1 Discrete Probability

(The definitions follow Babai’s Notes, the examples are mostly from Rosen.)

**Definition 1.** A finite probability space is a finite set $\Omega$ of atomic events or elementary events together with a function $Pr : \Omega \rightarrow R^+$ (the probability distribution), that maps atomic events to positive reals. The probability distribution function is defined on all subsets $E$ of $\Omega$ (called events) by $Pr(E) = \sum_{\omega \in E} Pr(\omega)$. $Pr()$ has the following properties:

(i) for all atomic events $\omega$ $Pr(\omega) > 0$

(ii) $Pr(\Omega) = 1$

(iii) $Pr(\emptyset) = 0$

$\Omega$ is also called sample space.

**Definition 2.** The uniform distribution assigns the same probability to all atomic events ($= \frac{1}{|\Omega|}$)

The uniform distribution on coin toss is often called ‘unbiased coin’, the uniform distribution on the faces as outcome of a toss of a die, a ‘fair die’. In general, this is also called a ‘random choice’.

**Example 3.** Consider a random choice of 3 fair dice. The sample space is all triples $(a, b, c) : a, b, c \in \{1, 2, 3, 4, 5, 6\}$ Each atomic event has probability $\frac{1}{216}$.
To compute the probability of the event $E = \text{`the sum of the three dice is 9’}$ we need to count the number of outcomes where this happens:

\{
(6,1,2); (6,2,1); (5,3,1); (5,1,3); (5,2,2); (4,4,1); (4,1,4); (4,3,2); (4,2,3); (3,5,1);
(3,1,5); (3,4,2); (3,2,4); (3,3,3); (2,6,1); (2,1,6); (2,5,2); (2,2,5); (2,4,3); (2,3,4); (1,3,5); (1,5,3);
(1,6,2); (1,2,6)\}. So the probability is $\frac{24}{60} = \frac{2}{5}$

Of course, we want to use our knowledge of counting combinatorial configurations to avoid listing them.

**Example 4.** What is the probability of getting a hand with 4 of a kind? (Recall that the deck has 4 suits of 13 ‘kinds’ in each: a ‘kind’ is an A, or a 2, or a 3, or, ...)

The sample space $\Omega$ is the set of 5-tuples (subsets of size 5) taken from the deck. We assume the uniform distribution. The number of 5-tuples is $\binom{52}{5}$. To compute the number of hands that are 4 of a kind we

- choose a kind – we choose a subset of size 1 from a universe of 13 elements: $\binom{13}{1}$ ways
- We must take all 4 of the 4 cards of the kind in the deck – no choice, or, formally $\binom{4}{4}$
- We take a 5th card from the remaining 52-4 cards – $\binom{48}{1}$ ways

Using the product rule, the number of ‘good’ configurations is $\binom{13}{1} \times \binom{4}{4} \times \binom{48}{1}$, and so the probability is $\frac{\binom{13}{1} \times \binom{4}{4} \times \binom{48}{1}}{\binom{52}{5}}$.

**Example 5.** What is the probability of a ‘full house’ in poker? (A full house is 3 of a kind and 2 of a kind.) The sample space is the same as before.

Our choices for the good configurations are

- choose the two kinds (why must they be different?) – $\binom{13}{2}$ ways
- choose which 3 for the triple – $\binom{4}{3}$ ways
- choose which 2 for the pair – $\binom{4}{2}$ ways

So the probability is $\frac{\binom{13}{2} \times \binom{4}{3} \times \binom{4}{2}}{\binom{52}{5}}$. 

2
1.1 Union of Events

Assuming the uniform distribution, it is clear that for any events \( A, B \)
\[
Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)
\]
since when if count the atomic events in \( A \cup B \) by counting the elements of \( A \), then the elements in \( B \), we count the elements of \( A \cap B \) twice.

In particular, when \( A \cap B = \emptyset \) (\( A \) and \( B \) are disjoint) this simplifies to
\[
Pr(A \cup B) = Pr(A) + Pr(B)
\]

1.2 Conditional Probability, Independence, Correlation

**Definition 6.** Let \( F \neq \emptyset \).

The ‘conditional probability of \( E \) given \( F \)’ is \( Pr(E|F) = \frac{Pr(E \cap F)}{Pr(F)} \)

Looking at the definition (draw a Venn diagram!) we see that it corresponds at the probability mass in the intersection of \( E \) and \( F \), normalized by the probability of \( F \). One can think of it as looking at a new sample space, consisting of the event \( F \), and looking at the probability of the portion of \( E \) inside this space.

Note: the interpretation above fails if \( Pr(E|F) = 0 \) – but we explicitly excluded this possibility when defining conditional probability!

**Note on events of probability 0.** We could have considered elementary events with probability 0: the reason we forbade it in the definition is simply to make statements of definitions and theorems shorter. Also, one could then possibly have to consider an infinite sequence of 0 probability events... On the other hand, for the same reason nonatomic events can have 0 probability – if they are the empty set.

**Example 7.** Suppose a family has 2 children. The probabilities of the four possible boy (B) / girl (G) children – namely BB, BG, GB, and GG—are all the same. Given that one of the children is a boy, what is the probability that the family has 2 boys?

We have that \( E=\{BB\} \), \( F=\{BG, GB, BB\} \), \( Pr(E|F) = \frac{Pr(E \cap F)}{Pr(F)} = \frac{1/4}{3/4} = \frac{1}{3} \)