Definitions & Limit Theorems for Spatial Statistics

Spatial Stat RIT Talk, Eric Slud, Sept. 25, 2017

1 Background: Process Definitions

<u>**Definition:**</u> A **random process** $X_t = X_t(\omega)$ with index-set T is a family of random variables (measurable real-valued functions) defined on the same probability space $(\Omega, \mathcal{F}, P)$.

The index-sets $T = \mathbb{R}^d$ and $\mathbb{Z}^d$ are of greatest interest in Spatial Statistics, and the process is then called a **d-dimensional random field**.

Examples: (a) scan data (environmental remote sensing, biomedical imaging) generally come in systematic blocks ('windows') in $\mathbb{R}^d$ or $\mathbb{Z}^d$.

(b) in geostatistics, e.g. ground-based rainfall measurements or mineral prospecting data or spatial econometrics, data come as $X(\mathbf{t}_k)$, $k = 1, \ldots, N$, at irregular spacings.

<u>**Definitions:**</u> when $T = \mathbb{R}^d$ or $\mathbb{Z}^d$, the random field $(X(\mathbf{t}),\ \mathbf{t} \in T)$ is called **stationary** if for any finite $A \subset T$ and $\mathbf{s} \in T$, $(X(\mathbf{s}+\mathbf{t}) :\ \mathbf{t} \in A) \overset{\mathcal{D}}{=} (X(\mathbf{t}) :\ \mathbf{t} \in A)$

covariance-stationary or **wide-sense stationary** if

$$EX(\mathbf{t}) = \mu \qquad \text{and} \qquad \mathrm{Cov}(X(\mathbf{t}), X(\mathbf{t}+\mathbf{s})) \equiv C(\mathbf{s}) \qquad \text{do not depend on } \mathbf{t}$$

intrinsic if $E\{X(\mathbf{t}+\mathbf{s}) - X(\mathbf{t})\}^2 \equiv \gamma(\mathbf{s})$ does not depend on $\mathbf{t}$

Examples: (i) partial sum processes $\sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} Z_{i_1,i_2,\ldots,i_d}$ of iid random variables are *intrinsic*, not (strict or wide-sense) stationary

(ii) Brownian motion (Wiener process) and Brownian sheet in $\mathbb{R}^d$ is intrinsic (stationary-increment in $d = 1$) not stationary.

Random Field Examples: (1°) Gaussian & (2°) Markov Random Fields

2 Limit Theorems under Weak-Dependence Conditions

General Problem: when will Laws of Large Numbers or CLT's hold for

$$\sum_{\mathbf{1} \le \mathbf{t} \le \mathbf{n}} X(\mathbf{t}) \qquad \textit{or} \qquad \int_{\mathbf{t} \in D} X(\mathbf{t})\, \mathbf{dt} \qquad \textit{or} \qquad \sum_{k=1}^{N} X(\mathbf{t}_k) \qquad ?$$

2.1 Ergodicity & Mixing

When D_N with $\mathrm{diam}(D_N) \to \infty$ is an increasing and convex sequence of bounded sets in $T = \mathbb{R}^d$ or $\mathbb{Z}^d$ and $X(\mathbf{t})$ is stationary with 2nd moments, then generally (SLLN, extension of Birkhoff Ergodic Theorem, see Gaeton & Guyon Appendix B Thm. B.1)

$$S_N = \frac{1}{|D_N|} \sum_{\mathbf{1} \le \mathbf{t} \in D_N} X(\mathbf{t}) \to Y \quad a.s. \text{ and } L^2 \ , \quad Y \text{ translation-invariant}$$

Ergodicity is the property that **the only translation-invariant events have probability 0 or 1**, usually derived from Mixing. When it holds, $Y = E(X(\mathbf{0})$ is constant a.s.

2.2 Mixing Conditions

Definition: a stochastic process is **(strong-) mixing** if whenever $A(r), B(r)$ are index-sets such that $d(A, B) \ge r$ and $r \to \infty$, and

$$\mathcal{A}(r) = \sigma(X(\mathbf{s}) : \ \mathbf{s} \in A(r)), \ \ \mathcal{B}(r) = \sigma(X(\mathbf{t}) : \ \mathbf{t} \in B(r))$$

then

$$\sup_{A(r), B(r),\, F \in \mathcal{A}(r),\, G \in \mathcal{B}(r)} |P(F \cap G) - P(F)P(G)| \equiv \alpha(r) \to 0 \quad \text{as} \quad r \to \infty$$

The property of **$\boldsymbol{\varphi}$-mixing** says

$$\sup_{A(r), B(r),\, F \in \mathcal{A}(r),\, G \in \mathcal{B}(r)} |P(G|F) - P(G)| \equiv \varphi(r) \to 0 \quad \text{as} \quad r \to \infty$$

Stationary Random Fields on $\mathbb{Z}^d$ with finite variances are known to have the spectral representation of covariances:

$$C(\mathbf{s}) \;=\; \int_{(-\pi,\pi]^d} e^{it\underline{\lambda}}\, dF(\underline{\lambda})$$

When $X(\mathbf{t})$ is Gaussian, mixing holds for $X(\mathbf{t})$ if F has a continuous and positive density f (Rosenblatt 1985 book, and many other places), implying in particular $\sup_{\mathbf{s}:\,|\mathbf{s}|\ge r}\, C(\mathbf{s}) \to 0$ as $r \to \infty$.

But rates of convergence to 0 for $\alpha(r)$ or $\varphi(r)$ can be hard to compute, even for Gaussian examples with explicit parametric covariance functions like $C(\mathbf{s}) \;=\; \exp(-\,a\|\mathbf{s}\|^b)$.

Further Facts about Ergodicity and Mixing:

Suppose that $X(\mathbf{t})$ is stationary, and g is a measurable mapping from the whole random field to $\mathbb{R}$, $g : \mathbb{R}^T \mapsto \mathbb{R}$, and

$$Y(\mathbf{s}) \;\equiv\; g(\{X(\mathbf{s}\,+\,\mathbf{t}) \;:\;\; \mathbf{t} \in T\}) \quad , \qquad \text{for all} \quad \mathbf{s} \in T$$

Then $Y(\cdot)$ is an ergodic (respectively strong-mixiing) stationary random field if $X(\cdot)$ is.

Examples of this sort, with $X(\mathbf{t})$ an iid family of random variables for $\mathbf{t} \in T = \mathbb{Z}^d$, provide a class of random fields extending the Moving Average or Linear processes of Time Series, e.g. by

$$Y(\mathbf{s}) \;=\; \sum_{\mathbf{t}\in\mathbb{Z}^d} c_{\mathbf{t}}\, X(\mathbf{s}+\mathbf{t}) \qquad , \qquad \sum_{\mathbf{t}\in\mathbb{Z}^d} |c_{\mathbf{t}}| < \infty$$

Ergodic stationary $X(\mathbf{t})$ therefore has the property that for any $h : \mathbb{R} \to \mathbb{R}$ with $E(h(X(\mathbf{0}))^2) < \infty$,

$$\frac{1}{|D_N|} \sum_{\mathbf{t}\in D_N} h(X(\mathbf{t})) \;\stackrel{a.s.}{\to}\; E(h(X(\mathbf{t})) \quad \text{as} \quad \text{diam}(D_N) \to \infty$$

Similar result holds for functions $K(X(\mathbf{t}),\, X(\mathbf{t}+\mathbf{u}))$ used in estimating covariances.

3 Statistics – Increasing-Window Asymptotics

Generalized Method of Moment Estimators (GMMs): Suppose the probability law of stationary random-field data $(X(\mathbf{t}) : \mathbf{t} \in D_N)$ is *parametric* based on unknown $\vartheta \in \Theta \subset \mathbb{R}^p$.

In most spatial examples, likelihoods are very hard to compute, so we can try estimation based on GMMs. Suppose $G : \mathbb{R} \mapsto \mathbb{R}^k$ is a known function for which $E(G(X(\mathbf{t}))) = q(\vartheta)$ can be computed explicitly or closely approximated as continuous functions of the unkown ϑ.

Then estimate by solving for ϑ components the relations

$$\frac{1}{|D_N|} \sum_{\mathbf{t} \in D_N} G(X(\mathbf{t})) \; = \; q(\hat{\vartheta}) \qquad (*)$$

Similar equations involving functions $K(X(\mathbf{t}), X(\mathbf{t}+\mathbf{u}))$ at pairs of locations can be used in estimating parameter components from covariances.

Defining the estimators and establishing consistency is a topic for SLLN (or WLLN is generally good enough) as long as functions G over-determining ϑ through $q(\vartheta)$ can be found.

Looking for reference distributions is a matter of proving CLT's for the LHS of (*) and applying the Delta Method (when $q(\cdot)$ is a smooth function).

4 Examples of Known CLT's and Weak LLNs

(I) Suppose stationary data $X(\mathbf{t}_k)$ are observed, $k = 1, \ldots, N$, for which

$$\inf_{j,k:\, j \neq k} \|\mathbf{t}_j - \mathbf{t}_k\| \geq \Delta > 0 \quad , \qquad \sup_{\mathbf{s}:\, \|\mathbf{s}\| \geq r} C(\mathbf{s}) \to 0 \text{ as } r \to \infty$$

Then it can be checked easily by counting and bounding covariance terms that $\text{Var}(\sum_{k=1}^N X(\mathbf{t}_k)) = o(N^2)$, and Chebychev ineq. implies the Weak Law of Large Numbers (convergence to 0 in Probability) of $N^{-1} \sum_{k=1}^N (X(\mathbf{t}_k) - \mu)$, where $\mu = E(X(\mathbf{0}))$.

If $\sum_{\mathbf{s} \in \mathbb{Z}^d} |C(\mathbf{s})| < \infty$, then check as in Time Series case that Var$= O(N)$.

Various CLTs and LLNs are proved under Mixing Conditions. Restrict attention to strong-mixing (not necessarily stationary) random fields on $\mathbb{Z}^d$ with mean 0, finite variances.

Guyon (Prop. B.1 in Appendix B of his book) states a result that includes the following, for

$$S_N \;=\; \sum_{k=1}^{N} X(\mathbf{t}_k) \quad , \qquad \sigma_N^2 \;=\; \mathrm{Var}(S_N)$$

Theorem 1 *If there exists* $\delta > 0$ *such that* $\sup_{\mathbf{t}} E|X(\mathbf{t})|^{2+\delta} < \infty$ *and*

$$\alpha(r) \;=\; o(r^{-d}) \quad \textit{and} \qquad \sum_{r \geq 1} r^{d-1} (\alpha(r))^{\delta/(2+\delta)} \;<\; \infty$$

and if $\liminf_N \, \sigma_N^2/N \;>\; 0,$ *then* $S_N \,/\, \sigma_N \;\stackrel{\mathcal{D}}{\to}\; \mathcal{N}(0,1).$

CLTs under stationarity and φ-mixing have been proved by various authors (Rosenblatt 1970, Deo 1975), e.g.

Theorem 2 *Let* $\mathbf{n} = (n_1, \ldots, n_d)$ *and* $N = |\mathbf{n}| = \prod_{j=1}^{d} n_j$ *and* $D_N = \{\mathbf{t} \in \mathbb{Z}^d : \;\; 1 \leq t_j \leq n_j \;\; for \;\; j = 1, \ldots, d\}$*, and let* $X(\mathbf{t})$ *be a stationary random field on* $\mathbb{Z}^d$ *with mean 0 and finite variance.*

If $X(\cdot)$ *is* φ*-mixing with* $\sum_{r \geq 1} r^{d-1} \, (\varphi(r))^{1/2} \;<\; \infty,$ *then*

$$\sigma_N^2/N \;\to\; \sum_{\mathbf{t} \in \mathbb{Z}^d} C(\mathbf{t}) \;\equiv\; \sigma^2$$

If $\sigma > 0$*, then as* $N \to \infty,$ $S_N \,/\, \sigma_N \;\stackrel{\mathcal{D}}{\to}\; \mathcal{N}(0,1).$