Reasoning with Probabilities

Eric Pacuit Joshua Sack

July 27, 2009

Plan for the Course

Day 1: Introduction and Background

Day 2: Probabilistic Epistemic Logics

Day 3: Dynamic Probabilistic Epistemic Logics

Day 4: Reasoning with Probabilities

Day 5: Conclusions and General Issues

Both **epsitemic logic** and **probability** have proven to be powerful tools to reason about agents beliefs in a dynamic environment.

The goal of this course is to investigate logical systems that incorporate both probabilistic and modal operators.

What we are *not* doing in this course

- Bayesian Epistemology
- Foundations of Probability
- Reasoning about Uncertainty

What we are *not* doing in this course

- Bayesian Epistemology

A. Hájek and S. Hartmann. *Bayesian Epistemology*. manuscript (2008).

L. Bovens and S. Hartmann. *Bayesian Epistemology*. Oxford University Press (2003).

- Foundations of Probability

- Reasoning about Uncertainty

What we are *not* doing in this course

- Bayesian Epistemology

A. Hájek and S. Hartmann. *Bayesian Epistemology*. manuscript (2008).

L. Bovens and S. Hartmann. *Bayesian Epistemology*. Oxford University Press (2003).

- Foundations of Probability

- Reasoning about Uncertainty

What we are *not* doing in this course

- Bayesian Epistemology

A. Hájek and S. Hartmann. *Bayesian Epistemology*. manuscript (2008).

L. Bovens and S. Hartmann. *Bayesian Epistemology*. Oxford University Press (2003).

- Foundations of Probability

- Reasoning about Uncertainty

J. Halpern. *Reasoning About Uncertainty*. MIT Press (2003).

F. Huber. *Formal Models of Beliefs*. `plato.stanford.edu/entries/formal-belief` (2009).

Plan for today

1. Motivating Examples
   - Agreeing to disagree
   - Monty Hall puzzle
2. Background

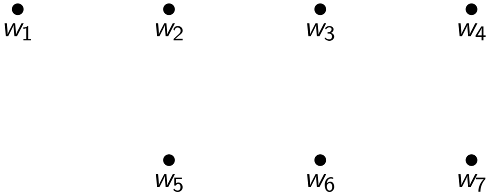
Agreeing to Disagree

Theorem: Suppose that n agents share a common prior and have different private information. If there is common knowledge in the group of the posterior probabilities, then the posteriors must be equal.

Robert Aumann. *Agreeing to Disagree.* Annals of Statistics **4** (1976).

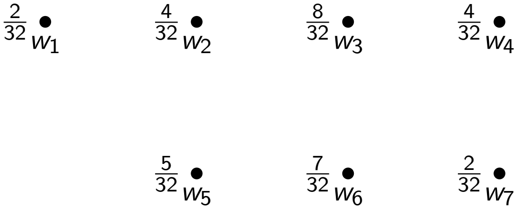
G. Bonanno and K. Nehring. *Agreeing to Disagree: A Survey*. (unpublished) 1997.

2 Scientists Perform an Experiment



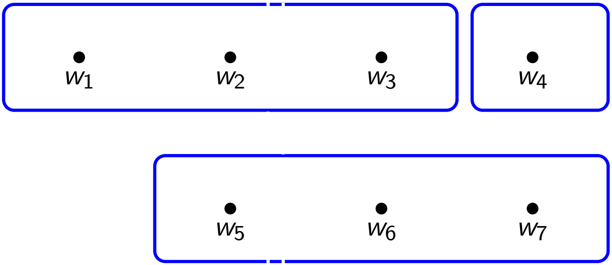
They agree the true state is one of seven different states.

2 Scientists Perform an Experiment



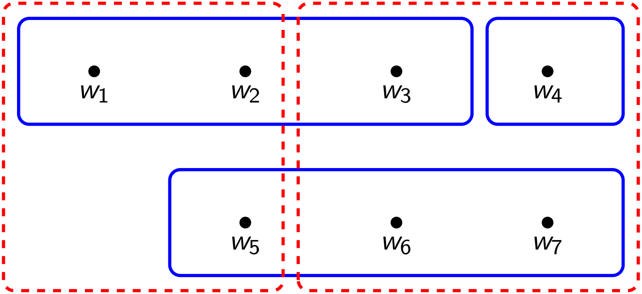
They agree on a common prior.

2 Scientists Perform an Experiment



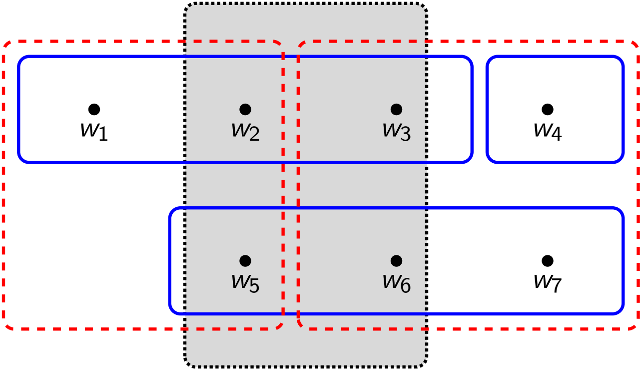
They agree that Experiment 1 would produce the blue partition.

2 Scientists Perform an Experiment



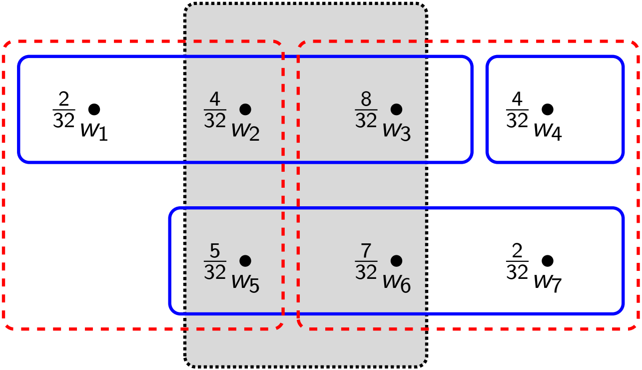
They agree that Experiment 1 would produce the blue partition and Experiment 2 the red partition.

2 Scientists Perform an Experiment



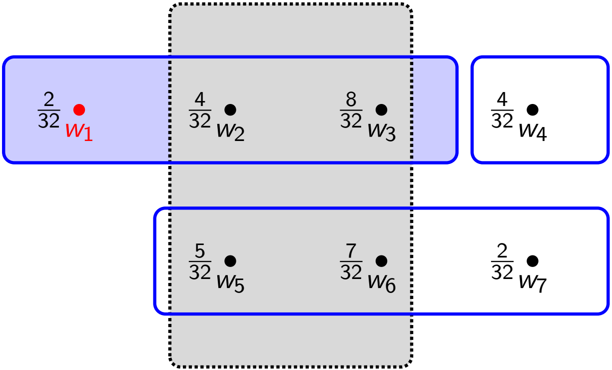
They are interested in the truth of $E = \{w_2, w_3, w_5, w_6\}$.

2 Scientists Perform an Experiment



So, they agree that $P(E) = \frac{24}{32}$.

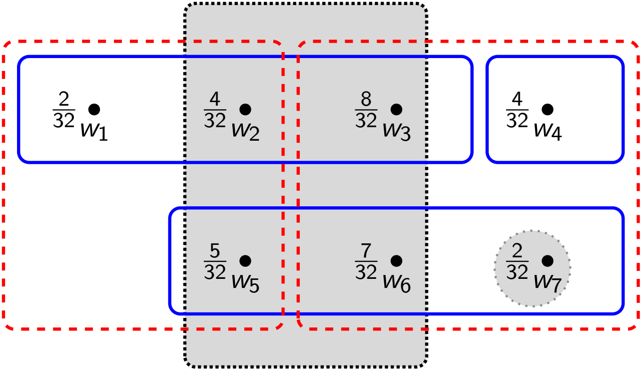
2 Scientists Perform an Experiment



Also, that *if the true state is* w_1, then Experiment 1 will yield

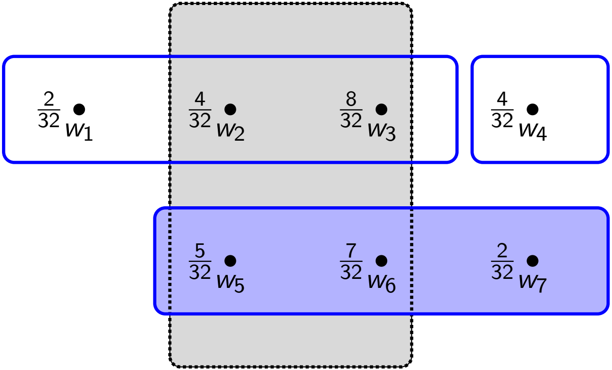
$$P(E|I) = \frac{P(E \cap I)}{P(I)} = \frac{12}{14}$$

2 Scientists Perform an Experiment



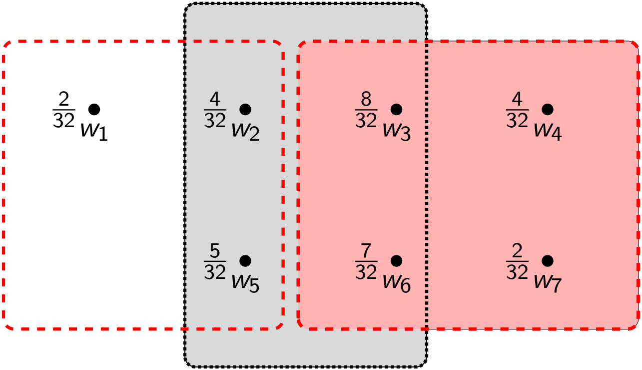
Suppose the true state is w_7 and the agents preform the experiments.

2 Scientists Perform an Experiment



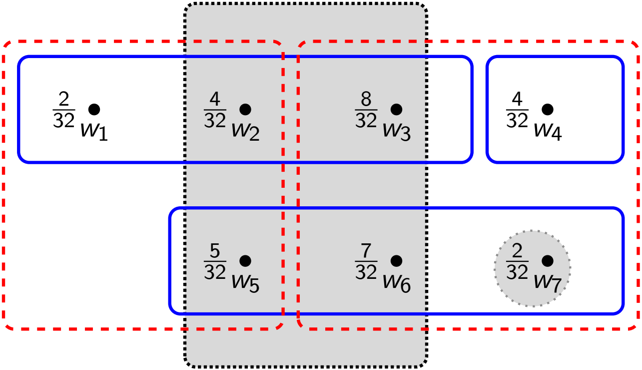
Suppose the true state is w_7, then $Pr_1(E) = \frac{12}{14}$

2 Scientists Perform an Experiment



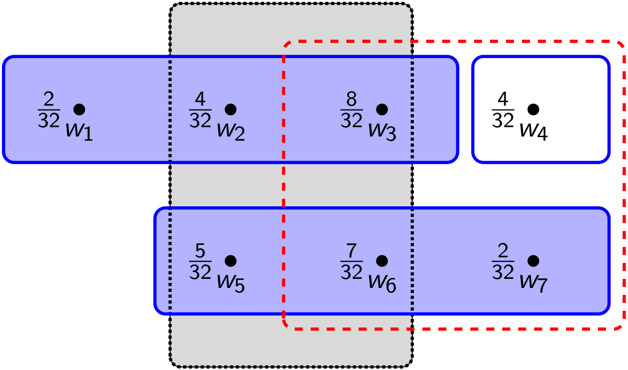
Then $Pr_1(E) = \frac{12}{14}$ and $Pr_2(E) = \frac{15}{21}$

2 Scientists Perform an Experiment



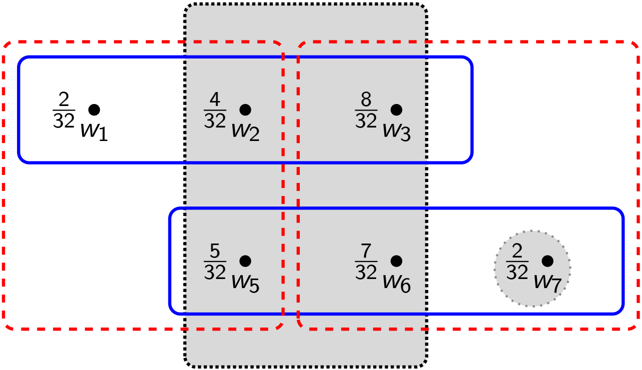
Suppose they exchange emails with the new subjective probabilities: $Pr_1(E) = \frac{12}{14}$ and $Pr_2(E) = \frac{15}{21}$

2 Scientists Perform an Experiment



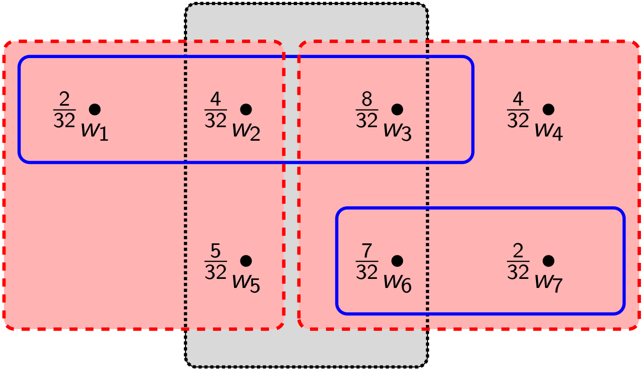
Agent 2 learns that w_4 is **NOT** the true state (same for Agent 1).

2 Scientists Perform an Experiment



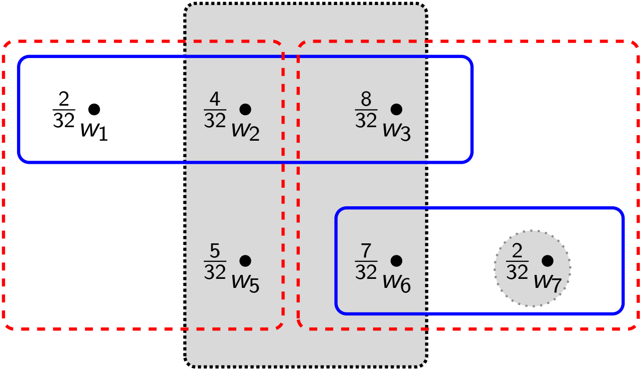
Agent 2 learns that w_4 is **NOT** the true state (same for Agent 1).

2 Scientists Perform an Experiment



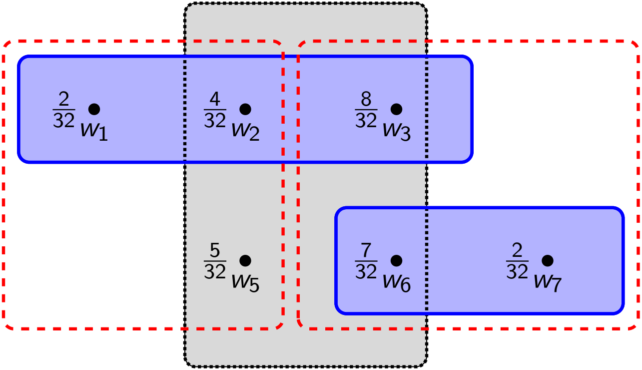
Agent 1 learns that w_5 is **NOT** the true state (same for Agent 1).

2 Scientists Perform an Experiment



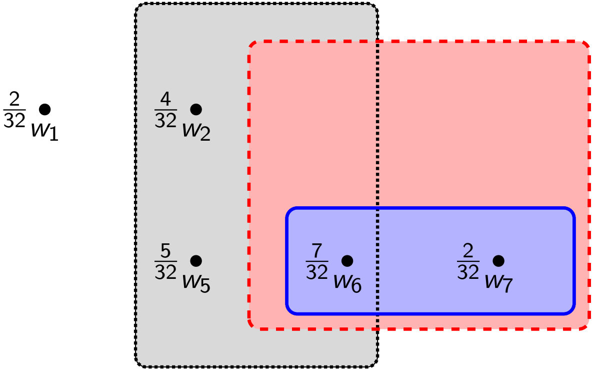
The new probabilities are $Pr_1(E|I') = \frac{7}{9}$ and $Pr_2(E|I') = \frac{15}{17}$

2 Scientists Perform an Experiment



After exchanging this information ($Pr_1(E|I') = \frac{7}{9}$ and $Pr_2(E|I') = \frac{15}{17}$), Agent 2 learns that w_3 is **NOT** the true state.

2 Scientists Perform an Experiment



No more revisions are possible and the agents agree on the posterior probabilities.

Dissecting Aumann's Theorem

- "No Trade" Theorems (Milgrom and Stokey); from probabilities of events to aggregates (McKelvey and Page); Common Prior Assumption, etc.

- How do the posteriors *become* common knowledge?

J. Geanakoplos and H. Polemarchakis. *We Can't Disagree Forever.* Journal of Economic Theory (1982).

- What are the *states of knowledge* created in a group when communication takes place? What happens when communication is not the the whole group, but pairwise?

R. Parikh and P. Krasucki. *Communication, Consensus and Knowledge.* Journal of Economic Theory (1990).

Dissecting Aumann's Theorem

- "No Trade" Theorems (Milgrom and Stokey); from probabilities of events to aggregates (McKelvey and Page); Common Prior Assumption, etc.

- How do the posteriors *become* common knowledge?

J. Geanakoplos and H. Polemarchakis. *We Can't Disagree Forever*. Journal of Economic Theory (1982).

- What are the *states of knowledge* created in a group when communication takes place? What happens when communication is not the the whole group, but pairwise?

R. Parikh and P. Krasucki. *Communication, Consensus and Knowledge*. Journal of Economic Theory (1990).

Dissecting Aumann's Theorem

- "No Trade" Theorems (Milgrom and Stokey); from probabilities of events to aggregates (McKelvey and Page); Common Prior Assumption, etc.

- How do the posteriors *become* common knowledge?

J. Geanakoplos and H. Polemarchakis. *We Can't Disagree Forever*. Journal of Economic Theory (1982).

- What are the *states of knowledge* created in a group when communication takes place? What happens when communication is not the the whole group, but pairwise?

R. Parikh and P. Krasucki. *Communication, Consensus and Knowledge*. Journal of Economic Theory (1990).

Dissecting Aumann's Theorem

- Qualitative versions: like-minded individuals cannot agree to make different decisions.

M. Bacharach. *Some Extensions of a Claim of Aumann in an Axiomatic Model of Knowledge*. Journal of Economic Theory (1985).

J.A.K. Cave. *Learning to Agree*. Economic Letters (1983).

D. Samet. *The Sure-Thing Principle and Independence of Irrelevant Knowledge*. 2008.

C. Dègrèmont and O. Roy. *Agreement Theorems in Dynamic-Epistemic Logic*. TARK 2009.

Monty Hall Puzzle

Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

Monty Hall Puzzle

Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

Monty Hall Puzzle

Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

Monty Hall Puzzle

Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

Monty Hall Puzzle

Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

Monty Hall Puzzle

Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

Background: (multiagent) epsitemic logic and probability theory

Qualitative Epistemic language

Let Φ be a set of proposition letters, and Agt a set of agents. Formulas:

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid [i]\varphi$$

where $p \in \Phi$, $i \in Agt$.

- $[i]\varphi$ is read "agents i knows/believes φ
- $\langle i\rangle\varphi \equiv \neg[i]\neg\varphi$ is read "agent i considers φ possible.

Epistemic Models and Semantics

Let Φ be set of proposition letters and Agt a set of agents. An epistemic model is a tuple $M = (W, R, \|\cdot\|)$, where

- W is a set of possible worlds
- R is a collection of relations $R_i \subseteq W^2$ for each $i \in Agt$.
- $\|\cdot\| : \Phi \to \mathcal{P}(W)$.

Define $I_i : \mathcal{P}(W) \to \mathcal{P}(W)$ to be such that $I_i(X) = \{x \in W \mid R_i(x) \subseteq X\}$. The semantics is given by a function $[\![\cdot]\!]$ from formulas to subsets of W.

$$
\begin{array}{lcl}
[\![\top]\!] & = & W \\
[\![p]\!] & = & \|p\| \\
[\![\neg\varphi]\!] & = & W - [\![\varphi]\!] \\
[\![\varphi \wedge \psi]\!] & = & [\![\varphi]\!] \cap [\![\psi]\!] \\
[\![[i]\varphi]\!] & = & I_i([\![\varphi]\!])
\end{array}
$$

Commonly accepted axioms

- (normality) $[i](\varphi \rightarrow \psi) \rightarrow ([i]\varphi \rightarrow [i]\psi)$
- (knowledge; reflexivity) $[i]\varphi \rightarrow \varphi$
- (positive introspection; transitivity) $[i]\varphi \rightarrow [i][i]\varphi$
- (negative introspection; Euclidean) $\neg[i]\varphi \rightarrow [i]\neg[i]\varphi$

Theorem

The above axioms (with Modus Ponens, Necessitation and substitution) are sound and strongly complete with respect to the class of Epistemic Models where each R_i is an equivalence relation.

Commonly accepted axioms

- (normality) $[i](\varphi \to \psi) \to ([i]\varphi \to [i]\psi)$
- (knowledge; reflexivity) $[i]\varphi \to \varphi$
- (positive introspection; transitivity) $[i]\varphi \to [i][i]\varphi$
- (negative introspection; Euclidean) $\neg[i]\varphi \to [i]\neg[i]\varphi$

Theorem

The above axioms (with Modus Ponens, Necessitation and substitution) are sound and strongly complete with respect to the class of Epistemic Models where each R_i is an equivalence relation.

Probability space

Definition (Probability space)

A probability space is a tuple $(S, \mathcal{A}, \mu)$, where

- S is a set:
  S is called the "sample space", its elements "outcomes"
- $\mathcal{A} \subseteq \mathcal{P}(S)$ is a σ-algebra:
  a non-empty set of subsets of S closed under complements and countable unions.
- $\mu : \mathcal{A} \to [0, 1]$ is a probability measure:
  a function satisfying
  - $\mu(S) = 1$ (normalize to 1)
  - $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ for every pairwise disjoint collection of sets $\{A_i\}_{i\in\mathbb{N}}$ in $\mathcal{A}$. (countable additivity)

The sets $A \in \mathcal{A}$ are called "events" or "measurable sets".

Measure space

Definition (Measure space)

A measure space is a tuple $(S, \mathcal{A}, \mu)$, where

- S is a set.
- $\mathcal{A}$ is a σ-algebra.
- $\mu : \mathcal{A} \to [0, \infty]$ is a measure:
  a function satisfying
  - $\mu(\emptyset) = 0$
  - $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ for every pairwise disjoint collection of sets $\{A_i\}_{i \in \mathbb{N}}$ in $\mathcal{A}$. (countable additivity)

A probability measure is just a measure normalized to 1.

Definition (Measurable space)

Given any measure space $(S, \mathcal{A}, \mu)$, the pair $(S, \mathcal{A})$ is a measurable space.

Uniform probability distributions

Definition

The uniform probability distribution μ over an interval $[a, b]$ is given by

$$\mu(A) = \int_A \frac{1}{b-a}\,\mathrm{d}\lambda$$

where λ is the Lebesgue measure.

Important properties of Lebesgue measure:

- $\lambda(A) = \lambda(B)$, whenever $B = \{a + t : a \in A\}$ for some $t \in \mathbb{R}$. (translation invariance)
- $\lambda(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \lambda(A_i)$ for every pairwise disjoint collection of sets $\{A_i\}_{i\in\mathbb{N}}$ in $\mathcal{A}$. (countable additivity)

Vitali sets

One motivation for σ-algeras are Vitali sets.

- Define equivalence $\sim$ over $\mathbb{R}$, where $a \sim b$ iff $a - b \in \mathbb{Q}$.
- Denote an equivalence class with representative a by $[a]$; let $\mathcal{E}$ be set of equivalence classes.
- Let $f : \mathcal{E} \to [0,1]$ be any function for which $[f(E)] = E$.
- A Vitali set is

$$V = \{f(E) : E \in \mathcal{E}\}.$$

- Let $V_q = \{x + q : x \in V\}$ for each $q \in [-1,1] \cap \mathbb{Q}$.
- Then $[0,1] \subseteq \bigcup V_q \subseteq [-1,2]$.

Continuity of a set function

Definition (Continuity from above)

A set function $\mu : \mathcal{A} \to [0, \infty]$ is continuous from above if for every non-increasing sequence $A_1 \supseteq A_2 \supseteq \cdots \in \mathcal{A}$,

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

We say that μ is continuous at $\emptyset$ if $\lim_{n\to\infty} \mu(A_n) = 0$, whenever $\bigcap A_n = \emptyset$.

Theorem

- *Any measure is continuous from above*
- *Any finitely additive set function $\mu : \mathcal{A} \to [0, \infty)$ that is continuous at $\emptyset$ is a measure (is countably additive).*

proof of part 1

- Let $\{A_n\}$ be a decreasing sequence of sets.
- Let $A = \bigcap A_n$.
- Let $B_n = X - A_n$.
- Let $D_0 = B_1$ and $D_{n+1} = B_{n+1} - B_n$.

$$
\begin{aligned}
1 - \mu(A) &= \mu(X) - \mu(A) \\
&= \mu(X - A) = \mu(\bigcup D_n) \\
&= \sum \mu(D_n) = \lim \mu(B_n) \\
&= \lim(\mu(X) - \mu(A_n)) = 1 - \lim \mu(A_n)
\end{aligned}
$$

$=$ from countable additivity (red)

$=$ from finite additivity (blue)

Proof of part 2

- Let $\{A_n\}$ be a sequence pairwise disjoint sets. there is a sequence of disjoint sets.
- Let $A = \bigcup_{i=1}^{\infty} A_i$
- Let $B_n = \bigcup_{i=1}^{n} A_i$
- Then $\bigcap(A - B_n) = \emptyset$
- Then $\lim \mu(A - B_n) = 0$ (by continuity at $\emptyset$)
- Then $\lim(\mu(A) - \mu(B_n)) = 0$ (by finite additivity)
- Then $\lim(\mu(A) - \sum_{i=1}^{n} \mu(A_n)) = 0$ (by finite additivity)
- Then $\lim \sum_{i=1}^{n} \mu(A_n) = \mu(A)$
- Thus $\sum_{i=1}^{\infty} \mu(A_n) = \mu(A) = \mu(\bigcup A_n)$

Product spaces

Definition

Given a family $(X_1, \mathcal{A}_1), \ldots, (X_n, \mathcal{A}_n)$ of measurable spaces, we define the product measurable space to be $(X, \mathcal{A})$, where

- $X = X_1 \times \cdots \times X_n$
- $\mathcal{A}$ is the σ-algebra generated by $\{A_1 \times \cdots \times A_n \mid A_i \in \mathcal{A}_i\}$.

If μ_i is a measure on $(X_i, \mathcal{A}_i)$, for each i, then we define the product measure μ to be

$$\mu(A) = \inf \left\{ \sum_{j=1}^{\infty} \prod_{i=1}^{n} \mu_i(A_i^j) \mid A_i^j \in \mathcal{A}_i \text{ and } A \subseteq \bigcup_{j=1}^{\infty} \prod_{i=1}^{n} A_i^j \right\}.$$

Outer measure

Definition (Outer measure)

Given a set S, an outer measure on S is a function $\mu : \mathcal{P}(S) \to [0, \infty]$, such that

- $\mu(\emptyset) = 0$
- $\mu(A_1) \leq \mu(A_2)$ whenever $A_1 \subseteq A_2$. (monotonicity)
- $\mu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$ for every collection of sets $\{A_i\}_{i \in \mathbb{N}}$ in $\mathcal{P}(S)$. (countable subadditivity)

Measurable sets

Definition (μ-measurable sets)

Given an outer measure $\mu : \mathcal{P}(X) \to [0, \infty]$, a set $A \subseteq X$ is μ-measurable if for every set $B \subseteq X$ if

$$B = \mu(B \cap A) + \mu(B - A).$$

Proposition

Given an outer measure μ, the set of μ-measurable sets forms a σ-algebra.

Definition (Measurable sets of a measure space)

Given a measure space $(S, \mathcal{A}, \mu)$, the set $\mathcal{A}$ consists of all the measurable sets of the space.

From outer measure to measure

Given an outer measure $\mathcal{P}(S) \to [0, \infty]$,

- let $\mathcal{A}$ be the set of μ-measurable sets,
- let $\mu' : \mathcal{A} \to [0, 1]$ such that $\mu'(A) = \mu(A)$ for all $A \in \mathcal{A}$.

It turns out that $(S, \mathcal{A}, \mu')$ is a measure space.

Example: Lebesgue measure

First define the Lebesgue outer measure $\lambda^* : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]$ by

$$\lambda^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \prod_{i=1}^{n} (b_i^j - a_i^j) \mid E \subseteq \bigcup_{j=1}^{\infty} \prod_{i=1}^{n} [a_i^j, b_i^j] \right\}.$$

We define the Lebesgue measure λ to be the restriction of λ^* to the λ^*-measurable sets.

From measure to outer measure

Definition (Outer Measure Extension of a Measure)

If $\mathcal{A}$ is a σ-algebra over S and $\mu : \mathcal{A} \to [0, \infty]$ is a measure, then the *outer measure extension of* μ is defined to be $\mu^* : \mathcal{P}(S) \to [0, \infty]$ given by

$$\mu^*(A) = \inf\{\mu(B) : B \in \mathcal{A}, A \subseteq B\}$$

Proposition

Let $\mathcal{A}$ be a σ-algebra over S, let $\mu : \mathcal{A} \to [0, \infty]$ be a measure, and let $\mu^* : \mathcal{P}(S) \to [0, \infty]$ be the outer measure extension of μ. Then

- $\mu^*(A) = \mu(A)$ *for every* $A \in \mathcal{A}$ *(μ^* does indeed extend μ).*
- μ^* *is an outer measure.*
- *If $\mathcal{A}'$ consists of the μ^*-measurable sets, then $\mathcal{A} \subseteq \mathcal{A}'$.*