Topic 1 — Probability spaces

1.1 Definition

In order to properly understand a statement like

> "the chance of getting a flush in five-card poker is about 0.2%",

we need to specify the underlying *probability space*. This has two components:

1. A *sample space* or *space of outcomes*.

   This is $\Omega = \{$all possible five-card hands$\}$.

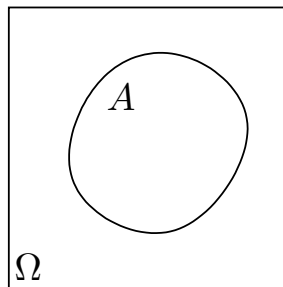2. The *probabilities of outcomes*.

   In this specific example, assuming the cards are dealt fairly, all outcomes are equally probable, so

   $$\Pr(\omega) = 1/|\Omega| \qquad \text{for all } \omega \in \Omega.$$

   In general, probabilities are nonnegative and sum to one:

   $$\sum_{\omega \in \Omega} \Pr(\omega) = 1.$$

Returning to the example, the *event* of interest is $A = \{\omega : \omega \text{ is a flush}\}$. This is a subset of Ω; that is, $A \subset \Omega$. Pictorially we can represent the situation thus:

The outer box is Ω; every point in it is a particular five-card hand ω. The inner set is A, and the probability that it occurs is

$$\Pr(A) = \sum_{\omega \in A} \Pr(\omega).$$

In this case, since all outcomes are equally likely, $\Pr(A) = |A|/|\Omega|$.

1.2 A first set of canonical examples

1. *Roll a die.* What is the chance of getting a number > 3?

   Probability space: sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$; probabilities $\Pr(\omega) = 1/6$.

   Event of interest: $A = \{4, 5, 6\}$; $\Pr(A) = 1/2$.

2. *Roll three dice.* What is the chance their sum is 3?

   The sample space is

$$
\begin{aligned}
\Omega &= \{(1,1,1), (1,1,2), \ldots, (6,6,6)\} \\
&= \{(d_1, d_2, d_3) : 1 \le d_1, d_2, d_3 \le 6\} \\
&= \Omega_o \times \Omega_o \times \Omega_o \quad = \quad \Omega_o^3 \quad \text{where } \Omega_o = \{1, 2, 3, 4, 5, 6\}.
\end{aligned}
$$

   The probabilities of outcomes $\omega \in \Omega$ are $\Pr(\omega) = 1/6^3 = 1/216$.

   The event of interest is $A = \{(1,1,1)\}$ whose probability is $\Pr(A) = 1/216$.

3. *Roll n dice.*

   Sample space $\Omega = \Omega_o^n$, where $\Omega_o = \{1, 2, 3, 4, 5, 6\}$. Each outcome $\omega \in \Omega$ has probability $1/|\Omega| = 1/6^n$.

4. *Socks in a drawer.* A drawer contains three blue socks and three red socks. You put your hand in and pick out a random sock. Then you put your hand in again and pick out another random sock. What's the chance the two of them match?

   There are several ways to set up the sample space, but one possibility is to have a tuple whose first coordinate is the color of the first sock and whose second coordinate is the color of the second sock. $\Omega = \{B, R\} \times \{B, R\}$. The probabilities of outcomes are

$$
\begin{aligned}
\Pr((B,B)) &= \frac{3}{6} \cdot \frac{2}{5} = \frac{1}{5} \\
\Pr((B,R)) &= \frac{3}{6} \cdot \frac{3}{5} = \frac{3}{10} \\
\Pr((R,B)) &= \frac{3}{6} \cdot \frac{3}{5} = \frac{3}{10} \\
\Pr((R,R)) &= \frac{3}{6} \cdot \frac{2}{5} = \frac{1}{5}
\end{aligned}
$$

   (Notice that they add up to 1.)

   The event of interest is $A = \{(B,B), (R,R)\}$, which has probability $2/5$.

5. *Socks in a drawer, again.* This time the drawer has three blue socks and four red socks.

   The sample space Ω is the same, but the probabilities of outcomes are different:

$$
\begin{aligned}
\Pr((B,B)) &= \frac{3}{7} \cdot \frac{2}{6} = \frac{1}{7} \\
\Pr((B,R)) &= \frac{3}{7} \cdot \frac{4}{6} = \frac{2}{7} \\
\Pr((R,B)) &= \frac{4}{7} \cdot \frac{3}{6} = \frac{2}{7} \\
\Pr((R,R)) &= \frac{4}{7} \cdot \frac{3}{6} = \frac{2}{7}
\end{aligned}
$$

   The event of interest is $A = \{(B,B), (R,R)\}$, which has probability $3/7$.

6. *Shuffling a deck of cards.* You randomly shuffle a deck of 52 cards and lay them out before you.

   Here $\Omega = \{$all possible orderings of 52 cards$\}$. One way to compute $|\Omega|$ is to reason that there are 52 choices for what the first card in the sequence will be, 51 choices for the second card, 50 choices for the third card, and so on. Therefore

   $$|\Omega| = 52 \cdot 51 \cdot 50 \cdots 2 \cdot 1.$$

   This expression is called 52! ("52 factorial"). It is the number of *permutations* of 52 elements.

1.3   Tossing a fair coin

Suppose you toss a fair coin. The sample space is $\Omega_o = \{H, T\}$ (heads or tails), and each of the two outcomes has probability exactly 1/2. What is more interesting is to toss the coin multiple times, independently.

1.3.1   Toss a fair coin 10 times

Now the sample space is $\Omega = \Omega_o^{10}$; it includes, for instance, the sequence $(H, T, H, T, H, T, H, T, H, T)$. Since $|\Omega| = 2^{10} = 1024$, each element in Ω has probability exactly 1/1024.

1. What is the chance that *none* of the coin tosses are heads?

   The event of interest is $\{(T, T, T, T, T, T, T, T, T, T)\}$, whose probability is 1/1024.

2. What is the chance of exactly *one* head?

   Now the event is

   $$A = \{(H, T, T, T, T, T, T, T, T, T), (T, H, T, T, T, T, T, T, T, T), \ldots, (T, T, T, T, T, T, T, T, T, H)\}.$$

   Each sequence in A is completely determined by the location of the H within it. There are 10 possible locations; therefore $|A| = 10$, whereupon $\Pr(A) = |A|/|\Omega| = 10/1024$.

3. What is the chance of exactly *nine* heads?

   Equivalently, what is the chance of exactly one tail? This is the same calculation as before, 10/1024.

4. What is the chance of exactly *two* heads?

   The sequences of interest are $A = \{\omega \in \Omega : \omega \text{ has exactly 2 heads}\}$. Each such sequence is specified by the locations of its two heads; write these as a pair (i, j), where $1 \leq i, j \leq 10$ and $i \neq j$. For instance, $(7, 5)$ refers to $(T, T, T, T, H, T, H, T, T, T)$.

   The number of such pairs is $10 \cdot 9 = 90$ (10 choices for i, and thereafter just 9 choices for j). But this ends up double-counting sequences, because for instance, the pair $(5, 7)$ also refers to $(T, T, T, T, H, T, H, T, T, T)$. Each sequence in A is counted twice – it corresponds to two pairs – and therefore $|A| = 45$.

   Finally, $\Pr(A) = |A|/|\Omega| = 45/1024$.

5. What is the chance of exactly *four* heads?

   This time, any sequence in $A = \{\omega : \text{has four heads}\}$ can be written as a 4-tuple (i, j, k, l), where $1 \leq i, j, k, l \leq 10$ and $i \neq j \neq k \neq l$. For instance, $(7, 2, 4, 9)$ denotes $(T, H, T, H, T, T, H, T, H, T)$. The number of such 4-tuples is $10 \cdot 9 \cdot 8 \cdot 7 = 10!/6!$.

But once again, there is overcounting. $(7, 2, 4, 9)$ refers to the same sequence as $(2, 4, 7, 9)$ and $(2, 7, 9, 4)$ and $(9, 7, 2, 4)$ and many other 4-tuples. How many of them? The number of permutations of the four elements $2, 4, 7, 9$, namely $4!$.

Therefore

$$|A| = \frac{10!}{4!6!} = 210$$

and $\Pr(A) = 210/1024$.

6. What is the chance of exactly *six* heads?

   This is the same as the chance of four tails, which is identical to the previous calculation.

In the last few calculations, we had to find the number of ways of choosing k positions out of n available slots (for instance, choosing 4 positions out of 10 slots where a H might occur). Generalizing the argument above, the number of ways to do this is

$$\frac{n!}{k!(n-k)!}, \text{ which we write as } \binom{n}{k}$$

(pronounced "n choose k"). This is also called a *binomial coefficient.*

1.3.2   Toss a fair coin n times

Now the sample space is $\Omega = \{H, T\}^n$, with each sequence of n outcomes having probability exactly $1/2^n$. Let A_k denote the event that the sequence has k heads.

Notice that the events $A_0, A_1, \ldots, A_n$ are *disjoint* (if A_i occurs then A_j cannot occur for $j \neq i$). Moreover,

$$\Omega = A_0 \cup A_1 \cup \cdots \cup A_n.$$

Therefore,

$$\sum_{k=0}^{n} \Pr(A_k) \quad = \quad \Pr(A_0) + \Pr(A_1) + \cdots + \Pr(A_n) \quad = \quad 1,$$

or equivalently,

$$\sum_{k=0}^{n} |A_k| \quad = \quad |A_0| + |A_1| + \cdots + |A_n| \quad = \quad |\Omega|.$$

In general, $|A_k|$ is the number of ways of placing k heads in a sequence of size n; we've seen that this is $\binom{n}{k}$. And since $|\Omega| = 2^n$, the last equality tells us that

$$\sum_{k=0}^{n} \binom{n}{k} \quad = \quad \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} \quad = \quad 2^n.$$

For example, take $n = 5$. Then

$$|A_0| = \binom{5}{0} = 1, \; |A_1| = \binom{5}{1} = 5, \; |A_2| = \binom{5}{2} = 10, \; |A_3| = \binom{5}{3} = 10, \; |A_4| = \binom{5}{4} = 5, \; |A_5| = \binom{5}{5} = 1$$

and these add up to $2^5 = 32$.

1.4   A second set of canonical examples

1. *Rooks on a chessboard.* You place 8 rooks at random on a chessboard. What is the chance that they are non-attacking (that is, no rook is attacking another)?

   To describe the sample space, number the squares in the chessboard as 1 through 64, and let the configuration of 8 rooks be given by a *set* of eight positions, $\omega \subset \{1, 2, \ldots, 64\}$. Thus:

   $$\Omega = \{\omega \subset \{1, 2, \ldots, 64\} : |\omega| = 8\}.$$

   You should check that

   $$|\Omega| \;=\; \binom{64}{8}$$

   and that of these, only 8! configurations are non-attacking.

   Therefore

   $$\Pr(\text{non-attacking configuration}) \;=\; \frac{8!\,8!\,56!}{64!}.$$

2. *Five-card poker.* You are dealt five cards at random from a deck of 52 cards. What is the chance of a flush? Of a straight flush? Of exactly one pair?

   Let $\Omega = \{\text{all possible 5-card hands}\}$. Then $|\Omega| = \binom{52}{5}$ and each $\omega \in \Omega$ occurs with probability $1/|\Omega|$.

   Define three events of interest:

   $$
   \begin{aligned}
   F &= \text{flush (all five cards of the same suit)} \\
   S &= \text{straight flush (same suit and consecutive)} \\
   P &= \text{the cards contain a single pair (eg. two 7s)}
   \end{aligned}
   $$

   Then $|F| = 4 \cdot \binom{13}{5}$ (first choose a suit, then pick 5 cards from that suit), $|S| = 4 \cdot 10$ (first choose a suit, then choose the starting card in the sequence), and $|P| = 13 \cdot \binom{4}{2} \cdot 4^3 \cdot \binom{12}{3}$ (first choose which card occurs in the pair, then choose the two suits for that pair, then choose the suits of the remaining three cards, then choose their values).

3. *Throw a dart at a dartboard.* Suppose for convenience that your dartboard has radius 1, and is centered at the origin. Its bullseye has radius 0.1. You throw a dart at it, which lands at a random location (all positions on the board are equally likely). What is the chance that it lands exactly at the origin? What is the chance that it lands in the bullseye?

   This differs from earlier examples in that the sample space is infinite and continuous. It is the set of all possible locations of the dart: any point in the circle. We can represent any such point by its (x, y) coordinates: $\Omega = \{(x, y) : x^2 + y^2 \leq 1\}$.

   The chance of landing exactly at the origin is 0, since there are infinitely many places the dart could land. It makes more sense to talk about landing in *regions* $A \subset \Omega$ rather than specific points $\omega \in \Omega$. In general

   $$\Pr(A) = \frac{\text{area of } A}{\text{area of } \Omega}$$

   and therefore $\Pr(\text{bullseye}) = (0.1)^2 = 0.01$.

4. *Birthday paradox.* A room contains n people. What is the chance that two of them have the same birthday?

The probability space is not properly specified, so we need to make some assumptions. First, we'll assume that the n birthdays are independent (that is, a person's birthday is not influenced by anyone else's birthday). Second, we'll assume that all days are equally likely – that is, the chance of a birthday falling on any particular day is exactly 1/365 (we're also ignoring the issue of leap years).

Number the people $1, 2, \ldots, n$, and number the days of the year $1, 2, \ldots, 365$. We will represent the birthdays of the people in the room by an n-tuple $(\omega_1, \ldots, \omega_n)$, where $\omega_i \in \{1, 2, \ldots, 365\}$ is the birthday of the ith person. Thus $\Omega = \{1, 2, \ldots, 365\}^n$ and each $\omega \in \Omega$ has probability exactly $1/365^n$.

The event of interest is

$$A \;\;=\;\; \{\omega : \omega_i = \omega_j \text{ for some } i \neq j\}.$$

This is a typical situation in which it is easier to analyze the *complement* of A than A itself (that is, it is easier to compute the probability that A *doesn't* occur than the probability that it occurs).

$$A^c \;\;=\;\; \Omega - A \;\;=\;\; \{\omega : \omega_1 \neq \omega_2 \neq \cdots \neq \omega_n\}.$$

In other words, A^c is the event that everyone's birthday is different. What is the size of A^c? There are 365 choices for ω_1, 364 for ω_2, and so on, whereupon

$$|A^c| \;\;=\;\; 365 \cdot 364 \cdot 363 \cdots (365 - n + 1) \;\;=\;\; \frac{365!}{(365 - n)!}.$$

Therefore

$$\Pr(A) \;\;=\;\; 1 - \Pr(A^c) \;\;=\;\; 1 - \frac{365!}{(365 - n)! 365^n}.$$

This is exactly correct, but it is a little hard to understand intuitively. So let's do the calculation a different way, using an approximation.

A very useful fact is that for small x (positive or negative), $e^x \approx 1 + x$. And in fact, $e^x \geq 1 + x$ no matter what x is. Now let's return to the event A^c.

$$
\begin{aligned}
\Pr(A^c) \;&=\; \Pr(\omega_2 \neq \omega_1) \cdot \Pr(\omega_3 \neq \omega_1, \omega_2) \cdots \Pr(\omega_n \neq \omega_1, \ldots, \omega_{n-1}) \\
&=\; \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right) \cdots \left(1 - \frac{n-1}{365}\right) \\
&\leq\; \exp(-1/365) \cdot \exp(-2/365) \cdots \exp(-(n-1)/365) \quad \text{where } \exp(x) \text{ means } e^x \\
&=\; \exp\left(-\frac{1}{365}\left(1 + 2 + \cdots + (n-1)\right)\right) \\
&=\; \exp\left(-\frac{n(n-1)}{730}\right).
\end{aligned}
$$

This upper bound is a very good approximation when n is much smaller than 365.

Interestingly, when $n = 23$, we find that $\Pr(A^c) \leq 0.5$, so $\Pr(A) \geq 0.5$. That is, if there are 23 people in the room, chances are that two of them have the same birthday!