18.175: Lecture 1

Probability spaces, distributions, random variables, measure theory

Scott Sheffield

MIT

Outline

Probability spaces and σ-algebras

Distributions on $\mathbb{R}$

Extension theorems

Outline

Probability spaces and σ-algebras

Distributions on $\mathbb{R}$

Extension theorems

Probability space notation

- **Probability space** is triple $(\Omega, \mathcal{F}, P)$ where Ω is sample space, $\mathcal{F}$ is set of events (the σ-algebra) and $P : \mathcal{F} \to [0, 1]$ is the probability function.

Probability space notation

- **Probability space** is triple $(\Omega, \mathcal{F}, P)$ where Ω is sample space, $\mathcal{F}$ is set of events (the σ-algebra) and $P : \mathcal{F} \to [0, 1]$ is the probability function.
- σ**-algebra** is collection of subsets closed under complementation and countable unions. Call $(\Omega, \mathcal{F})$ a measure space.

Probability space notation

- **Probability space** is triple $(\Omega, \mathcal{F}, P)$ where Ω is sample space, $\mathcal{F}$ is set of events (the σ-algebra) and $P : \mathcal{F} \to [0,1]$ is the probability function.
- σ-**algebra** is collection of subsets closed under complementation and countable unions. Call $(\Omega, \mathcal{F})$ a measure space.
- **Measure** is function $\mu : \mathcal{F} \to \mathbb{R}$ satisfying $\mu(A) \geq \mu(\emptyset) = 0$ for all $A \in \mathcal{F}$ and countable additivity: $\mu(\cup_i A_i) = \sum_i \mu(A_i)$ for disjoint A_i.

Probability space notation

- **Probability space** is triple $(\Omega, \mathcal{F}, P)$ where Ω is sample space, $\mathcal{F}$ is set of events (the σ-algebra) and $P : \mathcal{F} \to [0, 1]$ is the probability function.
- σ**-algebra** is collection of subsets closed under complementation and countable unions. Call $(\Omega, \mathcal{F})$ a measure space.
- **Measure** is function $\mu : \mathcal{F} \to \mathbb{R}$ satisfying $\mu(A) \geq \mu(\emptyset) = 0$ for all $A \in \mathcal{F}$ and countable additivity: $\mu(\cup_i A_i) = \sum_i \mu(A_i)$ for disjoint A_i.
- Measure μ is **probability measure** if $\mu(\Omega) = 1$.

Basic consequences of definitions

- **monotonicity**: $A \subset B$ implies $\mu(A) \subset \mu(B)$

Basic consequences of definitions

- **monotonicity**: $A \subset B$ implies $\mu(A) \subset \mu(B)$
- **subadditivity**: $A \subset \cup_{m=1}^{\infty} A_m$ implies $\mu(A) \leq \sum_{m=1}^{\infty} \mu(A_m)$.

Basic consequences of definitions

- **monotonicity**: $A \subset B$ implies $\mu(A) \subset \mu(B)$
- **subadditivity**: $A \subset \cup_{m=1}^{\infty} A_m$ implies $\mu(A) \leq \sum_{m=1}^{\infty} \mu(A_m)$.
- **continuity from below:** measures of sets A_i in increasing sequence converge to measure of limit $\cup_i A_i$

Basic consequences of definitions

- **monotonicity**: $A \subset B$ implies $\mu(A) \subset \mu(B)$
- **subadditivity**: $A \subset \cup_{m=1}^{\infty} A_m$ implies $\mu(A) \leq \sum_{m=1}^{\infty} \mu(A_m)$.
- **continuity from below:** measures of sets A_i in increasing sequence converge to measure of limit $\cup_i A_i$
- **continuity from above:** measures of sets A_i in decreasing sequence converge to measure of intersection $\cap_i A_i$

Why can't σ-algebra be all subsets of Ω?

- Uniform probability measure on $[0, 1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0, 1)$, their probabilities should be equal.

Why can't σ-algebra be all subsets of Ω?

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x + r) \mod 1$.

Why can't σ-algebra be all subsets of Ω?

- Uniform probability measure on $[0, 1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0, 1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x + r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.

Why can't σ-algebra be all subsets of Ω?

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x+r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.

Why can't σ-algebra be all subsets of Ω?

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x + r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.
- There are uncountably many of these classes.

Why can't σ-algebra be all subsets of Ω?

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x+r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.
- There are uncountably many of these classes.
- Let $A \subset [0,1)$ contain **one** point from each class. For each $x \in [0,1)$, there is **one** $a \in A$ such that $r = x - a$ is rational.

Why can't σ-algebra be all subsets of Ω?

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x+r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.
- There are uncountably many of these classes.
- Let $A \subset [0,1)$ contain **one** point from each class. For each $x \in [0,1)$, there is **one** $a \in A$ such that $r = x - a$ is rational.
- Then each x in $[0,1)$ lies in $\tau_r(A)$ for **one** rational $r \in [0,1)$.

Why can't σ-algebra be all subsets of Ω?

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x+r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.
- There are uncountably many of these classes.
- Let $A \subset [0,1)$ contain **one** point from each class. For each $x \in [0,1)$, there is **one** $a \in A$ such that $r = x - a$ is rational.
- Then each x in $[0,1)$ lies in $\tau_r(A)$ for **one** rational $r \in [0,1)$.
- Thus $[0,1) = \cup \tau_r(A)$ as r ranges over rationals in $[0,1)$.

Why can't σ-algebra be all subsets of Ω?

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x+r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.
- There are uncountably many of these classes.
- Let $A \subset [0,1)$ contain **one** point from each class. For each $x \in [0,1)$, there is **one** $a \in A$ such that $r = x - a$ is rational.
- Then each x in $[0,1)$ lies in $\tau_r(A)$ for **one** rational $r \in [0,1)$.
- Thus $[0,1) = \cup \tau_r(A)$ as r ranges over rationals in $[0,1)$.
- If $P(A) = 0$, then $P(S) = \sum_r P(\tau_r(A)) = 0$. If $P(A) > 0$ then $P(S) = \sum_r P(\tau_r(A)) = \infty$. Contradicts $P(S) = 1$ axiom.

Three ways to get around this

- 1. **Re-examine axioms of mathematics:** the very *existence* of a set A with one element from each equivalence class is consequence of so-called **axiom of choice**. Removing that axiom makes paradox goes away, since one can just suppose (pretend?) these kinds of sets don't exist.

Three ways to get around this

- 1. **Re-examine axioms of mathematics:** the very *existence* of a set A with one element from each equivalence class is consequence of so-called **axiom of choice**. Removing that axiom makes paradox goes away, since one can just suppose (pretend?) these kinds of sets don't exist.
- 2. **Re-examine axioms of probability:** Replace *countable additivity* with *finite additivity*? (Look up Banach-Tarski.)

Three ways to get around this

- 1. **Re-examine axioms of mathematics:** the very *existence* of a set A with one element from each equivalence class is consequence of so-called **axiom of choice**. Removing that axiom makes paradox goes away, since one can just suppose (pretend?) these kinds of sets don't exist.
- 2. **Re-examine axioms of probability:** Replace *countable additivity* with *finite additivity*? (Look up Banach-Tarski.)
- 3. **Keep the axiom of choice and countable additivity but don't define probabilities of all sets:** Restrict attention to some σ-algebra of measurable sets.

Three ways to get around this

- 1. **Re-examine axioms of mathematics:** the very *existence* of a set A with one element from each equivalence class is consequence of so-called **axiom of choice**. Removing that axiom makes paradox goes away, since one can just suppose (pretend?) these kinds of sets don't exist.
- 2. **Re-examine axioms of probability:** Replace *countable additivity* with *finite additivity*? (Look up Banach-Tarski.)
- 3. **Keep the axiom of choice and countable additivity but don't define probabilities of all sets:** Restrict attention to some σ-algebra of measurable sets.
- Most mainstream probability and analysis takes the third approach. But good to be aware of alternatives (e.g., **axiom of determinacy** which implies that all sets are Lebesgue measurable).

Borel σ-algebra

- The **Borel σ-algebra** $\mathcal{B}$ on a topological space is the smallest σ-algebra containing all open sets. In the case of $\mathbb{R}$, it is the smallest σ-algebra containing all open intervals.

Borel σ-algebra

- The **Borel σ-algebra** $\mathcal{B}$ on a topological space is the smallest σ-algebra containing all open sets. In the case of $\mathbb{R}$, it is the smallest σ-algebra containing all open intervals.
- Say that $\mathcal{B}$ is "generated" by the collection of open intervals.

Outline

Probability spaces and σ-algebras

Distributions on $\mathbb{R}$

Extension theorems

Outline

Probability spaces and σ-algebras

Distributions on $\mathbb{R}$

Extension theorems

How do we produce measures on $\mathbb{R}$?

- Write $F(a) = P\big((-\infty, a]\big)$.

How do we produce measures on $\mathbb{R}$?

- Write $F(a) = P\big((-\infty, a]\big)$.
- **Theorem:** for each right continuous, non-decreasing function F, tending to 0 at $-\infty$ and to 1 at ∞, there is a unique measure defined on the Borel sets of $\mathbb{R}$ with $P((a, b]) = F(b) - F(a)$.

How do we produce measures on $\mathbb{R}$?

- Write $F(a) = P\big((-\infty, a]\big)$.
- **Theorem:** for each right continuous, non-decreasing function F, tending to 0 at $-\infty$ and to 1 at ∞, there is a unique measure defined on the Borel sets of $\mathbb{R}$ with $P((a,b]) = F(b) - F(a)$.
- If we're given such a function F, then we know how to compute the measure of any set of the form $(a, b]$.

How do we produce measures on $\mathbb{R}$?

- Write $F(a) = P\big((-\infty, a]\big)$.
- **Theorem:** for each right continuous, non-decreasing function F, tending to 0 at $-\infty$ and to 1 at ∞, there is a unique measure defined on the Borel sets of $\mathbb{R}$ with $P((a, b]) = F(b) - F(a)$.
- If we're given such a function F, then we know how to compute the measure of any set of the form $(a, b]$.
- We would like to *extend* the measure defined for these subsets to a measure defined for the whole σ algebra generated by these subsets.

How do we produce measures on $\mathbb{R}$?

- Write $F(a) = P\big((-\infty, a]\big)$.
- **Theorem:** for each right continuous, non-decreasing function F, tending to 0 at $-\infty$ and to 1 at ∞, there is a unique measure defined on the Borel sets of $\mathbb{R}$ with $P((a, b]) = F(b) - F(a)$.
- If we're given such a function F, then we know how to compute the measure of any set of the form $(a, b]$.
- We would like to *extend* the measure defined for these subsets to a measure defined for the whole σ algebra generated by these subsets.
- Seems clear how to define measure of countable union of disjoint intervals of the form $(a, b]$ (just using countable additivity). But are we confident we can extend the definition to *all* Borel measurable sets in a consistent way?

Outline

Probability spaces and σ-algebras

Distributions on $\mathbb{R}$

Extension theorems

Outline

Probability spaces and σ-algebras

Distributions on $\mathbb{R}$

Extension theorems

Algebras and semi-algebras

- **algebra**: collection $\mathcal{A}$ of sets closed under finite unions and complementation.

Algebras and semi-algebras

- **algebra**: collection $\mathcal{A}$ of sets closed under finite unions and complementation.
- **measure on algebra**: Have $\mu(A) \geq \mu(\emptyset) = 0$ for all A in $\mathcal{A}$, and for disjoint A_i with union in $\mathcal{A}$ we have $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ (countable additivity).

Algebras and semi-algebras

- **algebra**: collection $\mathcal{A}$ of sets closed under finite unions and complementation.
- **measure on algebra**: Have $\mu(A) \geq \mu(\emptyset) = 0$ for all A in $\mathcal{A}$, and for disjoint A_i with union in $\mathcal{A}$ we have $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ (countable additivity).
- Measure μ on $\mathcal{A}$ is σ-**finite** if exists countable collection $A_n \in \mathcal{A}$ with $\mu(A_n) < \infty$ and $\cup A_n = \Omega$.

Algebras and semi-algebras

- **algebra**: collection $\mathcal{A}$ of sets closed under finite unions and complementation.
- **measure on algebra**: Have $\mu(A) \geq \mu(\emptyset) = 0$ for all A in $\mathcal{A}$, and for disjoint A_i with union in $\mathcal{A}$ we have $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ (countable additivity).
- Measure μ on $\mathcal{A}$ is σ-**finite** if exists countable collection $A_n \in \mathcal{A}$ with $\mu(A_n) < \infty$ and $\cup A_n = \Omega$.
- **semi-algebra**: collection $\mathcal{S}$ of sets closed under intersection and such that $S \in \mathcal{S}$ implies that S^c is a finite disjoint union of sets in $\mathcal{S}$. (Example: empty set plus sets of form $(a_1, b_1] \times \ldots \times (a_d, b_d] \in \mathbb{R}^d$.)

Algebras and semi-algebras

- **algebra**: collection $\mathcal{A}$ of sets closed under finite unions and complementation.
- **measure on algebra**: Have $\mu(A) \geq \mu(\emptyset) = 0$ for all A in $\mathcal{A}$, and for disjoint A_i with union in $\mathcal{A}$ we have $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ (countable additivity).
- Measure μ on $\mathcal{A}$ is σ-**finite** if exists countable collection $A_n \in \mathcal{A}$ with $\mu(A_n) < \infty$ and $\cup A_n = \Omega$.
- **semi-algebra**: collection $\mathcal{S}$ of sets closed under intersection and such that $S \in \mathcal{S}$ implies that S^c is a finite disjoint union of sets in $\mathcal{S}$. (Example: empty set plus sets of form $(a_1, b_1] \times \ldots \times (a_d, b_d] \in \mathbb{R}^d$.)
- One lemma: If $\mathcal{S}$ is a semialgebra, then the set $\overline{\mathcal{S}}$ of finite disjoint unions of sets in $\mathcal{S}$ is an algebra, called the **algebra generated by** $\mathcal{S}$.

π-systems and λ-systems

- Say collection of sets $\mathcal{P}$ is a π-system if closed under intersection.

π-systems and λ-systems

- Say collection of sets $\mathcal{P}$ is a π-system if closed under intersection.
- Say collection of sets $\mathcal{L}$ is a λ-system if

π-systems and λ-systems

- Say collection of sets $\mathcal{P}$ is a π-system if closed under intersection.
- Say collection of sets $\mathcal{L}$ is a λ-system if
  - $\Omega \in \mathcal{L}$

π-systems and λ-systems

- Say collection of sets $\mathcal{P}$ is a π-system if closed under intersection.
- Say collection of sets $\mathcal{L}$ is a λ-system if
  - $\Omega \in \mathcal{L}$
  - If $A, B \in \mathcal{L}$ and $A \subset B$, then $B - A \in \mathcal{L}$.

π-systems and λ-systems

- Say collection of sets $\mathcal{P}$ is a π-system if closed under intersection.
- Say collection of sets $\mathcal{L}$ is a λ-system if
  - $\Omega \in \mathcal{L}$
  - If $A, B \in \mathcal{L}$ and $A \subset B$, then $B - A \in \mathcal{L}$.
  - If $A_n \in \mathcal{L}$ and $A_n \uparrow A$ then $A \in \mathcal{L}$.

π-systems and λ-systems

- Say collection of sets $\mathcal{P}$ is a π-system if closed under intersection.
- Say collection of sets $\mathcal{L}$ is a λ-system if
  - $\Omega \in \mathcal{L}$
  - If $A, B \in \mathcal{L}$ and $A \subset B$, then $B - A \in \mathcal{L}$.
  - If $A_n \in \mathcal{L}$ and $A_n \uparrow A$ then $A \in \mathcal{L}$.
- THEOREM: If $\mathcal{P}$ is a π-system and $\mathcal{L}$ is a λ-system that contains $\mathcal{P}$, then $\sigma(\mathcal{P}) \subset \mathcal{L}$, where $\sigma(\mathcal{A})$ denotes smallest σ-algebra containing $\mathcal{A}$.

Caratheéodory Extension Theorem

- **Theorem:** If μ is a σ-finite measure on an algebra $\mathcal{A}$ then μ has a unique extension to the σ algebra generated by $\mathcal{A}$.

Carathéodory Extension Theorem

- **Theorem:** If μ is a σ-finite measure on an algebra $\mathcal{A}$ then μ has a unique extension to the σ algebra generated by $\mathcal{A}$.
- Detailed proof is somewhat involved, but let's take a look at it.

Carattheéodory Extension Theorem

- **Theorem:** If μ is a σ-finite measure on an algebra $\mathcal{A}$ then μ has a unique extension to the σ algebra generated by $\mathcal{A}$.
- Detailed proof is somewhat involved, but let's take a look at it.
- We can use this extension theorem prove existence of a unique translation invariant measure (Lebesgue measure) on the Borel sets of $\mathbb{R}^d$ that assigns unit mass to a unit cube. (Borel σ-algebra $\mathcal{R}^d$ is the smallest one containing all open sets of $\mathbb{R}^d$. Given any space with a topology, we can define a σ-algebra this way.)

Extension theorem for semialgebras

- Say $\mathcal{S}$ is semialgebra and μ is defined on $\mathcal{S}$ with $\mu(\emptyset = 0)$, such that μ is finitely additive and countably subadditive. [This means that if $S \in \mathcal{S}$ is a finite disjoint union of sets $S_i \in \mathcal{S}$ then $\mu(S) = \sum_i \mu(S_i)$. If it is a countable disjoint union of $S_i \in \mathcal{S}$ then $\mu(S) \leq \sum_i \mu(S_i)$.] Then μ has a unique extension $\bar{\mu}$ that is a measure on the algebra $\overline{\mathcal{S}}$ generated by $\mathcal{S}$. If $\bar{\mu}$ is sigma-finite, then there is an extension that is a measure on $\sigma(S)$.