Note on Discrete probability

Note: This is a very brief introduction to discrete probability. There is much much more to probability than what was seen in this course.

When studying a situation using probability, we need to define the following two concepts.

Definition 1. The *sample space* is the set of all possible outcomes. It is denoted by Ω.

The *probability mass function* is a function $p : \Omega \to [0, 1]$ with the property

$$\sum_{e \in \Omega} p(e) = 1$$

An *event* is a subset E of Ω. The probability of an event is defined as

$$Pr[E] = \sum_{e \in E} p(e)$$

Note that Pr is a function from the subsets of Ω to $[0, 1]$ which can be obtain from p. Pr is known as the probability distribution.

Example 1. Say we throw an ordinary 6-sided die. We would like to know how likely it will be even.

Here, there are 6 possible outcomes for the value of the die, namely 1,2,3,4,5 and 6. So $\Omega = \{1, 2, 3, 4, 5, 6\}$ is our sample space.

We said that the die is ordinary so we obtain any specific value with probability 1/6. Therefore, the probability mass function is defined as $p(e) = \frac{1}{6}\ \forall e \in \Omega$.

We are interested in the event that we roll an even number. Thus, we want to look at the event $E = \{2, 4, 6\}$.

So, by definition, the probability of E is

$$Pr[E] = \sum_{e \in E} p(e) = \sum_{e \in \{2,4,6\}} p(e) = \sum_{e \in \{2,4,6\}} \frac{1}{6} = 3\frac{1}{6} = \frac{1}{2}$$

Note that in the previous example, the probability of every outcome was equally likely. When this happens, $p(e) = \frac{1}{|\Omega|}$ for all elements in Ω. This is since the sum of the probabilities of all outcomes is 1.

Definition 2. If

$$\forall e \in \Omega, p(e) = \frac{1}{|\Omega|}$$

then we say p is *uniform*.

When the probability mass function is uniform, we can calculate probability of events by simply computing sizes.

Theorem 1. *Let Ω be a sample space and p be an uniform probability mass function. Then for any event E,*

$$Pr[E] = \frac{|E|}{|\Omega|}$$

Proof.

$$Pr[E] = \sum_{e \in E} p(e) = \sum_{e \in E} \frac{1}{|\Omega|} = \frac{1}{|\Omega|} \sum_{e \in E} 1 = \frac{1}{|\Omega|}|E| = \frac{|E|}{|\Omega|}$$

□

Example 2. 8 people enter a restaurant leaving their hats at the front. When leaving, each person takes a random hat. Each person chooses each hat with equal probability. What is the probability that everyone gets someone else's hat?

Here the sample space Ω consists of all permutation from 8 people to 8 people (e.g., a permutation describe which hat each person got). Thus $|\Omega| = 8!$.

Since each hat is chosen with equal probability, the probability mass function in this case is uniform.

The event E consists of all derangements. Recall that a derangement is a permutation with no fixed point. From our calculations for the hats problem,

$$|E| = \sum_{i=0}^{n} (-1)^i \frac{n!}{i!}$$

Thus, the probability that no one takes their own hat is

$$\frac{|E|}{|\Omega|} = \sum_{i=0}^{n} (-1)^i \frac{1}{i!}$$

Two different problems could have the same sample space but different probability mass functions. For example, if we were throwing a loaded die rather than an ordinary ("fair") die, the outcomes are still 1,2,3,4,5 and 6 but not all outcomes are equally likely anymore.

Example 3. Suppose we throw an unusually 100 sided die where the probability of getting the number i is

$$\frac{i}{\sum_{j=1}^{100} j}$$

. We would like to know how likely are we to get a multiple of 2 or 5.

Note that theorems we have seen about sets such as the inclusion-exclusion principle carry over to probability theory. When the probability mass function is uniform, we can simply apply the corresponding theorem about sets. However, if the probability mass function is not uniform, it is necessary to reprove the theorem (usually, by following the same steps).

Example 4 (Monty Hall problem)**.** On a game show, a contestant must choose between 3 doors, one of which contains a prize. The host of the show then opens an unchosen door with no prize behind it. The contestant is asked if they wish to switch doors afterwards.

Should the contestant switch? What is the probability of winning if they switch? What is the probability of winning if they do not switch?