S520

Chad Burrus

2010/01/22

1 5.3. Conditional Probability

1. Probability given some extra information that changes the experiment.
2. $P(A|B)$ = (conditional) probability of A given B

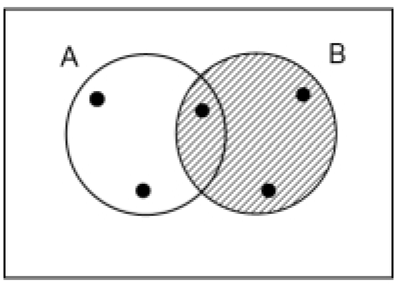

Figure 1: We restrict the possible values of A to those values also in B.

3. Basic idea : re-proportion the experiment so that B becomes the universe.
   (a) Replace S with B
   (b) Replace A with $A \cap B$
   (c) Therefore: $P(A|B) = \frac{P(A \cap B)}{P(B)}$
4. Multiplication Rule
   (a) $P(A \cap B) = P(B) \times P(A|B)$
   (b) $P(A \cap B) = P(A) \times P(B|A)$

2 Bayes Theorem Example Using Tree Diagrams

1. Bayes Theorem is confusing to look at and apply.
2. Tree diagrams:
   (a) Make the Bayes Theorem work out automatically, when done properly.
   (b) Require you to pay attention to what you're given when setting up the tree.
3. Hypertension Example (see Example 3.10 in the book for a similar problem)
   (a) Assume 20% of adults are hypertensive (HT)

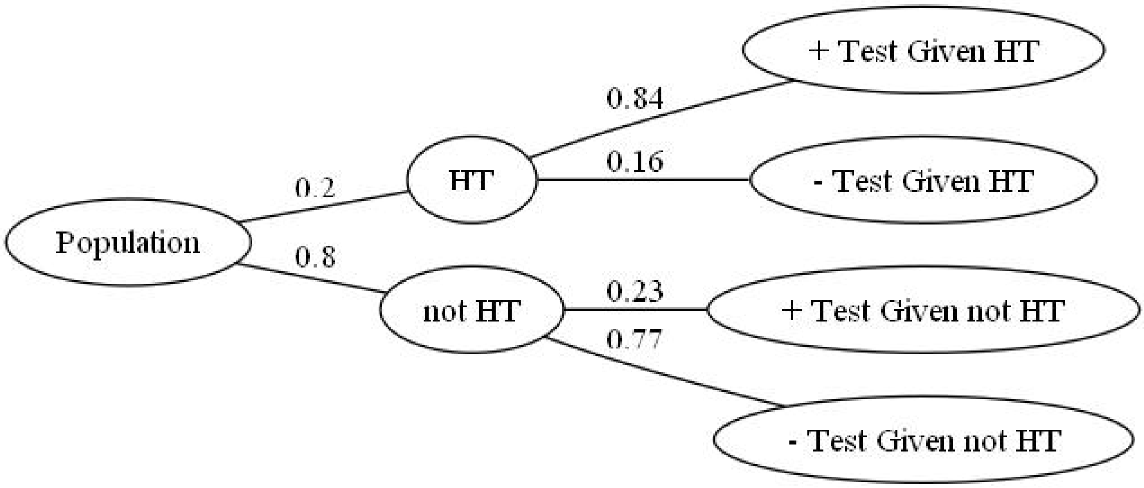

Figure 2: The HT tree diagram.

(b) Assume a testing machine diagnoses 84% of HT adults correctly and 23% of $notHT$ adults incorrectly as HT.

4. $P(HT \cap + \text{ test}) = P(HT) \times P(+\text{ test}|HT) = 0.2 \times 0.16 = \boxed{0.32}$
5. Aside: Know the Multiplication Rule, but if your tree diagram is set up right, it should just work out automatically.
6. Back to math:

$$\begin{aligned} P(HT \cap + \text{ test}) &= 0.2 \times 0.84 = 0.168 \\ P(HT \cap \text{- test}) &= 0.2 \times 0.16 = 0.320 \\ P(notHT \cap + \text{ test}) &= 0.8 \times 0.23 = 0.184 \\ P(notHT \cap \text{- test}) &= 0.8 \times 0.77 = 0.616 \end{aligned}$$

7. $P(+\text{ test})$ is now easy to calculate–add them up.

$$P(+\text{ test}) = P(HT \cap + \text{ test}) + P(notHT \cap + \text{ test}) = 0.168 + 0.184 = \boxed{0.352} \quad (1)$$

8. So what's the $P(HT|+\text{ test})$?

$$P(HT|+\text{ test}) = \frac{P(HT \cap + \text{ test})}{P(+\text{ test})} = \frac{0.168}{0.352} = 0.477 \approx 50\% \quad (2)$$

9. Lesson: When P(having a disease) is low, it's not always a good idea to have a test done, because the number of false positives outweighs the number of true positives. If the chance of having a disease is 1 in a million, the test is pretty inaccurate. If the chance is 1 in 10, that's a different story. (I.e., if your friend tells you to get a test, probably it's a bad idea. If your doctor does, maybe you should listen to him/her.)

3 Independence

1. Intuitive definition: If the conditioning has no effect, then we say that the events A and B are independent.

$$P(A|B) = P(A) \text{ and } P(B|A) = P(B) \quad (3)$$

2. The mathematical rule states that if A and B are independent, then:

$$P(A \cap B) = P(A) \times P(B) \tag{4}$$

3. There are two main uses of equation 4.

   (a) To check independence by checking the equation—proving something is independent—this is not the way we'll typically use it—probabilists mainly use it this way.

   (b) Assume independence (requires scientific knowledge), then we use the equation appropriately—statisticians typically use it this way.

4. Example: Decide whether or not two events A and B are independent. (Exercise 3.12 in the book)

   (a) Gender distribution in chess clubs.

$$\begin{aligned} S &= \text{US citizens} \\ A &= \text{a randomly selected person is in a chess club} \\ B &= \text{a randomly selected person is female} \end{aligned}$$

   Note: we're talking about patterns of occurrence, not free will—be careful! Compare :

$$\text{Very small!} \approx P(B|A) \text{ vs. } P(B) \approx 50\% \tag{5}$$

   Since $P(B|A) \neq P(B)$, A and B are not independent.

   i. if $P(B|A) = P(B)$, and I mean really equal, they are independent.

   ii. if they're close, the events are still dependent.