Chapter 2 Theory from Sections 2.4 and 2.5

A simple example is useful for understanding conditioning – after this I will survey the theory of section 2.4 and 2.5.

An bowl contains 4 White balls and 2 Black balls. One ball is drawn at random and its colour is called C1. It is not replaced. Then a second ball is drawn at random from the four that are left and its colour is called C2. The sample space for this experiment is{(W,W),(W,B),(B,W),(B,B)}. The probabilities associated with these outcomes are, respectively, (4/6)(3/5), (4/6)(2/5), (2/6)(4/5),(2/6)(1/5) which simplifies to 6/15,4/15,4/15,1/15. Suppose we want P(C2=W after C1=B is known), and we write this P(C2=W|C1=B). The condition forces us to look only at the last two possible outcomes, (B,W) and (B,B). Note that these are NOT equally likely outcomes, and in fact their probabilities are 4/15 and 1/15. So in this subset, (B,W) occurs 80% of the time and (B,B) occurs 20% of the time. We might alternatively do the calculation this way ...
P(C2=W|C1=B) = (4/15)/(4/15+1/15) = .80

Is there a formalism that will organize this kind of calculation in a more complex situation?

On p 77 we can use definition (2.3) to write

P(C2=W|C1=B) = P(C2=W AND C1=B)/P(C1=B) = (4/15)/P(C1=B)

But P(C1=B) = P({C1=B,C2=W} or {C1=B,C2=B}) = 4/15 + 1/15 = 1/3

So P(C2=W|C1=B) = (4/15)/(1/3) = .80

What if you wanted P(C1=B|C2=W)? Is it possible to go backwards in time? Yes!

P(C1=B|C2=W) = P(C1=B AND C2=W)/P(C2=W) = (4/15)/P(C2=W)

But, P(C2=W) = P({C2=W and C1=W}or {C2=W and C1=B}) = 6/15 + 4/15 =2/3

So, P(C1=B|C2=W) = (4/15)/(2/3)=.40

So we see that P(C2=W|C1=B) = .80 but P(C1=B|C2=W)=.40

Be careful with the interpretation. Can we conclude that W follows B more often than B follows W? No. They both occur with equal frequency (4/15). The conditional probabilities assume the condition has occurred already. The conditional probability does not say what happens unconditionally.

---

p 77 (2.3) defines P(A|B). A look at the definition should suggest why it is verbalized as the "probability of A given B is know to occur", or, more briefly, "probability of A given

B". Of all the LRRF associated with B, we want the proportion of LRRF that is also associated with A.

p 78 The "multiplication rule" is really just a re-expression of definition (2.3). But the reason it is called that is, *in the special case that* $P(A|B) = P(A)$, it is then true that $P(A \text{ and } B) = P(A)\ P(B)$

p 81. Since it seems so obviously true, we already used the fact that

$A = (A \cap B) \cup (A \cap B')$.
The outcomes in A are exactly those outcomes that occur with B, or else without B.
Note that $(A \cap B)$ and $(A \cap B')$ are mutually exclusive so that
$P(A) = P(A \cap B) + P(A \cap B')$
and using the defintion of conditional probability
$= P(A|B)P(B) + P(A|B')P(B')$
which essentially verifies (2.5). This is a very useful relationship. If you think about it long enough, you will see that it is intuitively obvious.

In the introductory paragraph for these notes, we computed something like $P(A|B)$ and then $P(B|A)$. Note that we could express one of these in terms of the other using

$P(B|A) = P(A \cap B)/P(A)$ and then using the same relationship again $= P(A|B)P(B)/P(A)$

This looks very handy but the catch is that $P(A)$ is not always readily available. We need to use (2.5) for it. When we do the result is Bayes Theorem – see p 82.

Section 2.5

The special case of the multiplication rule, mentioned above,

in the special case that $P(A|B) = P(A)$, it is then true that $P(A \text{ and } B) = P(A)\ P(B)$

is the case of *independence* of A and B.

Note that if A and B are independent, and B and C are independent, it is NOT necessarily true that A and C are independent. A counterexample emerges when C is the same event as A. This is why we need a definition of mutual independence p 89.

The next posting will include some worked examples that we will also cover in class.