Random sets

Limit theorems

Ilya Molchanov

University of Bern, Switzerland

Definitions

- $\mathcal{F}$ family of closed subsets in locally compact Hausdorff second countable space $\mathbb{E}$ (usually $\mathbb{E} = \mathbb{R}^d$).
- X is a random closed set in $\mathbb{E}$.
- $X_1, X_2, \dots$ are i.i.d. copies of X.
- $\|X\| = \sup\{\|x\| : x \in X\}$ the norm of X.
- $\rho_{\mathrm{H}}(K, L) = \inf\{r > 0 : K \subset L^r, L \subset K^r\}$ the Hausdorff metric, where K^r and L^r are r-neighbourhoods of K and L.

Operations with sets

- Two basic operations with sets:

- Minkowski sum $K + L = \{x + y : \; x \in K, \, y \in L\}$
- union $K \cup L$

- If K and L are singletons, their Minkowski sum is the usual addition.

- If $K = (-\infty, x]$ and $L = (-\infty, y]$, then the union corresponds to the maximum $x \vee y$.

However

- For convex compact sets the Minkowski addition is similar to the usual addition, but $K + X = L$ does not necessarily have a solution X.

- The Minkowski sum of two convex (non-compact) sets is not necessarily closed.

- For non-convex sets the Minkowski addition may have more peculiar properties. E.g. $K + X = K + Y$ does not imply $X = Y$ and $X + X$ is not necessarily $2X$, e.g. if $X = \{0, 1\}$.

- The union is idempotent operation, i.e. $X \cup X = X$.

Characters

- Closed sets with a particular (commutative) operation form a semigroup.

- Idea:

  - define a family of characters, i.e. homomorphisms between this semigroup and the complex unit disk with multiplication $\chi(x+y)=\chi(x)\chi(y)$;
  - take the expectation in the unit complex disk $\mathbf{E}\,\chi(X)$ for random X.

- The family of characters should be rich enough, namely if $\chi(x)=\chi(y)$ for all χ, then necessarily $x=y$.

Implementation

- $\mathbb{R}_+$ with addition: $\chi_t(x) = e^{-tx}$. Then $\mathbf{E}\, e^{-t\xi}$ is Laplace transform of random variable ξ.

- $\mathbb{R}$ with addition: $\chi_t(x) = e^{itx}$. Then $\mathbf{E}\, e^{it\xi}$ is the characteristic function.

- Closed sets and unions: $\chi_K(X) = \mathbb{1}_{X\cap K=\emptyset}$. Then $\mathbf{E}\,\mathbb{1}_{X\cap K=\emptyset} = 1 - T_X(K)$ is the avoidance functional of X.

- Minkowski sums of convex compact sets. Support function

$$h(K,u) = \sup\{\langle x,u\rangle : \ x \in K\}\,.$$

Then $h(K+L,u) = h(K,u) + h(L,u)$. If all sets contain the origin, the characters are

$$\chi_u(K) = e^{-h(K,u)}.$$

How to normalise the sum?

- Multiplicative normalisation

$$\frac{X_1 + \cdots + X_n}{a_n} \quad \text{or} \quad \frac{X_1 \cup \cdots \cup X_n}{a_n}$$

- Additive normalisation

$$X_1 + \cdots + X_n + K_n \quad \text{or} \quad X_1 \cup \cdots \cup X_n + K_n \ (\text{or } \cup K_n)$$

- Note that subtraction is not well defined and is generally impossible.

However $X_1 + \cdots + X_n$ generally "gets bigger" if n increases!

Two main examples

- Minkowski sums
- Unions

Minkowski sums of compact sets

- Reduction to sums of convex compact sets

Shapley-Folkman-Starr Theorem. If $K_1, \dots, K_n$ are compact subsets of $\mathbb{R}^d$, then

$$\rho_{\mathrm{H}}(K_1 + \cdots + K_n, \operatorname{co}(K_1 + \cdots + K_n)) \le \sqrt{d} \max_{1 \le i \le n} \|K_i\| \,.$$

- Special case $K_1 = \cdots = K_n = K$

$$\rho_{\mathrm{H}}(\frac{1}{n}(K+\cdots+K), \frac{1}{n}\operatorname{co}(K+\cdots+K)) \le \frac{\sqrt{d}}{n}\|K\| \to 0 \quad \text{as } n \to \infty \,.$$

- K is infinitely divisible for Minkowski sums (i.e. $K = L_n + \cdots + L_n$ for all $n \ge 2$) if and only if K is convex.

Normalised sums

$X_1, X_2, \dots$ are i.i.d. copies of random compact set X

- Idea:

$$h(\frac{1}{n}(X_1 + \cdots + X_n), u) = \frac{1}{n}(h(X_1, u) + \cdots + h(X_n, u))$$

The right-hand side is the arithmetic sum of functions on the unit sphere, and so it is possible to use the law of large numbers in the Banach space of functions.

Expectation

- A real function $f(u)$, $u \in \mathbb{R}^d$, is a support function of a convex compact set if and only if f is homogeneous (i.e. $f(cu) = cf(u)$) and sublinear (i.e. $f(u+v) \leq f(u) + f(v)$).

- Assume X is a random compact set with $\mathbf{E}\,\|X\| < \infty$ (X is called integrably bounded).

- $h(X, u)$ is homogeneous and sublinear almost surely.

- The expectation $\mathbf{E}\,h(X, u)$ is also homogeneous and sublinear. Thus, there exists a compact convex set denoted by $\mathbf{E}\,X$ such that $h(\mathbf{E}\,X, u) = \mathbf{E}\,h(X, u)$.

- Note that $\mathbf{E}\,X$ is always convex.

Definition using selections

- Assume that the underlying probability space is non-atomic. Then $\mathbf{E}\,X$ can be equivalently defined as

$$\mathbf{E}\,X = \{\mathbf{E}\,\xi : \ \xi \in X \text{ a.s.}\} .$$

So defined $\mathbf{E}\,X$ is also always convex on non-atomic probability spaces.

- If $\|X\|$ is integrable, then all selections are integrable. It suffices however to require that X possesses at least one integrable selection. Then X is not necessarily compact and $\mathbf{E}\,X$ may be unbounded.

Law of large numbers

Theorem (Artstein-Vitale). If $\mathbf{E}\,\|X\| < \infty$ (X is integrably bounded), then

$$\rho_{\mathrm{H}}(n^{-1}(X_1 + \cdots + X_n), \mathbf{E}\,X) \to 0 \quad \text{as } n \to \infty\,.$$

❑ Proof.

- If X is convex, apply LLN for support functions in $C(\mathbb{S}^{d-1})$.
- In general, use the Shapley-Folkman-Starr theorem

$$\rho_{\mathrm{H}}(n^{-1}(X_1+\cdots+X_n), n^{-1}\mathrm{co}\,(X_1+\cdots+X_n)) \le \frac{\sqrt{d}}{n} \max_{1\le i\le n} \|X_i\|$$

Use the fact that $\|X\|$ is integrable.

An application: zonoids

- If $X = [0, \xi]$ is a random segment, then $\mathbf{E}\, X$ is a zonoid (zonoid of ξ).

- The law of large numbers implies that zonoids appear as limits for sums of segments (zonotopes).

- Zonoid of ξ does not characterise the distribution of ξ, but it carries useful information about it (multivariate statistics).

- The lift zonoid of ξ is $\mathbf{E}\, Y$ for $Y = [0, (1, \xi)] \subset \mathbb{R}^{d+1}$. The lift zonoid characterises uniquely the distribution of ξ, since

$$h(\mathbf{E}\, Y, (t, u)) = \mathbf{E} \max\{0, (t + \langle \xi, u \rangle)\}$$

for all $t \in \mathbb{R}$ suffices to determine the distribution of $\langle \xi, u \rangle$ for each u.

Centring

$$\frac{X_1 + \dots + X_n}{n} \to \mathbf{E}\, X$$

- It is not possible to "subtract" $\mathbf{E}\, X$.
- If $\mathbf{E}\, X = \{0\}$, then X is a singleton with mean zero.

So need to deal with centring in a different way. Compare

$$\sqrt{n}\left(\frac{\xi_1 + \dots + \xi_n}{n} - \mathbf{E}\,\xi\right) \to \mathcal{N}(0, \sigma^2)$$

and

$$\sqrt{n}\rho(\frac{\xi_1 + \dots + \xi_n}{n}, \mathbf{E}\,\xi) \to |\mathcal{N}(0, \sigma^2)|\,.$$

Central Limit theorem

Theorem (Weil). Let $X, X_1, X_2, \dots$ be square integrable, i.e. $\mathbf{E}\,\|X\|^2 < \infty$. Then

$$\sqrt{n}\rho_{\mathrm{H}}(n^{-1}(X_1 + \cdots + X_n), \mathbf{E}\,X) \to \sup_{\|u\|=1} |\zeta(u)|\,,$$

where $\zeta(u)$ is a centred Gaussian field on the unit sphere (or ball) with the covariance

$$\mathbf{E}\,\zeta(u)\zeta(v) = \mathbf{E}[h(X,u)h(X,v)] - \mathbf{E}\,h(X,u)\,\mathbf{E}\,h(X,v)\,.$$

Difficulties

- The random field ζ usually cannot be represented as a support function. If this is the case, then necessarily $X = \xi + K$.
  Open problem: Find a useful geometric interpretation of ζ.

- Open problem. Asymptotic expansions of the type

$$\frac{X_1 + \cdots + X_n}{n} = \mathbf{E}\, X + n^{-1/2} Z + \cdots$$

- ζ is multivariate field with a complicated covariance structure, so its maximum does not have a simple distribution.
  Tools: consider values of ζ on a finite subset of $\mathbb{S}^{d-1}$;
  use bootstrap-related techniques.

- Aim: get confidence intervals for $\mathbf{E}\, X$.

An application (Beresteanu & Molinari, 2006)

Incompletely specified models in econometrics. Explanatory variable x, response interval-valued $Y = [y_l, y_u]$.

Define random set $G = \{(y, xy) : \ y \in Y\}$.

Best linear predictor solves

$$\mathbf{E}(y) = \theta_1 + \theta_2 \, \mathbf{E}(x) \, , \quad \mathbf{E}(xy) = \theta_1 \, \mathbf{E}(x) + \theta_2 \, \mathbf{E}(x^2) \, ,$$

where y is a selection of Y. Set-valued best linear predictor is

$$\Theta = \begin{bmatrix} 1 & \mathbf{E}(x) \\ \mathbf{E}(x) & \mathbf{E}(x^2) \end{bmatrix}^{-1} \mathbf{E}\, G \, .$$

The theoretical expectation is then replaced by Minkowski averages for $G_i = \{(y, x_i y) : \ y \in Y_i\}$.

Stability and limit theorems

Random variables $\xi_1, \dots, \xi_n$

$$\frac{\xi_1 + \cdots + \xi_n}{n^\alpha} \to \zeta$$

The limits are necessarily stable, i.e.

$$a^{1/\alpha}\zeta_1 + b^{1/\alpha}\zeta_2 \stackrel{\mathcal{D}}{=} (a+b)^{1/\alpha}\zeta$$

The Gaussian limit appears if $\alpha = 2$.

Although for random sets the limit theorem is formulated differently (for the Hausdorff metric), it is possible to define the stability concept. However, stable random sets are no longer immediately related to the limits, since the limits are not set-valued.

Gaussian and stable random sets

Random convex compact set X is strictly α-stable iff

$$a^{1/\alpha}X_1 + b^{1/\alpha}X_2 \stackrel{\mathcal{D}}{=} (a+b)^{1/\alpha}X$$

The Gaussian case $\alpha = 2$ iff $h(X, u)$ is a Gaussian random function of u.

Theorem. X is α-stable with $\alpha \in [1, 2]$ if and only if $X = \xi + K$ where K is deterministic and ξ is α-stable vector.

Reason (Gaussian case). Consider linear functional

$$s(X) = \frac{1}{\kappa_d} \int_{\mathbb{S}^{d-1}} h(X, u) u du$$

(the Steiner point of X). Then $s(X) = \xi$ is Gaussian and $Y = X - s(X)$ contains the origin a.s., so that $h(Y, u)$ is Gaussian and non-negative, thus deterministic.

Extensions

- law of iterated logarithm
- three series theorem
- renewal theorem
- ergodic theorems
- large deviation results

Unions of random sets

- The natural tool for unions is the capacity functional (because the avoidance functional actually is the Laplace transform in this case).

- Need to figure out the cases when the Laplace transform vanishes. In particular,

$$F_X = \{x : \mathbf{P}\{x \in X\} = 1\}$$

is the set of fixed points for X. Note that $F_X \subset X$ a.s. and $1 - T(K) = 0$ if $K \cap F_X \neq \emptyset$.

- X is said to be infinite divisible for unions if, for each $n \geq 2$,

$$X \stackrel{\mathcal{D}}{=} X_{n1} \cup \cdots \cup X_{nn}$$

for i.i.d. random closed sets $X_{n1}, \ldots, X_{nn}$.

Infinite divisible sets

- **Theorem** (Matheron). X is union infinitely divisible iff

$$T_X(K) = 1 - e^{-\Psi(K)},$$

where Ψ is a completely alternating upper semicontinuous functional (like T_X but with values in $[0, \infty)$) such that $\Psi(K) < \infty$ if $K \cap F_X = \emptyset$.

- Ψ defines a measure ν on $\mathcal{F}$ such that $\Psi(K) = \nu\{F : \ F \cap K \neq \emptyset\}$. If we consider a Poisson process on $\mathcal{F}$ with intensity measure ν, then X is the union of points (actually sets) which are elements of this process.

Sketch of Proof

- If X is infinitely divisible, then $T = 1 - (1 - T_n)^n$ with T_n being a capacity functional.
- $T(K) < 1$ for all K such that $K \cap F_X = \emptyset$.
- $nT_n(K) = n(1-(1-T(K))^{1/n}) \to -\log(1-T(K)) = \Psi(K)$, so that Ψ is completely alternating and upper semicontinuous.
- $T = 1 - e^{\Psi}$ is a capacity functional, and $1 - (1 - T(K))^{1/n} = 1 - e^{n^{-1}\Psi(K)}$ is a capacity functional.

Union-stable random sets

X is said to be union-stable if, for every $n \geq 2$,

$$a_n X \stackrel{\mathcal{D}}{=} X_1 \cup \cdots \cup X_n ,$$

where $X_1, \ldots, X_n$ are i.i.d. and have the same distribution as X.

Theorem. A non-deterministic X is union-stable iff $T_X = 1 - e^{-\Psi}$ with

$$\Psi(sK) = s^{\alpha}\Psi(K) , \quad K \cap F_X = \emptyset,\ s > 0 ,$$

and $sF_X = F_X$.

Idea of the proof

- The stability definition implies that

$$T(K) = 1 - (1 - T(a_n K))^n\,, \quad n \geq 1\,,$$

  i.e. $n\Psi(a_n K) = \Psi(K)$.

- Then $s\Psi(a(s)K) = \Psi(K)$ for all rational s and $\Psi(a(s)a(s_1)K) = \psi(a(ss_1)K)$. Need to conclude that $a(s)a(s_1) = a(ss_1)$.

- This is not immediate! $X \overset{\mathcal{D}}{=} cX$ (and so $T_X(K) = T_X(cK)$) is possible for $c \neq 1$, e.g. zero sets of the Wiener process. However, the stability property excludes this option.

- Show that a is continuous on rational numbers. Then $a(s) = s^\gamma$ with $\gamma = -1/\alpha$.

Examples

- If X does not have fixed points, then $\alpha > 0$. If $\alpha < 0$, then $0 \in F_X$.
- If X is stationary, then $\alpha > 0$.
- $X = (-\infty, \xi]$ is union-stable iff ξ is max-stable.
- If X is Poisson process with intensity measure Λ, then X is infinitely divisible for unions and $\Lambda(K) = \Psi(K)$. So X is union-stable iff $\Lambda(sK) = s^\alpha \Lambda(K)$. E.g. the stationary process in $\mathbb{R}^d$ is stable with $\alpha = d$ (Λ is the Lebesgue measure).

Weak convergence results

$a_n^{-1}(X_1 \cup \cdots \cup X_n)$ weakly converges to a (necessarily union-stable) random closed set Z.

Basic condition: regular variation of $f(x) = T(xK)$ as a function of $x > 0$ at $x \to \infty$. Note that $f(x)$ is regular varying iff

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} = x^\alpha$$

for each $x > 0$.

Extensions

- Similar results exist for convex hulls of sets in terms of their containment functionals. The alternative tool is provided by the support function, since

$$h(\operatorname{co}(K \cup L), u) = \max(h(K,u), h(L,u)) .$$

- It is also possible to work with intersections of random closed sets.

References

[1] Z. Artstein and R. A. Vitale. A strong law of large numbers for random compact sets. *Ann. Probab.*, 3:879–882, 1975.

[2] A. Beresteanu and F. Molinari Asymptotic properties for a class of partially identified models. Working paper, Cornell University, 2006.

[3] S. Li, Y. Ogura, and V. Kreinovich. *Limit Theorems and Applications of Set-Valued and Fuzzy Set-Valued Random Variables*. Kluwer, Dordrecht, 2002.

[4] I. Molchanov. *Theory of Random Sets*. Springer, London, 2005.

[5] W. Weil. An application of the central limit theorem for Banach-space-valued random variables to the theory of random sets. *Z. Wahrsch. verw. Gebiete*, 60:203–208, 1982.