Chan and Khandani
Conditional Probability and Independence

Introduction

So far, we've gone over how to create sample spaces and how to find the probability of certain combinations of events within that same space.

But what about conditions? What's the probability of the entire human race being wiped out if a meteor hits the earth? How about the probability of the ratio in ECE being good if 5 times the amount of girls applied (this is a trick question, the latter's probability is zero)?

In this chapter, we'll learn what conditional probability is, and how to calculate it.

Conditional Probability

Formally, this is defined as

$$P(E|F) = \frac{P(E \cap F)}{P(F)}, \quad P(F) \neq 0$$

Translated, this means the probability of E occuring, if F occurs, is the intersection's probability divided by the probability of the dependent event.

Note that the events here are still within the same sample space - it doesn't make sense to calculate unrelated probabilities, for example, the probability of me stubbing my toe if an airplane has crossed my sights.

ex. So let's take a look at a simple example:

1	2	3
4	5	6

Again, we just roll a die once.

So what is the probability that we roll a 6, GIVEN THAT the roll is even?

Of course, we already know this is just $1/3$. But for the sake of practice, we'll use the previous formula.

$P(\text{even number}) = 1/2$ $\qquad$ $P(\text{even} \cap 6) = 1/6$

$P(6) = 1/6$

$$\therefore \quad P(6 \mid \text{even}) = \frac{P(\text{even} \cap 6)}{P(\text{even})} = \frac{1/6}{1/2} = \frac{1}{6} \cdot \frac{2}{1} = 1/3$$

We've essentially shrunken our sample space to just the even event: $\{2, 4, 6\}$, and calculated the probability of a 6 based on that.

The Multiplication Rule and Bayes' Formula

Sometimes, (and by sometimes I mean contrived algebraic proof questions) it's useful to deconstruct a certain probability into a linear combination of other probabilities. This is what these formulae are for.

$$P(E_1 \cap E_2 \cap \ldots \cap E_n) = P(E_1)\,P(E_2 \mid E_1)\,P(E_3 \mid E_2 E_1)\ldots$$
$$P(E_n \mid E_{n-1} \ldots E_1)$$

The multiplication rule gives you another way to calculate the probability of the intersection of multiple events

We can think about why this works in a qualitative way: remember that intersections are about satisfying every event involved. As such, for this overall probability not to equal zero at the end, all of these events must have shared elementary elements; that's where the conditional relationships come in...

ex. Again, a simple example with a die roll

$E_1 = \{1, 2\}$
$E_2 = \{2, 3\}$
$E_3 = \{2, 6\}$

All of these events only share the elementary event 2: as such we can confirm our final answer is $1/6$.

$$P(E_1) = 2/6 = 1/3$$
$$P(E_2 \mid E_1) = \frac{P\{2\}}{P\{2,3\}} = 1/2$$
$$P(E_3 \mid E_2 \cap E_1) = \frac{P\{2\}}{P\{2\}} = 1$$

$$\therefore P(E_1 E_2 E_3) = (1/3)(1/2)(1) = 1/6$$

So we continuously calculate probabilities of events based on the intersections of other events, leading to a product equalling the overall probability.

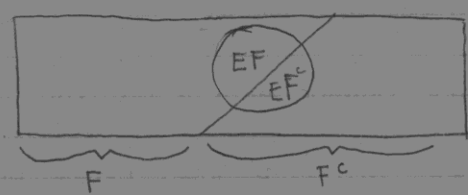
You could also just rearrange the original conditional probability equation and get the same thing

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

$$P(E \cap F) = P(E|F)\,P(F)$$

Bayes' formula also gives us a way to represent the probability of one event using two.

$$P(E) = P(E \cap F) + P(E \cap F^c)$$



$$= P(E|F)\,P(F) + P(E|F^c)\,P(F^c)$$

Notice that this is the exact same thing with the multiplication rule applied.

$$= P(E|F)\,P(F) + P(E|F^c)\,(1 - P(F))$$

And here, we've replaced $P(F^c)$ with $1 - P(F)$. Wow. Bayes' formula, everyone.

The Odds Ratio

What a great and meaningful name. The odds ratio states the ratio of an event happening versus it not happening.

It is defined as:

$$\frac{P(A)}{P(A^c)} = \frac{P(A)}{1-P(A)}$$

The odds ratio of a coin flip is 1. The odds ratio of a die roll is 1/5. It doesn't quite make sense in this context, but it makes much more sense if we expand this to a conditional event.

If we have some event A that is based on B:

$$\underbrace{P(A|B)}_{\text{A occurs}} = \frac{P(A \cap B)}{P(B)} \quad \text{and} \quad \underbrace{P(A^c|B)}_{A^c \text{ doesn't occur}} = \frac{P(A^c \cap B)}{P(B)}$$

$$= \frac{P(A)P(B|A)}{P(B)} \qquad\qquad = \frac{P(A^c)P(B|A^c)}{P(B)}$$

by multiplication rule

$\therefore$ the odds ratio is

$$\frac{P(A|B)}{P(A^c|B)} = \frac{\frac{P(A)P(B|A)}{P(B)}}{\frac{P(A^c)P(B|A^c)}{P(B)}}$$

$$= \left[\frac{P(A)P(B|A)}{\cancel{P(B)}}\right]\left[\frac{\cancel{P(B)}}{P(A^c)P(B|A^c)}\right]$$

$$= \underbrace{\frac{P(A)}{P(A^c)}}_{\text{A's odds ratio}} \; \underbrace{\frac{P(B|A)}{P(B|A^c)}}_{\text{new additional odds ratio}}$$

Now, this still probably doesn't make much sense, so let's use some real events.

A: a person has a heart attack (HA)
B: a person has high blood pressure (HBP)

So $$P(HA \mid HBP) = \frac{P(HA)}{P(HA^c)} \; \frac{P(HBP \mid HA)}{P(HBP \mid HA^c)}$$

odds ratio of a heart attack

odds ratio of having high blood pressure, with or without heart attack

Now, let's assign these some numbers:

$P(HA) = 35\%$ $\quad P(HBP \mid HA) = 90\%$ chance that a person has HBF if they have a HA

$P(HA^c) = 65\%$ $\quad P(HBP \mid HA^c) = 10\%$ chance of having HBF if they don't have a HA

$$\therefore \; P(HA \mid HBP) = \left(\frac{0.35}{0.65}\right)\left(\frac{0.9}{0.1}\right)$$
$$= 4.85$$

In statistics, since the OR > 1, this states that having high blood pressure raises the risk of heart attack.

Another Variant of Bayes' Formula

If we have some set of events $F_1, \ldots, F_n$, such that they partition the sample space,

$$P(F_j \mid E) = \frac{P(E \cap F_j)}{P(E)}$$
$$= \frac{P(E \mid F_j)P(F_j)}{\sum_{i=1}^{n} P(E \mid F_i)P(F_i)}$$

So all we've really done is just apply the multiplication rule again.

<u>Independent (aka Disjoint) Events</u>

As before, if two events are independent,

$$P(E \cap F) = P(E)P(F)$$

And as follows, three independent events would be defined as

$$P(EFG) = P(E)P(F)P(G)$$
$$P(EF) = P(E)P(F)$$
$$P(FG) = P(F)P(G)$$
$$P(EG) = P(E)P(G)$$

And if we want to stretch this to n events;

E_i, where $i = 1, \ldots, n$ is independent if and only if any collection of r indices $(\alpha_1, \ldots, \alpha_r)$ chosen from i creates

$$P(E_{\alpha_1} \cap \ldots \cap E_{\alpha_r}) = P(E_{\alpha_1}) \ldots P(E_{\alpha_r})$$

which is really just a fancy way of saying the probability of the intersections should be equal to the product of the probabilities.

Time for examples!

ex. In a batch of 50 products, 5 are defective. One is selected at random, and then one more from the remaining 49. Find the probability that both are defective.

This is fairly simple. The probability of the first being defective (A) is just 5/50.

$$P(A) = 5/50$$

Then, we have 4 defective units in the remaining batch, so the probability of the second being defective (B), is

$$P(B) = 4/49$$

Since these are disjoint (the first selected unit doesn't have an effect on which of the remaining we select)

$$P(A \cap B) = \frac{5}{50}\ \frac{4}{49} = \frac{2}{245}$$

If we then wanted to pick another, working unit (C), this would just become

$$P(ABC) = \frac{5}{50}\ \frac{4}{49}\ \frac{\overset{\text{(working}}{45}}{\underset{\uparrow\text{remaining}}{48}}$$

$$= \frac{3}{392}$$

ex. A flips 3 coins, and Betty 2. A wins if he has more heads than B. What's the probability A wins?

A:		B:	
H H H	- 3	H H	- 2
H H T	- 2	H T	- 1
H T H	- 2	T H	- 1
H T T	- 1	T T	- 0
T H H	- 2		
T H T	- 1		
T T H	- 1		
T T T	- 0		

So A has: $P(3) = 1/8$ B has: $P(2) = 1/4$

$P(2) = 3/8$ $P(1) = 2/4$

$P(1) = 3/8$ $P(0) = 1/4$

$P(0) = 1/8$

Assuming B flips first (doesn't really matter, honestly)

if B gets 0 $\rightarrow$ A has a $7/8$ chance to win
B gets 1 $\rightarrow$ A has $4/8$
B gets 2 $\rightarrow$ A has $1/8$

So if we define the event A as A winning, and B_i as the number of heads B gets,

$$P(A \mid B_0) = 7/8$$
$$P(A \mid B_1) = 4/8$$
$$P(A \mid B_2) = 1/8$$

Notice that AB_i, where B flips i heads but A wins regardless, partition the entire sample space of A, where A wins.

As such, we can simply add the probabilities of AB_i for all valid i to generate our final result.

But we don't actually have AB_i, just $A|B_i$. Thankfully the multiplication rule comes in handy.

$$\begin{aligned} P(AB_0) &= P(A)\,P(B_0|A) \\ &= P(B_0)\,P(A|B_0) \\ &= (1/4)(7/8) \\ &= 7/32 \end{aligned}$$

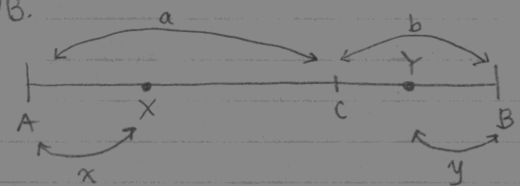
$$\begin{aligned} P(AB_1) &= P(B_1)\,P(A|B_1) \\ &= (1/2)(4/8) \\ &= 1/4 \end{aligned}$$

$$\begin{aligned} P(AB_2) &= P(B_2)\,P(A|B_2) \\ &= (1/4)(1/8) \\ &= 1/32 \end{aligned}$$

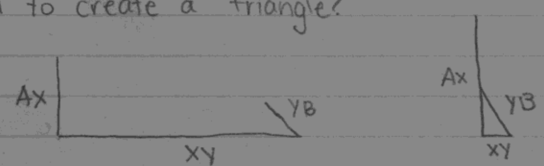
$$\begin{aligned} \therefore\ P(A) &= P(AB_0) + P(AB_1) + P(AB_2) \\ &= \frac{7}{32} + \frac{1}{4} + \frac{1}{32} \\ &= \frac{8}{32} + \frac{8}{32} \\ &= 1/2 \end{aligned}$$

A has a 50% chance of winning.

ex. Consider a line AB divided into two by some point C, where $AC > CB$. Suppose a random point X is chosen on AC, and Y on CB, where x and y are the distances from A/B.



What is the probability that AX, XY, and YB can fold to create a triangle?



The most important thing to realize here is that the triangle can only be formed if each line's length must be shorter than the sum of the other two:

Length of AX: x
XY: $(a+b-x-y)$
YB: y

$$\therefore \quad x < a+b-x-y+y$$
$$< a+b-x$$

$$(a+b-x-y) < x+y$$

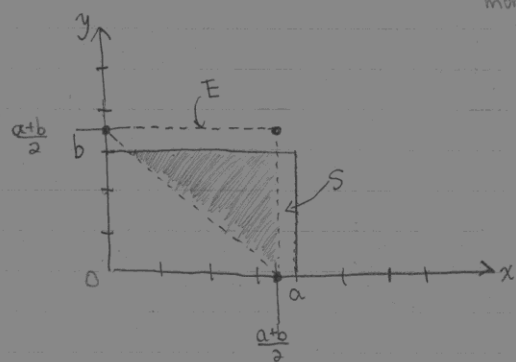
$$y < a+b-x-y+x$$
$$< a+b-y$$

So now, note that if any of x, y, or $a+b-x-y$ exceeds half of the length of AB, one of our conditions will always be violated

$$\therefore \quad x < \frac{a+b}{2} \qquad y < \frac{a+b}{2}$$

$$x + y > \frac{a+b}{2}$$

↑ if this is true, there's no way $a+b-x-y$ can exceed more than half as well



This rectangle now represents all possible lengths of a and b, effectively creating a graphical representation of our sample space

Then, since we've restricted x and y to be less than $\frac{a+b}{2}$, we can plot those points as well and construct our event E: a proper triangle can be made.

$\therefore$ $P(A \cap E)$ will give us the final answer.

$$\text{The area of } A \cap E = \tfrac{1}{2}(\text{base})(\text{height}) = \tfrac{1}{2} b^2$$

because this is an isosceles triangle: its base = height.

$$\therefore P(A \cap E) = \frac{\frac{1}{2}b^2}{ab} = \frac{b}{2a}$$

ex. Alright, last one. A gambler has 9 dice - 3 fair, 6 that have a 50% chance of rolling a 6.

A: 6 appears on the first roll
B: 6 appears on the second roll

One die is chosen at random and rolled twice: are these events independent?

Die picked $\xrightarrow{\text{fair}}$ rolls 2 6s: $P(AB \mid \text{fair})\,P(\text{fair})$ by mult. rule

Die picked $\xrightarrow{\text{loaded}}$ rolls 2 6s: $P(AB \mid \text{loaded})\,P(\text{loaded})$

$$P(AB \mid \text{fair})\,P(\text{fair}) = \left[\left(\frac{1}{6}\right)\left(\frac{1}{6}\right)\right]\left(\frac{3}{9}\right)$$
$$= 3/324$$
$$= 1/108$$

$$P(AB \mid \text{loaded})\,P(\text{loaded}) = \left[\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\right]\left(\frac{6}{9}\right)$$
$$= 6/36$$
$$= 18/108$$

Since $P(AB) = P(AB \mid \text{fair})\,P(\text{fair}) + P(AB \mid \text{loaded})\,P(\text{loaded})$
$$= 19/108$$

$P(A)\,P(B)$ must be equal to $P(AB)$ if the events are independent.

$$P(A) = P(A \mid \text{fair})\,P(\text{fair}) + P(A \mid \text{loaded})\,P(\text{loaded})$$
$$= \left(\frac{1}{6}\right)\left(\frac{3}{9}\right) + \left(\frac{1}{2}\right)\left(\frac{6}{9}\right)$$
$$= \frac{7}{18}$$

Of course, $P(A) = P(B)$ due to them being essentially the exact same event (roll a 6).

$$\therefore \; P(A)\,P(B) = \left(\frac{7}{18}\right)\left(\frac{7}{18}\right)$$
$$= \frac{49}{324}$$
$$\neq \frac{19}{108}$$
$$\neq P(AB)$$

and as such A and B are not independent.