Handout 3
Probability II: "Dutch Books" &
Conditionalization

DECISIONS, GAMES &
RATIONAL CHOICE

"Dutch Books"

Why is it so important that we can derive the "laws" of probability from a small set of axioms? One reason is that it makes it much easier to *justify* them. We only need to justify the axioms themselves, and then justification for what follows from the axioms is immediate.

There are two ways of thinking of probability.

> ***Objective Probability***: The chance, independent of any agent's evidence, that some event occurs.
>
> ***Subjective Probability***: The degree of belief a particular agent does or should have in an event occurring.

Quantum mechanics might provide examples of objective probabilities. On the other hand, consider how likely you think it is that it will rain tomorrow, and how likely a particular meteorologist thinks it is that it will rain tomorrow. You might attach different subjective probabilities—also called ***credences***—to the same event. What's more, it seems like you *both* can be rational in doing so (because of your different bodies of evidence).

There is a very interesting way of trying to justify the axioms of probability as rules governing rational *subjective* probabilities or credences. It proceeds by first linking credences to betting behavior.

Let's say that a ***bet with a:b odds on e*** is a bet such that if the better puts an amount proportional to $b in a pot, and her opponent puts an amount proportional to $a in the pot, if e happens the better takes the pot, and if e doesn't happen the opponent takes the pot. In other words, it's a bet which on which you lose $n if e doesn't happen, but on which you gain (a/b) x $n if e happens.

> *Assumption*: let's suppose that someone who believes event *e* will occur to degree P(e), will always and only be willing to take a bet with odds of at least
>
> 1-P(e):P(e) on e …or…
> P(e):1-P(e) on not e.

Now, let's define the following

> A series of bets with a person X constitutes a ***dutch book***, if X is sure to lose money on the bets overall, no matter what the outcome.

For example, suppose I get you to agree to the following bets:

> $10 on rain at 2:1 odds.
> $8 on not rain at 2:1 odds.

If it rains you lose more than you win. If it doesn't rain you lose more than you win. So you lose money no matter what. You've been "Dutch Booked".

Interesting fact

> **Dutch Book Theorem.** *Given our assumption about the relationship between credence and betting, X can be "dutch booked" if and only if X's credences fail to obey the probability axioms.*

Example 1

Jane believes it's only 99% likely that either it will rain tomorrow or it won't.

Given our assumption Jane will take a bet that (it will rain or not) at odds up to

$$1\text{-}.99{:}.99 = \ .01{:}.99 \ = \ 1{:}99$$

Let's not use that fact. Jane will also take a bet that [it will not (rain or not rain)] at odds no less than 99:1.
In particular, she should be willing to take the bet:

> $10 on [not (rain or not rain)] at 99:1

But since it will definitely rain or not rain, she'll lose $10 no matter what.

Example 2

Jones believes it is 25% likely that Alice will win the race, and 80% likely that Betty will, and that it is (100%) certain that one and only one of them can win.

Given our assumption, Jones will take a bet on Alice winning on at least 1-.25:.25 = .75:.25 = 3:1 odds
Jones will take a bet on Betty winning on at least 1-.8:8 = .2:.8 = 1:4 odds

So Jones should be willing to take the following two bets:

> $10 on Alice (alone) winning at 3:1
> $32 on Betty (alone) winning at 1:4

Again this guarantees a loss.

I'll spare you the full details of the proof. More importantly: how strong is this justification? Will a rational person always bet along their credences as per our assumption?

Conditionalization

We've just seen reasons why you might think that the axioms of probability might provide rules for how a rational person should adjust their beliefs *at a single time*. This kind of rationality has its own name.

> ***Synchronic Rationality of Belief***: the rationality of a belief state as ascertained at a single point in time.
>
> ***Diachronic Rationality of Belief***: the rationality of a belief state as it evolves over time.

Diachronic rationality is incredibly important, because we change our beliefs constantly in response to new *evidence*. One of the most important notions to understand belief change is the following...

The ***probability of d conditional on e***, written P(d|e), is defined as

$$P(d|e) = \frac{P(e \text{ and } d)}{P(e)}$$

Note that this is just a definition. It can't be "wrong".

<u>Example 1</u>
What is the probability that you'll role a 5 on a fair six sided die, conditional on your rolling an odd number?

<u>Example 2</u>
I flip a coin twice. What is the probability of it landing heads on the second flip, conditional on its landing heads on the first flip?

<u>Example 3</u>
Suppose 25% of the class aced the first exam, but only 10 percent of the class aced both the first and the second. What's the probability of a given student having aced the second exam, conditional on their having aced the first?

As I just noted, conditional probability is just a defined quantity. But we bother to define it because it seems like a good candidate to play a special role: that of helping us understand *diachronic rationality*.

The key case of rational belief change involves the acquisition of *evidence.* E.g.: You think it's unlikely to rain. Then you step outside and see dark clouds on the horizon. Now you think it's likely to rain. What's the relationship here?

More formally, suppose you start with some ***prior*** credence distribution given by P. Then you learn, conclusively, that E. It seems you should have a new credence distribution which we can call P_E. What *should* P_E be? Suggestion: the function always delivers the prior probabilities of events *conditional on* E.

> ***Rule of Conditionalization***: Given a credence distribution P, if a rational agent acquires conclusive evidence that E, then for any event e,
>
> $$P_E(e) = P(e|E)$$

Note: this is *not* a definition, and isn't even a triviality. It's a substantive claim about diachronic rationality.

Also note that this kind of rationality of belief is kind of like instrumental rationality in action. It tells you how to get from a rational output *holding fixed* (the rationality of?) an input.

Note a nice feature of conditionalization: if P obeys the probability axioms, then P_E will as well. So the constraint on diachronic rationality *preserves* synchronic rationality.