edX

Course › Newco... › Maximi... › Back to...

Back to Newcomb

What does the Principle of Expected Value Maximization say we should do in a Newcomb situation?

Recall that you can choose between the following two options:

$1B$	$2B$
One-Box	Two-Box

Recall, moreover, that there are two relevant states of the world:

F	E
The large box is full	The large box is empty

This means that there are four possible outcomes, depending on whether you One-Box or Two-Box and on whether the large box is full or empty.

The value of each outcome is as follows:

	Large box full (F)	**Large box empty (E)**
One-Box ($1B$)	\$1,000,000	\$0
Two-Box ($2B$)	\$1,001,000	\$1,000

When the predictor is assumed to be 99% accurate, we work with the following probabilities:

	Large box full (F)	**Large box empty (E)**
One-Box ($1B$)	99%	1%
Two-Box ($2B$)	1%	99%

As before, each cell in this matrix corresponds to a *conditional probability*. For example, the probability in the lower right corner corresponds to the probability of finding the large box empty, given that you two-box.

The expected value of one-boxing and two-boxing can then calculated as follows:

$$\begin{array}{llllllllll} EV(1B) & = & v(1B\,F) & \cdot & p(F|1B) & + & v(1B\,E) & \cdot & p(E|1B) & \\ & = & 1000000 & \cdot & 0.99 & + & 0 & \cdot & 0.01 & = & 990,000 \end{array}$$

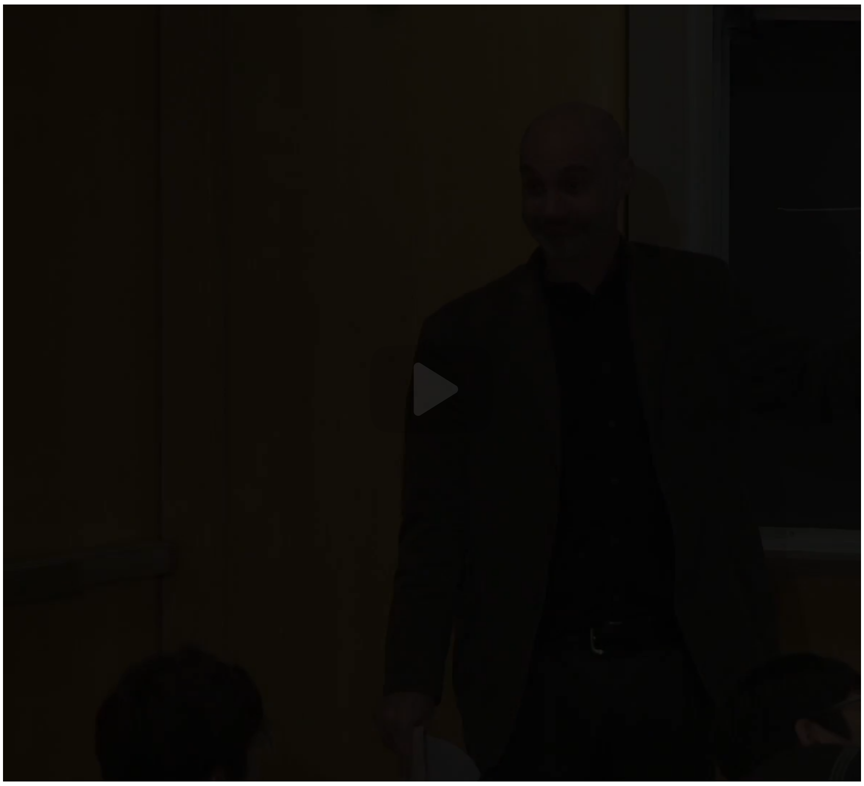
$$\begin{array}{llllllllll} EV(2B) & = & v(2B\,F) & \cdot & p(F|2B) & + & v(2B\,E) & \cdot & p(E|2B) & \\ & = & 1001000 & \cdot & 0.01 & + & 1000 & \cdot & 0.99 & = & 11,000 \end{array}$$

The expected value of one-boxing ($990,000) is much greater than the expected value of two-boxing ($11,000). So according to the Principle of Expected Utility Maximization, we should one-box.

Is that the right result?

We'll try to tackle that question in the next section.

Video Review: Calculating EV in the Newcomb Problem



That is 0.8.

And that gives us 200,100.

So on this way of thinking about decision theory,

we get the recommendation to one-box,

because that's what maximizes expected value.

So if you are right, and I think that you are right,

standard decision theory is wrong.

3:27 / 3:27 1.50x

End of transcript. Skip to the start.

Video

Download video file

Transcripts

Download SubRip (.srt) file

Download Text (.txt) file

Problem 1

2/2 points (ungraded)

Assume the predictor has an 80% chance of making the right prediction.

What is the expected value of one-boxing?

800000

✔ **Answer:** 800000

800000

What is the expected value of two-boxing?

201000

✔ **Answer:** 201000

201000

Explanation

When $p = 80\%$, the expected value of one-boxing is

$$1,000,000 \cdot 0.8 + 0 \cdot 0.2 = 800,000$$

the expected value of two-boxing is

$$1,001,000 \cdot 0.2 + 1,000 \cdot 0.8 = 200,200 + 800 = 201,000$$

Submit

Answers are displayed within the problem

Problem 2

1/1 point (ungraded)

How accurate does the predictor need to be in order for the Principle of Expected Value Maximization to entail that one should one-box?

In other words: identify the smallest number x such that, as long as the predictor has an accuracy of more than x, the Principle of Expected Value Maximization entails that one should one-box.

(As usual, assume that the small box contains $1000 and that the large box contains $1M or nothing.)

0.5005

Answer: .5005

0.5005

Explanation

If p is the probability that the expert's prediction is accurate, the expected value of one-boxing is

$$1,000,000 \cdot p + 0 \cdot (1 - p)$$

and the expected value of two-boxing is

$$1,001,000 \cdot (1 - p) + 1,000 \cdot p$$

These two expected values are equal when $p = 0.5005$:

$$\begin{aligned}
1,000,000 \cdot p &= 1,001,000 \cdot (1 - p) + 1,000 \cdot p \\
1,000,000 \cdot p - 1,000 \cdot p &= 1,001,000 \cdot (1 - p) \\
999,000 \cdot p &= 1,001,000 - 1,001,000 \cdot p \\
2,000,000 \cdot p &= 1,001,000 \\
p &= \frac{1,001,000}{2,000,000} \\
p &= .5005
\end{aligned}$$

So as long as p is greater than 50.05%, the expected value of one-boxing will be greater than the expected value of two-boxing.

This means that as long as the predictor is at least 50.05% accurate, the Principle of Expected Value Maximization entails that one should be a one-boxer.

(That is pretty remarkable. Even if experts that are accurate 99% of the time exist only in science fiction, it is not hard to find a predictor that is accurate 50.05% of the time. I myself have an 80% success rate when performing this experiment on my students!)

Submit

Answers are displayed within the problem

Problem 3

1/1 point (ungraded)

Is the following true or false?

> As long as one has a predictor who does better than chance, it is possible to find payoffs that generate a Newcomb Problem.
>
> More precisely: for any small-box value s and any positive value ϵ ($0 < \epsilon \leq 0.5$), one can find an large-box value l such that when the predictor has an accuracy of $0.5 + \epsilon$, the Principle of Expected Value Maximization entails that one should one-box in a Newcomb Scenario based on s and l.

True

Answer: True

Explanation

When the predictor is $0.5 + \epsilon$ accurate, the expected value of one-boxing is

$$l \cdot (0.5 + \epsilon) + 0 \cdot (0.5 - \epsilon)$$

and the expected value of two-boxing is

$$(l + s) \cdot (0.5 - \epsilon) + s \cdot (0.5 + \epsilon)$$

This means that the Principle of Expected Value Maximization will recommend one-boxing if and only if:

$$\begin{aligned}
(l + s) \cdot (0.5 - \epsilon) + s \cdot (0.5 + \epsilon) &< l \cdot (0.5 + \epsilon) \\
(l + s) \cdot (0.5 - \epsilon) &< (l - s) \cdot (0.5 + \epsilon) \\
l \cdot (0.5 - \epsilon) + s \cdot (0.5 - \epsilon) &< l \cdot (0.5 + \epsilon) - s \cdot (0.5 + \epsilon) \\
s \cdot (0.5 - \epsilon) + s \cdot (0.5 + \epsilon) &< l \cdot (0.5 + \epsilon) - l \cdot (0.5 - \epsilon) + \\
s \cdot (0.5 - \epsilon + 0.5 + \epsilon) &< l \cdot (0.5 + \epsilon - 0.5 + \epsilon) \\
s &< l \cdot 2\epsilon \\
\frac{s}{2\epsilon} &< l
\end{aligned}$$

This means that as long as $\epsilon > 0$, we can guarantee that the Principle of Expected Value Maximization recommends one-boxing by making l sufficiently big.
(Note that when $s = 1,000$ and $\epsilon = 0.0005$, which are the operative values in the previous exercise, the Principle Expected Value Maximization will recommend one-boxing as long as $l > 1,000,000$.)

Submit

Answers are displayed within the problem

Problem 4

1/1 point (ungraded)

Consider a variant of the Newcomb case that works as follows:

> You can choose either the large box or the small box, but not both.
>
> If the predictor predicts that you would choose the large box, then she left the large box empty, and placed $100 in the small box.
>
> If the predictor predicted that you will choose the small box, then she left the small box empty, and placed $1000 in the large box.

According to the Principle of Expected Value Maximization, which of the two boxes should you choose? (Assume the predictor is not perfectly accurate.)

- (•) Choose the large box.
- ( ) Choose the small box.
- ( ) None of the above.

✓

Explanation

The expected value of choosing the large box is

$$(1000 \cdot (1 - p)) + (0 \cdot p) = 1000 \cdot (1 - p)$$

The expected value of choosing the small box is

$$(100 \cdot (1 - p)) + (0 \cdot p) = 100 \cdot (1 - p)$$

So the Principle of Expected Value Maximization entails that one should choose the large box (for every case except $p = 1$; in that case, both options have expected value 0, so the Principle of Expected Value Maximization doesn't entail that you ought to choose one option over the other).

Submit

Answers are displayed within the problem

Discussion

Topic: Week 5 / Back to Newcomb

Hide Discussion

Add a Post

Show all posts by recent activity

- These conditional probabilities do not make sense to me 21
- Agustin experiment in the class 3
  So far it seems that Agustin is 9/11 in predicting correctly what the student would do. I want Agustin ...
- The best outcome for problem 4 would be choose based on a coin flip. 4
  Expected value based on a coin flip is 550, which is better than the EV if the predictor is more than 5...
- Possible error in Problem 4 3