Application Of Set Theory To Probability

The ideas outlined in Set Theory are central to the study of Probability Theory. We all believe (or least if we had to put our own money on it, would bet) that if a card is drawn at **random** from a **standard deck** of 52 cards, then

$$\text{the probability of drawing an ace is } \frac{4}{52}.$$

Our argument that this should be the case is simply that;

> because there are 4 aces in the deck and, since the cards are drawn at random, each card is **equally likely** of being drawn.

This illustrates the following most important set-up in Probability Theory.

The Sample Space Of Equally Likely Outcomes: We begin with a **finite set** S called a sample space. For example, S is the set of cards in a standard deck. We imagine that if we were to draw any element from S then all elements $s \in S$ would have the same probability $(= 1/|S|)$ of being drawn. If E is any subset of S then the probability that an element drawn from S is in fact from E (denoted by $\boldsymbol{P}(E)$) is given by

$$\boldsymbol{P}(E) = \frac{|E|}{|S|}.$$

Note, $\boldsymbol{P}(E)$ is called **The Probability Of The EVENT** E. The following are now immediate:

(i) $\boldsymbol{P}(S) = 1$ and $\boldsymbol{P}(\emptyset) = 0$.

(ii) If the events (i.e. subsets of S) $E_1, E_2, \ldots, E_m$ are **disjoint** then

$$\begin{aligned} \boldsymbol{P}(E_1 \cup E_2 \cup \ldots \cup E_m) &= \frac{|\, E_1 \cup E_2 \cup \ldots \cup E_m \,|}{|S|} \\ &= \frac{|E_1| + |E_2| + \cdots + |E_m|}{|S|} \\ &= \boldsymbol{P}(E_1) + \boldsymbol{P}(E_2) + \cdots + \boldsymbol{P}(E_m). \end{aligned}$$

(iii) For any event E (with complement $\bar{E}$) we have the **disjoint union**

$$\begin{aligned} & E \cup \bar{E} &= S \\ \Rightarrow\quad & \boldsymbol{P}(E \cup \bar{E}) &= \boldsymbol{P}(S) \\ \Rightarrow\quad & \boldsymbol{P}(E) + \boldsymbol{P}(\bar{E}) &= 1 \\ \Rightarrow\quad & \boldsymbol{P}(\bar{E}) &= 1 - \boldsymbol{P}(E). \end{aligned}$$

Though all of these identities are very useful for computing probabilities, we will find that computation in general is much facilitated by the following notion.

Conditional Probability : As motivation, consider the following example. Suppose we put the names of all the students in this class into a hat and draw one at random. In this case, the

$$\text{sample space } S = \{\text{ all people in this class }\}.$$

Consider now the events

$$Front = \{\text{person whose name is drawn is in the } \textbf{front row}\}$$

and

$$Fe = \{\text{person whose name is drawn is } \textbf{female}\}.$$

As explained already we know how to calculate the probability that the name drawn is that of a person in the front row, in fact, this is just

$$\boldsymbol{P}(Front) = \frac{|Front|}{|S|}.$$

Now, suppose that I look at the name and tell you that the name is that of a female, how does this change your calculation? The answer is as follows. What you want to calculate now is:

> The probability of the event *Front* **given that** the event *Fe* has occurred. This is usually denoted by

We usually write

$$\boldsymbol{P}[\,Front \,|\, Fe] = \left\{ \begin{array}{c} \text{probability of the event } Front \\ \textbf{given that} \\ \text{the event } Fe \text{ has occurred.} \end{array} \right.$$

The main observation is that

the sample space has changed from S to Fe

and, in particular,

when considering $\boldsymbol{P}[\,Front \,|\, Fe]$ we must
count only the females from the front row.

That is, we must count only $|\,(\,Fe\,) \cap (\,Front\,)\,|$.

The calculation now is:

$$\boldsymbol{P}[\,Front \mid Fe\,] \;=\; \frac{|\,(\,Fe\,) \cap (\,Front\,)\,|}{|\,Fe\,|}$$

$$= \frac{\left[\dfrac{|\,(\,Fe\,) \cap (\,Front\,)\,|}{|\,S\,|}\right]}{\left[\dfrac{|\,Fe\,|}{|S|}\right]}$$

$$= \frac{\boldsymbol{P}[\,(\,Fe\,) \cap (\,Front\,)\,]}{\boldsymbol{P}[\,Fe\,]}.$$

That is;

$$\frac{\boldsymbol{P}[\,(\,Fe\,) \cap (\,Front\,)\,]}{\boldsymbol{P}[\,Fe\,]} \;=\; \boldsymbol{P}[\,Front \mid Fe\,]$$

and, on multiplying across by $\boldsymbol{P}[\,Fe\,]$ we get

$$\boldsymbol{P}[\,(\,Fe\,) \cap (\,Front\,)\,] \;=\; \boldsymbol{P}[\,Fe\,]\,.\,\boldsymbol{P}[\,Front \mid Fe\,]$$

The exact same (counting) argument holds for any two events A and B taken from any sample space of equally likely outcomes so we have

FACT : For any two events A and B **the conditional probability of** B **given** A, which is denoted by $\boldsymbol{P}[B|A]$, is given by

$$\boldsymbol{P}[B|A] = \frac{\boldsymbol{P}(B \cap A)}{\boldsymbol{P}(A)}.$$

This formula is often written in the form

$$\boldsymbol{P}(B \cap A) = \boldsymbol{P}(A)\,\boldsymbol{P}[B|A]$$

and using the fact that $B \cap A = A \cap B$ we can write

$$\boxed{\boldsymbol{P}(A \cap B) = \boldsymbol{P}(A)\,\boldsymbol{P}[B|A]}$$

which is the form in which Conditional Probability is most often applied. This latter formula you might find easier to remember if you call A the 1^{st}-event and call B the 2^{nd}-event, then the formula is:

$$\boxed{\boldsymbol{P}[\,(1^{st}) \cap (2^{nd})\,] = \boldsymbol{P}[\,(1^{st})\,]\,.\,\boldsymbol{P}[\,(2^{nd}) \mid (1^{st})\,]}$$

Example [1] : If two cards are drawn (one after the other without replacement) from a standard deck (of 52) what is the probability that **the first is an Ace and the second is a Jack**?

Solution : It is convenient to put

$$\begin{aligned} A_1 &= \text{the event that the first card drawn is an Ace} \\ \text{and} \\ J_2 &= \text{the event that the second card drawn is a Jack} \end{aligned}$$

so that the probability we require is $\boldsymbol{P}(A_1 \cap J_2)$ which we calculate as follows:

$$\boldsymbol{P}(A_1 \cap J_2) = \boldsymbol{P}(A_1)\,\boldsymbol{P}[J_2|A_1]$$

where $\boldsymbol{P}(A_1) = 4/52$ (because there are 4 Aces and a total of 52 cards) while $\boldsymbol{P}[J_2|A_1] = 4/51$ (because there are 4 Jacks and a total of 51 cards remaining given that the first card drawn was an Ace). Therefore,

$$\boldsymbol{P}(A_1 \cap J_2) = \boldsymbol{P}(A_1)\,\boldsymbol{P}[J_2|A_1] = \frac{4}{52} \cdot \frac{4}{51}$$

Example [2] : We modify Example 1 in the following way. A pair of cards is drawn (without replacement) from a standard deck, what is the probability that **the pair drawn consist of an Ace and a Jack**?

Solution : To facilitate computation we put

$$\begin{aligned} A_i &= \text{the event that the } i^{\text{th}} \text{ card drawn is an Ace} \\ J_i &= \text{the event that the } i^{\text{th}} \text{ card drawn is a Jack} \\ \text{and} \\ E &= \text{the event that the pair drawn is an Ace and a Jack.} \end{aligned}$$

We can think of the pair being drawn one at a time, so the event E occurs if and only if we draw **EITHER** (an Ace first and then a Jack) **OR** (a Jack first and then an Ace). Therefore, the event E is given by the **disjoint union**

$$E = (A_1 \cap J_2) \cup (J_1 \cap A_2).$$

Using the result in Example 2, we calculate

$$\begin{aligned} \boldsymbol{P}(E) &= \boldsymbol{P}[(A_1 \cap J_2) \cup (J_1 \cap A_2)] \\ &= \boldsymbol{P}(A_1 \cap J_2) + \boldsymbol{P}(J_1 \cap A_2) && \ldots \text{ disjoint union} \\ &= \boldsymbol{P}(A_1)\,\boldsymbol{P}[J_2|A_1] + \boldsymbol{P}(J_1)\,\boldsymbol{P}[A_2|J_1] \\ &= \frac{4}{52} \cdot \frac{4}{51} + \frac{4}{52} \cdot \frac{4}{51} && \ldots \text{ Example 1.} \end{aligned}$$

Example [3] : In a certain small college students are enroled in one of three Faculties: Arts, Business or Science. The percentage of students in each Faculty is: Arts (50%), Business (30%) and Science (20%). If 40% of Arts students study Mathematics, 60% of Business students study Mathematics and 90% of Science students study Mathematics, what is the probability that a student picked at random from this college studies Mathematics?

Solution : Here we take the sample space to be

$$\Omega = \{\text{all students who are enroled in this college}\}$$

and define the following events obtained by picking a student, at random, from Ω:

$$\begin{aligned} A &= \{\text{the student is enroled in Arts}\} \\ B &= \{\text{the student is enroled in Business}\} \\ Sc &= \{\text{the student is enroled in Science}\} \\ M &= \{\text{the student studies Mathematics}\}. \end{aligned}$$

Accordingly, we have the **disjoint union** $\Omega = A \cup B \cup Sc$ and, therefore,

$$\begin{aligned} M &= \Omega \cap M \\ &= (A \cup B \cup Sc) \cap M \\ &= (A \cap M) \cup (B \cap M) \cup (Sc \cap M) \end{aligned}$$

Remark: You are strongly advised to present the information gathered thus far by means of a Venn diagram.

Since this union is disjoint, it follows that

$$\begin{aligned} \mathbb{P}[M] &= \mathbb{P}[A \cap M] + \mathbb{P}[B \cap M] + \mathbb{P}[Sc \cap M] \\ &= \mathbb{P}[A]\,\mathbb{P}[M \mid A] + \mathbb{P}[B]\,\mathbb{P}[M \mid B] + \mathbb{P}[Sc]\,\mathbb{P}[M \mid Sc] \\ &= (0.5).(0.4) + (0.3).(0.6) + (0.2).(0.9) \\ &= 0.56 \end{aligned}$$

Example [4] : Three dice (Red, Blue and **Green**) are thrown. What is the probability that the Red die shows an even number and that the SUM of the numbers shown on all three dice is eight?

Solution : Let us write $R =$ Red, $B =$ Blue, $\mathbf{G} =$ Green and let $S =$ Sum. We wish to calculate

$$\mathbb{P}\,[\,(R = \text{even}) \cap (S = 8)\,].$$

To do this we partition the event $R = \text{even}$ according to whether

$$R = 2 \quad \text{or} \quad R = 4 \quad \text{or} \quad R = 6$$

and, thus, we have:

$$\mathbb{P}\,[(R = \text{even}) \cap (S = 8)]$$

$$= \mathbb{P}\,[\,\{\,(\mathbf{R} = \mathbf{2}) \cup (\mathbf{R} = \mathbf{4}) \cup (\mathbf{R} = \mathbf{6})\,\} \cap (\mathbf{S} = \mathbf{8})\,]$$

$$= \mathbb{P}\,[\,\{\,(\mathbf{R} = \mathbf{2}) \cap (\mathbf{S} = \mathbf{8})\,\} \cup \{\,(\mathbf{R} = \mathbf{4}) \cap (\mathbf{S} = \mathbf{8})\,\} \cup \{\,(\mathbf{R} = \mathbf{6}) \cap (\mathbf{S} = \mathbf{8})\,\}\,]$$

$$= \mathbb{P}\,[\,(R = 2) \cap (S = 8)\,] + \mathbb{P}\,[\,(R = 4) \cap (S = 8)\,] + \mathbb{P}\,[\,(R = 6) \cap (S = 8)\,]$$

$$= \mathbb{P}\,[\,R = 2\,] \,.\, \mathbb{P}\,[\,S = 8 \,|\, R = 2] + \mathbb{P}\,[\,R = 4\,] \,.\, \mathbb{P}\,[\,S = 8 \,|\, R = 4] + \mathbb{P}\,[\,R = 6\,] \,.\, \mathbb{P}\,[\,S = 8 \,|\, R = 6]$$

$$= \mathbb{P}\,[\,R = 2\,] \,.\, \mathbb{P}\,[\,B + \mathbf{G} = 6] + \mathbb{P}\,[\,R = 4\,] \,.\, \mathbb{P}\,[\,B + \mathbf{G} = 4] + \mathbb{P}\,[\,R = 6\,] \,.\, \mathbb{P}\,[\,B + \mathbf{G} = 2]$$

$$= \frac{1}{6}\,\mathbb{P}\,[\,(B,\,\mathbf{G}) \in \{(1,\,\mathbf{5}),\,(2,\,\mathbf{4}),\,(3,\,\mathbf{3}),\,(4,\,\mathbf{2}),\,(5,\,\mathbf{1})\}\,] + \frac{1}{6}\,\mathbb{P}\,[\,(B,\,\mathbf{G}) \in \{(1,\,\mathbf{3}),\,(2,\,\mathbf{2}),\,(3,\,\mathbf{1})\}\,] + \frac{1}{6}\,\mathbb{P}\,[\,(B,\,\mathbf{G}) \in \{(1,\,\mathbf{1})\}\,]$$

$$= \frac{1}{6} \cdot \frac{5}{36} + \frac{1}{6} \cdot \frac{3}{36} + \frac{1}{6} \cdot \frac{1}{36}$$

$$= \frac{1}{24}$$

Independent Events : In everyday life we say that

the event E is **independent** of the event F	if and only if	the occurrence of F does **not** affect the occurrence of E

For example you might say that the event

$$\text{E} = \{\text{I will win the next lottery jackpot}\}$$

is independent of the event

$$\text{F} = \{\text{I will watch television tonight}\}.$$

However, it is most likely you would say that the event

$$\text{E} = \{\text{I will win the next lottery jackpot}\}$$

is **not** independent of the event

$$\text{F} = \{\text{I will buy a lottery ticket before the next draw}\}.$$

Mathematically, we can write this as:

$$\boxed{E \text{ is independent of } F} \iff \boxed{\boldsymbol{P}[\,E \mid F\,] = \boldsymbol{P}\,[\,E\,]}$$

If we substitute this into the main formula

$$\boxed{\boldsymbol{P}(F \cap E) = \boldsymbol{P}(F)\,\boldsymbol{P}[E|F]}$$

we see that

$$E \text{ is independent of } F \iff \boldsymbol{P}(F \cap E) = \boldsymbol{P}(F)\,\boldsymbol{P}(E).$$

Now, look more closely at the last line and you will see the symmetry (in E and F) on the right hand side of '$\iff$'. Thus, E is independent of F if and only if F is independent of E and in either case we simply say that E and F are independent. Because of its importance we record this as

FACT:

$$\boxed{\text{Events } E \text{ and } F \text{ are independent}} \iff \boxed{\boldsymbol{P}(E \cap F) = \boldsymbol{P}(E)\,\boldsymbol{P}(F)}$$

Example [5] : Mrs. Murphy has two children what is the probability that one is a girl and the other is a boy?

Solution : We solve this in two ways.

(i) **Method 1:** We list the sample space

$$S = \{GG, GB, BG, BB\}$$

where, for example, GB means Mrs. Murphy gave birth to a girl first and then a boy. It seems reasonable that these four outcomes are equally likely so each has probability 1/4. Therefore,

$$\begin{aligned} \boldsymbol{P}(\text{one is a girl and the other is a boy}) &= \boldsymbol{P}(\{GB, BG\}) \\ &= \boldsymbol{P}(GB) + \boldsymbol{P}(BG) \\ &= \frac{1}{4} + \frac{1}{4}. \end{aligned}$$

(ii) **Method 2:** We use the same notation as in Solution 1 but, without considering the full sample space, we can say immediately that

$$\begin{aligned} \boldsymbol{P}(\text{one is a girl and the other is a boy}) &= \boldsymbol{P}(\{GB, BG\}) \\ &= \boldsymbol{P}(GB) + \boldsymbol{P}(BG) \end{aligned}$$

Now there is no reason for us to believe that the sex of Mrs. Murphy's second child is in any way determined by that of her first. Thus, having a boy second is **independent** of having a girl first and vice versa. Accordingly,

$$\boldsymbol{P}(GB) = \boldsymbol{P}(G).\boldsymbol{P}(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

and similarly $\boldsymbol{P}(BG)$ =1/4 from which the answer follows.

Example [6] : Mrs. Murphy has two children. If we are told that one is a boy what is the probability that the other is a girl?

Solution : Let us put

$$\begin{aligned} \tilde{G} &= \{\text{one of the pair is a girl}\} \\ \text{and} \quad \tilde{B} &= \{\text{one of the pair is a boy}\}, \end{aligned}$$

Then the probability we require is

$$\begin{aligned} \boldsymbol{P}[\tilde{G}|\tilde{B}] &= \frac{\boldsymbol{P}(\tilde{G} \cap \tilde{B})}{\boldsymbol{P}(\tilde{B})} \\ &= \frac{\boldsymbol{P}(\{GB,\, BG\})}{\boldsymbol{P}(\{BB,\, BG,\, GB\})} \\ &= \frac{1/4 + 1/4}{1/4 + 1/4 + 1/4} = \frac{2}{3} \end{aligned}$$

If you prefer, you could solve this problem by listing **the relevant sample space of equally likely outcomes**. Since we're given that one of the children is a boy, this space is

$$S^* = \{BB^*,\, BG^*,\, GB^*\}$$

Where the '$*$' is to avoid confusion with the previous sample space of this type. Now there are only **three** equally likely outcomes so each has probability $= 1/3$. Therefore,

$$\boldsymbol{P}[\tilde{G}|\tilde{B}] = \boldsymbol{P}(\{BG^*,\, GB^*\}) = \frac{1}{3} + \frac{1}{3}$$

Example [7] (Optional) : Mrs. Murphy has two children. A visitor calls to her home and knocks on the door. If the door is opened by Mrs. Murphy's son what is the probability that Mrs. Murphy's other child is a daughter?

Solution : A convenient sample space here is

$$S = \{\, BBb,\, BGb,\, GBb,\, GGg,\, GBg,\, BGg \,\}$$

where, for example, BGb means that Mrs. Murphy gave birth to a boy first **and** then a girl **and** that a boy opened the door. We can calculate the probabilities of these outcomes as follows (let's just take the first two):

$$
\begin{aligned}
\boldsymbol{P}(BBb) &= \boldsymbol{P}((B_1B_2)b) \qquad \begin{cases} B_1 = \text{first child is a boy} \\ B_2 = \text{second child is a boy} \end{cases} \\
&= \boldsymbol{P}(B_1B_2)\boldsymbol{P}[b|B_1B_2] \\
&= \boldsymbol{P}(B_1)\boldsymbol{P}[B_2|B_1]\boldsymbol{P}[b|B_1B_2] \\
&= \frac{1}{2}\cdot\frac{1}{2}\cdot 1 \qquad = \frac{1}{4}.
\end{aligned}
$$

and

$$
\begin{aligned}
\boldsymbol{P}(BGb) &= \boldsymbol{P}((B_1G_2)b) \qquad \begin{cases} B_1 = \text{first child is a boy} \\ G_2 = \text{second child is a girl} \end{cases} \\
&= \boldsymbol{P}(B_1G_2)\boldsymbol{P}[b|B_1G_2] \\
&= \boldsymbol{P}(B_1)\boldsymbol{P}[G_2|B_1]\boldsymbol{P}[b|B_1G_2] \\
&= \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} \qquad = \frac{1}{8}.
\end{aligned}
$$

We list all the probabilities as follows:

	BBb	BGb	GBb	GGg	GBg	BGg
$\boldsymbol{P}$	1/4	1/8	1/8	1/4	1/8	1/8

For the sake of clarity we put

$$\tilde{G} = \{\text{one of the pair is a girl}\}$$

and

$$b = \{\text{a boy (Mrs. Murphy's son) opens the door}\}.$$

Thus, the probability we require is:

$$
\begin{aligned}
\boldsymbol{P}[\tilde{G}|b] &= \frac{\boldsymbol{P}(\tilde{G}\cap b)}{\boldsymbol{P}(b)} \\
&= \frac{\boldsymbol{P}(\{BGb,\ GBb\})}{\boldsymbol{P}(\{BBb,\ BGb,\ GBb\})} \\
&= \frac{1/8+1/8}{1/4+1/8+1/8} = \frac{1}{2}
\end{aligned}
$$

As an alternative method, you might like to construct **a sample space of equally likely outcomes** which is appropriate for this example.

Example [8] : Two fair dice (a RED and a BLUE, say) are rolled as a pair, n times one after the other. After each roll of the pair, the sum of the numbers showing is computed. Let

$$\boxed{S_j = \text{the sum of the numbers showing on the } j^{\text{th}} \text{ roll}}$$

and do the following:

(i) Calculated $\mathbb{P}[S_1 = 5]$.

(ii) Calculated $\mathbb{P}[S_1 \neq 7]$.

(iii) Calculated $\mathbb{P}[S_1$ is neither a 5 nor a 7$]$.

(iv) Calculated $\mathbb{P}[E_j]$, where

$$E_j = \boxed{\begin{array}{c}\text{the event that neither a 5 nor a 7 appears before} \\ \text{the } j^{\text{th}} \text{ roll, but a 5 appears on the } j^{\text{th}} \text{ roll}\end{array}}\,.$$

(v) Calculated $\mathbb{P}[A]$, where

$$A = \boxed{\begin{array}{c}\text{the event that in a succession of 10 rolls,} \\ \text{a sum} = 5 \text{ appears before a sum} = 7.\end{array}}\,.$$

Solution :

(i) $\mathbb{P}[S_1 = 5] = \mathbb{P}[\,(R,\, B) \in \{(1,\,4),\, (2,\,3),\, (3,\,2),\, (4,\,1)\,\}\,] = \dfrac{4}{36} = \dfrac{1}{9}\cdot$

(ii)
$$\begin{aligned}\mathbb{P}[S_1 \neq 7] &= 1 - \mathbb{P}[S_1 = 7] \\ &= 1 - \mathbb{P}[\,(R,\, B) \in \{(1,\,6),\, (2,\,5),\, \ldots\,,(5,\,2),\, (6,\,1)\,\}\,] \\ &= 1 - \frac{6}{36} = \frac{5}{6}\cdot\end{aligned}$$

(iii)
$$\begin{aligned}\mathbb{P}[S_1 \text{ is neither a 5 nor a 7}] &= \mathbb{P}[(S_1 \neq 5) \cap (S_1 \neq 7)] \\ &= \mathbb{P}[\overline{(S_1 = 5)} \cap \overline{(S_1 = 7)}] \\ &= \mathbb{P}[\overline{(S_1 = 5) \cup (S_1 = 7)}] \\ &= 1 - \mathbb{P}[(S_1 = 5) \cup (S_1 = 7)] \\ &= 1 - \{\,\mathbb{P}[S_1 = 5] + \mathbb{P}[S_1 = 7]\,\} \\ &= 1 - \left\{\frac{1}{9} + \frac{1}{6}\right\} = \frac{13}{18}\cdot\end{aligned}$$

(iv)
$$\begin{aligned}
\mathbb{P}[E_j] &= \mathbb{P}[(S_1 \notin \{5,\, 7\}) \cap (S_2 \notin \{5,\, 7\}) \cap \ldots \cap (S_{j-1} \notin \{5,\, 7\}) \cap (S_j = 5)\,] \\
&= \mathbb{P}[S_1 \notin \{5,\, 7\}] \times \mathbb{P}[S_2 \notin \{5,\, 7\}] \times \ \ldots \ \times \mathbb{P}[S_{j-1} \notin \{5,\, 7\}] \times \mathbb{P}[S_j = 5] \\
&= \underbrace{\frac{13}{18} \times \frac{13}{18} \times \ \ldots \ \times \frac{13}{18}}_{(j-1) \text{ times}} \times \frac{1}{9} \\
&= \frac{1}{9}\left(\frac{13}{18}\right)^{(j-1)}.
\end{aligned}$$

(v) Observe that the event A in question can be written as the **disjoint union**

$$A = E_1 \cup E_2 \cup E_3 \cup \ldots \cup E_9 \cup E_{10}$$

and, therefore,

$$\begin{aligned}
\mathbb{P}[A] &= \mathbb{P}[E_1] + \mathbb{P}[E_2] + \mathbb{P}[E_3] + \ldots + \mathbb{P}[E_9] + \mathbb{P}[E_{10}] \\
&= \sum_{j=1}^{10} \mathbb{P}[E_j] \ = \ \sum_{j=1}^{10} \frac{1}{9}\left(\frac{13}{18}\right)^{(j-1)} \ = \ \frac{1}{9}\sum_{k=0}^{9}\left(\frac{13}{18}\right)^{k} \\
&= \frac{1}{9}\,\frac{\left[1 - \left(\frac{13}{18}\right)^{10}\right]}{\left[1 - \left(\frac{13}{18}\right)\right]} \ = \ 0.384556\ldots
\end{aligned}$$

Example [9] : Mary and John decide to play a sequence of three tennis matches against one another. Mary is regarded as being the better of the two players and the probability of her winning any one of the three matches is reckoned to be as follows:

- The probability that Mary will win the first match is 0.6 .
- The probability that Mary will win any match after the first is again 0.6 if she has won the previous match but drops to just 0.5 if she has lost the previous match.

Now, determine the probability that Mary will win exactly two of the three matches.

Solution : The event (call it E) that Mary wins exactly two of the three matches can represented by:

$$E = \{ M_1 M_2 J_3,\ M_1 J_2 M_3,\ J_1 M_2 M_3 \}$$

where, for example, $M_1 M_2 J_3$ means that (Mary wins the first match) and (Mary wins the second match) and (John wins the third match). Thus

$$\begin{aligned}
\mathbb{P}[E] &= \mathbb{P}[M_1 M_2 J_3] + \mathbb{P}[M_1 J_2 M_3] + \mathbb{P}[J_1 M_2 M_3] \\
&= \mathbb{P}[M_1] \,.\, \mathbb{P}[M_2 \mid M_1] \,.\, \mathbb{P}[J_3 \mid M_1 M_2] \\
&\qquad + \mathbb{P}[M_1] \,.\, \mathbb{P}[J_2 \mid M_1] \,.\, \mathbb{P}[M_3 \mid M_1 J_2] \\
&\qquad\qquad + \mathbb{P}[J_1] \,.\, \mathbb{P}[M_2 \mid J_1] \,.\, \mathbb{P}[M_3 \mid J_1 M_2] \\
&= (0.6) \,.\, (0.6) \,.\, (0.4) \\
&\qquad + (0.6) \,.\, (0.4) \,.\, (0.5) \\
&\qquad\qquad + (0.4) \,.\, (0.5) \,.\, (0.6) \\
&= 0.384
\end{aligned}$$

A CURIOUS OBSERVATION : Those of you who have looked closely at the basic formula (for conditional probability)

$$\boxed{\boldsymbol{P}(A \cap E) = \boldsymbol{P}(A)\, \boldsymbol{P}[E|A]}$$

will have noticed the following

$$\begin{aligned}
\boldsymbol{P}(A)\, \boldsymbol{P}[E|A] &= \boldsymbol{P}(A \cap E) \\
&= \boldsymbol{P}(E \cap A) \\
&= \boldsymbol{P}(E)\, \boldsymbol{P}[A|E].
\end{aligned}$$

That is,

$$\boldsymbol{P}(A)\,\boldsymbol{P}[E|A] \;=\; \boldsymbol{P}(E)\,\boldsymbol{P}[A|E]$$

and, therefore,

$$\boxed{\boldsymbol{P}[E|A] = \frac{\boldsymbol{P}(E)\,\boldsymbol{P}[A|E]}{\boldsymbol{P}(A)}}\,.$$

Now, what is curious or interesting about this latter formula is that (without its aid) it is sometimes very difficult to calculate $\boldsymbol{P}[E|A]$ though it may be very easy to calculate $\boldsymbol{P}[A|E]$ and, indeed, it is often the difficult one that is of most interest to us. Consider the following example.

Example [10] : In a certain population it is estimated that 45% of the people smoke tobacco while it is known from medical statistics that 6% of this population die of lung cancer. Furthermore, it is also known from medical statistics that 80% of those people who die of lung cancer were smokers of tobacco. The question that may be of interest to a person in this population is:

> If I am a tobacco smoker, what is the probability that I will die of lung cancer?

Solution : Let us fix the following notation:

$$\begin{aligned} TS &= \{\,\text{tobacco smokers}\,\} \\ \text{and} \qquad & \\ DL &= \{\,\text{die from lung cancer}\,\}. \end{aligned}$$

The following information is known to us:

$$\begin{aligned} \boldsymbol{P}(TS) &= 45/100 &= 0.45 \\ \boldsymbol{P}(DL) &= 6/100 &= 0.06 \\ \boldsymbol{P}[TS|DL] &= 80/100 &= 0.8 \end{aligned}$$

The probability we require is

$$\boldsymbol{P}[DL|TS] = \frac{\boldsymbol{P}(DL)\,\boldsymbol{P}[TS|DL]}{\boldsymbol{P}(TS)} = \frac{(0.06)\,(0.8)}{0.45} = 0.10666\ldots \simeq 10.7\%$$

Partitioning The Sample Space : Our curious formula

$$\boldsymbol{P}[E|A] = \frac{\boldsymbol{P}(E)\,\boldsymbol{P}[A|E]}{\boldsymbol{P}(A)}$$

allows of a modest, though useful, generalization when the sample space S is **partitioned** into a disjoint union by the events (subsets) E_1, E_2, ... E_m. That is, when

$$S = E_1 \cup E_2 \cup \ldots \cup E_m \quad (\textbf{ disjoint union}).$$

In this situation, any event A can be written as a **disjoint union**

$$A = S \cap A = (E_1 \cap A) \cup (E_2 \cap A) \cup \ldots \cup (E_m \cap A).$$

Diagram.

Accordingly,

$$\begin{aligned} \boldsymbol{P}(A) &= \boldsymbol{P}(E_1 \cap A) \cup (E_2 \cap A) \cup \ldots \cup (E_m \cap A) \quad (\textbf{ disjoint union}) \\ &= \boldsymbol{P}(E_1 \cap A) + \boldsymbol{P}(E_2 \cap A) + \cdots + \boldsymbol{P}(E_m \cap A) \\ &= \boldsymbol{P}(E_1)\,\boldsymbol{P}[A|E_1] + \boldsymbol{P}(E_2)\,\boldsymbol{P}[A|E_2] + \cdots + \boldsymbol{P}(E_m)\,\boldsymbol{P}[A|E_m] \end{aligned}$$

and now our curious formula

$$\boldsymbol{P}[E|A] = \frac{\boldsymbol{P}(E)\,\boldsymbol{P}[A|E]}{\boldsymbol{P}(A)}$$

(with the event E replaced by any partitioning event E_i) can be written in the form:

$$\boldsymbol{P}[E_i|A] = \frac{\boldsymbol{P}(E_i)\,\boldsymbol{P}[A|E_i]}{\boldsymbol{P}(E_1)\,\boldsymbol{P}[A|E_1] + \boldsymbol{P}(E_2)\,\boldsymbol{P}[A|E_2] + \cdots + \boldsymbol{P}(E_m)\,\boldsymbol{P}[A|E_m]}$$

Thus we have established:

Bayes' Theorem: If A is any event and if E_1, E_2, ... E_m is any **partition** of the sample space, then

$$\boldsymbol{P}[E_i|A] = \frac{\boldsymbol{P}(E_i)\,\boldsymbol{P}[A|E_i]}{\boldsymbol{P}(E_1)\,\boldsymbol{P}[A|E_1] + \boldsymbol{P}(E_2)\,\boldsymbol{P}[A|E_2] + \cdots + \boldsymbol{P}(E_m)\,\boldsymbol{P}[A|E_m]}$$

Example [11] : (See, *New England Journal of Medicine*, 1974, Vol. 291, p1115) It is estimated that in a diagnosis of renovascular hypertension the intravenous pyelogram gives a **positive indication** in 100% of patients **who have the disease**, but also gives a **positive indication** in 10% of those patients **who do not have the disease**. If in a certain population it is estimated that 10% actually have the disease, find the probability that a person from this population who tests positive for renovascular hypertension does indeed have the disease.

Solution : Let us fix the following notation:

$$P = \{\text{people who test positive}\}$$

and

$$D = \{\text{people who have the disease}\}.$$

As usual $\bar{P}$ and $\bar{D}$ will denote the **complements** of P and D, respectively. The following information is known to us:

$$\boldsymbol{P}[P|D] = 1, \quad \boldsymbol{P}[P|\bar{D}] = 0.1 \quad \text{and} \quad \boldsymbol{P}(D) = 0.1$$

We **partition the sample space** according to

$$S = D \cup \bar{D} \quad \textbf{diojoint union}$$

and apply **Bayes' Theorem** as follows:

$$
\begin{aligned}
\boldsymbol{P}[D|P] &= \frac{\boldsymbol{P}(D)\,\boldsymbol{P}[P|D]}{\boldsymbol{P}(D)\,\boldsymbol{P}[P|D] + \boldsymbol{P}(\bar{D})\,\boldsymbol{P}[P|\bar{D}]} \\
&= \frac{(0.1)\,(\,1\,)}{(0.1)\,(\,1\,) + (0.9)\,(0.1)} \\
&= \frac{0.1}{0.19} \;=\; 0.5263\ldots
\end{aligned}
$$