CNT 5412 – Network Security   **Spring 2025**

Scribe 1: Probabilistic Analysis

Instructor: Viet Tung Hoang

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

1.1 A brief review of probability

You should practice the following questions, and then compare your answers with the provided ones.

Exercise 1.1 *Suppose that we toss a fair coin n times. What is the chance that we get exactly k Heads (and $n - k$ Tails)?*

Solution: There are $\binom{n}{k}$ ways to choose k positions out of n tosses. For each of those k positions, the chance that the coin lands Head is $1/2$, and thus the chance that the coin lands Head in all those k positions is 2^{-k}. For each of the remaining $n - k$ positions, the chance that the coin lands Tail is $1/2$, and thus the chance that the coin lands Tail in all those $n - k$ positions is $2^{-(n-k)}$. Hence the probability that we have exactly k Heads and $n - k$ Tails is

$$\binom{n}{k} \cdot 2^{-k} \cdot 2^{-(n-k)} = \binom{n}{k} \cdot 2^{-n} \ .$$

Exercise 1.2 *Alice throws 6 dice and wins if she scores at least one ace. Bob throws 12 dice and wins if he scores at least two aces. Who has the greater probability to win?*

Solution: Alice *loses* if none of her dice produces an ace, which happens with probability

$$\left(\frac{5}{6}\right)^6 \ .$$

Hence Alice wins with probability

$$1 - \left(\frac{5}{6}\right)^6 \approx 0.665 \ .$$

Next, Bob loses if either (1) none of his dice produces an ace, or (2) exactly one of his dice produces an ace. On the one hand, event (1) happens with probability

$$\left(\frac{5}{6}\right)^{12} \ .$$

On the other hand, event (2) happens with probability

$$12 \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{11} \ .$$

Since events (1) and (2) are disjoint (meaning that they can't simultaneously happen), the chance that Bob wins is

$$1 - \left(\left(\frac{5}{6}\right)^{12} + 12 \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{11}\right) \approx 0.619 \ .$$

Thus Alice has better chance of winning.

Exercise 1.3 *Suppose that we throw n balls uniformly at random to n bins. What's the chance that exactly one bin remains empty?*

Solution: If exactly one bin remains empty, then among the other $n-1$ bins, one bin will have exactly 2 balls, and each other has exactly 1 ball. There are $\binom{n}{2}$ ways to pick the two balls that are destined to fall within the same bin. Among these balls, the first ball has n choices for its bin, but the second ball has just a single choice. For the other $n-2$ balls, the first ball has $n-1$ choices for the bin that it will fall into, the second ball has just $n-2$ choices, and so on. Hence there are totally $\binom{n}{2} \cdot n!$ ways that the balls can fall to leave exactly one empty bin, among n^n possible ways of throwing balls into bins. Hence the chance that exactly one bin remains empty is

$$\binom{n}{2} \cdot \frac{n!}{n^n} \ .$$

Exercise 1.4 *Suppose that we pick X and Y uniformly and independently from $\{0,1\}^n$ (the set of all n-bit binary strings). What's the chance that $X = Y$?*

Solution: For each string $C \in \{0,1\}^n$, the chance that $X = Y = C$ is 2^{-2n}. Summing this over 2^n possible choices for C, the chance that $X = Y$ is 2^{-n}.

Alternatively, you can view X as fixed, and Y as uniformly random over $\{0,1\}^n$. Thus $\Pr[Y = X] = 2^{-n}$.

Exercise 1.5 *A bias coin has probability of landing Head as p, where p is a small number, say $p = 2^{-10}$. How many times do we need to toss it to get at least one Head with reasonable probability?*

Solution: An intuitive answer is to toss $1/p$ times. Indeed, if the coin is fair ($p = 1/2$), we expect to toss twice. But what's the success probability?

To answer this question, suppose that we toss the coin t times. The chance that it always lands in Tail is $(1-p)^t$, and thus the chace that there's at least one Head is $1 - (1-p)^t$. If we choose $t \approx 1/p$, then $1 - (1-p)^t \approx 1 - 1/e \approx 0.63$. If we instead pick $t \approx 4/p$ then $1 - (1-p)^t \approx 1 - 1/e^4 \approx 0.98$.

1.2   Probabilistic inequalities

UNION BOUND. Suppose that you have n events $E_1, \ldots, E_n$. The Union bound states that

$$\Pr[E_1 \cup E_2 \cup \cdots \cup E_n] \leq \sum_{i=1}^{n} \Pr[E_i] \ ,$$

where the union $E_1 \cup \cdots \cup E_n$ means that "at least one of the events E_i occurs". Note that here the events $E_1, \ldots, E_n$ are *arbitrary*, and do *not* need to be independent.

To understand the union bound, consider the illustration in Figure 1.1 for $n = 3$. The union bound says that the area (i.e., the probability mass) in the union of the circles is bounded above by the sum of the areas of the circles.

BIRTHDAY BOUND. Suppose that we sample q numbers $X_1, \ldots, X_q$ independently and uniformly from a set S of N elements. Let Bad be the event that we have some $i < j$ such that $X_i = X_j$.

We can compute the exact value of $\Pr[\mathsf{Bad}]$ as follows. In fact, it is easier to compute $\Pr[\overline{\mathsf{Bad}}] = 1 - \Pr[\mathsf{Bad}]$. Suppose that we sample $X_1, \ldots, X_q$ in that order, and $\overline{\mathsf{Bad}}$ happens. First, when we pick $X_2 \leftarrow_{\$} S$, the
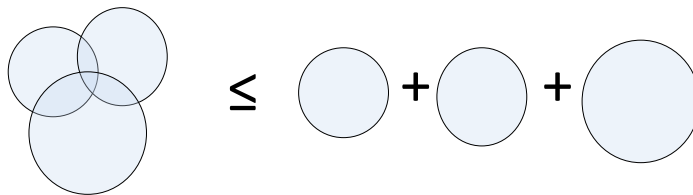
Figure 1.1: Area of union is bounded by sum of areas.

chance that it doesn't fall into $\{X_1\}$ is $(N-1)/N$. Next, when we pick $X_3 \leftarrow_\$ S$, the chance that it doesn't fall into $\{X_1, X_2\}$ is $(N-2)/N$, and so on. Hence

$$\Pr[\overline{\mathsf{Bad}}] = \prod_{i=1}^{q-1} \frac{N-i}{N} \quad,$$

and thus

$$\Pr[\mathsf{Bad}] = 1 - \prod_{i=1}^{q-1} \frac{N-i}{N} \quad.$$

The exact formula above is however complicated. We therefore want to give a simple, tight upper bound instead. For $i \le q-1$, let Bad_i be the event that when we sample $X_{i+1} \leftarrow_\$ S$, it falls into $\{X_1, \ldots, X_i\}$. Then by using the union bound,

$$\Pr[\mathsf{Bad}] = \Pr[\mathsf{Bad}_1 \cup \cdots \cup \mathsf{Bad}_{q-1}] \le \sum_{i=1}^{q-1} \Pr[\mathsf{Bad}_i] \quad.$$

Now, for each $i \le q-1$, since there are at most i elements in the set $\{X_1, \ldots, X_i\}$, if we pick $X_{i+1} \leftarrow_\$ S$, the chance that it falls into the set $\{X_1, \ldots, X_i\}$ is at most i/N. In other words, $\Pr[\mathsf{Bad}_i] \le i/N$. Hence

$$\Pr[\mathsf{Bad}] \le \sum_{i=1}^{q-1} \frac{i}{N} \le \frac{q^2}{N} \quad.$$