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

A gentle introduction to elementary probability theory - II

In this lecture we continue refreshing our concepts of probability theory.

1 Probability of union of n events

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 now state and prove the following theorem.

Theorem 1.1 *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,*

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

Proof: Our proof strategy will be similar to that for the case of union of two or three events discussed in the previous lecture. Consider any elementary event $\omega \in E_i$. We shall show that ω contributes $\mathbf{P}[w]$ in the expression $S_1 - S_2 + ... + (-1)^{n+1} S_n$.

Let ω belong to exactly k events from the set $\{\mathcal{E}_1, \ldots, \mathcal{E}_n\}$. Furthermore, let these k events be $\mathcal{E}_{j_1}, ..., \mathcal{E}_{j_k}$. Let us now try to see the contribution of Ω in each S_i. To start with, what is the contribution of ω in S_1? Well, it is $k\mathbf{P}[w]$ since ω belongs to exactly k events.

What is contribution of ω in S_2. Here, ω will contribute $\mathbf{P}[\omega]$ for each pair of events from the set $\{\mathcal{E}_{j_1}, \ldots, \mathcal{E}_{j_k}\}$. There are total $\binom{k}{2}$ pairs. Hence ω contributes $\binom{k}{2}\mathbf{P}[\omega]$ to S_2.

Along similar arguments, we can see that ω contributes $\binom{k}{i}\mathbf{P}[\omega]$ to S_i for any given $i \le k$. Moreover, it is obvious that ω does not contribute to any S_i for $i > k$. Hence the net contribution of the elementary event ω to R.H.S. is

$$\left[\binom{k}{1} - \binom{k}{2} + \cdots + (-1)^{k+1}\binom{k}{k}\right]\mathbf{P}[\omega]$$

Using binomial expansion of $(1-1)^k$, it is easy to follow that the summation with in the [] parenthesis is exactly equal to 1. Hence ω contributes $\mathbf{P}[w]$ to R.H.S. and we are done.

2 Conditional Probability

Consider a probability space $(\Omega, \mathbf{P})$. Probability of an event may get influenced by occurrence of another events. Consider the following examples which most of us will agree with. Probability of *"rain on a day given that the day is cloudy"* is more than just the probability of *"rain on a day"*. We may sometimes be interested in probability of *"event A conditioned on event B"*, which actually means the probability of event A given that event B has happened. (Clearly $\mathbf{P}[B] > 0$ in such cases). First note that the happening of event B, in effect, has reduced the sample space under consideration. It is no more Ω.

Instead it is the subset of Ω corresponding to the event B. In this new sample set, we are interested in those elementary events which belong to A. This set is clearly $A \cap B$. Therefore the probability of A conditioned on event B, can be expressed as

$$\mathbf{P}[A|B] = \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]} \tag{1}$$

In the later portion, we shall introduce the concept of random variable and expected value of a random variable.

Probability of happening of two events simultaneously

Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be any two events defined over a probability space $(\Omega, \mathbf{P})$. Equation 1 can be reformulated as follows (depending upon the requirement)

$$\mathbf{P}[\mathcal{E}_1 \cap \mathcal{E}_2] = \mathbf{P}[\mathcal{E}_1|\mathcal{E}_2]\ \mathbf{P}[\mathcal{E}_2] = \mathbf{P}[\mathcal{E}_2|\mathcal{E}_1]\ \mathbf{P}[\mathcal{E}_1] \tag{2}$$

The above equation is used to compute probability of simultaneous happening of two (or more) events.

3 Three Important theorems

We shall first state two important theorems. The proofs of these theorems will follow from elementary set theory and definition of probability space. **To get a better insight and confidence, the reader is encouraged to construct these proofs.**

Theorem 3.1 *(**Union theorem**) Let $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ be n events defined over a probability space $(\Omega, \mathbf{P})$. Then*

$$\mathbf{P}[\cup_{i=1}^{n} \mathcal{E}_i] \leq \sum_{i=1}^{i=n} \mathbf{P}[\mathcal{E}_i]$$

For the following theorem, it is important to understand the notion of partition of a probability space.

Definition 3.1 *A set of events $S = \{\mathcal{E}_1, \ldots, \mathcal{E}_n\}$ defined over a probability space $(\Omega, \mathbf{P})$ form a partition of the sample space Ω if the following conditions are satisfied.*

- *For each $i \neq j$, $\mathcal{E}_i \cap \mathcal{E}_j = \emptyset$.*
- $\cup_{i \leq n} \mathcal{E}_i = \Omega$

Theorem 3.2 *Let $(\Omega, \mathbf{P})$ be a given probability space, and let $S = \{\mathcal{E}_1, \ldots, \mathcal{E}_n\}$ be a partition of the sample space Ω. Then for any event $\mathcal{E}$,*

$$\mathbf{P}[\mathcal{E}] = \sum_{i=1}^{n} \mathbf{P}[\mathcal{E} \cap \mathcal{E}_i]$$

Using Equation 2, the above theorem can be reformulated as

$$\mathbf{P}[\mathcal{E}] = \sum_{i=1}^{n} \mathbf{P}[\mathcal{E}|\mathcal{E}_i]\ \mathbf{P}[\mathcal{E}_i]$$

The two theorems mentioned above are so simple that it is quite natural to doubt their usefulness. However, as will become clear from the course, these theorems will play a crucial role in solving various problems related to analysis of various randomized algorithms.

We now proceed to third theorem called Bayes' theorem. You might have studied it in your past. Let us revisit it with fresh mind.

Bayes Theorem

Consider the following problem. We have two coins **I** and **II**. Probability that tossing the coin **I** gives HEAD is 1/2, while probability that tossing the coin **II** gives HEAD is 1/4. The random experiment starts with picking a random coin from { **I,II** }, tossing the coin, and reporting the outcome (HEAD or TAIL). Furthermore it is given that we pick coin **I** with probability 2/3 (and pick coin **II** with probability 1/3). Now if you are told that the outcome of the experiment was a HEAD, then what is the probability that coin **I** was picked in the experiment ?

One can observe that the question asks for calculating a conditional probability : given that the outcome was a HEAD, what is the probability that it was coin **I** which was picked during the experiment.

Let us formalize it.

What is the sample space ? It has four elementary events. Let us formalize the events which we need to consider in the given problem. There are the following three events.

1. $\mathcal{E}_I$: the event that the coin **I** is picked
2. $\mathcal{E}_{II}$: the event that the coin **II** is picked.
3. **H** : the event that the outcome of the experiment was a HEAD.

It can be observed that the two events $\mathcal{E}_I$ and $\mathcal{E}_{II}$ define a partition of the sample space. What is it that we need to calculate ? We need to calculate $\mathbf{P}[\mathcal{E}_I|\mathbf{H}]$. Using the equation for the conditional probability,

$$\mathbf{P}[\mathcal{E}_I|\mathbf{H}] = \frac{\mathbf{P}[\mathcal{E}_I \cap \mathbf{H}]}{\mathbf{P}[H]}$$

Now how to calculate the terms in the numerator and denominator of the above equation. Now using Equation 2 for the intersection of two events

$$\mathbf{P}[\mathcal{E}_I \cap \mathbf{H}] = \mathbf{P}[\mathbf{H}|\mathcal{E}_I]\mathbf{P}[\mathcal{E}_I] = 1/2 \times 2/3 = 1/3$$

Also using Partition theorem and then Equation 2, we get

$$\begin{aligned} \mathbf{P}[\mathbf{H}] &= \mathbf{P}[\mathbf{H} \cap \mathcal{E}_I] + \mathbf{P}[\mathbf{H} \cap \mathcal{E}_{II}] \\ &= \mathbf{P}[\mathbf{H}|\mathcal{E}_I]\mathbf{P}[\mathcal{E}_I] + \mathbf{P}[\mathbf{H}|\mathcal{E}_I]\mathbf{P}[\mathcal{E}_I] \\ &= 1/2 \times 2/3 + 1/4 \times 1/3 = 5/12 \end{aligned}$$

Hence

$$\mathbf{P}[\mathcal{E}_I|\mathbf{H}] = (1/3)/(5/12) = 4/5$$

We can summarize the above discussion in the following theorem.

Theorem 3.3 *(Bayes' Theorem) Let the events $\mathcal{E}_1, ..., \mathcal{E}_k$ form a partition of the sample space Ω, and let $\mathcal{E}$ be any event. Then*

$$\mathbf{P}[\mathcal{E}_i|\mathcal{E}] = \frac{\mathbf{P}[\mathcal{E}|\mathcal{E}_i]\mathbf{P}[\mathcal{E}_i]}{\sum_j \mathbf{P}[\mathcal{E}|\mathcal{E}_j]\mathbf{P}[\mathcal{E}_j]}$$

This theorem is usually used when it is not easy to calculate $\mathbf{P}[\mathcal{E}_i|\mathcal{E}]$ directly, but it is easy to calculate $\mathbf{P}[\mathcal{E}|\mathcal{E}_i]$ for various values of i.

Homework Problems

1. There are n different letters and n envelopes with unique address written on each envelope. Each letter has to be placed into the right envelope (with right address). But we assign them randomly : We mix the pile of letters and the pile of envelopes separately thoroughly and then pick a letter and an envelope, place the letter in the envelope, and post it. We do so until no letter is left. What is the probability that at least one recipient receives the right letter ?
2. There are n balls and n bins. We throw each ball uniformly randomly and independently (of other balls) into any of the n bins. What is the probability that there will be at least one empty bin ?

3. Attempt all the problems given in the practice sheet of problems posted at the site

   `http://web.cse.iitk.ac.in/users/cs648/Assignments/Practice-sheet-1.pdf`

4. We throw n balls randomly uniformly and independently into n bins. Show that with high probability the maximum numbers of balls in the most heavily loaded bin will be $O(\log n)$. Here with high probability means $1 - n^{-c}$ for any arbitrary constant $c > 1$.

5. Let there be two events $\mathcal{E}_1$ and $\mathcal{E}_2$ defined over a probability space $(\Omega, \mathbf{P})$. We say that the two events are independent if happening of one event does not influences happening of the other event. This can be expressed formally as

   $$\mathbf{P}[\mathcal{E}_1|\mathcal{E}_2] = \mathbf{P}[\mathcal{E}_1] = \mathbf{P}[\mathcal{E}_1|\neg\mathcal{E}_2]$$

   Independent events are much easier to handle. In particular, the probability of simultaneous happening of two independent events is equal to the product of the probability of happening of each of them. That is,

   $$\mathbf{P}[\mathcal{E}_1 \cap \mathcal{E}_2] = \mathbf{P}[\mathcal{E}_1]\mathbf{P}[\mathcal{E}_2]$$

   For the case when each elementary event of the sample space has same probability, try to see how two independent events will look like in the graphical description of their corresponding sets ?

6. **A gem from the world of probability/randomized-algorithms :**

   This problem is not to be solved at this stage. Consider the following theorem.

   It is possible to place n points in a unit square so that the each triangle defined by any three points among them will have area at least $\Omega(1/n^2)$.

   Can you believe that there is a short and simple proof for the above theorem using randomization ? The world of randomization is full of such pleasant surprises ....

Important advice : Please never generate negativity or aversion in your mind if you are not able to solve any problem in this course. Every subject demands investment of persistent efforts and time to master it. Please approach these (practice as well as bed time problems) with open mind, full of scientific spirit (without worrying about the course). Just imagine as if these problems have been given to you by some of your close friends. You will sooner or later develop the insight which is very important to design and analyse any randomized algorithm.

Happy problem solving...:-)