Modeling distances between ignitions.

André Dabrowski
University of Ottawa

May 25, 2005

Abstract

Forest fires ignitions by lightning are frequently represented as a point process, and one can adopt a variety of methods to study the characteristics of such a process. Here we look at inter-point distances, $\|X_1 - X_2\|$ and the tail index α defined by

$$P[\|X_1 - X_2\| \leq x] \sim x^\alpha$$

as $x \downarrow 0$. This index is can be estimated using the extreme least order statistics of the inter-point distances, and we illustrate those methods on a sample data set of ignitions.

Modeling distances between ignitions.

1. A data set of ignitions.

2. Inter-point distances and a power law.

3. A limit theorem for minimal inter-point distances.

4. Application to data.

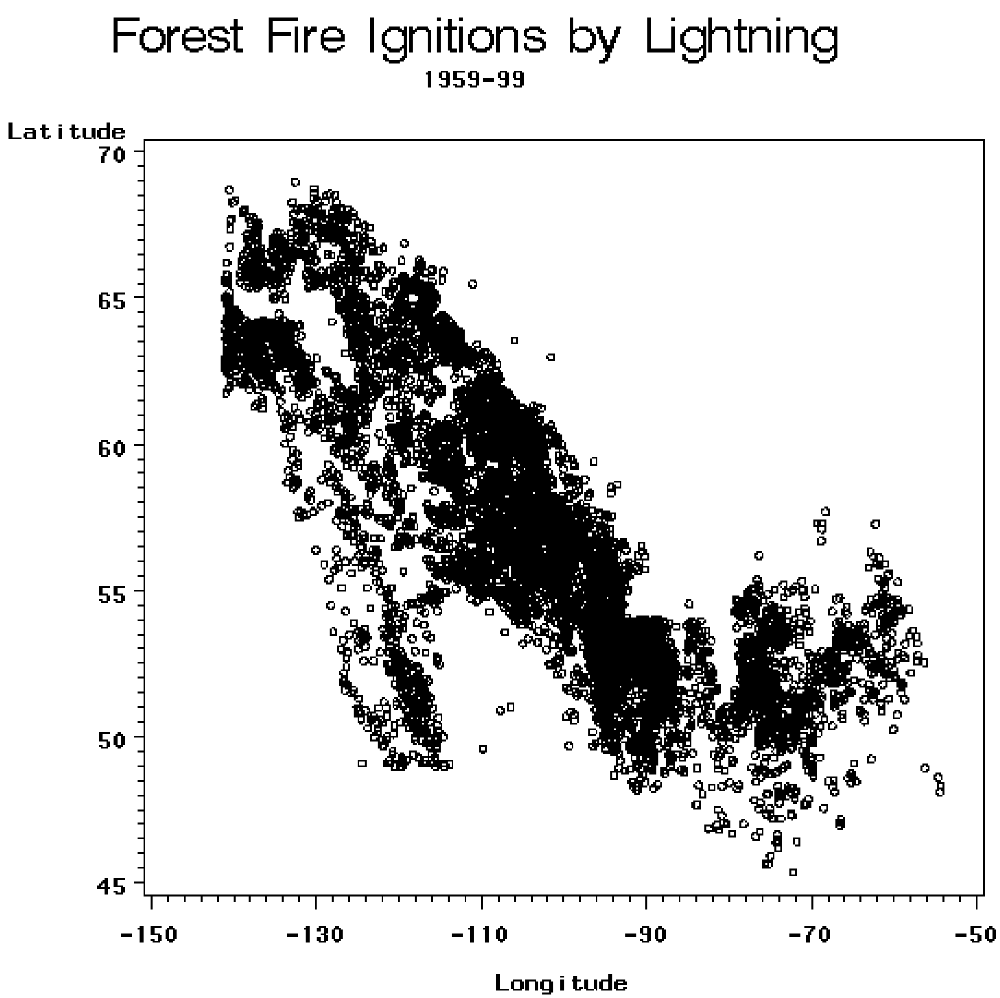
1. A data set of ignitions.

`http://cwfis.cfs.nrcan.gc.ca/en/historical/ha_lfdb_maps_e.php`

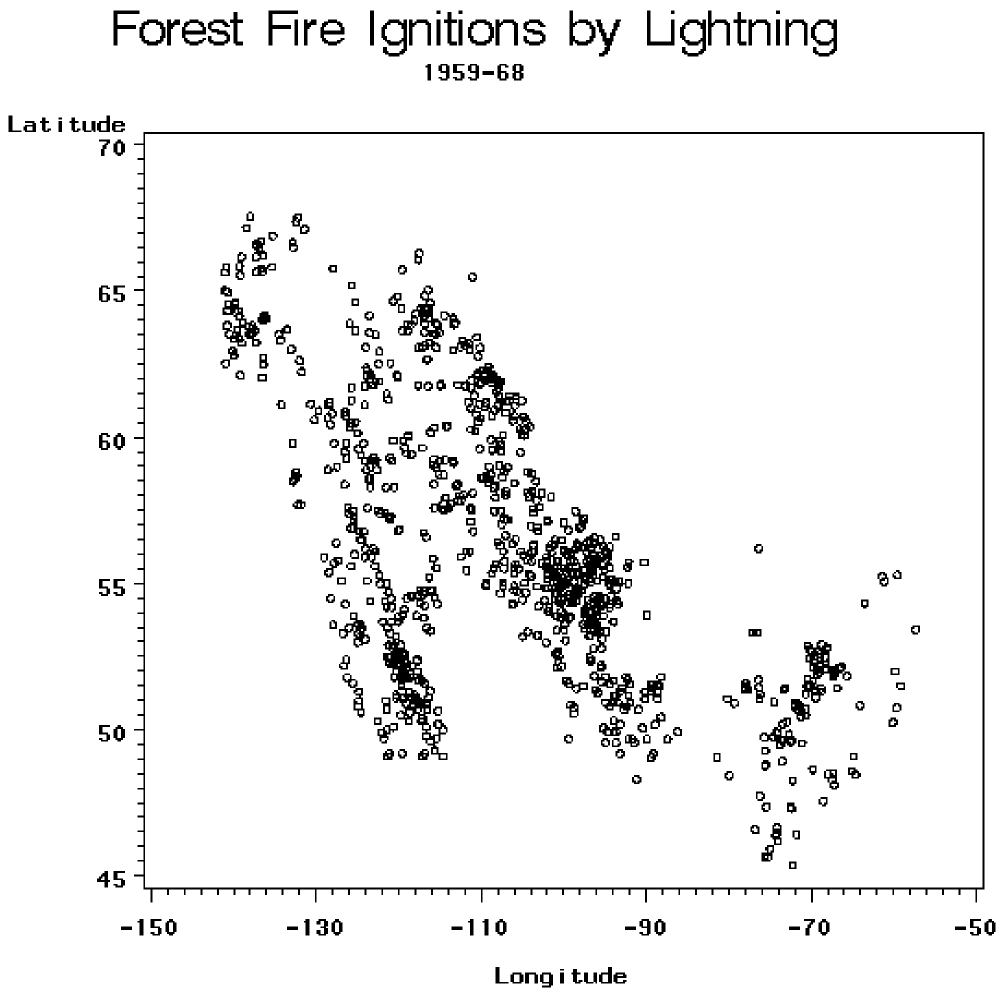
A data set of lightning-induced fires across Canada of at least 200*ha* from 1959 to 1999.

Retained Data:
longitude,
latitude,
detection date

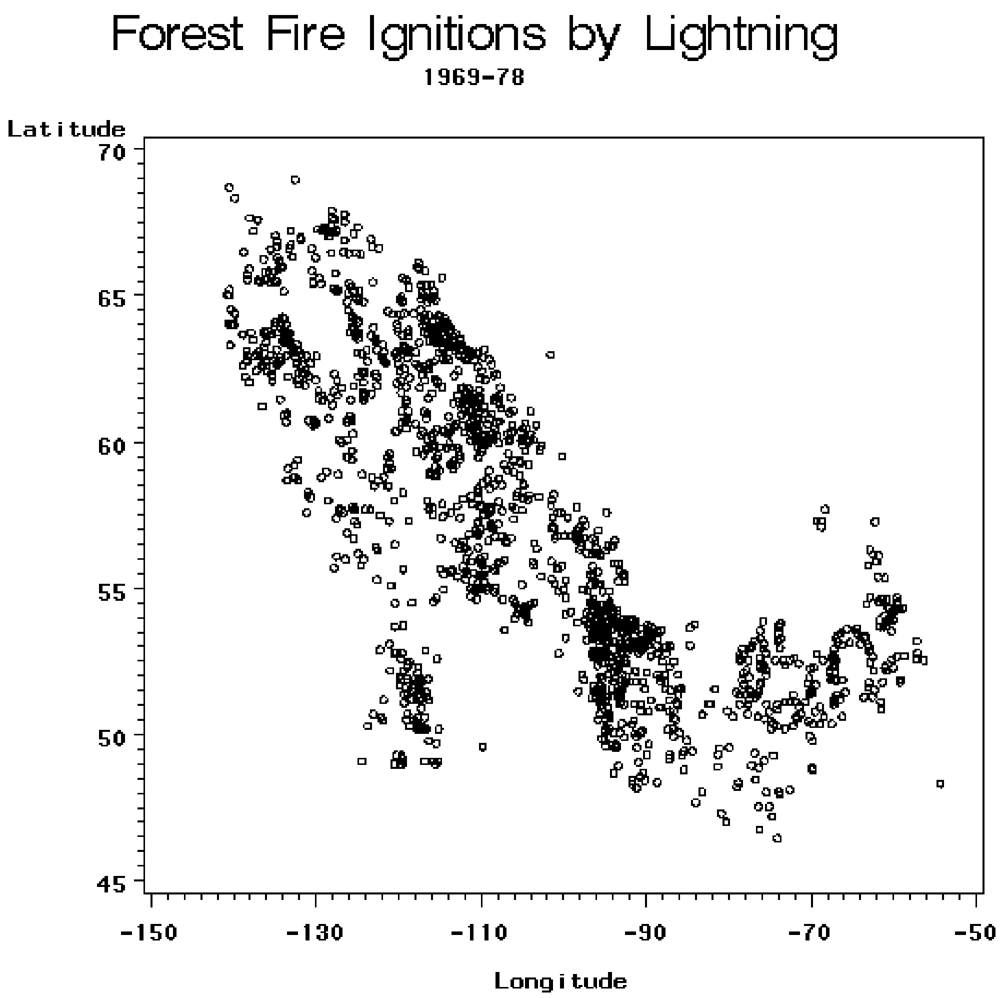
We can see regional variations in numbers of ignitions.



We also see variations in time.



We also see variations in time.



Nonetheless, we can think of the data set as the realization of a spatio-temporal point process, and can seek to characterize some of its properties.

- Poisson-ness (homogeneous, cluster, . . .)
- intensity measure (exogenous, integrate-and-fire, . . .)
- inter-point distances (spatial structure, K function, . . .)

2. Inter-point distances and a power law.

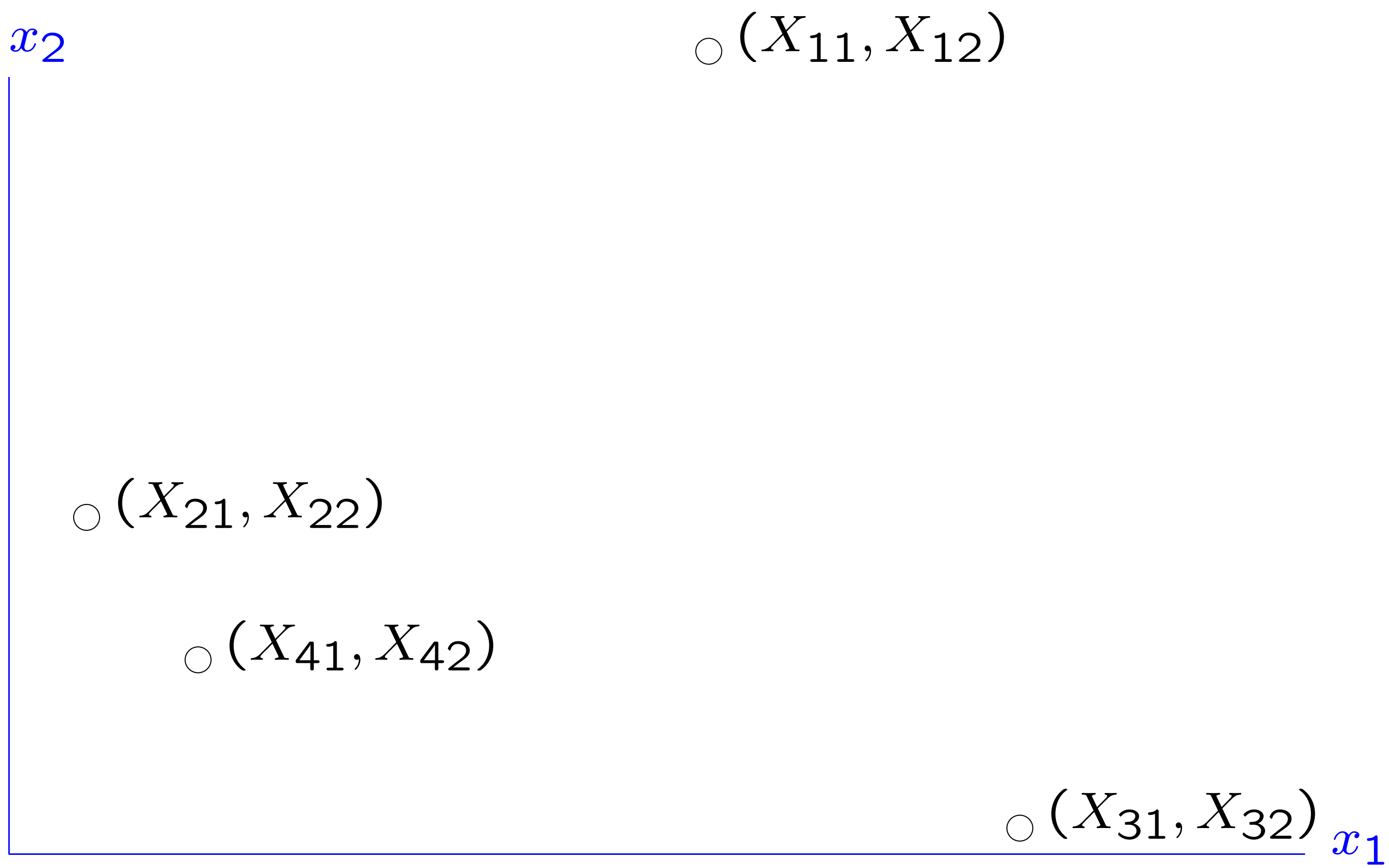
Denote the location of the kth ignition by X_k. In general the location can be in any dimension (e.g. $\Re^d$-valued) but for simplicity we will assume that $X_k = (X_{k1}, X_{k2}) \in \Re^2$. The inter-point distance between X_i and X_j is defined as $h(X_i, X_j)$, where h is a positive symmetric function.

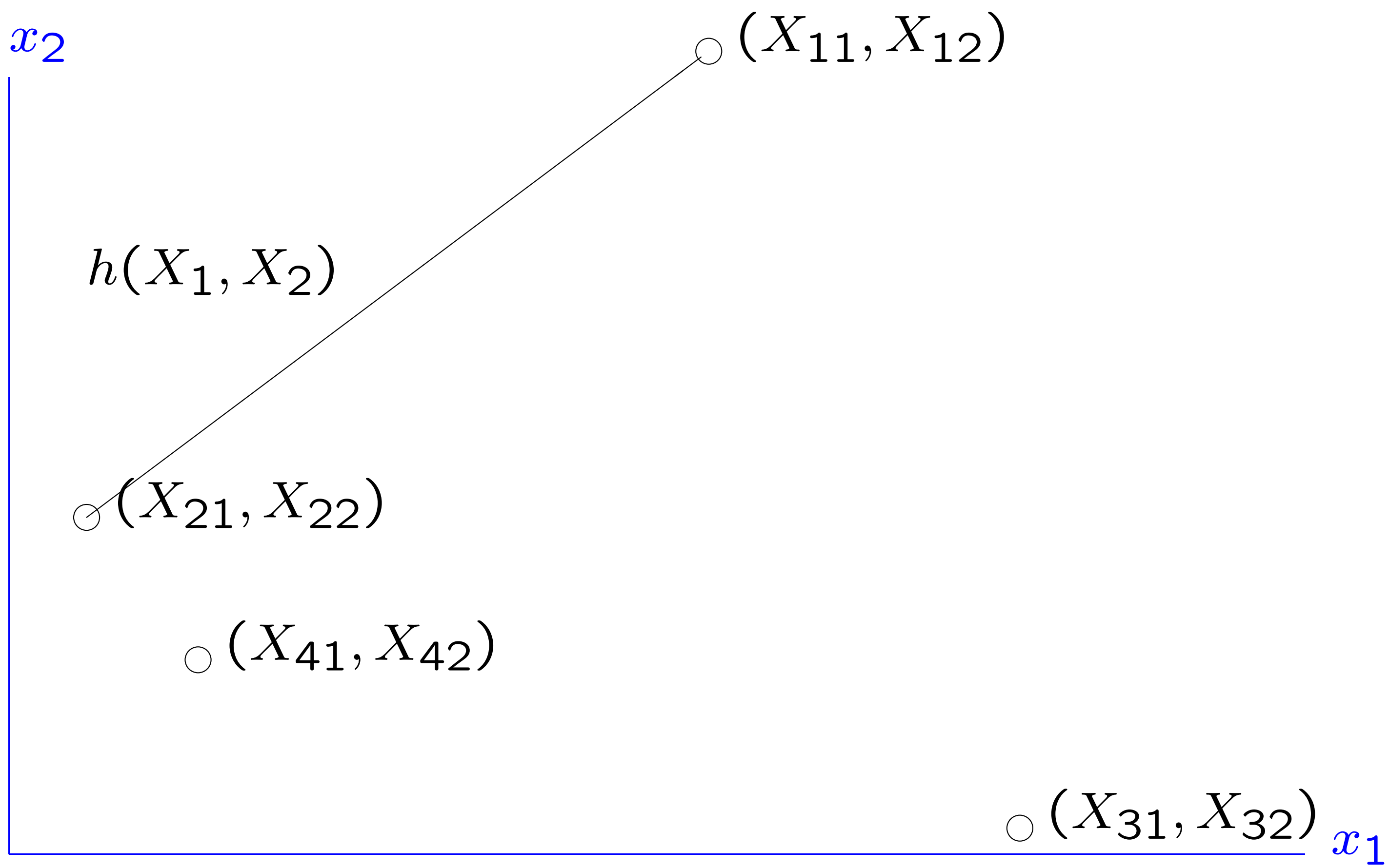
We will assume that $\{X_k \ : \ k \geq 1\}$ are independent and identically distributed random vectors.

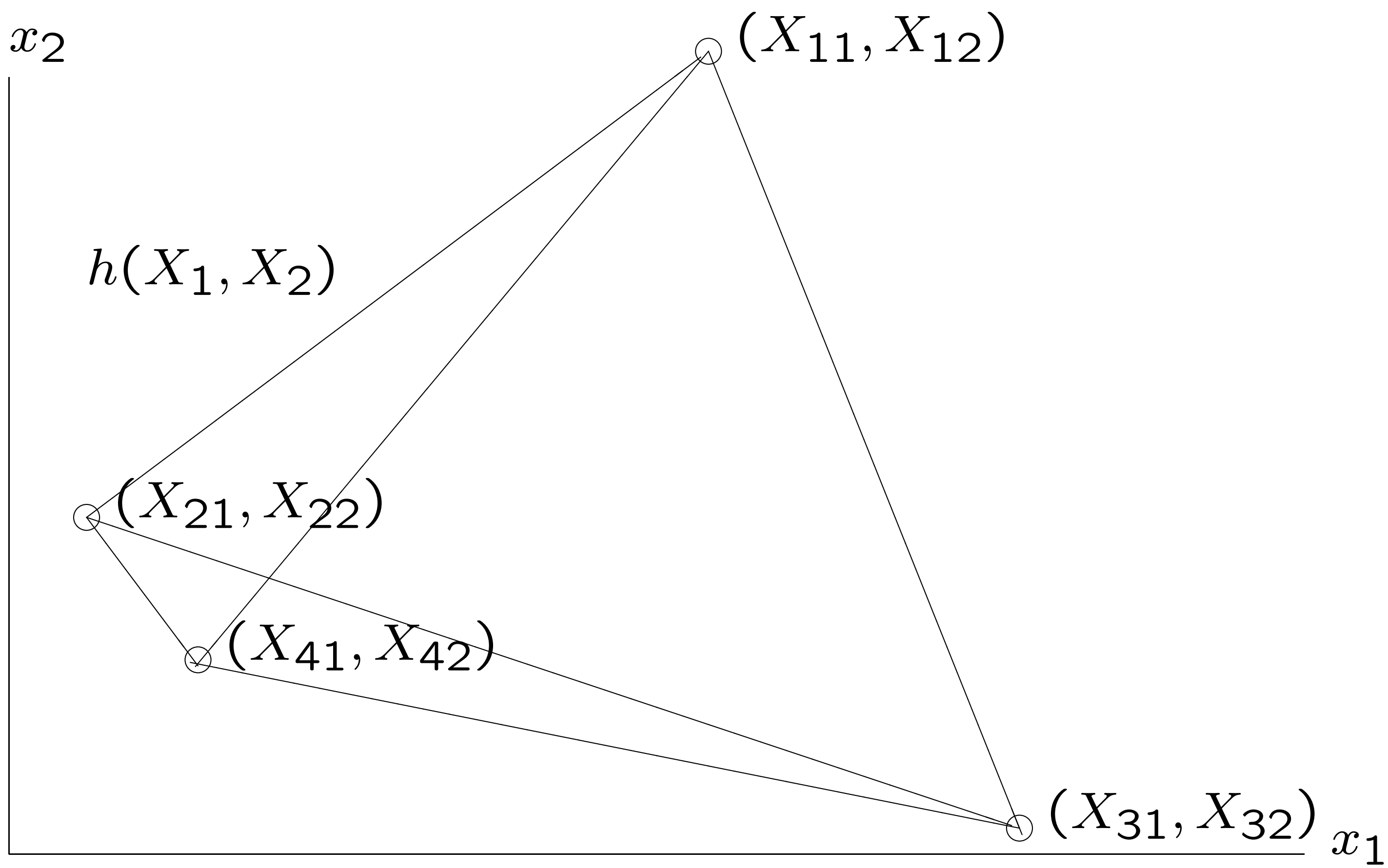
We further assume that

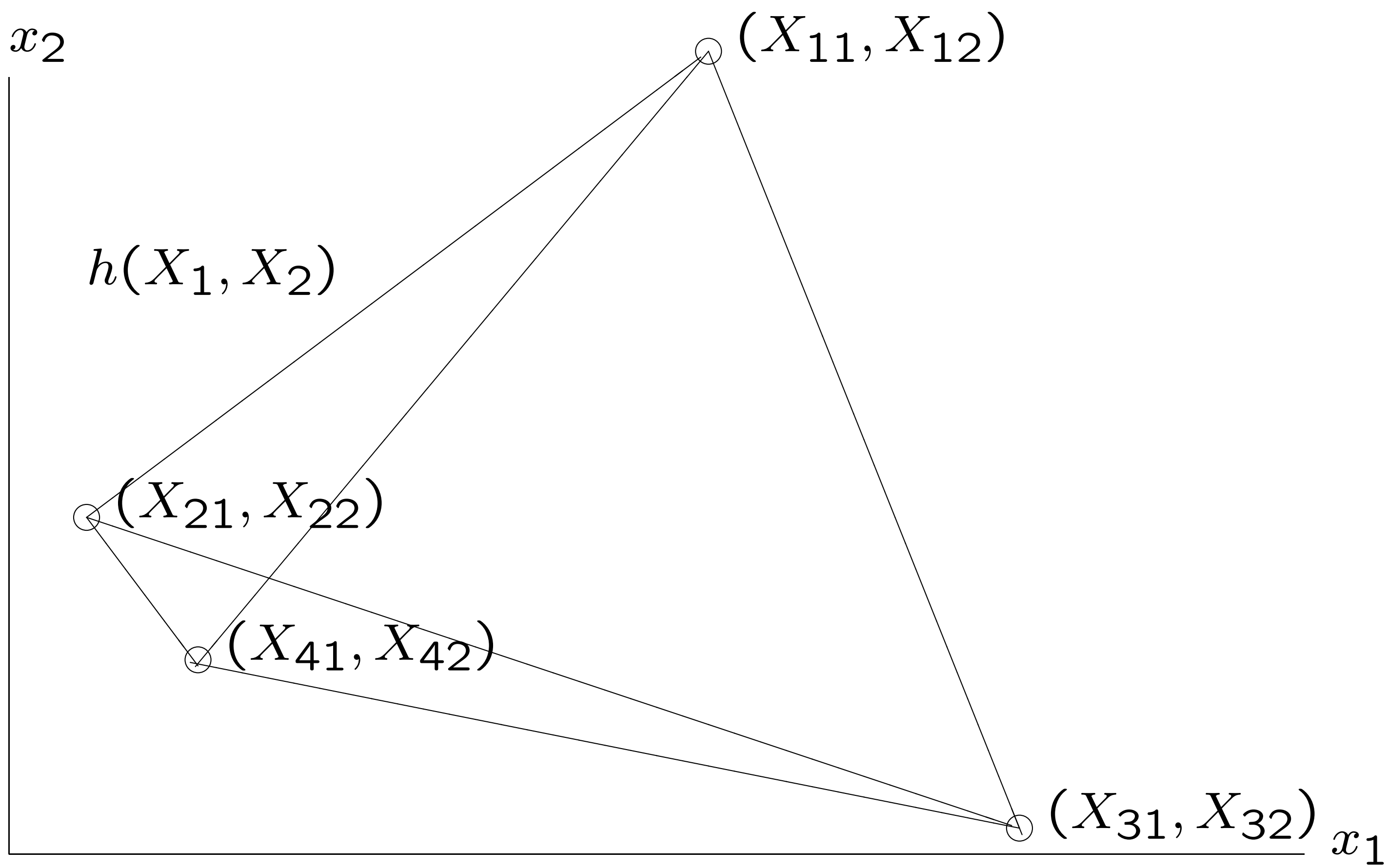
$$P[h(X_1, X_2) \leq x] \ \sim L(x^{-1})x^{\alpha}$$

for a slowly varying function L and some $\alpha > 0$.


x_2
(X_{11}, X_{12})
(X_{21}, X_{22})
(X_{41}, X_{42})
(X_{31}, X_{32})
x_1


x_2
(X_{11}, X_{12})
$h(X_1, X_2)$
(X_{21}, X_{22})
(X_{41}, X_{42})
(X_{31}, X_{32})
x_1


x_2
(X_{11}, X_{12})
$h(X_1, X_2)$
(X_{21}, X_{22})
(X_{41}, X_{42})
(X_{31}, X_{32})
x_1



The $h(X_i, X_j)$ are not independent.

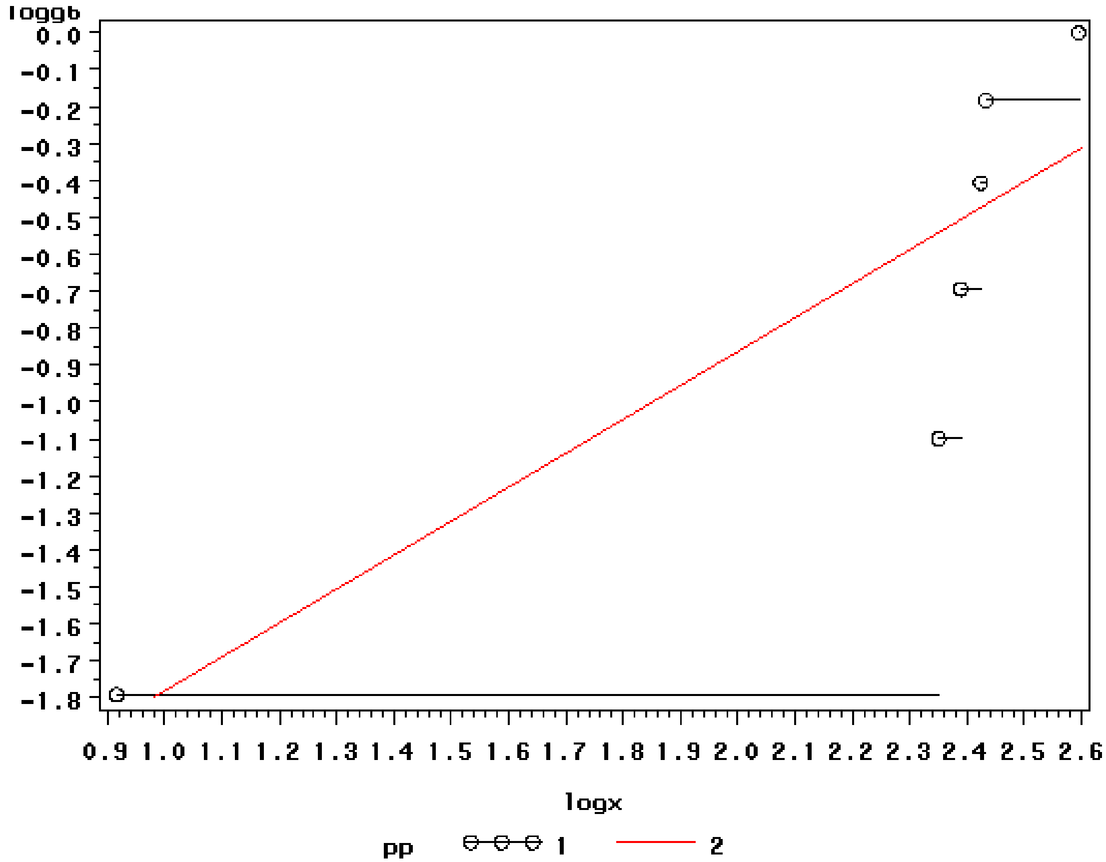
If we believe

$$P[h(X_1, X_2) \leq x] \; \sim x^{\alpha}$$

then we can write down the empirical distribution function, G_k, of the $h(X_i, X_j)$, try to fit it by x^{α} (or $\log G_k(x)$ by $\alpha \log x$), and so estimate α, i.e. as $x \downarrow 0$,

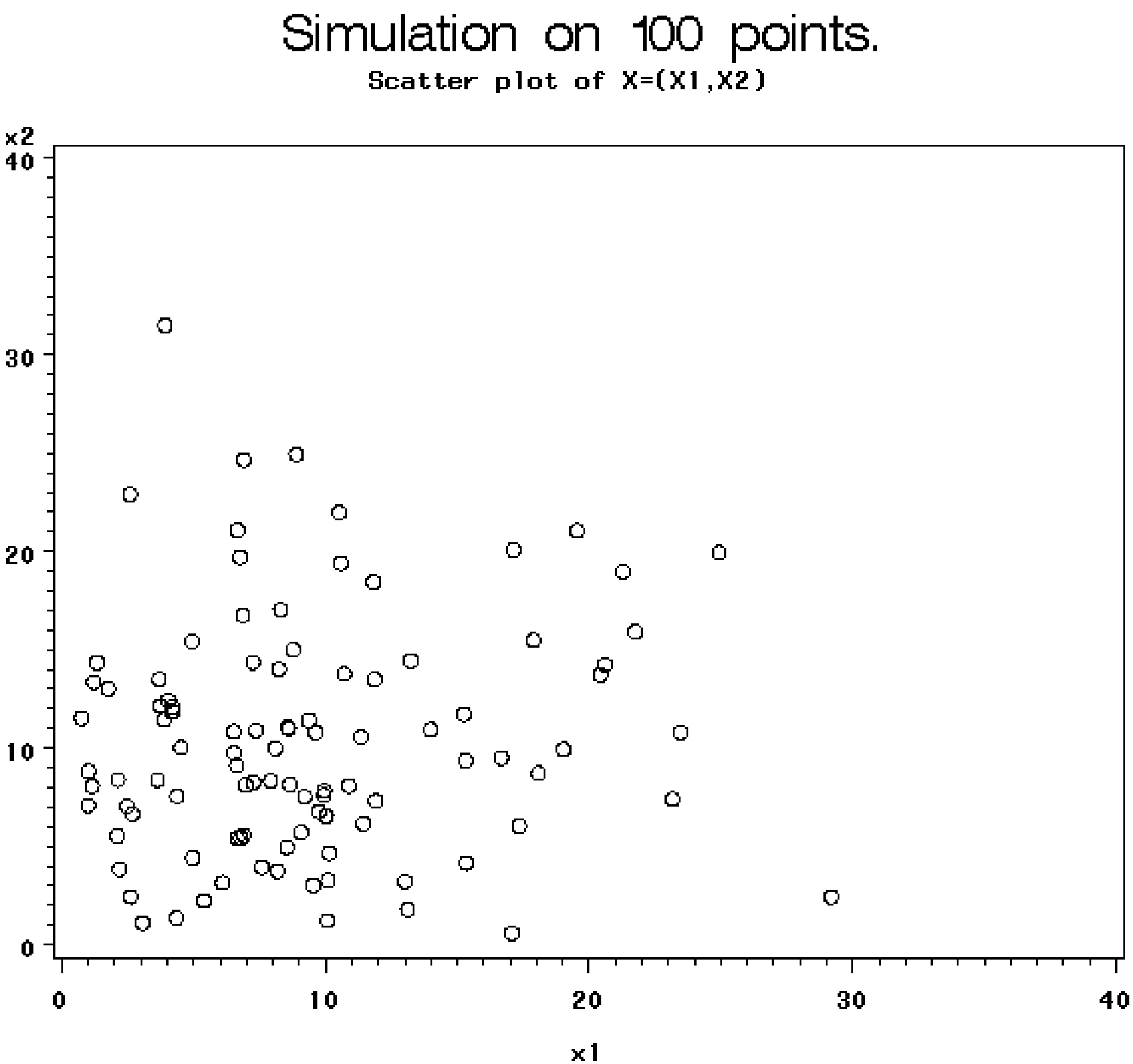
$$\log G(x) = \log P[h(X_1, X_2) \leq x] \; \sim \log c + \alpha \log x$$

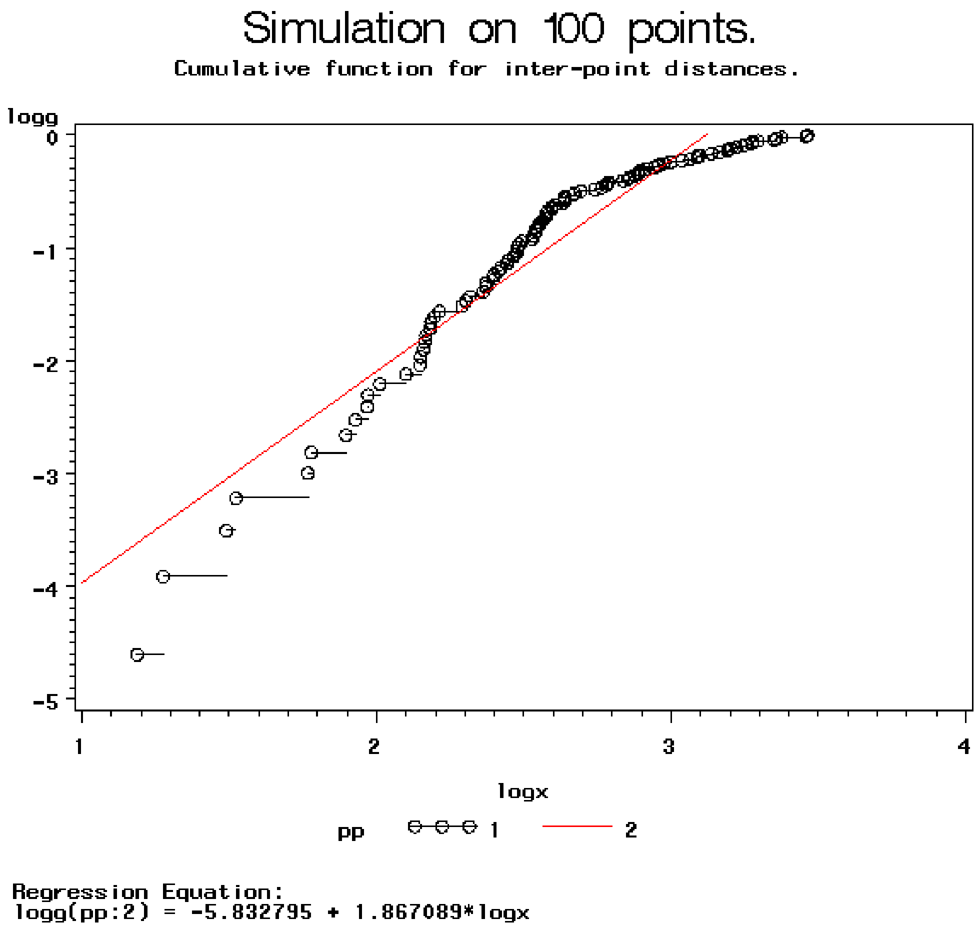
This is not too impressive on the 4-point data set:



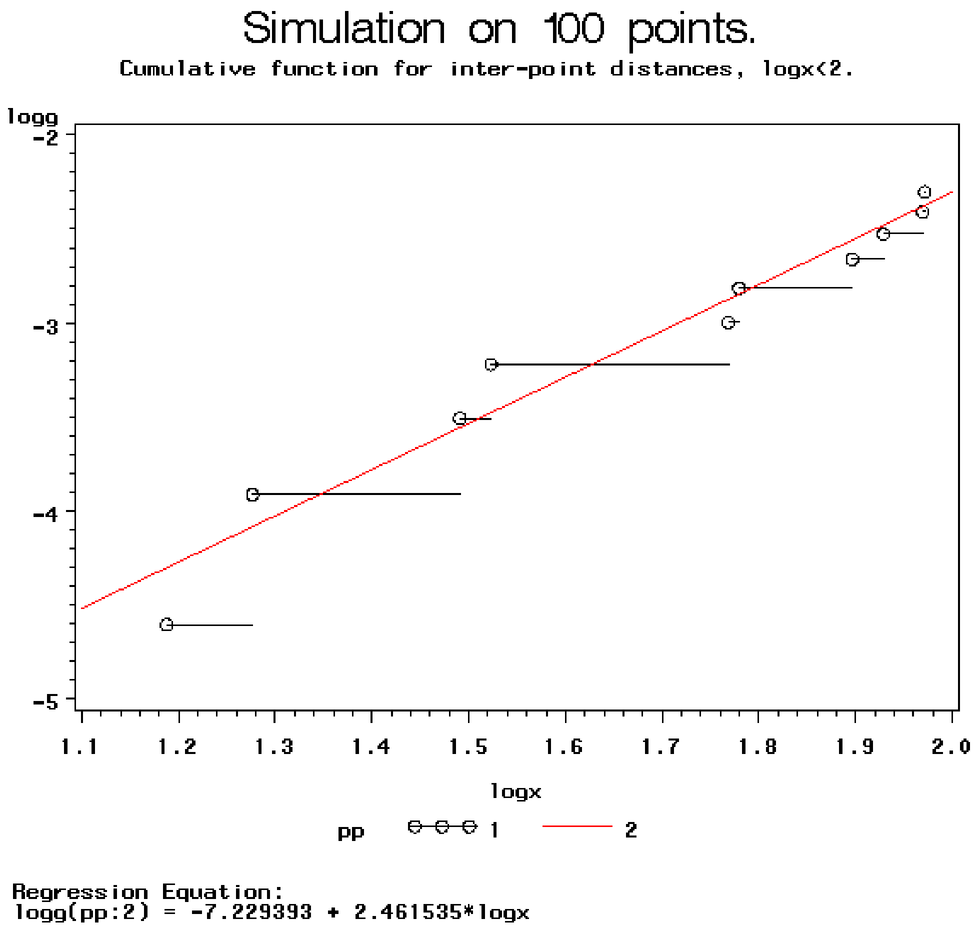
Regression Equation:
logg6(pp:2) = -2.702746 + 0.91877*logx

2.5
10.5
10.9
11.3
11.4
13.4


Simulation on 100 points.
Scatter plot of X=(X1,X2)
x2
40
30
20
10
0
0
10
20
30
40
x1

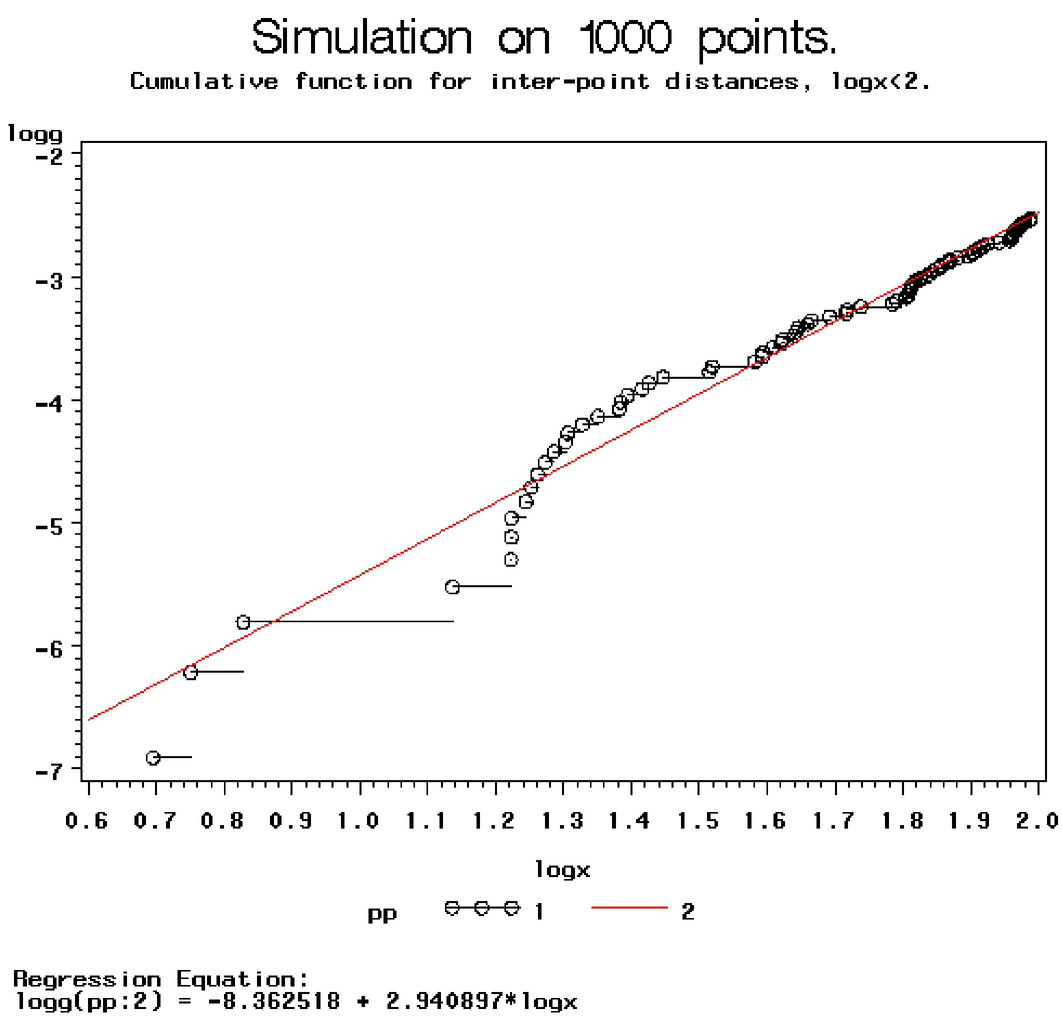


The straight line fit to all the $h(x_i, x_j)$.

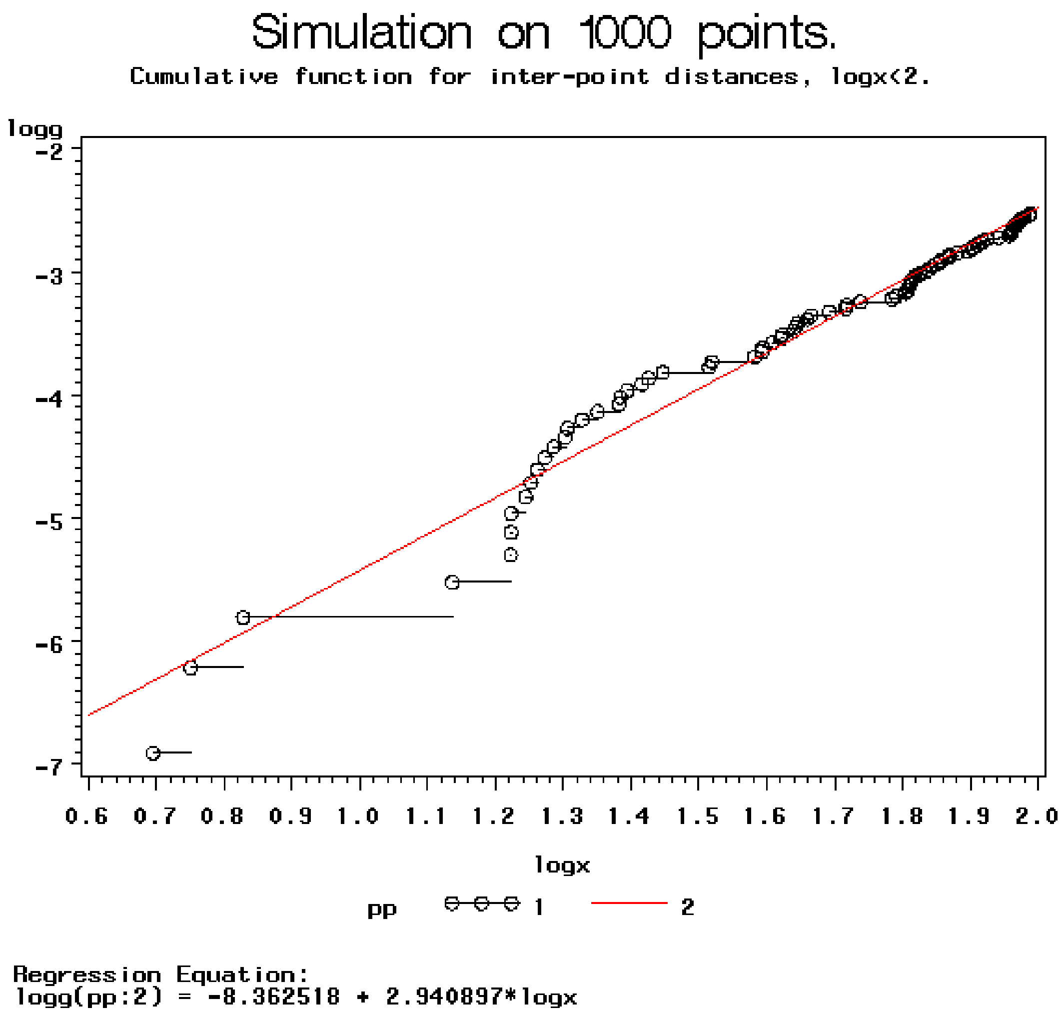


The straight line fit to a lower tail of the $h(x_i, x_j)$. Good here but ...

... sometimes this is not as convincing.



... and sometimes this is not as convincing.



Asymptotics?

Can we say anything about the limiting properties of such an estimator?

3. A limit theorem for minimal inter-point distances.

1. X_i iid, h symmetric non-negative kernel. For example, take $h(x,y) = |x-y|^\gamma$.

2. Regular variation condition: For $\alpha > 0$, as $x \to 0$

$$P[h(X_1, X_2) \le x] = L(x^{-1})x^\alpha.$$

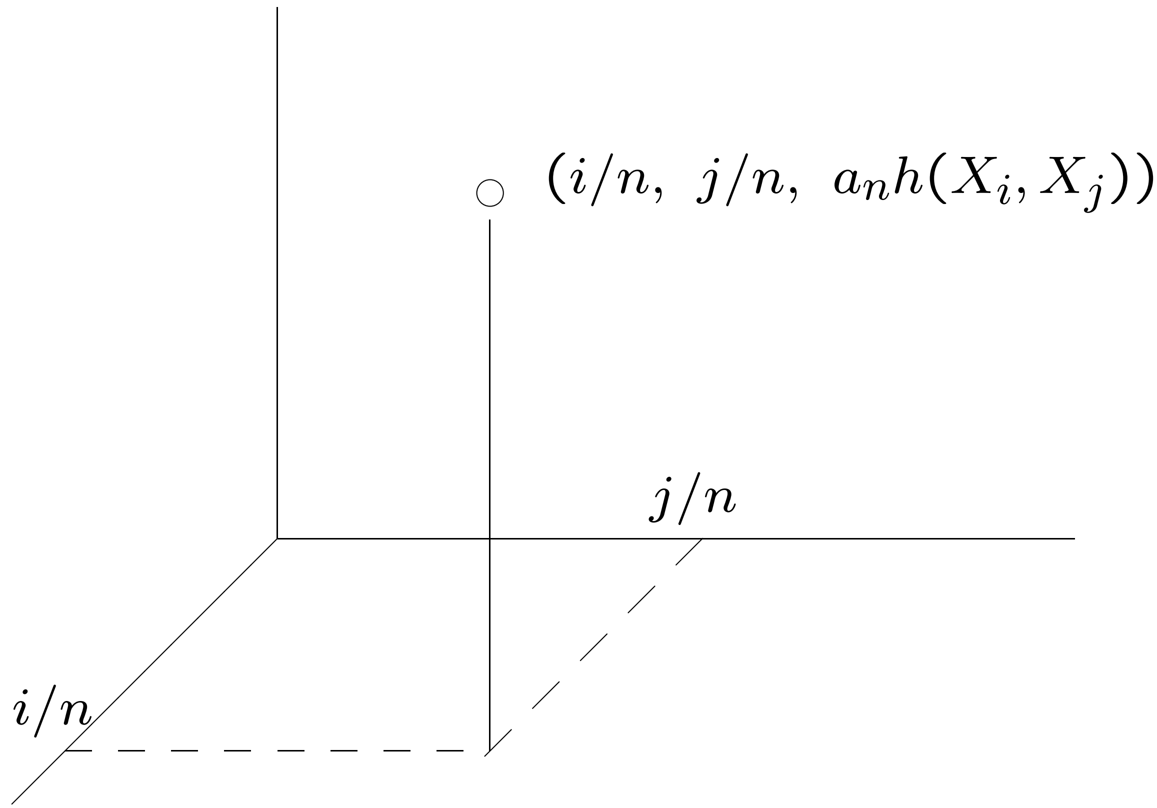
3. As $n \to \infty$, for all $x > 0$,

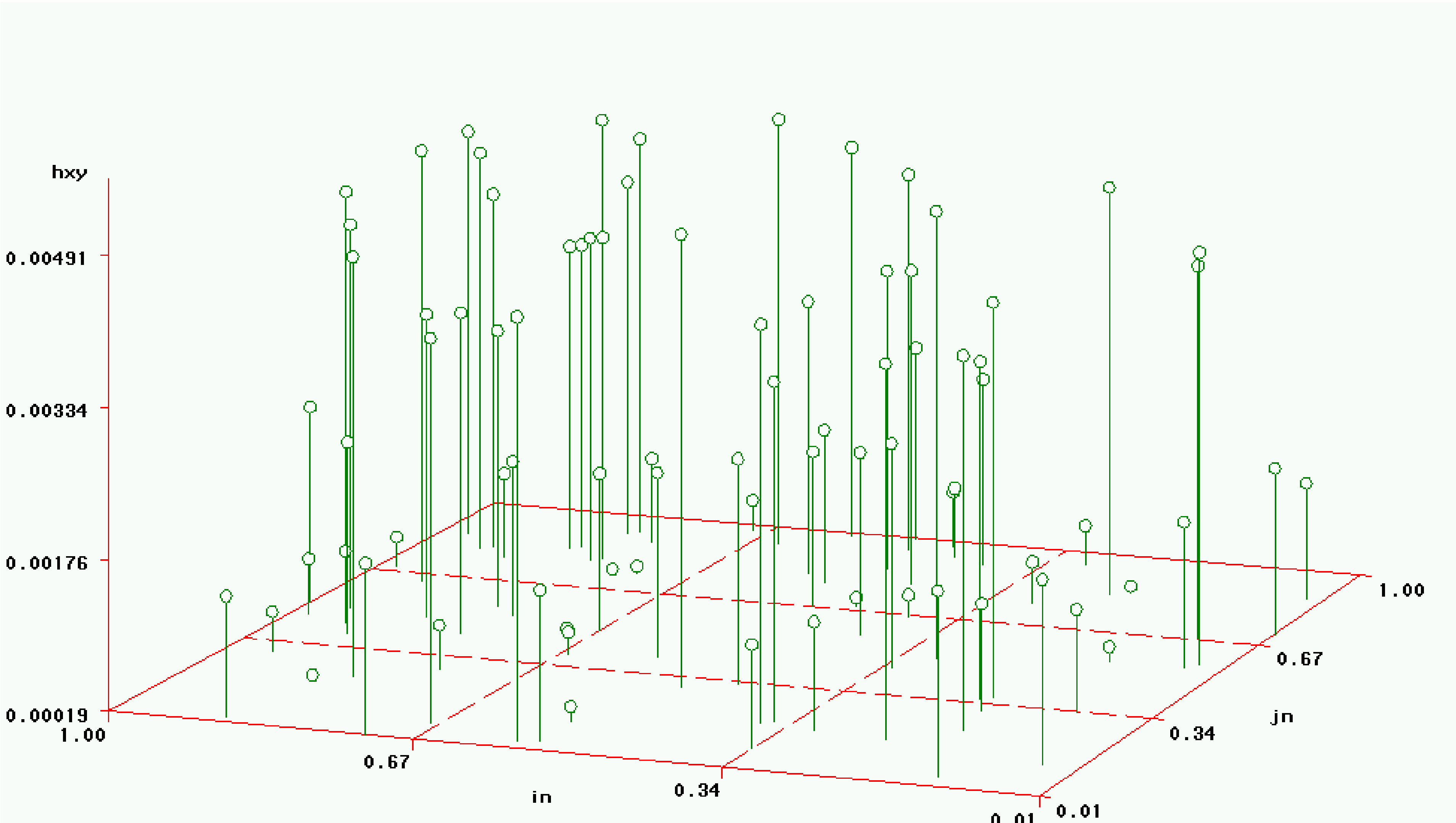
$$n^3 P\left[a_n h(X_1, X_2) \le x,\ a_n h(X_1, X_3) \le x\right] \to 0$$

In analogy with extremal processes we can define the point process

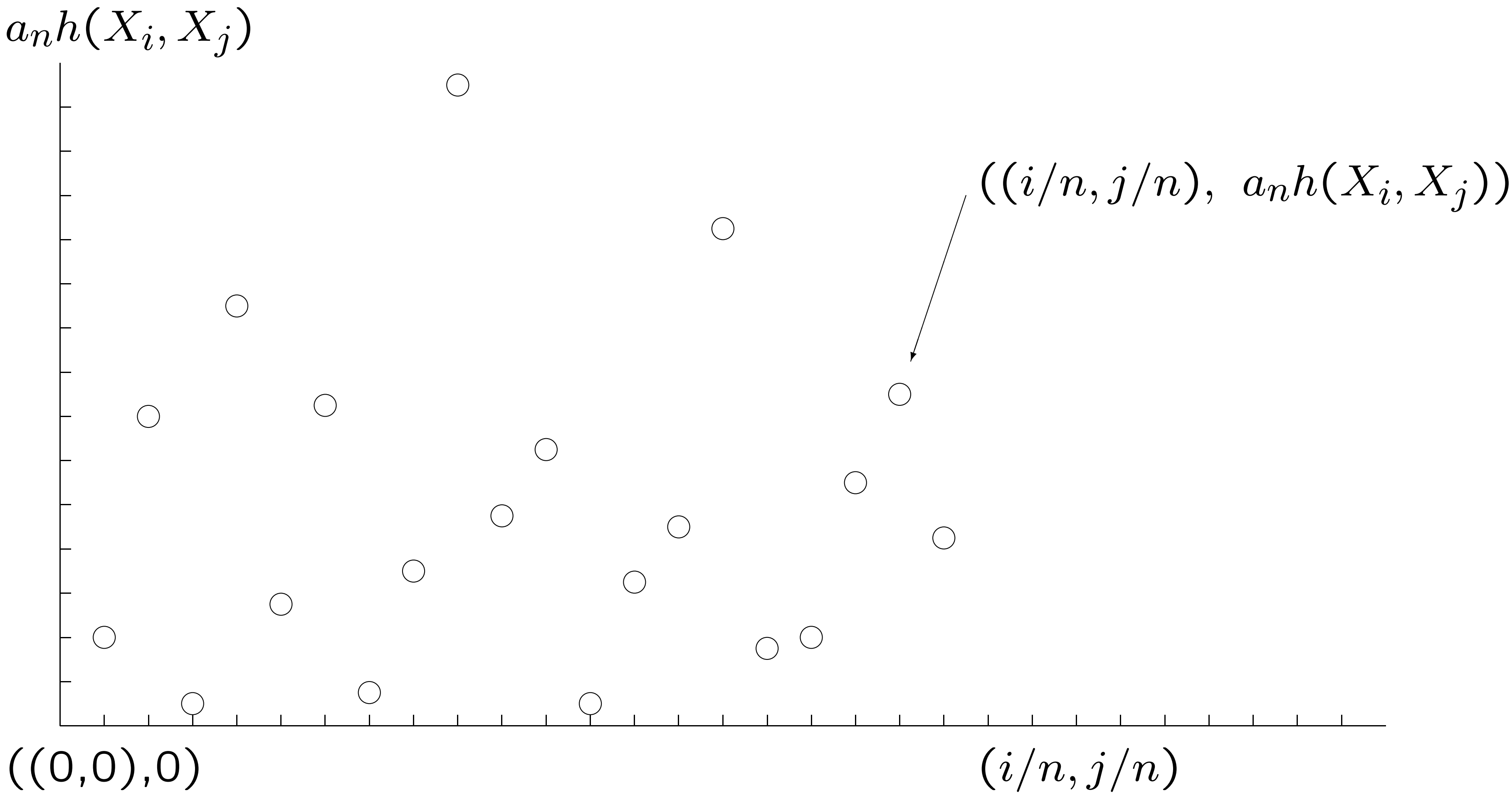
$$N_n(A\times B) = \#\{i < j \in \{1\ldots n\} : \Big((i/n,\ j/n), a_n h(X_i, X_j)\Big) \in A\times B\}$$

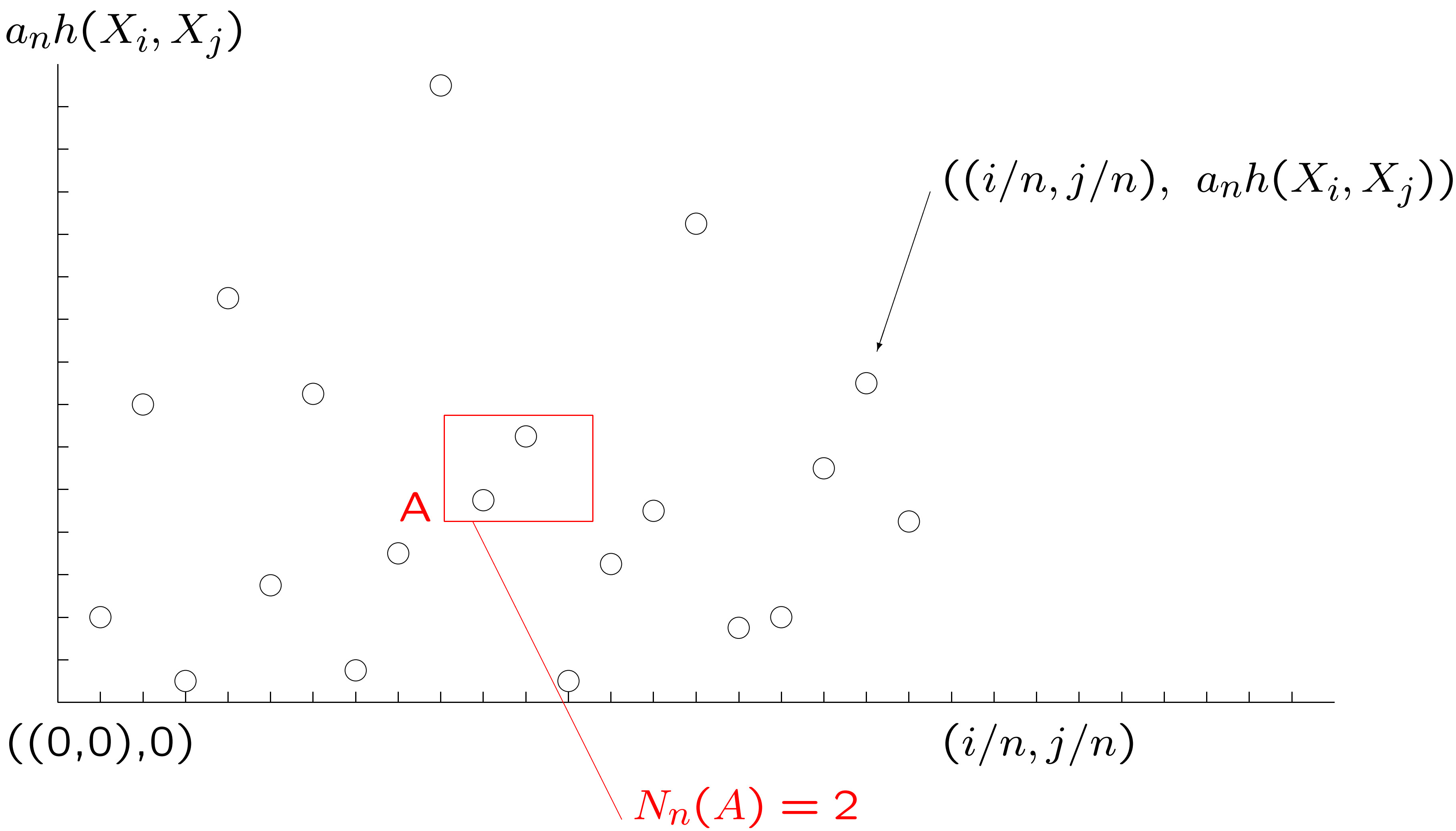
and look for its weak limit.


$(i/n,\ j/n,\ a_n h(X_i, X_j))$
j/n
i/n

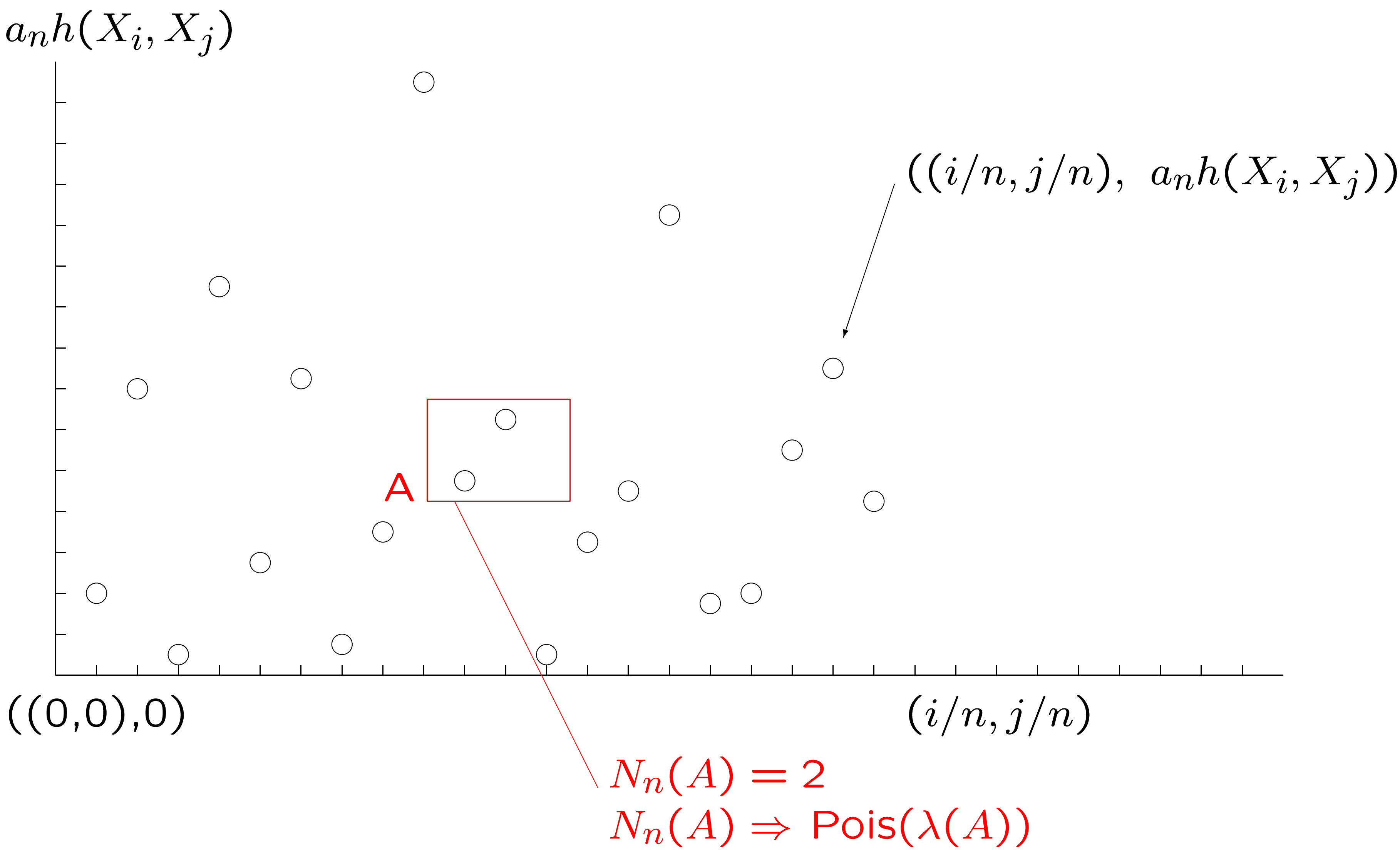


$h(X_i, X_j) = |X_i - X_j| < 0.05$ for uniform X_i.


$a_n h(X_i, X_j)$
$((i/n, j/n),\ a_n h(X_i, X_j))$
((0,0),0)
$(i/n, j/n)$



$a_n h(X_i, X_j)$
$((i/n, j/n),\ a_n h(X_i, X_j))$
A
((0,0),0)
$(i/n, j/n)$
$N_n(A) = 2$


$a_n h(X_i, X_j)$
$((i/n, j/n),\ a_n h(X_i, X_j))$
A
((0,0),0)
$(i/n, j/n)$
$N_n(A) = 2$
$N_n(A) \Rightarrow \text{Pois}(\lambda(A))$

Theorem. [*]*Assume* $P[h(X_1, X_2) \le x] = L(x^{-1})x^{\alpha}$ *and other conditions. Then*

$$N_n \Rightarrow N$$

where N *is a point process on*

$$\{(x, y):\ 0 < x \le 1,\ 0 < y < x\} \times \Re^{+}$$

with intensity

$$\eta\left((a_1, b_1] \times (a_2, b_2] \times (a_3, b_3]\right) = 2(b_1 - a_1)(b_2 - a_2)(b_3^{\alpha} - a_3^{\alpha})$$

[*]*Poisson limits for* U*–statistics*, AD, Herold Dehling, Thomas Mikosch, Olimjon Sharipov, *Stoch. Proc. Appl.* **99**, 137-157, (2002).

How does this help in estimating α?

How does this help in estimating α?

There are several point estimates from several contexts.

- Takens' estimator for the correlation dimension.

- The spatial K-function near zero.

- Hill estimator.

How does this help in estimating α?

There are several point estimates from several contexts.

- Takens' estimator for the correlation dimension.
- The spatial K-function near zero.
- Hill estimator.

We can see how well each of these potential estimators performs.

Takens' estimator for the correlation dimension

As an alternative to the Grassberger-Procaccia estimator, Takens * introduced a dimension estimator motivated by the maximum likelihood principle. Assume

$$P(\|\mathbf{X}_1 - \mathbf{X}_2\| \le x) = x^{\alpha}\,, \quad \text{for } 0 \le x \le \delta.$$

Takens' estimator modified with $\delta_n = \delta / a_n$.

$$\widehat{\alpha}_T = \left[\frac{- \sum_{i=2}^{n} \sum_{j=1}^{i-1} \log(\|\mathbf{X}_i - \mathbf{X}_j\| / \delta_n) I_{[0,\delta_n]}(\|\mathbf{X}_i - \mathbf{X}_j\|)}{\sum_{i=2}^{n} \sum_{j=1}^{i-1} I_{[0,\delta_n]}(\|\mathbf{X}_i - \mathbf{X}_j\|)} \right]^{-1}.$$

*Takens, F. (1985) In *Lecture Notes in Math.* **1125**, pp. 99–106

By a continuous mapping argument, exploiting a representation in terms of gamma variables, and by simple facts, we identify the limit distribution for the modified Takens' estimator:
$\widehat{\alpha}_T^{-1}$ has asymptotic expectation

$$\alpha^{-1}$$

and variance

$$\alpha^{-2}\left[P(N(\delta^{\alpha})=0)+E[N(\delta^{\alpha})I_{\{N(\delta^{\alpha}\})>0}]^{-1}\right].$$

As $\delta \to 0$, the variance is of the order $\alpha^{-2}[1+\delta^{\alpha}]$. Note that it does not shrink to 0.

Poisson convergence of the K-function

In the spatial analysis of point patterns the K-function is used as a measure of spatial dependence*. A sample version of it is given by the U-statistic

$$K_n(\delta) = \sum_{i=2}^{n} \sum_{j=1}^{i-1} I_{[0,\delta]}(a_n \|\mathbf{X}_i - \mathbf{X}_j\|) \,.$$

Thus we have the kernel

$$h(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \,,$$

and so we may conclude that $K_n(\delta) = N_n(\mathbf{E}_1 \times [0, \delta])$ converges in distribution to a Poisson random variable with mean δ^α.

*Cressie (1993) *Statistics for Spatial Data.* Wiley, New York.

More generally, the K_n-processes converge in distribution in $M_p(\mathbf{R}_+)$ to a Poisson process K with mean measure $\alpha\, x^{\alpha-1}\, dx$:

$$K_n(\cdot) = N_n(\mathbf{E}_1 \times \cdot) = \sum_{i=2}^{n} \sum_{j=1}^{i-1} \varepsilon_{a_n|\mathbf{X}_i - \mathbf{X}_j|}(\cdot) \stackrel{d}{\rightarrow} K(\cdot)\,. \qquad (1)$$

Writing $K(\delta) = K([0,\delta])$, it follows that there is distributional convergence near zero,

$$(K_n(\delta))_{\delta \geq 0} \stackrel{d}{\rightarrow} (K(\delta))_{\delta \geq 0}\,.$$

The continuous mapping theorem for $\mathbf{D}[\Delta_0, \Delta_1]$ with $0 < \Delta_0 < \Delta_1 < \infty$ mapping to $\mathbf{C}[0, b]$ yields convergence of squared error between the data and a linear fit to logarithms;

$$\begin{aligned} B_n &= \left(\int_{\Delta_0}^{\Delta_1} (\log^+ K_n(\delta) - (\beta_0 + \beta \log \delta))^2 \, d\delta\right)_{\beta \in [0,b]} \\ &\xrightarrow{d} B = \left(\int_{\Delta_0}^{\Delta_1} (\log^+ K(\delta) - (\beta_0 + \beta \log \delta))^2 \, d\delta\right)_{\beta \in [0,b]}, \end{aligned}$$

in $\mathbf{C}[0, b]$.

Another application of the continuous mapping shows that the LS minimizer $\beta(n)$ of B_n on $[0, b]$ converges to the minimizer $\widehat{\beta}$ of B on $[0, b]$ (where $\log^+ x = \log(\max(1, x))$):

$$\begin{aligned}
\beta(n) &= \frac{\int_{\Delta_0}^{\Delta_1} (\log\delta - \overline{\log\delta})(\log^+ K_n(\delta) - \overline{\log^+ K_n(\delta)})\, d\delta}{\int_{\Delta_0}^{\Delta_1} (\log\delta - \overline{\log\delta})^2\, d\delta} \\
\xrightarrow{d} \widehat{\beta} &= \alpha + \text{ an expression in } (\alpha, K, \Delta_0, \Delta_1). \qquad (2)
\end{aligned}$$

So the best linear fit to extremes in the K-function is an asymptotically biased estimator of α.

Hill estimation of α.

Write

$$h_{(1)} \leq \cdots \leq h_{(n(n-1)/2)}$$

for the order statistics of the sample $h(\mathbf{X}_i, \mathbf{X}_j)$, $i = 2, \ldots, n, j = 1, \ldots, i-1$. A classical estimator of α in the univariate case is *Hill's estimator** given by

$$\widehat{\alpha}_{n,m} = -\left(\frac{1}{m} \sum_{i=1}^{m} \log(h_{(i)}/h_{(m)})\right)^{-1}$$

for $m \geq 1$;

Theorem. *Under regular variation conditions, if $m = m_n \to \infty$ and $\sqrt{m_n}/n \to 0$, then Hill's estimator is consistent, i.e. $\widehat{\alpha}_{n,m} \xrightarrow{P} \alpha$.*

*Hill, B.M. (1975)*Ann. Statist.* **3**, 1163–1174.

- Takens' estimator.

$$\alpha_T = \left[\frac{-\sum_{i=2}^{n}\sum_{j=1}^{i-1} \log(\|\mathbf{X}_i - \mathbf{X}_j\|/\delta_n) I_{[0,\delta_n]}(\|\mathbf{X}_i - \mathbf{X}_j\|)}{\sum_{i=2}^{n}\sum_{j=1}^{i-1} I_{[0,\delta_n]}(\|\mathbf{X}_i - \mathbf{X}_j\|)}\right]^{-1}.$$

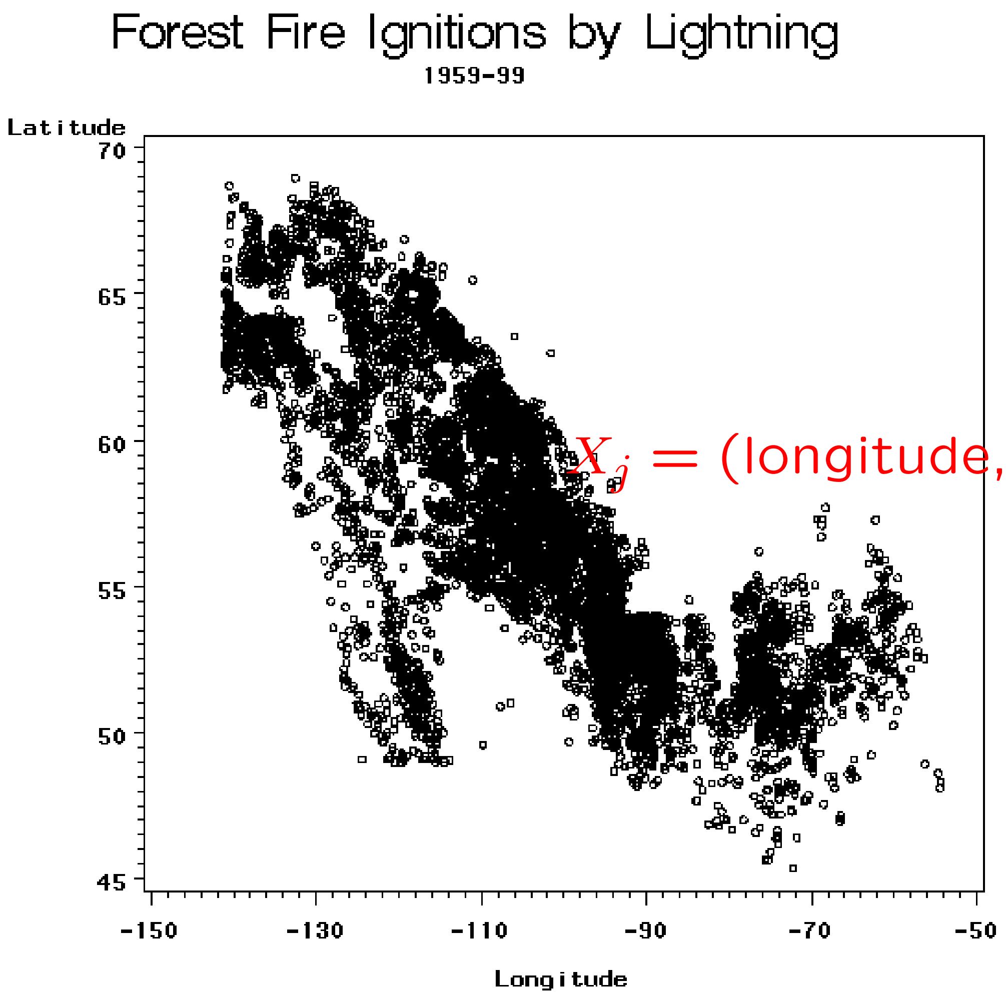
- The spatial K-function near zero.

$$\beta(n) = \frac{\int_{\Delta_0}^{\Delta_1} (\log\delta - \overline{\log\delta})(\log^+ K_n(\delta) - \overline{\log^+ K_n(\delta)})\, d\delta}{\int_{\Delta_0}^{\Delta_1} (\log\delta - \overline{\log\delta})^2\, d\delta}$$

- Hill's estimator.

$$\alpha_{n,m} = -\left(\frac{1}{m}\sum_{i=1}^{m} \log(h_{(i)}/h_{(m)})\right)^{-1}$$

4. Application to data.



$X_j =$ (longitude,latitude,detection date)

To illustrate the estimators, we can compute the estimators of α on this data.

We assume the space-time data on ignitions to be iid observations from a single density.

$$h((x,y,t)_1,(x,y,t)_2) = \sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(t_1-t_2)^2}$$

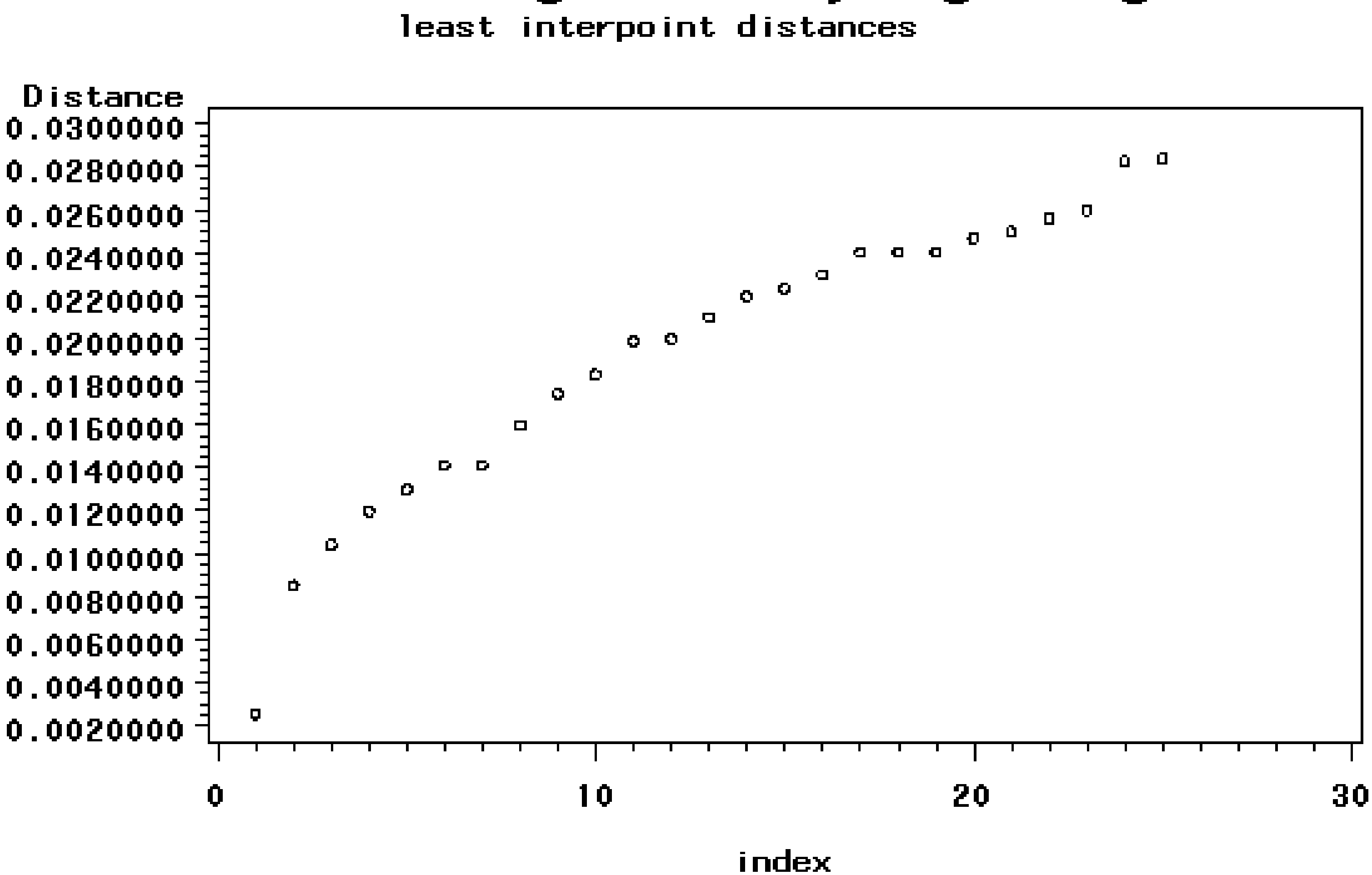
Here x is longitude, y is latitude, and t is time in years$*$1000.

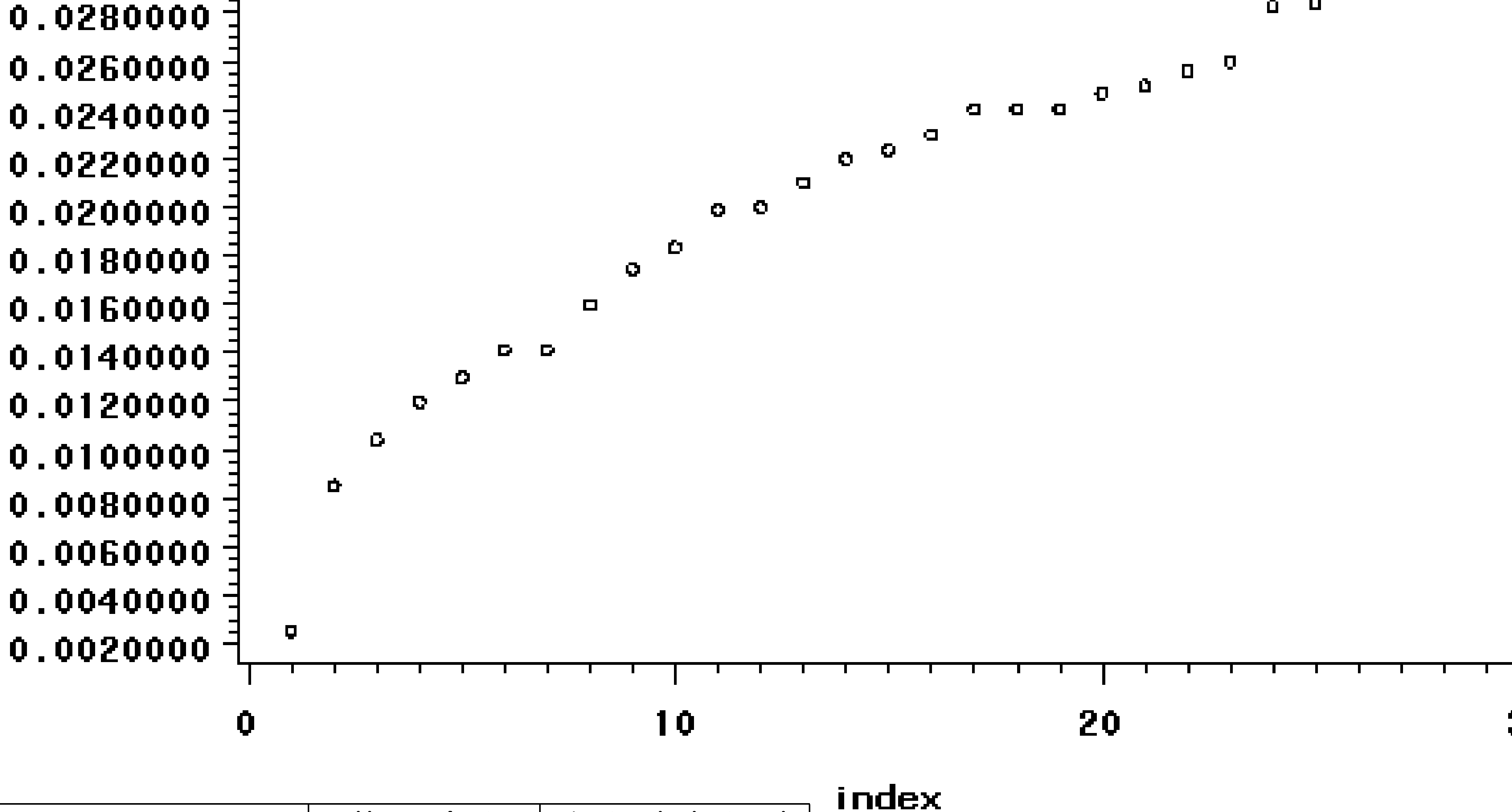
As the index is estimated on just a few small inter-point distances; the estimate will effectively be determined by the most intense "areas".

Inter-point distances of 0 were deleted (38 out of 8050).

Conditional on using the same number of extreme values, Takens' and Hill's estimators yield the same values. The Takens estimator employs essentially a fixed number of minimal inter-point values, the Hill estimator a slowly increasing number.

Forest Fire Ignitions by Lightning





	all points	1st deleted
Takens $\delta = 0.03$	1.90	2.32
K function	1.54	2.71
Hill $m = 25$	2.11	2.61

Here we seem to have (approximately)

$$P(\|\mathbf{X}_1 - \mathbf{X}_2\| \leq x) \sim x^2$$

Here we seem to have (approximately)

$$P(\|\mathbf{X}_1 - \mathbf{X}_2\| \le x) \sim x^2$$

$$1/\left[\widehat{\alpha}^{-1} + 2\sqrt{\widehat{\alpha}^{-2}(1+\delta^{\alpha})}\right] = 0.63$$

Here we seem to have (approximately)

$$P(\|\mathbf{X}_1 - \mathbf{X}_2\| \leq x) \sim x^2$$

Simulating 3-dimensional standard normal . . .

Takens/Hill	K fn	
2.5	2.9	
3.5	2.4	
3.8	2.9	
4.1	4.7	
3.6	3.5	
3.50	3.28	averages

Here we seem to have (approximately)

$$P(\|\mathbf{X}_1 - \mathbf{X}_2\| \leq x) \sim x^2$$

Simulating 3-dimensional standard normal . . .

Takens/Hill	K fn	
2.5	2.9	
3.5	2.4	
3.8	2.9	
4.1	4.7	
3.6	3.5	
3.50	3.28	averages
3.00		exact

Here we seem to have (approximately)

$$P(\|\mathbf{X}_1 - \mathbf{X}_2\| \leq x) \sim x^2$$

Some from 2-dimensional normal . . .

Takens/Hill	K fn	
2.0	1.4	
3.2	2.4	
3.5	5.0	
1.7	1.4	
2.1	2.1	
2.50	2.16	averages
2.00		exact

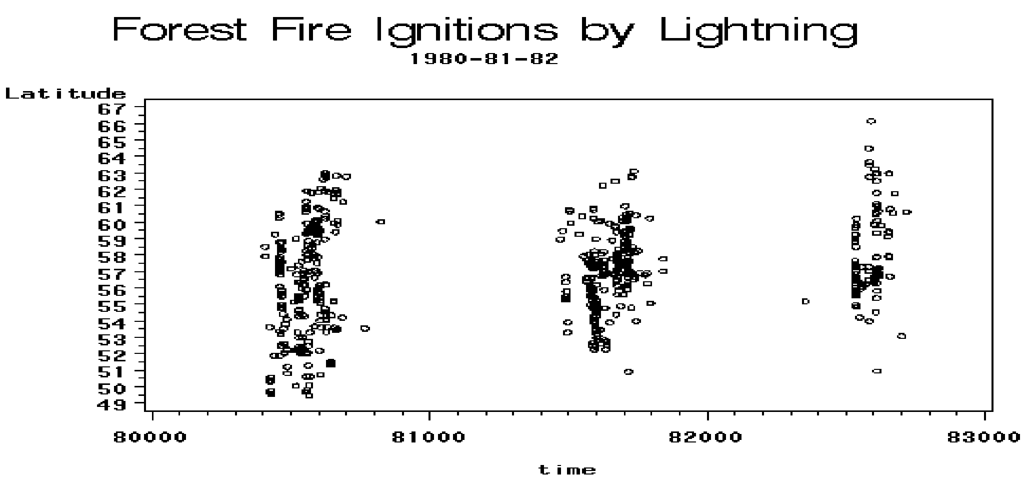
It seems that $P(\|\mathbf{X}_1 - \mathbf{X}_2\| \leq x)$ for the ignition data is more consistent with $\alpha = 2$ than $\alpha = 3$.

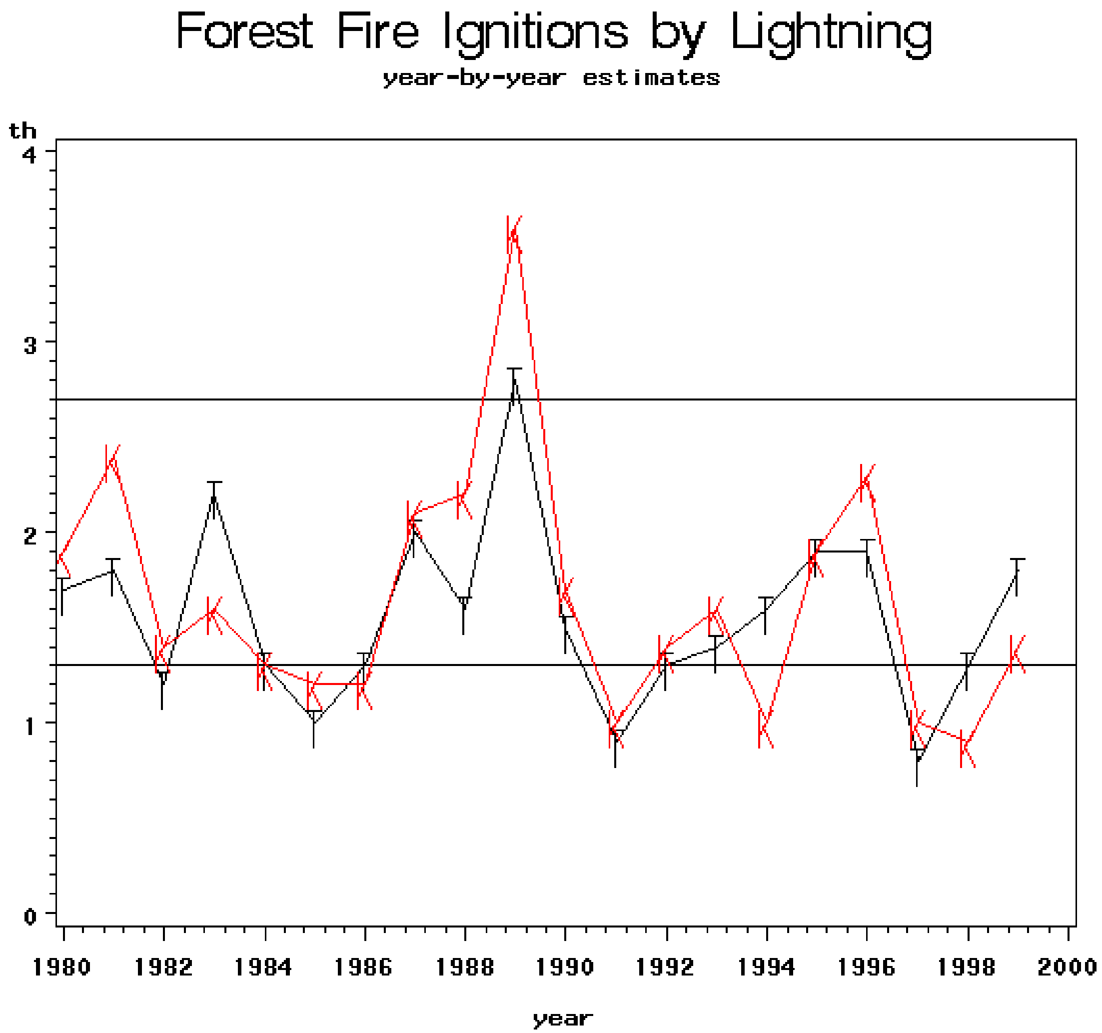
The best approximate CI we have is $[0.63, \infty]$.

It seems that $P(\|\mathbf{X}_1 - \mathbf{X}_2\| \leq x)$ for the ignition data is more consistent with $\alpha = 2$ than $\alpha = 3$.

The best approximate CI we have is $[0.63, \infty]$.

If we treat each year (summer) as an independent sample, we can compute an estimate for each year, and then compute an approximate CI based on the independent estimates.





$\overline{\alpha} = 1.57$

$[1.34,\ 1.79]$

How to interpret $\overline{\alpha} = 1.57$?

This observation seems to indicate that (in the heart of most intense lightning storms) that the process of ignitions seems to behave more like a random process in dimension 1.5 rather than a random process in dimension 3. This could arise, for example, if an ignition spawned "daughter" ignitions (either by "spotting" or by clustering of the underlying lightning strikes) only along a branching path downwind of the initial site.

In a more practical vein, this value and $P[h(X_1, X_2) \leq x] \simeq x^{\alpha}$ can be employed to estimate the chance of a second ignition in close proximity to the first.

Thanks to the organizers for a stimulating conference.