Outline

Probability spaces and $\sigma$-algebras

Distributions on $\mathbb{R}$
Probability spaces and $\sigma$-algebras

Distributions on $\mathbb{R}$
Probability space notation

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

- $\sigma$-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: $\mu(\bigcup_i A_i) = \sum_i \mu(A_i)$ for disjoint $A_i$.

- Measure $\mu$ is probability measure if $\mu(\Omega) = 1$. 

18.175 Lecture 1
Probability space notation

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

- $\sigma$-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 $\Omega$ is sample space, $\mathcal{F}$ is set of events (the $\sigma$-algebra) and $P : \mathcal{F} \rightarrow [0, 1]$ is the probability function.

- **$\sigma$-algebra** is collection of subsets closed under complementation and countable unions. Call $(\Omega, \mathcal{F})$ a measure space.

- **Measure** is function $\mu : \mathcal{F} \rightarrow \mathbb{R}$ satisfying $\mu(A) \geq \mu(\emptyset) = 0$ for all $A \in \mathcal{F}$ and countable additivity: $\mu(\bigcup_i A_i) = \sum_i \mu(A_i)$ for disjoint $A_i$. 
Probability space notation

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

- **\(\sigma\)-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: \(\mu(\bigcup_i A_i) = \sum_i \mu(A_i)\) for disjoint \(A_i\).

- Measure \(\mu\) is **probability measure** if \(\mu(\Omega) = 1\).
Basic consequences of definitions

- **monotonicity**: \( A \subset B \) implies \( \mu(A) \leq \mu(B) \)
Basic consequences of definitions

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

- **monotonicity**: $A \subset B$ implies $\mu(A) \leq \mu(B)$
- **subadditivity**: $A \subset \bigcup_{m=1}^{\infty} A_m$ 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 $\bigcup_i A_i$.
Basic consequences of definitions

- **monotonicity**: $A \subset B$ implies $\mu(A) \leq \mu(B)$
- **subadditivity**: $A \subset \bigcup_{m=1}^{\infty} A_m$ 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 $\bigcup_i A_i$
- **continuity from above**: measures of sets $A_i$ in decreasing sequence converge to measure of intersection $\bigcap_i A_i$
Why can’t $\sigma$-algebra be all subsets of $\Omega$?

- 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 $\sigma$-algebra be all subsets of $\Omega$?

- 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** $\tau_r(x) = (x + r) \mod 1$. 

---

18.175 Lecture 1
Why can’t $\sigma$-algebra be all subsets of $\Omega$?

- 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 $B$. 

18.175 Lecture 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.
Why can’t \( \sigma \)-algebra be all subsets of \( \Omega \)?

- 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 \(\sigma\)-algebra be all subsets of \(\Omega\)?

- 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 $\sigma$-algebra be all subsets of $\Omega$?

- 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 \( \sigma \)-algebra be all subsets of \( \Omega \)?

- 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 $\sigma$-algebra be all subsets of $\Omega$?

- 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 $[0, 1) = \bigcup \tau_r(A)$ 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 $\sigma$-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 $\sigma$-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).
The **Borel \( \sigma \)-algebra** \( \mathcal{B} \) is the smallest \( \sigma \)-algebra containing all open intervals.
The **Borel \( \sigma \)-algebra** \( \mathcal{B} \) is the smallest \( \sigma \)-algebra containing all open intervals.

Say that \( \mathcal{B} \) is “generated” by the collection of open intervals.
The **Borel σ-algebra** \( \mathcal{B} \) is the smallest σ-algebra containing all open intervals.

Say that \( \mathcal{B} \) is “generated” by the collection of open intervals.

Why does this notion make sense? If \( \mathcal{F}_i \) are σ-fields (for \( i \) in possibly uncountable index set \( I \)) does this imply that \( \bigcap_{i \in I} \mathcal{F}_i \) is a σ-field?
Probability spaces and σ-algebras

Distributions on $\mathbb{R}$
Outline

Probability spaces and $\sigma$-algebras

Distributions on $\mathbb{R}$
Can we classify set of all probability measures on $\mathbb{R}$?

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

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

18.175 Lecture 1
Can we classify set of all probability measures on $\mathbb{R}$?

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

18.175 Lecture 1