Reasoning under uncertainty

David “davidad” Dalrymple
davidad@alum.mit.edu

August 20, 2017
http://espr.cf

Reasoning under uncertainty (Foundations of probability theory)

David "davidad" Dalrymple
davidad@alum.mit.edu

August 20, 2017
http://espr.cf

Reasoning without uncertainty

Reasoning without uncertainty

1. All synapses are either electrical or chemical

Reasoning without uncertainty

1. All synapses are either electrical or chemical
2. All chemical synapses are intercellular

Reasoning without uncertainty

1. All synapses are either electrical or chemical
2. All chemical synapses are intercellular
3. If a synapse is not intercellular, it is not an electrical synapse

Reasoning without uncertainty

1. All synapses are either electrical or chemical
2. All chemical synapses are intercellular
3. If a synapse is not intercellular, it is not an electrical synapse
4. Are all synapses intercellular?

Reasoning without uncertainty

1. All synapses are either electrical or chemical
2. All chemical synapses are intercellular
3. If a synapse is not intercellular, it is not an electrical synapse
4. Are all synapses intercellular (yes or no)?

Reasoning without uncertainty

Reasoning without uncertainty

1. 3 boxes are labeled “Red balls,” “Green balls,” and “Red and green balls.”

Reasoning without uncertainty

1. 3 boxes are labeled “Red balls,” “Green balls,” and “Red and green balls.”
2. All three labels are incorrect;

Reasoning without uncertainty

1. 3 boxes are labeled “Red balls,” “Green balls,” and “Red and green balls.”
2. All three labels are incorrect; each belongs on a different box.

Reasoning without uncertainty

1. 3 boxes are labeled “Red balls,” “Green balls,” and “Red and green balls.”
2. All three labels are incorrect; each belongs on a different box.
3. You may ask me to take one ball from one box.

Reasoning without uncertainty

1. 3 boxes are labeled “Red balls,” “Green balls,” and “Red and green balls.”
2. All three labels are incorrect; each belongs on a different box.
3. You may ask me to take one ball from one box.
4. How can you label the boxes correctly?

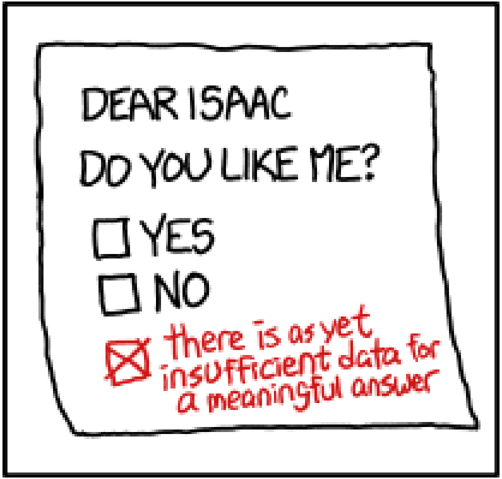
Reasoning with uncertainty

1. I have a full deck of 52 playing cards, and take the top card.

Reasoning with uncertainty

1. I have a full deck of 52 playing cards, and take the top card.
2. Is it a queen?

What's the next best thing?



http://xkcd.com/1448; reference is to “The Last Question” by Isaac Asimov

What's the next best thing?

Different framings of this question:

1. **Betting strategy**
2. **Subjective belief**
3. **Objective frequency**
4. **Many (deterministic) worlds**

What's the next best thing?

Different framings of this question:

1. **Betting strategy**
   - You're required to set prices for bets. How can you not let an equally-informed opponent guarantee a profit off you?
2. **Subjective belief**
3. **Objective frequency**
4. **Many (deterministic) worlds**

What's the next best thing?

Different framings of this question:

1. Betting strategy
   - You're required to set prices for bets. How can you not let an equally-informed opponent guarantee a profit off you?
2. **Subjective belief**
   - Assign quantitative "degrees of belief" instead of truth values, as an extension of logic.
3. Objective frequency
4. Many (deterministic) worlds

What's the next best thing?

Different framings of this question:

1. Betting strategy
   - You're required to set prices for bets. How can you not let an equally-informed opponent guarantee a profit off you?
2. Subjective belief
   - Assign quantitative "degrees of belief" instead of truth values, as an extension of logic.
3. **Objective frequency**
   - Consider hypothetically repeating a random experiment (infinitely) many times, and counting the number of outcomes where a proposition is true.
4. Many (deterministic) worlds

What's the next best thing?

Different framings of this question:

1. Betting strategy
   - You're required to set prices for bets. How can you not let an equally-informed opponent guarantee a profit off you?
2. Subjective belief
   - Assign quantitative "degrees of belief" instead of truth values, as an extension of logic.
3. Objective frequency
   - Consider hypothetically repeating a random experiment (infinitely) many times, and counting the number of outcomes where a proposition is true.
4. **Many (deterministic) worlds**
   - Any given world would have enough information for deduction, but we don't know which world we're in. Count the number of worlds where a proposition is true.

What's the next best thing?

Different framings of this question:

1. **Betting strategy**
   - You're required to set prices for bets. How can you not let an equally-informed opponent guarantee a profit off you?
2. **Subjective belief**
   - Assign quantitative "degrees of belief" instead of truth values, as an extension of logic.
3. **Objective frequency**
   - Consider hypothetically repeating a random experiment (infinitely) many times, and counting the number of outcomes where a proposition is true.
4. **Many (deterministic) worlds**
   - Any given world would have enough information for deduction, but we don't know which world we're in. Count the number of worlds where a proposition is true.

What's the next best thing?

Different framings of this question:

1. **Betting strategy**
   - You're required to set prices for bets. How can you not let an equally-informed opponent guarantee a profit off you?
2. **Subjective belief**
   - Assign quantitative "degrees of belief" instead of truth values, as an extension of logic.
3. **Objective frequency**
   - Consider hypothetically repeating a random experiment (infinitely) many times, and counting the number of outcomes where a proposition is true.
4. **Many (deterministic) worlds**
   - Any given world would have enough information for deduction, but we don't know which world we're in. Count the number of worlds where a proposition is true.

Big takeaway: these framings **all justify equivalent formal theories**, i.e. probability!

What's the next best thing?

Different framings of this question:

1. **Betting strategy**
   - You're required to set prices for bets. How can you not let an equally-informed opponent guarantee a profit off you?
2. **Subjective belief**
   - Assign quantitative "degrees of belief" instead of truth values, as an extension of logic.
3. **Objective frequency**
   - Consider hypothetically repeating a random experiment (infinitely) many times, and counting the number of outcomes where a proposition is true.
4. **Many (deterministic) worlds**
   - Any given world would have enough information for deduction, but we don't know which world we're in. Count the number of worlds where a proposition is true.

Big takeaway: these framings **all justify equivalent formal theories**, i.e. probability!

Note: there are other framings that justify different theories, like Dempster-Shafer theory, but this concordance is still surprising.

Boolean algebra: operators and notation

Logic		Algebra		Sets	
conjunction	$\wedge$	multiplication	$\cdot$	intersection	$\cap$
disjunction	$\vee$	addition	$+$	union	$\cup$
negation	$\neg$	negation	$\neg$	complement	$\Omega\backslash$

Boolean algebra: axioms

$\top$ means true; $\perp$ means false

$$A \vee \perp = A \qquad A \wedge \top = A$$

$$A \vee B = B \vee A \qquad A \wedge B = B \wedge A$$

$$A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C) \qquad A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$$

$$A \vee \neg A = \top \qquad A \wedge \neg A = \perp$$

Some probability notation

- $\mathbb{P}(A|B)$ is the "conditional probability" of a proposition (or "event") A given certain knowledge of another proposition B.

Some probability notation

- $\mathbb{P}(A|B)$ is the "conditional probability" of a proposition (or "event") A given certain knowledge of another proposition B.
- $\mathbb{P}(A)$ is the "unconditional probability" of a proposition, often interpreted as $\mathbb{P}(A|\Omega)$, where Ω is interpreted as a set of assumptions considered to be unconditional.

Some probability notation

- $\mathbb{P}(A|B)$ is the "conditional probability" of a proposition (or "event") A given certain knowledge of another proposition B.
- $\mathbb{P}(A)$ is the "unconditional probability" of a proposition, often interpreted as $\mathbb{P}(A|\Omega)$, where Ω is interpreted as a set of assumptions considered to be unconditional.
- Comma is sometimes used for $\wedge$, e.g. $\mathbb{P}(A,B|\Omega) := \mathbb{P}(A \wedge B|\Omega)$.

Derivations from each justification

In each framing/justification, I'll state the specific laws of probability, then go over the argument.

Betting strategy: laws of probability

1. (Nonnegativity) $\mathbb{P}(A|\Omega) \geq 0$

Betting strategy: laws of probability

1. (Nonnegativity) $\mathbb{P}(A|\Omega) \geq 0$
2. (Tautology) $\mathbb{P}(A|\Omega) = 1$ if A is a logical truth given Ω

Betting strategy: laws of probability

1. (Nonnegativity) $\mathbb{P}(A|\Omega) \geq 0$
2. (Tautology) $\mathbb{P}(A|\Omega) = 1$ if A is a logical truth given Ω
3. (Countable additivity) For arbitrary disjoint A_i,
   $\mathbb{P}\left(\bigvee_{i=1}^{\infty} A_i \middle| \Omega\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i|\Omega)$

Betting strategy: laws of probability

1. (Nonnegativity) $\mathbb{P}(A|\Omega) \geq 0$
2. (Tautology) $\mathbb{P}(A|\Omega) = 1$ if A is a logical truth given Ω
3. (Countable additivity) For arbitrary disjoint A_i,
   $\mathbb{P}\left(\bigvee_{i=1}^{\infty} A_i \middle| \Omega\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i|\Omega)$
4. (Conditional) $\mathbb{P}(A|B,\Omega) = \frac{\mathbb{P}(A \wedge B|\Omega)}{\mathbb{P}(B|\Omega)}$

Betting strategy: premises

A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Betting strategy: premises

A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

The rules of the game:

Betting strategy: premises

A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

The rules of the game:

1. Agent must state $p(A)$ for all A, such that Agent would buy or sell arbitrarily (possibly fractionally) many bets on A with probability $p(A)$

Betting strategy: premises

A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

The rules of the game:

1. Agent must state $p(A)$ for all A, such that Agent would buy or sell arbitrarily (possibly fractionally) many bets on A with probability $p(A)$
2. Bookie may then buy or sell Agent's bets on any As

Betting strategy: premises

A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

The rules of the game:

1. Agent must state $p(A)$ for all A, such that Agent would buy or sell arbitrarily (possibly fractionally) many bets on A with probability $p(A)$
2. Bookie may then buy or sell Agent's bets on any As
3. Finally, an observation is made and the bets are settled

Betting strategy: premises

A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

The rules of the game:

1. Agent must state $p(A)$ for all A, such that Agent would buy or sell arbitrarily (possibly fractionally) many bets on A with probability $p(A)$
2. Bookie may then buy or sell Agent's bets on any As
3. Finally, an observation is made and the bets are settled

The goal as Agent is to set your probabilities in such a way that Bookie cannot *guarantee* that they will make a profit from you.

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A|\Omega) \not\geq 0$.
How can you guarantee a profit as Bookie?

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A|\Omega) \not\geq 0$.
How can you guarantee a profit as Bookie?

- Bookie buys Agent's bet for $\mathbb{P}(A|\Omega)$ dollars, a strictly negative amount,

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A|\Omega) \ngeq 0$.
How can you guarantee a profit as Bookie?

- Bookie buys Agent's bet for $\mathbb{P}(A|\Omega)$ dollars, a strictly negative amount, which means effectively Agent must pay Bookie a strictly positive amount.

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A|\Omega) \not\geq 0$.
How can you guarantee a profit as Bookie?

- Bookie buys Agent's bet for $\mathbb{P}(A|\Omega)$ dollars, a strictly negative amount, which means effectively Agent must pay Bookie a strictly positive amount.
  - If A turns out to be true, Agent must pay Bookie an additional dollar, but in either case Bookie has made a profit.

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(\text{tautology}|\Omega) > 1$.
How can you guarantee a profit as Bookie?

Betting strategy: opponent strategies

For reference:

A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(\text{tautology}|\Omega) > 1$.
How can you guarantee a profit as Bookie?

- Bookie sells this bet to Agent. Agent must pay strictly more than 1 dollar up front. Since it's a tautology, Bookie must return 1 dollar, but still makes a strictly positive profit.

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(\text{tautology}|\Omega) < 1$.
How can you guarantee a profit as Bookie?

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(\text{tautology}|\Omega) < 1$.
How can you guarantee a profit as Bookie?

- Bookie buys Agent's bet for less than 1 dollar, and Agent must pay 1 dollar.

Betting strategy: opponent strategies

For reference:

A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A \vee B|\Omega) > \mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$ for disjoint A and B. How can you guarantee a profit as Bookie?

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A \vee B|\Omega) > \mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$ for disjoint A and B. How can you guarantee a profit as Bookie?

- Bookie buys Agent's bets for A and B at a cost of $\mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$. Bookie then sells Agent a bet for $A \vee B$ and Agent pays $\mathbb{P}(A \vee B|\Omega)$.

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A \vee B|\Omega) > \mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$ for disjoint A and B. How can you guarantee a profit as Bookie?

- Bookie buys Agent's bets for A and B at a cost of $\mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$. Bookie then sells Agent a bet for $A \vee B$ and Agent pays $\mathbb{P}(A \vee B|\Omega)$.
- Bookie has made a profit, and if A and B both turn out to be false then no further money changes hands, but if either do turn out to be true then Bookie and Agent owe each other a dollar.

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A \vee B|\Omega) > \mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$ for disjoint A and B. How can you guarantee a profit as Bookie?

- Bookie buys Agent's bets for A and B at a cost of $\mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$. Bookie then sells Agent a bet for $A \vee B$ and Agent pays $\mathbb{P}(A \vee B|\Omega)$.
- Bookie has made a profit, and if A and B both turn out to be false then no further money changes hands, but if either do turn out to be true then Bookie and Agent owe each other a dollar.
  - Why do A and B have to be disjoint?

Betting strategy: opponent strategies

For reference:

A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A \vee B|\Omega) > \mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$ for disjoint A and B. How can you guarantee a profit as Bookie?

- Bookie buys Agent's bets for A and B at a cost of $\mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$. Bookie then sells Agent a bet for $A \vee B$ and Agent pays $\mathbb{P}(A \vee B|\Omega)$.
- Bookie has made a profit, and if A and B both turn out to be false then no further money changes hands, but if either do turn out to be true then Bookie and Agent owe each other a dollar.
  - Why do A and B have to be disjoint?
- What if Agent's $\mathbb{P}(A \vee B|\Omega) < \mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$?

Betting strategy: opponent strategies

For reference:

A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A \vee B|\Omega) > \mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$ for disjoint A and B. How can you guarantee a profit as Bookie?

- Bookie buys Agent's bets for A and B at a cost of $\mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$. Bookie then sells Agent a bet for $A \vee B$ and Agent pays $\mathbb{P}(A \vee B|\Omega)$.
- Bookie has made a profit, and if A and B both turn out to be false then no further money changes hands, but if either do turn out to be true then Bookie and Agent owe each other a dollar.
  - Why do A and B have to be disjoint?
- What if Agent's $\mathbb{P}(A \vee B|\Omega) < \mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$?
  - Swap buying with selling.

Betting strategy: opponent strategies

For reference:
A **bet** on A with probability p is this contract: buyer pays \$$p$ up front to seller; if A is observed to be true, seller pays \$1 to buyer (else, nothing)

Suppose that Agent's $\mathbb{P}(A \vee B|\Omega) > \mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$ for disjoint A and B. How can you guarantee a profit as Bookie?

- Bookie buys Agent's bets for A and B at a cost of $\mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$. Bookie then sells Agent a bet for $A \vee B$ and Agent pays $\mathbb{P}(A \vee B|\Omega)$.
- Bookie has made a profit, and if A and B both turn out to be false then no further money changes hands, but if either do turn out to be true then Bookie and Agent owe each other a dollar.
  - Why do A and B have to be disjoint?
- What if Agent's $\mathbb{P}(A \vee B|\Omega) < \mathbb{P}(A|\Omega) + \mathbb{P}(B|\Omega)$?
  - Swap buying with selling.
- This can be generalized to arbitrarily many disjoint events.

Betting strategy: opponent strategies

Suppose that if evidence B is observed, Agent will update to a $\mathbb{P}(A|B,\Omega) \neq \frac{\mathbb{P}(A \wedge B|\Omega)}{\mathbb{P}(B|\Omega)}$.
How can you guarantee a profit as Bookie?

Betting strategy: opponent strategies

Suppose that if evidence B is observed, Agent will update to a $\mathbb{P}(A|B,\Omega) < \frac{\mathbb{P}(A \wedge B|\Omega)}{\mathbb{P}(B|\Omega)}$. How can you guarantee a profit as Bookie?

- Bet 1. Bookie sells Agent a bet for $A \wedge B$; Agent pays $\mathbb{P}(A \wedge B|\Omega)$.
- Bet 2. Bookie sells Agent $\frac{\mathbb{P}(A \wedge B|\Omega)}{\mathbb{P}(B|\Omega)}$ of a bet for $\neg B$; Agent pays $\mathbb{P}(A \wedge B|\Omega)\,\frac{\mathbb{P}(\neg B|\Omega)}{\mathbb{P}(B|\Omega)}$.
- Bet 3. Bookie sells Agent $\frac{\mathbb{P}(A \wedge B|\Omega)}{\mathbb{P}(B|\Omega)} - \mathbb{P}(A|B,\Omega)$ of a bet for B; Agent pays $\mathbb{P}(A \wedge B|\Omega) - \mathbb{P}(A|B,\Omega)\,\mathbb{P}(B|\Omega)$.
- If B is false, then all the bets settle. Bookie pays back $\frac{\mathbb{P}(A \wedge B|\Omega)}{\mathbb{P}(B|\Omega)}$ for Bet 2. Due to countable additivity, Bets 1 and 2 exactly cancel out and the price for Bet 3 remains as profit.
- If B is true, then Bets 2 and 3 are settled, with Bookie paying back $\frac{\mathbb{P}(A \wedge B|\Omega)}{\mathbb{P}(B|\Omega)} - \mathbb{P}(A|B,\Omega)$ for Bet 3.
- Bet 4. Bookie then buys from Agent a bet for A at Agent's new probability of $\mathbb{P}(A|B,\Omega)$.
- At this point, whether or not A turns out to be true, Bookie and Agent owe each other a net 0 dollars, and Bookie has made a net profit.

Betting strategy: laws of probability

1. (Nonnegativity) $\mathbb{P}(A|\Omega) \geq 0$
2. (Tautology) $\mathbb{P}(A|\Omega) = 1$ if A is a logical truth given Ω
3. (Countable additivity) For arbitrary disjoint A_i,
   $\mathbb{P}\left(\bigvee_{i=1}^{\infty} A_i \middle| \Omega\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i|\Omega)$
4. (Conditional) $\mathbb{P}(A|B,\Omega) = \frac{\mathbb{P}(A \wedge B|\Omega)}{\mathbb{P}(B|\Omega)}$

Subjective belief: laws of probability

1. (Bounding) $0 = \mathbb{P}(\bot|\Omega) \leq \mathbb{P}(A|\Omega) \leq \mathbb{P}(\Omega|\Omega) = 1$

Subjective belief: laws of probability

1. (Bounding) $0 = \mathbb{P}(\perp|\Omega) \leq \mathbb{P}(A|\Omega) \leq \mathbb{P}(\Omega|\Omega) = 1$
2. (Negation) $\mathbb{P}(A|\Omega) = 1 - \mathbb{P}(\neg A|\Omega)$

Subjective belief: laws of probability

1. (Bounding) $0 = \mathbb{P}(\bot|\Omega) \leq \mathbb{P}(A|\Omega) \leq \mathbb{P}(\Omega|\Omega) = 1$
2. (Negation) $\mathbb{P}(A|\Omega) = 1 - \mathbb{P}(\neg A|\Omega)$
3. (Product rule) $\mathbb{P}(A \wedge B|\Omega) = \mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega)$

Subjective belief: laws of probability

1. (Bounding) $0 = \mathbb{P}(\bot|\Omega) \leq \mathbb{P}(A|\Omega) \leq \mathbb{P}(\Omega|\Omega) = 1$
2. (Negation) $\mathbb{P}(A|\Omega) = 1 - \mathbb{P}(\neg A|\Omega)$
3. (Product rule) $\mathbb{P}(A \wedge B|\Omega) = \mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega)$
4. (Countable additivity) For arbitrary disjoint A_i, $\mathbb{P}\left(\bigvee_{i=1}^{\infty} A_i \middle| \Omega\right) B = \sum_{i=1}^{\infty} \mathbb{P}(A_i|\Omega)$

Subjective belief: premises

1. $\mathbb{P}(A|B)$ is a real number, and A and B are elements of a Boolean algebra

Subjective belief: premises

1. $\mathbb{P}(A|B)$ is a real number, and A and B are elements of a Boolean algebra
    - Technically, a complete Boolean algebra

Subjective belief: premises

1. $\mathbb{P}(A|B)$ is a real number, and A and B are elements of a Boolean algebra
   - Technically, a complete Boolean algebra, which essentially means that countably infinite disjunctions (and/or conjunctions) are allowed.

Subjective belief: premises

1. $\mathbb{P}(A|B)$ is a real number, and A and B are elements of a Boolean algebra
   - Technically, a complete Boolean algebra, which essentially means that countably infinite disjunctions (and/or conjunctions) are allowed.
   - Any (incomplete) Boolean algebra (or any topology!) generates a canonical complete Boolean algebra.

Subjective belief: premises

1. $\mathbb{P}(A|B)$ is a real number, and A and B are elements of a Boolean algebra
   - Technically, a complete Boolean algebra, which essentially means that countably infinite disjunctions (and/or conjunctions) are allowed.
   - Any (incomplete) Boolean algebra (or any topology!) generates a canonical complete Boolean algebra.
2. $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ such that $A_i \nearrow A$ implies $\mathbb{P}(A_i|\Omega) \nearrow \mathbb{P}(A|\Omega)$.

Subjective belief: premises

1. $\mathbb{P}(A|B)$ is a real number, and A and B are elements of a Boolean algebra
   - Technically, a complete Boolean algebra, which essentially means that countably infinite disjunctions (and/or conjunctions) are allowed.
   - Any (incomplete) Boolean algebra (or any topology!) generates a canonical complete Boolean algebra.
2. $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ such that $A_i \nearrow A$ implies $\mathbb{P}(A_i|\Omega) \nearrow \mathbb{P}(A|\Omega)$.
3. $\mathbb{P}(A \wedge B|\Omega) = F[\mathbb{P}(A|B,\Omega), \mathbb{P}(B|\Omega)]$ for some F.

Subjective belief: premises

1. $\mathbb{P}(A|B)$ is a real number, and A and B are elements of a Boolean algebra
   - Technically, a complete Boolean algebra, which essentially means that countably infinite disjunctions (and/or conjunctions) are allowed.
   - Any (incomplete) Boolean algebra (or any topology!) generates a canonical complete Boolean algebra.
2. $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ such that $A_i \nearrow A$ implies $\mathbb{P}(A_i|\Omega) \nearrow \mathbb{P}(A|\Omega)$.
3. $\mathbb{P}(A \wedge B|\Omega) = F[\mathbb{P}(A|B,\Omega), \mathbb{P}(B|\Omega)]$ for some F.
4. $\mathbb{P}(\neg A|\Omega) = N[\mathbb{P}(A|\Omega)]$ for some N.

Subjective belief: proof

- Cox “proved” something like this in 1946, known as Cox’s Theorem.

Subjective belief: proof

- Cox "proved" something like this in 1946, known as Cox's Theorem.
- It turns out Cox's Theorem is wrong!

Subjective belief: proof

- Cox “proved” something like this in 1946, known as Cox’s Theorem.
- It turns out Cox’s Theorem is wrong!
  - Factoid of the hour: “Cox’s Theorem is false; Halpern in 1999 provided a counterexample”

Subjective belief: proof

- Cox "proved" something like this in 1946, known as Cox's Theorem.
- It turns out Cox's Theorem is wrong!
  - Factoid of the hour: "Cox's Theorem is false; Halpern in 1999 provided a counterexample"
  - The version on Wikipedia is also wrong

Subjective belief: proof

- Cox "proved" something like this in 1946, known as Cox's Theorem.
- It turns out Cox's Theorem is wrong!
  - Factoid of the hour: "Cox's Theorem is false; Halpern in 1999 provided a counterexample"
  - The version on Wikipedia is also wrong
- But the laws of probability I showed here *do* follow from the premises I showed here, in a modern proof by Terenin and Draper

Subjective belief: laws of probability

1. (Bounding) $0 = \mathbb{P}(\bot|\Omega) \leq \mathbb{P}(A|\Omega) \leq \mathbb{P}(\Omega|\Omega) = 1$
2. (Negation) $\mathbb{P}(A|\Omega) = 1 - \mathbb{P}(\neg A|\Omega)$
3. (Product rule) $\mathbb{P}(A \wedge B|\Omega) = \mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega)$
4. (Countable additivity) For arbitrary disjoint A_i,
   $\mathbb{P}\left(\bigvee_{i=1}^{\infty} A_i \middle| \Omega\right) B = \sum_{i=1}^{\infty} \mathbb{P}(A_i|\Omega)$

Objective frequency: laws of probability

1. (Nonnegativity) $\mathbb{P}(A) \geq 0$

Objective frequency: laws of probability

1. (Nonnegativity) $\mathbb{P}(A) \geq 0$
2. (Normalization) $\mathbb{P}(\Omega) = 1$

Objective frequency: laws of probability

1. (Nonnegativity) $\mathbb{P}(A) \geq 0$
2. (Normalization) $\mathbb{P}(\Omega) = 1$
3. (Countable additivity) For arbitrary disjoint A_i,
   $\mathbb{P}\left(\bigvee_{i=1}^{\infty} A_i \middle| \Omega\right) B = \sum_{i=1}^{\infty} \mathbb{P}(A_i|\Omega)$

Objective frequency: laws of probability

1. (Nonnegativity) $\mathbb{P}(A) \geq 0$
2. (Normalization) $\mathbb{P}(\Omega) = 1$
3. (Countable additivity) For arbitrary disjoint A_i,
   $\mathbb{P}\left(\bigvee_{i=1}^{\infty} A_i \middle| \Omega\right) B = \sum_{i=1}^{\infty} \mathbb{P}(A_i|\Omega)$
4. (Conditional) $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \wedge B)}{\mathbb{P}(B)}$

Objective frequency: premises

- Consider set of outcomes of (infinitely) repeated independent experiments

Objective frequency: premises

- Consider set of outcomes of (infinitely) repeated independent experiments
- "Probability" $\mathbb{P}(A)$ is defined as the ratio N_A/N of outcomes where A is true (N_A) to total number of experiments (N).

Objective frequency: proofs

- Nonnegativity

Objective frequency: proofs

- Nonnegativity
  - Both the total number of experiments, and the number of outcomes in which the proposition is true, cannot be negative, therefore their ratio cannot be negative.

Objective frequency: proofs

- Nonnegativity
  - Both the total number of experiments, and the number of outcomes in which the proposition is true, cannot be negative, therefore their ratio cannot be negative.
- Normalization

Objective frequency: proofs

- Nonnegativity
  - Both the total number of experiments, and the number of outcomes in which the proposition is true, cannot be negative, therefore their ratio cannot be negative.
- Normalization
  - The number of experiments in which an experiment was performed is identical to the total number of experiments, so their ratio is 1.

Objective frequency: proofs

- Nonnegativity
  - Both the total number of experiments, and the number of outcomes in which the proposition is true, cannot be negative, therefore their ratio cannot be negative.
- Normalization
  - The number of experiments in which an experiment was performed is identical to the total number of experiments, so their ratio is 1.
- Countable additivity

Objective frequency: proofs

- Nonnegativity
  - Both the total number of experiments, and the number of outcomes in which the proposition is true, cannot be negative, therefore their ratio cannot be negative.
- Normalization
  - The number of experiments in which an experiment was performed is identical to the total number of experiments, so their ratio is 1.
- Countable additivity
  - The number of outcomes in which one of many mutually exclusive alternatives is true is precisely the number of outcomes in which at least one is true.

Objective frequency: proofs

- Conditional

Objective frequency: proofs

- Conditional
  - We can justify this by the principle of rejection:

Objective frequency: proofs

- Conditional
  - We can justify this by the principle of rejection: to perform an experiment "conditional" on the knowledge of B is simply to reject (exclude from consideration) every experiment in which B is false.

Objective frequency: proofs

- Conditional
    - We can justify this by the principle of rejection: to perform an experiment "conditional" on the knowledge of B is simply to reject (exclude from consideration) every experiment in which B is false.
    - Thus, the (effective) total number of experiments is the number of experiments in which B is true, and the (effective) total number of experiments in which A is true is those in which both A and B are true.

Objective frequency: proofs

- Conditional
  - We can justify this by the principle of rejection: to perform an experiment "conditional" on the knowledge of B is simply to reject (exclude from consideration) every experiment in which B is false.
  - Thus, the (effective) total number of experiments is the number of experiments in which B is true, and the (effective) total number of experiments in which A is true is those in which both A and B are true.
  - So

$$\mathbb{P}(A|B) = \frac{N_{A \wedge B}}{N_B} = \frac{N_{A \wedge B}/N}{N_B/N} = \frac{\mathbb{P}(A \wedge B)}{\mathbb{P}(B)}$$

Many (deterministic) worlds

- Instead of considering a sequence of independent experiments, we can just as well consider an un-ordered set of possible worlds.

Many (deterministic) worlds

- Instead of considering a sequence of independent experiments, we can just as well consider an un-ordered set of possible worlds.
- The same arguments as for objective frequency go through.

Many (deterministic) worlds

- Instead of considering a sequence of independent experiments, we can just as well consider an un-ordered set of possible worlds.
- The same arguments as for objective frequency go through.
- You may prefer one or the other philosophically.

Many (deterministic) worlds

- Instead of considering a sequence of independent experiments, we can just as well consider an un-ordered set of possible worlds.
- The same arguments as for objective frequency go through.
- You may prefer one or the other philosophically.
- Bernoulli even used both: one layer of uncertainty due to ignorance, and one due to randomness.

Many (deterministic) worlds

- Instead of considering a sequence of independent experiments, we can just as well consider an un-ordered set of possible worlds.
- The same arguments as for objective frequency go through.
- You may prefer one or the other philosophically.
- Bernoulli even used both: one layer of uncertainty due to ignorance, and one due to randomness.
- Formally, they behave the same and blend together.

Sample space

- If we take the objective-frequency or many-worlds interpretations, we call the experimental outcomes or possible worlds as elements of a **sample space**, which is notated Ω and is an object of set theory.

Sample space

- If we take the objective-frequency or many-worlds interpretations, we call the experimental outcomes or possible worlds as elements of a **sample space**, which is notated Ω and is an object of set theory.
- The subjective-belief and betting-strategy frameworks are sufficiently abstract that they don't come with a concrete sample space, but a canonical one can be constructed (via the Loomis-Sikorski representation theorem).

Derivation of Bayes' Theorem

Using the **Product Rule:** $\mathbb{P}(A \wedge B | \Omega) = \mathbb{P}(A | B, \Omega) \cdot \mathbb{P}(B | \Omega)$

Derivation of Bayes' Theorem

Using the **Product Rule:** $\mathbb{P}(A \wedge B|\Omega) = \mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega)$

$$\mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega) = P(A \wedge B) = P(B \wedge A) = \mathbb{P}(B|A,\Omega) \cdot \mathbb{P}(A|\Omega)$$

$$\mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega) = \mathbb{P}(B|A,\Omega) \cdot \mathbb{P}(A|\Omega)$$

Derivation of Bayes' Theorem

Using the **Product Rule:** $\mathbb{P}(A \wedge B|\Omega) = \mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega)$

$$\mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega) = P(A \wedge B) = P(B \wedge A) = \mathbb{P}(B|A,\Omega) \cdot \mathbb{P}(A|\Omega)$$

$$\mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega) = \mathbb{P}(B|A,\Omega) \cdot \mathbb{P}(A|\Omega)$$

Bayes: $\mathbb{P}(A|B,\Omega) = \dfrac{\mathbb{P}(B|A,\Omega) \cdot \mathbb{P}(A|\Omega)}{\mathbb{P}(B|\Omega)}$

Derivation of Bayes' Theorem

Using the **Product Rule:** $\mathbb{P}(A \wedge B|\Omega) = \mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega)$

$$\mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega) = P(A \wedge B) = P(B \wedge A) = \mathbb{P}(B|A,\Omega) \cdot \mathbb{P}(A|\Omega)$$

$$\mathbb{P}(A|B,\Omega) \cdot \mathbb{P}(B|\Omega) = \mathbb{P}(B|A,\Omega) \cdot \mathbb{P}(A|\Omega)$$

Bayes: $\mathbb{P}(A|B,\Omega) = \dfrac{\mathbb{P}(B|A,\Omega) \cdot \mathbb{P}(A|\Omega)}{\mathbb{P}(B|\Omega)}$

Why is this useful?

Generative vs. discriminative models

Suppose you're a geologist examining a rock. You can measure its density but want to know its composition.

Generative vs. discriminative models

Suppose you're a geologist examining a rock. You can measure its density but want to know its composition. Let's set propositions

$$A = \text{the rock contains iron.}$$

$$B = \text{the rock's density is } 3{,}160\ \mathrm{kg/m^3}.$$

Generative vs. discriminative models

Suppose you're a geologist examining a rock. You can measure its density but want to know its composition. Let's set propositions

$$\begin{aligned} A &= \text{the rock contains iron.} \\ B &= \text{the rock's density is } 3{,}160\ \mathrm{kg/m^3}. \end{aligned}$$

We say that $\mathbb{P}(A|B,\Omega)$ is a **discriminative model** (which is usually what you want), and $\mathbb{P}(B|A,\Omega)$ is a **generative model** (which is usually easier to make scientifically).

Generative vs. discriminative models

Suppose you're a geologist examining a rock. You can measure its density but want to know its composition. Let's set propositions

$$A = \text{the rock contains iron.}$$
$$B = \text{the rock's density is } 3{,}160\ \mathrm{kg/m^3}.$$

We say that $\mathbb{P}(A|B,\Omega)$ is a **discriminative model** (which is usually what you want), and $\mathbb{P}(B|A,\Omega)$ is a **generative model** (which is usually easier to make scientifically).

Bayes lets you convert between the two.

$$\textbf{Bayes:}\ \mathbb{P}(A|B,\Omega) = \frac{\mathbb{P}(B|A,\Omega)\cdot\mathbb{P}(A|\Omega)}{\mathbb{P}(B|\Omega)}$$

Bayes caveat

$$A = \text{the rock contains iron.}$$
$$B = \text{the rock's density is } 3{,}160\ \text{kg/m}^3.$$

$$\textbf{Bayes: } \mathbb{P}(A|B,\Omega) = \frac{\mathbb{P}(B|A,\Omega)\cdot\mathbb{P}(A|\Omega)}{\mathbb{P}(B|\Omega)}$$

In the cases where we want to use Bayes, $\mathbb{P}(B|\Omega)$ is usually impossible to compute, because one must consider *all possible hypotheses* for how B might be true, and we're bounded agents.

Bayes caveat

$$A = \text{the rock contains iron.}$$
$$B = \text{the rock's density is } 3{,}160 \text{ kg/m}^3.$$

$$\textbf{Bayes: } \mathbb{P}(A|B,\Omega) = \frac{\mathbb{P}(B|A,\Omega)\cdot\mathbb{P}(A|\Omega)}{\mathbb{P}(B|\Omega)}$$

In the cases where we want to use Bayes, $\mathbb{P}(B|\Omega)$ is usually impossible to compute, because one must consider *all possible hypotheses* for how B might be true, and we're bounded agents. Instead, we generally use Bayes to *compare* hypotheses.

$$\frac{P(A_1|B)}{P(A_2|B)} = \frac{P(B|A_1)\cdot P(A_1)/P(B)}{P(B|A_2)\cdot P(A_2)/P(B)}$$
$$\frac{P(A_1|B)}{P(A_2|B)} = \frac{P(B|A_1)}{P(B|A_2)}\cdot\frac{P(A_1)}{P(A_2)}$$

Bayes caveat

$$A = \text{the rock contains iron.}$$
$$B = \text{the rock's density is } 3{,}160\ \text{kg/m}^3.$$

$$\textbf{Bayes: } \mathbb{P}(A|B,\Omega) = \frac{\mathbb{P}(B|A,\Omega)\cdot\mathbb{P}(A|\Omega)}{\mathbb{P}(B|\Omega)}$$

In the cases where we want to use Bayes, $\mathbb{P}(B|\Omega)$ is usually impossible to compute, because one must consider *all possible hypotheses* for how B might be true, and we're bounded agents. Instead, we generally use Bayes to *compare* hypotheses.

$$\frac{P(A_1|B)}{P(A_2|B)} = \frac{P(B|A_1)\cdot P(A_1)/P(B)}{P(B|A_2)\cdot P(A_2)/P(B)}$$
$$\frac{P(A_1|B)}{P(A_2|B)} = \frac{P(B|A_1)}{P(B|A_2)}\cdot\frac{P(A_1)}{P(A_2)}$$

Generating good candidate hypotheses is much of the difficulty in most real inference problems.

Definition of expected value

- If V is a real-valued function defined on the sample space, the **expected value** $\mathbb{E}(V)$ is defined as the Lebesgue integral $\int_\Omega V(x)\, d\mathbb{P}(x|\Omega)$.

Definition of expected value

- If V is a real-valued function defined on the sample space, the **expected value** $\mathbb{E}(V)$ is defined as the Lebesgue integral $\int_\Omega V(x)\, d\mathbb{P}(x|\Omega)$.
- In cases where a discrete summation $\sum_{x\in\Omega} V(x) \cdot \mathbb{P}(x|\Omega)$ applies, this agrees with the Lebesgue integral.

Definition of expected value

- If V is a real-valued function defined on the sample space, the **expected value** $\mathbb{E}(V)$ is defined as the Lebesgue integral $\int_{\Omega} V(x)\,\mathrm{d}\mathbb{P}(x|\Omega)$.
- In cases where a discrete summation $\sum_{x\in\Omega} V(x)\cdot\mathbb{P}(x|\Omega)$ applies, this agrees with the Lebesgue integral.
  - In such a case, $\mathbb{P}(x|\Omega)$ is known as a **probability mass function**.

Definition of expected value

- If V is a real-valued function defined on the sample space, the **expected value** $\mathbb{E}(V)$ is defined as the Lebesgue integral $\int_\Omega V(x)\,\mathrm{d}\mathbb{P}(x|\Omega)$.
- In cases where a discrete summation $\sum_{x\in\Omega} V(x)\cdot\mathbb{P}(x|\Omega)$ applies, this agrees with the Lebesgue integral.
  - In such a case, $\mathbb{P}(x|\Omega)$ is known as a **probability mass function**.
- Also, in cases where a Riemann integral $\int_\Omega V(x)\cdot\mathrm{pdf}(x)\,\mathrm{d}x$ applies (i.e. where $\Omega\subseteq\mathbb{R}^n$), this agrees with the Lebesgue integral as well.

Definition of expected value

- If V is a real-valued function defined on the sample space, the **expected value** $\mathbb{E}(V)$ is defined as the Lebesgue integral $\int_{\Omega} V(x)\,\mathrm{d}\mathbb{P}(x|\Omega)$.
- In cases where a discrete summation $\sum_{x\in\Omega} V(x)\cdot\mathbb{P}(x|\Omega)$ applies, this agrees with the Lebesgue integral.
  - In such a case, $\mathbb{P}(x|\Omega)$ is known as a **probability mass function**.
- Also, in cases where a Riemann integral $\int_{\Omega} V(x)\cdot\mathrm{pdf}(x)\,\mathrm{d}x$ applies (i.e. where $\Omega\subseteq\mathbb{R}^n$), this agrees with the Lebesgue integral as well.
  - $\mathrm{pdf}(x)=\frac{d\mathbb{P}(\omega\leqq x|\Omega)}{dx}$ is known as a **probability density function**.

Definition of expected value

- If V is a real-valued function defined on the sample space, the **expected value** $\mathbb{E}(V)$ is defined as the Lebesgue integral $\int_{\Omega} V(x)\,\mathrm{d}\mathbb{P}(x|\Omega)$.
- In cases where a discrete summation $\sum_{x\in\Omega} V(x)\cdot\mathbb{P}(x|\Omega)$ applies, this agrees with the Lebesgue integral.
  - In such a case, $\mathbb{P}(x|\Omega)$ is known as a **probability mass function**.
- Also, in cases where a Riemann integral $\int_{\Omega} V(x)\cdot\mathrm{pdf}(x)\,\mathrm{d}x$ applies (i.e. where $\Omega\subseteq\mathbb{R}^n$), this agrees with the Lebesgue integral as well.
  - $\mathrm{pdf}(x)=\frac{d\mathbb{P}(\omega\leqq x|\Omega)}{dx}$ is known as a **probability density function**.
  - $\mathbb{P}(\omega\leqq x|\Omega)$ itself, also a function of x, is known as a **cumulative distribution function** (or simply **distribution function**), and uniquely specifies the probability of all events depending on ω.

Properties of expectation

- The expected value of an indicator variable 1_A (defined to be 1 if A is true and 0 otherwise) is $\mathbb{E}(1_A) = \mathbb{P}(A|\Omega)$.

Properties of expectation

- The expected value of an indicator variable 1_A (defined to be 1 if A is true and 0 otherwise) is $\mathbb{E}(1_A) = \mathbb{P}(A|\Omega)$.
- The most important property of expectation is linearity:

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

$$\mathbb{E}(a \cdot X) = a \cdot \mathbb{E}(X)$$

Properties of expectation

- The expected value of an indicator variable 1_A (defined to be 1 if A is true and $\mathtt{0}$ otherwise) is $\mathbb{E}(1_A) = \mathbb{P}(A|\Omega)$.
- The most important property of expectation is linearity:

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

$$\mathbb{E}(a \cdot X) = a \cdot \mathbb{E}(X)$$

These follow from the linearity of the Lebesgue integral.

Invariance arguments

- If you can assume probability mass functions are invariant under arbitrary permutations of a finite set of events, then the probabilities of those events are **uniform**, i.e. equal.

Invariance arguments

- If you can assume probability mass functions are invariant under arbitrary permutations of a finite set of events, then the probabilities of those events are **uniform**, i.e. equal.
- From countable additivity, they must equal $1/N$ (where N is the number of events that can be permuted in such a way).

Invariance arguments

- If you can assume probability mass functions are invariant under arbitrary permutations of a finite set of events, then the probabilities of those events are **uniform**, i.e. equal.
- From countable additivity, they must equal $1/N$ (where N is the number of events that can be permuted in such a way).
    - Thus we obtain probability $1/2$ for a "fair" (symmetric) coin landing on either side, $1/6$ for a "fair" (symmetric) cubical die landing on any face, etc.

Invariance arguments

- If you can assume probability mass functions are invariant under arbitrary permutations of a finite set of events, then the probabilities of those events are **uniform**, i.e. equal.
- From countable additivity, they must equal $1/N$ (where N is the number of events that can be permuted in such a way).
  - Thus we obtain probability $1/2$ for a "fair" (symmetric) coin landing on either side, $1/6$ for a "fair" (symmetric) cubical die landing on any face, etc.
- Similar arguments can be applied to density functions in continuous cases, usually to justify uniform density over an interval (scaled so that the integral of the density is 1, as required by the laws of probability).

Puzzles

1. Are conditional probabilities probabilities of implications?
    - Either prove that $\mathbb{P}(B|A,\Omega) = \mathbb{P}(A \rightarrow B|\Omega)$, or find a counterexample.

Puzzles

1. Are conditional probabilities probabilities of implications?
   - Either prove that $\mathbb{P}(B|A,\Omega) = \mathbb{P}(A \to B|\Omega)$, or find a counterexample.
2. From each system of "laws of probability", prove all the laws of all the other systems.
   - Note: there's a lot of overlapping laws, which are trivially provable from each other.

Puzzles

1. Are conditional probabilities probabilities of implications?
   - Either prove that $\mathbb{P}(B|A,\Omega) = \mathbb{P}(A \to B|\Omega)$, or find a counterexample.
2. From each system of "laws of probability", prove all the laws of all the other systems.
   - Note: there's a lot of overlapping laws, which are trivially provable from each other.
3. A meter-stick is cut at a uniformly sampled point along its length. What is the expected value of the larger piece?

Puzzles

1. Are conditional probabilities probabilities of implications?
   - Either prove that $\mathbb{P}(B|A,\Omega) = \mathbb{P}(A \rightarrow B|\Omega)$, or find a counterexample.
2. From each system of "laws of probability", prove all the laws of all the other systems.
   - Note: there's a lot of overlapping laws, which are trivially provable from each other.
3. A meter-stick is cut at a uniformly sampled point along its length. What is the expected value of the larger piece?
   - Bonus puzzle: if the stick is instead cut simultaneously at two uniformly sampled points, what is the expected value of the largest piece?