Probability & Statistics

Dongjiao Ge

djge@cityu.edu.mo

Session 03

Chapter 2 Conditional probability

We will develop the technique of conditional probability to deal with cases where events are not independent.

2.1 What is conditional probability?

Alice and Bob are going out to dinner. They toss a fair coin 'best of three' to decide who pays: if there are more heads than tails in the three tosses then Alice pays, otherwise, Bob pays.

Clearly each has a 50% chance of paying. The sample space is

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\},$$

and the events 'Alice pays' and 'Bob pays' are respectively

$$A = \{HHH, HHT, HTH, THH\},$$
$$B = \{HTT, THT, TTH, TTT\}.$$

They toss the coin once and the result is heads; call this event E. How should we now reassess their chances? We have

$$E = \{HHH, HHT, HTH, HTT\},$$

and if we are given the information that the result of the first toss is heads, then E now becomes the sample space of the experiment, since the outcomes not in E are no longer possible. In the new experiment, the outcomes ‘Alice pays’ and ‘Bob pays’ are

$$A \cap E = \{HHH, HHT, HTH\},$$
$$B \cap E = \{HTT\}.$$

Thus the new probabilities that Alice and Bob pay for dinner are $\frac{3}{4}$ and $\frac{1}{4}$ respectively.

In general, suppose that we are given that an event E has *occurred*, and we want to compute the probability that another event A occurs. In general, we can no longer count, since the outcomes may not be equally likely. The correct definition is as follows.

Let E be an event with non-zero probability, and let A be any event. The ***conditional probability*** of A given E is defined as

$$P(A \mid E) = \frac{P(A \cap E)}{P(E)}.$$

- Remark that this is the definition.
- If you are asked for the definition of conditional probability, it is not enough to say "the probability of A given that E has occurred", although this is the best way to understand it. There is no reason why event E should occur before event A!
- **P(E) is not equal to zero**, since we have to divide by it, and this would make no sense if P(E) = 0.

To check the formula in our example:

$$P(A \mid E) = \frac{P(A \cap E)}{P(E)} = \frac{3/8}{1/2} = \frac{3}{4},$$
$$P(B \mid E) = \frac{P(B \cap E)}{P(E)} = \frac{1/8}{1/2} = \frac{1}{4}.$$

Example A random car is chosen among all those passing through Trafalgar Square on a certain day. The probability that the car is yellow is 3/100: the probability that the driver is blonde is 1/5; and the probability that the car is yellow and the driver is blonde is 1/50. Find the conditional probability that the driver is blonde given that the car is yellow.

Solution: If Y is the event 'the car is yellow' and B the event 'the driver is blonde', then we are given that $P(Y) = 0.03$, $P(B) = 0.2$, and $P(Y \cap B) = 0.02$. So

$$P(B \mid Y) = \frac{P(B \cap Y)}{P(Y)} = \frac{0.02}{0.03} = 0.667$$

There is a connection between conditional probability and independence:

Proposition 2.1 *Let A and B be events with $P(B) \neq 0$. Then A and B are independent if and only if $P(A \mid B) = P(A)$.*

This proposition is most likely what people have in mind when they say 'A and B are independent means that B has no effect on A'.

Proof The words 'if and only if' tell us that we have two jobs to do: we have to show that if A and B are independent, then $P(A \mid B) = P(A)$; and that if $P(A \mid B) = P(A)$, then A and B are independent.

So first suppose that A and B are independent. Remember that this means that $P(A \cap B) = P(A) \cdot P(B)$. Then

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A),$$

that is, $P(A \mid B) = P(A)$, as we had to prove.

Now suppose that $P(A \mid B) = P(A)$. In other words,

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

using the definition of conditional probability. Now clearing fractions gives

$$P(A \cap B) = P(A) \cdot P(B),$$

which is just what the statement 'A and B are independent' means.

2.2 Genetics

How genes code eye colour? Assuming only two colours of eyes.

Each person has two genes for eye colour. Each gene is either B or b. A child receives one gene from each of its parents. The gene it receives from its father is one of its father's two genes, each with probability 1/2; and similarly for its mother. The genes received from father and mother are independent. If your genes are BB or Bb or bB, you have brown eyes; if your genes are bb, you have blue eyes.

Example Suppose that John has brown eyes. So do both of John's parents. His sister has blue eyes. What is the probability that John's genes are BB?

Solution *John's sister has genes bb, so one b must have come from each parent. Thus each of John's parents is Bb or bB; we may assume Bb. So the possibilities for John are (writing the gene from his father first) BB,Bb,bB,bb each with probability 1/4. (For example, John gets his father's B gene with probability 1/2 and his mother's B gene with probability 1/2, and these are independent, so the probability that he gets BB is 1/4. Similarly for the other combinations.)*

Let X be the event 'John has BB genes' and Y the event 'John has brown eyes'. Then X = {BB} and Y = {BB,Bb,bB}. The question asks us to calculate $P(X \mid Y)$. *This is given by*

$$P(X \mid Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{1/4}{3/4} = 1/3.$$

2.3 The Theorem of Total Probability

Sometimes we are faced with a situation where we do not know the probability of an event B, but we know what its probability would be if we were sure that some other event had occurred.

Example An ice cream seller has to decide whether to order more stock for the Bank Holiday weekend. He estimates that, if the weather is sunny, he has a 90% chance of selling all his stock; if it is cloudy, his chance is 60%; and if it rains, his chance is only 20%. According to the weather forecast, the probability of sunshine is 30%, the probability of cloud is 45%, and the probability of rain is 25%. (We assume that these are all the possible outcomes, so that their probabilities must add up to 100%.) What is the overall probability that the salesman will sell all his stock?

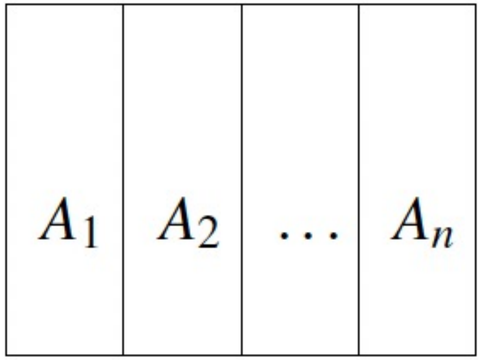
This problem is answered by the Theorem of Total Probability.

Definition of the partition.
The events $A_1, A_2, \ldots, A_n$ form a ***partition*** of the sample space if the following two conditions hold

(a) the events are pairwise disjoint, that is, $A_i \cap A_j = \emptyset$ for any pair of events A_i and A_j;

(b) $A_1 \cup A_2 \cup \cdots \cup A_n = \mathcal{S}$.

Another way of saying the same thing is that every outcome in the sample space lies in exactly one of the events $A_1, A_2, \ldots, A_n$. The picture shows the idea of a partition.



Theorem 2.2 *Let $A_1, A_2, \ldots, A_n$ form a partition of the sample space with $P(A_i) \neq 0$ for all i, and let B be any event. Then*

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i) \cdot P(A_i).$$

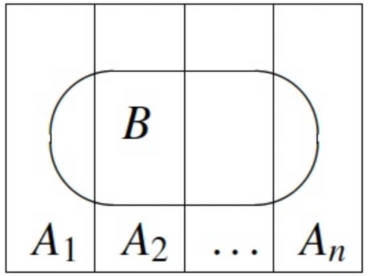
Proof By definition, $P(B \mid A_i) = P(B \cap A_i)/P(A_i)$. Multiplying up, we find that

$$P(B \cap A_i) = P(B \mid A_i) \cdot P(A_i).$$

Now consider the events $B \cap A_1, B \cap A_2, \ldots, B \cap A_n$. These events are pairwise disjoint; for any outcome lying in both $B \cap A_i$ and $B \cap A_j$ would lie in both A_i and A_j, and by assumption there are no such outcomes. Moreover, the union of all these events is B, since every outcome lies in one of the A_i. So, by Axiom 3, we conclude that

$$\sum_{i=1}^{n} P(B \cap A_i) = P(B).$$

Substituting our expression for $P(B \cap A_i)$ gives the result.



Back to the ice cream salesman example: Let A_1 be the event 'it is sunny', A_2 the event 'it is cloudy', and A_3 the event 'it is rainy'. Then A_1, A_2 and A_3 form a partition of the sample space, and we are given that

$$P(A_1) = 0.3, \quad P(A_2) = 0.45, \quad P(A_3) = 0.25.$$

Let B be the event 'the salesman sells all his stock'. The other information we are given is that

$$P(B \mid A_1) = 0.9, \quad P(B \mid A_2) = 0.6, \quad P(B \mid A_3) = 0.2.$$

By the Theorem of Total Probability,

$$P(B) = (0.9 \times 0.3) + (0.6 \times 0.45) + (0.2 \times 0.25) = 0.59.$$

One special case of the Theorem of Total Probability is very commonly used, and is worth stating in its own right. For any event A, the events A and A' form a partition of S. To say that both A and A' have non-zero probability is just to say that $P(A) \neq 0, 1$. Thus we have the following corollary:

Corollary 2.3 *Let A and B be events, and suppose that* $P(A) \neq 0, 1$. *Then*

$$P(B) = P(B \mid A) \cdot P(A) + P(B \mid A') \cdot P(A').$$

2.4 Sampling revisited

We can use the notion of conditional probability to treat sampling problems involving ordered samples.

Example I have two red pens, one green pen, and one blue pen. I select two pens without replacement.
(a) What is the probability that the first pen chosen is red?
(b) What is the probability that the second pen chosen is red?

For the first pen, there are four pens of which two are red, so the chance of selecting a red pen is $2/4 = 1/2$.

For the second pen, we must separate cases. Let A_1 be the event 'first pen red', A_2 the event 'first pen green' and A3 the event 'first pen blue'. Then $P(A_1) = 1/2, P(A_2) = P(A_3) = 1/4$ (arguing as above). Let B be the event 'second pen red'.

If the first pen is red, then only one of the three remaining pens is red, so that $P(B \mid A_1) = 1/3$. On the other hand, if the first pen is green or blue, then two of the remaining pens are red, so $P(B \mid A_2) = P(B \mid A_3) = 2/3$.

By the Theorem of Total Probability,

$$\begin{aligned} P(B) &= P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + P(B \mid A_3)P(A_3) \\ &= (1/3) \times (1/2) + (2/3) \times (1/4) + (2/3) \times (1/4) \\ &= 1/2. \end{aligned}$$

Q1 (15 minutes)Toss a coin 10 times. If you know (a) that exactly 7 Heads are tossed, (b) that at least 7 Heads are tossed, what are the probabilities that your first toss is Heads given (a) and (b), respectively?

2.5 Bayes' Theorem

Obviously, $P(A \mid B)$ and $P(B \mid A)$ are different.

Suppose that a new test is developed to identify people who are liable to suffer from some genetic disease in later life.

No test is perfect --- there will be some carriers of the defective gene who test negative, and some non-carriers who test positive.
e.g., let A be the event '*the patient is a carrier*', and B the event '*the test result is positive*'.

The scientists are concerned with the probabilities that the test result is wrong,
i.e., $P(B \mid A')$ and $P(B' \mid A)$.
However, *a patient* who has taken the test has different concerns.

- *If I tested positive, what is the chance that I have the disease? i.e.,* $P(A \mid B)$
- *If I tested negative, how sure can I be that I am not a carrier? i.e.,* $P(A' \mid B')$.

These conditional probabilities are related --- by Bayes' Theorem

Bayes' Theorem

Theorem 2.4 *Let A and B be events with non-zero probability. Then*

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}.$$

Proof.

$$P(A \mid B) \cdot P(B) = P(A \cap B) = P(B \mid A) \cdot P(A),$$

using the definition of conditional probability twice. (Note that we need both A and B to have non-zero probability here.) Now divide this equation by $P(B)$ to get the result.

If $P(A) \neq 0$ and $P(B) \neq 0$ we use corollary 2.3 to present $P(B)$, we will get another form of Bayes' theorem

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B \mid A) \cdot P(A) + P(B \mid A') \cdot P(A')}.$$

Remark

Bayes' Theorem is often stated in this form.

Example Consider the ice cream salesman from Section 2.3. Given that he sold all his stock of ice cream, what is the probability that the weather was sunny?

A_1 *--- the event 'it is sunny'*
B *--- the event 'the salesman sells all his stock'.*

We are asked to compute $P(A_1 \mid B)$. *We were given that* $P(B \mid A_1) = 0.9$ *and that* $P(A_1) = 0.3$, *and we calculated that* $P(B) = 0.59$. *So by Bayes' Theorem*

$$P(A_1 \mid B) = \frac{P(B \mid A_1)P(A_1)}{P(B)} = \frac{0.9 \times 0.3}{0.59} = 0.46$$

Example Consider the clinical test described at the start of this section. Suppose that 1 in 1000 of the population is a carrier of the disease. Suppose also that the probability that a carrier tests negative is 1%, while the probability that a noncarrier tests positive is 5%. (A test achieving these values would be regarded as very successful.)

Let A be the event 'the patient is a carrier', and B the event 'the test result is positive'. We are given that $P(A) = 0.001$ (so that $P(A') = 0.999$), and that

$$P(B \mid A) = 0.99, \qquad P(B \mid A') = 0.05.$$

(a) A patient has just had a positive test result. What is the probability that the patient is a carrier?

(b) A patient has just had a negative test result. What is the probability that the patient is a carrier?

(a) A patient has just had a positive test result. What is the probability that the patient is a carrier?

This is to calculate $P(A|B)$. Using Bayes' theorem, we have

$$\begin{aligned} P(A \mid B) &= \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A')P(A')} \\ &= \frac{0.99 \times 0.001}{(0.99 \times 0.001) + (0.05 \times 0.999)} \\ &= \frac{0.00099}{0.05094} = 0.0194. \end{aligned}$$

(b) A patient has just had a negative test result. What is the probability that the patient is a carrier?

It is to compute the probability $P(A|B')$. Using Bayes' theorem, we have

$$\begin{aligned} P(A \mid B') &= \frac{P(B' \mid A)P(A)}{P(B' \mid A)P(A) + P(B' \mid A')P(A')} \\ &= \frac{0.01 \times 0.001}{(0.01 \times 0.001) + (0.95 \times 0.999)} \\ &= \frac{0.00001}{0.94095} = 0.00001. \end{aligned}$$

Note that these calculations assume that the patient has been ***<u>selected at random from the population</u>***. If the patient has a family history of the disease, the calculations would be quite different.

Q2 (15 minutes) 2% of the population have a certain blood disease in a serious form; 10% have it in a mild form; and 88% don't have it at all. A new blood test is developed; the probability of testing positive is 9/10 if the subject has the serious form, 6/10 if the subject has the mild form, and 1/10 if the subject doesn't have the disease. I have just tested positive. What is the probability that I have this serious form of the disease?

2.6 Iterated conditional probability

The conditional probability of C, given that both A and B have occurred, is just $P(C \mid A \cap \mathrm{B})$. Sometimes instead we just write $P(C \mid A, B)$. It is given by

$$P(C \mid A,B) = \frac{P(C \cap A \cap B)}{P(A \cap B)},$$

So,

$$P(A \cap B \cap C) = P(C \mid A,B)P(A \cap B).$$

By the definition of conditional probability we also have

$$P(A \cap B) = P(B \mid A)P(A),$$

so assuming that $P(A \cap \mathrm{B}) \neq 0$, we have

$$P(A \cap B \cap C) = P(C \mid A,B)P(B \mid A)P(A).$$

This can be generalized to any number of events.

Proposition 2.5 *Let $A_1, \ldots, A_n$ be events. Suppose that $P(A_1 \cap \cdots \cap A_{n-1}) \neq 0$. Then*

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_n \mid A_1, \ldots, A_{n-1}) \cdots P(A_2 \mid A_1) P(A_1).$$

We will apply Proposition 2.5 to the *birthday paradox* stated as follows:

If there are 23 or more people in a room, then the chances are better than even that two of them have the same birthday.

To simplify the analysis, we ignore 29 February, and assume that the other 365 days are all equally likely as birthdays of a random person.

Suppose that we have n people $p_1, p_2, \ldots, p_n$.
A_2 --- the event 'p_2 has a different birthday from p_1'.

$P(A_2) = 1 - 1/365$, since whatever p_1's birthday is, there is a 1 in 365 chance that p_2 will have the same birthday.

A_3 --- the event 'p_3 has a different birthday from p_1 and p_2'.

It is not straightforward to evaluate $P(A_3)$, since we have to consider whether p_1 and p_2 have the same birthday or not. But we can calculate that $P(A_3 | A_2) = 1 - 2/365$, since if A_2 occurs then p_1 and p_2 have birthdays on different days, and A_3 will occur only if p_3's birthday is on neither of these days. So

$$P(A_2 \cap A_3) = P(A_2)P(A_3 \mid A_2) = (1 - \tfrac{1}{365})(1 - \tfrac{2}{365}).$$ *(By the definition of conditional probability)*

$A_2 \cap A_3$ --- *the event that all three people have birthdays on different days.*

Now we extend this process.
A_i --- the event 'p_i's birthday is not on the same day as any of $p_1, \ldots, p_{i-1}$', then

$$P(A_i \mid A_1, \ldots, A_{i-1}) = 1 - \tfrac{i-1}{365},$$

Using Proposition 2.5, we have

$$P(A_1 \cap \cdots \cap A_i) = (1 - \tfrac{1}{365})(1 - \tfrac{2}{365}) \cdots (1 - \tfrac{i-1}{365}).$$

We use q_i to present the above probability value; it is the probability that all of the people $p_1, \ldots, p_i$ have their birthdays on different days.

The numbers q_i decrease, since at each step we multiply by a factor less than 1.
So there will be some value of n such that

$$q_{n-1} > 0.5, \qquad q_n \leq 0.5,$$

n --- the smallest number of people for which the probability that they all have different birthdays is less than 1/2, that is, the probability of at least one coincidence is greater than 1/2.

By calculation, we find that $q_{22} = 0.5243$, $q_{23} = 0.4927$; so 23 people are enough for the probability of coincidence to be greater than 1/2.

As we mentioned that $P(A_3)$ is not that obvious to obtain, so now we return to answering the question “What is $P(A_3)$”? (This is the event that p_3 has a different birthday from both p_1 and p_2.) We have two cases:

- If p_1 and p_2 have different birthdays, the probability $P(A_3 \mid A_2)$ is 1− 2/365 : this is the calculation we already did.
- If p_1 and p_2 have the same birthday, then the probability $P(A_3 \mid A_2{}')$ is 1− 1/365.

So, by the Theorem of Total Probability,

$$
\begin{aligned}
P(A_3) &= P(A_3 \mid A_2)P(A_2) + P(A_3 \mid A_2')P(A_2') \\
&= \left(1-\tfrac{2}{365}\right)\left(1-\tfrac{1}{365}\right) + \left(1-\tfrac{1}{365}\right)\tfrac{1}{365} \\
&= 0.9945
\end{aligned}
$$

Problem How many people would you need to pick at random to ensure that the chance of two of them being born in the same month are better than even?

Assuming all months equally likely, if B_i is the event that p_i is born in a different month from any of $p_1, \ldots, p_{i-1}$, then as before we find that

$$P(B_i \mid B_1, \cdots, B_{i-1}) = 1 - \tfrac{i-1}{12},$$

We have 12 months a year.

We can also get

$$P(B_1 \cap \cdots \cap B_i) = (1 - \tfrac{1}{12})(1 - \tfrac{2}{12})(1 - \tfrac{i-1}{12}).$$

When $i = 4$, we calculate that this probability is

$$(11/12) \times (10/12) \times (9/12) = 0.5729$$

When $i = 5$, we calculate that this probability is

$$(11/12) \times (10/12) \times (9/12) \times (8/12) = 0.3819$$

So, with five people, it is more likely that two will have the same birth month.

Some years ago, in a probability class with only **ten** students, the lecturer started discussing the **Birthday Paradox**.

He said to the class, “I bet that no two people in the room have the same birthday”. He should have been on safe ground, since $q_{11} = 0.859$. (Remember that there are eleven people in the room!)

However, a student in the back said “I’ll take the bet”, and after a moment all the other students realised that the lecturer would **certainly lose his wager**.

The wining probability looks large, why the lecturer still lose?

---------- The **expectation** (see next session)

Exercises

EX1 Toss two fair coins, blindfolded. Somebody tells you that you tossed at least one Heads. What is the probability that both tosses are Heads?

EX2 Flip a fair coin. If you toss Heads, roll 1 die. If you toss Tails, roll 2 dice. Compute the probability that you roll exactly one 6.

EX3 A factory has three machines, M_1, M_2 and M_3, that produce items (say, lightbulbs). It is impossible to tell which machine produced a particular item, but some are defective. Here are the known numbers:

machine	proportion of items made	prob. any made item is defective
M_1	0.2	0.001
M_2	0.3	0.002
M_3	0.5	0.003

You pick an item, test it, and find it is defective. What is the probability that it was made by machine M_2?

EX4 An insurance company examines its pool of auto insurance customers and gathers the following information:

(i) All customers insure at least one car.
(ii) 70% of the customers insure more than one car.
(iii) 20% of the customers insure a sports car.
(iv) Of those customers who insure more than one car, 15% insure a sports car.

Calculate the probability that a randomly selected customer insures exactly one car and that car is not a sports car.

EX5 The Land of Nod lies in the monsoon zone, and has just two seasons, Wet and Dry. TheWet season lasts for 1/3 of the year, and the Dry season for 2/3 of the year. During the Wet season, the probability that it is raining is 3/4; during the Dry season, the probability that it is raining is 1/6.

(a) I visit the capital city, Oneirabad, on a random day of the year. What is the probability that it is raining when I arrive?

(b) I visit Oneirabad on a random day, and it is raining when I arrive. Given this information, what is the probability that my visit is during the Wet season?

(c) I visit Oneirabad on a random day, and it is raining when I arrive. Given this information, what is the probability that it will be raining when I return to Oneirabad in a year's time?

EX6 An actuary is studying the prevalence of three health risk factors, denoted by A, B, and C, within a population of women. For each of the three factors, the probability is 0.1 that a woman in the population has only this risk factor (and no others). For any two of the three factors, the probability is 0.12 that she has exactly these two risk factors (but not the other). The probability that a woman has all three risk factors, given that she has A and B, is 1/3. Calculate the probability that a woman has none of the three risk factors, given that she does not have risk factor A.

Ex7 Many casinos allow you to bet even money on the following game. Two dice are rolled and the sum S is observed.

If $S \in \{7,11\}$, you win immediately. If $S \in \{2,3,12\}$, you lose immediately.

If $S \in \{4, 5, 6, 8, 9,10\}$, the pair of dice is rolled repeatedly until one of the following happens:

- S repeats, in which case you win.
- 7 appears, in which case you lose.

What is the winning probability?