Total Probability and Bayes' Theorem

35.4

Introduction

When the ideas of probability are applied to engineering (and many other areas) there are occasions when we need to calculate conditional probabilities other than those already known. For example, if production runs of ball bearings involve say, four machines, we might know the probability that any given machine produces faulty ball bearings. If we are inspecting the total output prior to distribution to users, we might need to know the probability that a faulty ball bearing came from a particular machine. Even though we do not address the area of statistics known as Bayesian Statistics here, it is worth noting that Bayes' theorem is the basis of this branch of the subject.

Prerequisites

Before starting this Section you should ...

- understand the ideas of sets and subsets.
- understand the concepts of probability and events.
- understand the addition and multiplication laws and the concept of conditional probability.

Learning Outcomes

On completion you should be able to ...

- understand the term 'partition of a sample space'
- understand the special case of Bayes' theorem arising when a sample space is partitioned by a set and its complement
- be able to apply Bayes' theorem to solve basic engineering related problems

1. The theorem of total probability

To establish this result we start with the definition of a partition of a sample space.

A partition of a sample space

The collection of events $A_1, A_2, \dots A_n$ is said to **partition** a sample space S if

(a) $A_1 \cup A_2 \cup \cdots \cup A_n = S$

(b) $A_i \cap A_j = \emptyset$ for all i, j

(c) $A_i \neq \emptyset$ for all i

In essence, a partition is a collection of non-empty, non-overlapping subsets of a sample space whose union is the sample space itself. The definition is illustrated by Figure 10.

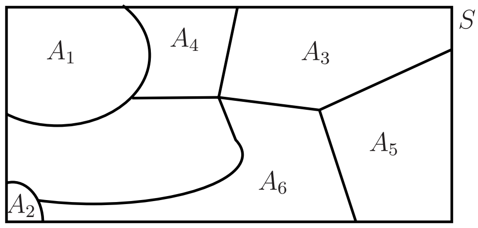


Figure 10

If B is any event within S then we can express B as the union of subsets:

$$B = (B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n)$$

The definition is illustrated in Figure 11 in which an event B in S is represented by the shaded region.

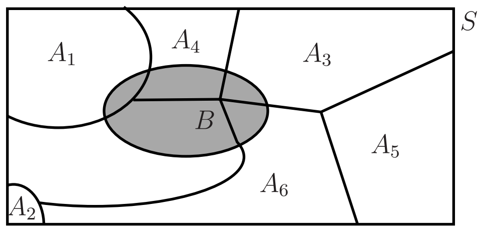


Figure 11

The bracketed events $(B \cap A_1), (B \cap A_2) \dots (B \cap A_n)$ are mutually exclusive (if one occurs then none of the others can occur) and so, using the addition law of probability for mutually exclusive events:

$$\mathsf{P}(B) = \mathsf{P}(B \cap A_1) + \mathsf{P}(B \cap A_2) + \cdots + \mathsf{P}(B \cap A_n)$$

Each of the probabilities on the right-hand side may be expressed in terms of conditional probabilities:

$$\mathsf{P}(B \cap A_i) = \mathsf{P}(B|A_i)\mathsf{P}(A_i) \qquad \text{for all } i$$

Using these in the expression for $\mathsf{P}(B)$, above, gives:

$$\begin{aligned} \mathsf{P}(B) &= \mathsf{P}(B|A_1)\mathsf{P}(A_1) + \mathsf{P}(B|A_2)\mathsf{P}(A_2) + \cdots + \mathsf{P}(B|A_n)\mathsf{P}(A_n) \\ &= \sum_{i=1}^{n} \mathsf{P}(B|A_i)\mathsf{P}(A_i) \end{aligned}$$

This is the theorem of Total Probability. A related theorem with many applications in statistics can be deduced from this, known as Bayes' theorem.

2. Bayes' theorem

We again consider the conditional probability statement:

$$\mathsf{P}(A|B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)} = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B|A_1)\mathsf{P}(A_1) + \mathsf{P}(B|A_2)\mathsf{P}(A_2) + \cdots + \mathsf{P}(B|A_n)\mathsf{P}(A_n)}$$

in which we have used the theorem of Total Probability to replace $\mathsf{P}(B)$. Now

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

Substituting this in the expression for $\mathsf{P}(A|B)$ we immediately obtain the result

$$\mathsf{P}(A|B) = \frac{\mathsf{P}(B|A) \times \mathsf{P}(A)}{\mathsf{P}(B|A_1)\mathsf{P}(A_1) + \mathsf{P}(B|A_2)\mathsf{P}(A_2) + \cdots + \mathsf{P}(B|A_n)\mathsf{P}(A_n)}$$

This is true for *any* event A and so, replacing A by A_i gives the result, known as Bayes' theorem as

$$\mathsf{P}(A_i|B) = \frac{\mathsf{P}(B|A_i) \times \mathsf{P}(A_i)}{\mathsf{P}(B|A_1)\mathsf{P}(A_1) + \mathsf{P}(B|A_2)\mathsf{P}(A_2) + \cdots + \mathsf{P}(B|A_n)\mathsf{P}(A_n)}$$

3. Special cases

In the case where we consider A to be an event in a sample space S (the sample space is partitioned by A and A') we can state simplified versions of the theorem of Total Probability and Bayes theorem as shown below.

The theorem of total probability: special case

This special case enables us to find the probability that an event B occurs taking into account the fact that another event A may or may not have occurred.
The theorem becomes

$$\mathsf{P}(B) = \mathsf{P}(B|A) \times \mathsf{P}(A) + \mathsf{P}(B|A') \times \mathsf{P}(A')$$

The result is easily seen by considering the general result already derived or it may be derived directly as follows. Consider Figure 12:

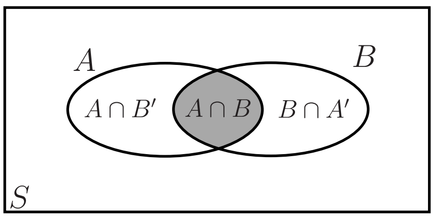


Figure 12

It is easy to see that the event B consists of the union of the (disjoint) events $A \cap B$ and $B \cap A'$ so that we may write B as the union of these disjoint events. We have

$$B = (A \cap B) \cup (B \cap A')$$

Since the events $A \cap B$ and $B \cap A'$ are disjoint, they must be independent and so

$$\mathsf{P}(B) = \mathsf{P}(A \cap B) + \mathsf{P}(B \cap A')$$

Using the conditional probability results we already have we may write

$$\begin{aligned} \mathsf{P}(B) &= \mathsf{P}(A \cap B) + \mathsf{P}(B \cap A') \\ &= \mathsf{P}(B \cap A) + \mathsf{P}(B \cap A') \\ &= \mathsf{P}(B|A) \times \mathsf{P}(A) + \mathsf{P}(B|A') \times \mathsf{P}(A') \end{aligned}$$

The result we have derived is

$$\mathsf{P}(B) = \mathsf{P}(B|A) \times \mathsf{P}(A) + \mathsf{P}(B|A') \times \mathsf{P}(A')$$

Bayes' theorem: special case

This result is obtained by supposing that the sample space S is partitioned by event A and its complement A' to give:

$$\mathsf{P}(A|B) = \frac{\mathsf{P}(B|A) \times \mathsf{P}(A)}{\mathsf{P}(B|A) \times \mathsf{P}(A) + \mathsf{P}(B|A') \times \mathsf{P}(A')}$$

Example 13

At a certain university, 4% of men are over 6 feet tall and 1% of women are over 6 feet tall. The total student population is divided in the ratio 3:2 in favour of women. If a student is selected at random from among all those over six feet tall, what is the probability that the student is a woman?

Solution

Let M={Student is Male}, F={Student is Female}.
Note that M and F partition the sample space of students.
Let T={Student is over 6 feet tall}.

We know that $\mathsf{P}(M) = 2/5$, $\mathsf{P}(F) = 3/5$, $\mathsf{P}(T|M) = 4/100$ and $\mathsf{P}(T|F) = 1/100$.

We require $\mathsf{P}(F|T)$. Using Bayes' theorem we have:

$$\begin{aligned} \mathsf{P}(F|T) &= \frac{\mathsf{P}(T|F)\mathsf{P}(F)}{\mathsf{P}(T|F)\mathsf{P}(F) + \mathsf{P}(T|M)\mathsf{P}(M)} \\ &= \frac{\frac{1}{100} \times \frac{3}{5}}{\frac{1}{100} \times \frac{3}{5} + \frac{4}{100} \times \frac{2}{5}} \\ &= \frac{3}{11} \end{aligned}$$

Example 14

A factory production line is manufacturing bolts using three machines, A, B and C. Of the total output, machine A is responsible for 25%, machine B for 35% and machine C for the rest. It is known from previous experience with the machines that 5% of the output from machine A is defective, 4% from machine B and 2% from machine C. A bolt is chosen at random from the production line and found to be defective. What is the probability that it came from

(a) machine A (b) machine B (c) machine C?

Solution

Let

D={bolt is defective},

A={bolt is from machine A},

B={bolt is from machine B},

C={bolt is from machine C}.

We know that $\mathsf{P}(A) = 0.25$, $\mathsf{P}(B) = 0.35$ and $\mathsf{P}(C) = 0.4$.

Also

$\mathsf{P}(D|A) = 0.05, \mathsf{P}(D|B) = 0.04, \mathsf{P}(D|C) = 0.02.$

A statement of Bayes' theorem for three events A, B and C is

$$\begin{aligned}\mathsf{P}(A|D) &= \frac{\mathsf{P}(D|A)\mathsf{P}(A)}{\mathsf{P}(D|A)\mathsf{P}(A) + \mathsf{P}(D|B)\mathsf{P}(B) + \mathsf{P}(D|C)\mathsf{P}(C)} \\ &= \frac{0.05 \times 0.25}{0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.4} \\ &= 0.362\end{aligned}$$

Similarly

$$\begin{aligned}\mathsf{P}(B|D) &= \frac{0.04 \times 0.35}{0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.4} \\ &= 0.406 \\ \mathsf{P}(C|D) &= \frac{0.02 \times 0.4}{0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.4} \\ &= 0.232\end{aligned}$$

Task

An engineering company advertises a job in three newspapers, A, B and C. It is known that these papers attract undergraduate engineering readerships in the proportions 2:3:1. The probabilities that an engineering undergraduate sees and replies to the job advertisement in these papers are 0.002, 0.001 and 0.005 respectively. Assume that the undergraduate sees only one job advertisement.

(a) If the engineering company receives only one reply to it advertisements, calculate the probability that the applicant has seen the job advertised in place A.

(i) A, (ii) B, (iii) C.

(b) If the company receives two replies, what is the probability that both applicants saw the job advertised in paper A?

Your solution

Answer
Let

$A = \{\text{Person is a reader of paper } A\}$,

$B = \{\text{Person is a reader of paper } B\}$,

$C = \{\text{Person is a reader of paper } C\}$,

$R = \{\text{Reader applies for the job}\}$.

We have the probabilities

(a)

$$\begin{aligned}
\mathsf{P}(A) &= 1/3 \qquad \mathsf{P}(R|A) = 0.002\\
\mathsf{P}(B) &= 1/2 \qquad \mathsf{P}(R|B) = 0.001\\
\mathsf{P}(C) &= 1/6 \qquad \mathsf{P}(R|C) = 0.005\\
\mathsf{P}(A|R) &= \frac{\mathsf{P}(R|A)\mathsf{P}(A)}{\mathsf{P}(R|A)\mathsf{P}(A) + \mathsf{P}(R|B)\mathsf{P}(B) + \mathsf{P}(R|C)\mathsf{P}(C)} = \frac{1}{3}
\end{aligned}$$

Similarly

$$\mathsf{P}(B|R) = \frac{1}{4} \qquad \text{and} \qquad \mathsf{P}(C|R) = \frac{5}{12}$$

(b) Now, assuming that the replies and readerships are independent

$$\begin{aligned}
\mathsf{P}(\text{Both applicants read paper} A) &= \mathsf{P}(A|R) \times \mathsf{P}(A|R)\\
&= \frac{1}{3} \times \frac{1}{3}\\
&= \frac{1}{9}
\end{aligned}$$

Exercises

1. Obtain the sample space of an experiment that consists of a fair coin being tossed four times. Consider the following events:

   A is the event 'all four results are the same.'

   B is the event 'exactly one Head occurs.'

   C is the event 'at least two Heads occur.'

   Show that $\text{P}(A) + \text{P}(B) + \text{P}(C) = \dfrac{17}{16}$ and explain why $\text{P}(A) + \text{P}(B) + \text{P}(C) > 1$.

2. The table below show the number of complete years a group of people have been working in their current employment.

Years of Employment	Number of People
0 or 1 year	15
2 or 3 years	12
4 or 5 years	9
6 or 8 years	6
8 to 11 years	6
12 years and over	2

   What is the probability that a person from the group, selected at random;

   (a) is in the modal group
   (b) has been working there for less than 4 years
   (c) has been working there for at least 8 years.

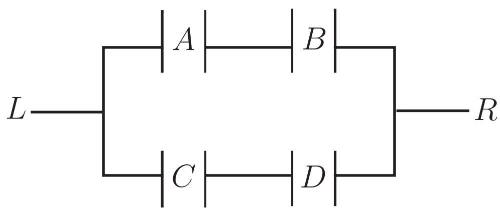
3. It is a fact that if A and B are independent events then it is also true that A' and B' are independent events. If A and B are independent events such that the probability that they both occur simultaneously is $\dfrac{1}{8}$ and the probability that neither of them will occur is $\dfrac{3}{8}$, find:

   (a) the probability that event A will occur
   (b) the probability that event B will occur.

4. If A and B are two events associated with an experiment and $\text{P}(A) = 0.4$, $\text{P}(A \cup B) = 0.7$ and $\text{P}(B) = p$, find:

   (a) the choice of p for which A and B are mutually exclusive
   (b) the choice of p for which A and B are independent.

5. The probability that each relay closes in the circuit shown below is p. Assuming that each relay functions independently of the others, find the probability that current can flow from L to R.



6. From a batch of 100 items of which 20 are defective, exactly two items are chosen, one at a time, without replacement. Calculate the probabilities that:

   (a) the first item chosen is defective

   (b) both items chosen are defective

   (c) the second item chosen is defective.

7. A garage mechanic keeps a box of good springs to use as replacements on customers cars. The box contains 5 springs. A colleague, thinking that the springs are for scrap, tosses three faulty springs into the box. The mechanic picks two springs out of the box while servicing a car. Find the probability that:

   (a) the first spring drawn is faulty (b) the second spring drawn is faulty.

8. Two coins are tossed. Find the conditional probability that two Heads will occur given that at least one occurs.

9. Machines A and B produce 10% and 90% respectively of the production of a component intended for the motor industry. From experience, it is known that the probability that machine A produces a defective component is 0.01 while the probability that machine B produces a defective component is 0.05. If a component is selected at random from a day's production and is found to be defective, find the probability that it was made by

   (a) machine A (b) machine B.

Answers

1. $\mathsf{P}(A) = \dfrac{2}{16}$, $\mathsf{P}(B) = \dfrac{4}{16}$, $\mathsf{P}(C) = \dfrac{11}{16}$, $\mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C) = \dfrac{17}{16}$

   A, B and C are not mutually exclusive since events A and C have outcomes in common. This is the reason why $\mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C) = \dfrac{17}{16}$; we are adding the probabilities corresponding to common outcomes more than once.

2. (a) $\mathsf{P}(\text{person falls in the modal group}) = \dfrac{15}{50}$

   (b) $\mathsf{P}(\text{person has been working for less than 4 years}) = \dfrac{27}{50}$

   (c) $\mathsf{P}(\text{person has been working for more than 8 years}) = \dfrac{8}{50}$

3. $\mathsf{P}(A) \times \mathsf{P}(B) = \dfrac{1}{8}$ and $(1 - \mathsf{P}(A)) \times (1 - \mathsf{P}(B)) = \dfrac{3}{8}$

   Treat these equations as $xy = \dfrac{1}{8}$ and $(1 - x)(1 - y) = \dfrac{3}{8}$ and solve to get:

   $\mathsf{P}(A) = \dfrac{1}{2}$ (or $\dfrac{1}{4}$) and $\mathsf{P}(B) = \dfrac{1}{4}$ (or $\dfrac{1}{2}$)

4. (a) $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B)$ so $0.7 = 0.4 + p$ implying $p = 0.3$

   (b) $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A) \times \mathsf{P}(B)$ so $0.7 = 0.4 + p - 0.4 \times p$ implying $p = 0.5$.

Answers

5. $$\begin{aligned}\text{P}((A\cap B)\cup(C\cap D)) &= \text{P}(A\cap B)+\text{P}(C\cap D)-\text{P}(A\cap B\cap C\cap D)\\ &= p^2+p^2-p^4\\ &= 2p^2-p^4\end{aligned}$$

6. Let A={first item chosen is defective}, B ={second item chosen is defective}

(a) $\text{P}(A)=\dfrac{20}{100}=\dfrac{1}{5}$

(b) $\text{P}(A\cap B)=\text{P}(A|B)\text{P}(A)=\dfrac{19}{99}\times\dfrac{20}{100}=\dfrac{19}{495}$

(c) $\text{P}(B)=\text{P}(B|A)\text{P}(A)+\text{P}(B|A')\text{P}(A')=\dfrac{19}{99}\times\dfrac{20}{100}+\dfrac{20}{99}\times\dfrac{80}{100}=\dfrac{198}{990}=\dfrac{1}{5}$

7. Let A ={first spring chosen is faulty}, B ={second spring chosen is faulty}

(a) $\text{P}(A)=\dfrac{3}{8}$

(b) $\text{P}(B)=\text{P}(B|A)\text{P}(A)+\text{P}(B|A')\text{P}(A')=\dfrac{2}{7}\times\dfrac{3}{8}+\dfrac{3}{7}\times\dfrac{5}{8}=\dfrac{21}{56}=\dfrac{3}{8}$

8. Let A = {at least one Head occurs}, B = {two Heads occur}

$$\text{P}(B)=\frac{\text{P}(A\cap B)}{\text{P}(A\cup B)}=\frac{\text{P}(A)\times\text{P}(B)}{\text{P}(A)+\text{P}(B)-\text{P}(A)\times\text{P}(B)}=\frac{\frac{1}{2}\times\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}-\frac{1}{2}\times\frac{1}{2}}=\frac{1}{3}$$

9. Let A = {item from machine A}, B = {item from machine B}, D = {item is defective}.
We know that: $\text{P}(A)=0.1,\ \text{P}(B)=0.9,\ \text{P}(D|A)=0.01,\ \text{P}(D|B)=0.05.$

(a)

$$\begin{aligned}\text{P}(A|D) &= \frac{\text{P}(D|A)\text{P}(A)}{\text{P}(D|A)\text{P}(A)+\text{P}(D|B)\text{P}(B)}\\ &= \frac{0.01\times 0.1}{0.01\times 0.1+0.05\times 0.9}\\ &= 0.02\end{aligned}$$

(b) Similarly $\text{P}(B|D)=0.98$