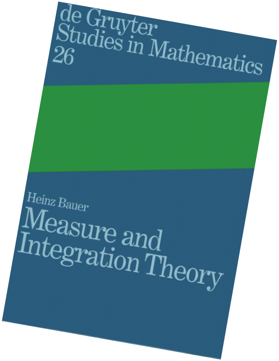
Three Chapters of Measure Theory in Isabelle/HOL

Johannes Hölzl   Armin Heller

Technische Universität München

ITP 2011
23 August 2011

Three Chapters of Measure Theory



Heinz Bauer: Measure and Integration Theory

(Most of the first three chapters)

Motivation: Program Analysis

- Randomness sources: input, scheduling, random number generator, ...
- Measure sets of sequences of random numbers, scheduling decisions, ...
  ($\mathbb{N} \Rightarrow \alpha$, non-countable spaces)

Concepts

Existence	Show existence of the measure spaces of traces
Lebesgue integral	Expected output size, ...
Density	Information flowing from input to output (Information theory)
Product	Combine different sources in one measure space

Concepts

Existence	Show existence of the measure spaces of traces
Lebesgue integral	Expected output size, ...
Density	Information flowing from input to output (Information theory)
Product	Combine different sources in one measure space

σ-Algebra

$$
\begin{aligned}
&\texttt{record}\ \alpha\ \mathit{set\text{-}system} = \\
&\qquad \mathit{space} \quad :: \quad \alpha\ \mathit{set} \qquad \text{“}\Omega\text{”} \\
&\qquad \mathit{sets} \quad :: \quad \alpha\ \mathit{set}\ \mathit{set} \qquad \text{“}\mathscr{A}\text{”}
\end{aligned}
$$

σ-Algebra

```
record α set-system =
    space :: α set       "Ω"
    sets  :: α set set   "𝒜"
```

- Set systems are introduced as locales

locale *algebra*

$\mathscr{A} \subseteq \mathscr{P}(\Omega)$,
$\emptyset, \Omega / A$,
$A \cup B \in \mathscr{A}$

↑

locale *σ-algebra*

$\bigcup_{\mathbb{N}} F \in \mathscr{A}$

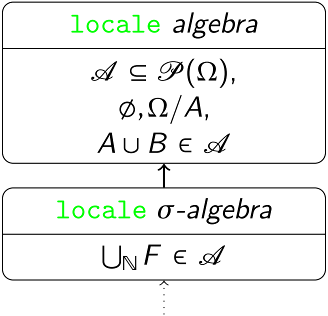
⋮

σ-Algebra

record α *set-system* =

space	::	α *set*	"Ω"
sets	::	α *set set*	"$\mathscr{A}$"

- Set systems are introduced as locales
- $\sigma(\mathscr{G})$:= smallest σ-algebra containing $\mathscr{G}$

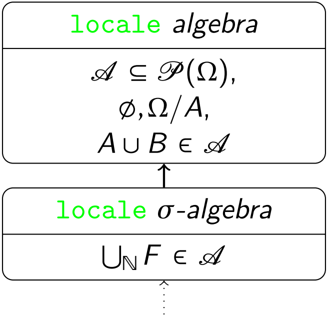


σ-Algebra

record *α set-system* =

space	::	*α set*	"Ω"
sets	::	*α set set*	"$\mathscr{A}$"

- Set systems are introduced as locales
- $\sigma(\mathscr{G})$:= smallest σ-algebra containing $\mathscr{G}$
- Example: Traces $\sigma(\{\{t | t \textit{ starts with } ts\} | ts\})$



Measure Spaces

datatype $\overline{\mathbb{R}} = -\infty \mid$ **ereal** $\mathbb{R} \mid \infty$

record α *measure-space* =
  α *set-system* +
  measure :: α *set* $\Rightarrow \overline{\mathbb{R}}$ "μ"

locale σ-*algebra*

↑

locale *measure-space*

$\mu(\emptyset) = 0 \leq \mu(A)$,
disj $F \rightarrow \mu(\bigcup F) = \sum_i \mu(F_i)$

↑

locale σ-*finite-measure-space*

$\exists F.\ \bigcup_{\mathbb{N}} F = \Omega \ \wedge\ \mu(F_i) < \infty$

↑

locale *finite-measure-space*

$\mu(\Omega) < \infty$

Measure Spaces

`datatype` $\overline{\mathbb{R}} = -\infty \mid$ **ereal** $\mathbb{R} \mid \infty$

`record` α *measure-space* =
  α *set-system* +
  measure :: α *set* $\Rightarrow \overline{\mathbb{R}}$ "μ"

Caratheodory: Extends a measure on an algebra $\mathscr{G}$ to its σ-algebra $\sigma(\mathscr{G})$

`locale` σ-*algebra*

`locale` *measure-space*

$\mu(\emptyset) = 0 \leq \mu(A)$,
disj $F \rightarrow \mu(\bigcup F) = \sum_i \mu(F_i)$

`locale` σ-*finite-measure-space*

$\exists F.\ \bigcup_{\mathbb{N}} F = \Omega \ \wedge\ \mu(F_i) < \infty$

`locale` *finite-measure-space*

$\mu(\Omega) < \infty$

Measure Spaces

datatype $\overline{\mathbb{R}} = -\infty \mid \textbf{ereal}\ \mathbb{R} \mid \infty$

record α *measure-space* =
  α *set-system* +
  measure :: α *set* $\Rightarrow$ $\overline{\mathbb{R}}$ "μ"

Caratheodory Extends a measure on an algebra $\mathcal{G}$ to its σ-algebra $\sigma(\mathcal{G})$

Uniqueness σ-finite measures equal on an $\cap$-stable set $\mathcal{G}$ are equal on its σ-algebra $\sigma(\mathcal{G})$

locale σ-*algebra*

↑

locale *measure-space*

$\mu(\emptyset) = 0 \leq \mu(A)$,
$\text{disj } F \to \mu(\bigcup F) = \sum_i \mu(F_i)$

↑

locale σ-*finite-measure-space*

$\exists F.\ \bigcup_{\mathbb{N}} F = \Omega\ \wedge\ \mu(F_i) < \infty$

↑

locale *finite-measure-space*

$\mu(\Omega) < \infty$

Concepts

Existence	Show existence of the measure spaces of traces
Lebesgue integral	Expected output size, . . .
Density	Information flowing from input to output (Information theory)
Product	Combine different sources in one measure space

Lebesgue Integral

$$\begin{aligned}
\int^{+} \cdot \quad &:: \quad (\alpha \Rightarrow \overline{\mathbb{R}}) \Rightarrow \overline{\mathbb{R}} \\
\int \cdot \quad &:: \quad (\alpha \Rightarrow \mathbb{R}) \Rightarrow \mathbb{R} \\
\int_x f(x) \quad &= \quad \int_x^{+} f^{+}(x) - \int_x^{+} f^{-}(x) \\
\text{integrable } f \quad &\longleftrightarrow \quad \text{measurable } f \ \wedge \\
&\qquad \int_x^{+} f^{+}(x) < \infty \ \wedge \ \int_x^{+} f^{-}(x) < \infty
\end{aligned}$$

Indicator $\int^{+} \chi_A = \mu(A)$

Lebesgue Integral

$$\begin{aligned}
\int^{+} \cdot \quad &:: \quad (\alpha \Rightarrow \overline{\mathbb{R}}) \Rightarrow \overline{\mathbb{R}} \\
\int \cdot \quad &:: \quad (\alpha \Rightarrow \mathbb{R}) \Rightarrow \mathbb{R} \\
\int_x f(x) \quad &= \quad \int_x^{+} f^{+}(x) - \int_x^{+} f^{-}(x) \\
\text{integrable } f \quad &\longleftrightarrow \quad \text{measurable } f \;\wedge \\
& \qquad \int_x^{+} f^{+}(x) < \infty \;\wedge\; \int_x^{+} f^{-}(x) < \infty
\end{aligned}$$

Indicator $\int^{+} \chi_A = \mu(A)$

Monotonicity $f \leq g \text{ ae } \longrightarrow \int^{+} f \leq \int^{+} g$

Lebesgue Integral

$$\begin{aligned}
\int^{+} \cdot \quad &:: \quad (\alpha \Rightarrow \overline{\mathbb{R}}) \Rightarrow \overline{\mathbb{R}} \\
\int \cdot \quad &:: \quad (\alpha \Rightarrow \mathbb{R}) \Rightarrow \mathbb{R} \\
\int_x f(x) \quad &= \quad \int_x^{+} f^{+}(x) - \int_x^{+} f^{-}(x) \\
\text{integrable } f \quad &\longleftrightarrow \quad \text{measurable } f \wedge \\
&\qquad \int_x^{+} f^{+}(x) < \infty \ \wedge \ \int_x^{+} f^{-}(x) < \infty
\end{aligned}$$

Indicator $\int^{+} \chi_A = \mu(A)$

Monotonicity $f \leq g \text{ ae} \longrightarrow \int^{+} f \leq \int^{+} g$

Convergence (f_i measurable, f mono): $\sup_i(\int^{+} f_i) = \int^{+}(\sup_i f_i)$

Lebesgue Integral

$$
\begin{aligned}
\int^{+} \cdot \quad &:: \quad (\alpha \Rightarrow \overline{\mathbb{R}}) \Rightarrow \overline{\mathbb{R}} \\
\int \cdot \quad &:: \quad (\alpha \Rightarrow \mathbb{R}) \Rightarrow \mathbb{R} \\
\int_x f(x) \quad &= \quad \int_x^{+} f^{+}(x) - \int_x^{+} f^{-}(x) \\
\text{integrable } f \quad &\longleftrightarrow \quad \text{measurable } f \ \wedge \\
&\qquad \int_x^{+} f^{+}(x) < \infty \ \wedge \ \int_x^{+} f^{-}(x) < \infty
\end{aligned}
$$

Indicator $\int^{+} \chi_A = \mu(A)$

Monotonicity $f \le g \text{ ae } \longrightarrow \int^{+} f \le \int^{+} g$

Convergence (f_i measurable, f mono): $\sup_i(\int^{+} f_i) = \int^{+}(\sup_i f_i)$

Linearity (f, g measurable): $\int^{+}(c \cdot f + d \cdot g) = c \cdot \int^{+} f + d \cdot \int^{+} g$

Concepts

Existence	Show existence of the measure spaces of traces
Lebesgue integral	Expected output size, ...
Density	Information flowing from input to output (Information theory)
Product	Combine different sources in one measure space

Density (Radon-Nikodým Derivative)

Density := The measurable function f to represent ν as the integral over μ:

$$\forall A. \qquad \nu(A) = \int_X^+ f(x) \cdot \chi_A(x) d\mu$$

Density (Radon-Nikodým Derivative)

Density := The measurable function f to represent ν as the integral over μ:

$$\forall A. \qquad \nu(A) = \int_X^+ f(x) \cdot \chi_A(x) d\mu$$

Existence If $\mu(A) = 0 \longrightarrow \nu(A) = 0$ then a density f exists. (*Radon-Nikodým* derivative)

Density (Radon-Nikodým Derivative)

Density := The measurable function f to represent ν as the integral over μ:

$$\forall A. \qquad \nu(A) = \int_{X}^{+} f(x) \cdot \chi_A(x) d\mu$$

Existence If $\mu(A) = 0 \longrightarrow \nu(A) = 0$ then a density f exists. (*Radon-Nikodým* derivative)

Uniqueness Two densities from μ to ν are almost everywhere equal (σ-finite).

Concepts

Existence	Show existence of the measure spaces of traces
Lebesgue integral	Expected output size, ...
Density	Information flowing from input to output (Information theory)
Product	Combine different sources in one measure space

Product Measure

Sets := $\mathscr{A} \otimes \mathscr{B}$ is the σ-closure of all $A \times B$

Measure := $\mu \otimes \nu$ (σ-finite μ and ν) on $\mathscr{A} \otimes \mathscr{B}$ with:

$$\mu \otimes \nu(A \times B) = \mu(A) \cdot \nu(B)$$

$$\mu \otimes \nu(C) = \int_x^+ \left(\int_y^+ \chi_C(x,y) d\nu\right) d\mu$$

Product Measure

Sets := $\mathscr{A} \otimes \mathscr{B}$ is the σ-closure of all $A \times B$

Measure := $\mu \otimes \nu$ (σ-finite μ and ν) on $\mathscr{A} \otimes \mathscr{B}$ with:
$\mu \otimes \nu(A \times B) = \mu(A) \cdot \nu(B)$
$\mu \otimes \nu(C) = \int_x^+ (\int_y^+ \chi_C(x,y) d\nu) d\mu$

Fubini
$$\int_{(x,y)} f(x,y) d(\mu \otimes \nu) = \int_x \left(\int_y f(x,y) d\nu \right) d\mu = \int_y (\int_x f(x,y) d\mu) d\nu$$

Product Measure

Sets := $\mathscr{A} \otimes \mathscr{B}$ is the σ-closure of all $A \times B$

Measure := $\mu \otimes \nu$ (σ-finite μ and ν) on $\mathscr{A} \otimes \mathscr{B}$ with:
$\mu \otimes \nu(A \times B) = \mu(A) \cdot \nu(B)$
$\mu \otimes \nu(C) = \int_x^+ (\int_y^+ \chi_C(x,y) d\nu) d\mu$

Fubini
$$\begin{aligned} \int_{(x,y)} f(x,y) d(\mu \otimes \nu) &= \int_x \left(\int_y f(x,y) d\nu\right) d\mu \\ &= \int_y \left(\int_x f(x,y) d\mu\right) d\nu \end{aligned}$$

Symmetry Symmetric variants are proven by the isomorphism swap $(\lambda(x,y).(y,x))$

Product Measure

Sets := $\mathscr{A} \otimes \mathscr{B}$ is the σ-closure of all $A \times B$

Measure := $\mu \otimes \nu$ (σ-finite μ and ν) on $\mathscr{A} \otimes \mathscr{B}$ with:
$\mu \otimes \nu(A \times B) = \mu(A) \cdot \nu(B)$
$\mu \otimes \nu(C) = \int_x^+ (\int_y^+ \chi_C(x,y) d\nu) d\mu$

Fubini
$$\int_{(x,y)} f(x,y) d(\mu \otimes \nu) = \int_x \left(\int_y f(x,y) d\nu\right) d\mu = \int_y (\int_x f(x,y) d\mu) d\nu$$

Symmetry: Symmetric variants are proven by the isomorphism swap $(\lambda(x,y).(y,x))$

Iterated Product: Over a finite I: $\bigotimes_{i \in I} \mathscr{A}_i :: (\iota \Rightarrow \alpha)$ *measure-space*

Concepts

Existence	Show existence of the measure spaces of traces
Lebesgue integral	Expected output size, ...
Density	Information flowing from input to output (Information theory)
Product	Combine different sources in one measure space

Concepts

Existence	Show existence of the measure spaces of traces
Lebesgue integral	Expected output size, ...
Density	Information flowing from input to output (Information theory)
Product	Combine different sources in one measure space

But measure theory is even more ...

Multivariate Analysis

- $\sigma\left(\left\{\,[a,b[\;\middle|\,a,b\in\mathbb{R}\right\}\right)$ – called Borel sets
- On euclidean spaces equivalent to $\sigma(\{X|\text{ open }X\})$
  $\implies$ define Borel sets on *topological-space*s ($\mathbb{R}$, $\mathbb{R}^n$, $\overline{\mathbb{R}}$, ...)

Multivariate Analysis

- $\sigma\left(\left\{\, [a,b[\,\middle|\, a,b \in \mathbb{R} \right\}\right)$ – called Borel sets
- On euclidean spaces equivalent to $\sigma(\{X|\text{ open } X\})$
  $\Longrightarrow$ define Borel sets on *topological-spaces* ($\mathbb{R}$, $\mathbb{R}^n$, $\overline{\mathbb{R}}$, ...)
- Lebesgue measure λ: Extension of the function which assigns each interval its length: $\lambda([a,b[) = b - a$

Multivariate Analysis

- $\sigma\left(\left\{ [a,b[\,\middle|\, a,b \in \mathbb{R}\right\}\right)$ – called Borel sets
- On euclidean spaces equivalent to $\sigma(\{X|\text{ open } X\})$
  $\Longrightarrow$ define Borel sets on *topological-spaces* ($\mathbb{R}$, $\mathbb{R}^n$, $\overline{\mathbb{R}}$, . . .)
- Lebesgue measure λ: Extension of the function which assigns each interval its length: $\lambda([a,b[) = b - a$
- Existence by Gauge integral over χ_A, *not Caratheodory*

Multivariate Analysis

- $\sigma\left(\left\{ [a,b[\,\middle|\, a,b \in \mathbb{R}\right\}\right)$ – called Borel sets
- On euclidean spaces equivalent to $\sigma(\{X|\text{ open }X\})$
  $\Longrightarrow$ define Borel sets on *topological-space*s ($\mathbb{R}$, $\mathbb{R}^n$, $\overline{\mathbb{R}}$, ...)
- Lebesgue measure λ: Extension of the function which assigns each interval its length: $\lambda([a,b[) = b - a$
- Existence by Gauge integral over χ_A, *not Caratheodory*
- Euclidean space equal to products:
  $\bigotimes_{i<n} \lambda \approx \lambda^n$
- Lebesgue integral subset of Gauge integral:

$$\begin{aligned} \int_X^G f(x) &= \int_X f(x) d\lambda^n \\ &= \int_X f(x) d\left(\bigotimes_{i<n} \lambda\right) \end{aligned}$$

Related Work

	HOL							
	Hurd	Richter	Coble	Mhamdi	Lester (PVS)	Mizar	HOL-Light	Isabelle
$\overline{\mathbb{R}}$				ITP'11	✓	✓		✓
Borel Sets (Open)				✓	✓			✓
Leb. Integral		✓	✓	✓	✓	✓	✓	✓
Leb. Measure	$\mathbb{N} \Rightarrow \mathbb{B}$				✓	✓	$\lambda \cdot \in \mathbb{R}$	✓
Products			finite		✓		$\mathbb{R}^{\alpha+\beta}$	✓
Dynkin						✓		✓
Radon-Nikodým			finite	ITP'11				✓

Applications (Not in Paper)

Information Theory Use Radon-Nikodým derivative to formalize relative entropy (Kullback-Leibler divergence), Mutual Information and Entropy

Applications (Not in Paper)

Information Theory	Use Radon-Nikodým derivative to formalize relative entropy (Kullback-Leibler divergence), Mutual Information and Entropy
Probability Theory	Existence of infinite products ($\mathbb{N} \Rightarrow [0,1[$), independence of random variables

Applications (Not in Paper)

Information Theory Use Radon-Nikodým derivative to formalize relative entropy (Kullback-Leibler divergence), Mutual Information and Entropy

Probability Theory Existence of infinite products ($\mathbb{N} \Rightarrow [0,1[$), independence of random variables

Markov Chains Formalize measure of Markov chains traces