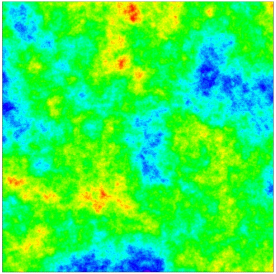
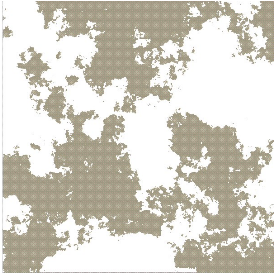




CLT for the volume of excursion sets of stationary associated random fields

Joint work with Vadim Demichev

Judith Schmidt | 6.9.2013
Probability, Analysis and Geometry, Ulm

Overview

- Motivation
- Random fields and their characteristics
- Excursion sets of random fields
- CLT about the volume of excursion sets for growing excursion level
- Proof of the CLT
- Example: Gaussian random field
- Example: Random field on lattices

Motivation

A.V. Ivanov and N.N. Leonenko proved in their book ”Statistical analysis of random fields” a CLT for isotropic Gaussian random fields.

Random fields

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $T \subseteq \mathbb{R}^d$. A collection of random variables $X = \{X(t), t \in T\}$ indexed by elements of T is called a random field.

A random field $X = \{X(t), t \in T\}$ is measurable if for each $A \in \mathcal{B}(\mathbb{R})$ it holds that $\{(\omega, t) : X(\omega, t) \in A\} \in \mathcal{F} \otimes \mathcal{B}(T)$, where $\mathcal{F} \otimes \mathcal{B}(T)$ is the product σ-algebra of $\mathcal{F}$ and $\mathcal{B}(T)$.

A random field $X = \{X(t), t \in T\}$ is called stationary if all finite-dimensional distributions of X are invariant with respect to translations, that is, for all $k \in \mathbb{N}$, $t_1, \dots, t_k$, $s \in T$ it holds that

$$(X(t_1 + s), \dots, X(t_k + s)) \stackrel{d}{=} (X(t_1), \dots, X(t_k)).$$

Random fields

A random field $X = \left\{X(t), t \in \mathbb{R}^d\right\}$ is called positively (**PA**) or negatively (**NA**) associated if

$$\mathrm{Cov}\left(f\left(X_I\right), g\left(X_J\right)\right) \geq 0 \quad (\leq 0, \text{ resp.})$$

for all finite disjoint subsets $I, J \subset \mathbb{R}^d$, and for any bounded coordinatewise non–decreasing functions $f : \mathbb{R}^{\mathrm{card}(I)} \to \mathbb{R}$, $g : \mathbb{R}^{\mathrm{card}(J)} \to \mathbb{R}$, where $X_I = \{X(t), t \in I\}$, $X_J = \{X(t), t \in J\}$.

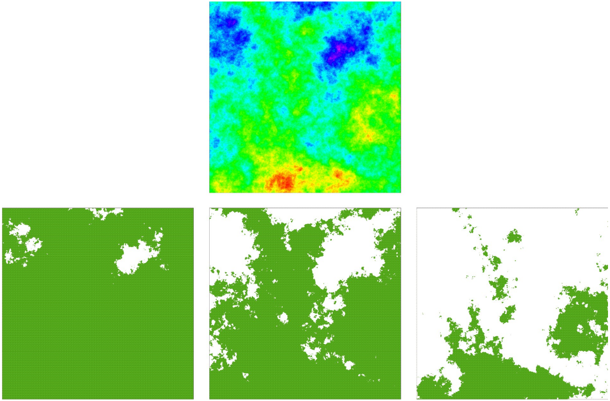
Excursion set

Let $W_n = [0, n]^d$, $n \in \mathbb{N}$ be a sequence of observation windows. Let $v_d(B)$ be the volume of a measurable set $B \in \mathcal{B}(\mathbb{R}^d)$ and let $\mathbb{1}(C)$ denote the indicator function of a set C. Let $|\cdot|$ denote the Euclidean norm in $\mathbb{R}^d$.

An excursion set of random field X at level $u \in \mathbb{R}$ is defined by

$$A_u(X, T) = \{t \in T : X(t) \geq u\}.$$

Excursion set



Centered Gaussian random field on $[0, 1]^2$,
$r(t) = \exp(-\|t\|_2 / 0.3)$,
Levels: $u = -1.0, 0.0, 1.0$

Excursion set

Fix a sequence of increasing excursion levels $\{u_n\}$, $u_n \to \infty$.
The volume of the excursion set of X at level u_n is

$$S_n = v_d(A_{u_n}(X, [0, n]^d)) = \int_{[0,n]^d} \mathbb{1}(X(t) \geq u_n)\, dt.$$

If X is stationary then by Fubini's theorem, the mean of S_n is

$$\mathrm{E}[S_n] = n^d\, \mathbb{P}(X(0) \geq u_n).$$

Similarly, the variance of S_n can be rewritten as

$$\mathrm{Var}(S_n) = \int_{[0,n]^d} \int_{[-x,n-x]^d} \mathrm{Cov}(\mathbb{1}(X(0) \geq u_n), \mathbb{1}(X(t) \geq u_n))\, dtdx.$$

Theorem

Let $X = \{X(t), t \in \mathbb{R}^d\}$ be a measurable, stationary and positively associated random field, with bounded density f and continuous covariance function $\mathrm{Cov}(X(x), X(0))$ satisfying the following conditions

1. $\mathrm{Var}(S_n) \to \infty$ as $n \to \infty$.
2. It exists $\mu > 3d$ such that $|\mathrm{Cov}(X(0), X(t))| = \mathcal{O}(|t|^{-\mu})$ as $|t| \to \infty$.
3. It exists $\{u_n\} \subset \mathbb{R}, u_n \to \infty$ such that $\delta_n := \frac{n^d \gamma^{\frac{2}{3}}(u_n)}{(\mathrm{Var}(S_n))^{\frac{1}{3}(\mu+3)}} \to 0$, as $n \to \infty$, where $\gamma(x) = \sup_{t \geq x} f(t)$.

Then one gets

$$\frac{S_n - \mathrm{E}[S_n]}{\sqrt{\mathrm{Var}(S_n)}} \xrightarrow{d} X \sim \mathrm{N}(0,1).$$

Proof of the CLT

Under the above assumptions, one can find a sequence $m_n \to \infty$ such that for $n \to \infty$

$$\frac{m_n^d}{\mathrm{Var}(S_n)} \to 0 \text{ and } \frac{n^d}{(\mathrm{Var}(S_n))^2}\gamma^{\frac{2}{3}}(u_n)m_n^{d-\frac{\mu}{3}} \to 0.$$

To see this, consider for example

$$m_n = \max\left\{\sqrt{\mathrm{Var}(S_n)}, \delta_n^{\frac{1}{\mu-3}}\,\mathrm{Var}(S_n)\right\}.$$

Set $q_n^d = \sqrt{\mathrm{Var}(S_n)m_n^d}$ and $r_n = \left\lfloor \frac{n}{q_n} \right\rfloor$. Note that for $n \to \infty$

$$\frac{m_n}{q_n} \to 0, \frac{q_n^d}{\mathrm{Var}(S_n)} \to 0 \text{ and } \frac{n}{q_n} \to \infty.$$

It holds that $\mathrm{Var}(S_n) \le const. \cdot n^d$. Set

$$0 < m_n^d \le q_n^d \le \mathrm{Var}(S_n).$$

Proof

Define $\widetilde{S}_n = v_d(A_{u_n}(X, [0, r_n q_n]^d))$ and $Z_n = \sum_{k \in \mathbb{Z}^d \cap [1, r_n]^d} \xi_{n,k}$,

where $\{\xi_{n,k}\}$ are i.i.d. and $\xi_{n,k} \overset{d}{=} v_d(A_{u_n}(X, [0, q_n]^d))$.

It holds that

$$\frac{\mathrm{Var}(Z_n)}{\mathrm{Var}(S_n)} \to 1, \text{ as } n \to \infty \ (*).$$

It suffices to show that

$$\frac{\widetilde{S}_n - \mathsf{E}[\widetilde{S}_n]}{\sqrt{\mathrm{Var}(S_n)}} \overset{d}{=} \frac{Z_n - \mathsf{E}[Z_n]}{\sqrt{\mathrm{Var}(S_n)}}, \quad n \to \infty \tag{1}$$

and that $\frac{Z_n - \mathsf{E}[Z_n]}{\sqrt{\mathrm{Var}(S_n)}} \xrightarrow{d} N(0,1)$ as $n \to \infty$ (2)

Proof

To show (1) it is sufficient to show that

$$\left| \mathsf{E}\left[\exp\left(it\frac{\widetilde{S}_n - \mathsf{E}[\widetilde{S}_n]}{\sqrt{\mathrm{Var}(S_n)}}\right)\right] - \mathsf{E}\left[\exp\left(it\frac{Z_n - \mathsf{E}[Z_n]}{\sqrt{\mathrm{Var}(S_n)}}\right)\right]\right| \to 0, n \to \infty.$$

This can be shown by using Newman's inequality, $(*)$ and the requirement for m_n.

To show (2) one can apply the CLT of Lindeberg by replacing $\mathrm{Var}(S_n)$ with $\mathrm{Var}(Z_n)$ by $(*)$.

Example: Gaussian random field

Let X be a standard Gaussian random field with $\mathrm{Cov}(X(x), X(0)) = \exp(-|x|)$. Then the requirements of the theorem are fulfilled for the level $u_n^2 = \log\log n$.
In this case it holds that

$$\mathrm{Var}(S_n) = \int_{[0,n]^d} \int_{[0,n]^d} \frac{1}{2\pi} \int_0^{r(t_1,t_2)} \frac{1}{\sqrt{1-r^2}} \exp\left(-\frac{u_n^2}{1+r}\right) dr dt_1 dt_2,$$

where $r(t_1, t_2) = \exp(-|t_1 - t_2|)$. This term is greater than or equal to

$$\frac{1}{2\pi} d!(1-\exp(-n))n^d \exp(-u_n^2)$$

and thus the first requirement is shown.

Example: Gaussian random field

To show that

$$\delta_n := \frac{n^d \, \gamma^{\frac{2}{3}}(u_n)}{(\mathrm{Var}(S_n))^{\frac{1}{3}(\mu+3)}} \to 0, \text{ as } n \to \infty,$$

note that

$$\gamma(u_n) = \frac{1}{2\pi} \, \exp\left(-\frac{1}{2} \, u_n^2\right).$$

It holds that

$$n^{-\frac{1}{3}\mu d} \, \exp\left(\frac{1}{3}(\mu+2)u_n^2\right) \to 0,$$

as $n \to \infty$ and for $u_n^2 = \log\log n$. Thus it follows that $\delta_n \to 0$, as $n \to \infty$.

Example: Random field on lattices

Consider a random field $X = \{X(t), t \in \mathbb{Z}^d\}$. Let X be a stationary and positively associated random field with bounded density f. In that case, $\mathrm{E}[S_n] \to \infty$ implies $\mathrm{Var}(S_n) \to \infty$.
To show this, note that due to the stationarity we get

$$\mathrm{Var}(S_n) = \sum_{k \in \mathbb{Z}^d \cap [0,n]^d} \sum_{l \in \mathbb{Z}^d \cap [0,n]^d} \mathrm{Cov}(\mathbb{1}(X(l-k) \geq u_n), \mathbb{1}(X(0) \geq u_n)).$$

Splitting this sum and taking the advantage of the association then the term above is greater than or equal to

$$(n+1)^d \ \mathrm{Var}(\mathbb{1}(X(0) \geq u_n)).$$

This variance can be easily calculated and we get

$$\mathrm{Var}(S_n) \geq (n+1)^d \, \mathbb{P}(X(0) \geq u_n) \, P(X(0) < u_n).$$

Example: Random field on lattices

For the exponential distribution with parameter $\delta < \frac{\mu}{\mu+3}$, set $u_n = d\log(n)$. Note that

$$\varphi(u_n) := \mathbb{P}(X(0) \geq u_n) = \exp(-\delta u_n) \text{ and } \gamma(u_n) = \delta\varphi(u_n).$$

Take into account that,

$$\mathrm{Var}(S_n) \geq n^d \varphi(u_n)\mathbb{P}(X(0) < u_n),$$

and therefore,

$$\frac{n^d\ \gamma^{\frac{2}{3}}(u_n)}{\left(\mathrm{Var}(S_n)\right)^{\frac{1}{3}(\mu+3)}} \leq \delta^{\frac{2}{3}} n^{-\frac{1}{3}d\mu} \left(\varphi(u_n)\right)^{-\frac{1}{3}(\mu+3)} \left(\mathbb{P}(X(0) < u_n)\right)^{1/3(\mu+3)}.$$

Thus, $\delta_n \to 0$ and $\mathrm{E}(S_n) \to \infty$, as $n \to \infty$.

Open problem

What happens if $\mathrm{E}S_n \to \lambda$, as $n \to \infty$?
To discuss this question consider a random field on lattices, where X_1, X_2,... are i.i.d. random variables. Therefore,

$$S_n = \sum_{i=1}^{n} \mathbb{1}(X(i) \geq u_n).$$

Thus, S_n is the sum of Bernoulli-distributed random variables and therefore binomial distributed with parameters n and $\mathbb{P}(X(0) \geq u_n)$.
If $\mathrm{E}S_n = n\,\mathbb{P}(X(0) \geq u_n) \to \lambda$ for $n \to \infty$ and $P(X(0) \geq u_n) \to 0$, then $S_n \xrightarrow{d} Z \sim \mathrm{Poi}(\lambda)$, as $n \to \infty$, see the Poisson limit theorem.

References

- A.V. Ivanov and N.N. Leonenko. Statistical analysis of random fields. Kluwer, Dordrecht, 1989.
- A. Bulinski, E. Spodarev, and F. Timmermann. Central limit theorems for the excursion sets volumes of weakly dependent random fields. Bernoulli, 18(1):100 - 118, 2012.
- E. Spodarev. Limit theorems for excursion sets of stationary random fields. submitted, 2013.
- R.J. Adler and J.E. Taylor. Random Fields and Geometry. Springer, New York, 2007.
- L. Heinrich. In Spodarev, E. (ed.): Stochastic Geometry, Spatial Statistics and Random Fields. Asymptotics Methods, Lecture Notes in Mathematics, volume 2068, chapter Asymptotic methods in statistics of random point processes, pages 115 - 150. Springer, Berlin, 2013.
- G. Last and M.D. Penrose. Probability Theory and Related Fields, volume 150, chapter Poisson process Fock space representation, chaos expansion and covariance inequalities, pages 663 - 690. Springer, Berlin, Heidelberg, 2011.

Thank you for your attention!