Random sets

Distributions, capacities and their applications

Ilya Molchanov

University of Bern, Switzerland
Definitions

- $\mathbb{E}$ locally compact Hausdorff second countable space (usually $\mathbb{E} = \mathbb{R}^d$)

- $\mathcal{F}$ family of closed subsets of $\mathbb{E}$

- $(\Omega, \mathcal{F}, P)$ probability space

- $X : \Omega \mapsto \mathcal{F}$ is a random closed set if $\{\omega : X(\omega) \cap K \neq \emptyset\} \in \mathcal{F}$ for all $K \in \mathcal{K}$ (compact sets in $\mathbb{E}$).

- In other words, $X$ is measurable with respect to the $\sigma$-algebra on $\mathcal{F}$ generated by families of sets

  \[ \{F \in \mathcal{F} : F \cap K \neq \emptyset\} \text{ for } K \in \mathcal{K}. \]
Simple examples

- \( X = \{\xi\} \) random singleton

- \( X = B_\eta(\xi) \) random ball. Check measurability
  \[
  \{X \cap K \neq \emptyset\} = \{\rho(K, \xi) \leq \eta\}
  \]
  and use the fact that \( \rho(K, \xi) \) is a random variable

- \( X = \triangle_{\xi_1,\xi_2,\xi_3} \) random triangle

- \( X = \{x \in \mathbb{R}^d : \langle x, \xi \rangle \leq t\} \) random half-space

- \( X = \{t \geq 0 : \xi_t = 0\} \) for a continuous stochastic process \( \xi_t \)
Random variables associated with random closed sets

- The norm $\|X\| = \sup\{\|x\| : x \in X\}$ is a random variable (with possibly infinite values).
- For every $x \in \mathbb{E}$ the indicator $\mathbb{I}_X(x)$ is a random variable.
- If $\mu$ is a locally finite Borel measure on $\mathbb{E}$, then $\mu(X)$ is a random variable. This follows directly from Fubini's theorem since $\mu(X) = \int \mathbb{I}_X(x) \mu(dx)$.
If $X$ is a random closed set, then

$$T_X(K) = \mathbb{P}\{X \cap K \neq \emptyset\}$$

is called the capacity functional of $X$. Also write $T(K)$ instead of $T_X(K)$.

Usually $K$ belongs to the family $\mathcal{K}$ of compact sets, but $T_X$ can be extended for more general arguments $K$. 
Simple examples

- If $X = \{\xi\}$ random singleton, then

$$T_X(K) = P\{\xi \in K\}$$

is the probability distribution of $\xi$. This is the only case of an additive $T_X$.

- If $X = B_r(\xi)$ (ball of radius $r$ centred at $\xi$), then

$$T_X(K) = P\{\xi \in K^r\},$$

where $K^r$ is the $r$-neighbourhood of $K$. Note that $K_1^r$ and $K_2^r$ are not necessarily disjoint even if $K_1$ and $K_2$ are.

- If $X = [\xi, \infty)$ on $\mathbb{R}^1$, then

$$T_X(K) = P\{\xi \leq \sup K\}$$
Capacity functional (properties)

- Monotonic $T(K_1) \leq T(K_2)$ if $K_1 \subseteq K_2$.
- Subadditive $T(K_1 \cup K_2) \leq T(K_1) + T(K_2)$.
- Concave $T(K_1 \cap K_2) + T(K_1 \cup K_2) \leq T(K_1) + T(K_2)$.
- Semicontinuous $T(K_n) \downarrow T(K)$ if $K_n \downarrow K$. 
Complete alternation

Successive differences for $T = T_X$ are defined as

$$
\Delta_{K_1} T(K) = T(K) - T(K \cup K_1),
$$
$$
\Delta_{K_n} \cdots \Delta_{K_1} T(K) = \Delta_{K_{n-1}} \cdots \Delta_{K_1} T(K)
$$
$$
- \Delta_{K_{n-1}} \cdots \Delta_{K_1} T(K \cup K_n), \quad n \geq 2.
$$

Then $\Delta_{K_1} T(K) = P\{X \cap K = \emptyset, \ X \cap K_1 \neq \emptyset\}$ and

$$
- \Delta_{K_n} \cdots \Delta_{K_1} T(K)
$$
$$
= P\{X \cap K = \emptyset, \ X \cap K_i \neq \emptyset, \ i = 1, \ldots, n\} \geq 0.
$$

$T$ is said to be completely alternating.
The Choquet theorem

Let $E$ be a locally compact Hausdorff second countable space.

**Theorem:** A functional $T: K \rightarrow [0, 1]$ such that $T(\emptyset) = 0$ is the capacity functional of a (necessarily unique) random closed set in $E$ if and only if $T$ is upper semicontinuous and completely alternating.

**Reason:** $\chi_K(F) = \mathbb{1}_{F \cap K = \emptyset}$ satisfies

$\chi_K(F_1 \cup F_2) = \chi_K(F_1) \chi_K(F_2)$ (compare $e^{it(x+y)} = e^{itx} e^{ity}$).

Then $E \chi_K(X) = 1 - T_X(K)$ is the Laplace transform of the distribution of $X$.

Complete alternation corresponds to the positive definiteness property.
Point processes

- $X$ in $\mathbb{R}^d$ is locally finite if $X \cap K$ is finite for each bounded set $K$.

- A locally finite random closed set is a point process.

- Then $N(K) = \text{card}(X \cap K)$ is a random variable (counting random measure).

- The Choquet theorem implies that the distribution of a simple (no multiple points) point process is uniquely determined by the avoidance probabilities $P\{X \cap K = \emptyset\} = P\{N(K) = 0\}$. 
Special capacities

- If $T(K) = \sup\{f(x) : x \in K\}$, then
  \[
  X = \{x : f(x) \geq \alpha\}
  \]
  with $\alpha$ uniformly distributed on $[0, 1]$.

- If $T(K) = 1 - e^{-\Lambda(K)}$, then $X$ is a Poisson point process with intensity measure $\Lambda$.

- If $T(K) = 1 - e^{-C(K)}$, where $C(K)$ is the Newton capacity. Then $X$ is related to graphs of Wiener processes.
Random sets vs fuzzy sets

A fuzzy set is a function \( f : \mathbb{E} \rightarrow [0, 1] \), so that \( f(x) \) is the “degree of membership” for a point \( x \).

If \( f \) is upper semicontinuous (usually assumed), then it is possible to write

\[
  f(x) = \mathbb{P}\{x \in X\}
\]

for

\[
  X = \{x : f(x) \geq \alpha\}.
\]

However, the distribution of \( X \) contains more information than its one-point covering probabilities.
Extension problems

- $T(K)$ is defined on some family of compact sets $K$.
  Aim: find if it is possible to extend it to a capacity functional on $\mathcal{K}$ and find (all possible) extensions.

- Trivial case: $T(\{x\}) = f(x)$ defined for all singletons $K$. Then define $X = \{x : f(x) \geq \alpha\}$, i.e. define the extension by $T(K) = \sup\{f(x) : x \in K\}$.

- Open problem: $T(\{x, y\}) = f(x, y)$. Characterise all possible functions $f$ such that $T$ is extendable to a capacity functional. If $X$ is stationary, then $T(\{x, y\}) = f(x - y)$. 

Measurability issues

The fundamental measurability theorem implies that the following are equivalent (in case $\mathbb{E}$ is Polish and the probability space is complete):

- $\{X \cap B \neq \emptyset\}$ is measurable event for all Borel $B$;
- $\{X \cap F \neq \emptyset\}$ is measurable event for all closed $F$;
- $\{X \cap F \neq \emptyset\}$ is measurable event for all open $G$;
- The distance function $\rho(y, X)$ is a random variable for all $y$;
- There exists a sequence of random singletons $\{\xi_n\}$ such that $X$ is the closure of $\{\xi_n, n \geq 1\}$;
- The graph $\{(\omega, x); x \in X(\omega)\}$ is measurable set in the product space $\Omega \times \mathbb{E}$. 
Notes to measurability issues

- If $\mathbb{E}$ is locally compact, all these statements are equivalent to the fact that $X$ is a random closed set.

- A random element $\xi$ satisfying $\xi \in X$ a.s. is called a selection of $X$.

- $X : \Omega \mapsto \mathcal{F}$ is called a multifunction, and

$$X^-(B) = \{\omega : X(\omega) \cap B \neq \emptyset\}$$

is its inverse.
Capacity functionals and properties of random sets

- $T_X$ determines the distribution of $X$ and so all features of $X$.

- E.g. if $\mu$ is a $\sigma$-finite measure, then

$$E\mu(X) = \int T_X(\{x\})\mu(dx) = \int \mathbb{P}\{x \in X\}\mu(dx)$$

(Robbins’ theorem).

- In general, the situation is more complicated, e.g. if $\mu$ is a Hausdorff measure.
Convex random sets

$X$ is said to be convex if its realisations are a.s. convex (check that the family of convex closed sets is measurable!).

**Theorem.** $X$ is convex if and only if

$$T_X(K_1 \cap K_2) + T_X(K_1 \cup K_2) = T_X(K_1) + T_X(K_2)$$

for all convex compact sets $K_1, K_2$ such that $K_1 \cup K_2$ is also convex.

**Reason:** If $X$ is convex, then

$$0 = \mathbb{P}\{X \cap K = \emptyset, X \cap K_1 \neq \emptyset, X \cap K_2 \neq \emptyset\} = -\Delta K_2 \Delta K_1 T_X(K).$$

In the reverse direction take $K_1 = [x, z]$ and $K_2 = [z, y]$. 
Weak convergence

- \( X_n \xrightarrow{d} X \) (weakly converges) iff the corresponding probability measures weakly converge.

- This is equivalent to

  \[ T_{X_n}(K) \to T_X(K) \]

  for all compact sets \( K \) such that \( T_X(K) = T_X(\text{Int } K) \) (continuity sets).

- Each sequence of random closed sets in a locally compact space possesses a weak convergent subsequence.
Capacities elsewhere

- \( \varphi(A) \), \( A \subseteq \mathbb{E} \), is a non-additive measure if \( \varphi(\emptyset) = 0 \), \( \varphi(\mathbb{E}) = 1 \) and \( \varphi \) is monotonic. Such non-additive measures (capacities) build a topological space, it is possible to study their random variants, weak convergence etc.

- Game theory: elements of \( \mathbb{E} \) are players, subsets of \( \mathbb{E} \) are coalitions, \( \varphi \) is a game. Dual \( \tilde{\varphi}(A) = 1 - \varphi(A^c) \).

- If \( \varphi(A) = \mathbb{P}\{X \subset A\} \) (containment functional of \( X \)), then \( \tilde{\varphi}(A) = T(A) = \mathbb{P}\{X \cap A \neq \emptyset\} \) is the capacity functional.
Core of non-additive measure

- The core of $\varphi$ is the family of finite-additive measures $\mu$ such that $\mu(A) \geq \varphi(A)$ for all $A$. In terms of random sets $\mathbb{P}\{X \subset A\} \leq \mu(A) \leq \mathbb{P}\{X \cap A \neq \emptyset\}$.

- If $\varphi$ is convex, i.e.

$$\varphi(A \cap B) + \varphi(A \cup B) \geq \varphi(A) + \varphi(B),$$

then the core is non-empty.

- If $\varphi$ is the containment functional of a.s. non-empty random set $X$, then all $\sigma$-additive measures from the core correspond to selections of $X$, i.e. $\mu$ is the distribution of $\xi$ such that $\xi \in X$ a.s.
Selections and comparison

- If $\xi \in X$, then $\mathbb{P}\{\xi \in K\} \leq T_X(K)$ for all $K \in \mathcal{K}$. This is also a sufficient condition for $\xi$ being a selection of $X$.

- If $Y \subset X$ a.s., then $T_Y(K) \leq T_X(K)$ for all $K$.

- However, the inequality between the capacity functionals does not suffice to deduce that $Y \subset X$. One needs inequalities between successive differences $\Delta(\cdots)$ of all orders.
Choquet integral

<table>
<thead>
<tr>
<th>Probability measure $\mu$</th>
<th>Non-additive measure $\varphi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expectation $\int \xi(\omega) \mu(d\omega)$</td>
<td>Choquet integral</td>
</tr>
</tbody>
</table>

Choquet integral is defined for a non-negative function $f$ as

$$\int f \, d\varphi = \int_0^\infty \varphi(\{x : f(x) \geq t\}) \, dt.$$

Compare: if $\mu = P$ is a probability measure and $\xi \geq 0$ a.s., then

$$E \xi = \int_0^\infty P\{\xi \geq t\} \, dt.$$

It is easy to see that

$$\int f \, dT_X = E \sup\{f(x) : x \in X\}.$$
Choquet integral and pricing

- The Lebesgue integral (expectation of a random variable $X$) $\Leftrightarrow$ linear pricing functional.

- Choquet integral $\Leftrightarrow$ sub-additive and comonotone additive, i.e.
  \[
  \int (f + g) d\varphi = \int f d\varphi + \int g d\varphi \text{ if } f \text{ and } g \text{ are comonotonic.}
  \]

Random variables $X$ and $Y$ are comonotonic if
\[
(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0 \text{ a.s. } \omega, \omega' \in \Omega
\]

Note that $f \mapsto \mathbb{E} \sup f(X)$ is comonotonic with respect to $f$ (but not necessarily additive)
Risk measures

\( \xi \in L^\infty \) represents a financial gain. Functional \( \rho : L^\infty \mapsto \mathbb{R} \).

**Translation invariance.** For all \( \xi \in L^\infty \) and \( t \in \mathbb{R} \): \( \rho(\xi + t) = \rho(\xi) - t \).

**Subadditivity.** For all \( \xi_1, \xi_2 \in L^\infty \): \( \rho(\xi_1 + \xi_2) \leq \rho(\xi_1) + \rho(\xi_2) \).

**Positive homogeneity.** For all \( c \geq 0 \) and \( \xi \in L^\infty \): \( \rho(c\xi) = c\rho(\xi) \).

**Monotonicity.** If \( \xi_1 \leq \xi_2 \) a.s., then \( \rho(\xi_1) \geq \rho(\xi_2) \).

Then \( \rho \) is called a coherent risk measure.
Representation of coherent risk measures

If \( \rho \) is coherent and lower semicontinuous in probability, then there exists a convex set \( \mathcal{A} \) of probability measures, such that

\[
\rho(\xi) = \sup\{ E_{\mu} \xi : \mu \in \mathcal{A} \}.
\]

Particularly important families of risk measures are given by Choquet integrals

\[
\rho(\xi) = \int \xi d\varphi.
\]

If \( \varphi \) is the capacity functional \( T_X \) with \( X \subset \Omega \), then

\[
\rho(\xi) = E \sup\{ \xi(\omega) : \omega \in X \}.
\]

Although too simple in the univariate case, this representation is useful for multivariate risk measures.
Belief and plausibility functions

<table>
<thead>
<tr>
<th>Containment functional $\mathbb{P}{X \subset K}$</th>
<th>Belief function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity functional $T_X$</td>
<td>Plausibility functional</td>
</tr>
</tbody>
</table>

Example: Likelihood plausibility $X = \{x : L(x) \geq \alpha\}$, where $L$ is the normalised likelihood.

Aim: Update of belief functions.

Open problem. For a given belief function find a probability measure that dominates it and has the maximal entropy. In other words, find a selection (of a random set) with the maximal entropy.
Capacities in robust statistics

“Contaminated” families of probability measures, e.g.

\[(1 - \varepsilon)P + \varepsilon Q\]

for all probability measures \(Q\) (contaminations).

The corresponding upper probability is a capacity functional

\[T_X(K) = \varepsilon + (1 - \varepsilon)P(K), \quad K \neq \emptyset,\]

so that \(X\) equals the whole space with probability \(\varepsilon\) and otherwise is the singleton with distribution \(P\).
Statistical issues for random sets themselves

- Reference measure (like Lebesgue or Gaussian) is not available.
- Likelihood-based methods are not available (so far).
- Models are scarce.
Random set models

- Point-based models \( X = \{\xi_1, \ldots, \xi_n\} \).

- Convex hull based models \( X = \text{co} (\xi_1, \ldots, \xi_n) \).
  E.g., random segments or triangles.

- More general setting: \( M : \mathbb{R}^m \rightarrow \mathcal{F} \) is a set-valued measurable function on an auxiliary space. Then \( X = M(\zeta) \) becomes a random closed set.
Stationary random sets

- Similar setting $X = M(\zeta)$, but $\zeta$ now is a point process.

- If $\Psi = \{x_1, x_2, \ldots\}$ is a point process and $X_1, X_2, \ldots$ i.i.d. random compact sets, then

$$\bigcup_{x_i \in \Psi} (x_i + X_i)$$

is a germ-grain model. Called Boolean model if the point process is Poisson.
Random fractal sets

Example. Cantor set construction.

At each step we delete the mid-parts of the remaining intervals (starting from $[0, 1]$), so that the left remaining part occupies the $C_1$-share and the right remaining part $C_2$-share of the whole length.

The usual Cantor set appears if $C_1 = C_2 = 1/3$. If $C_1$ and $C_2$ are random and i.i.d. at every step and every interval, then the Hausdorff dimension $s$ of the obtained random fractal set satisfies

$$E(C_1^s + C_2^s) = 1.$$
Sets related to stochastic processes

\[ \zeta_t, t \geq 0, \text{ is a stochastic process.} \]

- **Level set**, e.g.

  \[ X = \{ t \geq 0 : \zeta_t = 0 \} \]

  is a random closed set if \( \zeta \) is sample continuous.

- **Epigraph**

  \[ X = \{(t, x) : x \geq \zeta_t, t \geq 0, x \in \mathbb{R}\} = \text{epi } \zeta \]

  is a random closed set if \( \zeta \) is a.s. lower semicontinuous.
References


