Three Paths to Subjective Probability: Dutch Books, Scoring Rules, and Lindley's Measure

Conor Mayo-Wilson

Philosophy of Probability
June 17th, 2013

Review: Three objective Interpretations of Probability

- Frequentist: Probability is just a relative frequency.
- Propensity: Probability is a tendency towards an outcome.
- Logical: Probability is the measure of the degree to which a set of sentences support a conclusion.

Rest of Class: Subjective Theory of Probability

- Probability is a measure of an individual's strength of belief, or
- It is the strength of a rational individual's degree of belief.

Purported Virtues of Subjective Probability:

- **Admissibility:** There are many representation theorems that indicate that degrees of belief are (or ought to be) represented by probabilities.
- **Ascertainability:** Those same theorems often suggest a way to measure degree of belief
- **Applicability:** Subjective probability is one component (the other is utility) in the most widely applied theory of rational-decision making: subjective expected utility theory.

Outline

Outline

Outline

Outline

Outline

Outline

	Sun	Rain
Read	2	3
Biergarten	4	-2
Listen to Nickelback	-10	-10

Decision Matrices: Payoffs to the decision-maker depend upon the unknown state of nature and what action she chooses (in the rows).

Dominance

	Sun	Rain
Read	2	3
Biergarten	4	-2
Listen to Nickelback	-10	-10

	Sun	Rain
Read	2	3
Biergarten	4	-2
Listen to Nickelback	-10	-10

Dominance: If the outcome of some action a_1 (e.g., Listen to Nickelback) is worse than that of another a_2 (e.g., Read) regardless of the state of the world, do not choose a_1.

A little more precisely:

- An action a_1 is strictly dominated by a_2 if the payoff of a_2 is strictly better than that of a_1 in every state of the world.

A little more precisely:

- An action a_1 is strictly dominated by a_2 if the payoff of a_2 is strictly better than that of a_1 in every state of the world.
- An action a_1 is weakly dominated by a_2 if
  - The payoff of a_2 is strictly better than that of a_1 in **at least one** state of the world.
  - The payoff of a_2 is at least as good as a_1 in all remaining states of the world.

Worst-Case

	Sun	Rain
Read	2	3
Biergarten	4	-3

Worst-Case

	Sun	Rain
Read	2	3
Biergarten	4	-3

Worst-Case: Each action has a worst-case payoff. E.g., For Read, it's 2. For Biergarten, it's -3.

	Sun	Rain
Read	2	3
Biergarten	4	-3

	Sun	Rain
Read	2	3
Biergarten	4	-3

Minimax: Pick the action with the best worst-case payoff. Here, it's Read.

- But suppose you look outside, and it's a beautiful spring day in Munich.

- But suppose you look outside, and it's a beautiful spring day in Munich.
- You read the weather forecast, which claims the chance of rain is .5%.

- But suppose you look outside, and it's a beautiful spring day in Munich.
- You read the weather forecast, which claims the chance of rain is .5%.
- You believe the weather forecast.

- But suppose you look outside, and it's a beautiful spring day in Munich.
- You read the weather forecast, which claims the chance of rain is .5%.
- You believe the weather forecast.
- Minimax ignores the subjective probability of rain.

- But suppose you look outside, and it's a beautiful spring day in Munich.
- You read the weather forecast, which claims the chance of rain is .5%.
- You believe the weather forecast.
- Minimax ignores the subjective probability of rain.
- We'd like some decision rule that simultaneously considers payoffs/losses and probability.

Suppose you fully believe the weather forecast, which claims the chance of rain is .5%.

	Sun	Rain
Read	2	3
Biergarten	4	-3

The expected utility of Biergarten is:

$$
\begin{aligned}
\textsc{seu}(Biergarten) &= p(Sun)\cdot 4 + p(Rain)\cdot -3 \\
&= 995\cdot 4 + .005\cdot -3 \\
&= 3.965
\end{aligned}
$$

Suppose you fully believe the weather forecast, which claims the chance of rain is .5%.

	Sun	Rain
Read	2	3
Biergarten	4	-3

In contrast, expected utility of Read is:

$$\begin{aligned} \mathrm{SEU}(\mathit{Read}) &= p(\mathit{Sun}) \cdot 2 + p(\mathit{Rain}) \cdot 3 \\ &= 995 \cdot 2 + .005 \cdot 3 \\ &= 2.005 \end{aligned}$$

Three Decision Rules

- Maximize (subjective) expected utility (SEU)
- Dominance
- Minimax

Here are two simple observations about dominance:

Theorem

Suppose a is a dominant action. Then a is a minimax action and also maximizes subjective expected utility.

Dominant actions maximize expected utility:

	Sun	Rain
Frisbee	5	-1
Biergarten	4	-2

Dominant actions maximize expected utility:

	Sun	Rain
Frisbee	5	-1
Biergarten	4	-2

Suppose one believes the probability of rain is p. Then:

$$\begin{aligned}\mathrm{SEU}(\textit{Frisbee}) &= (5 \cdot p) + (-1 \cdot (1 - p)) \\ \mathrm{SEU}(\textit{Biergarten}) &= (4 \cdot p) + (-4 \cdot (1 - p))\end{aligned}$$

Each term in the sum of Frisbee is bigger than the corresponding term for Biergarten

Here are two simple observations about dominance:

1. **Fact:** Suppose *a* is a dominant action. Then *a* is a minimax action and also maximizes subjective expected utility.

Dominance and Probability

Here are two simple observations about dominance:

1. **Fact:** Suppose a is a dominant action. Then a is a minimax action and also maximizes subjective expected utility.
2. Dominance and minimax are well-defined decision rules even if
   - One does not assign states of the world probabilities; in fact, dominance does not even require qualitative comparison of the likelihood of outcomes.
   - One does not assign outcomes numerical payoffs; the decision rule makes sense even if outcomes can only be qualitatively compared.

Dutch Book Argument: Invented by [Ramsey, 1931] and made famous by [De Finetti, 1937].

Setup: Dutch Books

- You meet a bookie and will bet on some algebra of events $\mathcal{A}$.

Setup: Dutch Books

- You meet a bookie and will bet on some algebra of events $\mathcal{A}$.
- For each event in $E \in \mathcal{A}$, neither you nor the bookie knows whether the event has occurred.
  - E might be a future event. E.g. That it will rain tomorrow.
  - E might be a past event, E.g. That the Dodgers won the 1967 World Series.

Setup: Dutch Books

- You meet a bookie and will bet on some algebra of events $\mathcal{A}$.
- For each event in $E \in \mathcal{A}$, neither you nor the bookie knows whether the event has occurred.
  - E might be a future event. E.g. That it will rain tomorrow.
  - E might be a past event, E.g. That the Dodgers won the 1967 World Series.
- $\mathcal{A}$ can be any set of events such that, once bets are placed, it can be verified (perhaps after some known time delay) whether the events in $\mathcal{A}$ have transpired.
  - This isn't super rigorous. We'll come back to this point later.

Setup: Dutch Books

- For each $E \in \mathcal{A}$, you are asked to give a fair price $P(E)$ for a ticket that pays 1 €if E occurs.

Setup: Dutch Books

- For each $E \in \mathcal{A}$, you are asked to give a fair price $P(E)$ for a ticket that pays 1 €if E occurs.
- If your price $P(E) = 0,30$ €, then the bookie can buy a ticket from you for
  - If E occurs, you pay the bookie 1 €
  - If E does not occur, you keep my money.

Setup: Dutch Books

- For each $E \in \mathcal{A}$, you are asked to give a fair price $P(E)$ for a ticket that pays 1 €if E occurs.
- If your price $P(E) = 0,30$ €, then the bookie can buy a ticket from you for
  - If E occurs, you pay the bookie 1 €
  - If E does not occur, you keep my money.
- You must also **buy** tickets at your price. This is why the price is called fair. For example, if the bookie sells you a ticket at $0,30$ €, then
  - If E occurs, the bookie pays you 1 €.
  - If E does not occur, the bookie keeps your money.

Dominance in Betting:

- Let a_P denote the action “posting prices P”, and let a_N denote the action “post no odds” (i.e. abstaining from betting).

Dominance in Betting:

- Let a_P denote the action "posting prices P", and let a_N denote the action "post no odds" (i.e. abstaining from betting).
- A state of the world is an event $E \in \mathcal{A}$.

Dominance in Betting:

- Let a_P denote the action "posting prices P", and let a_N denote the action "post no odds" (i.e. abstaining from betting).
- A state of the world is an event $E \in \mathcal{A}$.
- We assume the bookie is cunning in the following sense:
    - If there is some collection of bets such that, given your odds, the bookie is guaranteed to win if she makes said bets,
    - Then the bookie will make said bets.

Dominance in Betting:

- A rational person ought not do a_P if it is strictly dominated by a_N.

Dominance in Betting:

- A rational person ought not do a_P if it is strictly dominated by a_N.
- Given our assumptions about the bookie, one should not offer prices such that one is guaranteed to lose money.

Dominance in Dutch Books

- Avoiding dominance in betting, on first glance, seems like a relatively weak criterion on rationality.

Dominance in Dutch Books

- Avoiding dominance in betting, on first glance, seems like a relatively weak criterion on rationality.
- Suppose you accept event odds on the event that "aliens land on earth tomorrow." Then you **are not guaranteed** to lose money. You just most likely will.

- Avoiding dominance in betting, on first glance, seems like a relatively weak criterion on rationality.
- Suppose you accept event odds on the event that "aliens land on earth tomorrow." Then you **are not guaranteed** to lose money. You just most likely will.
- But, in fact, it can be quite demanding ...

Dutch Book Theorem

Theorem (Dutch Book Theorem)

P is a (finitely additive) probability measure on $\mathcal{A}$ if and only if a_P is not strictly dominated by a_N.

Dutch Book Theorem

Here's why the Dutch Book Theorem is relevant to belief:

- Suppose your degrees of belief are **indicated** by your prices.
  - Note: There need be no number "in your head" or mental state that is numerically representable for this to be the case. All that is require is that there is some function from degrees of belief to prices.

Dutch Book Theorem

Here's why the Dutch Book Theorem is relevant to belief:

- Suppose your degrees of belief are **indicated** by your prices.
  - Note: There need be no number "in your head" or mental state that is numerically representable for this to be the case. All that is require is that there is some function from degrees of belief to prices.
- Say your degrees of belief are incoherent if the corresponding prices can be made sure losers.

Dutch Book Theorem

Here's why the Dutch Book Theorem is relevant to belief:

- Suppose your degrees of belief are **indicated** by your prices.
    - Note: There need be no number "in your head" or mental state that is numerically representable for this to be the case. All that is require is that there is some function from degrees of belief to prices.
- Say your degrees of belief are incoherent if the corresponding prices can be made sure losers.
- Then, by the Dutch Book Theorem, your degrees of belief are incoherent if and only if their corresponding prices are not probabilities.

Dutch Book Theorem

Kadane [2011] proof of this theorem is, I think, extremely clear.

I'm not sure I can improve the exposition in one direction, but I'll review it.

Claim: You must always post a price greater than or equal to zero, i.e., $P(E) \geq 0$ for all E.

Proof:

- Suppose you post a negative price, say $P(E) = -0,5$ €.
- The bookie **buys** a ticket from you for E.
- Since your price is negative, you pay the bookie $-0,5$ €.
- If E occurs, then you pay the bookie $1,50$ €in total.
- If E does not occur, then nothing happens. But the bookie still has your $0,5$.
- So you lost money either way, and would have preferred to abstain from betting.

Claim: You must always post a price greater less than or equal to one, i.e., $P(E) \leq 1$ for all E.
Proof:

- Suppose you post a price greater than 1, say $P(E) = 1,5$ €.

Claim: You must always post a price greater less than or equal to one, i.e., $P(E) \leq 1$ for all E.

Proof:

- Suppose you post a price greater than 1, say $P(E) = 1,5$ €.
- The bookie **sells** you a 1 €ticket for E.

Dutch Book Theorem

Claim: You must always post a price greater less than or equal to one, i.e., $P(E) \leq 1$ for all E.

Proof:

- Suppose you post a price greater than 1, say $P(E) = 1,5$ €.
- The bookie **sells** you a 1 €ticket for E.
- If E occurs, then the bookie pays you 1 €, but she still comes out $0,50$ €ahead..
- If E does not occur, then nothing happens. But the bookie still has your $1,5$ €.
- So you lost money either way, and would have preferred to abstain from betting.

Claim: Your price on the impossible event must be zero, i.e., $P(\emptyset) = 0$.

Proof:

- If not, you are willing to buy the bookie's bet on $E = \emptyset$ at some positive price q by the penultimate claim.

Claim: Your price on the impossible event must be zero, i.e., $P(\emptyset) = 0$.

Proof:

- If not, you are willing to buy the bookie's bet on $E = \emptyset$ at some positive price q by the penultimate claim.
- Since E cannot happen, you simply pay the bookie $q > 0$ and lose money for sure.

Dutch Book Theorem

Claim: If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.
Proof: Suppose $P(A) + P(B) > P(A \cup B)$.

- The bookie **sells** you a tickets on A and on B.

Dutch Book Theorem

Claim: If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.
Proof: Suppose $P(A) + P(B) > P(A \cup B)$.

- The bookie **sells** you a tickets on A and on B.
- The bookie **buys** your ticket for $A \cup B$.

Dutch Book Theorem

Claim: If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.
Proof: Suppose $P(A) + P(B) > P(A \cup B)$.

- The bookie **sells** you a tickets on A and on B.
- The bookie **buys** your ticket for $A \cup B$.
- Since $P(A) + P(B) > P(A \cup B)$, the bookie has made some positive amount of money $q > 0$ from this transaction. What happens when the events A, B, and $A \cup B$ are observed?

Claim: If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.
Proof:

- Suppose $A \cup B$ does not occur, then you do not owe the bookie any money for his purchase of $A \cup B$.

Claim: If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.
Proof:

- Suppose $A \cup B$ does not occur, then you do not owe the bookie any money for his purchase of $A \cup B$.
- But since $A \cup B$ does not occur, neither A nor B occurs.
- So the bookie owes you no money for your A ticket or for your B ticket.

Claim: If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.
Proof:

- Suppose $A \cup B$ does occur.
- Then you owe the bookie 1 €for his bet on $A \cup B$.
- Since $A \cup B$ occurs, however, the bookie owes you money as well for your bet on A or on B (or on both!).
- However, since $A \cap B = \emptyset$, the bookie owes you money for **at most one** of your bets.
- So the bookie owes you at most 1€.
- So you make no money.

Claim: If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.
Proof: In sum, if $P(A) + P(B) > P(A \cup B)$, then

- Regardless of whether $A \cup B$ occurs, you lost money in the initial transaction.
- Moreover, you gained no money when the events A, B and $A \cup B$ were observed.
- So you are a sure loser.

Dutch Book Theorem

A similar argument shows that it cannot be the case that $P(A) + P(B) < P(A \cup B)$

Dutch Book Theorem

- The previous argument shows that, if your prices are not probabilities, then you are a sure loser.
- A more complicated argument is required for the converse: if your prices are probabilities, then you are not a sure loser.
    - This "Converse Dutch Book Theorem" was first proven by Lehman [1955] and Kemeny [1955].

Moral: Considerations of dominance + Dutch Book Framework $\Rightarrow$ Degrees of belief are, in a particular sense, probabilities.

Is the Bookie Cunning?

- In the above argument, we assumed the bookie was "cunning" because she always took guaranteed wins.
- Suppose I post 2:1 odds on E and 2:1 odds on E^c. What does such a cunning bookie do with 2?

Is the Bookie Cunning?

- In the above argument, we assumed the bookie was “cunning” because she always took guaranteed wins.
- Suppose I post 2:1 odds on E and 2:1 odds on E^c. What does such a cunning bookie do with 2?
- She might place 1 bets on both E and its complement.

Is the Bookie Cunning?

- Suppose E is the event that aliens will land on earth tomorrow.
- The bookie considers E very unlikely ($p(E) = .001$).

Is the Bookie Cunning?

- Suppose E is the event that aliens will land on earth tomorrow.
- The bookie considers E very unlikely ($p(E) = .001$).
- Consider the following two acts:
  - a_1: Place 1 €bets on both E and its complement.
  - a_2: Place 2 €bet on E^c.
- Is either action dominant? Which action has higher expected utility? Which action is minimax?

Is the Bookie Cunning?

Moral: If the bookie is guaranteed to win, she cannot maximize expected utility.

So she must violate certain principles of rationality that many assume to be the consequences of arguments for subjective probability.

- The previous theorem suggests a way for measuring subjective degrees of belief: ask an individual for his prices!
- But one should be careful. Why did we ask for prices on a 1 €ticket?
  - In contrast, we could have asked for prices for 100 €tickets and divided the price by 100.

- With large amounts of money, people are risk-averse: I am unlikely to risk 1000 €in a gamble even if I think an outcome is very likely.

- With large amounts of money, people are risk-averse: I am unlikely to risk 1000 €in a gamble even if I think an outcome is very likely.
- With small amounts of money, people are risk-seeking: I am happy to engage in 10 cent wages for entertainment.

- With large amounts of money, people are risk-averse: I am unlikely to risk 1000 €in a gamble even if I think an outcome is very likely.
- With small amounts of money, people are risk-seeking: I am happy to engage in 10 cent wages for entertainment.
- 1 €seems about right to avoid both types of behavior,

Measuring Degrees of Belief

- The previous remarks, however, suggest a problem with eliciting beliefs via gambles of this type: money changes value.
  - Money has decreasing marginal utility: fifty dollars is way more valuable to me than for Bill Gates.
  - Loss of money has increasing risk: for me, losing a few hundred Euro is really upsetting, but losing a few thousand would be disastrous.

- The previous remarks, however, suggest a problem with eliciting beliefs via gambles of this type: money changes value.
  - Money has decreasing marginal utility: fifty dollars is way more valuable to me than for Bill Gates.
  - Loss of money has increasing risk: for me, losing a few hundred Euro is really upsetting, but losing a few thousand would be disastrous.
- Asking for prices, therefore, combines elicitation of two types of judgments: value and uncertainty.

- The previous remarks, however, suggest a problem with eliciting beliefs via gambles of this type: money changes value.
  - Money has decreasing marginal utility: fifty dollars is way more valuable to me than for Bill Gates.
  - Loss of money has increasing risk: for me, losing a few hundred Euro is really upsetting, but losing a few thousand would be disastrous.
- Asking for prices, therefore, combines elicitation of two types of judgments: value and uncertainty.
- And it may be difficult to determine how each type of judgment influences the prices we give.

Next Two Weeks: We'll discuss attempts to measure subjective value and probability simultaneously that provide reasons to separate the two types of judgments.

Today, we'll consider a second proposal, namely Lindley's, for measuring subjective probability that tries to divorce judgments of uncertainty from those of value.

Outline

Measurement Standards

In order to measure an individual's "personal" probability, Lindley argues that we should compare

- The individual's judgments concerning likelihood of arbitrary events, with
- The individual's judgments concerning likelihood of some fixed standard collection of events.

> *What does it mean to say that the distance is one foot? All it means is that somewhere there is a metal bar with two thin marks on it. The distance between these two marks is called a foot, and to say that the width of a table is one foot means only that, were the table and the bar placed together, the former would sit exactly between the two marks. In other words, there is a standard, a metal bar, and all measurements of distance refer to a comparison with this standard. . . . Our first task is therefore to develop a standard for uncertainty.*

Lindley [2006] Chapter 3.1. pp. 31-32.

Question: What is the standard Lindley recommends?

Question: What is the standard Lindley recommends?

Answer: Comparison with a single draw from an urn in which *the individual considers* each ball to be equally likely.

- The use of an urn rather than any other device, like a die or roulette wheel, is completely arbitrary. One just needs a device with many outcomes, each of which one considers to be equally likely.

How are the units of measurement defined for the standard?

> *We now make the first of the premises . . . [we] measure your belief that the random withdrawal of a ball from an urn with 100 balls, of which 30 are red, will result in a red ball, as the fraction 30/100.'*

Lindley [2006] Chapter 3.3. pp. 34.

Probability of arbitrary events are then calculated by comparison with the standard:

> *Consider any event that is uncertain for you. It is convenient to fix ideas and take the event of rain tomorrow (Example 1 of Section 1.2), but the discussion that follows applies to any uncertain event. Alongside that event, consider a second event that is also uncertain for you, namely the withdrawal at random of a red ball from an urn containing 100 balls of which some are red, the rest white. For the moment, the number of red balls is not stated. Were there no red balls, you would have higher belief in the event of rain, than in the impossible extraction of a red ball. At the other extreme, were all the balls red, you would have lower belief in rain than in the inevitable extraction of a red ball. Now imagine the number of red balls increasing steadily from 0 to 100. As this happens you have an increasing belief that a red ball will be withdrawn. Since your belief in red was less than your belief in rain at the beginning, yet was higher at the end with all balls red, there must be an intermediate number of red balls in the urn such that your beliefs in rain and in the withdrawal of a red ball are the same. This value must be unique, because if there were two values then they would have the same beliefs, being equal to that for rain tomorrow, which is nonsense as you have greater belief in red with the higher fraction.*

Lindley [2006] Chapter 3.4, pp. 35.

Coherence

Why should uncertainty so measured satisfy the probability axioms, according to Lindley?

Why should uncertainty so measured satisfy the probability axioms, according to Lindley?

Truthfully, I find Lindley's discussion of coherence a bit confusing. Here's his argument that $P(E) = 1 - P(E^c)$.

Consider the event of rain tomorrow (Example 1 of 1.2). Associated with this event is another event that it will not rain tomorrow; when the former is true, the latter is false and vice versa. Generally, for any event, the event which is true when the first is false, and false when it is true, is called the event that is complementary to the first. Just as we have discussed your belief in the event, expressed through a probability, so we could discuss your probability for the complementary event. How are these two probabilities related? This is easily answered by comparison with the withdrawal of a red ball from an urn. The event complementary to the removal of a red ball is that of a white one. The probability of red is the fraction of red balls in the urn and similarly, the probability of white is the fraction of white balls. But these two fractions always add to one, for there are no other colours of ball in the urn; if 30 are red out of 100, then 70 are white. Hence the standard event and its complement have probabilities that add to one. It follows by the comparison of any event with the urn that this will hold generally. If your belief in the truth of an event matches the withdrawal of a red ball, your belief in the falsity matches with a white ball. Stated formally it means the following:

The probability of the complementary event is one minus the probability of the original event.

Lindley [2006].

Two Different Types of Procedures

- The previous quotation suggests different procedures for measuring $p(E)$ and $p(E^c)$.
- But E^c is an event just like E. If $F = E^c$, then $F^c = E$.
- So Lindley seems to derive a fact about probability by changing the way uncertainty is measured.

Two Different Types of Procedures

- The previous quotation suggests different procedures for measuring $p(E)$ and $p(E^c)$.
- But E^c is an event just like E. If $F = E^c$, then $F^c = E$.
- So Lindley seems to derive a fact about probability by changing the way uncertainty is measured.
- A similar remark applies to his derivation of the additivity axiom.

- Lindley's measure of belief, on first glance, seems to work for any individual, rational or not.
  - The individual can be asked to compare her uncertainty of an event E with that of a drawn ball from different urns.
- But the derivation of $P(E) = 1 - P(E^c)$ and the additivity axiom seems to require some notion of coherence, which Lindley does not rigorously define.

Outline

Thus far, we've seen two procedures for eliciting degrees of belief: Lindley's and wagering.

Here's a third that works for rational individuals.

To motivate scoring rules, it's best to begin with a familiar improper scoring rule: 0/1 loss.

Scoring Rules

- Your high school teacher wants to test how much you know.

Scoring Rules

- Your high school teacher wants to test how much you know.
- She wants to design a test that makes sure you are not rewarded for guessing.

- Your high school teacher wants to test how much you know.
- She wants to design a test that makes sure you are not rewarded for guessing.
- So she wants to measure how **confident** you are in your answers.

- On your first exam, your instructor gives you a true/false exam.

- On your first exam, your instructor gives you a true/false exam.
- Suppose you are very confident a Statement 1 is true. What do you answer?

- On your first exam, your instructor gives you a true/false exam.
- Suppose you are very confident a Statement 1 is true. What do you answer?
- Suppose you are only very slightly more confident that Statement 1 is true than false. What do you answer?

- On your first exam, your instructor gives you a true/false exam.
- Suppose you are very confident a Statement 1 is true. What do you answer?
- Suppose you are only very slightly more confident that Statement 1 is true than false. What do you answer?
- So true/false tests are not the best vehicle for measuring you strength of belief.

- On your second exam, your instructor gets a bit creative.

- On your second exam, your instructor gets a bit creative.
- She asks you to write a number $r(E)$ that **R**eports your strength of belief in the statement E.

- On your second exam, your instructor gets a bit creative.
- She asks you to write a number $r(E)$ that **R**eports your strength of belief in the statement E.
- She tells you that you will be **penalized** according to the following distance between your strength of belief

$$\text{Penalty} = \begin{cases} r(E)^2 \text{ if } E \text{ is false.} \\ (1 - r(E))^2 \text{ if } E \text{ istrue.} \end{cases}$$

- Such a penalty function is called a scoring rule.

Scoring Rules

- On your second exam, your instructor gets a bit creative.
- She asks you to write a number $r(E)$ that **R**eports your strength of belief in the statement E.
- She tells you that you will be **penalized** according to the following distance between your strength of belief

$$\text{Penalty} = \begin{cases} r(E)^2 \text{ if } E \text{ is false.} \\ (1 - r(E))^2 \text{ if } E \text{ istrue.} \end{cases}$$

- Such a penalty function is called a scoring rule.
- Your goal is to minimize your penalty.

- Suppose for the moment that your strength of belief p is, in fact, a probability function.

- Suppose for the moment that your strength of belief p is, in fact, a probability function.
- **Question:** Is there any incentive for you to report your strength of belief is higher or lower than it actually is?
    - For example, if you are more confident than not that E is true, might you report that your degree of belief is 1 in order to minimize your penalty?

Definition: A scoring rule is proper if, for all events E, the unique subjective expected utility maximizing action is to report your true degree of belief $p(E)$.

Definition: A scoring rule is proper if, for all events E, the unique subjective expected utility maximizing action is to report your true degree of belief $p(E)$.

Squared distance (also called Briar score) is one such proper scoring rule.

What de Finetti shows in the section you read, therefore, is that a rational individual (i.e., one who maximizes subjective expected utility) will report her true degree of belief if penalized by squared error error, which is a proper scoring rule.

Theorem

Squared error is a proper scoring rule.

Proof:
Let E be any event, and let p be your degree of belief that E will occur. Recall, we are assuming that your degrees of belief are probabilities. Let a_r denote the action in which you report you degree of belief to be r. We want to show that if $r \neq p$, then:

$$\textsc{seu}(a_r) < \textsc{seu}(a_p)$$

In other words, you expect a higher penalty if you report something other than your degrees of belief.

Theorem

Squared error is a proper scoring rule.

Proof:

The expected utility of the action a_r is as follows:

$$\begin{aligned} seu(a_r) &= p \cdot \text{Penalty of reporting } r \text{ if } E \text{ is true} \\ &+ (1-p) \cdot \text{Penalty of reporting } r \text{ if } E \text{ is false} \\ &= -\left(p \cdot (1-r)^2 + (1-p) \cdot r^2\right) \end{aligned}$$

THEOREM

Squared error is a proper scoring rule.

Proof:

The expected utility of the action a_r is as follows:

$$\text{SEU}(a_r) = -\left(p \cdot (1-r)^2 + (1-p) \cdot r^2\right)$$

If $r = p$, then this equation simplifies as follows:

$$\begin{aligned} \text{SEU}(a_p) &= -\left(p \cdot (1-p)^2 + (1-p) \cdot p^2\right) \\ &= -\left(p - 2p^2 + p^3 + p^2 - p^3\right) \\ &= p^2 - p \end{aligned}$$

THEOREM

Squared error is a proper scoring rule.

Proof:

Suppose $r \neq p$. Then:

$$\begin{aligned}
\text{SEU}(a_p) - \text{SEU}(a_r) &= (p^2 - p) + \left(p \cdot (1-r)^2 + (1-p) \cdot r^2\right) \\
&= (p^2 - p) + \left(p - 2pr + pr^2 + (r^2 - pr^2)\right) \\
&= (p^2 - p) + \left(p - 2pr + r^2\right) \\
&= p^2 - 2pr + r^2 \\
&= (p - r)^2 \\
&> 0 \text{ because } r \neq p.
\end{aligned}$$

In other words:

$$\text{SEU}(a_p) > \text{SEU}(a_r)$$

□

References I

De Finetti, B. (1937). Foresight: its logical laws in subjective sources. In Kyburg, H. and Smokler, H., editors, *Studies in Subjective Probability*.

Kadane, J. B. (2011). *Principles of uncertainty*, volume 92. Chapman & Hall.

Kemeny, J. G. (1955). Fair bets and inductive probabilities. *The Journal of Symbolic Logic*, 20(3):263–273.

Lehman, R. S. (1955). On confirmation and rational betting. *The Journal of Symbolic Logic*, 20(3):251–262.

Lindley, D. V. (2006). *Understanding uncertainty*. Wiley-Interscience.

Ramsey, F. P. (1931). Truth and probability (1926). *The foundations of mathematics and other logical essays*, page 156198.