Lecture 3

Conditional Probability

By
Dr Sean Maudsley-Barton and Abdul Ali

Manchester Metropolitan University

Aims

- Understand the concept of conditional probability.
- Understand the relationship between conditional probability and independence.
- Calculate conditional probabilities for simple examples.

Conditional Probability

To motivate the concept of a conditional probability, consider the following simple example.

Example 3.1 (***Insurance Claims***)

An insurance broker is trawling through a sample of records and comes up with the following summary figures on claims in the previous 12 months for analysis.

- A figure the broker would be interested in is the **claim rate** in order that they may set premiums for the next year.

	Age		
	Under 25	25 and over	**Total**
No Claim	225	725	**950**
Claim	25	25	**50**
Total	**250**	**750**	1000

Conditional Probability

Overall, the claim rate is,

$$P(\text{Claim}) = \frac{50}{1000}$$
$$= 0.05$$

	Age		
	Under 25	25 and over	**Total**
No Claim	225	725	**950**
Claim	25	25	**50**
Total	**250**	**750**	1000

i.e. about 1 in 20 drivers might make a claim each year.

- However, this figure **hides a substantial difference** in the claim rates for young and old drivers.

Conditional Probability

- If we consider the 250 younger drivers separately we have,

$$P(\text{Claim}|\text{Under 25}) = \frac{25}{250}$$
$$= 0.10$$

whereas for the 750 older drivers we have,

$$P(\text{Claim}|\text{25 and over}) = \frac{25}{750}$$
$$= 0.03$$

	Age		
	Under 25	25 and over	Total
No Claim	225	725	950
Claim	25	25	50
Total	250	750	1000

Conditional Probability

Note: The notation | is read as ”**given that**”, i.e. a conditional statement, and these are **conditional probabilities**.

- The conditional probabilities show that the ratio of claim rates for younger drivers compared to older drivers is **0.10/0.03 = 3.3** times greater than that for older drivers.
- The value 3.3, in this case, is often referred to as a **relative risk** (**RR**) and is frequently quoted in the context of medical studies.

Conditional Probability

Example 3.2 (*Smoking and Lung Cancer*)

As another example, consider the following data from a study on male lung cancer patients carried out in 1950 in the UK, one of the earliest applications of epidemiology - the use of statistics to study disease patterns in populations.

	Smoker		
	No	Yes	**Total**
Lung Cancer	2	647	**649**
No Cancer	27	620	**647**
Total	**29**	**1267**	1296

Following the same procedure as before for the insurance data, we would calculate the relative risk for a smoker compared to a non-smoker as,

Conditional Probability

	Smoker		
	No	Yes	Total
Lung Cancer	2	647	**649**
No Cancer	27	620	**647**
Total	**29**	**1267**	1296

$$RR = \frac{647/1267}{2/29}$$

$$= 7.40$$

which seems to me a good enough reason to stop smoking!

Conditional Probability

Definition 3.1 (*Conditional Probability*)

The formal definition of the conditional probability of two arbitrary events A and B, say, is,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad \text{for } P(B) > 0$$

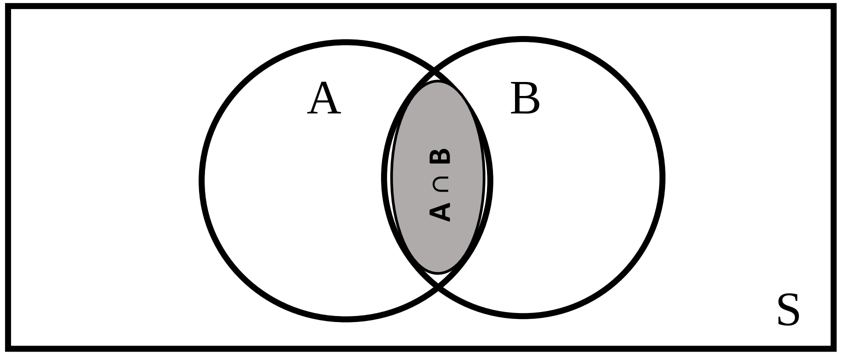
- where, as mentioned earlier, **|** we read as ”**given that**”.

- Thus the conditional probability $P(A|B)$ is the probability of A occurring when we know that B has already occurred.

Conditional Probability

Definition 3.1 (*Conditional Probability*)

- If we represent the events A and B in a Venn diagram, we can imagine their probabilities being represented by areas.
- In the following diagram the conditional probability, $P(A|B)$ is the proportion of the B set which is also in A and so is the $A \cap B$ area divided by the B area.



Conditional Probability and Kolmogorov's Axioms

It is straightforward to show that conditional probabilities satisfy Kolmogorov's axioms presented in chapter 2.

Theorem 3.1 *The conditional probability **P(A | B)** satisfies the first 3 Kolomogorov axioms* (*and hence other associated probability results*).

Proof:

- $P(A \mid B) \geq 0$, since,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \geq 0, \quad \text{since } P(A \cap B) \geq 0 \text{ and } P(B) \geq 0$$

- $P(S \mid B) = 1$ since,

$$P(S \mid B) = \frac{P(S \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

Conditional Probability and Kolmogorov's Axioms

- If $A_1 \cap A_2 = \emptyset$, then $P(A_1 \cup A_2 | B) = P(A_1 | B) + P(A_2 | B)$, since

$$P(A_1 \cup A_2 | B) = \frac{P((A_1 \cup A_2) \cap B)}{\mathrm{P}(B)}$$

$$= \frac{P((A_1 \cap B) \cup (A_2 \cap B))}{P(B)}$$

$$= \frac{P((A_1 \cap B)}{P(B)} + \frac{P((A_2 \cap B)}{P(B)}$$ (since $P(A_1 \cap B) \cap P(A_2 \cap B) = \emptyset$)

$$= P(A_1 | B) + P(A_2 | B)$$

Remember: The order in which we carry out the conditioning is crucial,

- both in terms of the **value** of the result and
- in terms of the **meaning** of the calculated probability.

The following example shows this:

Conditional Probability

	Age		
	Under 25	25 and over	**Total**
No Claim	225	725	**950**
Claim	25	25	**50**
Total	**250**	**750**	1000

Example 3.3 (***Insurance Claims - again***)

- Recall that,

P(Claim given driver under 25) = P(Claim | Under 25) = $\frac{25}{250}$ = 0.10, as before.

- We could, of course condition the other way round.
  - **Suppose we want to find P(Under 25 | Claimed),**

    {i.e., given that the person has made a claim, what is the probability they are under 25.}

- The condition tells us that we are dealing with one of the 50 policyholders in the second row. Of these, the number under 25 is 25 so that,

P(Claim given driver under 25) = P(Under 25| Claim) = $\frac{25}{50}$ = 0.5

- This example shows that, in general,

$$P(A \mid B) \neq P(B \mid A)$$

Conditional Probability

Example 3.3 (***Insurance Claims - again***)

Consider the same problem with the table displaying actual probabilities rather than frequencies. The values are in table,

	Age		
	Under 25	25 and over	**Total**
No Claim	0.225	0.725	**0.950**
Claim	0.025	0.025	**0.05**
Total	**0.250**	**0.750**	1.00

- To repeat the calculations above using the conditional probability formula,
- We need to identify the three probabilities,
  1. P(Claim and Under 25)=0.025,
  2. P(Under 25)=0.250 and
  3. P(Claim)=0.05.

Conditional Probability

Example 3.3 (***Insurance Claims - again***)

	Age		
	Under 25	25 and over	**Total**
No Claim	0.225	0.725	**0.950**
Claim	0.025	0.025	**0.05**
Total	**0.250**	**0.750**	1.00

$$P(\text{Claim} \mid \text{Under 25}) = \frac{P(\text{Claim and Under 25})}{P(\text{Under 25})}$$

$$= \frac{0.025}{0.250} = 0.10\text{, as before and}$$

$$P(\text{Under 25} \mid \text{Claim}) = \frac{P(\text{Claim and Under 25})}{P(\text{Claim})}$$

$$= \frac{0.025}{0.05} = 0.5\text{ again, as before.}$$

Note: It should be clear from this example that **P(A|B) is not necessarily the same as P(B|A)**, you need to be very careful about what event is being conditioned on.

Independence and Conditional Probability

The following theorem establishes a link between conditional probability and independent events.

Theorem 3.2 (*Independence and Conditional Probability*)

- Two events, A and B say, are independent if **P(A|B)=P(A)** or **P(B|A)=P(B)**, i.e. knowledge that one event has occurred does not affect the probability of the other event occurring.
- We also find by re-arranging the above formula that,

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

so that when events are independent, we find,

$$P(A \cap B) = P(A) \times P(B)$$

As given definition 2.1, (Chapter 2).

Independence and Conditional Probability

Example 3.4 (***Insurance Claims - again***)

- The data in the Example 3.1 is a good example to illustrate the connection between independence and conditional probabilities.
- It should be fairly clear that whether a policyholder makes a claim is very much dependent on their age.
- As we have already established that the RR is about 3.3 greater for younger drivers compared to older .
- In particular, we have

$$P(\text{Claim} \mid \text{Under 25}) = 0.10 \neq P(\text{Claim}) = 0.05$$

and similarly for P(Claim|25 and over).

Independence and Conditional Probability

Example 3.4 (***Insurance Claims - again***)

- Two questions to consider:
  - In a 2×2 table like the one for insurance claims, what pattern would you need to observe in the data to imply the independence of the events considered above?
  - What would this imply about the relative risk?

Independence and Conditional Probability

Manchester Metropolitan University

- Conditional probabilities are sometimes **counter-intuitive**. The following is often cited as an example.

Example 3.5 (***Birth Distributions***)

Consider the set of two children families and assume that all birth sequences i.e.{bb, bg, gb, gg} are equally likely. What is the probability that both children are boys given that at least one is a boy?

The usual (incorrect) answer given is 0.25. The correct answer is derived as follows:

Independence and Conditional Probability

Example 3.5 (***Birth Distributions***)

The Proof/Demonstration is:

Let

A = {both children are boys} = {bb}

B = {at least one child is a boy} = {bb, bg, gb}

Then,

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

$$= \frac{P(A)}{P(B)} \quad \textit{(why?)}$$

$$= \frac{1/4}{3/4} = \frac{1}{3}$$

Higher Order Conditional Probability

- Consider evaluating P(A ∩ B ∩ C). By considering A ∩ B as a single event D, say, we can write,

$$\begin{aligned} P(A \cap B \cap C) &= P(D \cap C) \\ &= P(C \mid D)P(D) \\ &= P(C \mid A \cap B)P(A \cap B) \\ &= P(C \mid A \cap B)P(B \mid A)P(A) \end{aligned}$$

- In an obvious fashion, we can extend this to n events $A_1, . . . , A_n$, and write,

$$\boxed{P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_n | A_1 \cap A_2 \cap \cdots \cap \mathrm{An{-}1}) \times P(A_{n-1} | A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \times \cdots \times P(A_2 | A_1)P(A_1)}$$

Higher Order Conditional Probability

Example 3.6 (***An Urn Problem***)

An urn contains five white balls, three black balls and four red balls. Four balls are drawn sequentially and without replacement. What is the probability of obtaining the sequence (W,W,R,B)?

We can represent the required sequence as follows:

$$(5W,3B,4R) \xrightarrow{W} (4W,3B,4R) \xrightarrow{W} (3W,3B,4R) \xrightarrow{R} (3W,3B,3R) \xrightarrow{B} (3W,2B,3R)$$

Then, the probability of obtaining the required sequence is,

$$P(WWRB) = P(B \mid WWR) \times P(R \mid WW) \times P(W \mid W) \times P(W)$$

$$= \frac{3}{9} \times \frac{4}{10} \times \frac{4}{11} \times \frac{5}{12}$$

$$= \frac{2}{99}$$

Conditional probabilities and Tree Diagrams

- When constructing tree diagrams the probabilities involved are usually conditional probabilities as there is a natural progression through the tree from left to right conditioning on what happened previously.

- The next example shows how conditioning works in an interesting technique to find the answer to embarrassing questions.

Conditional probabilities and Tree Diagrams

When conducting surveys of a sensitive or delicate nature, many respondents are often reluctant to divulge information. One simple way of reassuring them is to introduce a mask or shield in front of the awkward question.

Example 3.7 (***A question of confidentiality***)

Suppose a company wants to find what proportion of its employees have ever taken a day off sick to which they weren't entitled. The interviewer from the personnel department asks each employee to toss a coin and hide the result from them.

If the result was a head they were to answer the question *'Is your age in years an odd number?'* If the result was a tail they were to answer the question '*Have you ever taken a day off you shouldn't have?*'.

The key is that the interviewer doesn't know which question the employee is answering, so the employee needs to have no fear about answering either question truthfully.

Conditional probabilities and Tree Diagrams

Example 3.7 (***A question of confidentiality***)

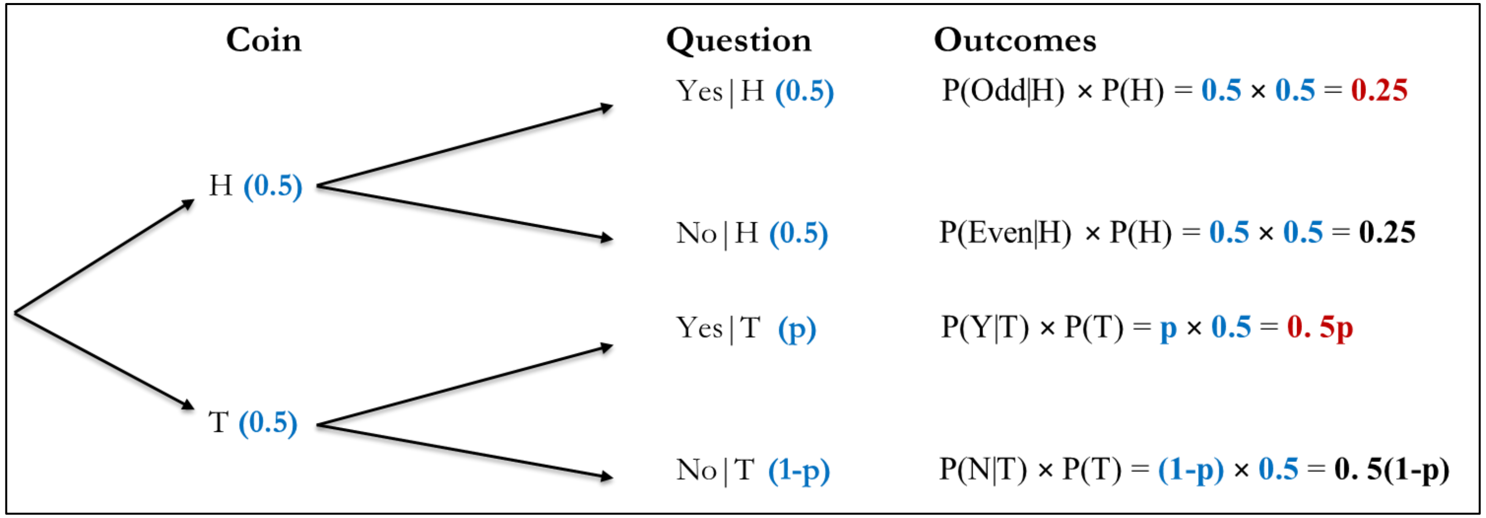
- Suppose that when interviewing a group of employees, 40% gave a '**yes**' answer. We now let **p = P(taken day off | Tail)**, i.e. the unknown proportion answering yes to the awkward question, where p is some number between 0 and 1.

- We assume that ages are randomly distributed so that the probabilities of even or odd ages are both $\frac{1}{2}$.

Conditional probabilities and Tree Diagrams

Example 3.7 (*A question of confidentiality*)

- We can then construct the following tree diagram,



- Because the result of the coin toss is hidden, we only observe the outcomes on the right hand side.
- The overall probability of answering '**yes**' is 0.25 + 0.5p and in the survey 40% answered 'yes'.
- We then have, $0.25 + 0.5p = 0.40$, so that, $0.5p = 0.15$ and $p=0.3$.
- Thus we estimate that 30% of employees have taken a day off.