CS3230 Design & Analysis of Algorithms	**January 28, 2021**

Discrete Probability Review

Prepared by: Fidella Widjojo, Audrey Felicio Anwar, Sng Weicong (Past TAs)

1.1 Introduction

We will be discussing Randomized Algorithms in Week 4, where algorithms are designed to avoid exact but expensive solutions, in favor of probabilistic but *orders* faster algorithms. A keen understanding of basic statistics concepts would be required to understand how to analyse these algorithms that have an associated probability of success.

1.2 Combinatorics

Probability is fundamentally a comparison between counts in the discrete case, or distribution mass in the continuous case. For the purpose of CS3230 we will focus only on the discrete setting. In many cases we compare the count of a set of events against the count of a superset of events. We will do a brief recap on counting, permutations and combinations in this section.

1.2.1 Counting

Consider a task that takes place in k stages, where the number of outcomes n_i at some stage i is independent of the outcomes of the previous stages. Then the total number of ways the task can be carried out is given by the product $= n_1 \cdot n_2 \cdot \ldots \cdot n_r$.

Example 1. How many possible 7-digit license plate numbers are there if the first 3 digits are alphabets and the next 4 digits are numbers? **Solution:** For each of the 3 alphabets, there are 26 ways. For each of the 4 numbers, there are 10 ways. Therefore, total number of possible license plates are $26^3 \times 10^4 = 175760000$.

Example 2. How many possible undirected 5-vertices graphs are there, assuming no loops and that the vertices are unique? **Solution:** We first count the number of possible edges, and there are $\binom{5}{2} = 10$ possible edges. Each edge can either be present or not, therefore the total number of possible graphs is 2^{10}.

1.2.2 Permutation

Suppose there are n distinct objects. The total number of distinct orderings are $n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1 = n!$ The intuition is that for the first object, we have n objects to choose from, and then for the second object, we have $n-1$ to choose from, ... and for the last object, we have 1 to choose from.

Example 3. There are 6 CS majors and 4 Math majors in a class. Students are ranked by their *unique* test scores. **a)** How many possible rankings are there? **b)** If CS majors are ranked among themselves (so are the Math majors), how many possible rankings are there? **Solution: a)** $10! = 3628800$ **b)** One can observe that CS major rankings and Math major rankings are independent of each other, and therefore we simply calculate the possible rankings of each bucket of students, and then multiply them to get the total number. $6! \times 4! = 17280$.

Example 4. How many ways are there to arrange 4 persons seated in a circle? **Solution:** While there are 4! ways of arranging 4 persons in a *line*, in the circle scenario, ABCD, BCDA, CDAB, DABC all result in the same seating arrangements. We are thus double-counting by a factor of 4 times, and therefore, total number of ways is $\frac{4!}{4} = 6$.

In the next example, we look at the case where there exist some objects that are indistinguishable from each other.

Example 5. How many different arrangements can be formed from the letters **a)** M I S S? **b)** M I S S I S S I P P I? **Solution:** **a)** If we treat each letter as unique, we would have obtained $4! = 24$. However one can observe that we have 2 indistinguishable 'S's, and we do not care about the ordering with respect to these 2 'S's. Out of the 24 orderings, there will be double-counting by a factor of $2! = 2$. The number of different arrangements is therefore $\frac{4!}{2!} = 12$. **b)** There are 4 'I's, 4 'S's, 2 'P's. Therefore, $\frac{11!}{4!4!2!} = 34650$.

1.2.3 Combination

We now look at the counting tasks where ordering does not matter.

Example 6. In how many ways can we choose 2 elements out of the set $S = \{a, b, c\}$? **Solution:** We first list all the possible subsets of S: $\phi, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}$. Out of these 8 possible subsets, 3 of them have 2 elements in the set. Therefore there are 3 ways.
We make the following observations:

1. We can also observe a triviality: each subset has k elements, where $k \in \{0, 1, 2, 3\}$.
2. In fact, for subsets with exactly 2 elements, there are $3 = \frac{3!}{2!(3-2)!} = \binom{3}{2}$ possible choices. One explanation for this is that we first pretend to care about ordering of the within-subset and outside-subset elements, which is just a naive count of the total number of permutations of 3 elements, $3! = 8$. We then remove for the double-counting as a result of caring for ordering within-subset (2!) and outside-subset ($(3 - 2)!$).
3. It is not by chance that there are 8 possible subsets, given that the set has 3 elements. In order to fully enumerate the possible subsets, one can observe that for a given element, it is either within or not within a subset. Since there are 3 elements, the total number of subsets is 2^3.

The following general equality links these 3 observations together:

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n \tag{1.1}$$

Example 7. A committee of 3 is to be formed from a group of 20 people. **a)** How many possible committees can be formed? **b)** Suppose Alice and Bob refuse to serve in the same committee. Under this restriction, how many possible committees can there be? **Solution:** **a)** $\binom{20}{3} = 1140$ **b)** We count the number of cases where both of them are in the same committee, and then take the complement. $\binom{20-2}{3-2} = 18$ cases where both of them are in the committee, therefore there are $1140 - 18 = 1122$ cases.

Example 8. In how many ways can we seat 30 males and 40 females in a lecture theatre with 120 seats? Assume that the students are indistinguishable apart from gender. **Solution:** We first allocate the males, which is $\binom{120}{30} = \frac{120!}{30!(120-30)!}$, and then we allocate the females among the remaining vacancies, which is $\binom{120-30}{40} = \frac{(120-30)!}{40!(120-30-40)!}$. Multiplying them together, we get $\frac{120!}{30!40!(120-30-40)!}$. This is in fact the multinomial coefficient $\binom{120}{30,40,50}$.

1.3 Probability

1.3.1 Axioms of probability

We start with a rigorous definition of probability:

Let S be a set (our full sample space). A function P is called *probability* if it assigns values to each subset of S, subject to:

- $\forall A \subset S, 0 \leq P(A) \leq 1$.
- $P(S) = 1$.
- If $A_1, A_2, ...$ are disjoint subsets of S, then $P(A_1 + A_2 + ...) = P(A_1) + P(A_2) + ...$

Some consequences from the above constraints:

- $P(\phi) = 0$.
- For $A \subset S$, let A^C be complement of A, then $P(A^C) = 1 - P(A)$.
- Let $A \subset B \subset S$, then $P(A) \leq P(B)$.

Example 9. Suppose we have an urn with 5 white balls and 4 black balls. We draw 3 balls *without* replacement. **a)** What is the probability of 1 ball being white and the other 2 balls being black? **b)** What is the probability that at most 2 balls are black? **Solution: a)** We pretend that the balls are distinct, and count the number of distinct samples of 2 black balls and 1 white ball, over the count of the number of distinct samples of 3 balls regardless of color. The numerator would be $\binom{4}{2} \times \binom{5}{1}$. The denominator would be $\binom{9}{3}$. The probability would thus be $\frac{\binom{4}{2}\times\binom{5}{1}}{\binom{9}{3}} = \frac{5}{14}$. **b)** It may be easier to use the complement. We first compute the number of ways one would end up drawing 3 black balls, which will be $\binom{4}{3}$. The complement would be "drawing at most 2 black balls". The probability is thus $\frac{\binom{9}{3}-\binom{4}{3}}{\binom{9}{3}} = \frac{20}{21}$. **Remarks:** There are many different and **correct** ways one can go about to get to the final answer.

1.3.2 Inclusion-Exclusion

We are often interested in knowing the probability of the union of n events, where the events are not necessarily disjoint. We outline a systematic way of computing such unions.

[**2 sets**] Let $A, B \subset S$. Then $P(A \cup B) =$
$P(A) + P(B)$
$- P(A \cap B)$.

[**3 sets**] Let $A, B, C \subset S$. Then $P(A \cup B \cup C) =$
$P(A) + P(B) + P(C)$
$- P(A \cap B) - P(A \cap C) - P(B \cap C)$
$+ P(A \cap B \cap C)$.

[**4 sets**] Let $A, B, C, D \subset S$. Then $P(A \cup B \cup C \cup D) =$
$P(A) + P(B) + P(C) + P(D)$
$- P(A \cap B) - P(A \cap C) - P(A \cap D) - P(B \cap C) - P(B \cap D) - P(C \cap D)$
$+ P(A \cap B \cap C) + P(A \cap B \cap D) + P(A \cap C \cap D) + P(B \cap C \cap D)$
$- P(A \cap B \cap C \cap D)$.

We can easily verify these by drawing the Venn diagrams and it should then be obvious that we are iteratively accounting for double counting of some subset. In fact, the following holds for $n \geq 2$ sets,

- For line k, there are $\binom{n}{k}$ terms.
- The signs are alternating for each line.

Example 10. (Hat Check Problem) 4 men enter a restaurant and check their hats. The hat-checker being absent-minded, mixes up their hats and returns them to the 4 men randomly upon them leaving the restaurant. What is the probability that none of them gets back his own hat? **Solution:** We let each event A_i be {Man i gets back his own hat}, and that the union of the 4 events $P(A_1 \cup A_2 \cup A_3 \cup A_4)$ would be {at least 1 man gets back his own hat}. We use the Inclusion-Exclusion formula to complete the probability of this union, $4 \times \frac{1}{4} - 6 \times \frac{2!}{4!} + 4 \times \frac{1!}{4!} - 1 \times \frac{1!}{4!} = \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!}$. The probability of that none of them getting back his own hat would be the complement $P(A_1^C \cap A_2^C \cap A_3^C \cap A_4^C) = 1 - (\frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!})$.

1.3.3 Conditional Probability

Conditional probability is a measure of the probability of an event occurring, given that another event (by assumption, presumption, assertion or evidence) has already occurred. If the event of interest is A and the event B is known or assumed to have occurred, "the conditional probability of A given B", or "the probability of A under the condition B", is usually written as $P(A \mid B)$.

For example, the probability that any given person has a cough on any given day may be only 5%. But if we know or assume that the person is sick, then they are much more likely to be coughing. For example, the conditional probability that someone unwell is coughing might be 75%, in which case we would have that $P(Cough) = 5\%$ and $P(Cough \mid Sick) = 75\%$.

Example 11. Suppose we have an urn with 5 white balls and 4 black balls. We draw 2 balls without replacement. What is the probability of getting 1 black and 1 white ball? **Solution:** $P(1B, 1W) = P(ball_1 = B, ball_2 = W) + P(ball_1 = W, ball_2 = B) = P(ball_1 = B) \cdot P(ball_2 = W \mid ball_1 = B) + P(ball_1 = W) \cdot P(ball_2 = B \mid ball_1 = W) = \frac{4}{9}\frac{5}{8} + \frac{5}{9}\frac{4}{8} = \frac{5}{9}$

1.3.4 Bayes Formulas

Bayes' theorem is stated mathematically as the following equation:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

where A and B are events and $P(B) \neq 0$.

1. $P(A \mid B)$ is a conditional probability: the probability of event A occurring given that B is true. It is also called the posterior probability of A given B.

2. $P(B \mid A)$ is also a conditional probability: the probability of event B occurring given that A is true. It can also be interpreted as the likelihood of A given a fixed B because $P(B \mid A) = L(A \mid B)$.

3. $P(A)$ and $P(B)$ are the probabilities of observing A and B respectively without any given conditions; they are known as the marginal probability or prior probability.

4. A and B must be different events.

The following may sometimes be useful when we do not have access to $P(B)$ directly.

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B \mid A_1)P(A_1) + \cdots + P(B \mid A_n)P(A_n)}$$

1.3.5 Independence

Let A and B be 2 events. We say they are *independent* when the probability of one event does not affect the probability of the other. For example, suppose we have 2 coin tosses. It should be obvious that the probability of the second toss being H does not depend on the first toss. This implies that $P(A \cap B) = P(A)P(B \mid A) = P(A)P(B)$

We say 2 events A and B are independent, $A \perp\!\!\!\perp B$, if

$$P(A \cap B) = P(A)P(B)$$

Some consequences of $A \perp\!\!\!\perp B$:

- $A^C \perp\!\!\!\perp B, A \perp\!\!\!\perp B^C, A^C \perp\!\!\!\perp B^C$
- If $P(A) > 0$, then $P(B \mid A) = P(B)$.
- If $P(B) > 0$, then $P(A \mid B) = P(A)$.

Example 12. If $A \perp\!\!\!\perp B, B \perp\!\!\!\perp C$, is $A \perp\!\!\!\perp C$ true? **Solution:** No. Let there be 2 coin tosses, let A be {1st toss is H}, B be {2nd toss is H}, C be {1st toss is T}. One can verify that $A \perp\!\!\!\perp B, B \perp\!\!\!\perp C, A \not\perp\!\!\!\perp C$.

We say 3 events A, B and C are *jointly* independent if the following holds:

- $A \perp\!\!\!\perp B, A \perp\!\!\!\perp C, B \perp\!\!\!\perp C$ (pairwise independence)
- $P(A \cap B \cap C) = P(A)P(B)P(C)$

Refer to `https://math.stackexchange.com/questions/1783225/example-of-pairwise-independent-but-not-jointly-independent-random-variables` for *pairwise AND NOT jointly* independent case.

1.4 Expectations

1.4.1 Random Variates

When we conduct an experiment, we are often more concerned with some aspect of the outcome rather than the full outcome. For example, when we roll 2 dice, we often care about the sum, instead of the granular individual dice-roll outcome. When we test a new drug on participants, we may only be interested in summary statistics like the mean difference in blood pressure before and after administering of the drug.

More formally, a *random variate* represents a mapping from some element of the sample space to a real number, $X : S \to R$. For example, let X denote the sum of 2 dice rolls, which is $X(dice_1, dice_2) = dice_1 + dice_2$. We can now write the following, $P(X = 12) = 1/36$.

We focus on *discrete* random variates, where the outcomes can take a finite or countably infinite number of outcomes. For this example, our $x_i \in X$ can lie within $\{2, 3, ..., 12\}$, and is thus a discrete random variable.

1.4.2 Expectation

We now define the notion of expected value. Let X be a discrete random variable taking values in $\mathbb{Z}$. Let x_i denote the possible realizations of X. The expected value of X is thus

$$E(X) = \sum_{x_i \in \mathbb{Z}} x_i P(X = x_i) \tag{1.2}$$

Note that when we are taking expectations, it is always with respect to some distribution with respect to a random variable, and in the case of discrete distributions the probabilities must sum to 1.

Example 13. Let X denote the value of a fair die. The expected value is $E(X) = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + ... + \frac{1}{6} \cdot 6 = 3.5$.

One can also compute expectations on functions of random variates.

$$E(f(X)) = \sum_{x_i \in \mathbb{Z}} f(x_i) P(X = x_i) \tag{1.3}$$

Some additional properties of expectations:

- $E(a) = a$, where a is some real constant.
- $E(aX + b) = aE(X) + b$, where a and b are some real constants.
- If A and B are independent, then $E(AB) = E(A)E(B)$.

Example 14. Let X denote the value of a fair die. The expected value of the square of the die roll is $E(X^2) = \frac{1}{6} \cdot 1^2 + \frac{1}{6} \cdot 2^2 + ... + \frac{1}{6} \cdot 6^2 = \frac{91}{6}$.

Example 15. Let X denote the value of a fair die. We roll 10 dice. Find the expected value of the product of the 10 dice-rolls. **Solution:** One can calculate this expectation by enumerating across all possible product outcomes along with the associated probability, but this is clearly tedious. We however, can exploit the independence property from above. This means we only need to calculate the expectation of a single die, which is 3.5 from an earlier example. The solution is thus $E(X_1 X_2 ... X_{10}) = E(X_1)E(X_2)...E(X_{10}) = 3.5^{10}$.

1.4.3 Linearity of Expectation

Linearity of expectation is the property that the expected value of the sum of random variables is equal to the sum of their individual expected values, regardless of whether they are independent.

For random variables $X_1, X_2, ..., X_n$ and constants $c1, c2, ..., c_n$.

$$E[\Sigma_{i=1}^{n} c_i X_i] = \Sigma_{i=1}^{n} c_i E[X_i]$$

1.5 Interesting Exercises

1. If I have 5 indistinguishable balls and 4 distinguishable urns, how many ways can I place the balls into the urns? **Ans:** 56 https://artofproblemsolving.com/wiki/index.php/Ball-and-urn
2. Similar to Q1, but what if the balls are distinguishable? **Ans:** 1024
3. https://en.wikipedia.org/wiki/Monty_Hall_problem
4. **(Tough)** https://xfcrs.blogspot.com/2019/10/in-ncaa-basketball-tournament-are-64.html
5. **(Tough)** https://math.stackexchange.com/questions/272927/expected-length-of-arc-in-a-randomly-divided-circle

1.6 Tips and tricks

- It is extremely common to double-count or under-count by some factor, due to symmetrical arguments.
- There are multiple ways of counting, and therefore multiple ways of computing probabilities for a given problem. One can thus reasonably check for correctness by going about counting using 2 different ways and checking if the same answer is obtained.
- When computing probabilities, we can use the fact that they sum up to 1 to check ourselves.
- For a more rigorous treatment of probability which is required knowledge for a full-fledged Randomized Algorithms course (CS5330), one may refer to this free online resource https://math.dartmouth.edu/~prob/prob/prob.pdf.