Probability and Statistics (IT302)

10th August 2020 (Monday) Class

Class-4

Total Probability Rule (Two Events)

For any events A and B,

$$P(B) = P(B \cap A) + P(B \cap A') = P(B|A)P(A) + P(B|A')P(A')$$

Total Probability Rule (Multiple Events)

Assume $E_1, E_2, \ldots, E_k$ are k mutually exclusive and exhaustive sets. Then

$$\begin{aligned} P(B) &= P(B \cap E_1) + P(B \cap E_2) + \cdots + P(B \cap E_k) \\ &= P(B|E_1)P(E_1) + P(B|E_2)P(E_2) + \cdots + P(B|E_k)P(E_k) \end{aligned}$$

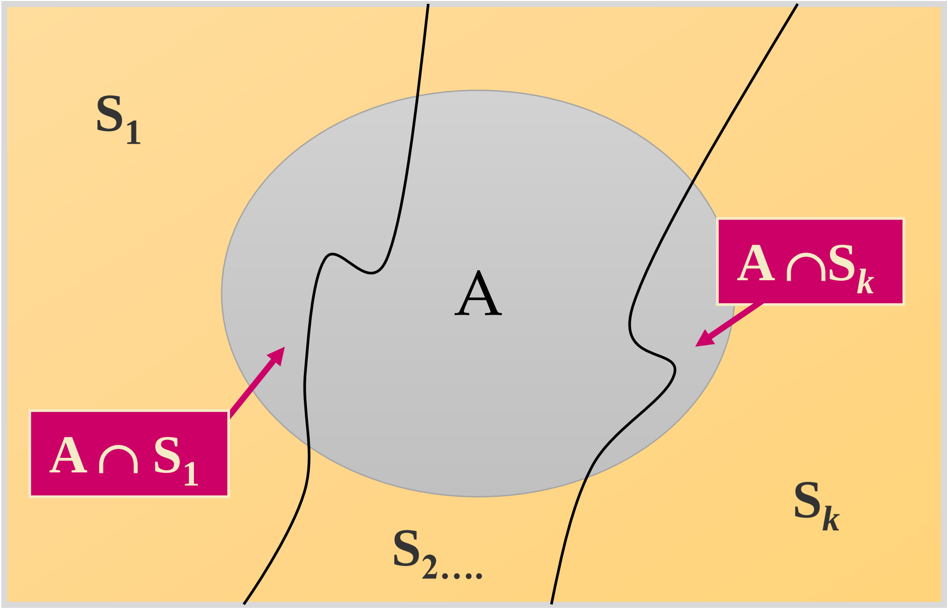
Source : Applied Statistics and Probability for Engineers by Douglas C.Montgomery and George C. Runger, 3rd Edition.

The Law of Total Probability

Let S_1 , S_2 , S_3 ,..., S_k be mutually exclusive and exhaustive events (that is, one and only one must happen). Then the probability of any event A can be written as

$$
\begin{aligned}
P(A) &= P(A \cap S_1) + P(A \cap S_2) + \ldots + P(A \cap S_k) \\
&= P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + \ldots + P(S_k)P(A|S_k)
\end{aligned}
$$

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt



Bayes' Rule

Let S_1 , S_2 , S_3 ,..., S_k be mutually exclusive and exhaustive events with prior probabilities $P(S_1)$, $P(S_2)$,…,$P(S_k)$. If an event A occurs, the posterior probability of S_i, given that A occurred is

$$P(S_i|A) = \frac{P(S_i)P(A|S_i)}{\sum P(S_i)P(A|S_i)} \text{ for } i = 1, 2, \ldots k$$

Proof

$$P(A|S_i) = \frac{P(AS_i)}{P(S_i)} \longrightarrow P(AS_i) = P(S_i)P(A|S_i)$$

$$P(S_i|A) = \frac{P(AS_i)}{P(A)} = \frac{P(S_i)P(A|S_i)}{\sum P(S_i)P(A|S_i)}$$

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Example-1

From a previous example, we know that 49% of the population are female. Of the female patients, 8% are high risk for heart attack, while 12% of the male patients are high risk. A single person is selected at random and found to be high risk. What is the probability that it is a male?

Define H: high risk   F: female   M: male

We know:

$P(F) =$ .49

$P(M) =$ .51

$P(H|F) =$ .08

$P(H|M) =$ .12

$$P(M|H) = \frac{P(M)P(H|M)}{P(M)P(H|M) + P(F)P(H|F)}$$

$$= \frac{.51\,(.12)}{.51\,(.12) + .49\,(.08)} = .61$$

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Example-2

Suppose a rare disease infects one out of every 1000 people in a population. Suppose that there is a good, but not perfect, test for this disease: if a person has the disease, the test comes back positive 99% of the time. On the other hand, the test also produces some false positives: 2% of uninfected people are also test positive. And someone just tested positive. What are his chances of having this disease?

Define A: has the disease B: test positive

We want to know $P(A|B)=?$

We Know

$P(A) = .001$ $\quad P(A^c) = .999$

$P(B|A) = .99$ $\quad P(B|A^c) = .02$

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

$$= \frac{.001 \times .99}{.001 \times .99 + .999 \times .02} = .0472$$

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Example-3

A survey of job satisfaction of teachers was taken, giving the following results

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	74	43	117
	High School	224	171	395
	Elementary	126	140	266
	Total	424	354	778

If all the cells are divided by the total number surveyed, 778, the resulting Table is a Table of empirically derived probabilities.

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.342
	Total	0.545	0.455	1.000

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Example-3 Contd.

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.342
	Total	0.545	0.455	1.000

For convenience, let C stand for the event that the teacher teaches college, S stand for the teacher being satisfied and so on. Let's look at some probabilities and what they mean.

$P(C) = 0.150$	is the proportion of teachers who are college teachers.
$P(S) = 0.545$	is the proportion of teachers who are satisfied with their job.
$P(C \cap S) = 0.095$	is the proportion of teachers who are college teachers and who are satisfied with their job.

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Example-3 Contd.

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.342
	Total	0.545	0.455	1.000

$$P(C|S) = \frac{P(C \cap S)}{P(S)} = \frac{0.095}{0.545} = 0.175$$

is the proportion of teachers who are college teachers given they are satisfied. Restated: This is the proportion of satisfied that are college teachers.

$$P(S|C) = \frac{P(S \cap C)}{P(C)} = \frac{P(C \cap S)}{P(C)} = \frac{0.095}{0.150} = 0.632$$

is the proportion of teachers who are satisfied given they are college teachers. Restated: This is the proportion of college teachers that are satisfied.

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Example-3 Contd.

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.342
	Total	0.545	0.455	1.000

Are C and S independent events?

$$P(C) = 0.150 \text{ and } P(C|S) = \frac{P(C \cap S)}{P(S)} = \frac{0.095}{0.545} = 0.175$$

$P(C|S) \neq P(C)$ so C and S are dependent events.

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Example-3 Contd.

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.658
	Total	0.545	0.455	1.000

$P(C \cup S)$?

$P(C) = 0.150$, $P(S) = 0.545$ and $P(C \cap S) = 0.095$, so

$$P(C \cup S) = P(C)+P(S) - P(C \cap S)$$
$$= 0.150 + 0.545 - 0.095$$
$$= 0.600$$

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Example-4

Tom and Dick are going to take a driver's test at the nearest DMV office. Tom estimates that his chances to pass the test are 70% and Dick estimates his as 80%. Tom and Dick take their tests independently.

Define D = {Dick passes the driving test}   T = {Tom passes the driving test}

T and D are independent. P (T) = 0.7,   P (D) = 0.8

What is the probability that at **most one of the two friends will pass the test?**

P(At most one person pass)

$= P(D^c \cap T^c) + P(D^c \cap T) + P(D \cap T^c)$

$= (1 - 0.8)(1 – 0.7) + (0.7)(1 – 0.8) + (0.8)(1 – 0.7)$

$= .44$

P(At most one person pass)

$= 1\text{-}P(\text{both pass}) = 1 - 0.8 \times 0.7 = .44$

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Example-4 Contd.

Tom and Dick are going to take a driver's test at the nearest DMV office. Tom estimates that his chances to pass the test are 70% and Dick estimates his as 80%. Tom and Dick take their tests independently.

Define D = {Dick passes the driving test}   T = {Tom passes the driving test}

T and D are independent. P (T) = 0.7,   P (D) = 0.8

What is the probability **that at least one of the two friends will pass the test?**

P(At least one person pass)

$= P(D \cup T)$

$= 0.8 + 0.7 - 0.8 \times 0.7$

$= .94$

P(At least one person pass)

$= 1 - P(\text{neither passes}) = 1 - (1-0.8) \times (1-0.7) = .94$

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Example-4 Contd.

Tom and Dick are going to take a driver's test at the nearest DMV office. Tom estimates that his chances to pass the test are 70% and Dick estimates his as 80%. Tom and Dick take their tests independently.

Define D = {Dick passes the driving test} T = {Tom passes the driving test}

T and D are independent. P (T) = 0.7, P (D) = 0.8

Suppose only **one of the two friends passed the test. What is the probability that it was Dick?**

$$P(D \mid \text{exactly one person passed})$$

$$= P(D \cap \text{exactly one person passed}) / P(\text{exactly one person passed})$$

$$= P(D \cap T^c) / (P(D \cap T^c) + P(D^c \cap T))$$

$$= 0.8 \times (1-0.7) / (0.8 \times (1-0.7) + (1-.8) \times 0.7)$$

$$= .63$$

Source: http://www.pstat.ucsb.edu/faculty/yuedong/5E/5Enote5.ppt

Theorem 2.13

If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A of S,

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)P(A|B_i).$$

Proof **:** Consider the Venn diagram of Figure, the event A is seen to be the union of the mutually exclusive events $B_1 \cap A, B_2 \cap A, \ldots, B_k \cap A$; that is, $A = (B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_k \cap A)$. Using Corollary 2.2 of Theorem 2.7 and Theorem 2.10, we have

$$\begin{aligned} P(A) &= P[(B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_k \cap A)] \\ &= P(B_1 \cap A) + P(B_2 \cap A) + \cdots + P(B_k \cap A) \\ &= \sum_{i=1}^{k} P(B_i \cap A) \\ &= \sum_{i=1}^{k} P(B_i)P(A|B_i). \end{aligned}$$

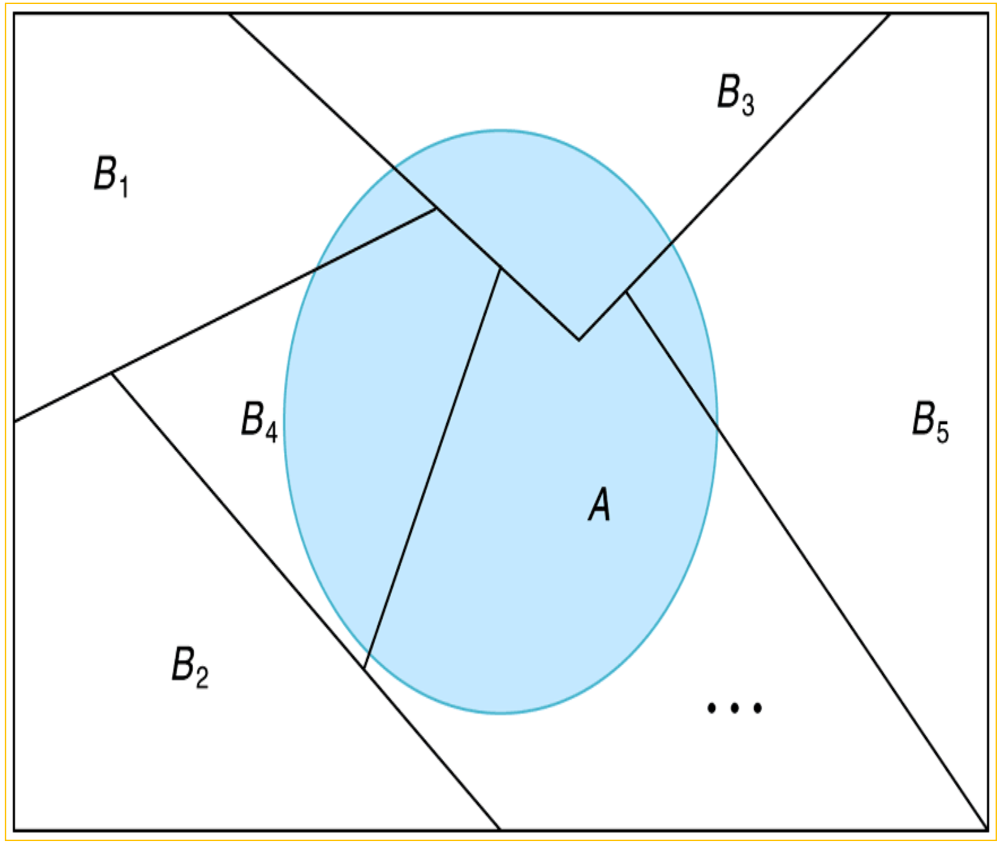


Figure 2.14 Partitioning the sample space S

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall

Example-2.14

Example 2.41: In a certain assembly plant, three machines, $B1$, $B2$, and $B3$, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected. What is the probability that it is defective?

Solution : Consider the following events:

A: the product is defective,

$B1$: the product is made by machine $B1$,

$B2$: the product is made by machine $B2$,

$B3$: the product is made by machine $B3$.

Applying the rule of elimination, we can write

$P(A) = P(B1)P(A|B1) + P(B2)P(A|B2) + P(B3)P(A|B3)$.

Referring to the tree diagram of Figure 2.15, we find that the three branches give the probabilities

$P(B1)P(A|B1) = (0.3)(0.02) = 0.006$,

$P(B2)P(A|B2) = (0.45)(0.03) = 0.0135$,

$P(B3)P(A|B3) = (0.25)(0.02) = 0.005$, and hence

$P(A) = 0.006 + 0.0135 + 0.005 = 0.0245$.

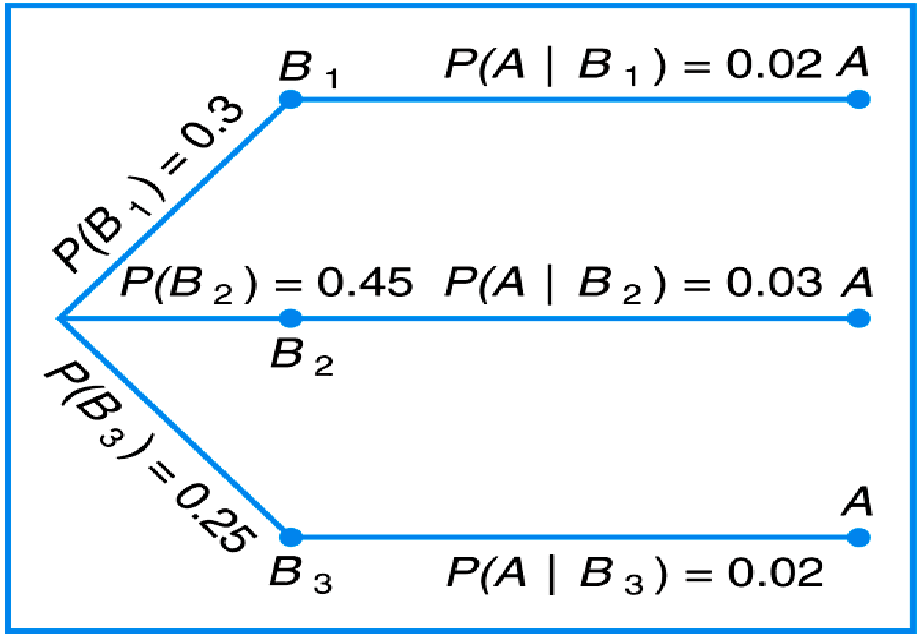


Figure 2.15 Tree diagram for Example 2.41

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall

Theorem 2.14

(Bayes' Rule) If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A in S such that $P(A) \neq 0$,

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)} \quad \text{for } r = 1, 2, \ldots, k.$$

Proof: By the definition of conditional probability,

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

and then using Theorem 2.13 in the denominator, we have

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)},$$

which completes the proof.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall

Example 2.42

Example 2.42: In a certain assembly plant, three machines, *B*1, *B*2, and *B*3, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective. if a product was chosen randomly and found to be defective, what is the probability that it was made by machine *B*3?

Solution :

A: the product is defective,

*B*1: the product is made by machine *B*1,

*B*2: the product is made by machine *B*2,

*B*3: the product is made by machine *B*3.

$$P(B_3|A) = \frac{0.005}{0.006 + 0.0135 + 0.005} = \frac{0.005}{0.0245} = \frac{10}{49}.$$

In view of the fact that a defective product was selected, this result suggests that it probably was not made by machine *B*3.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall

Example 2.43

Example 2.43: A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2, and 3 are used for 30%, 20%, and 50% of the products, respectively. The defect rate is different for the three procedures as follows: $P(D|P1) = 0.01$, $P(D|P2) = 0.03$, $P(D|P3) = 0.02$, where $P(D|Pj)$ is the probability of a defective product, given plan j. If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

Solution: From the statement of the problem

$$P(P_1) = 0.30, \quad P(P_2) = 0.20, \quad \text{and} \quad P(P_3) = 0.50,$$

we must find $P(P_j|D)$ for $j = 1, 2, 3$. Bayes' rule (Theorem 2.14) shows

$$\begin{aligned} P(P_1|D) &= \frac{P(P_1)P(D|P_1)}{P(P_1)P(D|P_1) + P(P_2)P(D|P_2) + P(P_3)P(D|P_3)} \\ &= \frac{(0.30)(0.01)}{(0.3)(0.01) + (0.20)(0.03) + (0.50)(0.02)} = \frac{0.003}{0.019} = 0.158. \end{aligned}$$

Similarly,

$$P(P_2|D) = \frac{(0.03)(0.20)}{0.019} = 0.316 \text{ and } P(P_3|D) = \frac{(0.02)(0.50)}{0.019} = 0.526.$$

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall

Exercise Problem (Question No.95)

In a certain region of the country it is known from past experience that the probability of selecting an adult over 40 years of age with cancer is 0.05. If the probability of a doctor correctly diagnosing a person with cancer as having the disease is 0.78 and the probability of incorrectly diagnosing a person without cancer as having the disease is 0.06, what is the probability that an adult over 40 years of age is diagnosed as having cancer?

Solution:

Consider the events: C : an adult selected has cancer,

D : the adult is diagnosed as having cancer.

- $P(C) = 0.05$,
- $P(D \mid C) = 0.78$,
- $P(C^J) = 0.95$ and $P(D \mid C^J) = 0.06$.
- So, $P(D) = P(C \cap D) + P(C^J \cap D) = (0.05)(0.78) + (0.95)(0.06) = 0.096$.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall

Additional Slide Related to 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

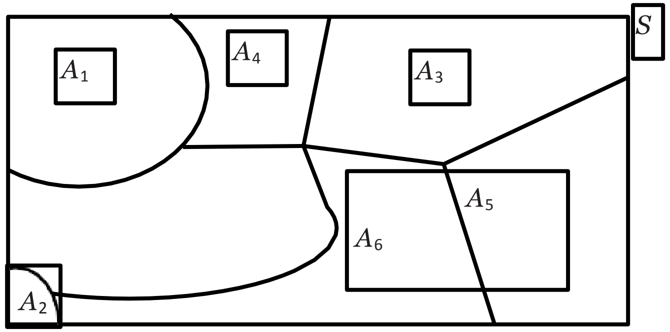
The collection of events $A_1, A_2, \ldots A_n$ is said to **partition** a sample space S if

(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 below Figure



Source:https://learn.lboro.ac.uk/archive/olmp/olmp_resources/pages/workbooks_1_50_jan2008/Workbook35/35_4_total_prob_bayes_thm.pdf Accessed on 7th August 2020

Additional Slide Related to Theorem of Total Probability Contd.

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 below in which an event B in S is represented by the shaded region.

The bracketed events $(B \cap A_1), (B \cap A_2) \ldots (B \cap A_n)$ 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:

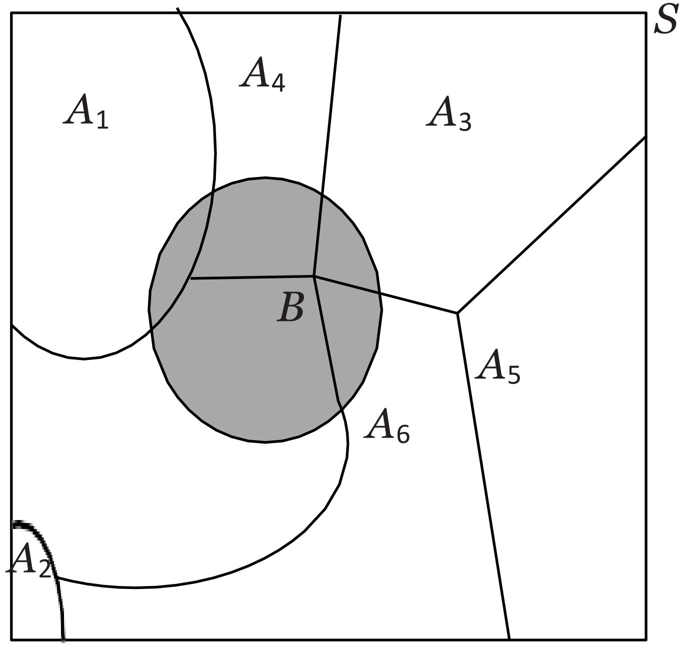
$$\mathsf{P}(B) = \mathsf{P}(B \cap A_1) + \mathsf{P}(B \cap A_2) + \cdots + \mathsf{P}(B \cap A_n)$$

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.

Source : https://learn.lboro.ac.uk/archive/olmp/olmp_resources/pages/workbooks_1_50_jan2008/Workbook35/35_4_total_prob_bayes_thm.pdf Accessed on 7th August 2020

Additional Slide Related to Bayes' Theorem

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

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

Source: https://learn.lboro.ac.uk/archive/olmp/olmp_resources/pages/workbooks_1_50_jan2008/Workbook35/35_4_total_prob_bayes_thm.pdf Accessed on 7th August 2020

Theorem of Total Probability (Special Case)

Special cases

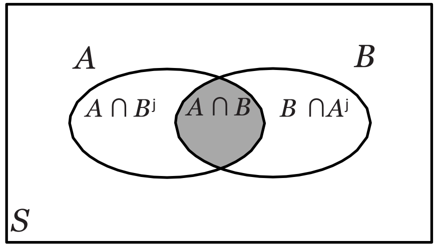
In the case where we consider A to be an event in a sample space S (the sample space is partitioned by A and A^j) we can state simplified versions of the theorem of Total Probability and Bayes theorem as shown below.

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

$$P(B) = P(B|A) \times P(A) + P(B|A^j) \times P(A^j)$$

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



It is easy to see that the event B consists of the union of the (disjoint) events $A \cap B$ and $B \cap A^j$ so

Source : https://learn.lboro.ac.uk/archive/olmp/olmp_resources/pages/workbooks_1_50_jan2008/Workbook35/35_4_total_prob_bayes_thm.pdf Accessed on 7th August 2020

Theorem of Total Probability (Special Case) Contd.

that we may write B as the union of these disjoint events. We have

$$B = (A \cap B) \cup (B \cap A^j)$$

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

$$\mathrm{P}(B) = \mathrm{P}(A \cap B) + \mathrm{P}(B \cap A^j)$$

Using the conditional probability results we already have we may write

$$\begin{aligned} \mathrm{P}(B) &= \mathrm{P}(A \cap B) + \mathrm{P}(B \cap A^j) \\ &= \mathrm{P}(B \cap A) + \mathrm{P}(B \cap A^j) \\ &= \mathrm{P}(B|A) \times \mathrm{P}(A) + \mathrm{P}(B|A^j) \times \mathrm{P}(A^j) \end{aligned}$$

The result we have derived is

$$\mathrm{P}(B) = \mathrm{P}(B|A) \times \mathrm{P}(A) + \mathrm{P}(B|A^j) \times \mathrm{P}(A^j)$$

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:

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

Source:https://learn.lboro.ac.uk/archive/olmp/olmp_resources/pages/workbooks_1_50_jan2008/Workbook35/35_4_total_prob_bayes_thm.pdf Accessed on 7th August 2020

Example-5

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?

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

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

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

Source : https://learn.lboro.ac.uk/archive/olmp/olmp_resources/pages/workbooks_1_50_jan2008/Workbook35/35_4_total_prob_bayes_thm.pdf Accessed on 7th August 2020

Example-6

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 $\mathrm{P}(A) = 0.25$, $\mathrm{P}(B) = 0.35$ and $\mathrm{P}(C) = 0.4$.

Also

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

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

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

Similarly

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

Source : https://learn.lboro.ac.uk/archive/olmp/olmp_resources/pages/workbooks_1_50_jan2008/Workbook35/35_4_total_prob_bayes_thm.pdf Accessed on 7th August 2020

Second Set of Additional Slides Related to Bayes' Theorem

Bayes' Theorem is a way of finding a probability when we know certain other probabilities.

The formula is: **P(A|B) = (P(A) × P(B|A)) / P(B)**

Which tells us : how often A happens given that B happens, written P(A|B)

When we know : how often B happens given that A happens, written P(B|A)

and how likely A is on its own, written P(A)

and how likely B is on its own, written P(B)

Let us say P(Fire) means how often there is fire, and P(Smoke) means how often we see smoke, then:

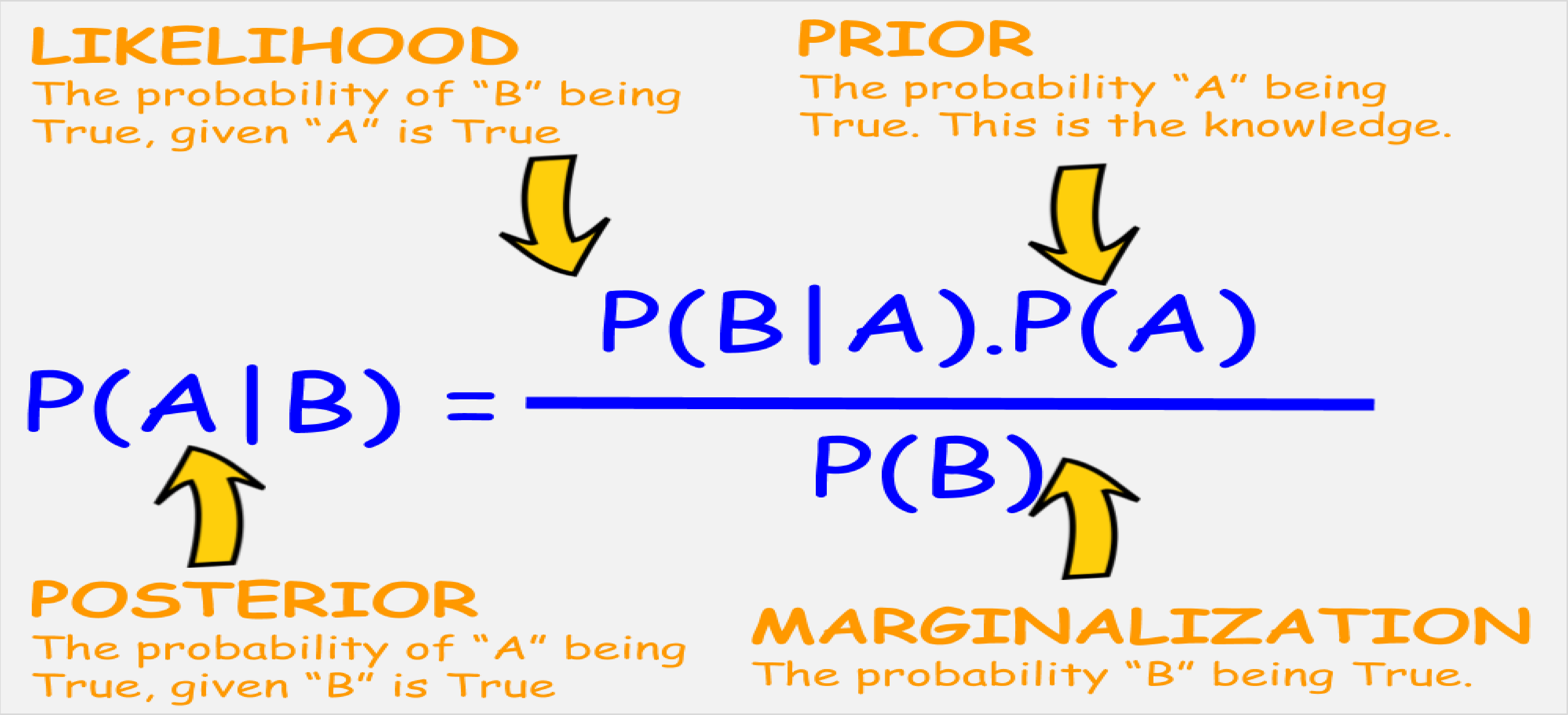
P(Fire|Smoke) means how often there is fire when we can see smoke

P(Smoke|Fire) means how often we can see smoke when there is fire

So the formula kind of tells us "forwards" P(Fire|Smoke) when we know "backwards" P(Smoke|Fire)

Source : https://www.mathsisfun.com/data/bayes-theorem.html Accessed on 7th August 2020

Second Set of Additional Slides Related to Bayes’ Theorem



Source : https://towardsdatascience.com/bayes-rule-with-a-simple-and-practical-example-2bce3d0f4ad0 Accessed on 7th August 2020

Second Set of Additional Slides Related to Bayes' Theorem Contd.

Example-7

- Dangerous fires are rare (1%)
- Smoke is fairly common (10%) due to barbecues,
- 90% of dangerous fires make smoke

Discover the **Probability of Dangerous Fire when there is Smoke**:

$$P(Fire|Smoke) = (P(Fire) \times P(Smoke|Fire)) / P(Smoke)$$
$$= 1\% \text{ x } 90\% / 10\%$$
$$= 9\%$$

So it is still worth checking out any smoke to be sure.

Source : https://www.mathsisfun.com/data/bayes-theorem.html Accessed on 7th August 2020

Second Set of Additional Slides Related to Bayes' Theorem Contd.

You are planning a picnic today, but the morning is cloudy. Oh no! 50% of all rainy days start off cloudy!. But, cloudy mornings are common (about 40% of days start cloudy) and this is usually a dry month (only 3 of 30 days tend to be rainy, or 10%). **What is the chance of rain during the day?** Rain to means rain during the day, and Cloud to mean cloudy morning.

The chance of Rain given Cloud is written P(Rain|Cloud)

So let's put that in the formula : P(Rain|Cloud) = (P(Rain) × **P(Cloud|Rain)**)/ **P(Cloud)**

- **P(Rain)** is Probability of Rain = 10%
- **P(Cloud|Rain)** is Probability of Cloud, given that Rain happens = 50%
- **P(Cloud)** is Probability of Cloud = 40%

P(Rain|Cloud) = 0.1 x 0.5 / **0.4** = .125 OR a 12.5% chance of rain.

Source : https://www.mathsisfun.com/data/bayes-theorem.html Accessed on 7th August 2020