1.7 Conditional Probabiliity

1.7.1 Definition

Let the event $B \in \mathcal{F}$ be such that $\mathrm{P}(B) > 0$. For any event $A \in \mathcal{F}$ the *conditional probability* of A given that B has occurred, denoted by $\mathrm{P}(A|B)$, is defined as

$$\boxed{\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}, \qquad (\mathrm{P}(B) > 0)} \tag{1.24}$$

(note that $\mathrm{P}(B|B) = 1$ as it should).

We read $\mathrm{P}(A|B)$ as 'the probability of A given B' or 'the probability of A conditioned on B'.

Theorem

If $B \in \mathcal{F}$ has $\mathrm{P}(B) > 0$, and $\mathrm{Q}(A) = \mathrm{P}(A|B)$, then $(\mathcal{S}, \mathcal{F}, \mathrm{Q})$ is a probability space.

Proof We have that

(i) $\mathrm{Q}(A) \geq 0$ for all $A \in \mathcal{F}$ (from (1.24) and probability axiom (a)).

(ii) $\mathrm{Q}(\mathcal{S}) = \mathrm{P}(\mathcal{S} \cap B)/\mathrm{P}(B) = \mathrm{P}(B)/\mathrm{P}(B) = 1$.

(iii) If $A_1, A_2, ...$ are mutually exclusive events in $\mathcal{F}$, so are $A_1 \cap B, A_2 \cap B, ..$: then

$$\begin{aligned}
\mathrm{Q}(\bigcup_i A_i) &= \frac{1}{\mathrm{P}(B)} \mathrm{P}((\bigcup_i A_i) \bigcap B) && \text{by definition of } Q \\
&= \frac{1}{\mathrm{P}(B)} \mathrm{P}(\bigcup_i (A_i \bigcap B)) && \text{using distributive law (1.6)} \\
&= \frac{1}{\mathrm{P}(B)} \sum_i \mathrm{P}(A_i \cap B) && \text{by probability axiom (1.10)} \\
&= \sum_i \mathrm{Q}(A_i).
\end{aligned}$$

Thus Q satisfies the three probability axioms and $(\mathcal{S}, \mathcal{F}, \mathrm{Q})$ is a probability space. ♢

It follows that all the results for $\mathrm{P}(A)$ carry over to $\mathrm{P}(A|B)$: e.g.

$$\mathrm{P}(A \cup B|C) = \mathrm{P}(A|C) + \mathrm{P}(B|C) - \mathrm{P}(A \cap B|C).$$

We often think of the conditioning event C as a *reduced sample space.*

1.7.2 Multiplication Rule

It follows from the definition (1.24) that, provided all stated conditioning events have positive probability,

$$\boxed{\mathrm{P}(A \cap B) = \mathrm{P}(A|B).\mathrm{P}(B) = \mathrm{P}(B|A).\mathrm{P}(A)} \tag{1.25}$$

and more generally, by repeated application of this result, that

$$\boxed{\mathrm{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathrm{P}(A_1).\mathrm{P}(A_2|A_1).\mathrm{P}(A_3|A_1 \cap A_2)...\mathrm{P}(A_n|A_1 \cap A_2 \cap ... \cap A_{n-1}).} \tag{1.26}$$

(In fact, since $A_1 \supseteq A_1 \cap A_2 \supseteq ... \supseteq A_1 \cap A_2 \cap ... \cap A_{n-1}$, it is sufficient, in view of (1.13), to require that this last event have positive probability.)

1.7.3 Law of Total Probability and Bayes' Theorem

Let $H_1, H_2, ..., H_n \in \mathcal{F}$ be a set of mutually exclusive and exhaustive events (i.e. a *partition* of the sample space $\mathcal{S}$): thus $\bigcup_{j=1}^{n} H_j = \mathcal{S}$ and $H_i \cap H_j = \emptyset, \quad i \neq j$. Suppose also that they are all possible events, i.e. $\mathrm{P}(H_j) > 0$ for $j = 1, ..., n$. Then for any event $A \in \mathcal{F}$

$$\begin{aligned}
\mathrm{P}(A) = \mathrm{P}(A \cap \mathcal{S}) &= \mathrm{P}\left(A \cap \left(\bigcup_{j=1}^{n} H_j\right)\right) \\
&= \mathrm{P}\left(\bigcup_{j=1}^{n} (A \cap H_j)\right) \quad \text{by (1.6)},\{A \cap H_j\} \text{ m.e.} \\
&= \sum_{j=1}^{n} \mathrm{P}(A \cap H_j) \quad \text{by (1.10)}.
\end{aligned}$$

Invoking (1.25), this becomes

$$\boxed{\mathrm{P}(A) = \sum_{j=1}^{n} \mathrm{P}(A|H_j)\mathrm{P}(H_j). \qquad (1.27)}$$

This is the *law of total probability* (also known as the completeness theorem or partition rule): it is of great importance in providing the basis for the solution of many problems by *conditioning*.

Furthermore, for any event A with $\mathrm{P}(A) > 0$, we have (using (1.24), (1.25) and (1.27))

$$\boxed{\mathrm{P}(H_k|A) = \frac{\mathrm{P}(H_k \cap A)}{\mathrm{P}(A)} = \frac{\mathrm{P}(A|H_k)\mathrm{P}(H_k)}{\sum_{j=1}^{n} \mathrm{P}(A|H_j)\mathrm{P}(H_j)}. \qquad (1.28)}$$

This is the famous *Bayes' Rule.*

1.7.4 Independence

Two events $A, B \in \mathcal{F}$ are said to be independent if and only if

$$\boxed{\mathrm{P}(A \cap B) = \mathrm{P}(A).\mathrm{P}(B). \qquad (1.29)}$$

Theorem

If A and $B(\in \mathcal{F})$ are independent events, then

$$\begin{aligned}
\mathrm{P}(A|B) &= \mathrm{P}(A) \text{ if } \mathrm{P}(B) > 0, \\
\mathrm{P}(B|A) &= \mathrm{P}(B) \text{ if } \mathrm{P}(A) > 0; \\
& A \text{ and } \overline{B} \text{ are independent events,} \\
& \overline{A} \text{ and } B \text{ are independent events,} \\
& \overline{A} \text{ and } \overline{B} \text{ are independent events.}
\end{aligned}$$

Proof To prove the first result:

$$\begin{aligned}
\mathrm{P}(A|B) &= \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \qquad \text{if } \mathrm{P}(B) > 0 \text{ [(1.24)]} \\
&= \frac{\mathrm{P}(A).\mathrm{P}(B)}{\mathrm{P}(B)} \qquad \text{since } A, B \text{ are independent [by (1.29)]} \\
&= \mathrm{P}(A).
\end{aligned}$$

The second result is proved similarly.

To prove the third result, we observe that

$$A = (A \cap B) \cup (A \cap \overline{B})$$

— the union of two mutually exclusive events. So

$$\begin{aligned} \mathrm{P}(A) &= \mathrm{P}(A \cap B) + \mathrm{P}(A \cap \overline{B}) \\ &= \mathrm{P}(A).\mathrm{P}(B) + \mathrm{P}(A \cap \overline{B}) \quad \text{by (1.29)} \\ \text{Hence } \mathrm{P}(A \cap \overline{B}) &= \mathrm{P}(A) - \mathrm{P}(A).\mathrm{P}(B) = \mathrm{P}(A).\mathrm{P}(\overline{B}) \quad \text{by (1.12).} \end{aligned}$$

So A and $\overline{B}$ are independent events.
The fourth and fifth results are proved similarly.

More generally, the events $A_1, A_2, ..., A_n$ are said to be *mutually independent* or *completely independent* if and only if

$$\boxed{\begin{aligned} \mathrm{P}(A_i \cap A_j) &= \mathrm{P}(A_i).\mathrm{P}(A_j), \qquad i \neq j, \\ \mathrm{P}(A_i \cap A_j \cap A_k) &= \mathrm{P}(A_i).\mathrm{P}(A_j).\mathrm{P}(A_k), \qquad i \neq j \neq k, \\ &\ldots\ldots\ldots\ldots \\ \mathrm{P}(A_1 \cap A_2 \cap ... \cap A_n) &= \mathrm{P}(A_1).\mathrm{P}(A_2)...\mathrm{P}(A_n). \end{aligned}} \qquad (1.30)$$

Hence pairwise independence does not imply complete independence.

1.8 Examples

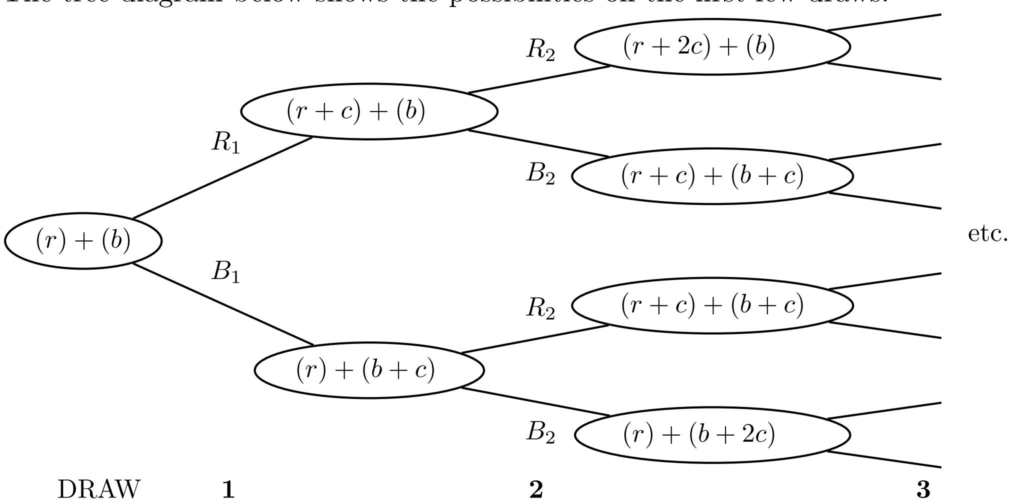
Many probability examples involving events occurring at specified times or trials can be solved by means of the multiplication rule for conditional or independent events, defining A_i to be the relevant event at time or trial i.

Example 1.6 *Polya's urn scheme.*
[The following model was proposed for the description of contagious phenomena where the occurrence of an event increases its probability of occurrence in the future.]
An urn initially contains r red balls and b black balls. A ball is drawn at random and replaced together with c balls of its own colour. This procedure is repeated many times. What is the probability that a red ball is obtained on the i^{th} draw?

Solution Introduce the events

R_i: *red* ball obtained on the i^{th} draw.
B_i: *black* ball obtained on i^{th} draw.

The tree diagram below shows the possibilities on the first few draws.



On the first draw,

$$\mathrm{P}(R_1) = \frac{r}{r+b}, \qquad \mathrm{P}(B_1) = \frac{b}{r+b}.$$

On the second draw:

$$\begin{aligned}
\mathrm{P}(R_2) &= \mathrm{P}((R_1 \cap R_2) \cup (B_1 \cap R_2)) \\
&= \mathrm{P}(R_1 \cap R_2) + \mathrm{P}(B_1 \cap R_2) \qquad \text{(m.e.)} \\
&= \mathrm{P}(R_2|R_1)\mathrm{P}(R_1) + \mathrm{P}(R_2|B_1)\mathrm{P}(B_1) \\
&= \frac{r+c}{r+c+b}.\frac{r}{r+b} + \frac{r}{r+c+b}.\frac{b}{r+b} = \frac{r}{r+b}.
\end{aligned}$$

Hence

$$\mathrm{P}(B_2) = \frac{b}{r+b}.$$

On the third draw, we have:

$$\begin{aligned}
\mathrm{P}(R_3) &= \mathrm{P}(R_1 \cap R_2 \cap R_3) + \mathrm{P}(R_1 \cap B_2 \cap R_3) \\
&\quad + \mathrm{P}(B_1 \cap R_2 \cap R_3) + \mathrm{P}(B_1 \cap B_2 \cap R_3) \\
&= \mathrm{P}(R_3|R_1 \cap R_2)\mathrm{P}(R_1 \cap R_2) + \mathrm{P}(R_3|R_1 \cap B_2)\mathrm{P}(R_1 \cap B_2) \\
&\quad + \mathrm{P}(R_3|B_1 \cap R_2)\mathrm{P}(B_1 \cap R_2) + \mathrm{P}(R_3|B_1 \cap B_2)\mathrm{P}(B_1 \cap B_2)
\end{aligned}$$

Now

$$\mathrm{P}(R_3|R_1 \cap R_2) = \frac{r+2c}{r+2c+b}, \qquad \text{etc.}$$

and

$$\mathrm{P}(R_1 \cap R_2) = \mathrm{P}(R_2|R_1)\mathrm{P}(R_1) = \frac{r+c}{r+c+b}.\frac{r}{r+b}, \qquad \text{etc.}$$

from which we deduce that

$$\mathrm{P}(R_3) = \frac{r}{r+b}$$

and hence that

$$\mathrm{P}(B_3) = \frac{b}{r+b}.$$

This naturally leads us to the *conjecture*:

$$\mathrm{P}(R_i) = \frac{r}{r+b} \text{ for } i = 1, 2, 3, ... \ ?$$

This can be proved by induction, as follows. It has already been shown to be true for $i = 1, 2, 3$. Suppose it is true for $i = n$. Then

$$\mathrm{P}(R_{n+1}) = \mathrm{P}(R_{n+1}|R_1)\mathrm{P}(R_1) + \mathrm{P}(R_{n+1}|B_1)\mathrm{P}(B_1).$$

Now clearly R_{n+1} given R_1 is equivalent to starting with $(r + c)$ red balls and b black balls and obtaining a red on the nth draw, and from the above supposition it follows that

$$\mathrm{P}(R_{n+1}|R_1) = \frac{r+c}{r+c+b}.$$

Similarly

$$\mathrm{P}(R_{n+1}|B_1) = \frac{r}{r+c+b}.$$

So

$$\mathrm{P}(R_{n+1}) = \frac{r+c}{r+c+b}.\frac{r}{r+b} + \frac{r}{r+c+b}.\frac{b}{r+b} = \frac{r}{r+b}.$$

Hence, by induction, the conjecture is proved. ♢

Notes:

(i) Probabilities such as $\mathrm{P}(R_1 \cap R_2), \mathrm{P}(R_2|R_1), \mathrm{P}(R_1 \cap R_2 \cap R_3), \mathrm{P}(R_3|R_2)$ etc. can be calculated using the standard results and by enumerating the outcomes at draws $1, 2, 3$ etc.

(ii) This approach cannot be used for the calculation of other probabilities after n draws, when n is not small. One approach is to use *recurrence relations* – an approach which is taken up in the next section.

Example 1.7 *The Lift Problem*

A simple form of this problem is as follows:

A lift has three occupants A, B and C, and there are three possible floors (1, 2 and 3) at which they can get out. Assuming that each person acts independently of the others and that each person is equally likely to get out at each floor, calculate the probability that exactly one person will get out at each floor.

Solution We use a sequential approach. Let

F_i: one person gets off at floor i.

A_i: A gets off at floor i (events B_i and C_i are defined similarly).

Then the required probability is

$$\mathrm{P}(F_1 \cap F_2 \cap F_3) = \mathrm{P}(F_1)\mathrm{P}(F_2|F_1)\mathrm{P}(F_3|F_1 \cap F_2) \quad \text{by (1.26).}$$

Now

$$\begin{aligned} \mathrm{P}(F_1) &= \mathrm{P}((A_1 \cap \overline{B_1} \cap \overline{C_1}) \cup (\overline{A_1} \cap B_1 \cap \overline{C_1}) \cup (\overline{A_1} \cap \overline{B_1} \cap C_1)) \\ &= 3.(\tfrac{1}{3})(\tfrac{2}{3})^2 = \tfrac{4}{9} \quad \text{(invoking independence and (1.30))} \\ \mathrm{P}(F_2|F_1) &= 2.(\tfrac{1}{2})(\tfrac{1}{2}) = \tfrac{1}{2} \qquad \text{(by similar argument)} \\ \mathrm{P}(F_3|F_1 \cap F_2) &= 1 \end{aligned}$$

So the required probability is $\frac{4}{9}.\frac{1}{2}.1 = \frac{2}{9}$. ♢

Let's consider the generalisation of this problem to n persons and n floors. Let p_n denote the probability that exactly one person gets off at each floor. We have

$$\begin{aligned} p_n &= \mathrm{P}(F_1 \cap F_2 \cap \cdots \cap F_n) \\ &= \mathrm{P}(F_2 \cap F_3 \cap \cdots \cap F_n|F_1)\mathrm{P}(F_1) \quad \text{by (1.25)} \\ &= p_{n-1}\left\{n.\left(\frac{1}{n}\right).\left(\frac{n-1}{n}\right)^{n-1}\right\} \\ &= \left(\frac{n-1}{n}\right)^{n-1} p_{n-1}, \qquad n > 1. \end{aligned}$$

(The crucial step here is the recognition that, in view of the independence assumption, the conditional probability is identical to p_{n-1}). We note that $p_1 = 1$. The recurrence formula is easily solved:

$$\begin{aligned} p_n &= \left(\frac{n-1}{n}\right)^{n-1} p_{n-1} \\ &= \left(\frac{n-1}{n}\right)^{n-1}\left(\frac{n-2}{n-1}\right)^{n-2} p_{n-2} \\ &= \ldots\ldots\ldots\ldots\ldots\ldots \\ &= \left(\frac{n-1}{n}\right)^{n-1}\left(\frac{n-2}{n-1}\right)^{n-2} \cdots \left(\frac{1}{2}\right)^{1} p_1 \\ &= \frac{(n-1)!}{n^{n-1}} = \frac{n!}{n^n}, \qquad n \geq 1. \end{aligned}$$

In particular $p_3 = \frac{3!}{3^3} = \frac{2}{9}$ in agreement with our more restricted discussion at the start.

1.9 Conditioning and Recurrence Relations

The argument at the end of Example 1.6 involved conditioning on the result of the first draw, and then recognising the relationship between the conditional probabilities involved and the probability under discussion. Here is another example of this kind of argument.

Example 1.8

A random experiment has three outcomes, A, B and C, with probabilities p_A, p_B and p_C respectively, where $p_C = 1 - p_A - p_B$. What is the probability that, in independent performances of the experiment, A will occur before B?

Solution 1 (by decomposition into the union of m.e. events). The event

$$D: \qquad A \text{ occurs before } B$$

can occur in any of the following mutually exclusive ways: $A, CA, CCA, CCCA, ...$ So its probability is

$$\begin{aligned} \mathrm{P}(D) = p_A + p_C.p_A + p_C^2.p_A + ... &= p_A(1 + p_C + p_C^2 + ...) \\ &= p_A/(1 - p_C) \\ &= \frac{p_A}{p_A + p_B}. \end{aligned}$$

Solution 2 (by conditioning). Condition on the result of the first trial (A_1, B_1 or C_1). Thus

$$\begin{aligned} \mathrm{P}(D) &= \mathrm{P}(D|A_1)\mathrm{P}(A_1) + \mathrm{P}(D|B_1)\mathrm{P}(B_1) + \mathrm{P}(D|C_1)\mathrm{P}(C_1) \\ &= \mathrm{P}(D|A_1)p_A + \mathrm{P}(D|B_1)p_B + \mathrm{P}(D|C_1)p_C \end{aligned}$$

Now

$$\mathrm{P}(D|A_1) = 1; \qquad \mathrm{P}(D|B_1) = 0 \quad \text{(obviously)}$$

while

$$\mathrm{P}(D|C_1) = \mathrm{P}(D),$$

since in this case the problem after the first trial is exactly as at the start (in view of the independence of the trials). So we have

$$\mathrm{P}(D) = p_A + p_C.\mathrm{P}(D)$$

which solves to give

$$\mathrm{P}(D) = \frac{p_A}{1 - p_C} = \frac{p_A}{p_A + p_B}$$

as before. ♢

[There is a simple third method of solving the problem. Consider the critical trial at which the sequence of trials ends with either A or B. Then

$$\mathrm{P}(A|A \cup B) = \frac{\mathrm{P}(A \cap (A \cup B))}{\mathrm{P}(A \cup B)} = \frac{p_A}{p_A + p_B}$$

as before.]

Now we extend this idea to situations where the desired probability has an associated parameter, and conditioning leads to a recurrence formula.

Example 1.9

In a series of independent games, a player has probabilities $\frac{1}{3}$, $\frac{5}{12}$ and $\frac{1}{4}$ of scoring 0, 1 and 2 points respectively in each game. The scores are added; the series terminates when the player scores 0 in a game. Obtain a recurrence relation for p_n, the probability that the player obtains a final score of n points.

Solution We have here a discrete but infinite sample space: an outcome is a sequence of 1's and/or 2's ending with a 0, or just 0 on the first game. Clearly

$$\begin{aligned} p_0 &= \tfrac{1}{3} \\ p_1 &= \tfrac{5}{12} \cdot \tfrac{1}{3} = \tfrac{5}{36}. \end{aligned}$$

Conditioning on the result of the first game, we have (for $n \geq 2$)

$$\begin{aligned} p_n &= \mathrm{P}(n \text{ points in total}) \\ &= \mathrm{P}(n \text{ points in total}|1 \text{ on first game})\mathrm{P}(1 \text{ on first game}) \\ &\quad +\mathrm{P}(n \text{ points in total}|2 \text{ on first game})\mathrm{P}(2 \text{ on first game}) \\ &= \tfrac{5}{12}p_{n-1} + \tfrac{1}{4}p_{n-2}, \qquad n \geq 2. \end{aligned}$$

(Again, the independence assumption is essential when we come to re-interpret the conditional probabilities in terms of p_{n-1} and p_{n-2}.)

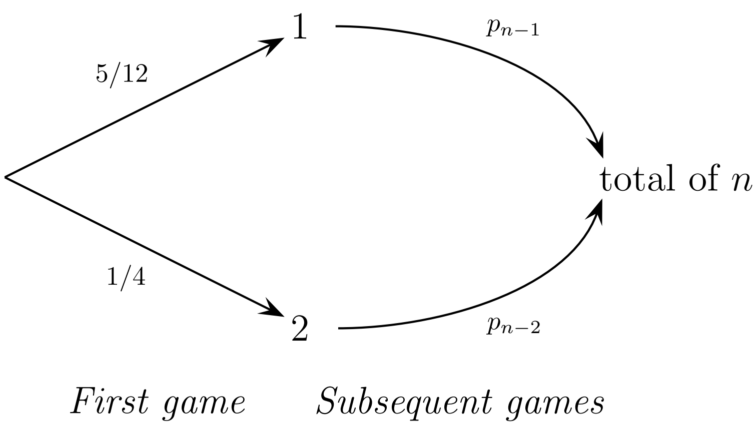
To get an expression for p_n, this has to be solved subject to the conditions

$$p_0 = \frac{1}{3}, \qquad p_1 = \frac{5}{36}.$$

The result (either from the theory of difference equations or by induction) is

$$p_n = \frac{3}{13}\left(\frac{3}{4}\right)^n + \frac{4}{39}\left(-\frac{1}{3}\right)^n, \qquad n \geq 0.$$

A diagram such as that given below is often helpful in writing down the recurrence formula required. ♢



Finally, we return to a problem considered earlier, and solve it by means of a recurrence formula.

Example 1.10 *The Matching Problem (Revisited)*

For a description of the problem, see §1.5.1 above. Let

$$\begin{array}{rl} B: & \text{no matches occur.} \\ A_1: & \text{there is a match in position 1} \end{array}$$

Let p_n denote the probability of no matches occurring when n numbered cards are randomly distributed over n similarly numbered positions. Then conditioning on what happens in the *first* position, we have

$$p_n = \mathrm{P}(B) = \mathrm{P}(B|A_1)\mathrm{P}(A_1) + \mathrm{P}(B|\overline{A_1})\mathrm{P}(\overline{A_1}).$$

Now clearly $\mathrm{P}(B|A_1) = 0$, so

$$p_n = \mathrm{P}(B|\overline{A_1}).\frac{n-1}{n}$$

The probability $\mathrm{P}(B|\overline{A_1})$ is the probability that no matches occur when $(n-1)$ cards (numbered 2 to n but with k, say, missing and replaced by 1) are randomly distributed over positions 2 to n. This can happen in two mutually exclusive ways:

(i) card 1 falls on a position other than k, and none of the other cards make a match;

(ii) card 1 falls on position k, and none of the other cards make a match.

We deduce that

$$\mathrm{P}(B|\overline{A_1}) = p_{n-1} + \frac{1}{n-1}p_{n-2}$$

(the first term being derived by temporarily regarding position k as being the 'matching position' for card 1). Hence we obtain the recurrence relation

$$p_n = \frac{n-1}{n}p_{n-1} + \frac{1}{n}p_{n-2}$$

i.e.

$$p_n - p_{n-1} = -\frac{1}{n}(p_{n-1} - p_{n-2}).$$

Now it is readily seen that

$$p_1 = 0; \qquad p_2 = \frac{1}{2},$$

so by repeated application of the recurrence relation we get

$$\begin{array}{rcll} p_3 - p_2 & = & -\dfrac{(p_2 - p_1)}{3} = -\dfrac{1}{3!} & \text{or } p_2 = \dfrac{1}{2!} - \dfrac{1}{3!} \\ p_4 - p_3 & = & \dfrac{(p_3 - p_2)}{4} = \dfrac{1}{4!} & \text{or } p_3 = \dfrac{1}{2!} - \dfrac{1}{3!} + \dfrac{1}{4!} \end{array}$$

and more generally

$$p_n = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots + \frac{(-1)^n}{n!}$$

– the result obtained previously. ♢

1.10 Further Examples

Example 1.11 *The 'Monty Hall' Problem*

A game show contestant is shown 3 doors, one of which conceals a valuable prize, while the other 2 are empty. The contestant is allowed to choose one door (note that, regardless of the choice made, at least one of the remaining doors is empty). The show host (who knows where the prize is) opens one of the remaining doors to show it empty (it is assumed that if the host has a choice of doors, he selects at random). The contestant is now given the opportunity to switch doors. Should the contestant switch?

Solution Number the contestant's chosen door '1', and the other doors '2' and '3'. Let

$$\begin{aligned} A_i: &\quad \text{the prize is behind door } i (i = 1, 2, 3) \\ D: &\quad \text{door 2 opened by host.} \end{aligned}$$

We assume $\mathrm{P}(A_1) = \mathrm{P}(A_2) = \mathrm{P}(A_3) = \frac{1}{3}$. Then

$$\mathrm{P}(D|A_1) = \frac{1}{2}, \quad \mathrm{P}(D|A_2) = 0, \quad \mathrm{P}(D|A_3) = 1.$$

Then by Bayes' Rule (1.28):

$$\mathrm{P}(A_3|D) = \frac{\frac{1}{3}\cdot 1}{\frac{1}{3}\cdot\frac{1}{2} + \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot 1} = \frac{2}{3}.$$

So the contestant *should* switch and will gain the prize with probability $\frac{2}{3}$. ♢

(*Comment:* This problem is called after the host in a US quiz show, and has given rise to considerable debate. The point which is often overlooked is that the host sometimes has the choice of *two* doors and sometimes *one*.)

Example 1.12 *Successive Heads*

A biased coin is such that the probability of getting a head in a single toss is p. Let v_n be the probability that two successive heads are *not* obtained in n tosses of the coin. Obtain a recurrence formula for v_n and verify that, in the case where $p = \frac{2}{3}$,

$$v_n = 2\left(\frac{2}{3}\right)^{n+1} + \left(-\frac{1}{3}\right)^{n+1}.$$

Solution Conditioning on the result of the first toss, we have (in an obvious notation)

$$\begin{aligned} v_n &= \mathrm{P}(\text{no pairs in } n \text{ tosses}) \\ &= \mathrm{P}(\text{no pairs in } n \text{ tosses}|H_1)\mathrm{P}(H_1) + \mathrm{P}(\text{no pairs in } n \text{ tosses}|T_1)\mathrm{P}(T_1) \\ &= p\mathrm{P}(\text{tail followed by no pairs (in } n-2 \text{ tosses)}) + (1-p)\mathrm{P}(\text{no pairs in } n-1 \text{ tosses}) \\ &= p(1-p)v_{n-2} + (1-p)v_{n-1} \qquad n \geq 2 \end{aligned}$$

We note that $v_0 = v_1 = 1$.

When $p = \frac{2}{3}$: $\quad v_n = \frac{2}{9}v_{n-2} + \frac{1}{3}v_{n-1}, \quad n \geq 2$. The given result may be proved by induction (or obtained directly from the theory of difference equations). ♢