1. Counting

Chris Piech and Mehran Sahami

May 2017

1 Introduction

Although you may have thought you had a pretty good grasp on the notion of counting at the age of three, it turns out that you had to wait until now to learn how to really count. Aren't you glad you took this class now?! But seriously, below we present some properties related to counting which you may find helpful in the future.

The ideas presented in this chapter are core to probability. Counting is like the foundation of a house (where the house is all the great things we will do later in probability for computer scientists, such as machine learning). Houses are awesome. Foundations on the other hand are pretty much just concrete in a hole. But don't make a house without a foundation. Trust me on that.

2 Basic Building Blocks

Counting with STEPS

Product Rule of Counting:
If an experiment has two parts, where the first part can result in one of m outcomes and the second part can result in one of n outcomes regardless of the outcome of the first part, then the total number of outcomes for the experiment is mn.

Rewritten using set notation, the Product Rule states that if an experiment with two parts has an outcome from set A in the first part, where $|A| = m$, and an outcome from set B in the second part (regardless of the outcome of the first part), where $|B| = n$, then the total number of outcomes of the experiment is $|A||B| = mn$.

Counting with OR

Sum Rule of Counting:
If the outcome of an experiment can either be one of m outcomes **or** one of n outcomes, where none of the outcomes in the set of m outcomes is the same as the any of the outcomes in the set of n outcomes, then there are m + n possible outcomes of the experiment.

Rewritten using set notation, the Sum Rule states that if the outcomes of an experiment can either be drawn from set A or set B, where $|A| = m$ and $|B| = n$, and $A \cup B = \emptyset$, then the number of outcomes of the experiment is $|A| + |B| = m + n$.

The Inclusion Exclusion Principle

Inclusion-Exclusion Principle:
If the outcome of an experiment can either be drawn from set A or set B, and sets A and B may potentially overlap (i.e., it is not guaranteed that $A \cup B = \emptyset$), then the number of outcomes of the experiment is $|A \cup B| = |A| + |B| - |A \cap B|$.

Note that the Inclusion-Exclusion Principle generalizes the Sum Rule of Counting for arbitrary sets A and B. In the case where $A \cap B = \emptyset$, the Inclusion-Exclusion Principle gives the same result as the Sum Rule of Counting since $|\emptyset| = 0$.

Double Counting and Constraints

There are many reasons for having counted some elements more than once (aka "double counted"). One common case, is that there is a constraint in the problem that you must contend with. It goes without saying that if you over-count, then you have to subtract off the number of elements that were double counted. If you did something along the lines of: count every element some multiple, then you can divide your total number of elements by that multiple to get the correct final answer.

3 Combinatorics

Counting problems can be approached from the basic building blocks described in the first section. However some counting problems are so ubiquitous in the world of probability that it is worth knowing a few higher level counting abstractions. When solving problems, if you can cast them into the format of these

Permutations of Distinct Objects

Permutation Rule: A permutation is an ordered arrangement of n distinct object. Those n objects can be permuted in n x (n – 1) x (n – 2) x ... x 2 x 1 = n! ways.

This changes slightly if you are permuting a subset of distinct objects, or if some of your objects are indistinct. We will handle those cases shortly!

Permutations of Indistinct Objects

Permutation of Indistinct Objects: Generally when there are n objects and

n_1 are the same (indistinguishable) and
n_2 are the same and
...
n_r are the same, then there are $\dfrac{n!}{n_1!n_2!\dots n_r!}$ permutations

Combinations of Distinct Objects

Combinations: A combination is an unordered selection of r objects from a set of n objects. If all objects are distinct, then the number of ways of making the selection is:

$$\frac{n!}{r!(n-r)!} = \binom{n}{r} \text{ ways}$$

This is often stated as “n choose r”

Consider this general way to product combinations: To select r distinct, unordered objects from a set of n distinct objects, E.g. ”7 choose 3”,

1. First consider permutations of all n objects. There are n! ways to do that.
2. Then select the first r in the permutation. There is one way to do that.
3. Note that the order of r selected objects is irrelevant. There are r! ways to permute them. The selection remains unchanged.
4. Note that the order of (n-r) unselected objects is irrelevant. There are (n-r)! ways to permute them. The selection remains unchanged.

$$\text{total} = \frac{n!}{r!(n-r)!} = \binom{n}{r} = \binom{n}{n-r}$$

The total ways to chose 3 objects from a set of 7 distinct objects is:

$$\text{total} = \binom{7}{3} = \frac{7!}{3!(7-3)!} = 35$$

4 Group Assignment

You have probably heard about the dreaded “balls and urns” probability examples. What are those all about? They are the many different ways that we can think of stuffing elements into containers. I looked up why people called their containers urns. It turns out that Jacob Bernoulli was into voting and ancient Rome. And in ancient Rome they used urns for ballot boxes. Group assignment problems are useful metaphors for many counting problems.

Note that there are many flavors of group assignment (eg with replacement, without replacement).

Assignment of Distinct Objects

Problem: Say you want to put n distinguishable balls into r urns. (no wait don’t say that). Ok fine. No urns. Say we are going to put n strings into r buckets of a hashtable where all outcomes are equally likely. How many possible ways are there of doing this? Answer: You can think of this as n independent experiments each with r outcomes. Using our friend the Product Rule of Counting this comes out to r^n.

Assignment of Indistinct Objects

Divider Method:
A divider problem is one where you want to place n indistinguishable items into r containers. The divider method works by imagining that you are going to solve this problem by sorting two types of objects, your n original elements and $(r–1)$ dividers. Thus you are permuting $n+r–1$ objects, n of which are same (your elements) and $r–1$ of which are same (the dividers). Thus:

$$\text{Total ways } = \frac{(n+r-1)!}{n!(r-1)!} = \binom{n+r-1}{r-1}$$

5 Counting and Computer Science

Counting is important in the world of computer science for a few reasons. Primarily (1) in order to understand probability on a fundamental level, it is useful to first understand counting (2) while computers are fast, some problems require so much work that they would take an unreasonable amount of time to complete. Using counting we can better estimate how much work we have to do (3) some counting problems are so large and complex that we benefit from computation to solve them.