CS648 : Randomized Algorithms
Semester I, 2011-12, CSE, IIT Kanpur

A gentle introduction to elementary probability theory - I

In this chapter, we started a gentle introduction to discrete probability theory. We introduced some fundamental terms and formulas. The reader is advised to internalize all these basic concepts.

1 Introduction

Happening of unexpected events around us is not uncommon. It therefore makes sense to abandon the idea of a world with certainties, and accept a world where we can, at best, associate likelihood to each event. Probability is a quantitative measure of likelihood of an event. We are all familiar with these phrases and we have also fair amount of understanding about them :

- Probability of getting a head in a toss of a fair coin is 1/2.
- Probability of a human getting infected by HIV during blood transfer is 0.000000001.
- Probability that six appears in two consecutive throws of a fair dice is 1/36.

For each such random experiment, there is a well defined set of all possible events (outcomes) which the experiment can results in. It is also quite natural to say that probability of any event has to be non-negative. Based on these understandings, we can define the concept of a probability space which formalizes the notion of probability as follows. We shall deal with discrete probability theory for most of the times, so we follow a simple set theoretic approach to introduce discrete probability theory.

Definition 1.1 (Probability Space)
Probability space is a pair $(\Omega, \mathbf{P})$ where Ω is a set called the sample space, and $\mathbf{P}$ is a real-valued function on the elements of Ω with the following rules.

1. $\mathbf{P}(\omega) \geq 0$*, for each* $\omega \in \Omega$.
2. $\sum_{\omega\in\Omega} \mathbf{P}(\omega) = 1$.

The elements of the sample space Ω are called the sample points or elementary events.

Though each experiment will produce exactly one elementary event, we are sometimes interested in some event which is essentially a collection of elementary events. For example, if we toss a coin six times, there will be 2^6 elementary events, and an event can be defined as *"at least 3 heads appear"*.

Definition 1.2 *(***Event***)*
Given a probability space $(\Omega, \mathbf{P})$, an event is a subset of set Ω. It is quite natural to define probability of an event $\mathcal{E}$ as

$$\mathbf{P}(\mathcal{E}) = \sum_{\omega\in\mathcal{E}} \mathbf{P}(\omega)$$

In many cases, each elementary event is equally likely. In such cases, it is quite easy to compute the probability of an event : If there are k elementary events defining the sample space Ω and each event is equally likely, then $\mathbf{P}(w) = 1/k$; the probability of an event $\mathcal{E}$ can thus be calculated by multiplying $1/k$ with the number of elementary events belonging to $\mathcal{E}$. A few Examples are as follows.

1. **Throw of two fair coins**
   The sample space is {HH, HT, TT, TH} and the probability of each elementary event is 1/4. The events of interest could be : *"at least one head appears in the outcome"*, or *"head appears both the times"*, *"two outcomes are different"*.

2. **Throw of two dice**
   The sample space is $\{(i, j)|1 \leq i, j \leq 6\}$ and the probability of each elementary event is 1/36. A few events of interest are : *"the sum of the numbers appearing on top faces is even"*, *"the numbers appearing on the top faces are different"*, *"the numbers appearing on the top face differ at least by 2"*.

3. **Sampling Coupons**
   A bin has coupons labeled from 1 to n. We sample k coupons uniformly randomly without replacement. We are interested in the probability that the smallest value of the coupons is greater than some number s. What is the sample space and what is the probability of this event?

4. m **balls into** n **bins**
   There are m balls and n bins. All m balls are thrown randomly uniformly and independently into n bins. (This means that each bin selects its destination bin randomly uniformly from n bins and this selecton is not influenced by the selection made by other balls). We may be interested in the event : *at least one bin is empty* or the event *at least 5 bins are having two balls each*.

While analyzing any random experiment, it is very important to have a clear picture of the sample space associated with it. To realize this importance through an insightful example, visit the following site :

`http://en.wikipedia.org/wiki/Bertrand_paradox_probability`

2 Combination of Events

The definition of probability space and event as described above are based on elementary set theory. Therefore, the notions of union of events and intersection of events are defined quite naturally as follows : Let $\mathcal{E}_1, \mathcal{E}_2$ be two events defined over a probability space $(\Omega, \mathbf{P})$.

Union of events: The event $\mathcal{E}_1 \cup \mathcal{E}_2$ corresponds to the set $\{\omega \in \Omega | \omega \in \mathcal{E}_1 \text{ or } \omega \in \mathcal{E}_2\}$.

Intersection of events: The event $\mathcal{E}_1 \cap \mathcal{E}_2$ corresponds to the set $\{\omega \in \Omega | \omega \in \mathcal{E}_1 \text{ and } \omega \in \mathcal{E}_2\}$.

Given a probability space $(\Omega, \mathbf{P})$, let $\mathcal{E}_1$ and $\mathcal{E}_2$ be any two events.

$$\mathbf{P}[\mathcal{E}_1 \cup \mathcal{E}_2] = \mathbf{P}[\mathcal{E}_1] + Pr[\mathcal{E}_2] - Pr[\mathcal{E}_1 \cap \mathcal{E}_2]$$

From elementary set theory one can easily verify the above equation as follows. An easy way to prove it is to show that each $\omega \in \mathcal{E}_1 \cup \mathcal{E}_2$, contributes exactly $\mathbf{P}[]$ in the expression on the R.H.S. of this equation. (It is quite obvious to see that an event $\omega \notin \mathcal{E}_1 \cup \mathcal{E}_2$ does not contribute any thing in the R.H.S.). For showing this, a simple case analysis depending upon whether ω belongs to one or both of $\mathcal{E}_1$ and $\mathcal{E}_2$ is required (done in the class). In a similar way try to show that for any three events $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$,

$$\mathbf{P}[\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3] = \sum_{1 \leq i \leq 3} \mathbf{P}[\mathcal{E}_i] - \sum_{1 \leq i < j \leq 3} \mathbf{P}[\mathcal{E}_i \cap \mathcal{E}_j] + \mathbf{P}[\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3] \tag{1}$$

(Hint : There are basically three cases : ω is contained in exactly one of the three events (sets) or is contained in exactly two of the three events, or is contained in all three events. For each of these possible cases, verify that ω should contribute $\mathbf{P}[\omega]$ exactly once on right side of above equation.

3 Homeworks

1. Visit the folowing website

   `http://en.wikipedia.org/wiki/Bertrand_paradox_probability`

   to realize that it is important to have a clear picture of the probability space underlying any randomized algorithm/experiment which we wish to analyse.

2. Although we shall use discrete probability theory for most of the times in the course, sometimes we shall study some examples which deal with continuous probability theory. But we shall use basic common sense and intuition to work on these examples. Try to see if you can solve the following problem which deals with continuous probability theory.

   We select a point randomly uniformly from the interval $[0, 1]$. This will split the interval into two intervals. What is the average length of the smaller interval formed ?

   There is a simple calculation free answer for the above question. The following fact is quite hard to prove but you are encouraged to appreciate it (and not solve it) : If we select n points randomly uniformly from $[0, 1]$, the average length of the smallest interval will be $1/(n + 1)^2$.

3. How can we extend the formula given in Equation 1 to more than three events? In particular, what is the formula for $\mathbf{P}[\mathcal{E}_1 \cup \mathcal{E}_2 \cup \cdots \cup \mathcal{E}_n]$ for any arbitrary n? Let us define the following notations.

$$
\begin{aligned}
S_1 &= \sum_{1 \le i \le n} \mathbf{P}[\mathcal{E}_i] \\
S_2 &= \sum_{1 \le j_1 < j_2 \le n} \mathbf{P}[\mathcal{E}_{j_1} \cap \mathcal{E}_{j_2}] \\
S_3 &= \sum_{1 \le j_1 < j_2 < \cdots < j_i \le n} \mathbf{P}[\mathcal{E}_{j_1} \cap \mathcal{E}_{j_2} \cap \cdots \cap \mathcal{E}_{j_i}]
\end{aligned}
$$

   We shall prove the following theorem tomorrow. But it is strongly recommended that you make sincere attempt to solve it on your own before coming to tomorrow's class.

   Theorem 3.1 *For any n events $\mathcal{E}_1, ..., \mathcal{E}_n$ defined over a probability space $(\Omega, \mathbf{P})$, and S_i's as defined above,*

$$\mathbf{P}[\mathcal{E}_1 \cup \mathcal{E}_2 \cup \cdots \cup \mathcal{E}_n] = S_1 - S_2 + S_3 ... + (-1)^{n+1} S_n$$