Chapter 1.4 Conditional Probability

Review

- What is conditional probability?
- Multiplication Law of Probability
- How to calculate probabilities using the two equations?

A Game: Monty Hall problem

The Monty Hall problem is a probability puzzle based on the American television game show ***Let’s make a deal.*** *Suppose you're on this show.*

You're given the choice of three doors. You know previously that,
behind one door, there is a car;
behind the other two doors, there is a goat behind each of the two doors.

You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

http://math.ucsd.edu/~crypto/Monty/monty.html

Conditional Probability

- **Conditional probabilities** reflect how the probability of an event can change if we know that some other event has occurred or is true.
- The probability that a random student will get an “A” on the midterm (event B) is different if we choose from students who studied (event A) or students who didn’t (event C).

The conditional probability of event A given event B is: (provided that P(B) ≠ 0)

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

Simple Example

Example: Suppose that P(A)=0.7, P(B)=0.3, P(A and B)=0.2

$P(A|B)=$ ____________?
$P(B|A)=$____________?

Probability axiom for conditional probability

Conditional probabilities range from 0 (no chance of the event) to 1 (the event has to happen).

For any event A and B,
$0 \le P(A|B) \le 1$

The conditional probability of the complete sample space must equal 1.

$P(B|B) = 1$
How about $P(\emptyset\ |B)=?$

For any positive k, If A1, A2,…,Ak are disjoint events, then the additive rule still applies.

If $A_i \cap A_j = \emptyset$ for any i and j, then
$P(A_1 \cup A_2 \cup \ldots \cup A_k|B)=P(A_1|B)$
$+P(A_2|B)+\ldots+P(A_k|B)$

General multiplication rule

The probability that any two events, A and B, both occur is

- $P(A \cap B) = P(A \text{ and } B) = P(B|A) \times P(A)$
- $P(A \cap B) = P(A \text{ and } B) = P(A|B) \times P(B)$

Example: Artificial pond with 10 male and 10 female frogs.

Probability of successively capturing at random two male frogs:
P(2 males) = P(first is male) P(second is male | first is male)
$= (10/20)(9/19) \approx 0.237$

With the first picked frog returned to the pond.
P(2 males)=P(first is male)P(second is male)=1/4.

Example : Draw Cards

From a deck of 52 playing cards, we randomly draw cards one-by-one and without replacement. What is the probability that the 3rd spade appears on the 6th draw?

- E1: We draw two spade in the first 5 draws.
- E2: We draw one spade in the 6th draw.
- Want to calculate

  $P(E2 \text{ and } E1)=P(E2|E1)\times P(E1)=11/47\times 2C13\times 3C39/5C52.$

Monty Hall problem Revisit
What is going to happen if switch?

There are three possible scenarios, each with equal probability (1/3):

1. The player picks goat number 1. The game host picks the other goat. Switching will win the car.
2. The player picks goat number 2. The game host picks the other goat. Switching will win the car.
3. The player picks the car. The game host picks either of the two goats. Switching will lose
4. $P(win|switch)=2/3$

To Switch is better than to stay

Assume that you choose door 3 and the game host open door 1.

Is it to your advantage to switch your choice?

P(car in door 2| host open door 1)=________________.

Consider the position when door 3 has been chosen and no door has been opened. The probability that the car is behind door 2, p(C2), is plainly 1/3, as it may equally well be in any of the three places.

The probability that the game host will open door 1, p(O1), is 1/2 since there is equal probability the car is behind door 1 (forcing the host to open door 2) or door 2 (forcing the host to open door 1) and if the car is behind neither door by the given assumptions the host opens one of them randomly.

But when the car is behind door 2, the game host will certainly open door 1, by the assumptions; that is, p(O1|C2) = 1. Hence the probability that the car is behind door 2 *given that the game host opens door 1 is*

$$P(C2|O1) = \frac{P(O1|C2)P(C2)}{P(O1)} = \frac{1 \times \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

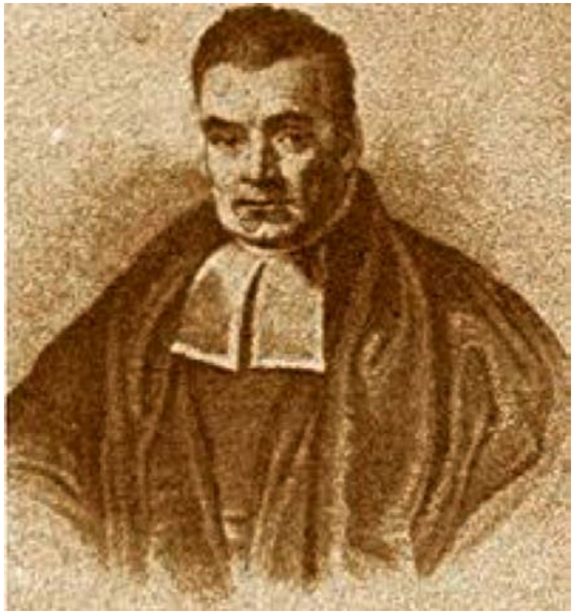
Chapter 1.6 Bayes' Theorem

Consider $P(A \cap B) = P(B \cap A)$. Then

$$P(B|A)P(A) = P(A|B)P(B)$$

Thus, we have

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



Sir. Thomas Bayes
(c. 1702 – 1761)

Bayes' Theorem

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(B) * P(A \mid B)}{P(B) * P(A \mid B) + P(B') * P(A \mid B')}$$

P(B) is called **prior probability**
P(B|A) is called **posterior probability**

Note: Bayes' theorem allows us to "invert" the conditioning.
For example, it might be easy to calculate,

P(later event | earlier event)

but we might only observe the later event and wish to deduce the probability that the earlier event occurred,

P(earlier event | later event)

Bayes' Theorem (general version)

Let B1, B2,…Bm be mutually exclusive and exhaustive events.

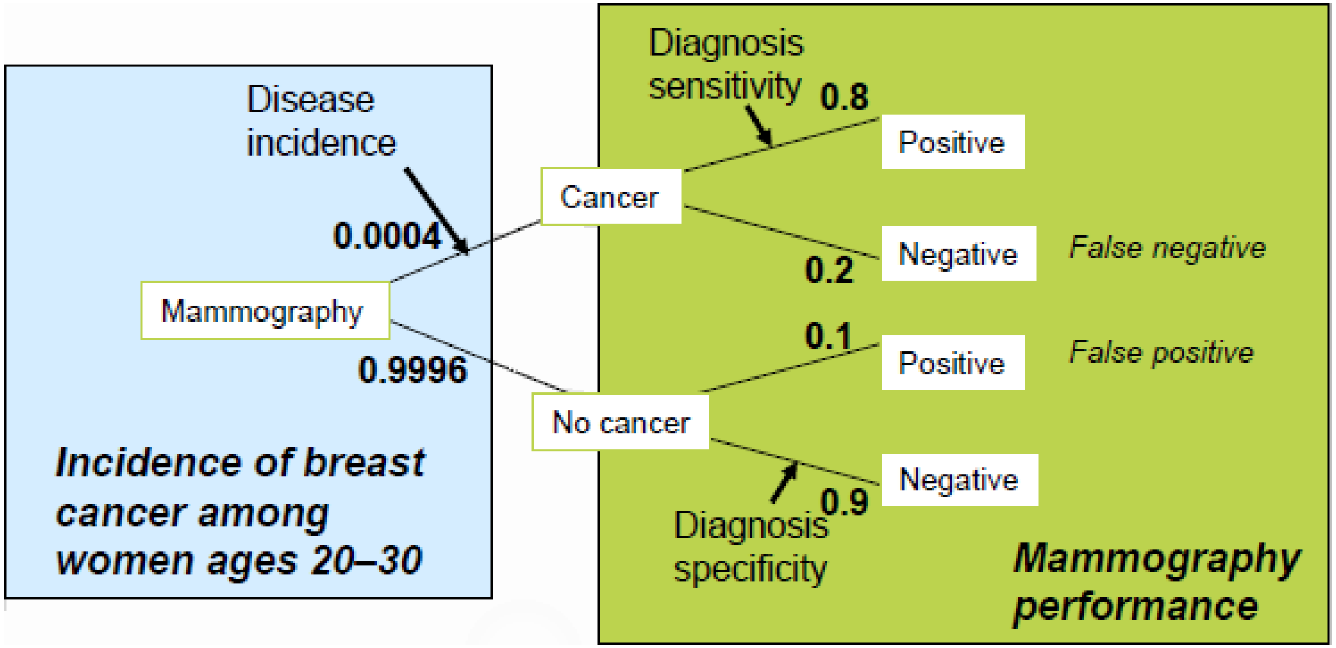
If A is an event, then we have

$$P(A)=P(B1\cap A)+P(B2\cap A)+\ldots+P(Bm\cap A)$$
$$=P(B1)*P(A|B1)+P(B2)*P(A|B2)+\ldots+P(Bm)*P(A|Bm)$$

Therefore,

$$P(B_k|A)=\frac{P(B_k)P(A|B_k)}{\Sigma_{i=1}^{m}P(B_i)P(A|B_i)},\ k=1,\cdots,m.$$

Example: Breast Cancer



- Possible outcomes given the positive diagnosis: positive test and breast cancer, or positive test but not cancer (false positive).
- P(cancer | test positive) = ______________________________?
- This value is called the positive predictive value, or PV+. It is an important piece of information but, unfortunately, is rarely communicated to patients.

The case of Sally Clark -- Revisit

Sally Clark, a British woman who was accused in 1998 of having killed her first child at 11 weeks of age, then conceived another child and allegedly killed it at 8 weeks of age.

According to the expert witness, for an affluent non-smoking family like the Clarks, the probability of a single cot death was 1 in 8,543, so the probability of two cot deaths in the same family was around "1 in 73 million" (8543×8543).

Sally Clark (1964 –2007).

The case of Sally Clark (Cont'd)

Formulate the events:

- *A= " Sally Clark's 2 babies died unexpectedly"*
- *B* =" Sally Clark is innocent "
- P(A | B)= P (*2 cot deaths in a family*) = *1 in 73 million (??)*
- P(B | A)= ___________?

- Would need the prior information of P(Sally Clark is innocent).

(Compare with the previous example of breast cancer.)

Example: Ketchup bottles

Ketchup bottles are produced in 3 different factories, accounting for 50%, 30%, and 20% of the total output respectively. The percentage of defective bottles from the 3 factories is respectively 0.*4%, 0.6% and 1.2%. A* statistic lecturer finds a defective bottle. What is the probability that it came from Factory 1?

1. Formulate the events:

Ai ="bottle comes from Factory i" (i=1,2,3)
B ="bottle is defective"

2. Information given:

$P(A1)$ =____, $P(A2)$ =____, $P(A3)$ =________,
$P(B|A1)$ =____, $P(B|A2)$ =____, $P(B|A3)$ =_____.

3. Looking for:

4. Applying Bayes' Theorem