# CMPSCI 575/MATH 513

## Combinatorics and Graph Theory

Lecture #17: Distributions
(Tucker Section 5.4)
David Mix Barrington
21 October 2016

# Distributions

- Objects in Boxes
- Unlimited Repetition: All Functions
- No Repetition: One-to-One Functions
- Restricted Repetition
- Counting Onto Functions
- The Twelvefold Way

# Objects in Boxes

- Today we're going to look at a new framework to describe our basic counting problems, that of **distributions**.
- The general picture is that we have r objects that we are going to place into n boxes. We are thus choosing a function $f: O \to B$, where O has size r and B has size n.
- But there are multiple versions of the problem of counting these functions, depending on our definitions.

# Objects in Boxes

- Suppose the objects in O are labeled with the numbers 1,…,r. (They are **distinct**.) Then if we put object 1 in box f(1), 2 in f(2), and so forth, our function is defined by the sequence of numbers f(1)f(2)…f(r).
- But if the objects are **identical**, then each of our functions is characterized by just the **number** of objects we put in each box.

# Examples of Objects in Boxes

- Suppose we have 100 diplomats who must each be assigned to one of five continents. If the diplomats are distinct, there are $5^{100}$ ways to do this.
- If we further insist that there be 20 diplomats assigned to each continent, we can consider that the 100 diplomats are being mapped into 100 positions by a bijection, in one of 100! ways, and then correct for the over counting to get $100!/(20!)^5$, which Tucker also writes as $P(100; 20, 20, 20, 20, 20)$.

# Examples of Objects in Boxes

- In **bridge**, the 52 cards are dealt 13 to each player. There are P(52; 13, 13, 13, 13) ways to divide the cards into four distinct hands.
- What's the probability that West gets all 13 spades? We could calculate this as P(39; 13, 13, 13)/P(52; 13, 13, 13, 13), but it is easier to see that West gets a uniform random *subset* of 13 cards, so one particular subset occurs with probability 1/C(52, 13).

# With Repetition: All Functions

- If we map distinct objects into the boxes, we are choosing one of the $n^r$ functions from O to B, or equivalently one of the $n^r$ sequences of r elements of B.
- On the other hand, when we map identical objects into the boxes, we are choosing a multiset of size n, whose elements are the boxes, with one member for each element mapped into the box. As we have seen, there are C(n+r-1, r) of these multisets.

# Integer Solutions

- There is an important alternate characterization of the number of multisets of size r with elements taken from B.
- Each such multiset is a solution to the equation $x_1 + \ldots + x_n = r$, where each of the $x_i$'s is a non-negative integer. Here $x_i$ represents the number of objects in the $i^{th}$ box.
- We often need to convert from counting multisets to number of integer solutions to number of distributions of identical elements.

# No Repetition: Injections

- If we are not allowed to map more than one copy of the same object into a box, we are choosing an **injection** (a **one-to-one** function) from the objects to the boxes.
- With distinct objects, there are P(n, r) of these mappings, one for each sequence of objects with no repeats. Of course this number is 0 if r > n, since no such sequence exists in that case.

# No Repetition: Injections

- If we map r identical objects into the n boxes, the only question is which *subset* of r boxes receive objects. There are C(n, r) such subsets, and C(n, r) is also 0 if r > n.
- If we represent these subsets as strings, we need a bit to represent the presence or absence of an object in each box. This gives us a binary string of length n, with exactly r ones and n-r zeros.

# Restricted Repetition

- If we think of elements of O as types of objects rather than just objects, we are allowed one object of each type in the no-repetition case and an unlimited number in the general case.
- If we have a specific number $k_i$ of each distinct object $o_i$, and n is the sum of the $k_i$'s, our function is an arrangement of the multiset with $k_i$ copies of each $o_i$.
- There are $P(n; k_1,\ldots,k_r)$ of these.

# Counting Onto Functions

- We've counted *all* functions from O to B, and all *one-to-one* functions, in the case of both distinct and identical objects.
- What about *onto* functions or **surjections**?
- One case is easy. If I map r identical objects into n boxes by a surjection, I first must have $r \geq n$.
- I can put one object into each box, and then the other r-n into the boxes in C(r-1, n-1) ways. (We're just picking a multiset of size r-n from n possible items.)

# Counting Onto Functions

- The case of mapping distinct objects is more complicated. We're now picking a **partition** of the r objects into n non-empty blocks.
- The **Stirling number of the second kind**, or S(r, n), is the number of partitions when we don't care about the order of the blocks. If we do care about that order, the number of different onto functions is n!S(r, n).

# Counting Onto Functions

- Clearly $S(0, 0) = 1$, $S(r, 0) = 0$ for $n > 0$, $S(r, 1) = 1$, and $S(r, n) = 0$ when $r < n$. Also $S(r, r) = 1$, as there is only one way to put one in each block.
- $S(r, 2)$ is $2^{r-1} - 1$. Consider all the subsets of O, remove ∅ and O itself, and consider putting each other set in the first of two blocks. This counts every two-block partition exactly twice.
- $S(r, r-1)$ is just $C(r, 2)$ because we pick a pair of elements to be put in the same block.

# Stirling Numbers

- We aren't ready to compute general Stirling numbers yet, but we'll see in Chapter 6 that we can describe all of them by a generating function. Here's some values of S(r, n):

| | n=0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| r=0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 3 | 0 | 1 | 3 | 1 | 0 | 0 | 0 |
| 4 | 0 | 1 | 7 | 6 | 1 | 0 | 0 |
| 5 | 0 | 1 | 15 | 25 | 10 | 1 | 0 |
| 6 | 0 | 1 | 31 | 90 | 65 | 15 | 1 |

# The Twelvefold Way

- We can count all functions, just injections, or just surjections. (Counting bijections is either easy ($n!$, if $n = r$) or trivial (0, if $n \neq r$).)
- We can then have our objects be distinct or not, and have our boxes be distinct or not. This gives a total of twelve problems, whose solutions are organized in a table called the **Twelvefold Way**.
- We're only ready here to tackle a few of these problems.

# The Twelvefold Way

| O is: | B is: | Any function | Injection | Surjection |
|---|---|---|---|---|
| dist | dist | $n^r$ | P(n, r) | n!S(r, n) |
| ident | dist | C(n+r-1, r) | C(n, r) | C(r-1, r-n) |
| dist | ident | sum of S's | 0 or 1 | S(r, n) |
| ident | ident | sum of p's | 0 or 1 | $p_n(r)$ |

Here $p_n(r)$ is a **partition number**, the number of ways to divide r identical objects into n identical nonempty groups. We'll see these again in Section 6.3 of Tucker, with a generating function.