Part I
Basic probability

1 Discrete outcomes

1.1 A uniform distribution

Lest men suspect your tale untrue,
Keep probability in view.

J. Gay (1685–1732), English poet

In this section we use the simplest (and historically the earliest) probabilistic model where there are a finite number m of possibilities (often called *outcomes*) and each of them has the same probability $1/m$. A collection A of k outcomes with $k \leq m$ is called an *event* and its probability $\mathbb{P}(A)$ is calculated as k/m:

$$\mathbb{P}(A) = \frac{\text{the number of outcomes in } A}{\text{the total number of outcomes}}. \tag{1.1}$$

An empty collection has probability zero and the whole collection one. This scheme looks deceptively simple: in reality, calculating the number of outcomes in a given event (or indeed, the total number of outcomes) may be tricky.

Problem 1.1 You and I play a coin-tossing game: if the coin falls heads I score one, if tails you score one. In the beginning, the score is zero. (i) What is the probability that after $2n$ throws our scores are equal? (ii) What is the probability that after $2n+1$ throws my score is three more than yours?

Solution The outcomes in (i) are all sequences $HHH\ldots H, THH\ldots H, \ldots, TTT\ldots T$ formed by $2n$ subsequent letters H or T (or, 0 and 1). The total number of outcomes is $m = 2^{2n}$, each carries probability $1/2^{2n}$. We are looking for outcomes where the number of Hs equals that of Ts. The number k of such outcomes is $(2n)!/n!n!$ (the number of ways to choose positions for n Hs among $2n$ places available in the sequence). The probability in question is $\frac{(2n)!}{n!n!} \times \frac{1}{2^{2n}}$.

In (ii), the outcomes are the sequences of length $2n+1$, 2^{2n+1} in total. The probability equals

$$\frac{(2n+1)!}{(n+2)!(n-1)!} \times \frac{1}{2^{2n+1}}. \quad \square$$

Problem 1.2 A tennis tournament is organised for 2^n players on a knock-out basis, with n rounds, the last round being the final. Two players are chosen at random. Calculate the probability that they meet (i) in the first or second round, (ii) in the final or semi-final, and (iii) the probability they do not meet.

Solution The sentence 'Two players are chosen at random' is crucial. For instance, one may think that the choice has been made after the tournament when all results are known. Then there are 2^{n-1} pairs of players meeting in the first round, 2^{n-2} in the second round, two in the semi-final, one in the final and $2^{n-1}+2^{n-2}+\cdots+2+1=2^n-1$ in all rounds.

The total number of player pairs is $\binom{2^n}{2}=2^{n-1}(2^n-1)$. Hence the answers:

$$\text{(i)}\quad \frac{2^{n-1}+2^{n-2}}{2^{n-1}(2^n-1)}=\frac{3}{2(2^n-1)},\qquad \text{(ii)}\quad \frac{3}{2^{n-1}(2^n-1)},$$

and

$$\text{(iii)}\,\frac{2^{n-1}(2^n-1)-(2^n-1)}{2^{n-1}(2^n-1)}=1-\frac{1}{2^{n-1}}.\quad\square$$

Problem 1.3 There are n people gathered in a room.

(i) What is the probability that two (at least) have the same birthday? Calculate the probability for $n=22$ and 23.

(ii) What is the probability that at least one has the same birthday as you? What value of n makes it close to 1/2?

Solution The total number of outcomes is 365^n. In (i), the number of outcomes not in the event is $365\times 364\times\cdots\times(365-n+1)$. So, the probability that all birthdays are distinct is $\big(365\times 364\times\cdots\times(365-n+1)\big)/365^n$ and that two or more people have the same birthday

$$1-\frac{365\times 364\times\cdots\times(365-n+1)}{365^n}.$$

For $n=22$:

$$1-\frac{365}{365}\times\frac{364}{365}\times\cdots\times\frac{344}{365}=0.4927,$$

and for $n=23$:

$$1-\frac{365}{365}\times\frac{364}{365}\times\cdots\times\frac{343}{365}=0.5243.$$

In (ii), the number of outcomes not in the event is 364^n and the probability in question $1-(364/365)^n$. We want it to be near 1/2, so

$$\left(\frac{364}{365}\right)^n\approx\frac{1}{2},\quad\text{i.e. } n\approx-\frac{1}{\log_2(364/365)}\approx 252.61.\quad\square$$

Problem 1.4 Mary tosses $n+1$ coins and John tosses n coins. What is the probability that Mary gets more heads than John?

Solution 1 We must assume that all coins are unbiased (as it was not specified otherwise). Mary has 2^{n+1} outcomes (all possible sequences of heads and tails) and John 2^n; jointly 2^{2n+1} outcomes that are equally likely. Let H_{M} and T_{M} be the number of Mary's heads and tails and H_{J} and T_{J} John's, then $H_{\text{M}}+T_{\text{M}}=n+1$ and $H_{\text{J}}+T_{\text{J}}=n$. The events $\{H_{\text{M}}>H_{\text{J}}\}$ and $\{T_{\text{M}}>T_{\text{J}}\}$ have the same number of outcomes, thus $\mathbb{P}(H_{\text{M}}>H_{\text{J}})=\mathbb{P}(T_{\text{M}}>T_{\text{J}})$.

On the other hand, $H_{\text{M}}>H_{\text{J}}$ if and only if $n-H_{\text{M}}<n-H_{\text{J}}$, i.e. $T_{\text{M}}-1<T_{\text{J}}$ or $T_{\text{M}}\le T_{\text{J}}$. So event $H_{\text{M}}>H_{\text{J}}$ is the same as $T_{\text{M}}\le T_{\text{J}}$, and $\mathbb{P}(T_{\text{M}}\le T_{\text{J}})=\mathbb{P}(H_{\text{M}}>H_{\text{J}})$.

But for any (joint) outcome, either $T_{\text{M}}>T_{\text{J}}$ or $T_{\text{M}}\le T_{\text{J}}$, i.e. the number of outcomes in $\{T_{\text{M}}>T_{\text{J}}\}$ equals 2^{2n+1} minus that in $\{T_{\text{M}}\le T_{\text{J}}\}$. Therefore, $\mathbb{P}(T_{\text{M}}>T_{\text{J}})=1-\mathbb{P}(T_{\text{M}}\le T_{\text{J}})$. To summarise:

$$\mathbb{P}(H_{\text{M}}>H_{\text{J}})=\mathbb{P}(T_{\text{M}}>T_{\text{J}})=1-\mathbb{P}(T_{\text{M}}\le T_{\text{J}})=1-\mathbb{P}(H_{\text{M}}>H_{\text{J}}),$$

whence $\mathbb{P}(H_{\text{M}}>H_{\text{J}})=1/2$.

Solution 2 (Fallacious, but popular with some students.) Again assume that all coins are unbiased. Consider pair $(H_{\text{M}},H_{\text{J}})$, as an outcome; there are $(n+2)(n+1)$ such possible pairs, and they all are equally likely (wrong: you have to have biased coins for this!). Now count the number of pairs with $H_{\text{M}}>H_{\text{J}}$. If $H_{\text{M}}=n+1$, H_{J} can take any value $0,1,\dots,n$. In general, $\forall l\le n+1$, if $H_{\text{M}}=l$, H_{J} will take values $0,\dots,l-1$. That is, the number of outcomes where $H_{\text{M}}>H_{\text{J}}$ equals $1+2+\cdots+(n+1)=\frac{1}{2}(n+1)(n+2)$. Hence, $\mathbb{P}(H_{\text{M}}>H_{\text{J}})=1/2$. $\square$

Problem 1.5 You throw $6n$ dice at random. Show that the probability that each number appears exactly n times is

$$\frac{(6n)!}{(n!)^6}\left(\frac{1}{6}\right)^{6n}.$$

Solution There are 6^{6n} outcomes in total (six for each die), each has probability $1/6^{6n}$. We want n dice to show one dot, n two, and so forth. The number of such outcomes is counted by fixing first which dice show one: $(6n)!\big/[n!(5n)!]$. Given n dice showing one, we fix which remaining dice show two: $(5n)!\big/[n!(4n)!]$, etc. The total number of desired outcomes is the product that equals $(6n)!(n!)^6$. This gives the answer. $\square$

In many problems, it is crucial to be able to spot recursive equations relating the cardinality of various events. For example, for the number f_n of ways of tossing a coin n times so that successive tails never appear: $f_n=f_{n-1}+f_{n-2}$, $n\ge 3$ (a Fibonacci equation).

Problem 1.6 (i) Determine the number g_n of ways of tossing a coin n times so that the combination HT never appears. (ii) Show that $f_n = f_{n-1} + f_{n-2} + f_{n-3}$, $n \geq 3$, is the equation for the number of ways of tossing a coin n times so that three successive heads never appear.

Solution (i) $g_n = 1 + n$; 1 for the sequence $HH \ldots H$, n for the sequences $T \ldots TH \ldots H$ (which includes $T \ldots T$).

(ii) The outcomes are 2^n sequences $(y_1, \ldots, y_n)$ of H and T. Let A_n be the event {no three successive heads appeared after n tosses}, then f_n is the cardinality $\#A_n$. Split: $A_n = B_n^{(1)} \cup B_n^{(2)} \cup B_n^{(3)}$, where $B_n^{(1)}$ is the event {no three successive heads appeared after n tosses, and the last toss was a tail}, $B_n^{(2)} =$ {no three successive heads appeared after n tosses, and the last two tosses were TH} and $B_n^{(3)} =$ {no three successive heads appeared after n tosses, and the last three tosses were THH}.

Clearly, $B_n^{(i)} \cap B_n^{(j)} = \emptyset$, $1 \leq i \neq j \leq 3$, and so $f_n = \#B_n^{(1)} + \#B_n^{(2)} + \#B_n^{(3)}$.

Now drop the last digit y_n: $(y_1, \ldots y_n) \in B_n^{(1)}$ iff $y_n = T$, $(y_1, \ldots y_{n-1}) \in A_{n-1}$, i.e. $\#B_{n-1}^{(1)} = f_{n-1}$. Also, $(y_1, \ldots y_n) \in B_n^{(2)}$ iff $y_{n-1} = T$, $y_n = H$, and $(y_1, \ldots y_{n-2}) \in A_{n-2}$. This allows us to drop the two last digits, yielding $\#B_n^{(2)} = f_{n-2}$. Similarly, $\#B_n^{(3)} = f_{n-3}$. The equation then follows. □

1.2 Conditional Probabilities. The Bayes Theorem. Independent trials

Probability theory is nothing but common sense
reduced to calculation.

P.-S. Laplace (1749–1827), French mathematician

Clockwork Omega

(From the series '*Movies that never made it to the Big Screen*'.)

From now on we adopt a more general setting: our outcomes do not necessarily have equal probabilities $p_1, \ldots, p_m$, with $p_i > 0$ and $p_1 + \cdots + p_m = 1$.

As before, an *event* A is a collection of outcomes (possibly empty); the *probability* $\mathbb{P}(A)$ of event A is now given by

$$\mathbb{P}(A) = \sum_{\text{outcome } i \in A} p_i = \sum_{\text{outcome } i} p_i I(i \in A). \tag{1.2}$$

($\mathbb{P}(A) = 0$ for $A = \emptyset$.) Here and below, I stands for the *indicator function*, viz.:

$$I(i \in A) = \begin{cases} 1, & \text{if } i \in A, \\ 0, & \text{otherwise.} \end{cases}$$

The probability of the total set of outcomes is 1. The total set of outcomes is also called the whole, or full, event and is often denoted by Ω, so $\mathbb{P}(\Omega) = 1$. An outcome is often denoted by ω, and if $p(\omega)$ is its probability, then

$$\mathbb{P}(A) = \sum_{\omega \in A} p(\omega) = \sum_{\omega \in \Omega} p(\omega) I(\omega \in A). \tag{1.3}$$

As follows from this definition, the probability of the union

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) \tag{1.4}$$

for any pair of disjoint events A_1, A_2 (with $A_1 \cap A_2 = \emptyset$). More generally,

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = \mathbb{P}(A_1) + \cdots + \mathbb{P}(A_n) \tag{1.5}$$

for any collection of pair-wise disjoint events (with $A_j \cap A_{j'} = \emptyset \ \forall j \neq j'$). Consequently, (i) the probability $\mathbb{P}(A^c)$ of the complement $A^c = \Omega \setminus A$ is $1 - \mathbb{P}(A)$, (ii) if $B \subseteq A$, then $\mathbb{P}(B) \le \mathbb{P}(A)$ and $\mathbb{P}(A) - \mathbb{P}(B) = \mathbb{P}(A \setminus B)$, and (iii) for a general pair of events A, B: $\mathbb{P}(A \setminus B) = \mathbb{P}\big(A \setminus (A \cap B)\big) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$.

Furthermore, for a general (not necessarily disjoint) union:

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) \le \sum_{i=1}^{n} \mathbb{P}(A_i);$$

a more detailed analysis of the probability $\mathbb{P}(\cup A_i)$ is provided by the exclusion–inclusion formula (1.12); see below.

Given two events A and B with $\mathbb{P}(B) > 0$, the *conditional probability* $\mathbb{P}(A|B)$ of A given B is defined as the ratio

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{1.6}$$

At this stage, the conditional probabilities are important for us because of two formulas. One is the formula of complete probability: if $B_1, \ldots, B_n$ are pair-wise disjoint events partitioning the whole event Ω, i.e. have $B_i \cap B_j = \emptyset$ for $1 \le i < j \le n$ and $B_1 \cup B_2 \cup \cdots \cup B_n = \Omega$, and in addition $\mathbb{P}(B_i) > 0$ for $1 \le i \le n$, then

$$\mathbb{P}(A) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) + \cdots + \mathbb{P}(A|B_n)\mathbb{P}(B_n). \tag{1.7}$$

The proof is straightforward:

$$\mathbb{P}(A) = \sum_{1 \le i \le n} \mathbb{P}(A \cap B_i) = \sum_{1 \le i \le n} \frac{\mathbb{P}(A \cap B_i)}{\mathbb{P}(B_i)} \mathbb{P}(B_i) = \sum_{1 \le i \le n} \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

The point is that often it is conditional probabilities that are given, and we are required to find unconditional ones; also, the formula of complete probability is useful to clarify the nature of (unconditional) probability $\mathbb{P}(A)$. Despite its simple character, this formula is an extremely powerful tool in literally all areas dealing with probabilities. In particular, a large portion of the theory of Markov chains is based on its skilful application.

Representing $\mathbb{P}(A)$ in the form of the right-hand side (RHS) of (1.7) is called conditioning (on the collection of events $B_1, \ldots, B_n$).

Another formula is the *Bayes formula* (or the *Bayes Theorem*) named after T. Bayes (1702–1761), an English mathematician and cleric. It states that *under the same assumptions as above, if in addition* $\mathbb{P}(A) > 0$, *then the conditional probability* $\mathbb{P}(B_i|A)$ *can be expressed in terms of probabilities* $\mathbb{P}(B_1), \ldots, \mathbb{P}(B_n)$ *and conditional probabilities* $\mathbb{P}(A|B_1), \ldots, \mathbb{P}(A|B_n)$ *as*

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum\limits_{1\le j\le n} \mathbb{P}(A|B_j)\mathbb{P}(B_j)}. \tag{1.8}$$

The proof is the direct application of the definition and the formula of complete probability:

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

and

$$\mathbb{P}(A) = \sum_j \mathbb{P}(A|B_j)\mathbb{P}(B_j).$$

A standard interpretation of equation (1.8) is that it relates the *posterior probability* $\mathbb{P}(B_i|A)$ (conditional on A) with *prior probabilities* $\{\mathbb{P}(B_j)\}$ (valid before one knew that event A occurred).

In his lifetime, Bayes finished only two papers: one in theology and one called 'Essay towards solving a problem in the doctrine of chances'; the latter contained the Bayes Theorem and was published two years after his death. Nevertheless he was elected a Fellow of The Royal Society. Bayes' theory (of which the above theorem is an important part) was for a long time subject to controversy. His views were fully accepted (after considerable theoretical clarifications) only at the end of the nineteenth century.

Problem 1.7 Four mice are chosen (without replacement) from a litter containing two white mice. The probability that both white mice are chosen is twice the probability that neither is chosen. How many mice are there in the litter?

Solution Let the number of mice in the litter be n. We use the notation $\mathbb{P}(2) = \mathbb{P}(\text{two white chosen})$ and $\mathbb{P}(0) = \mathbb{P}(\text{no white chosen})$. Then

$$\mathbb{P}(2) = \binom{n-2}{2} \Big/ \binom{n}{4}.$$

Otherwise, $\mathbb{P}(2)$ could be computed as:

$$\frac{2}{n}\frac{1}{n-1} + \frac{2}{n}\frac{n-2}{n-1}\frac{1}{n-2} + \frac{2}{n}\frac{n-2}{n-1}\frac{n-3}{n-2}\frac{1}{n-3} + \frac{n-2}{n}\frac{2}{n-1}\frac{1}{n-2}$$
$$+ \frac{n-2}{n}\frac{n-3}{n-1}\frac{2}{n-2}\frac{1}{n-3} + \frac{n-2}{n}\frac{2}{n-1}\frac{n-3}{n-2}\frac{1}{n-3} = \frac{12}{n(n-1)}.$$

On the other hand,

$$\mathbb{P}(0) = \binom{n-2}{4} \Big/ \binom{n}{4}.$$

Otherwise, $\mathbb{P}(0)$ could be computed as follows:

$$\mathbb{P}(0) = \frac{n-2}{n}\frac{n-3}{n-1}\frac{n-4}{n-2}\frac{n-5}{n-3} = \frac{(n-4)(n-5)}{n(n-1)}.$$

Solving the equation

$$\frac{12}{n(n-1)} = 2\frac{(n-4)(n-5)}{n(n-1)},$$

we get $n = (9 \pm 5)/2$; $n = 2$ is discarded as $n \geq 6$ (otherwise the second probability is 0). Hence, $n = 7$. □

Problem 1.8 Lord Vile drinks his whisky randomly, and the probability that, on a given day, he has n glasses equals $\mathrm{e}^{-1}/n!$, $n = 0, 1, \ldots$ Yesterday his wife Lady Vile, his son Liddell and his butler decided to murder him. If he had no whisky that day, Lady Vile was to kill him; if he had exactly one glass, the task would fall to Liddell, otherwise the butler would do it. Lady Vile is twice as likely to poison as to strangle, the butler twice as likely to strangle as to poison, and Liddell just as likely to use either method. Despite their efforts, Lord Vile is not guaranteed to die from any of their attempts, though he is three times as likely to succumb to strangulation as to poisoning.

Today Lord Vile is dead. What is the probability that the butler did it?

Solution Write $\mathbb{P}(\text{dead}|\text{strangle}) = 3r$, $\mathbb{P}(\text{dead}|\text{poison}) = r$, and

$$\mathbb{P}(\text{drinks no whisky}) = \mathbb{P}(\text{drinks one glass}) = \frac{1}{\mathrm{e}},$$
$$\mathbb{P}(\text{drinks two glasses or more}) = 1 - \frac{2}{\mathrm{e}}.$$

Next:

$$\mathbb{P}(\text{strangle}|\text{Lady V}) = \frac{1}{3},\ \mathbb{P}(\text{poison}|\text{Lady V}) = \frac{2}{3},$$
$$\mathbb{P}(\text{strangle}|\text{butler}) = \frac{2}{3},\ \mathbb{P}(\text{poison}|\text{butler}) = \frac{1}{3},$$

and

$$\mathbb{P}(\text{strangle}|\text{Liddell}) = \mathbb{P}(\text{poison}|\text{Liddell}) = \frac{1}{2}.$$

Then the conditional probability $\mathbb{P}(\text{butler}|\text{dead})$ is

$$\frac{\mathbb{P}(\text{d}|\text{b})\mathbb{P}(\text{b})}{\mathbb{P}(\text{d}|\text{b})\mathbb{P}(\text{b})+\mathbb{P}(\text{d}|\text{LV})\mathbb{P}(\text{LV})+\mathbb{P}(\text{d}|\text{Lddl})\mathbb{P}(\text{Lddl})}$$
$$=\frac{\left(1-\frac{2}{\text{e}}\right)\left(\frac{3r\times 2}{3}+\frac{r}{3}\right)}{\left(1-\frac{2}{\text{e}}\right)\left(\frac{3r\times 2}{3}+\frac{r}{3}\right)+\frac{1}{\text{e}}\left(\frac{3r}{3}+\frac{r\times 2}{3}\right)+\frac{1}{\text{e}}\left(\frac{3r}{2}+\frac{r}{2}\right)}$$
$$=\frac{\text{e}-2}{\text{e}-3/7}\approx 0.3137. \quad \square$$

Problem 1.9 At the station there are three payphones which accept 20p pieces. One never works, another always works, while the third works with probability $1/2$. On my way to the metropolis for the day, I wish to identify the reliable phone, so that I can use it on my return. The station is empty and I have just three 20p pieces. I try one phone and it does not work. I try another twice in succession and it works both times. What is the probability that this second phone is the reliable one?

Solution Let A be the event in the question: the first phone tried did not work and second worked twice. Clearly:

$$\mathbb{P}(A|\text{1st reliable})=0,$$
$$\mathbb{P}(A|\text{2nd reliable})=\mathbb{P}(\text{1st never works}|\text{2nd reliable})$$
$$+\frac{1}{2}\times\mathbb{P}(\text{1st works half-time}|\text{2nd reliable})$$
$$=\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}=\frac{3}{4},$$

and the probability $\mathbb{P}(A|\text{3rd reliable})$ equals

$$\frac{1}{2}\times\frac{1}{2}\times\mathbb{P}(\text{2nd works half-time}|\text{3rd reliable})=\frac{1}{8}.$$

The required probability $\mathbb{P}(\text{2nd reliable})$ is then

$$\frac{1/3\times 3/4}{1/3\times(0+3/4+1/8)}=\frac{6}{7}. \quad \square$$

Problem 1.10 Parliament contains a proportion p of Labour Party members, incapable of changing their opinions about anything, and $1-p$ of Tory Party members changing their minds at random, with probability r, between subsequent votes on the same issue. A randomly chosen parliamentarian is noticed to have voted twice in succession in the same way. Find the probability that he or she will vote in the same way next time.

Solution Set

$$A_1 = \{\text{Labour chosen}\},\ \ A_2 = \{\text{Tory chosen}\},$$

$$B = \{\text{the member chosen voted twice in the same way}\}.$$

We have $\mathbb{P}(A_1) = p$, $\mathbb{P}(A_2) = 1 - p$, $\mathbb{P}(B|A_1) = 1$, $\mathbb{P}(B|A_2) = 1 - r$. We want to calculate

$$\mathbb{P}(A_1|B) = \frac{\mathbb{P}(A_1 \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_1)\mathbb{P}(B|A_1)}{\mathbb{P}(B)}$$

and $\mathbb{P}(A_2|B) = 1 - \mathbb{P}(A_1|B)$. Write

$$\mathbb{P}(B) = \mathbb{P}(A_1)\mathbb{P}(B|A_1) + \mathbb{P}(A_2)\mathbb{P}(B|A_2) = p \cdot 1 + (1-p)(1-r).$$

Then

$$\mathbb{P}(A_1|B) = \frac{p}{p + (1-r)(1-p)},\ \ \mathbb{P}(A_2|B) = \frac{(1-r)(1-p)}{p + (1-r)(1-p)},$$

and the answer is given by

$$\mathbb{P}\big(\text{the member will vote in the same way}\big|B\big) = \frac{p + (1-r)^2(1-p)}{p + (1-r)(1-p)}. \quad \square$$

Problem 1.11 The Polya urn model is as follows. We start with an urn which contains one white ball and one black ball. At each second we choose a ball at random from the urn and replace it together with one more ball of the same colour. Calculate the probability that when n balls are in the urn, i of them are white.

Solution Denote by $\mathbb{P}_n$ the conditional probability given that there are n balls in the urn. For $n = 2$ and 3

$$\mathbb{P}_n(\text{one white ball}) = \begin{cases} 1, & n = 2 \\ \frac{1}{2}, & n = 3, \end{cases}$$

and

$$\mathbb{P}_n(\text{two white balls}) = \tfrac{1}{2},\ \ n = 3.$$

Make the induction hypothesis

$$\mathbb{P}_k(i \text{ white balls}) = \frac{1}{k-1},$$

$\forall\ k = 2, \dots, n-1$ and $i = 1, \dots, k-1$. Then, after $n-1$ trials (when the number of balls is n),

$$\begin{aligned} &\mathbb{P}_n(i \text{ white balls}) \\ &= \mathbb{P}_{n-1}(i-1 \text{ white balls}) \times \frac{i-1}{n-1} + \mathbb{P}_{n-1}(i \text{ white balls}) \times \frac{n-1-i}{n-1} \\ &= \frac{1}{n-1},\ \ i = 1, \dots, n-1. \end{aligned}$$