

CS1231 Part 9 - Counting and Probability

Based on lectures by Terence Sim and Aaron Tan

Notes taken by Andrew Tan

AY18/19 Semester 1

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

1 Introduction

Probability is a measure of the likelihood that an event will occur, and counting in probability refers to the techniques used to count the set of possible outcomes.

To say that a process is **random** means that when it occurs, one outcome from a set of outcomes is sure to occur, but it is impossible to predict with certainty which outcome that will be.

We will introduce several definitions and ideas in our analysis of counting and probability:

- A **sample space** is the set of all possible outcomes of a random process or experiment.
- An **event** is a subset of a sample space.

For a finite set A , $N(A)$ denotes the number of elements in A .

If S is a finite sample space in which all outcomes are equally likely and E is an event in S , then the **probability** of E , denoted $P(E)$, is

$$P(E) = \frac{\text{The number of outcomes in } E}{\text{The total number of outcomes in } S} = \frac{N(E)}{N(S)}$$

From here, we can see that calculating the probability of certain events may be as simple as counting the elements of a list.

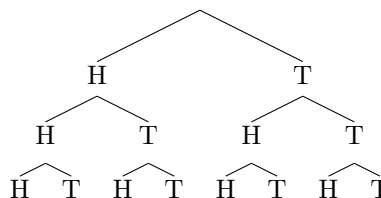
For example, we know that given two integers m and n and $m \leq n$, then there are $n - m + 1$ integers from m to n inclusive. Now let's find the probability that a randomly chosen 3-digit integer is divisible by 5:

We can see that the number of multiples of 5 from 100 to 999 is equal to the number of integers from 20 to 199 inclusive. Hence there are $199 - 20 + 1 = 180$ 3-digit integers that are divisible by 5. Furthermore, the total number of integers from 100 to 999 = $999 - 100 + 1 = 900$. Hence, the probability that a randomly chosen 3-digit integer is divisible by 5 is $180/900 = 1/5$ by our definition.

2 Possibility trees and the Multiplication rule

A **tree structure** is a useful tool for keeping systematic track of all possibilities in situations in which events happen in order. By drawing out the tree structure, we can easily identify specific properties that we are looking for.

For example, let us see the simple case of flipping 3 coins:



The number of ways that we can flip 3 coins is represented by the number of distinct paths from the *root* (the top) to a *leaf* (a terminal point). Note that we can easily see that the number of possible outcomes is simply the number of leaves in the tree, which in this case is 8.

To find the probability of obtaining 2 heads and 1 tail, for example, we can simply go through each path accordingly and list them explicitly. From here we show that there are 3 possible outcomes: HHT, HTH, THH. And thus the probability of obtaining 2 heads and 1 tail is $3/8$.

2.1 Multiplication rule

The Multiplication rule states that if an operation consists of k steps, and

- the first step can be performed in n_1 ways,
- the second step can be performed in n_2 ways (regardless of how the first step was performed),
- ...
- the k^{th} step can be performed in n_k ways (regardless of how the preceding steps were performed),

then the entire operation can be performed in

$$n_1 \times n_2 \times \cdots \times n_k \text{ ways.}$$

Note that it is often difficult or impossible to apply the multiplication rule when the number of ways to perform or choose the ways for a step *depends* on preceding steps. In these instances, it is usually clearer to simply use possibility trees, or reorder the steps to ensure that we can apply the multiplication rule.

2.2 Permutations

A **permutation** of a set of objects is an ordering of the objects in a row. For example, the set of elements a, b, c has 6 permutations:

$$abc \ acb \ bac \ bca \ cab \ cba$$

In general, given a set of n objects, we can derive the number of permutations using the multiplication rule: we can think of forming a permutation as an n -step operation:

- choose an element to be the first: n ways,
- choose another element to be the second: $n - 1$ ways,
- ...
- choose the remaining element to be the n^{th} : 1 ways,

By the multiplication rule, we have $n \times (n - 1) \times (n - 1) \times \cdots \times 2 \times 1 = n!$ ways to perform the entire operation.

Lets say now we wish to only select r elements from a set of n elements in order. We call this an **r-permutation** of a set of n elements, which is an ordered selection of r elements taken from the set. The number of r -permutations of a set of n elements is denoted as $P(n, r)$.

We can see that $P(n, r) = n \times (n - 1) \times (n - 2) \times \cdots \times (n - r + 1) = \frac{n!}{(n-r)!}$ by the multiplication rule.

3 Counting elements of disjoint sets

The basic rule underlying the calculation of the number of elements in a union or difference or intersection is the **addition rule**. This rule states that the number of elements in a union of mutually disjoint finite sets equals the sum of the number of elements in each of the component sets.

More formally, suppose a finite set A equals the union of k distinct mutually disjoint subsets A_1, A_2, \dots, A_k . the **addition rule** states that

$$N(A) = N(A_1) + N(A_2) + \dots + N(A_k)$$

The addition rule also implies the **difference rule**, which states that if A is a finite set and B is a subset of A , then

$$N(A - B) = N(A) - N(B)$$

Moreover, recall that the complement of a set is simply all the element in the universe that are not in the set. In the context of probabilities, the **complement of an event** is the set of outcomes in the sample space that is not in the event. We can see that since the probabilities of all outcomes add up to 1, if S is a finite sample space and A is an event in S , then

$$P(A^c) = 1 - P(A)$$

3.1 The Inclusion/Exclusion rule

The addition rule only states the number of elements are in a union of sets if the sets are mutually disjoint. If there exists an overlap between two sets, if we add the number of elements between them, we will be counting some elements more than once due to the overlap.

To get an accurate count of the elements in $A \cup B$, it is necessary to subtract the number of elements that are in both A and B (which intuitively is $A \cap B$). Thus, the **inclusion/exclusion** rule states that, if A, B , and C are any finite sets, then

$$N(A \cup B) = N(A) + N(B) - N(A \cap B), \text{ and} \\ N(A \cup B \cup C) = N(A) + N(B) + N(C) - N(A \cap B) - N(A \cap C) - N(B \cap C) + N(A \cap B \cap C).$$

Note the addition of the intersection of A, B , and C in the second equation. We add these for the same reason that we subtract the intersections between each two sets. As we subtract these intersections, we will subtract $N(A \cap B \cap C)$ one too many times, and thus we will have to add it back. Also, we've only explored the inclusion/exclusion rule for 2 or 3 sets, but it is not hard to generalize it to more sets (although the calculations themselves may become tedious).

4 The Pigeonhole Principle

The **Pigeonhole Principle** states that a function from one finite set to a smaller finite set cannot be one-to-one: there must be at least 2 elements in the domain that have the same image in the co-domain.

To break it down, lets say we have n pigeons and m pigeonholes, and $n > m$. We can easily see that not all pigeons will be able to go into the pigeonholes neatly; there will be at least one hole that contains two or more pigeons!

This is the essence of the pigeonhole principle. If we have a greater number of elements in the domain than the range, then some elements in the domain will have to map to the same element in the range.

We can further generalize this to the **Generalized Pigeonhole Principle**, which states that, for any function f from a finite set X with n elements to a finite set Y with m elements and for any positive integer k , if $k < n/m$, then there is some $y \in Y$ such that y is the image of at least $k + 1$ distinct elements of X .

Lets break this down similarly with an example: lets say we have 25 people and we want to know at least how many people the same birth month. As there are 12 months in a year, we see that $25/12 = 2.0833... > 2$. Hence, there has to be *at least* $2 + 1 = 3$ people that share a birth month.

4.1 Using Pigeonhole Principle on properties of functions

Recall the definitions of injective and surjective functions. We can understand them in the context of the pigeonhole principle:

- For any function f from a finite set X with n elements to a finite set Y with m elements, if $n > m$, then f is not one-to-one.
- Let X and Y be finite sets with the same number of elements and suppose f is a function from X to Y . Then f is one-to-one if, and only if, f is onto.