Basic probability theory: Conditional probability

- What if a dice is biased so that it rolls 6 twice as often as every other number? How can we deal with 'uneven' *base rates*?
- Why should we care?
  - Because real life uses biased dice
  - eg. the *conditional probability* of being schizophrenic, given that a person has an appointment with a doctor who specializes in schizophrenia, is quite different from the *unconditional probability* that a person has schizophrenia (the *base rate*)

Conditional Probability:

The Atonement Lecture

Conditional probability

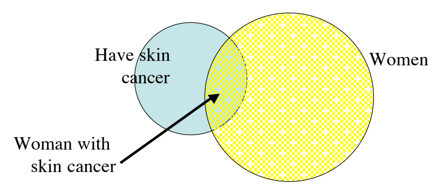
- Conditional probability arises when one probability P(A) depends on another probability P(B) *which is defined over the same population of events* (VERY IMPORTANT!)
- We say that we want to know P(A) *given* P(B); notationally, P(A|B)
- We can say that the word 'given' defines a subset of the population of events: namely that subset that depends on P(B)
- For example, if we care the incidence of skin cancer in woman only, we might want to know P(skin cancer | female).
- Intuitively, what this is saying is: first, pick out all the females, and second, figure out the probability they will get skin cancer

Conditional probability

- More formally, what P(A|B) says is: Pick out the events to which both P(A) and P(B) apply, and consider them as part of the subset of events to which only P(B) applies: hence

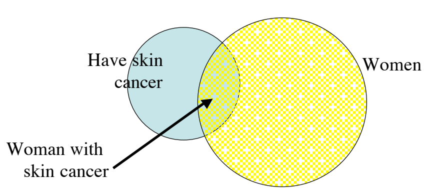
P(A|B) = P(A and B) | P(B)

= P(A,B) | P(B) [A notational change only]



Conditional probability

- What we are doing with P(A|B) = P(A and B) | P(B) is *selecting out the same subset of the event population in both the numerator and the denominator*: in this case, only women



Conditional probability

- What we are doing with P(A|B) = P(A and B) | P(B) is *selecting out the same subset of the event population in both the numerator and the denominator*: in this case, only women
- The reason it is so important to do this is because in some cases P(A) in the selected subset of an event population is very different from P(A) in the whole event population
- For example, P(Have healthcare | Canadian) is very different from P(Have healthcare)
  - Everyone in Canada has healthcare: P(Have healthcare | Canadian) = 1
  - Most people in the world do not have healthcare: P(Have healthcare) < 0.01

A generalization: Bayes' theorem

P(A|B) = P(B|A) P(A) / P(B)

• Some people find Bayes' Theorem helpful, since in some cases it can clarify the problem being considered
• Some people find it more confusing than helpful,
• The important point to understand is that Bayes' Theorem is just a re-statement of the definition of conditional probability, not a new finding

[Note: More complex versions of Bayes' Theorem are defined which deal with multiple possible outcomes; we don't use them in this class, in which we only consider two-possibility problems.]

A generalization: Bayes' theorem

P(A|B) = P(B|A) P(A) / P(B)

• Note that all this says is: P(A and B) = P(B|A) P(A)
• eg. The probability of having being a Canadian and having health care is the same as the probability of having health care given that you are Canadian X the probability of being Canadian
  • In this example we can easily see that this is true: Since all Canadians have healthcare, the probability of being a Canadian and having health care is the same as the probability that you are Canadian: P(Heath care | Canadian) = 1

A generalization: Bayes' theorem

P(A|B) = P(B|A) P(A) / P(B)

• Note that all this says is: P(A and B) = P(B|A) P(A)
• Another intuitive example: The probability of winning the lottery and buying a ticket is dependent both on the probability of winning the lottery when [= given that] you have a ticket and on the probability that you buy a ticket
• A person with a low probability of buying a lottery ticket has a lower probability of holding a winning ticket than a person that buys more tickets: even though P(Winning|Ticket)- the odds of any particular ticket being a winning ticket- is the same for both people

A generalization: Bayes' theorem

P(A|B) = P(B|A) P(A) / P(B)

Proof: By definition, (1.) P(A|B) = P(A,B) / P(B)
(2.) P(B|A) = P(A,B) / P(A)
(3.) P(B|A) P(A) = P(A,B) [Multiply (2.) by P(A)]
(4.) P(A|B) P(B) = P(B|A) P(A) [Substitute (1.) in (3.)]
(5.) P(A|B) = P(B|A) P(A) / P(B)

Why I am so confused (and how you can avoid it)

- In the last class, I crashed and burned on a simple question in conditional probability: P(Female | Second row)
- My error was a classic error in probabilistic reasoning, & therefore instructive: I accidentally changed event populations without noticing
- The relevant numbers are something like this:
  - # in class = 35
  - Females in class = 30
  - Number in row 2: 7
  - Males in row 2: 2

Why I am so confused (and how you can avoid it)

- The relevant numbers are something like this:
  # in class = 35 Females in class = 30
  Number in row 2 = 7 Males in row 2 = 2
- I began by noting that P (Female and Row 2) = 5/7. **This is not correct since *the proper population is the whole class***: it should be 5/35!
- With this step I changed the problem and just helped myself to the very answer I was supposed to be computing!
- I failed to correctly calculate because I changed the event population in my head from the whole class (as it should be) to Row 2 alone.
- Properly the solution I intended to show you was:
  P(F|2) = P(F &2) / P(2)
  = 5/35 / 7/35
  = 5/7