Probability Theory with Continuous Model Theory

H. Jerome Keisler
University of Wisconsin-Madison

July 15, 2025

1. Motivation

Objective: Give an efficient way to apply continuous model theory to study random variables and stochastic processes in probability theory.

The probability literature often works in a setting with an underlying probability space and studies random variables with values in a complete separable metric space. The Loeb measure construction has been successfully applied to probability theory. That worked because Loeb probability spaces have a property called saturation, and good things happen when the underlying probability space is saturated. Continuous logic is well-suited for probability theory because statements in probability theory are often naturally expressed by sets of continuous formulas.

Idea: Work in a continuous pre-metric structure for which:

(1) The elements are random variables on a **saturated** probability space.

(2) The theory is tame from the point of view of Shelah's classification theory.

(3) The important properties in the probability literature are naturally expressible by sets of formulas in continuous logic.

2. Continuous Model Theory

Truth values Formulas have truth values in $[0, 1]$, where 0 means **true**.

Vocabulary: A set of **sorts**. predicate, function, and constant **symbols**.
A **distance predicate** instead of equality in each sort.
Atomic formulas: Same as in first-order logic.
Connectives: Continuous functions from $[0, 1]^n$ into $[0, 1]$.
Quantifiers: $\sup, \inf$.

Signature: vocabulary and modulus of uniform continuity for each symbol.
Metric models: A complete metric space of diameter ≤ 1 in each sort.
Pre-metric models: Pseudo-metrics instead of metrics.

Much of first-order model theory carries over (e.g. compactness, ultraproducts, saturated models, classification theory).

3. The pre-ordering $\mathcal{K} \trianglelefteq \mathcal{M}$

An ultrafilter U is over I is κ-regular if there is a set $X \subseteq U$ of power $|X| = \kappa$ such that each $i \in I$ belongs to only finitely many sets in X.

Definition

Let $\mathcal{K}, \mathcal{M}$ be first-order (1967) or pre-metric structures. $\mathcal{K} \trianglelefteq \mathcal{M}$ iff for every κ and κ-regular ultrafilter U, if $\prod_U \mathcal{M}$ is κ^+-saturated then $\prod_U \mathcal{K}$ is κ^+-saturated.

If $\mathcal{K} \equiv \mathcal{M}$ then $\mathcal{K} \trianglelefteq \mathcal{M}$. So $\trianglelefteq$ is a pre-ordering on complete theories.
$\mathcal{K}$ is $\trianglelefteq$-minimal iff for every κ and κ-regular U, $\prod_U \mathcal{K}$ is κ^+-saturated.

Theorem

(Shelah 1972, Keisler 2024) In either first-order or continuous logic with countable vocabularies, every $\trianglelefteq$-minimal theory is stable, and the first two $\trianglelefteq$-equivalence classes are the class of $\trianglelefteq$-minimal theories and the class of stable non-$\trianglelefteq$-minimal theories.

4. Some notation in probability theory

Probability spaces: $\mathbf{\Omega} = (\Omega, \mathcal{F}, P)$.

Let $(\mathbb{M}, d_{\mathbb{M}})$ be a compact metric space of diameter ≤ 1.
$\mathbb{M}$ is a complete separable metric space.

$(\text{Meas}(\mathbb{M}), \pi_{\mathbb{M}})$ is the space of probability measures on the Borel subsets of $\mathbb{M}$ with the Prohorov metric.
$(\text{Meas}(\mathbb{M}), \pi_{\mathbb{M}})$ is also a compact metric space of diameter ≤ 1.

$L^0(\mathbf{\Omega}, \mathbb{M}) = \{\mathbb{M}\text{-valued random variables (measurable functions) on } \Omega\}$
with the pseudo-metric $\rho_{\mathbb{M}}(x, y) = E(d_{\mathbb{M}}(x, y))$ of **convergence in probability**.
$(L^0(\mathbf{\Omega}, \mathbb{M}), \rho_{\mathbb{M}})$ is complete but can be large and not separable.

law: $L^0(\mathbf{\Omega}, \mathbb{M}) \to \text{Meas}(\mathbb{M})$ where $(\text{law}(x))(B) = P((x^{-1}(B))$.
If $\mathbf{\Omega}$ is atomless, law is continuous and surjective.

5. Saturated probability spaces

Definition

(Hoover and Keisler 1984). $\mathbf{\Omega}$ is saturated if $\mathbf{\Omega}$ is atomless, and for every compact metric space $\mathbb{M}$ and $x, x' \in L^0(\mathbf{\Omega}, \mathbb{M})$ with $\text{law}(x) = \text{law}(x')$, $\forall y \in L^0(\mathbf{\Omega}, \mathbb{M})\ \exists y' \in L^0(\mathbf{\Omega}, \mathbb{M})$ with $\text{law}(x, y) = \text{law}(x', y')$.

Examples Atomless Loeb probability spaces are saturated.
Uncountable powers of $[0, 1]$ are saturated.
The Lebesgue unit interval is not saturated.
If $\mathbb{M}$ is separable and $\mu \in \text{Meas}(\mathbb{M})$, $(\mathbb{M}, \mu)$ is not saturated.

Saturated probability spaces are "very rich". Yeneng Sun showed that random variables with respect to saturated probability spaces have many desirable properties and are useful for proving existence theorems in probability theory.

Arguments from nonstandard analysis that worked for Loeb spaces often carry over for arbitrary saturated probability spaces.

6. The pre-metric structure $\mathcal{R}(\mathbf{\Omega}, \mathcal{C})$

Idea (repeated): Work in a continuous pre-metric structure for which:

(1) The elements are random variables on a **saturated** probability space.
(2) The theory is tame from the point of view of Shelah's classification theory.
(3) The important properties in the probability literature are naturally expressible by sets of formulas in continuous logic.

Let $\mathbf{\Omega}$ be a saturated probability space.
Let $\mathcal{C}$ be a set consisting of a copy of each compact metric space of diameter ≤ 1.

Definition

$\mathcal{R}(\mathbf{\Omega}, \mathcal{C})$ is the pre-metric structure with:

- For each $\mathbb{M} \in \mathcal{C}$, a sort $\mathbb{S}_{\mathbb{M}}$ with universe $L^0(\mathbf{\Omega}, \mathbb{M})$ (thus $2^{\aleph_0}$ sorts)
- and the distance predicate $\rho_{\mathbb{M}}$ of convergence in probability.
- For each $\mathbb{M}, \mathbb{P} \in \mathcal{C}$ and continuous $f : \mathbb{M} \to \mathbb{P}$, a function symbol $\boldsymbol{f}$ where $(\boldsymbol{f}(x))(w) = (f(x(w)))$ for each x of sort $\mathbb{S}_{\mathbb{M}}$ and $w \in \mathbf{\Omega}$.

Ben Yaacov (2013) studied the theory of the analogous structure with only the sort $\mathbb{S}_{[0,1]}$. Then took a different approach—randomizations of metric structures.

7. $\mathcal{R}(\mathbf{\Omega}, \mathcal{C})$ is rich and tame

Theorem ($\aleph_1$-compactness)

If $\mathbf{\Omega}$ is saturated, then $\mathcal{R}(\mathbf{\Omega}, \mathcal{C})$ is $\aleph_1$-compact, that is, every countable finitely satisfiable set of formulas with parameters in $\mathcal{R}(\mathbf{\Omega}, \mathcal{C})$ is satisfiable.

This gives a large variety of existence theorems in probability theory.

Definition

Let T be the set of continuous sentences that are true in $\mathcal{R}(\mathbf{\Omega}, \mathcal{C})$ for every saturated $\mathbf{\Omega}$. Let $\mathcal{L}$ be the signature of T.

Theorem (Tameness)

(i) T is a complete theory. $T = \mathrm{Th}(\mathcal{R}(\mathbf{\Omega}, \mathcal{C}))$ for all saturated $\mathbf{\Omega}$.
(ii) T admits elimination of quantifiers.
(iii) For every countable $\mathcal{L}_0 \subseteq \mathcal{L}$, the $\mathcal{L}_0$-part of T is $\aleph_0$-stable.
(iv) T is $\trianglelefteq$-minimal (but not $\aleph_2$-categorical).

8. Expressive power of $\mathcal{R}(\mathbf{\Omega}, \mathcal{C})$

If $\mathbf{\Omega}$ is saturated, most properties from the probability theory literature can be readily expressed by countable sets of continuous formulas in $\mathcal{R}(\mathbf{\Omega}, \mathcal{C})$.

For example:
Constant r.v.'s: For each $m \in \mathbb{M} \in \mathcal{C}$, $\boldsymbol{m}(w) = m$ for all $w \in \mathbf{\Omega}$.
Expected value, $E(x) = \rho_{[0,1]}(x, \mathbf{0})$ when x has sort $\mathbb{S}_{[0,1]}$.
Variance, $var(x) = E(x - E(x))^2$ when x has sort $\mathbb{S}_{[0,1]}$.

$P(x(w) \in B) = r$ where x has sort $\mathbb{S}_{\mathbb{M}}$ and $B \subseteq \mathbb{M}$ is closed.

x and y are independent.

x_n converges to x almost surely, in r-th mean, in probability, in distribution.

Can express properties of an r.v. with values in a complete separable $\mathbb{M}$ using the fact that any such r.v. x is tight, so there is a sequence of compact $\mathbb{M}_n \subseteq \mathbb{M}$ and $x_n \in L^0(\mathbf{\Omega}, \mathbb{M}_n)$ such that x_n converges to x in probability.

Can also treat stochastic processes, which are indexed families of r.v.'s.

References

Ben Yaacov. On theories of random variables. Israel J. Math. (2013).

Ben Yaacov, Berenstein, Henson and Usvyatsov. Model theory for metric structures, (2008).

Berenstein and Henson. Model theory of probability spaces. To appear.

Billingsley. Convergence of Probability Measures, Second Edition. Wiley (1999).

Fajardo and Keisler. Model Theory of Stochastic Processes. Lecture Notes in Logic, (2002).

Hoover and Keisler. Adapted probability distributions. TAMS (1984).

Keisler. Using ultrapowers to compare continuous structures. APAL 175 B (2024).

Keisler and Sun. Why saturated probability spaces are necessary. Adv. in Math. (2009).

Shelah. Saturation of ultrapowers and Keisler's order. AML 4 (1972).

Shelah. Classification Theory. North-Holland, (1978) and (1990).