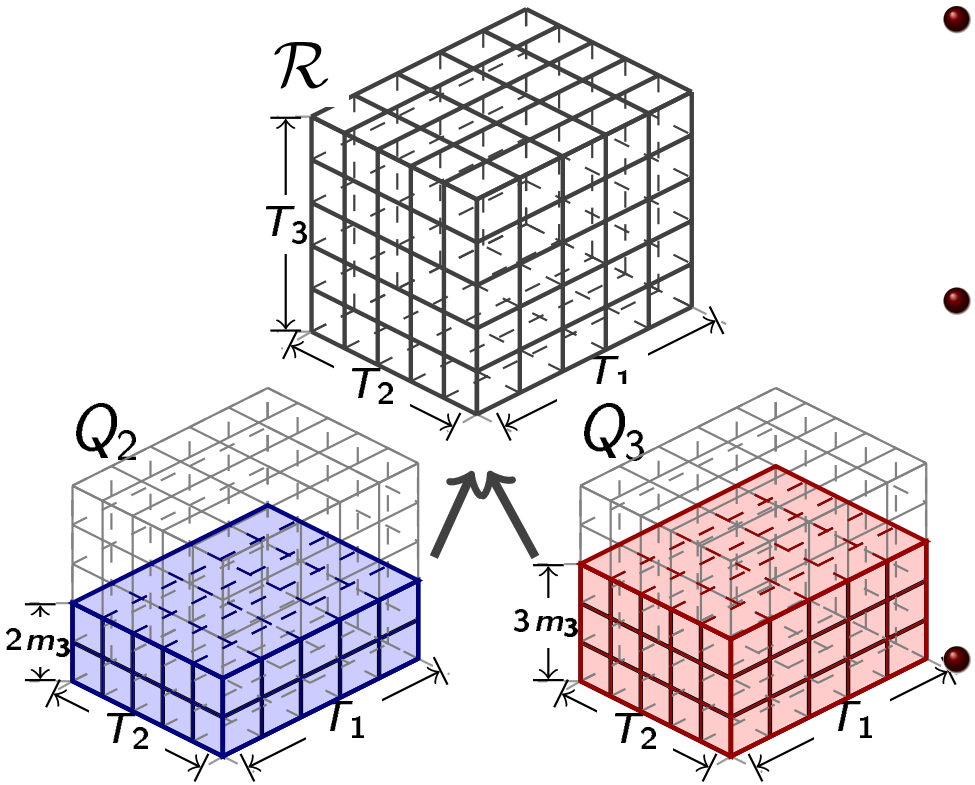
- Define for $l \in \{1,2,\ldots,L_3-1\}$

$$Z_l = \max_{\substack{1\le i_1\le (L_1-1)m_1+1\\ 1\le i_2\le (L_2-1)m_2+1\\ (l-1)m_3+1\le i_3\le lm_3+1}} Y_{i_1 i_2 i_3}$$

- $(Z_l)_l$ is 1-dependent and stationary
- Observe

$$S_{m_1,m_2,m_3} = \max_{1\le l\le L_3-1} Z_l$$

Haiman Type Approximation: First Step Approximation



- Define
$$Q_2 = \mathbb{P}(Z_1 \leq k)$$
$$Q_3 = \mathbb{P}(Z_1 \leq k, Z_2 \leq k)$$
- If $1 - Q_2 \leq \alpha_1 < 0.1$ the (first) approximation
$$\mathbb{P}(S \leq k) \approx \frac{2Q_2 - Q_3}{\left[1 + Q_2 - Q_3 + 2(Q_2 - Q_3)^2\right]^{L_3 - 1}}$$
where $S = S_{m_1, m_2, m_3}(T_1, T_2, T_3)$
- Approximation error
$$(L_3 - 1)E(\alpha_1, L_3 - 1)(1 - Q_2)^2$$

Haiman Type Approximation: Approximation for Q_2 and Q_3

Q_2 :

- For $l \in \{1, 2, \ldots, L_2 - 1\}$

$$Z_l^{(2)} = \max_{\substack{1 \le i_1 \le (L_1-1)m_1+1 \\ (l-1)m_2+1 \le i_2 \le lm_2+1 \\ 1 \le i_3 \le m_3+1}} Y_{i_1 i_2 i_3}$$

- $Q_2 = \mathbb{P}\left(\max_{1 \le l \le L_2 - 1} Z_l^{(2)} \le n \right)$
- Define
$$Q_{22} = \mathbb{P}(Z_1^{(2)} \le n)$$
$$Q_{32} = \mathbb{P}(Z_1^{(2)} \le n, Z_2^{(2)} \le n)$$
- Approximation $(1 - q_{22} \le \alpha_2)$

$$Q_2 \approx \frac{2Q_{22} - Q_{32}}{[1 + Q_{22} - Q_{32} + 2(Q_{22} - Q_{32})^2]^{L_2 - 1}}$$

- $(L_2 - 1)E(\alpha_2, L_2 - 1)(1 - Q_{22})^2$

Q_3 :

- For $l \in \{1, 2, \ldots, L_2 - 1\}$

$$Z_l^{(3)} = \max_{\substack{1 \le i_1 \le (L_1-1)m_1+1 \\ (l-1)m_2+1 \le i_2 \le lm_2+1 \\ 1 \le i_3 \le 2m_3+1}} Y_{i_1 i_2 i_3}$$

- $Q_3 = \mathbb{P}\left(\max_{1 \le l \le L_2 - 1} Z_l^{(3)} \le n \right)$
- Define
$$Q_{23} = \mathbb{P}(Z_1^{(3)} \le n)$$
$$Q_{33} = \mathbb{P}(Z_1^{(3)} \le n, Z_2^{(3)} \le n)$$
- Approximation $(1 - q_{23} \le \alpha_2)$

$$Q_3 \approx \frac{2Q_{23} - Q_{33}}{[1 + Q_{23} - Q_{33} + 2(Q_{23} - Q_{33})^2]^{L_2 - 1}}$$

- $(L_2 - 1)E(\alpha_2, L_2 - 1)(1 - Q_{23})^2$

Haiman Type Approximation: Last Step (Approximating Q_{ts})

Applying again the second part of the Main Theorem...

- For $s, t \in \{2, 3\}$ and $j \in \{1, 2, \ldots, L_1 - 1\}$ define

$$Z_j^{(ts)} = \max_{\substack{(j-1)m_1+1 \le i_1 \le jm_1+1 \\ 1 \le i_2 \le (t-1)m_2+1 \\ 1 \le i_3 \le (s-1)m_3+1}} Y_{i_1 i_2 i_3}$$

- Observe

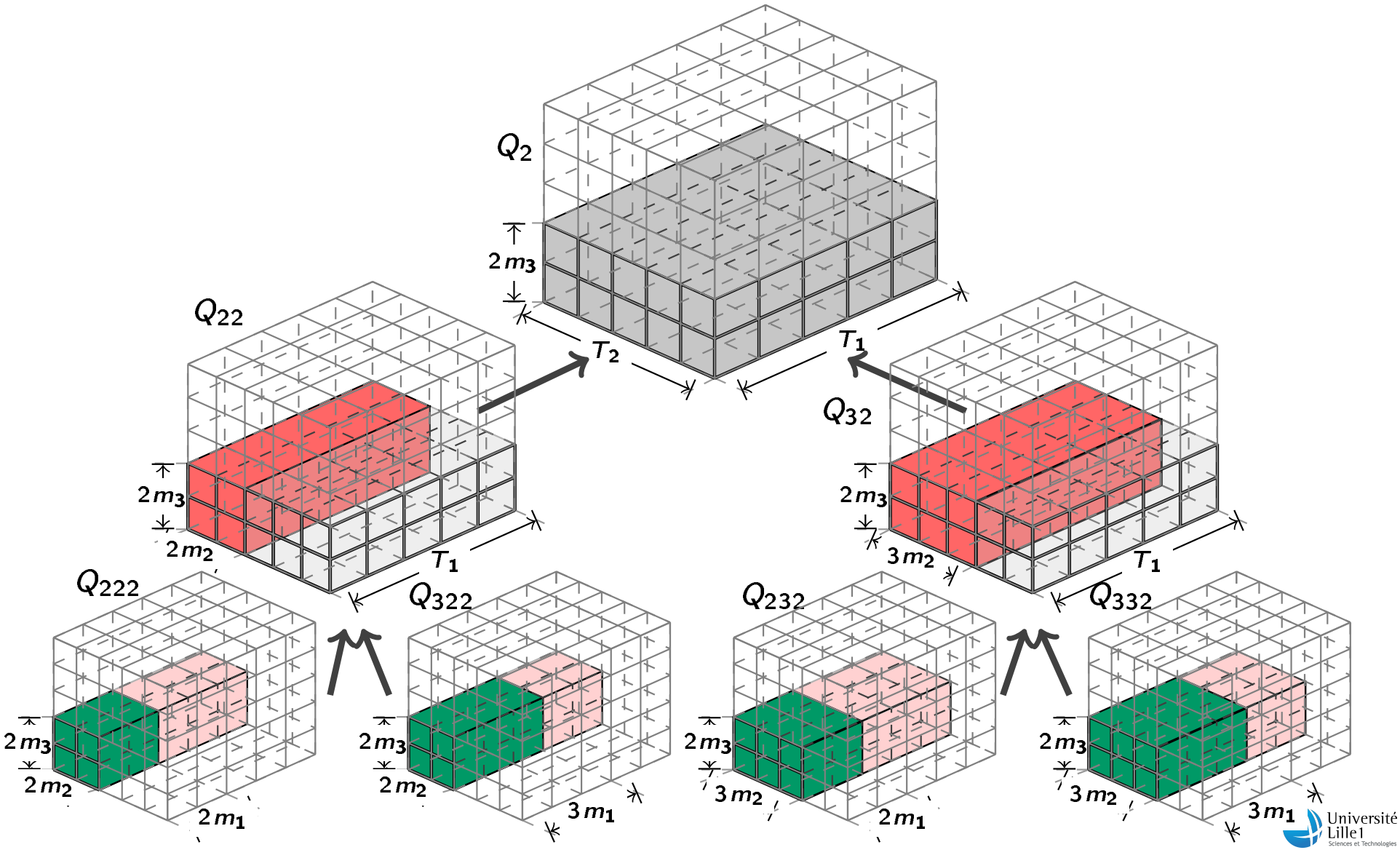
$$Q_{ts} = \mathbb{P}\left(\max_{1 \le j \le L_1 - 1} Z_j^{(ts)} \le n\right)$$

- Define for $r, s, t \in \{2, 3\}$, $Q_{rts} = \mathbb{P}\left(\cap_{j=1}^{r-1} \{Z_j^{(ts)} \le n\}\right)$
- If $1 - Q_{2ts} \le \alpha_3$ then the approximation and the error

$$Q_{ts} \approx \frac{2Q_{2ts} - Q_{3ts}}{\left[1 + Q_{2ts} - Q_{3ts} + 2(Q_{2ts} - Q_{3ts})^2\right]^{L_1 - 1}}$$

$$(L_1 - 1)E(\alpha_3, L_1 - 1)(1 - Q_{2ts})^2$$

An Illustration of the Approximation Chain (Q_2)



Outline

Bernoulli Case

Comparing with existing results:

Table 5 : $n = 1, p = 0.00005, m_1 = m_2 = m_3 = 5, T_1 = T_2 = T_3 = 60, It_{App} = 10^5$

k	$\hat{\mathbb{P}}(S \leq k)$	Glaz et al. Product type	Our Approximation	Approximation Error	Simulation Error	Total Error
1	0.841806	0.841424	0.851076	0.011849	0.064889	0.076738
2	0.999119	0.999142	0.999192	0.000000	0.000170	0.000170
3	0.999997	0.999998	0.999997	0.000000	3×10^{-7}	3×10^{-7}

Table 6 : $n = 1, p = 0.0001, m_1 = m_2 = m_3 = 5, T_1 = T_2 = T_3 = 60, It_{App} = 10^5$

k	$\hat{\mathbb{P}}(S \leq k)$	Glaz et al. Product type	Our Approximation	Approximation Error	Simulation Error	Total Error
2	0.993294	0.993241	0.993192	0.000010	0.001367	0.001377
3	0.999963	0.999964	0.999963	0.000000	0.000005	0.000005
4	0.999999	0.999999	0.999999	0.000000	2×10^{-9}	2×10^{-9}

- Glaz, J., Naus, J., Wallenstein, S.: Scan statistic. *Springer* (2001).
- Glaz, J., Pozdnyakov, V., Wallenstein, S.: Scan statistic: Methods and Applications. *Birkhauser* (2009).
- Amarioarei, A.: Approximation for the distribution of extremes of one dependent stationary sequences of random variables, arXiv:1211.5456v1 (submitted)
- Amarioarei, A., Preda, C.: Approximation for the distribution of three dimensional discrete scan statistic, *Methodol Comput Appl Probab* (2013) DOI:10.1007/s11009-013-9382-3.
- Amarioarei, A., Preda, C.: Approximation for two dimensional discrete scan statistics in some block-factor type dependent models, arXiv:1401.2822 (submitted)
- Boutsikas, M.V., Koutras, M.: Reliability approximations for Markov chain imbeddable systems. *Methodol Comput Appl Probab* **2** (2000), 393–412.

- Boutsikas, M. and Koutras, M. *Bounds for the distribution of two dimensional binary scan statistics*, Probability in the Engineering and Information Sciences, 17, 509–525, 2003.
- Chen, J. and Glaz, J. *Two-dimensional discrete scan statistics*, Statistics and Probability Letters 31, 59–68, 1996.
- Glaz, J., Guerriero, M., Sen, R.: Approximations for three dimensional scan statistic. *Methodol Comput Appl Probab* **12** (2010), 731–747.
- Haiman, G.: First passage time for some stationary sequence. *Stoch Proc Appl* **80** (1999), 231–248.
- Haiman, G.: Estimating the distribution of scan statistics with high precision. *Extremes* **3** (2000), 349–361.
- Haiman, G., Preda, C.: A new method for estimating the distribution of scan statistics for a two-dimensional Poisson process. *Methodol Comput Appl Probab* **4** (2002), 393–407.


Haiman, G., Preda, C.: Estimation for the distribution of two-dimensional scan statistics. *Methodol Comput Appl Probab* **8** (2006), 373–381.


Haiman, G.: Estimating the distribution of one-dimensional discrete scan statistics viewed as extremes of 1-dependent stationary sequences. *J. Stat Plan Infer* **137** (2007), 821–828.


Kuai, H., Alajaji, F., Takahara, G.: A lower bound on the probability of a finite union of events. *Discrete Mathematics* **215** (2000), 147–158.

Introducing the Model

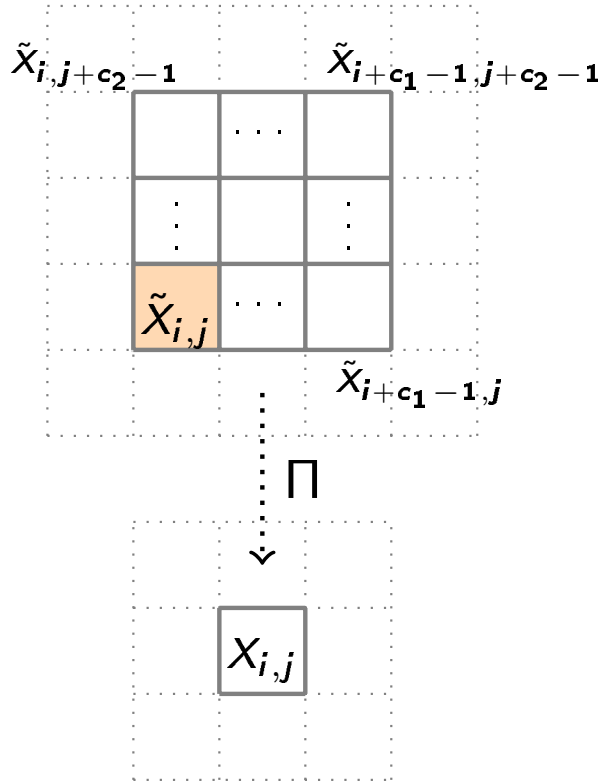
Let $1 \leq c_s \leq \tilde{T}_s$, $s \in \{1, 2\}$ integers

- $\left(\tilde{X}_{ij}\right)_{\substack{1 \leq i \leq \tilde{T}_1 \\ 1 \leq j \leq \tilde{T}_2}}$ i.i.d. r.v.'s
- configuration matrix in (i, j)

$$C_{(i,j)} = \left(C_{(i,j)}(k, l)\right)_{\substack{1 \leq k \leq c_2 \\ 1 \leq l \leq c_1}}$$

$$C_{(i,j)}(k, l) = \tilde{X}_{i+l-1, j+c_2-k}$$

- transformation $\Pi : \mathcal{M}_{c_2, c_1}(\mathbb{R}) \to \mathbb{R}$



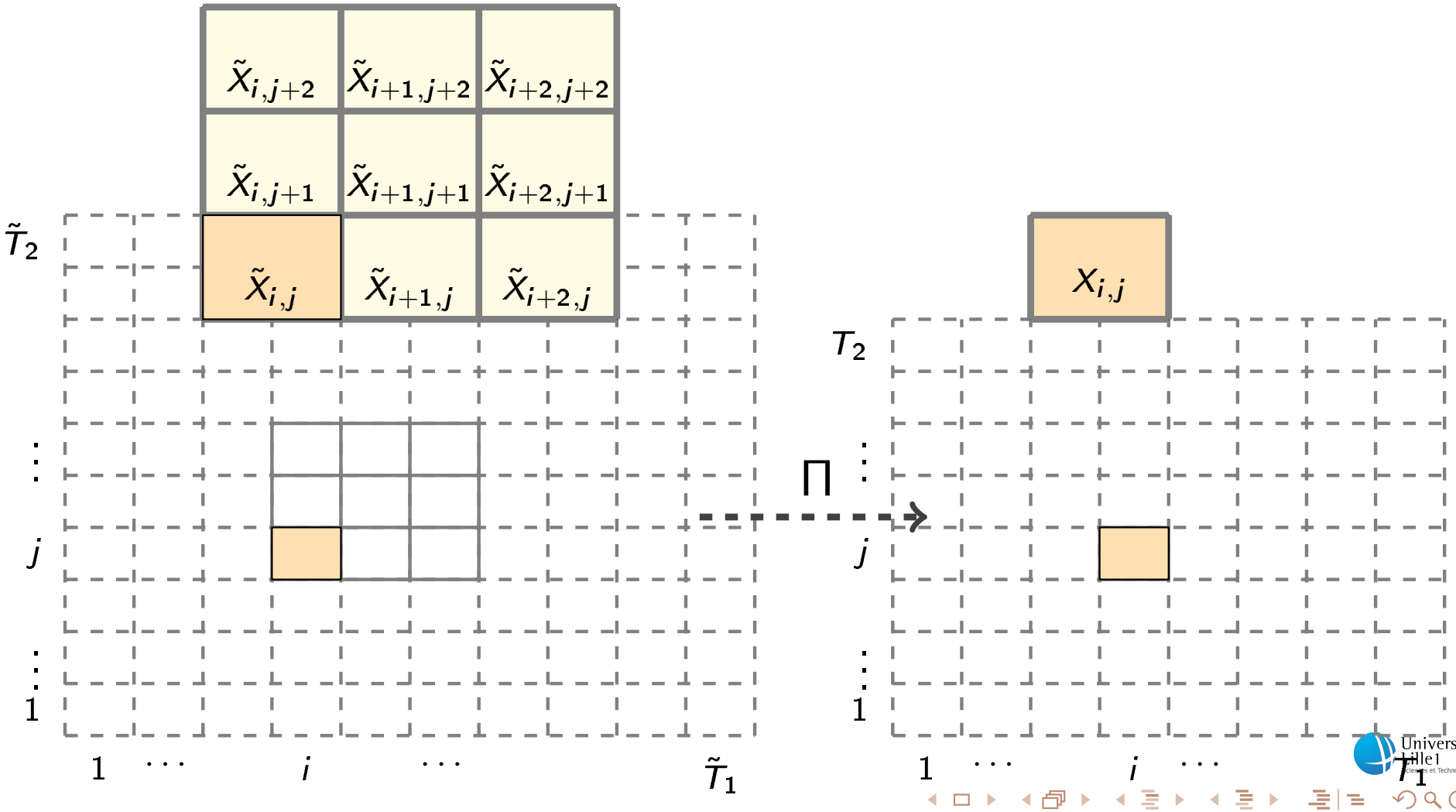
Define the block-factor model, $T_1 = \tilde{T}_1 - c_1 + 1$, $T_2 = \tilde{T}_2 - c_2 + 1$

$$X_{i,j} = \Pi\left(C_{(i,j)}\right), \quad \begin{matrix} 1 \leq i \leq T_1 \\ 1 \leq j \leq T_2 \end{matrix}$$

Return

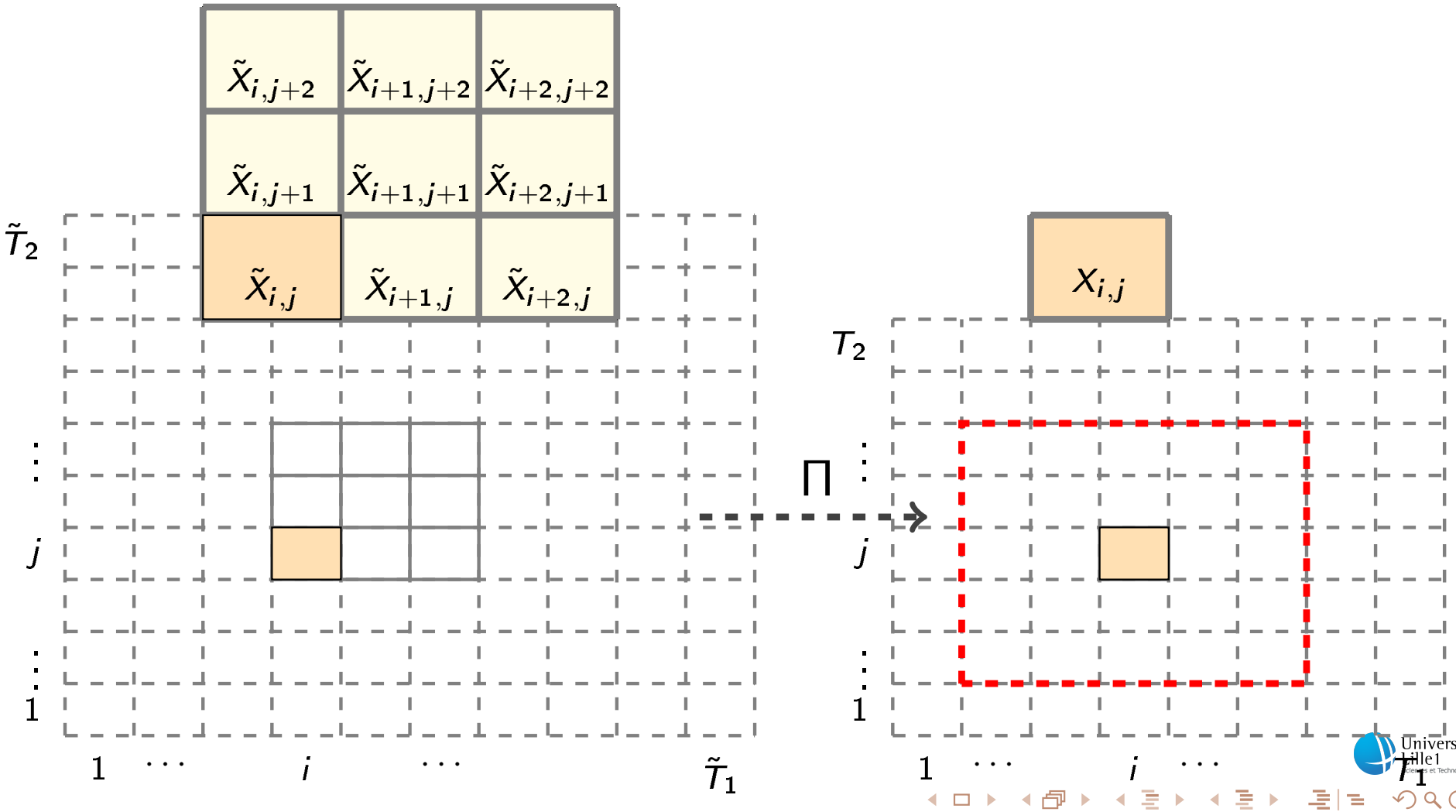
Model: case $c_1 = c_2 = 3$

- To simplify the presentation we consider $c_1 = c_2 = 3$

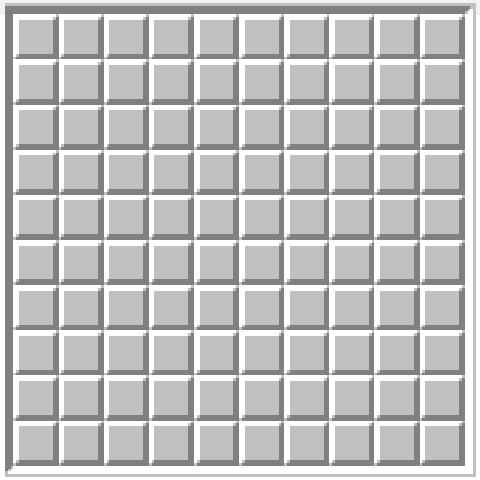


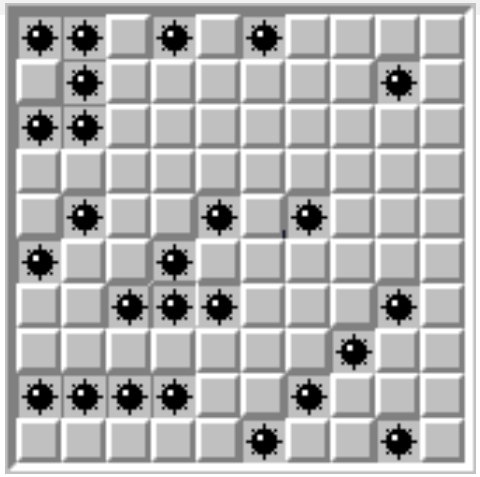
Model: case $c_1 = c_2 = 3$

- To simplify the presentation we consider $c_1 = c_2 = 3$



Example: A game of minesweeper





Model:

- $\tilde{X}_{i,j} \sim \mathcal{B}(p)$ (presence, absence of a mine)
- number of neighboring mines

$$T\left(C_{(i,j)}\right) = \sum_{\substack{(s,t)\in\{0,1,2\}^2 \\ (s,t)\neq(1,1)}} \tilde{X}_{i+s,j+t}$$

  - $X_{i,j} = \Pi\left(C_{(i,j)}\right)$

Return

Computation of $Q(m_1+1, m_2)$ and $Q(m_1+1, m_2+1)$

Consider $X_{ij} \sim B(n,p)$ and the notation $Y_{i_1,i_2}^{j_1,j_2} = \sum_{i_1=1}^{j_1} \sum_{i_2=1}^{j_2} X_{ij}$,

$$\mathbb{P}\left(S_{m_1,m_2}(m_1+1,m_2) \leq k\right) = \sum_{y=0}^{k \wedge (m_1-1)m_2 n} \mathbb{P}^2\left(Y_{1,1}^{1,m_2} \leq k-y\right) \mathbb{P}\left(Y_{2,1}^{m_1,m_2} = y\right)$$

$$\begin{aligned}
\mathbb{P}\left(S_{m_1,m_2}(m_1+1,m_2+1) \leq k\right) = & \sum_{y_1=0}^{k \wedge (m_1-1)(m_2-1)n} \sum_{y_2=0}^{(k-y_1) \wedge (m_2-1)n} \sum_{y_3=0}^{(k-y_1) \wedge (m_2-1)n} \\
& \sum_{y_4=0}^{[k-y_1-y_2 \vee y_3] \wedge (m_1-1)n} \sum_{y_5=0}^{[k-y_1-y_2 \vee y_3] \wedge (m_1-1)n} \mathbb{P}\left(Y_{1,1}^{1,1} \leq a_1\right) \\
& \times \mathbb{P}\left(Y_{1,m_2+1}^{1,m_2+1} \leq a_2\right) \mathbb{P}\left(Y_{m_1+1,1}^{m_1+1,1} \leq a_3\right) \mathbb{P}\left(Y_{m_1+1,m_2+1}^{m_1+1,m_2+1} \leq a_4\right) \\
& \times \mathbb{P}\left(Y_{2,2}^{m_1,m_2} = y_1\right) \mathbb{P}\left(Y_{1,2}^{1,m_2} = y_2\right) \mathbb{P}\left(Y_{m_1+1,2}^{m_1+1,m_2} = y_3\right) \\
& \times \mathbb{P}\left(Y_{2,1}^{m_1,1} = y_4\right) \mathbb{P}\left(Y_{2,m_2+1}^{m_1,m_2+1} = y_5\right)
\end{aligned}$$

$$a_1 = k - y_1 - y_2 - y_4, \quad a_2 = k - y_1 - y_2 - y_5,$$
$$a_3 = k - y_1 - y_3 - y_4, \quad a_4 = k - y_1 - y_3 - y_5,$$

Return

Karwe Naus recursive methods

Define

$$b^k_{2(m)}(y) = \mathbb{P}\left(S_m(2m) \leq k, Y_{m+1}(m) = y\right)$$
$$f(y) = \mathbb{P}(X_1 = y)$$
$$Q^k_{2m} = \mathbb{P}\left(S_m(2m) \leq k\right)$$

We have the recurrence relations

$$b^k_{2(1)}(y) = \left(\sum_{j=0}^{k} f(j)\right) f(y)$$
$$b^k_{2(m)}(y) = \sum_{\eta=0}^{y} \sum_{\nu=0}^{k-y+\eta} b^{k-\nu}_{2(m-1)}(y-\eta) f(\nu) f(\eta)$$
$$Q^k_{2m} = \sum_{y=0}^{k} b^k_{2(m)}(y)$$
$$Q^k_{2m-1} = \sum_{x=0}^{k} f(x) Q^{k-x}_{2(m-1)}$$

Return

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

α	$K(\alpha)$	$\Gamma(\alpha)$
0.1	38.63	480.69
0.05	21.28	180.53
0.025	17.56	145.20
0.01	15.92	131.43

Table 7 : Selected values for $K(\alpha)$ and $\Gamma(\alpha)$

Return