# 24.118: Paradox and Infinity, Spring 2019
# Problem Set 6: Probability

How these problems will be graded:

- In Part I there is no need to justify your answers. Assessment will be based on whether your answers are correct.
- In Part II you must justify your answers. Assessment will be based both on whether you give the correct answer and on how your answers are justified. (In some problem sets I will ask you to answer questions that don't have clear answers. In those cases, assessment will be based entirely on the basis of how your answer is justified. Even if it is unclear whether your answer is correct, it should be clear whether or not the reasons you have given in support of your answer are good ones.)
- *No answer may consist of more than 150 words.* Longer answers will not be given credit. (Showing your work in a calculation, a proof, or a computer program does not count towards the word limit.)
- You may consult published literature and the web. You must, however, credit all sources. Failure to do so constitutes plagiarism and can have serious consequences. For advice about when and how to credit sources see: https://integrity.mit.edu/ (You do not need to credit course materials.)

## Preliminaries

Recall that a **probability function**, $p(\ldots)$, is an assignment of real numbers between 0 and 1 to propositions that satisfies the following two coherence conditions:

**Necessity** $p(A) = 1$ whenever $A$ is a necessary truth

**Additivity** $p(A \vee B) = p(A) + p(B)$ whenever $A$ and $B$ are incompatible propositions

We take logically equivalent sentences to express the same proposition, which means that "$p(A) = p(B)$" is true whenever $A$ and $B$ are logically equivalent. (So, for instance, $p(A) = p(AB \vee A\overline{B})$.)

Throughout this problem set, we will assume that the credence function of a rational agent is always a probability function, and always satisfies the following principle:

**Bayes' Law** $p(AB) = p(A) \cdot p(B|A)$

In addition, we will assume the following:

**Update by Conditionalization** If $S$ is rational, she will update her credences as follows, upon learning that $B$:

$$p^{new}(A) = p^{old}(A|B)$$

where $p^{old}$ is the function describing $S$'s credences before she learned that $B$, and $p^{new}$ is the function describing her credences after she learned that $B$.

Finally, we will assume the following connection between objective and subjective probabilities:

**The Objective-Subjective Connection** The objective probability of $A$ at time $t$ is the subjective probability that a perfectly rational agent would assign to $A$, if she had perfect information about the way the world is before $t$ and no information about the way the world is after $t$.[1]

It'd be nice if we could also assume the Principle of Indifference:

**Principle of Indifference** Consider a set of propositions and suppose one knows that exactly one of them is true. Suppose, moreover, that one has no more reason to believe any one of them than any other. Then, insofar as one is rational, one should assign equal credence to each proposition in the set.

But, as you may recall from Section 6.2.4 of the lecture notes, the Principle of Indifference leads to absurd results. At the same time, it's not clear what to put in its place. So we'll leave it in place warily: we'll try not to use it, but we will if we must.

**Logical Notation**

- $\overline{A}$ is the negation of $A$;
- $AB$ is the conjunction of $A$ and $B$;
- $A \vee B$ is the disjunction of $A$ and $B$.

# Part I

[No Questions in Part I. Go straight to Part II.]

[1] *Nerdy observation:* Here I am tacitly presupposing that a perfectly rational agent is always certain about the objective probabilities at $t$, given full information about how the world is before $t$. So, in particular, for each complete history of the world up to $t$, $H_t$, there is a specification $P_t$ of the objective probabilities at $t$ such that the agent treats $H_t$ and $H_1 P_t$ as equivalent. This assumption is potentially controversial but adds simplicity to our discussion.

## Part II

1. Assume $p(B|A) = p(\overline{B}|\overline{A})$ and $p(A) = p(\overline{A})$. Use Bayes' Law and the fact that $p(\dots)$ is a probability function to prove each of the following results. You must display every step of your proofs. Earlier results may be used in your proofs of later results. (5 points each)

   (a) $p(BA) = p(\overline{B}\,\overline{A})$

   (b) $p(C|A) = 1 - p(\overline{C}|A)$, for arbitrary $C$

   (c) $p(\overline{B}A) = p(B\overline{A})$

   (d) $p(B) = p(\overline{B})$

2. For each of the questions below, there is only one right answer. Half of your grade will be determined by whether the given answer is correct. The other half will depend on whether you are able to justify your answer using one or more of the principles mentioned in "Preliminaries". Do your best to make your justification water-tight. If you don't succeed, then at least make sure that you identify the right principles from "Preliminaries" and use them sensibly. Since the Principle of Indifference leads to absurd results, use it only as a last resort.

   (a) Let $x$ be a particle of $^{235}$U. What credence should you assign to the proposition that $x$ will decay sometime within the next $7.04 \cdot 10^8$ years? The half life of $^{235}$U is $7.04 \cdot 10^8$ years. (10 points)

   (b) A standard deck of 52 cards has been thoroughly shuffled. What is the objective probability that the card at the top of the deck is the three of hearts? (10 points)

   (c) Assuming you have no information beyond what's stated in (2b) (and basic background information), what credence should you assign to the proposition that the card at the top of the deck is the three of hearts? (10 points)

   (d) As in (2b), except that this time a time traveler informs you that she has seen the card at the top of the deck being drawn and that it is, indeed, the three of hearts. Assume that there is no funny business going on (e.g. no slight of hand) and that the time-traveller is completely reliable. What is the objective probability that the card at the top of the deck is the three of hearts? (10 points)

   (e) As in (2d). What should be your credence in the proposition that the card at the top of the deck is the three of hearts? (You may assume that you are *certain* that there is no funny business going on and that the time-traveller is completely reliable.) (10 points)

3. An urn is filled using the following procedure. At each time $t_i$ ($0 \leq i \leq 2$), two coins are tossed. The first toss is used to decide what kind of object to add to the urn: a marble or a die. The second toss is used to decide the color of the selected object:

black or white. At the end of the process, there are three objects in the box, each a marble or a die, each black or white.

For each of the questions below, make sure you justify your answers using the assumptions listed in "Preliminaries".

(a) What should be your initial credence in the proposition that every die in the urn is black (equivalently: the proposition that no object in the urn is a white die)? (5 points)

(b) Your friend draws an object at random from the urn, looks at it, and places it back in the urn. She informs you that the object she looked at is a die, but she does not clarify whether it's black or white. You are certain that your friend speaks truly. What should your updated credence be in the proposition that every die in the urn is black? (5 points)

(c) Suppose your friend had made a different announcement in the situation described in (3b). Rather than announcing that she drew a die, she announces that she drew a white marble. As before, you are certain that she speaks truly. What should your updated credence be in the proposition that every die in the urn is black? (5 points)

4. A fair coin will be tossed until it lands heads. If the coin lands heads on the $n$th toss, you get $\$2^n$. (If the coin never lands Heads, no money exchanges hands.)

(a) What is the expected dollar value of playing the game? (5 points)

(b) A standard assumption in decision theory is that if the expected dollar value of an option is greater than $\$m$, then you should be willing to pay $\$m$ for the privilege of taking that option. On this assumption, how much should one be willing to pay for the privilege of playing the game? (5 points)

(c) Does the standard assumption above deliver the right result in this case? If so, explain why. If not, explain what is wrong with the standard assumption? (5 points)