Probability and Philosophy

A.C. Paseau

Trinity 2025

Lecture 1, Wednesday 30 April 2025

Lectures

1. **Introduction to Probability Theory**
2. Dutch Book Arguments
3. Dutch Book Arguments Assessed and Empirical Matters
4. Miscellanea
5. Bayesianism
6. Bayesianism and the Problem of Induction

Location: Lecture Room, Radcliffe Humanities (2nd floor)

Time: 9.30 to 11 am on Wednesdays, weeks 1–4 and 7-8 of Trinity Term. NB no lectures in weeks 5–6.

Papers

- This is a set of six 1.5-hour undergraduate lectures on probability from a philosophical perspective.
- Our main topic beyond the first lecture will be credences or degrees of belief.
- The topic is on the syllabus of the following undergraduate papers:
  - Philosophy of Science (FHS)
  - Philosophy of Science and Social Science (FHS)
  - Philosophical Topics in Logic and Probability (FPE).
- It is also very relevant to any papers on metaphysics and epistemology.
- All other students (BPhil, MSt, DPhil, visiting, etc.) very welcome.

Set-up I

- We normally wish to assign probabilities to possible outcomes of some process. For example, a coin toss or a throw of the dice. More generally: to any way the world might be.
- Call the things to which we wish to assign probability *events*. We can think of events as collections of *outcomes*, which are maximally specific as far as the modelling situation goes.
- Example: when rolling a die, the relevant outcomes might be that the die lands on 1, 2, 3, 4, 5 or 6. (We don't care about e.g. the manifold ways it might land on a 6.) The following outcomes are all events: $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}$ or $\{6\}$. As are the following: $\{1, 3\}$ (throwing a 1 or a 3), or $\{1, 2, 3, 4, 5, 6\}$ (the trivial event), or by courtesy the impossible event $\emptyset$.

Set-up II

- The set of outcomes (recall that sets of these make up events) is known as the *sample space* and is usually denoted Ω. It is assumed to be non-empty.
- Each event is a subset of Ω.
- The set of events is assumed to have a certain set-theoretic structure: it contains Ω (the trivial event); it is closed under complementation (it contains the 'negation' of any event); and if $E_1, E_2, \ldots, E_n, \ldots$ are events so is the event that is their 'disjunction'.
- In other words, the set of events Σ is a σ-algebra, that is:
  - $\Omega \in \Sigma$
  - If $E \in \Sigma$ then $\Omega \setminus E \in \Sigma$
  - If $E_1, E_2, \ldots, E_n, \cdots \in \Sigma$ then $\bigcup_{i=1}^{\infty} E_i \in \Sigma$.
- The ordered pair $\langle \Omega, \Sigma \rangle$ is known as a *measurable* or *Borel* space.

σ-algebras I

- If Ω is non-empty then $\langle \Omega, \mathbb{P}(\Omega) \rangle$, where $\mathbb{P}(\Omega)$ is the power set of Ω (i.e. the set of all its subsets), is a measurable space.
- If Ω is non-empty then $\langle \Omega, \{\Omega, \emptyset\} \rangle$ is a measurable space, i.e. $\{\Omega, \emptyset\}$ is a σ-algebra over Ω.
- Suppose $\Omega = \{1, 2, 3, 4\}$. Let Σ be the set whose four members are:

  $\emptyset, \{1, 2\}, \{3, 4\}, \{1, 2, 3, 4\}$

  Then Σ is a σ-algebra over Ω. Incidentally, this shows that the singleton of an outcome need not be an event, i.e. $\omega \in \Omega$ does not imply that $\{\omega\} \in \Sigma$.

Finite, Countable and Uncountable

- It is important for this course to be clear on what all these terms mean. We define them informally.
- A *finite* set is of size n for n a natural number. So it is either empty or the same size as the set $\{0, 1, \dots n-1\}$.
- A *countably infinite* set is the same size as $\mathbb{N} = \{0, 1, \dots\}$, the set of natural numbers. The set of rational numbers $\mathbb{Q}$ is another example of a countably infinite set.
- A *countable* set is either finite or countably infinite.
- An *uncountable* set is larger than the set of natural numbers. Examples of uncountable sets include $\mathbb{R}$ or $[0, 1]$, the set of real numbers between 0 and 1 inclusive.
- We won't need this fact in this course, but in case you don't know it: two sets are of the same size just when there is a one-to-one and onto function from one to the other. A one-to-one and onto function is also known as a *bijection*.

σ-algebras II

- If Σ is a σ-algebra then it is closed under countable intersection: if $E_1, E_2, \ldots, E_n, \cdots \in \Sigma$ then $\bigcap_{i=1}^{\infty} E_i \in \Sigma$. It is also closed under relative complement: if $E_1, E_2 \in \Sigma$ then $E_1 \setminus E_2 \in \Sigma$.
- Let $\mathbb{R}$ be the set of real numbers. Let $\mathcal{B}$ be the *Borel* subsets of $\mathbb{R}$, i.e. the smallest set of subsets containing all the intervals of the form (a, b), where $a < b$ are both real numbers, and closed under complementation and countable union.
- *Exercises*: show that $\mathcal{B}$ is a σ-algebra over $\mathbb{R}$; and show that $\mathcal{B}$ contains all finite and countably infinite subsets of $\mathbb{R}$.
- The argument at the end of this lecture will imply that there are subsets of the real numbers that are not Borel, i.e. $\mathcal{B} \neq \mathbb{P}(\mathbb{R})$.

Probability/Kolmogorov Axioms

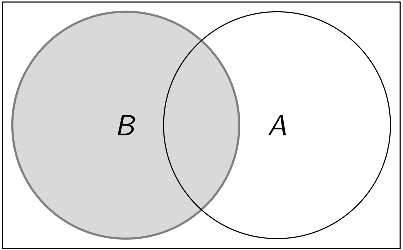
- Suppose $\langle \Omega, \Sigma \rangle$ is a measurable space. As usual, $\mathbb{R}$ is the set of real numbers. A *probability function p* is a map from Σ to $\mathbb{R}$ such that:
  - **Non-Negativity:** for any E, $p(E) \geq 0$.
    The probability of any event is non-negative.
  - **Normalisation:** $p(\Omega) = 1$.
    The probability of the sample space/trivial event is 1.
  - **Countable Additivity:** For any $E_1, E_2, \ldots, E_n, \cdots \in \Sigma$: if $E_i \cap E_j = \emptyset$ for all $i \neq j$ then $p(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} p(E_i)$.
    The probability of the countable union of any countable sequence of pairwise disjoint sets is the sum of their respective probabilities.
- We then say that $\langle \Omega, \Sigma, p \rangle$ is a probability space.

Ratio Formula I

- What is the probability that a fair die will land on a multiple of 6 *given* that it lands on a multiple of 3?
- The two events we need to consider are that of landing on a multiple of 6, i.e. the event $\{6\}$, and that of landing on a multiple of 3, i.e. the event $\{3, 6\}$.
- Using the symbol '$|$', we write this conditional probability as $p(\{6\}|\{3, 6\})$.
- Note that a conditional probability is assigned to a pair of events, not a single event. The second event is supposed to have occurred and the probability of the first event is evaluated on that supposition.
- In this example, the conditional probability is readily seen to be $\frac{1}{2}$.

Ratio Formula II

More generally, we need to consider the conditional probability of A given B, which we write as $p(A|B)$. The picture is that we restrict attention to the B-part of the Venn diagram (which represents the sample space).



We then consider the proportion of this region that corresponds to the event A.

Ratio Formula III

- This yields the following formula:

$$p(A|B) = \frac{p(A \cap B)}{p(B)}, \text{ if } p(B) \neq 0$$

.

- NB Mathematicians often treat this formula as a *definition* of conditional probability. We won't do so here. We take the ratio formula to be not a definition but a constraint on probabilities that may or may not be violated. This will become particularly clear when we think of probabilities as credences.
- The ratio formula can be considered a supplementary axiom to the Kolmogorov ones listed earlier.
- As an aside, note that the ratio formula leaves it open what $p(A|B)$ is when $p(B)$ is 0.

Version with Finite Additivity

- A small number of probability theorists have held that the additivity condition should take a finite rather than countably infinite form. An example is Bruno de Finetti (1906–1985). In many applications of probability theory, especially elementary ones, the finite form is the only one required. We present it here and return to question of Finite vs Countable Additivity in a later lecture.
- Suppose $\mathcal{F}$ is an *algebra* over the sample space Ω, i.e. satisfies the following three conditions:
  - $\Omega \in \mathcal{F}$.
  - If $E \in \mathcal{F}$ then $\Omega \setminus E \in \mathcal{F}$.
  - If $E_1, E_2, \ldots, E_n \in \mathcal{F}$ (where n is finite) then $\bigcup_{i=1}^{n} E_i \in \mathcal{F}$.

  A *probability function* p is a map from $\mathcal{F}$ to $\mathbb{R}$ such that:
  - **Non-Negativity:** $p(E) \geq 0$.
  - **Normalisation:** $p(\Omega) = 1$.
  - **Finite Additivity:** For any $E_1, E_2, \ldots, E_n \in \mathcal{F}$: if $E_i \cap E_j = \emptyset$ for all $i \neq j$ then $p(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} p(E_i)$.

(Mere) Algebra vs Σ-Algebra I

- All σ-algebras are algebras.
- And Finite Additivity follows from Countable Additivity and Normalisation.
- We could simply modify Countable Additivity to take in the finite case too, and rebrand it Countable-or-Finite Additivity.
- Or observe that setting $E_1 = \Omega$ and $E_2, E_3, \ldots, E_n, \cdots = \emptyset$ and using Normalisation and Countable Addivity yields $p(\emptyset) = 0$.
- So now set $E_{n+1} = E_{n+2} = \cdots = \emptyset$ in Countable Additivity to derive Finite Additivity.

(Mere) Algebra vs Σ-Algebra II

- All σ-algebras are algebras, but not all algebras are σ-algebras.
- Suppose Ω is a countably infinite set, e.g. the set of natural numbers.
- Let $\mathcal{F}$ be the set of finite subsets of Ω. Let $\mathcal{CF}$ be the set of cofinite subsets of Ω, i.e. the set of all sets whose complement in Ω is finite.
- Then $\mathcal{F} \cup \mathcal{CF}$ is *not* a σ-algebra over Ω, as it does not satisfy the condition of countably infinite closure.
- However, $\mathcal{F} \cup \mathcal{CF}$ *is* an *algebra* over Ω because it satisfies the condition of finite closure.

Four Interpretations of Probability [1]

- *Count*: relative frequency, i.e. the ratio of some events of a certain kind relative to a broader class of events.
- *Credence*: the strength or degree of someone's belief. This is an epistemic notion, and the one we will focus on in later lectures.
- *Chance*: a mind-independent feature of the natural world, e.g. the half-life of some radioactive material.
- *Confirmation*: the degree to which one proposition confirms or supports another. This is also known as the *logical* interpretation.
- No commitment is made to all of these being primitive/irreducible, nor conversely to there being no other types of probability.

[1] Here I follow Adam Caulton's classification and terminology in his PTLP notes, for in-house consistency.

A Brief History of Probability I

- The mathematical theory of probability is standardly taken to begin in 1654. This is the year Pascal and Fermat began a correspondence analysing some gambling problems. Inspired by this work, Huygens published a treatise on probability in 1657.
- Precursors include Cardano (1501–76), whose c. 1564 handbook for gamblers was only published in 1663, as well as Galileo (1564–1642).
- Jakob Bernoulli's 1713 *Ars Conjectandi* was a landmark. In it, Bernoulli proved the first limit theorem, which today would be seen as a special case of the law of large numbers.
- Laplace's *Philosophical Essay on Probabilities* of 1814 was particularly influential. Hacking (2001, p. 45) calls it 'the first introductory college textbook about probability'.
- The probability axioms in more or less their above form are owed to Kolmogorov in 1933.

A Brief History of Probability II

- It's interesting to speculate why probability arose so late in the day, comparatively speaking. For example, why did the ancient Greeks, who were such skilled mathematicians, not develop probability theory?
- We can extract three reasons from the discussion in Gillies (2000, pp. 22–23). The first is that Greek mathematical expertise lay mainly in geometry. Yet probability theory requires arithmetic and algebra.
- The second is that Greek notation prevented them from doing so. It's true that the early 17th century was a watershed moment in the development of mathematical notation. It's not clear, however, how notation-heavy probability theory is compared to other branches of mathematics.
- The third is that gambling in the ancient world was carried out not with regular dice but with irregular astragali (= small bones in the heels of sheep or deer).

The Classical Theory I

- Laplace's 1814 text established the so-called classical theory of probability as the canonical one. It reigned supreme for more than 100 years.
- Laplace believed that the world was deterministic, so for him probabilities were epistemic. He defined them as follows:

  The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favourable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favourable cases and whose denominator is the number of all the cases possible. (Laplace 1814, pp. 6–7; Truscott & Emory transl.).

The Classical Theory II

J.M Keynes (1883–1946) called the principle the classical theory is based on *The Principle of Indifference*. (Jakob Bernoulli had called it the Principle of Non-Sufficient Reason.)

> *The principle of indifference asserts that if there is no known reason for predicating of our subject one rather than another of several alternatives, then relatively to such knowledge the assertions of each of these alternatives have an equal probability. Thus equal probabilities must be assigned to each of several arguments, if there is an absence of positive ground for assigning unequal ones. (Keynes 1921, p. 45)*

The Classical Theory III

- We have a mixture a water and wine. All we know is that there is at most 3 times more of one than the other.
- By the Principle of Indifference, the ratio of wine to water has a uniform probability density in the interval $[\frac{1}{3}, 3]$.
- Ditto for the ratio of water to wine.
- $Pr(wine : water \leq 2) = \frac{2-\frac{1}{3}}{3-\frac{1}{3}} = \frac{\frac{5}{3}}{\frac{8}{3}} = \frac{5}{8}$
- $Pr(water : wine \geq \frac{1}{2}) = \frac{3-\frac{1}{2}}{3-\frac{1}{3}} = \frac{\frac{5}{2}}{\frac{8}{3}} = \frac{15}{16}$
- Yet these are the same event, just described differently.
- Applying the Principle of Indifference to the event described one way gives a different result to applying to the same event described another way.

The Classical Theory IV

- As we have seen, one of the main problems with the classical theory is that the probabilities depend on how we carve up the problem.
- Another example: you might argue that the probability of a coin landing heads nine times in nine tosses is $\frac{1}{10}$ (since the possible outcomes are 0, 1, 2, 3, 4, 6, 7, 8, or 9). Or if you consider each coin toss individually, you might argue that it is $(\frac{1}{2})^9$.
- Also, how does the classical theory deal with a biased coin or die? It seems incapable of doing so.
- It also seems incapable of handling scenarios in which there are infinitely many outcomes, since we cannot divide by infinity.
- For these sorts of reasons, the classical theory has been abandoned.

A Limitative Result I

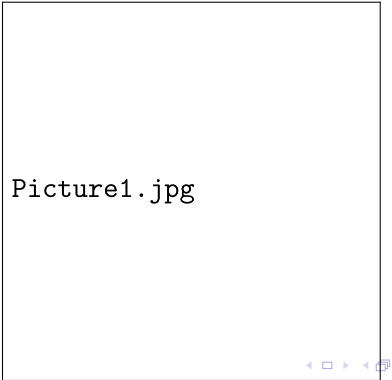
- We now present a limitative result.
- What it shows is that some events that you might think could be assigned a probability cannot be.
- The question is whether this damages the mathematical theory as a good applied model of probabilistic phenomena.

A Limitative Result II

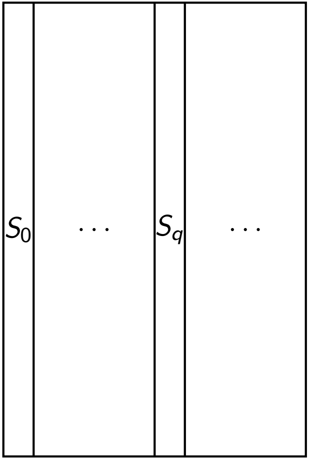
- Consider the measurable space $\langle [0,1], \mathbb{P}([0,1]) \rangle$. Is it possible to define a 'natural' probability function on this space?
- What is a natural probability function?
- If $a, b \in [0,1]$ with $a < b$, then the probability of (a, b) should be equal to $b - a$.
- And the translation of any set of points 'modulo the integers' should have the same probability as the original set of points.
- This second condition implies that e.g. if we think of $[0,1)$ as a circle then the rotation of any set of points should have the same probability as the original set of points.
- Call such a probability function natural.
- The answer to the question is negative. Using the Axiom of Choice (one of the principles of standard set theory), one can prove that there is *no* natural probability function p.

A Limitative Result III

- There are subsets of $[0,1)$, and therefore of $[0,1]$, that cannot be assigned a probability. Proof idea: think of $[0,1)$ as a circle and partition the points on it into equivalence classes.
- Points are in the same equivalence class iff the arc determined by them is a rational fraction of the circumference. Figure showing a few members of an equivalence class from Isaacs *et al.* (2022):



A Limitative Result IV



- The Venn diagram contains all the points in $[0, 1)$. The horizontal strips are the $S_q = \{x + q : x \in S_0\}$, i.e. S_q is S_0 shifted by q, for rational q such that $0 \leq q < 1$.
- Each S_q is uncountable.
- There's a countable infinity of sets S_q, one for each rational q such that $0 \leq q < 1$.

Countable Additivity now implies a contradiction. Either $p(S_q) = 0$ for all rational q or $p(S_q) > 0$ for all rational q. In the first case, the unit circle has probability 0. In the second case, its probability is undefined (infinite).

A Limitative Result V

- History behind this: the Italian mathematician Giuseppe Vitali showed how to construct non-measurable sets, i.e. subsets of $[0, 1]$ (or $\mathbb{R}$, $\mathbb{R}^2$, etc.) that cannot be measured. A measure is like a probability function except that normalisation may be waived and the measure must be natural in the above sense or its analogue in higher dimensions.
- Summary: if $\langle [0, 1], \Sigma, p \rangle$ is a probability space, where p is a natural probability function in the specified sense then the σ-algebra Σ must be a proper subset of $\mathbb{P}[0, 1]$.

A Limitative Result VI

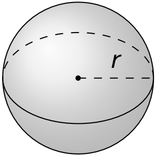
- You might take this to be bad news for the mathematical account of probability described on earlier slides, assuming it is construed as modelling what we would intuitively call probabilistic situations. (As a mathematical theory *per se*, forgetting about applications, it is of course impeccable.)
- The argument might be: we should be able to assign probability to every subset of $[0,1]$ in such a way that all intervals have probability their length and probabilities are not affected by rotations (thinking of the unit interval as the unit circle).
- This could be reinforced by the following thought experiment: suppose you throw a dart repeatedly at the unit interval $[0,1]$. The frequency with which the dart lands on the subinterval (a,b), where $a<b$, tends to $b-a$ and should be translationally symmetric (modulo the integers). Given a subset X of $[0,1]$, surely the frequency with which the dart hits an element of X tends to some number.

A Limitative Result VII

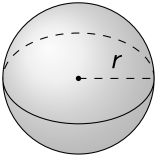
- My suggestion is that fans of probability theory should bite the bullet. There is no probability to speak of here.
- It's not as if this response could be disproved empirically, since we could not build a real-numbered circle/interval in the physical world and measure where a dart lands on it with sufficient precision to establish within reasonable doubt that a limit does exist.
- Compare this limitative result with the Banach-Tarski Paradox.

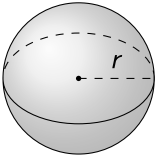
A Limitative Result VIII: The Banach-Tarski Paradox

- Take a ball B with radius r in $\mathbb{R}^3$.



- B can be decomposed into a finite number of disjoint subsets that can be put back together to form two solid balls B_1 and B_2, each of which is of radius r, i.e. is the same size as the original ball B.





A Limitative Result IX

- The usual response to the Banach-Tarski Paradox is that our geometric and physical intuitions are shaped by the physical world. In this world, we have never encountered non-measurable shapes or figures. Our intuitions about them—according to which a single ball cannot be decomposed and reassembled into two identical ones when the pieces involved are non-measurable—are not reliable.
- A similar response can be given to the first limitative result. Our intuitions about the frequency of a dart hitting a target are shaped by experiences in the physical world—a world in which we have never encountered a real interval or non-measurable segments. Consequently, we should not rely on intuitions that suggest a dart will land in a non-measurable subset of the real line with a well-defined frequency, and thus a well-defined probability.

Bibliography for Lecture 1

D. Gillies (2000), *Philosophical Theories of Probability*, Routledge.

I. Hacking (2001), *An Introduction to Probability and Inductive Logic*, Cambridge University Press.

Y. Isaacs, A. Hájek & J. Hawthorne (2022), 'Non-Measurability, Imprecise Credences, and Imprecise Chances', *Mind* 131, pp. 894–918.

J.M. Keynes (1921/1978), *A Treatise on Probability*, Vol. 8 of *The Collected Writings of John Maynard Keynes*, Royal Economic Society.

P-S Laplace (1814/1951), *A Philosophical Essay on Probabilities*, Dover.

Probability and Philosophy

A.C. Paseau

Trinity 2025

Lecture 2, Wednesday 7 May 2025

Lectures

1. Introduction to Probability Theory
2. **Dutch Book Arguments**
3. Dutch Book Arguments Assessed and Empirical Matters
4. Miscellanea
5. Bayesianism
6. Bayesianism and the Problem of Induction

Location: Lecture Room, Radcliffe Humanities (2nd floor)

Time: 9.30 to 11 am on Wednesdays, weeks 1–4 and 7-8 of Trinity Term. NB no lectures in weeks 5–6.

Recap

- Recall the properties of a probability function. It is a map from a σ-algebra over a set of outcomes to the set of real numbers. And it is non-negative; normalised; and countably additive.
- More precisely, $\langle \Omega, \Sigma, p \rangle$ is a probability space just when Σ is a σ-algebra over Ω and p is a map from Σ to $\mathbb{R}$ with the following properties.
  - **Non-Negativity:** for any E in Σ, $p(E) \geq 0$.
  - **Normalisation:** $p(\Omega) = 1$.
  - **Countable Additivity:** For any $E_1, E_2, \ldots, E_n, \cdots \in \Sigma$: if $E_i \cap E_j = \emptyset$ for all $i \neq j$ then $p(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} p(E_i)$.

Credences

- One way to interpret the probability calculus is in terms of *credences* or *degrees of belief* of an agent.
- *Probabilism* is the thesis that (i) an agent's beliefs come in degrees—these are what we call credences; (ii) credences are rationally required to obey the probability axioms.
- How do we measure a person's credences? A question asked in Ramsey (1926). Ramsey observes that
  - We lack what he calls a 'psychogalvanometer', i.e. a measuring device we could attach to or insert in a subject's brain to read off their credences.
  - Strength of feeling is not a good guide to credence (cf. my credence that Lima is the capital of Peru).
  - However, credences do play a role in guiding our actions and help us make good decisions. This gives us a way in.

Expected Utility I

- A rational agent is assumed to have a utility function, which quantifies how valuable they find something. A constraint on such a function U_A is that $U_A(x) > U_A(y)$ iff the agent A prefers x to y. (The agent subscript will from now on usually be implicit.) If the agent's preferences satisfy certain fairly natural assumptions, a utility function can simply be thought of as a way of encoding their preferences.
- Agents are also assumed to have credences defined over all possible combinations of *acts* and *outcome states*. We represent the former as a and the latter as s.

Expected Utility II

- The expected utility of a possible act a is defined as follows:

$$EU(a) = \sum_{s \in S} U(s|a)cr(s|a)$$

  where $U(s|a)$ is the utility of the state s obtaining given that the agent takes action a and $cr(s|a)$ is the agent's credence that s will obtain given that they take action a. S here is the set of all possible states. If the utility of s is independent of which action the agent takes, we may write $U(s)$.
- The Expected Utility Hypothesis says that a subject should perform the act with the greatest expected utility (for them).

Representing a Rational Action

- We use decision tables to represent decision problems. Rows correspond to possible acts and columns to possible states. Entries represent the agent's utilities.

	state 1	state 2
act 1	10	20
act 2	20	10

- In this example, the subject is better off performing act 1 if the word is in state 2, and better off performing act 2 if the world is in state 1.

Credences and Betting Quotients: the Informal Idea I

- Suppose your credence in E is $c(E)$.
- Consider a bet that pays you £100 if E is true and £0 if E is false. We assume throughout that you prefer more money to less.
- You are offered this bet for X pounds.
- It seems that if $X < 100c(E)$, you should accept the bet.
- If $X = 100c(E)$, it seems you may rationally either take the bet or refuse it.
- And if $X > 100c(E)$, it seems you should refuse it.
- Similarly but the other way around if you are the one offering the bet: offer it for X pounds if $X > 100c(E)$; if $X = 100c(E)$, you should be indifferent between offering and not; do not offer it for X pounds if $X < 100c(E)$.

Credences and Betting Quotients: the Informal Idea II

- The idea is that the more confident you are in an event occurring, the more you should be prepared to pay for a bet on it.
- If you're sure it won't occur (or has not occurred), pay nothing for it. If you're sure it has, pay the stake. And when your credence does not take one of these extreme forms, your betting quotient increases linearly with your credence.
- Similarly the other way around, i.e. if you're the one offering the bet.
- Thus it seems that betting quotients = credences. Or if you like, credences in their action-guiding or pragmatic role are betting quotients. (We will come back to this equation in the next lecture.)

Credences are Betting Quotients by Definition?

- Some C20 writers thought of an agent's betting quotients as constitutive of their credences. de Finetti (1937/1964, pp. 101–2) wrote:

  Let us suppose that an individual is obliged to evaluate the rate p at which he would be ready to exchange the possession of an arbitrary sum S (positive or negative) dependent on the occurrence of a given event E, for the possession of the sum pS; we will say by definition that this number p is the measure of the degree of probability attributed by the individual considered to the event E or, more simply, that p is the probability of E (according to the individual considered).

- But you don't have to think that credences are *by definition* equal to betting quotients to think they are equal to them.

Dutch Book Arguments I

- The *bettor* is the agent placing the bet. The *bookie* (short for bookmaker) offers the bet and chooses the stakes.
- In the i^{th} bet, the bettor bets on whether state E_i obtains. The bets are all indexed by i in some index set I.
- The bettor has to choose a number q_i for each bet. This is their *betting quotient* on E_i.
- The bettor *then* sets a stake S_i for the i^{th} bet.
- S_i may be positive or negative or zero. (We thereby combine the two cases of being offered a bet and offering a bet in one.)
- The bettor hands £$q_i S_i$ over to the bookie to place the bet.

Dutch Book Arguments II

- If E_i obtains, the bettor gets the stake £S_i. If E_i does not obtain, she gets nothing.
- So in the first case (E_i obtains), the bettor ends up with £$(S_i - q_i S_i) =$ £$S_i(1 - q_i)$ more than they did prior to the bet.
- And in the second case (E_i does not obtain), the bettor ends up with £$-q_i S_i$ more.
- We may tabulate this as follows (entries are pound values):

	E_i	$\Omega \setminus E_i$
i^{th} bet	$S_i(1 - q_i)$	$-q_i S_i$

- Let's assume for the time being that utilities are equal to monetary payoffs.

Ramsey–de Finetti Theorem: General

- A note on the word 'coherent'. We'll take it to mean that a subject's credences satisfy the axioms of probability.
- Some have given it another meaning. For them, a subject's credences are coherent if their betting quotients are such that the bookie has *no* profit-guaranteeing strategy. Such a strategy would involve placing different bets that the bettor must accept given her betting quotients and yet, no matter what happens, the bettor is guaranteed to lose money to the bookie, i.e. the bookie is guaranteed a profit.
- To avoid any ambiguity, we will say that a bettor is 'Dutch Bookable' (or just 'DB-able') to mean that the bookie has a profit-guaranteeing strategy against the bettor. We could also use the word 'exploitable'.
- The version of the theorem we will prove is the following: If an agent's bettinq quotients violate the axioms of probability then they are Dutch Bookable.

Ramsey – de Finetti Theorem: Non-Negativity & Boundedness

	E_i	$\Omega \setminus E_i$
i^{th} bet	$S_i(1-q_i)$	$-q_i S_i$

Non-Negativity Fails $\Rightarrow$ *DB-able*

If $q_i < 0$ then the bookie can set $S_i < 0$, so that both $S_i(1-q_i), -q_i S_i < 0$.

Not Bounded by 1 $\Rightarrow$ *DB-able*

If $q_i > 1$ then the bookie can set $S_i > 0$, so that both $S_i(1-q_i), -q_i S_i < 0$.

Ramsey – de Finetti Theorem: Normalisation

	Ω	$\emptyset$
i^{th} bet	$S_i(1 - q_i)$	$-q_i S_i$

Not-Normalised $\Rightarrow$ *DB-able*

We have already shown that if the subject is not Dutch Bookable then $0 \leq q_i \leq 1$.

If $q_i < 1$ then the bookie can set $S_i < 0$, so that $S_i(1 - q_i) < 0$. The left column is the only outcome.

Ramsey – de Finetti Theorem: Additivity I

As stated earlier, (Countable) Additivity applied to a collection of mutually exclusive events. It is easier to prove a coherence result for the special case of the axiom that applies to a collection of mutually exclusive and jointly exhaustive events, namely:

Additivity*. For any $E_1, E_2, \ldots, E_n, \cdots \in \Sigma$: if $E_i \cap E_j = \emptyset$ for all $i \neq j$ and $\bigcup_{i \in \mathbb{N}} E_i = \Omega$ then $\sum_{i \in \mathbb{N}} p(E_i) = 1$.

Proof of Equivalence. One direction (Additivity $\Rightarrow$ Additivity*) is trivial in light of Normalisation. For the other, given mutually exclusive E_i, apply Additivity* to events E_i in the index set plus the event $\Omega \setminus (\bigcup_{i \in \mathbb{N}} E_i)$. This gives:

$$\sum_{i \in \mathbb{N}} p(E_i) + p(\Omega \setminus \bigcup_{i \in \mathbb{N}} E_i) = 1$$

Rearranging and using the fact that $1 - p(\Omega \setminus \bigcup_{i \in \mathbb{N}} E_i) = p(\bigcup_{i \in \mathbb{N}} E_i)$ (which follows from Additivity*) gives Additivity.

Ramsey – de Finetti Theorem: Additivity II

Suppose the E_i are mutually exclusive and jointly exhaustive. As above, the i^{th} bet takes the form:

	E_i	$\Omega \setminus E_i$
i^{th} bet	$S_i(1-q_i)$	$-q_iS_i$

*Not-Additivity** $\Rightarrow$ *DB-able*

Suppose the outcome is E_j. Then the bettor gains

$$S_j(1-q_j) - \sum_{i \neq j} S_i q_i$$

If $\sum_i q_i \neq 1$, the bookie sets all stakes $= S$. The bettor's gain is

$$S(1 - q_j - \sum_{i \neq j} q_i) = S(1 - \sum_i q_i)$$

If $\sum_i q_i < 1$, the bookie sets $S < 0$, so that $S(1 - \sum_i q_i) < 0$.

If $\sum_i q_i > 1$, the bookie sets $S > 0$, so that $S(1 - \sum_i q_i) < 0$.

What Kind of Additivity?

- An interesting feature of the argument just given is that we haven't insisted on Countable Additivity.
- We want all the sums to be finite in order to manipulate expressions such as $\sum_i S_i q_i$. But this can hold when the index set I, from which the i are drawn, is countably infinite. Or even uncountable, if only countably many of these terms is non-zero.
- More generally, we may be as liberal as we like about the size of I as long as we impose the condition that $\sum_i |S_i q_i|$ and $\sum_i q_i$ are finite.

Converse? I

- We have shown that if you're not Dutch Bookable then your betting quotients must satisfy Non-Negativity, Normalisation and Additivity.
- A natural question now is whether the converse holds, i.e. whether the fact that your betting quotients satisfy the axioms of probability implies that you are not Dutch Bookable.
- Whether this is true depends on what a Dutch book is. We'll show shortly that if we allow Dutch books that exploit the Ratio Formula, the claim is in fact false.
- First, though, we'll show that in the scenario mentioned on the previous slide, if the subject's betting quotients sum to 1 then they are indeed *not* Dutch-bookable.

Converse? II

Suppose a bookie tries to extract a guaranteed profit from a bettor whose betting quotients satisfy the probability axioms by offering them a series of bets based on the mutually exclusive and jointly exhaustive events E_i. As above, the i^{th} bet takes the form:

	E_i	$\Omega \setminus E_i$
i^{th} bet	$S_i(1-q_i)$	$-q_i S_i$

*Additivity** (and *Non-negativity*) $\Rightarrow$ *Not DB-able*

Proceed by reductio. If the outcome is E_j, the bettor gains

$$S_j(1-q_j) - \sum_{i \neq w} S_i q_i$$

Thus, by the assumption that this is a Dutch Book, for all j:

$$S_j(1-q_j) - \sum_{i \neq j} S_i q_i < 0$$

Converse? III

Rearranging shows that, for all j:

$$S_j < \sum_i S_i q_i$$

Hence, since each q_j is non-negative,

$$q_j S_j \leq q_j (\sum_i S_i q_i),$$

and at least one of these inequalities must be strict since at least one of the q_j is positive. Summing all these inequalities (at least one of which is strict) for all the j in the index set yields

$$\sum_j q_j S_j < (\sum_j q_j)(\sum_i S_i q_i)$$

But since $\sum_j q_j = 1$, by Additivity*, this yields: $\sum_j q_j S_j < \sum_i S_i q_i$.
Contradiction! (The two sides of the inequality are the same value.)

Converse? IV

- We will introduce conditional bets shortly. What about unconditional bets on events that are not mutually exclusive?
- Let's consider the finite case first. Suppose a bookie tries to extract a guaranteed profit from a bettor by offering them a series of bets based on events E_i for $1 \leq i \leq n$, with respective stakes S_i. These events may or may not be mutually exclusive and may or may not be jointly exhaustive.
- This is equivalent to a series of bets on the 2^n events $F_1 \cap \cdots \cap F_i \cdots \cap F_n$ where each F_i is either E_i or $\Omega \setminus E_i$. And these events are mutually exclusive and jointly exhaustive, so we can apply the previous reasoning.
- Suppose for example that the bookie pays the bettor S_1 if E_1 and S_2 if E_2 ($n = 2$). This is equivalent to the following bet:
  - payoff of $S_1 + S_2$ if $E_1 \cap E_2$
  - payoff of S_1 if $E_1 \cap \Omega \setminus E_2$
  - payoff of S_2 if $\Omega \setminus E_1 \cap E_2$
  - payoff of 0 if $\Omega \setminus E_1 \cap \Omega \setminus E_2$

Converse? V

- What if the bets are placed on infinitely many events, say a countable infinity of events? Can we reason in the same way?
- The problem is that if there is a countable infinity of events E_i then this leads to an uncountable infinity of events that is the intersection of each of the events or its complement, i.e. there are uncountably many

  $$\bigcap_{i=1}^{\infty} F_i$$

  where each F_i is E_i or $\Omega \setminus E_i$. (The formal result is that '2 to an infinite power' is uncountable; this can be stated precisely and proved.)
- So we would need more than Countable Additivity to apply exactly the same reasoning, i.e. to show that the bettor is not Dutch-Bookable. As we'll see in a later lecture, however, Uncountable Additivity is usually rejected.
- And of course if there are infinitely many stakes S_i, we have to ensure that all the sums are well-defined.

Ratio Formula Dutch Book I

- Recall that in its credal form, the Ratio Formula is

$$cr(A|B) = \frac{cr(A \cap B)}{cr(B)},$$

  assuming $cr(B) \neq 0$. A Dutch Book argument can also be given for this principle. A *conditional* bet on A given B is a bet that is called off if B does not occur, pays the bettor $S_{A|B}$ if A and B occur and nothing if B occurs but A doesn't. The bettor's betting quotient is $q_{A|B}$. The betting quotients $q_{A \cap B}$ and q_B and stakes $S_{A \cap B}$ and S_B have their obvious meanings: they are the respective betting quotients and stakes for respective bets on $A \cap B$ and B.
- Consider a bettor who places the three bets above with a bookie: the conditional bet on A given B; the bet on $A \cap B$; and the bet on B.

Ratio Formula Dutch Book II

- There are three possible cases.
  - A and B both occur. The bettor's payoff is $S_{A|B}(1 - q_{A|B}) + S_{A\cap B}(1 - q_{A\cap B}) + S_B(1 - q_B)$
  - A does not occur but B does. The bettor's payoff is $-S_{A|B}q_{A|B} - S_{A\cap B}q_{A\cap B} + S_B(1 - q_B)$
  - B does not occur. The bettor's payoff is $-S_{A\cap B}q_{A\cap B} - S_B q_B$
- Set $S_B = q_{A|B}S_{A|B}$ and $S_{A|B} = -S_{A\cap B}$. For brevity, let $S_{A\cap B}$ be S. Then in each case the three (identical!) payoffs are:
  - $-S(1-q_{A|B})+S(1-q_{A\cap B})-Sq_{A|B}(1-q_B) = S(q_{A|B}q_B-q_{A\cap B})$
  - $Sq_{A|B} - Sq_{A\cap B} - Sq_{A|B}(1 - q_B) = S(q_{A|B}q_B - q_{A\cap B})$
  - $-Sq_{A\cap B} + Sq_{A|B}q_B = S(q_{A|B}q_B - q_{A\cap B})$
- If $q_{A|B}q_B < q_{A\cap B}$, the bookie can choose $S > 0$ to guarantee themself a profit.
- If $q_{A|B}q_B > q_{A\cap B}$, the bookie can choose $S < 0$ to guarantee themself a profit.
- So if the bettor is not Dutch bookable, $q_{A|B}q_B = q_{A\cap B}$.

Ratio Formula Converse Dutch Book Argument I

We now give a converse Dutch Book argument involving the ratio formula. Namely: if $q_{A\cap B} = q_{A|B}q_B$ then the bettor is not Dutch Bookable with this kind of bet. (Assuming their betting quotients also satisfy the other axioms—this will be implicit.)

Let's rewrite the three payoffs in the following way, availing ourselves of matrix multiplication.

$$\begin{pmatrix} \text{1st payoff} \\ \text{2nd payoff} \\ \text{3rd payoff} \end{pmatrix} = \begin{pmatrix} 1 - q_{A|B} & 1 - q_{A\cap B} & 1 - q_B \\ -q_{A|B} & -q_{A\cap B} & 1 - q_B \\ 0 & -q_{A\cap B} & -q_B \end{pmatrix} \begin{pmatrix} S_{A|B} \\ S_{A\cap B} \\ S_B \end{pmatrix}$$

And let's now write $q_{A|B}$ as λ and q_B as μ. The equation $q_{A\cap B} = q_{A|B}q_B$ may be rewritten as $q_{A\cap B} = \lambda\mu$. Note that at least one of $\lambda\mu, (1-\lambda)\mu$ and $1-\mu$ is non-zero. Also, if $\mu = 0$ there is no conditional bet; and since $1 \geq \mu \geq 0$, we may assume $1 \geq \mu > 0$. Finally, since $\lambda = \frac{q_{A\cap B}}{q_B}$ and the agent's unconditional betting quotients satisfy the probability axioms, it follows that $1 \geq \lambda \geq 0$. Substituting in, this gives:

Ratio Formula Converse Dutch Book Argument II

$$\begin{pmatrix} \text{1st payoff} \\ \text{2nd payoff} \\ \text{3rd payoff} \end{pmatrix} = \begin{pmatrix} 1-\lambda & 1-\lambda\mu & 1-\mu \\ -\lambda & -\lambda\mu & 1-\mu \\ 0 & -\lambda\mu & -\mu \end{pmatrix} \begin{pmatrix} S_{A|B} \\ S_{A\cap B} \\ S_B \end{pmatrix}$$

Let's now multiply the first row by $\lambda\mu$, the second by $(1-\lambda)\mu$ and the third by $1-\mu$. The right-hand side becomes

$$\begin{pmatrix} \lambda\mu(1-\lambda) & \lambda\mu(1-\lambda\mu) & \lambda\mu(1-\mu) \\ -\lambda(1-\lambda)\mu & -\lambda(1-\lambda)\mu^2 & (1-\lambda)\mu(1-\mu) \\ 0 & -\lambda\mu(1-\mu) & -\mu(1-\mu) \end{pmatrix} \begin{pmatrix} S_{A|B} \\ S_{A\cap B} \\ S_B \end{pmatrix}$$

A little algebra shows that the entries in each column sum to 0. For a Dutch Book, all three payoffs must be negative, so multiplying them by $\lambda\mu$, $(1-\lambda)\mu$ and $1-\mu$, all of which are non-negative and at least one of which is positive, and summing the entries in the resulting vector must yield a negative real number. But summing the entries in the column vector resulting from multiplying the matrix by the column vector of stakes on the right-hand side yields 0. Contradiction.

Ratio Formula Converse Dutch Book Argument II

To recap, the left-hand-side vector is:

$$\begin{pmatrix} \lambda\mu \times \text{1st payoff} \\ (1-\lambda)\mu \times \text{2nd payoff} \\ 1-\mu \times \text{3rd payoff} \end{pmatrix}$$

Whereas the right-hand-side one is:

$$\begin{pmatrix} \lambda\mu(1-\lambda) & \lambda\mu(1-\lambda\mu) & \lambda\mu(1-\mu) \\ -\lambda(1-\lambda)\mu & -\lambda(1-\lambda)\mu^2 & (1-\lambda)\mu(1-\mu) \\ 0 & -\lambda\mu(1-\mu) & -\mu(1-\mu) \end{pmatrix} \begin{pmatrix} \mathsf{S}_{A|B} \\ \mathsf{S}_{A\cap B} \\ \mathsf{S}_{B} \end{pmatrix}$$

The sum of the LHS vector's entries is negative (if the agent has been Dutch booked), whereas the sum of the RHS vector's entries is 0, since each column sums to 0. Contradiction. It follows that if $q_{A|B}q_B = q_{A\cap B}$ then the agent is not susceptible to this sort of Dutch Book.

Summary

- If the agent's betting quotients fail to satisfy Non-Negativity, or Normalisation, or Countable Additivity, she is Dutch-Bookable.
- If the agent's betting quotients satisfy Finite Additivity (and the other two Kolmogorov axioms), she is not Dutch-Bookable (with this kind of bet).
- If the agent's betting quotients don't satisfy the Ratio Formula, she is Dutch-Bookable.
- If the agent's betting quotients satisfy the Ratio Formula (and the Kolmogorov axioms), she is not Dutch-Bookable (with this kind of bet).

Bibliography for Lecture 2

B. de Finetti (1937), 'La prévision: ses lois logiques, ses sources subjectives', transl. as 'Foresight: Its Logical Laws, its Subjective Sources', in *Studies in Subjective Probability*, H. Kyburg Jr & H. Smokler (eds), Wiley, pp. 94–158.

A. Hájek (2008), 'Arguments for—or against—Probabilism', in F. Huber & C. Schmidt-Petri (eds), *Degrees of Belief*, Synthese Library, pp. 229–251.

F. Ramsey (1926), 'Truth and Probability', in his *The Foundations of Mathematics and other Logical Essays*, R.B. Braithwaite (ed.), Routledge and Kegan Paul, pp. 156–198.

Probability and Philosophy

A.C. Paseau

Trinity 2025

Lecture 3, Wednesday 14 May 2025

Lectures

1. Introduction to Probability Theory
2. Dutch Book Arguments
3. **Dutch Book Arguments Assessed and Empirical Matters**
4. Miscellanea
5. Bayesianism
6. Bayesianism and the Problem of Induction

Location: Lecture Room, Radcliffe Humanities (2nd floor)

Time: 9.30 to 11 am on Wednesdays, weeks 1–4 and 7-8 of Trinity Term. NB no lectures in weeks 5–6.

Recap

- Last time we looked at Dutch Book Arguments.
- Today, we're going to look at some criticisms levelled at Dutch Book arguments as arguments for probabilism. (There are more than we have time to go over.)
- The first kind target the link between credences and dispositions to bet.
- The second kind target the assumption of a linear relationship between money and utility. Dutch Book arguments seem to rely on this assumption.

Credences and Dispositions to Bet I

- The idea that behavioural dispositions are constitutive of mental states was very much in the air in the mid-twentieth century (e.g. Wittgenstein, Ryle, B.F. Skinner in psychology).
- But it is widely rejected, even derided, these days.
- Super-Spartans suffer pain uncomplainingly. They still feel pain even though they lack the associated behavioural dispositions. (This example is owed to Hilary Putnam.)
- On the prevailing contemporary picture, the link between behaviour and mental states is weaker, defeasible and holistic.

Credences and Dispositions to Bet II

- For the case of betting quotients (which captures an agent's dispositions to accept or reject certain bets) and credences, consider agents whose dispositions to place certain bets don't seem to be aligned with their corresponding credences.
- One example are agents who will never bet on anything. Perhaps they are very puritanical and have taken a religious vow never to gamble. So ingrained is this self-ordinance that they no longer even have the inclination to bet on anything (they don't even have a disposition that is trumped in some way). And yet they might well have credences, e.g. about whether it will be sunny tomorrow.
- At the other extreme, imagine agents who get a kick out of betting. Going to the casino every night is their favourite pastime. They're a taker for pretty much any gamble you can offer them. Even hypothetical bets are so enticing that they'll accept them for odds that don't reflect their credences.

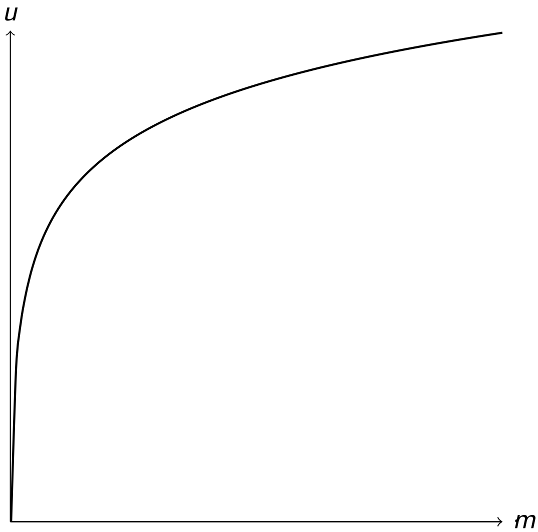
Credences and Dispositions to Bet III

- We may also imagine agents who, although disposed to accept any individual one of a collection of bets, would not be disposed to accept all of them.
- Following Hájek (2008, p. 236), we might call the *Package Principle* the requirement 'to value a set of bets at the sum of the values of the bets taken individually, or less specifically, to regard a set of bets as fair if one regards each bet individually as fair'.
- It's not clear why this should be true in general.
- It's even less clear in cases when there are interferences between the bets. Hájek's example: you're very confident that your partner is happy. You're fairly confident the Democrats will win the next election. But your partner hates you betting on anything and finds out as soon as you do. So packaging the bets together affects the odds you're willing to accept.

Money and Utility

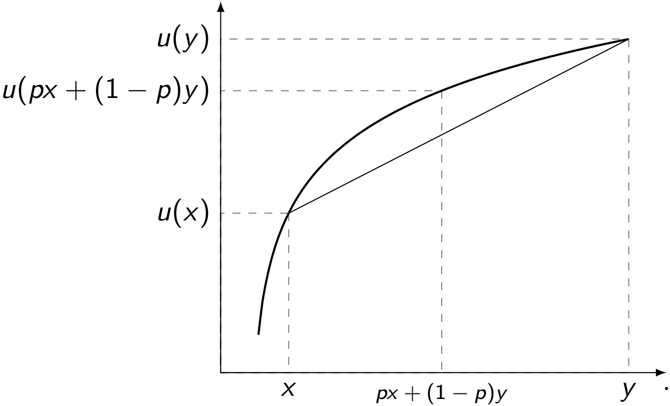
- Next, let's think about the relationship between money and utility.
- The payoffs in the last lecture were in terms of money. But an obvious objection to the Dutch Book arguments we gave there is that utilities do not match monetary payoffs that closely.
- For example, $£100,000$ is surely worth a lot more to a pauper than to a billionaire. Being gifted $£100,000$ would transform the former's life, whereas it would make very little difference to the latter. (Indeed, some billionaires earn that amount in a few seconds.)
- Also, intuitively it seems rational to prefer being given $£100,000$ for sure to a bet with a payoff of $£200,000$ with probability $\frac{1}{2}$ and 0 with probability $\frac{1}{2}$. This is known as *risk aversion*.
- Economists and others model the relationship between money (on the x-axis) and utility (on the y-axis) by means of a concave function, for example (some restriction of) a logarithmic function.

Concave Utility Function



Concave Utility Function II

The following diagram illustrates risk aversion.



.

Utilities I

- The moral of the story is that, when giving Dutch Book arguments and similar, we need to state payoffs in terms of the subject's utilities rather than money.
- We knew this all along: in the previous lecture, we first introduced Expected Utility Theory and then made the simplifying assumption that utilities can be equated with monetary payoffs.
- Unfortunately, utilities cannot be observed directly but have to be inferred. This seems to scupper the idea that credences can be measured by looking at a subject's willingness to accept various bets.
- In fact, we can determine an agent's utilities *and* credences simultaneously, under certain assumptions, by considering which acts they prefer. Frank Ramsey in the paper cited in the previous lecture was the first to sketch how that can be done.

Utilities II

- Also, if we assume that utility functions can be approximated locally by linear functions—as would be the case for example if utility functions are logarithmic—then the equation of utilities with monetary payoffs is acceptable, so long as the stakes are small.
- (A potentially relevant result here is Lebesgue's Theorem, which implies that a continuous and increasing function on an interval is differentiable almost everywhere.)
- Forgetting about the first set of problems (regarding dispositions to bet), perhaps we could define credences as the limit of a sequence of betting quotients for bets with payoffs that tend to zero.

Violations of Expected Utility Theory

Moving on from Dutch Book Arguments, we're now going to look at three apparent violations of Expected Utility Theory. These are:

- The St Petersburg Paradox
- The Allais Paradox
- The Reflection Principle

We'll then end with two outright fallacies: the Conjunction Fallacy and the Base Rate Fallacy. In the next lecture, we'll look at an attempt to draw some philosophical morals from the latter.

The St Petersburg Paradox I

- You are invited to play a game involving a fair coin. The coin is tossed until it lands on heads. If it lands on heads for the first time after n tosses, you win 2^n.
- Q1: How much *should* you pay to play this game?
- Answer to Q1: The probability of the coin landing heads for the first time after n tosses is $\frac{1}{2^n}$, since the sequence has to be $n-1$ tails followed by heads. Your expected gain is

$$\sum_{n \in \{1,2,\dots\}} \frac{1}{2^n} 2^n = 1 + 1 + \dots + 1 + \dots$$

- So the answer to Q1, according to Expected Utility Theory: any finite amount is worth paying since the payoff is infinite.
- Q2: How much *would* you pay to play this game according to Expected Utility Theory?
- Answer to Q2: pretty much no one would be prepared to pay any amount to play this game.

The St Petersburg Paradox II

- Unlike the fallacies discussed below, the verdict in the St Petersburg Paradox seems robust. Learning some probability theory does *not* make people think that the intuitive initial response was clearly mistaken, in the way that it would for a fallacy.
- One might appeal to the diminishing marginal utility of money, i.e. the fact that the utility function is concave. But this in itself won't block the argument, since we could set the payoffs to be 2^n in utils. For example, if $u(\pounds 2^n) = n$ then consider payoffs of $\pounds 2^{2^n}$ (with utility 2^n) if the coin lands on heads for the first time after n tosses.
- A better response is that utilities are bounded. There is a certain amount of money beyond which receiving even more might not be desirable. It might even be undesirable: think of storage issues; or the fact that being the richest person on the planet might bring you unwanted attention.

The Allais Paradox I

- Here is an example of an Allais paradox. The exposition is based on Kahneman & Tversky (1979).
- Subjects were first given a choice between:

  4000 with probability 0.8 vs 3000 (for sure)

  and then a choice between

  4000 with probability 0.2 vs 3000 with probability 0.25

  Which would you choose in each case?

The Allais Paradox I

- Here is an example of an Allais paradox. The exposition is based on Kahneman & Tversky (1979).
- Subjects were first given a choice between:

  4000 with probability 0.8 vs 3000 (for sure)

  and then

  4000 with probability 0.2 vs 3000 with probability 0.25

  Which would you choose in each case?

  –20% of subjects chose 4000 with probability 0.8 vs 80% who chose 3000 for sure.
  –65% of subjects chose 4000 with probability 0.2 vs 35% who chose 3000 with probability 0.25

The Allais Paradox II

Expected Utility Theory is inconsistent with the majority choice, since the following pair of inequalities is inconsistent:

$$0.8u(4000) < u(3000)$$

and

$$0.2u(4000) > 0.25u(3000)$$

as is readily seen by multiplying the bottom inequality by 4.

The Reflection Effect I

The following table is from Kahneman & Tversky (1979, p. 22).

Positive Prospects				**Negative Prospects**			
Problem 3:	(4,000, 0.80)	$<$	(3,000)	Problem 3':	(-4,000, 0.80)	$>$	(-3,000)
N = 95	[20]		[80]*	N = 95	[92]*		[8]
Problem 4:	(4,000, .20)	$>$	(3,000, .25)	Problem 4':	(-4,000, .20)	$<$	(-3,000, 0.25)
N = 95	[65]*		[35]	N = 95	[42]		[58]
Problem 7:	(3,000, .90)	$>$	(6,000, .45)	Problem 7':	(-3,000, 0.0)	$<$	(-6,000, .45)
N = 66	[86]*		[14]	N = 66	[8]		[92]*
Problem 8:	(3,000, .002)	$<$	(6,000, .001)	Problem 8':	(-3,000, 0.002)	$>$	(-6,000, .001)
N = 66	[27]		[73]*	N = 66	[70]*		[30]

The Reflection Effect II

- As Kahneman & Tversky comment (1979, p. 22): ‘In each of the four problems … the preference between negative prospects is the mirror image of the preference between positive prospects. Thus, the reflection of prospects around 0 reverses the preference order. We label this pattern *the reflection effect*.’
- They point out that risk aversion in the positive cases is accompanied by risk seeking in the negative ones.
- As in the case of the Allais Paradox, Problems 3 and 4 are inconsistent with Expected Utility Theory (multiplying factor $= 4$), as are Problems 7 and 8 (multiplying factor $= 45$). And they add: ‘The same psychological principle — the overweighting of certainty — favors risk aversion in the domain of gains and risk seeking in the domain of losses’ (Kahneman & Tversky 1979, p. 23).

The Reflection Effect III

- You might argue that what Problems 3 and 4 show is that people prefer prospects with high expected value and small variance. But Problems 3' and 4' go against that thought. Kahneman & Tversky (1979, p. 23) comment: ‘our data are incompatible with the notion that certainty is generally desirable. Rather, it appears that certainty increases the aversiveness of losses as well as the desirability of gains.’
- Their *prospect theory* (which we cannot present here) replaces probabilities with decision weights and assigns to each outcome a subjective value relative to a reference point (rather than final asset positions).

The Conjunction Fallacy and the Base Rate Fallacy

- H.G. Wells once famously wrote: 'A certain elementary training in statistical method is becoming as necessary for everyone living in this world of today as reading and writing' (cited in Gigenrenzer 2008, p. 127).
- Alas, almost all of us handle probabilities (and statistics) poorly.

The Conjunction Fallacy I

- The Conjunction Fallacy is Danny Kahneman and Amos Tversky's name for people's tendency to judge a conjunction of two events to be more probable than just one of the two. (See e.g. ch. 15 of the former's 2012 book.)
- In one of their most famous experiments, the two of them described a made-up woman by the name of Linda as follows:

  Linda is thirty-one years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations. (Kahneman 2012, p. 179)

- Experimental subjects were then asked which of a range of statements is more likely to be true. They included the statement that (i) Linda is a bank teller, and (ii) Linda is a bank teller and feminist.

The Conjunction Fallacy II

- People, including statistically highly educated people (e.g. doctoral students in decision science at top institutions), overwhelmingly thought (ii) more likely than (i).
- But it is easy to show from the probability axioms that $p(A \wedge B) \leq p(A)$. (To use conjunction—in our event setting we would use intersection.)
- Kahneman's explanation (2012, pp. 182–3):

  The judgments of probability that our respondents offered ... corresponded precisely to judgments of representativeness (similarity to stereotypes). Representativeness belongs to a cluster of closely related basic assessments that are likely to be generated together. The most representative outcomes combine with the personality description to produce the most coherent stories. The most coherent stories are not necessarily the most probable, but they are plausible, and the notions of coherence, plausibility, and probability are easily confused by the unwary.

The Conjunction Fallacy III

- Hacking (2001, p. 66) wonders whether people attend closely to the exact wording of the question. Maybe (to paraphrase him a little) they hear the question ‘Which is more probable?’ as 'Which is the most useful and instructive thing to say about Linda?’.
- Fielder (1988) claims that the percentage of people who commit the fallacy drops dramatically if the word ‘probability’ in the question is replaced by the word ‘frequency’ (and cognates).
- To explain the next fallacy, we must set out Bayes’s Rule.

Bayes's Rule

The following result, known as *Bayes's Rule* or *Bayes's Theorem*, is an immediate corollary of the probability axioms and the ratio formula $p(A|B) = \frac{p(A \cap B)}{p(B)}$, if $p(B) \neq 0$:

$$p(H|E) = \frac{p(E|H) \times p(H)}{p(E)}$$

The letter E is suggestive of 'evidence' and H of 'hypothesis'. The result is owed to the reverend Thomas Bayes (and/or Richard Price who edited and published his notes following Bayes's death) in the mid-18th century.

$p(E)$ and $p(H)$ are often known as *prior* probabilities, i.e. prior to anything being supposed or—in the temporal version—prior to learning evidence E. Similarly, $p(H|E)$ is known as the *posterior*. And (for historical reasons), $p(E|H)$ is known as the *likelihood*.

Law of Total Probability

A *partition* is a set of mutually exclusive and jointly exhaustive events. In other words, in a probabilistic context, a collection $\{E_i : i \in I\}$, each of which is an element of the sigma-algebra, such that $E_i \cap E_j = \emptyset$ if $i \neq j$, and $\bigcup_{i \in I} E_i = \Omega$.

Letting the index set I be finite, the Law of Total Probability may be stated as follows, where $E_1, \ldots, E_n$ is a partition (and we assume that $p(E_i) > 0$ for each i):

$$p(F) = \sum_{i=1}^{n} p(F|E_i)p(E_i)$$

The proof easily follows from the Ratio Formula, the fact that the event F is the union of the n mutually exclusive events $F \cap E_1, \ldots, F \cap E_n$, and the (Finite) Additivity axiom.

Bayes's Rule – Law of Total Probability Version

When there are lots of mutually exclusive and jointly exhaustive hypotheses $H_1, \ldots, H_n$ in play, we may avail ourselves of the Law of Total Probability and write Bayes's Rule as follows:

$$cr(H_1|E) = \frac{p(E|H_1) \times p(H_1)}{\sum_{i=1}^{n} p(E|H_i) \times p(H_i)}$$

Note that, H_n may be the catch-all hypothesis, i.e. we specify $H_1, \ldots, H_{n-1}$ and define H_n as $\neg(H_1 \vee \cdots \vee H_{n-1})$.

Reichenbach (1935/1949, pp. 94–5) on Bayes's Rule

The range of application for Bayes's rule is extremely wide, because nearly all inquiries into the causes of observed facts are performed in terms of this rule. The *method of indirect evidence*, as this form of inquiry is called, consists of inferences that on closer analysis can be shown to follow the structure of the rule of Bayes. The physician's inferences, leading from the observed symptoms to the diagnosis of a specified disease, are of this type; so are the inferences of the historian determining the historical events that must be assumed for the explanation of recorded observations; and, likewise, the inferences of the detective concluding criminal actions from inconspicuous observable data. In many instances the use of probability relations is not manifest because the probabilities occurring have either very high or very low values. Thus, when a corpse is found, it is virtually certain that a murder has been committed; and a fingerprint on the handle of a pistol may be considered as strict evidence for the assumption that a certain person X has fired the pistol. That even in such cases the inference has the structure of Bayes's rule is often seen from the fact that appraisals of the antecedent probabilities are made. Thus an inquiry by the detective into the motives of a crime is an attempt to estimate the antecedent probabilities of the case, namely, the probability of a certain person committing a crime of this kind, irrespective of the observed incriminating data. Similarly, the general inductive inference from observational data to the validity of a given scientific theory must be regarded as an inference in terms of Bayes's rule.

Base Rate Fallacy I

Gerd Gigerenzer reports the following:

> *...physicians with an average of 14 years professional experience were asked to imagine using the Haemoccult test to screen for colorectal cancer (Hoffrage & Gigerenzer, 1998). The prevalence of cancer was 0.3 percent, the sensitivity of the test was 50 percent, and the false positive rate was 3 percent. The doctors were asked: What is the probability that a person who tests positive actually has colorectal cancer? The correct answer is about 5 percent. However, the physicians' answers ranged from 1 percent to 99 percent, with about half estimating the probability as 50 percent (the sensitivity) and 47 percent (sensitivity minus false positive rate). If patients knew about this degree of variability and statistical innumeracy they would be justly alarmed. (2008, pp. 127–8)*

Base Rate Fallacy II

- Let's use Bayes' rule to work out the correct answer.

$$
\begin{aligned}
p(ca|+ve) &= \frac{p(+ve|ca) \times p(ca)}{p(+ve|ca) \times p(ca) + p(+ve|\neg ca) \times p(\neg ca)} \\
&= \frac{0.5 \times 0.003}{0.5 \times 0.003 + 0.03 \times 0.997} \\
&= \frac{0.0015}{0.0015 + 0.02991} \\
&\approx 4.8\%
\end{aligned}
$$

- Conditional probabilities refer to different classes—those with and without cancer in the above example—which, as Gigenrezer (2008, p. 132) observes, people have difficulty combining in their minds.

Bibliography for Lecture 3

K. Fielder (1988), 'The dependence of the conjunction fallacy on subtle linguistic factors', *Psychological Research* 50, pp. 123–129.

G. Gigerenzer (2008), *Rationality for Mortals*, Oxford UP.

I. Hacking (2001), *An Introduction to Probability and Inductive Logic*, Cambridge UP.

D. Kahneman (2012), *Thinking Fast and Slow*, Penguin.

D. Kahneman & A. Tversky (1979), 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica* 47, pp. 263–291, repr. in *Choices, Values, and Frames*, D. Kahneman & A. Tversky (eds), 2000, Cambridge UP, pp. 17–43.

H. Reichenbach (1935/1949), *The Theory of Probability*, expanded and revised version of *Wahrscheinlichkeitslehre*, transl. by E. Hutten & M. Reichenbach, U. of California Press.

Probability and Philosophy

A.C. Paseau

Trinity 2025

Lecture 4, Wednesday 21 May 2025

Lectures

1. Introduction to Probability Theory
2. Dutch Book Arguments
3. Dutch Book Arguments Assessed and Empirical Matters
4. **Miscellanea**
5. Bayesianism
6. Bayesianism and the Problem of Induction

Location: Lecture Room, Radcliffe Humanities (2nd floor)

Time: 9.30 to 11 am on Wednesdays, weeks 1–4 and 7-8 of Trinity Term. NB no lectures in weeks 5–6.

Loose Ends

In this lecture, we'll tie up some loose ends. We'll look at:

- Conditional probability
- An alleged philosophical implication of the base rate fallacy
- Thick credences
- Additivity

Conditional Probability I

- In the first lecture, we encountered the Ratio Formula: $p(A|B) = \frac{p(A \cap B)}{p(B)}$, if $p(B) \neq 0$, which we stressed is not a definition of conditional probability (unlike in mathematics) but a constraint on it.
- One reason is that there could be irrational subjects who have conditional probabilities that do not obey the Ratio Formula. Surely these are at least possible, if not actual.
- Another reason is that it *seems* possible for a conditional credence $cr(A \mid B)$ to be well-defined even when neither $cr(A \cap B)$ nor $cr(B)$ is defined. (Or at least one of the last two is not defined.) The following two examples are inspired by Sturgeon (2020, pp. 78–79).

Conditional Probability II: First Example

- You draw a ball from a bag about which you have no information other than the following: the bag contains some coloured balls; exactly three of them are blue; exactly one of the (three) blue balls has a red spot.
- If RS = the claim that you hold a ball with a red spot and B = the claim that you hold a blue ball, what should $cr(RS|B)$ be? It seems pretty clear that the answer is $\frac{1}{3}$.
- You don't seem to have an exact credence in RS nor one in B. If that's right, $cr(RS|B)$ cannot be defined as $\frac{cr(RS \cap B)}{cr(B)}$.
- A reply might be that $cr(RS \cap B) = \frac{1}{N}$ and $cr(B) = \frac{3}{N}$, where N is the (unknown) number of balls in the bag, so that the Ratio Formula's right-hand-side is defined (and of course $\frac{\frac{1}{N}}{\frac{3}{N}} = \frac{1}{3}$). Ideal subjects will have a credence for each of the possible values of N, and then using the Law of Total Probability, $cr(RS|B)$ can be derived from the Ratio Formula.

Conditional Probability III: Second Example

- A madman by the name of Tramp is standing for office. You have no idea what his chances of being elected are. You are 99% certain, however, that if elected he will wreak havoc.
- It seems that your credence in Tramp's wreaking havoc given that he is elected is 0.99, although you have no credence in (a) his being elected or in (b) his being elected and wreaking havoc.
- However, the same style of reply as in the previous example could be given here as well.
- A counter-reply might be that we should not insist on ideal agents having credences in all propositions. (There is a connection here with the Regularity Principle, to be discussed next time.) Moreover, we might be interested in the credences of non-ideal subjects. Or at least not ideal in this respect.

Credal Eliminativism I

Let's now move on to another topic: the base rate fallacy, discussed in the previous lecture. Let's connect it to more philosophical concerns. Some philosophers are suspicious of credences. Richard Holton is one of them. He thinks having credences is too cognitively demanding for creatures like us. Here is how he puts the point (Holton 2014, p. 14):

> *Unless their powers of memory and reasoning are very great, those who employ credences risk being overwhelmed by the huge mass of uncertainty that the approach generates. First, they will have to store very much more information: rather than just discarding the propositions that aren't believed and focusing on those that are, they will have to keep track of all of them and their associated credences. Second, they will have to be able to deploy the complicated methods needed for probabilistic reasoning. The problem will be all the worse if, in Bayesian fashion, they update their credences by conditionalization.*

Credal Eliminativism II

- He continues:

  ...if the Bayesian picture has a role, it is as an idealization. If it is to do that, however, we had better approximate Bayesian agents. It had better be that we can form credences and that we can conditionally update on them, even if we do not always do both perfectly. My contention ... is that even this minimal claim is false. I argue that we cannot form credences at all. The Bayesian approach is not an idealization of something we actually do. Instead, it is quite foreign to us. Just as our core native deliberative state is that of the simple intention, so our core native epistemic state is that of simple, all-out belief. (2014, p. 14)

Credal Eliminativism III

- Holton doesn't deny that beings with different cognitive lives to ours could have credences. He just thinks humans don't.
- And he does concede that humans have what he calls 'partial beliefs', which come in degrees, unlike full beliefs. He explains the contrast between the two as follows (2014, p. 28):

  All-Out Belief
  One all-out believes p if one takes p as a live possibility and does not take not-p as a live possibility.

  Partial Belief
  One partially believes p if one takes p as a live possibility and takes not-p as a live possibility.

- Interestingly, one of his arguments for credal eliminativism (my label, not his) exploits the base-rate fallacy.

Credal Eliminativism IV

- Holton thinks we sometimes believe there is a certain chance that something is true. But he thinks the probability is in the *content* rather than the *attitude*.
- Linguistic phenomena, although inconclusive, seem to support this view. For example he points out (2014, p. 17) that we say

  I believe that p is twice as likely as q

  rather than

  I believe p twice as much as I believe q.

Holton on the Base Rate Fallacy I

As Holton points out, also drawing on Gigerenzer (2008, ch. 9), for most people the following is a very difficult question.

> *The probability that a woman has breast cancer is 1%. If she has breast cancer, the probability that a mammogram will show a positive result is 80%. If a woman does not have breast cancer, the probability of a positive result is 10%. What is the probability that a woman who has a positive mammogram result has breast cancer? (Holton 2014, p. 22)*

Virtually no one gets the answer right. Fewer than one in 10 physicians sampled did!

Holton on the Base Rate Fallacy II

In contrast, the following is a much easier question (half of 12-year-olds can solve it).

> *10 out of every 1000 women have breast cancer. Of these 10 women with breast cancer, 8 will have a positive mammogram result. Of the remaining 990 women who do not have breast cancer, 99 will have a positive mammogram result. What is the probability that a woman who has a positive mammogram result has breast cancer? (Holton 2014, p. 23)*

The Base Rate Fallacy and Credal Eliminativism I

- Holton takes the base rate fallacy to support the idea that probabilistic elements are in the contents rather than attitudes.
- As he sees it, if we had credences, we would find the above calculations easier when they are presented as conditional probabilities rather than frequencies. We could just apply Bayes's rule whereas the frequency representation would require a further step as we would have to convert the data into conditional probabilities.
- In contrast, that the probabilities are represented in the content explains the data. Subjects just do some arithmetic on the content of their belief in the frequency case, and in the conditional probability case try to apply the more complicated Bayes's rule.

The Base Rate Fallacy and Credal Eliminativism II

- Holton's argument is certainly interesting. At best, however, I find it incomplete.
- It's incomplete because the background picture of how (a) we reason with credences, (b) whether this reasoning is explicit or implicit, etc. has been completely omitted. When the contours of that picture are better known, the force of the argument will be clearer.
- It's certainly not clear why the credence view should be committed to our performing Bayesian reasoning with explicit numerical values in a fast and competent way when presented with conditional probability data of the sort contained in the mammogram/cancer case.

The Base Rate Fallacy and Credal Eliminativism III

- Whatever we do, it's unlikely that we explicitly work with exact credences in most everyday situations. My guess—and I confess that I haven't performed the experiment—is that people would do much better if the example used qualitative language.

  The probability that a woman has breast cancer is extremely low. If she has breast cancer, the probability that a mammogram will show a positive result is very high. If a woman does not have breast cancer, the probability of a positive result is lowish. In qualitative terms, what do you think the probability is that a woman who has a positive mammogram result also has breast cancer?

- And the linguistic data might look different then too (e.g. 'I'm much more confident of p than q').

The Base Rate Fallacy and Credal Eliminativism IV

The example also contains what, for most subjects, is new information/data. Maybe we are much better at these sorts of qualitative calculations with more familiar information/data. As in the following structurally identical example to medical one, which uses qualitative terms and more familiar language.

> *If you're a British sportsperson, you're extremely unlikely to win Wimbledon in any given year. But if you do happen to win Wimbledon that year, you stand a very good chance of walking away with the BBC [British] Sports Personality of the Year award. Of course, if you're just any old British sportsperson, the chances of your becoming BBC Sports Personality of the Year that year are fairly low. On that basis, what would you say the chances are that the BBC Sports Personality of the Year is a Wimbledon winner?*

But of course here we have to exclude the hypothesis that people are simply using their background knowledge (or estimate) of the relevant frequency to assess the claim.

Thick Credences I

- Credences are point-valued subjective probabilities. They are hyper-precise.
- Are there also interval-valued subjective probabilities?
- You are told by an authority on the matter that 3-to-5 percent of Norwegians voters vote for the Green Party. How sure should you be that Frida, about whom you know nothing other than that she is a Norwegian voter, votes for the Green Party? It doesn't seem unreasonable to answer: 3 to 5%.
- You are told that exactly 80%-to-90% of balls in a box are red. You reach in the box and are about to grab a ball. How confident should you be that you hold a red ball? It doesn't seem unreasonable to answer: 80-to-90%. (Taken from Sturgeon 2020, p. 63.)
- If we take this at face value, it seems that there are *thick* (or *mushy* or *interval-valued* or *imprecise*) subjective probabilities. The alternative, non-face-value, approach would be to argue that the respective answers should be 4% and 85% respectively.

Thick Credences II

- How do we handle thick credences?
- For example, how do we add thick credences together?
- Sturgeon (2020, pp. 85–6) suggests using midpoints.
- In other words, if $E_1, \ldots E_n$ partition the sample space Ω, and each is assigned a confidence interval $[l_i, u_i]$, for $1 \leq i \leq n$ then $\sum_{i=1}^{n}(\frac{l_i+u_i}{2}) = 1$. Similarly for conditional probability (however we derive it from the n-many intervals $[l_i, u_i]$):

$$\frac{l_{A|B} + u_{A|B}}{2} = \frac{\frac{l_{A\cap B}+u_{A\cap B}}{2}}{\frac{l_B+u_B}{2}}$$

- But to my mind this is a disappointingly thin notion of thick confidence. The interval-valued credence $[l_i, u_i]$ effectively behaves as the point-valued one $\frac{l_i+u_i}{2}$.

Kolmogorov Axioms Again

- Recall from the first lecture the probability/Kolmogorov axioms, defined on a measurable space $\langle \Omega, \Sigma \rangle$:
  - **Non-Negativity:** for any E in Σ, $p(E) \geq 0$.
  - **Normalisation:** $p(\Omega) = 1$.
  - **Countable Additivity:** For any $E_1, E_2, \ldots, E_n, \cdots \in \Sigma$: if $E_i \cap E_j = \emptyset$ for all $i \neq j$ then $p(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} p(E_i)$.
- One structural feature is the use of real numbers $\mathbb{R}$. Another structural feature is that Σ is a σ-algebra over the set of outcomes Ω. This is related to the third axiom.
- An obvious question one could ask about the third axiom: why Countable Additivity? Why not Finite Additivity or Uncountable Additivity?

For Countable Additivity: Extrinsic Reasons

- The standard theory of probability is a deep and central area of mathematics.
- Countable Additivity is justified by its inclusion in this powerful and important theory. In particular, proofs of various limit theorems and laws of large numbers (which are not in the scope of this course) depend on it. Finite Addivity can't do this job.
- But on the same sorts of grounds, that is as far as we should go. Continuum-Sized Additivity would annihilate swathes of probability theory.
- Consider for example the uniform distribution on $[a, b]$, where $a < b$, whose probability density function is $\frac{1}{b-a}$. This distribution could not exist, since $1 =$ the probability of $[a, b]$ would have to be the sum of the uncountably many $p(x)$ for each $x \in [a, b]$, which would have to be the same.
- Call these sorts of reasons for Countable Additivity and against stronger or weaker forms *extrinsic* reasons.

Countable Additivity – Intrinsic Reasons

- There are also *intrinsic* reasons behind Countable Additivity.
- One is that it seems quite easy to conceive of assigning probabilities to countably infinite sets, i.e. a series of events indexed by the natural numbers
- Another is that some of the arguments for Finite Additivity generalise to Countable Additivity. An example is the Dutch Book Argument in an earlier lecture, as long as the sum of the absolute values involved is finite, as shown in Williamson (1999). In fact, our earlier argument did not specify the size of the index set.
- Let's look at an apparent argument against Countable Additivity.

Countably Infinite Fair Lotteries

- A countably infinite fair lottery seems conceivable.
- Countably infinite means that we can take the lottery's tickets to be (indexed by) natural numbers.
- Fair means that every ticket has the same chance of winning.
- Let's call this chance c. Fairness means that $p(i) = c$ for all $i \in \mathbb{N}$.
- By Countable Additivity,

$$p(\mathbb{N}) = \sum_{i \in \mathbb{N}} p(i) = \sum_{i \in \mathbb{N}} c = \begin{cases} 0 \text{ if } c = 0 \\ \text{undefined if } c > 0 \end{cases}$$

- Either way, this conflicts with Normalisation, i.e. the requirement that $p(\mathbb{N}) = 1$.

Response I

- It's not clear that Countable Additivity is the culprit.
- Suppose we retreat to Finite Additivity, i.e. for any $E_1, E_2, \ldots, E_n$ with n finite: if $E_i \cap E_j = \emptyset$ for all $i \neq j$ then $p(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} p(E_i)$.
- A countably infinite lottery in which each ticket has the same chance of winning *is* consistent with Finite Additivity. This chance has to be 0. (It can't be positive since in that case it would be greater than $\frac{1}{n}$ for some n, so that n tickets would have probability greater than 1.)
- Using some more advanced logic than you might have encountered (viz. the UltraFilter Lemma, which follows from the Axiom of Choice), we can show that a function p exists that is defined on $\mathbb{P}(\mathbb{N})$ and such that $p(F) = 0$ for any finite subset of $\mathbb{N}$, and in particular any singleton, and $p(CF) = 1$ for any cofinite subset of $\mathbb{N}$.

Response I Continued

- However, it's far from clear that this genuinely models a countably infinite fair lottery. Let O be the set of even numbers and E the set of even ones. Then either

$$p(O) = 1 \text{ and } p(E) = 0,$$

or

$$p(O) = 0 \text{ and } p(E) = 1.$$

- But this doesn't capture our notion of a countably infinite fair lottery. According to that notion, the sets O and E should have the same probability.
- One can remedy this. Carry out the same trick on each of O and E to obtain probability functions p_O and p_E that are finitely additive, 0 on all finite subsets and 1 on all cofinite sets. Then set $p(X) = \frac{p_O(X \cap O) + p_E(X \cap E)}{2}$, so that $p(O) = p(E) = \frac{1}{2}$.

Response I Continued

- But other intuitive principles aren't satisfied. It's easy to show that the principle 'if $X \subsetneq Y$ then $p(X) < p(Y)$' can't be satisfied by the above lottery.
- Nor more generally can the principle 'if X and Y are equinumerous then $p(X) = p(Y)$' (embodying the usual Cantorian criterion of size). Consider for example a partition $\mathbb{N} = A_1 \cup A_2$, where A_i is infinite and coinfinite, and $\mathbb{N} = B_1 \cup B_2 \cup B_3$, with the same condition. The principle forces $p(A_1) = p(A_2) = \frac{1}{2}$ and $p(B_1) = p(B_2) = p(B_3) = \frac{1}{3}$ but also e.g. $p(A_1) = p(B_2)$.
- The suspicion, then, is that the fact that Countable Additivity is incompatible with the existence of a countably infinite fair lottery is because there's something wrong with the latter notion rather than Countable Additivity.

Response II

- Easwaran (2013, p. 58) lays down a *Comparative Principle*. If P_1 and P_2 are probability distributions,

$$(P_1, E_1) \succ (P_2, E_2)$$

  denotes that E_1 is strictly more likely according to P_1 than E_2 is according to P_2. The Comparative Principle states that if A is a partition for two probability functions P_1 and P_2, it is not the case that for every member a of A, $(P_2, a) \succ (P_1, a)$.
- Consider the 'St Petersburg lottery' in which the n^{th} ticket has a probability of $\frac{1}{2^{n+1}}$ of winning. Every ticket is more likely to win under this lottery than the countably infinite fair one. Hence the Comparative Principle is violated.

Response III

- Rothery (2024) points out a very strange consequence of countably infinite fair lotteries if we assume that any finite initial segment has probability 0. (It can't be non-zero for the reasons given earlier.)
- You play such a lottery on each of the seven days of the week. Seven winning tickets are drawn, one on each day. Let's assume for simplicity (the moral is unaffected) that the tickets are different.
- What are the chances that the sequence of winning numbers is increasing?
- Since there are $N!$ permutations of N objects, you would imagine that the probability is $\frac{1}{N!}$.

Response III Continued

- But actually you should, with probability 1, expect to see an increasing sequence. Suppose Monday's winning ticket was k_1. Because $P(X \leq k_1) = 0$, you should have credence 1 that Tuesday's winning number will be larger than Monday's. Similarly, Wednesday's winning number should be larger than Tuesday's, etc.
- This result seems absurd.
- It also seems inconsistent. There is no reason not to consider the run of tickets in reverse order, starting from Sunday's then Saturday's etc. With probability 1, you should then expect the Sunday-Monday sequence to be increasing, i.e. the Monday-Sunday sequence to be decreasing.

Bibliography for Lecture 4

K. Easwaran (2013), 'Why countable additivity?', *Thought* 2, pp. 53–61.

R. Holton (2014), 'Intention as a Model for Belief', in M. Vargas & G. Yaffe (eds), *Rational Social Agency: The Philosophy of Michael Bratman*, pp. 12–37.

M. Rothery (2024), 'Enumerative Induction and Countable Conjectures in Mathematics', *BPhil Thesis, University of Oxford*.

S. Sturgeon (2020), *The Rational Mind*, Oxford University Press.

J. Williamson (1999), 'Countable Additivity and Subjective Probability', *The British Journal for the Philosophy of Science* 50, pp. 401–16.

Probability and Philosophy

A.C. Paseau

Trinity 2025

Lecture 5, Wednesday 11 June 2025

Lectures

1. Introduction to Probability Theory
2. Dutch Book Arguments
3. Dutch Book Arguments Assessed and Empirical Matters
4. Miscellanea
5. **Bayesianism**
6. Bayesianism and the Problem of Induction

Location: Lecture Room, Radcliffe Humanities (2nd floor)

Time: 9.30 to 11 am on Wednesdays, weeks 1–4 and 7-8 of Trinity Term. NB no lectures in weeks 5–6.

Bayesianism I

- To recap: according to probabilism, an agent's beliefs come in degrees called credences rationally required to obey the probability (i.e. Kolmogorov) axioms.
- Bayesianism, roughly, is the idea that an agent has credences that are rationally required to obey some probabilistic principles. These include the probability axioms and the Ratio Formula encountered earlier (so Bayesians are probabilists). But they may include other principles too.

Bayesianism II

Recall the three probability/Kolmogorov axioms on $\langle \Omega, \Sigma, p \rangle$:

- **Non-Negativity:** for any E in Σ, $p(E) \geq 0$.
- **Normalisation:** $p(\Omega) = 1$.
- **Countable Additivity:** For any $E_1, E_2, \ldots, E_n, \cdots \in \Sigma$: if $E_i \cap E_j = \emptyset$ for all $i \neq j$ then $p(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} p(E_i)$.
- And the **Ratio Formula**:

$$p(A|B) = \frac{p(A \cap B)}{p(B)}, \text{ if } p(B) \neq 0$$

.

- From which you'll remember **Bayes's Theorem** follows (when all terms are well-defined):

$$p(H|E) = \frac{p(E|H).p(H)}{p(E)}$$

Bayesianism III

- The obvious question is *what* other principles do or should Bayesians add to these? What are all the requirements on rational credence?
- Different types of Bayesianism result from adding different principles to these core ones. They all pretty much agree that an updating rule should be added. This is the rule to be discussed shortly which in its simple form is called Conditionalisation.

Objective vs Subjective Bayesianism

- Subjective vs Objective Bayesianism is not a binary distinction but a scale.
- At the (extreme) subjective end of the spectrum are those who think the only rational constraints are adherence to the probability axioms and the Ratio Formula.
- At the (extreme) objective end are those who adhere to a principle along these lines (cited in Titelbaum 2022, p. 129): Given any proposition and body of total evidence, there is exactly one attitude it is rationally permissible for agents with that body of total evidence to adopt towards that proposition.
- This is one of the issues we'll touch on (although certainly not resolve) in this lecture and the next.

Synchronic vs Diachronic Constraints

- Let's think about what principles credence might satisfy other than the four above if they are to be rational.
- In particular, so far we've looked at *synchronic* constraints, i.e. constraints governing a rational subject's credences at a given time.
- Let's now look at a *diachronic* constraint, i.e. a constraint governing a rational subject's credences across time.

Conditionalisation I

- Suppose you know a die to be fair. You then learn that it was thrown and landed on an even number. What should your credence be that it landed on a 2? Answer: $\frac{1}{3}$.
- Suppose more generally your credences are such that $cr(P|Q) = r$. What should your credence in P be if you learn that Q? Answer: r.
- Now think about the situation in which an agent has credences at time t_1 and then learns some facts represented by proposition E between t_1 and t_2. (We will continue to speak of events and propositions interchangeably. And let's assume throughout that E and H are both events in the sigma algebra and that at t_1, $cr(E) > 0$.)
- What should the agent's new credences be? How should she update them in light of this new evidence?
- Let's subscript the agent's (rational) credence function at different times: cr_1 at time t_1 and cr_2 at time t_2.
- The answer to the question just asked seems to be...

Conditionalisation II

- **Conditionalisation:** If E represents everything learnt by the agent between t_1 and later time t_2 then for any H:

$$cr_2(H) = cr_1(H|E)$$

- NB Although the two are closely related, *do not confuse Conditionalisation with the Ratio Formula*. The former is a diachronic constraint (t_1 and t_2 both appear), whereas the latter is synchronic (only one time appears). For comparison, here is the Ratio Formula applied to events H and E at t_1:

$$cr_1(H|E) = \frac{cr_1(H \cap E)}{cr_1(E)}$$

.

- The Ratio Formula concerns credences at a specific time, given by supposing that E occurs, while Conditionalisation deals with how credences are updated over a period of time in which the agent learns that E.

Conditionalisation III

- Conditionalisation is cumulative: conditionalising first on E_1 then on E_2 is equivalent to conditionalising on $E_1 \cap E_2$.
- Conditionalisation is also commutative: conditionalising first on E_1 then on E_2 is equivalent to conditionalising first on E_2 then on E_1.
- Since conditionalising on $E_1 \cap E_2$ is equivalent to conditionalising on $E_2 \cap E_1$, cumulativity implies commutativity.
- These are both desirable features of Conditionalisation: learning two facts is equivalent to learning their conjunction; and the order in which you learn them should not affect your (rational) credences,

Proof that Conditionalisation is Commutative and Cumulative

- Suppose the agent learns E_1 between t_1 and t_2 then E_2 between t_2 and t_3. (We assume all relevant credences are well-defined.)

$$cr_3(H) \underset{\text{Cond}}{=} cr_2(H \mid E_2) \underset{\text{RF2}}{=} \frac{cr_2(H \cap E_2)}{cr_2(E_2)} \underset{\text{Cond}}{=} \frac{cr_1((H \cap E_2) \mid E_1)}{cr_1(E_2 \mid E_1)}$$

$$\underset{\text{RF1}}{=} \frac{\frac{cr_1(H \cap E_2 \cap E_1)}{cr_1(E_1)}}{\frac{cr_1(E_2 \cap E_1)}{cr_1(E_1)}} = \frac{cr_1(H \cap E_2 \cap E_1)}{cr_1(E_2 \cap E_1)} \underset{\text{RF}}{=} cr_1(H \mid E_2 \cap E_1)$$

- Note that $E_2 \cap E_1 = E_1 \cap E_2$.
- So we have verified commutativity and cumulativity.

Jeffrey Conditionalisation I

- We might worry that Conditionalisation relies on an overly simple model of learning from experience.
- Often, a new piece of evidence is not known with certainty. For example, to use a famous example of Richard Jeffrey's, you may have seen a piece of cloth by candlelight but not be entirely sure whether it was green, blue or violet.
- You could say that the colour sense datum you received at the time was something you knew with certainty *then*. But (i) this is of little use to epistemology if you're no longer sure of its colour even a few seconds later; (ii) it's not clear there *are* sense data; (iii) even if there are, it's far from clear we take in their properties completely and infallibly (are you aware of *all* the shapes and colours in your visual field right now?).

Jeffrey Conditionalisation II

- So it's worth modifying Conditionalisation to allow for uncertain evidence.
- The more complex updating rule is known as *Jeffrey Conditionalisation*, first (clearly) stated in Jeffrey (1965).
- We'll give the version based on finite partitions, using credence functions at time t_1 and t_2 as above (where $t_1 < t_2$).
- As above, we assume that $E_1, \ldots, E_n$ and H are all events in the relevant sigma algebra and that $cr_1(E_i) > 0$ for all i.

Jeffrey Conditionalisation III

- **Jeffrey Conditionalisation:** Given a finite partition $E_1, \ldots, E_n$ (i.e. a mutually exclusive and jointly exhaustive set of events):

$$cr_2(H) = \sum_{i=1}^{n} cr_1(H|E_i).cr_2(E_i)$$

- Since it is an updating rule, Jeffrey Conditionalisation is diachronic: note the *cr*-subcripts '1' *and* '2'.
- NB Do not confuse Jeffrey Conditionalisation with the Law of Total Probability:

$$cr_2(H) = \sum_{i=1}^{n} cr_2(H|E_i).cr_2(E_i)$$

or

$$cr_1(H) = \sum_{i=1}^{n} cr_1(H|E_i).cr_1(E_i)$$

These two are synchronic whereas Jeffey Conditionalisation is diachronic, as should be clear from the subscripts.

Jeffrey Conditionalisation IV

- Jeffrey Conditionalisation is reasonable when updates 'originate' (as Jeffrey put it) in the partition $E_1, \ldots, E_n$, as in the cloth example. They are driven by changes in credences in this partition.
- As the equations on the previous slide make clear—look at two different expressions for $cr_2(H)$—Jeffrey Conditionalisation crucially relies on the following *Rigidity Condition*:

$$cr_1(H|E_i) = cr_2(H|E_i) \text{ for all } i$$

- In other words, conditional credences on the E_i do not change between t_1 and t_2, only credences in the E_i themselves.

More Principles

- Conditionalisation and its more sophisticated replacement Jeffrey Conditionalisation are updating rules. They have been proposed as supplements to the credence version of the probability axioms and the Ratio Formula.
- We'll introduce two further principles often argued to be rationally binding on credences: the first is the Principal Principle, the second the Regularity Principle.
- In the next lecture, we'll put Conditionalisation and the Regularity Principle to work to see how Bayesians might tackle the problem of induction.
- Before that, we'll discuss how we might quantify the degree to which evidence confirms a hypothesis.

The Principal Principle I

- Suppose you are told from an authoritative physicist (i.e. a physicist you have complete trust in) that the chance of an uranium atom decaying in the next hour is $\frac{1}{10}$. What should your credence that it will do so be? Answer: $\frac{1}{10}$.
- You hold a coin you know to be fair in your hand. What should your credence be that it will land heads if you toss it? Answer: the same as its chance, namely $\frac{1}{2}$.
- More generally, it seems that you should—if you are rational—adapt your credences to match the chances.
- This is what David Lewis's *Principal Principle* states. It is a chance-credence link.

The Principal Principle II

There are more or less precise ways of stating this. Here is Lewis's (1980, p. 266) formulation, with the notation slightly altered:

> **The Principal Principle.** *Let cr be any reasonable initial credence function. Let t be any time. Let x be any real number in the unit interval. Let* $Ch_t(A) = x$ *be the proposition that the chance, at time t, of A's holding equals x. Let E be any proposition compatible with* $Ch_t(A) = x$ *that is admissible at time t. Then*
>
> $$cr(A|Ch_t(A) = x \,\&\, E) = x.$$

Deference Principles

- The Principal Principle is a sort of deference principle. It says that you should tailor your credences to the objective chances. You should defer to them, if you like.
- As such, we can generalise it. Let Ex be any expert in the broadest sense, i.e. someone or something (it needn't be a person) you trust/defer to. Suppose *their* credence in A is x. What should yours be? Answer: x. In other words:

$$cr(A|Cr_{Ex}(A) = x) = x$$

- Of course we can complicate this analysis if there are several experts, each of whom you defer to partially regarding A.

The Regularity Principle I

- Recall the statement of Conditionalisation, upon learning E between t_1 and t_2: $cr_2(H) = cr_1(H|E)$. Notice that if H's prior probability is 0, its posterior probability must also be 0. (Prior = before updating; posterior = following it.)
- In particular, suppose H is not a logical contradiction. Then however much evidence in support of H the subject might gather, their credence in H will stay at 0.
- This seems counterintuitive, which is why some people have proposed:

The Regularity Principle: *In a rational credence distribution, no logically contingent proposition receives unconditional credence 0.*

(This is Titelbaum's 2022, p. 99 formulation in terms of propositions. Another name for the Regularity Principle is Cromwell's Rule.)

The Regularity Principle II

- The Regularity Principle implies that any partition of the sample space must be countable, since in an uncountable partition all but countably many events have probability 0.
- Argument sketch: there are countably many intervals of the form $I_n = (\frac{1}{n+1}, \frac{1}{n}]$ for n an integer ≥ 1. If the total probability is to sum to 1, the collection of events that has probability in each I_n must be finite. So the set of events with probability in some I_n or other is at most countably infinite and cannot be uncountable. Together with the Regularity Principle, this means that the collection of events cannot be uncountable. (We've assumed that uncountably many of these uncountably many events correspond to logically contingent propositions.)

A Confirmation Measure I

- It is reasonable to see evidence E as to some degree confirming H relative to Pr just in case $Pr(H|E) > Pr(H)$.
- This is known as positive probabilistic relevance.
- But can we measure this degree precisely, and if so how?
- Various *confirmation measures* have been proposed that try to quantify degree of confirmation.
- We'll look at one of these and show that it has some nice features. This does *not* mean that it is the uniquely correct confirmation measure.

A Confirmation Measure II

- The following measure is sometimes known as the log likelihood-ratio measure of confirmation. (We assume non-zero denominators throughout, which implies that $0 < Pr(H) < 1$.)

$$l(H, E) = \log\left[\frac{Pr(E|H)}{Pr(E|\neg H)}\right]$$

- The measure gets its name from the following two versions of Bayes's Theorem.

  $Pr(H|E) = Pr(H).\frac{Pr(E|H)}{Pr(E)}$. Here $\frac{Pr(E|H)}{Pr(E)}$ is the likelihood ratio.

  $Pr(\neg H|E) = Pr(\neg H).\frac{Pr(E|\neg H)}{Pr(E)}$. Here $\frac{Pr(E|\neg H)}{Pr(E)}$ is the likelihood ratio.
- Positive values of l represent confirmation.
- Negative values of l represent disconfirmation.
- A zero l-value represents irrelevance.

Measure l: 1st nice feature

- One nice feature of l is that E confirms H just as much as it disconfirms $\neg H$, i.e.

$$l(H,E) + l(\neg H, E) = 0$$

- This is because

$$\log\left[\frac{Pr(E|H)}{Pr(E|\neg H)}\right] + \log\left[\frac{Pr(E|\neg H)}{Pr(E|\neg\neg H)}\right] = \log\left[\frac{Pr(E|H)}{Pr(E|\neg H)} \cdot \frac{Pr(E|\neg H)}{Pr(E|\neg\neg H)}\right]$$

$$= \log\left[\frac{Pr(E|H)}{Pr(E|\neg H)} \cdot \frac{Pr(E|\neg H)}{Pr(E|H)}\right] = \log(1) = 0$$

Measure I: 2nd nice feature

- Following Carnap and others, it seems reasonable to take entailment to be the strongest form of confirmation and refutation the strongest form of disconfirmation.
- When E entails H, the denominator $Pr(E|\neg H)$ of $\log\left[\frac{Pr(E|H)}{Pr(E|\neg H)}\right]$ is 0, so $I(H,E)$ is infinite and hence larger than any finite amount.
- And the only way $I(H,E)$ can be infinite when all probabilities are well-defined is for $Pr(E|\neg H) = 0$. This amounts to E implying H as long we ignore events with probability 0. (This is reasonable, since an implication can never be fully captured by a confirmation measure based on Pr beyond 'up to probability zero'.)
- In short, I agrees with the intuitive idea that entailment is the strongest form of confirmation.
- And that refutation is the strongest form of disconfirmation.

Measure I: 3rd nice feature I

- Measure I makes confirmation by independent evidence additive.
- To see this, suppose evidence E_1 is probabilistically independent of evidence E_2 conditional on hypothesis H, i.e.

$$Pr(E_1|E_2 \cap H) = Pr(E_1|H)$$

which is equivalent to

$$Pr(E_1 \cap E_2|H) = Pr(E_1|H).Pr(E_2|H)$$

- This also shows by the way that E_1 is probabilistically independent of E_2 conditional on H iff E_2 is probabilistically independent of E_1 conditional on H.
- Suppose further that evidence E_1 is also probabilistically independent of evidence E_2 conditional on hypothesis $\neg H$, i.e.

$$Pr(E_1|E_2 \cap \neg H) = Pr(E_1|\neg H)$$

Measure I: 3rd nice feature II

- The additivity property is that in such a case,

$$I(H, E_1 \cap E_2) = I(H, E_1) + I(H, E_2),$$

  i.e. the total degree of confirmation is simply the sum of the individual degrees of confirmation.

- This is because

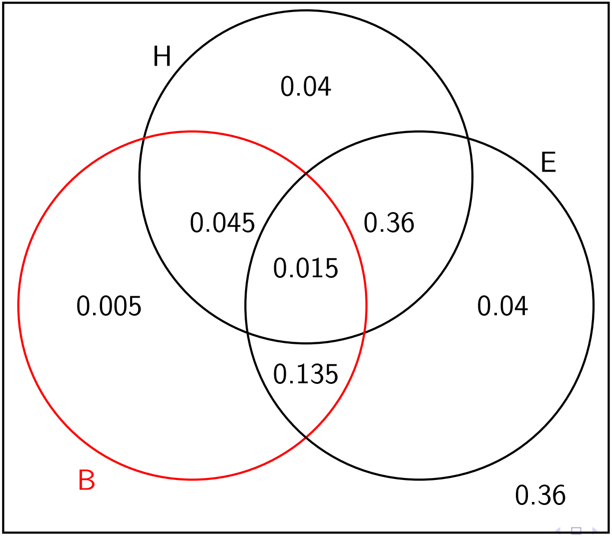
$$\log\left[\frac{Pr(E_1 \cap E_2 \mid H)}{Pr(E_1 \cap E_2 \mid \neg H)}\right] = \log\left[\frac{Pr(E_1|H).Pr(E_2|H)}{Pr(E_1|\neg H).Pr(E_2|\neg H)}\right]$$
$$= \log\left[\frac{Pr(E_1|H)}{Pr(E_1|\neg H)}\right] + \log\left[\frac{Pr(E_2|H)}{Pr(E_2|\neg H)}\right]$$

Confirmation and Jeffrey Conditionalisation I

- Positive probabilistic relevance—$Pr(H|E) > Pr(H)$—seems to be criterial of confirmation.
- That is, conditionalising on E should rationally make you more confident of H iff E confirms H.
- When using Jeffrey Conditionalisation, however, things are more complicated.
- It could be that $Pr(H|E) > Pr(H)$ yet an increase in your confidence in E *decreases* your confidence in H.
- The following example is taken from Titelbaum (2022, p. 449).

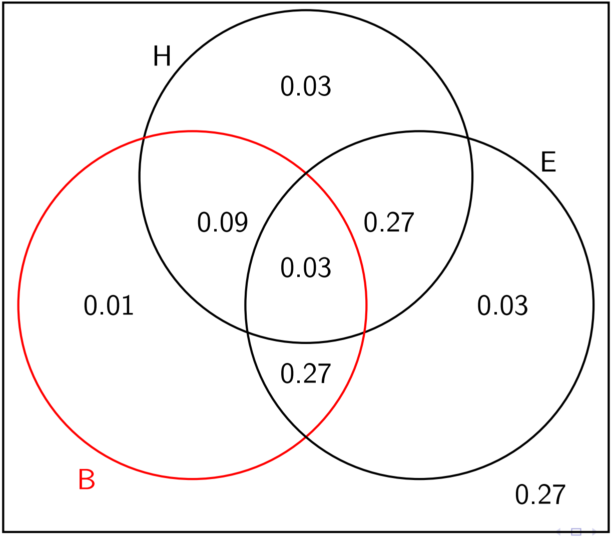
Confirmation and Jeffrey Conditionalisation II

Here's a credence distribution at time t_1. We observe that $pr_1(H) = 0.46$, $pr_1(B) = 0.2$, $pr_1(E) = 0.55$. We're going to J-Conditionalise on $B/\neg B$. $Pr_1(H|E) = \frac{0.375}{0.55} > 0.46 = Pr_1(H)$.



Confirmation and Jeffrey Conditionalisation III

The subject's confidence in B has increased from 0.2 to 0.4. After J-Conditionalising on $B/\neg B$, $Pr_2(E) = 0.6 > 0.55 = Pr_1(E)$ and $Pr_2(H) = 0.42 < 0.46 = Pr_1(H)$ even though $Pr_1(H|E) > Pr_1(H)$.



Confirmation and Jeffrey Conditionalisation IV

- This should not be surprising.
- We should not generally expect that

Your learning that E (i.e. updating your credence in E to 1)
$\Rightarrow$ your confidence in H increases

implies that

Your becoming more confident that E (i.e. revising your credence in E upwards but not all the way to 1)
$\Rightarrow$ your confidence in H increases

Bibliography for Lecture 5

R. Jeffrey (1965), *The Logic of Decision*, McGraw-Hill.

D. Lewis (1980), 'A Subjectivist's Guide to Objective Chance', in *Ifs*, W.L. Harper, R. Stalnaker & G. Pearce (eds), pp. 267–297.

M. Titelbaum (2022), *Fundamentals of Bayesian Epistemology 1: Introducing Credences*, Oxford University Press.

Probability and Philosophy

A.C. Paseau

Trinity 2025

Lecture 6, Wednesday 18 June 2025

Lectures

1. Introduction to Probability Theory
2. Dutch Book Arguments
3. Dutch Book Arguments Assessed and Empirical Matters
4. Miscellanea
5. Bayesianism
6. **Bayesianism and the Problem of Induction**

Location: Lecture Room, Radcliffe Humanities (2nd floor)

Time: 9.30 to 11 am on Wednesdays, weeks 1–4 and 7-8 of Trinity Term. NB no lectures in weeks 5–6.

Hume and the Problem of Induction

- Inductive reasoning is from observed instances to unobserved ones. It lies at the heart of the scientific method and everyday empirical inference.
- Here's a very brief summary of the sceptical problem Hume raised for inductive inference. (I assume you've encountered it before.)
- Hume argued that any justification for inductive reasoning is either a priori or is a posteriori.
- If a priori (resting on relations of ideas, as he called them, including pure mathematics and logic), it cannot convey any empirical information and therefore cannot justify our inductive practices.
- If a posteriori (based on assumptions about 'matters of facts') it must assume some principle linking the observed to the unobserved and therefore begs the question.
- The problem continues to vex philosophers to this day.

Hume's Historical Context and Bayesianism

- Hume was familiar with the mathematical theory of probability of his time. He is known to have read Jakob Bernoulli's *Ars Conjectandi* of 1713, mentioned in the first lecture. He mentions probabilistic ideas (in a non-mathematical way) in his 1739–40 *Treatise* (full title: *A Treatise of Human Nature*), in which he first posed the problem of induction. .
- Bayes's famous essay on probability was published in 1763, in Hume's lifetime (his dates are 1711–1776) but after his major works were published. It's unlikely that Hume read Bayes.
- Since then, various thinkers have argued that probabilistic methods may help counter Hume's objections to induction.
- We'll briefly look at Russell's idea at the start of the 20th century.
- Mainly, we'll focus on how Bayesians might respond to Hume's problem and assess their approach. I've partly drawn on Howson (2000, ch. 4).

Russell on Induction

Here is how Russell expressed one version of Hume's problem.

> *It has been argued that we have reason to know that the future will resemble the past, because what was the future has constantly become the past, and has always been found to resemble the past, so that we really have experience of the future, namely of times which were formerly future, which we may call past futures. But such an argument really begs the very question at issue. We have experience of past futures, but not of future futures, and the question is: Will future futures resemble past futures? This question is not to be answered by an argument which starts from past futures alone. We have therefore still to seek for some principle which shall enable us to know that the future will follow the same laws as the past. (Russell 1912, chapter 6)*

Russell's Inductive Principle I

Russell's solution appeals to what he calls the *principle of induction* and whose two parts he formulates as follows (1912, ch. 6)

> *(a) When a thing of a certain sort A has been found to be associated with a thing of a certain other sort B, and has never been found dissociated from a thing of the sort B, the greater the number of cases in which A and B have been associated, the greater is the probability that they will be associated in a fresh case in which one of them is known to be present;*
> *(b) Under the same circumstances, a sufficient number of cases of association will make the probability of a fresh association nearly a certainty, and will make it approach certainty without limit.*

Russell's Inductive Principle II

Russell (1912, ch. 6) takes the principle to be primitive.

> *The inductive principle, however, is equally incapable of being proved by an appeal to experience. ... All arguments which, on the basis of experience, argue as to the future or the unexperienced parts of the past or present, assume the inductive principle; hence we can never use experience to prove the inductive principle without begging the question. Thus we must either accept the inductive principle on the ground of its intrinsic evidence, or forgo all justification of our expectations about the future... Thus all knowledge which, on a basis of experience tells us something about what is not experienced, is based upon a belief which experience can neither confirm nor confute, yet which, at least in its more concrete applications, appears to be as firmly rooted in us as many of the facts of experience. The existence and justification of such beliefs...raises some of the most difficult and most debated problems of philosophy.*

Russell's Inductive Principle and Circularity

- Howson's dismissive comment about this and similar approaches (2000, p. 19):

  This is clearly not a critical response to Hume's argument, whose validity is the only question we are concerned with here; the inductive-principle advocates explicitly or implicitly acknowledge the validity of the argument.

- The question is whether Bayesianism or any other account of rationality can do better. To answer Hume, we must appeal to some fundamental principles about rationality, including reasoning about 'matters of fact'. Won't any such approach beg the question against him?
- We'll come back to this question of circularity below.

Bayesianism and Induction I

- Consider a stylised inference from finitely many instances of Fs that are observed to be G to the conclusion that all Fs are Gs.
- We may order the observed instances in time: the first is F_1, the second F_2, etc. The universal claim is then 'For all n, F_n is G'. Checking F_i to see whether it is G or not is an 'experiment' with a two-valued outcome: either it is G or it isn't.
- Some of the detail here is for specificity and not essential; e.g. nothing changes if the set of outcomes is finite rather than 2. Note that given human perceptual limitations, even aided by instruments, all measurements are finite.
- Let's call H the true hypothesis that for all n, F_n is G.

Bayesianism and Induction II

- Suppose we know that $F_1, \ldots, F_n$ are G. We then come across F_{n+1} and see that it is also G. Call this piece of evidence, to which we assume you did not attach credence 0 or 1 previously, E. On the Bayesian picture, learning E should boost rational credence in H.
- This fact follows immediately from Conditionalisation, which states that your posterior credence in H, upon learning E, is your prior credence in H given E. And from the Ratio Formula, the latter is $\frac{cr(H \cap E)}{cr(E)}$. Now H implies E, since H is a general claim of which E is an instance. Hence $cr(H \cap E) = cr(H)$. And since $0 < cr(E) < 1$, it follows that

$$\frac{cr(H \cap E)}{cr(E)} = \frac{cr(H)}{cr(E)} > cr(H)$$

  Hence hypothesis H is more likely after learning E than before.
- The result is very general and does not depend on H's form.

Bayesianism and Induction III

- Recall that H is the true hypothesis that for all n, F_n is G. An alternative hypothesis might be that no F is G. Or that all and only the even-numbered F_n are G. And so on.
- Call the set of hypotheses $\{A_i : i \in I\}$ where the index set $I =$ the set of countably infinite sequences each of whose members is a 0 or 1. (H corresponds to the sequence $1111\ldots$)
- By a similar argument to the one given in the last lecture, at most a countable infinity of these hypotheses has non-zero prior probability. Let's assume H is one of them. Renumbering if necessary, the others are $A_1, A_2, \ldots, A_n, \ldots$, where $div(A_i) \leq div(A_j)$ if $i < j$. Here $div(A_i)$ represents the first place where A_i and H diverge; think of A_i as a sequence with a 0 in the $div(A_i)^{th}$ place and 1s up to it (what happens beyond that is not relevant).

Bayesianism and Induction IV

- If as assumed H is the true hypothesis, the more evidence comes in, the greater rational credence in H will be. Indeed, that credence will tend to 1.
- Here's how to ensure $cr(H) \geq \frac{1}{2}$, for example. Suppose N is such that $\sum_{i=N+1}^{\infty} cr(A_i) \leq cr(H)$. Some such N must exist.
- If we check the first $div(A_N)$ Fs and find that they are all Gs (as we will if H is true) then we will have eliminated all the hypotheses up to and including A_N, leaving only those whose joint prior probability is no greater than H's prior probability.
- This argument can be formalised to show that H's posterior probability after this evidence has come in—i.e. after verifying that $F_1, F_2, \ldots, F_{div(A_N)}$ are all G—must be $\geq \frac{1}{2}$.
- A similar argument shows more generally that $cr(H)$ tends to 1 as more of the Fs are confirmed to be G.

Summary of the Bayesian Approach

- We have focused on the universal generalisation that all Fs are Gs, where the Fs are $F_0, F_1, \ldots F_n, \ldots$. This does duty for a large class of inductive inferences.
- Each of the rival hypotheses claims that one out of all the possible countably infinite sequences is the correct one. Each sequence has a 1 or 0 at each place.
- We showed that under the assumptions stated earlier, as the evidence comes in, one's rational credence in H will tend to 1, where H is the true hypothesis.
- Problem solved?

Objection 1: Infinite vs Finite I

- An obvious feature of the way we modelled induction is that the Fs are infinitely many.
- But in many cases, there are only finitely many Fs. (In a more probabilistic idiom: there are only finitely many events in the σ-algebra over our outcome space.) This will be true of *all* empirical cases if the universe is of finite temporal duration.
- Even if the universe is infinite and there are infinitely many Fs, we may make only finitely many predictions. At any rate, it's very unlikely that the human race or our successors (or successors' successors...) will live forever.

Objection 1: Infinite vs Finite II

- However, the convergence result essentially depends on us making infinitely many predictions about $F_0, F_1, \ldots, F_n, \ldots$.
- In the finite case where there are N Fs, rational credence in the true hypothesis H will increase as the evidence comes in and some rival hypotheses with non-zero prior get knocked out, assuming H too has non-zero prior probability. But the argument gives us *no* reason to suppose rational credence in H will tend to 1 as we observe more and more of the N Fs, as long as we haven't see all of them yet.

Objection 1: Infinite vs Finite III

- Still, we might be interested in an in-principle solution to Hume's problem.
- If you like, we might be interested in what the credences of ideal inductive agents who live forever (or an infinite succession of these) should be.
- And in that case, we may ignore the rather inconvenient fact of our finitude.

Objection 2: Zero Credences

- In the last lecture, we encountered the Principle of Regularity which states that in a rational credence distribution, no logically contingent proposition should receive unconditional credence 0.
- And we also saw its implication that no more than countably many events can have positive (i.e. non-zero) probability.
- The attempted Bayesian solution to the problem of induction assumes that initial credence in the true hypothesis H is positive. Otherwise, if it's 0, it can never get off the ground.
- So it looks like we've simply begged the question against the Humean sceptic by assuming that prior credence in $H > 0$.

First Response to the Zero-Credence Objection I

- A first response by the Bayesian: perhaps we can restrict the set of hypotheses to a countable set. For example, for each natural number n (including 0), let

  $$H_n = \text{the first divergence from the initial segment of 1s is at the } n^{th} \text{ place}$$

  To which we can add the (in our example true) hypothesis H, which we might call H_∞: at no point is there divergence from an initial segment of 1s, i.e. the sequence is 111....
- The resulting collection of hypotheses ($H_0, H_1, H_2, \ldots, H_n, \ldots$ plus H_∞) is a partition: either the sequence is entirely made up of 1s or some (possibly null) initial segment consists of 1s followed by a 0, which occurs for the first time at the n^{th} place. And since the partition is countable, every hypothesis can have a positive prior probability.

First Response to the Zero-Credence Objection II

- However, it doesn't seem that coarse-graining the rival hypotheses (so that each subsumes an uncountable number of finer-grained ones) and then comparing them with the fine-grained true one is acceptable. It treats the true hypothesis asymmetrically.
- How do we know, prior to enquiry, which hypotheses to treat individually and which to lump together?
- And what justifies this?

Another Response to the Zero-Credence Objection: Infinitesimals I

- The Bayesian might amend their theory to include infinitesimals. An infinitesimal i is a quantity smaller than any positive number but larger than 0.
- Early modern mathematics employed infinitesimals in the development of the calculus. The derivative of a function was interpreted as the finite (non-infinitesimal) part of the expression $\frac{f(x+\delta x)-f(x)}{\delta x}$, where δx represents an infinitesimally small quantity.
- In the late 19th century, mathematics abandoned the use of infinitesimals. They are not part of the real number line and are proscribed in mainstream mathematical practice.
- However, they can be put on a rigorous footing, as Abraham Robinson showed in the 1960s.

A Second Response to the Zero-Credence Objection: Infinitesimals II

- The Bayesian might try to block the zero-credence objection by arguing that all hypotheses may have non-zero probability if some of them merely have infinitesimal values.
- Even if this can be done compatibly with Countable Addivity, there seems to be a fatal problem for this objection.
- For if the true hypothesis H has infinitesimal prior probability i then its posterior $p(H|E) = \frac{p(H)}{p(E)}$ will be infinitesimal assuming $p(E)$ is real.
- Because infinitesimal $\times$ finite number $=$ infinitesimal.
- So however (finitely) much evidence supports H, its probability will remain infinitesimal if initially infinitesimal.
- In other words, even if the real numbers are extended, we have to assume that H must have non-infinitesimal positive prior probability. So we haven't made any progress by appealing to infinitesimals.

The Problem of the Priors I

- The Zero-Credence Ojection is really a species of the *The Problem of The Priors*.
- An agent's priors exert an ineliminable influence on their posteriors. We can understand an agent's posterior credences as the result of updating credences hypothetically held prior to acquiring any evidence on their total evidence as it comes in. (Some call these hypothetical priors the agent's epistemic standards.) This is a conceptual tool, not a claim about the agent's actual psychological history.
- But where do the priors come from? And how do we justify them?

The Problem of the Priors II

- The Bayesian argument we gave earlier that the rational agent's credence in the true hypothesis H tends to 1 as the evidence comes in is an example of the *washing out of priors*. As the evidence piles up, different agents' priors will converge.
- But as we also saw earlier, a crucial assumption is that the prior in H is non-zero.
- And we can also spin the argument another way.
- For any $\epsilon > 0$ and finite N, there are rational agents with positive prior credence in H who after seeing N cases of Fs being Gs have posterior credence less than ϵ in H.
- N can be huge, e.g. much larger than any instances we will collectively ever see of anything, and ϵ can be tiny.
- From this perspective, that two scientists agree in giving H high credence seems to be just as much an accident of biography as it is mandated by the evidence.
- Shared biology, or psychology, or the non-rational influence of someone else's priors, on your own, etc. don't take us beyond Hume's own observations about 'habit and custom'.

The Problem of the Priors III

- So far, we have assumed that Bayesians must be *subjective* to some degree.
- As we saw in the previous lecture, *objective* Bayesianism implies that only one set of priors is rational, whereas *subjective* Bayesianism allows for latitude here.
- If objective Bayesianism can be justified, it would support the objectivity of inductive reasoning, and, by extension, much of scientific reasoning.
- But the prospects for objective Bayesianism are not encouraging.
- What, for example, is the unique prior credence one should give to all elephants being grey prior to having seen any elephants? Or to having seen any animals?
- We saw in the first lecture that the classical theory of probability's reliance on the Principle of Indifference is problematic. Objective Bayesianism looks like it must rely on something like this principle.

Bayesianism and Circularity I

- Earlier, we sketched Russell's approach to the problem of induction. It seemed flatly circular (as Russell seemed to appreciate—sort of).
- Does the Bayesian solution do better than Russell on the circularity front?
- You could say: it's an all-encompassing theory of how to reason non-deductively.
- But Russell could say something similar about his principle of induction: it's the fundamental principle of non-deductive reasoning.
- Bayesianism is obviously a more elaborate, intricate and mathematically more precise theory than Russell's simple, if not simplistic, principle. But does any answer to the problem of induction on its basis, however otherwise successful, beg the question against the Hume sceptic?

Bayesianism's Scope: Mathematics? I

- What if Bayesianism applied to mathematics and logic, i.e. to what Hume called relations of ideas? If if it were a broader principle of rationality whose application to empirical cases was only a subset of its uses, perhaps it would be hard to accuse it of begging the question?
- So does Bayesianism apply to mathematics?
- Despite appearances to the contrary, mathematics is chock-full of uncertain reasoning.
- For example, mathematicians believe unproved conjectures to various degrees, or choose a research problem based on the chances it's amenable to proof, and so on.
- None of this is to deny the plain fact that in mathematics a result is only regarded as established if it's been proved.

Bayesianism's Scope: Mathematics? II

- Bayesianism, unfortunately, has trouble making sense of uncertain reasoning in mathematics.
- This is because it assumes that rational agents are logically omniscient. Normalisation implies that any logical truth must be believed with credence 1. Which implies that agents cannot give a theorem lower credence than the conjunction of the axioms it follows from.
- And this general assumption is crucial to the applications of Bayesianism outside mathematics.
- This is known as *The Problem of Logical Omniscience* for Bayesianism.
- Even setting this aside, a contemporary Humean might complain that the issue is not so much past vs future as observed vs unobserved. So the fact that inductive reasoning is used in mathematics does not address the Humean worry. It simply extends its scope to mathematics as well.

Bayesianism and Circularity II

- So does Bayesianism beg the question against the Humean sceptical challenge?
- That depends on how it's justified.
- As a theory of rational credences, we've given one sort of justification for it: Dutch Book Arguments. (And mentioned the extrinsic justification behind the mathematical theory of probability.)
- If the bets in a DBA were actual, real-life bets, there would be a problem of circularity. For example, to accept any bet from a bookie, you must form some beliefs about their future actions, and that will involve an inductive inference. To persuade a Humean sceptic to adjust their credences so as not to avoid a Dutch Book in real life, you'll have to convince them that the bookie will pay out according to the terms of the bet, for example, rather than just shower you with money whatever the outcome—and this argument will have to infer claims about the future from facts about the past.

Bayesianism and Circularity III

- What if we think of Dutch Books not as actual but hypothetical? The Humean sceptic, let's agree, can contemplate imaginary scenarios just as well as anybody else. (She can, for example, imagine a world in which inductive inference beyond the year 2025 holds. What she questions is the justification for thinking our world is like this.)
- She may agree that in the imaginary Dutch Book set-up, given its assumptions, a subject should have probabilistically coherent credences (i.e. credences that conform to the probability axioms).
- But from that she need draw no conclusions about what credences are rational in *our* non-imaginary/actual situation.
- 'If it were the case that X then I should ϕ; so I should ϕ' is surely an invalid argument. We need some reason for applying morals about a counterfactual case to an actual one.
- The question is whether Bayesianism can be justified in a different way to avoid this.

Summary

We discussed four problems for the Bayesian approach to Hume's problem of induction. (The second is an instance of the third.)

- Finite vs Infinite
- Zero Credence
- The Problem of Priors
- How to Justify Bayesianism

Bibliography for Lecture 6

C. Howson (2000), *Hume's Problem: Induction and the Justification of Belief*, Oxford University Press.

D. Hume (1739–40), *A Treatise of Human Nature*, various eds.

B. Russell (1912), *The Problems of Philosophy*, various eds.

Acknowledgements

Thanks to Andrew Rao, Tajei Puthucheary, Tianyang Zou and other students who attended these lectures for questions and comments that improved these slides and the thinking that went into them.