CMPT 210: Probability and Computing

Lecture 3

Sharan Vaswani

January 12, 2023

Recap - Counting

Product Rule: For sets A_1, $A_2 \ldots, A_m$, $|A_1 \times A_2 \times \ldots \times A_m| = \prod_{i=1}^{m} |A_i|$ (E.g: Selecting one course each from every subject.)

Sum rule: If $A_1, A_2 \ldots A_m$ are disjoint sets, then, $|A_1 \cup A_2 \cup \ldots \cup A_m| = \sum_{i=1}^{m} |A_i|$ (E.g Number of rainy, snowy or hot days in the year).

Generalized product rule: If S is the set of length k sequences such that the first entry can be selected in n_1 ways, after the first entry is chosen, the second one can be chosen in n_2 ways, and so on, then $|S| = n_1 \times n_2 \times \ldots n_k$. (E.g Number of ways n people can be arranged in a line $= n!$)

Division rule: $f : A \to B$ is a k-to-1 function, then, $|A| = k|B|$. (E.g. For arranging people around a round table, f : seatings $\to$ arrangements is an n-to-1 function).

Number of ways of choosing size k-subsets from a size n-set: $\binom{n}{k}$ (E.g. Number of n-bit sequences with exactly k ones).

Counting subsets - Example

Q: How many m-bit binary sequences contain exactly k ones?

Consider set $A = \{1, 2, \ldots, m\}$ and selecting S, a subset of size k. For example, say $m = 10, k = 3$ and $S = \{10, 3, 7\}$, then S records the positions of the 1's, and can mapped to the sequence 001 0001 001. Similarly, every m-bit sequence with exactly k ones can be mapped to a subset of size k. Hence, there is a bijection:
f : m-bit sequence with exactly k ones $\to$ subsets of size k from size m-set, and
$|m\text{-bit sequence with exactly } k \text{ ones}| = |\text{subsets of size } k| = \binom{m}{k}$.

Q: Suppose we want to buy 10 donuts. There are 5 donut varieties – chocolate, lemon-filled, sugar, glazed, plain. What is the number of ways to select the 10 donuts?

Recall that the number of ways of selecting 10 donuts with 5 varieties = number of 14-bit sequences with exactly 4 ones $= \binom{14}{4} = 1001$.

Q: What is the number of ways of choosing n things with k varieties?

Ans: Equal to the number of $n + k - 1$-bit sequences with exactly $k - 1$ ones $= \binom{n+k-1}{k-1}$.

Counting subsets - Example

Q: What is the number of n-bit binary sequences with at least k ones?

Ans: Set of n-bit binary sequences with at least k ones $=$ n-bit binary sequences with exactly k ones $\cup$ n-bit binary sequences with exactly $k+1$ ones $\cup \ldots \cup$ n-bit binary sequences with exactly n ones. By the sum rule for disjoint sets, number of n-bit binary sequences with at least k ones $= \sum_{i=k}^{n} \binom{n}{i}$.

Q: What is the number of n-bit binary sequences with less than k ones?

Ans: $\sum_{i=0}^{k-1} \binom{n}{i}$

Q: What is the total number of n-bit binary sequences?

Ans: 2^n

Total number of n-bit binary sequences $=$ number of n-bit binary sequences with at least k ones $+$ number of n-bit binary sequences with less than k ones.
Combining the above answers, we can conclude that, $\sum_{k=0}^{n} \binom{n}{k} = 2^n$. Have recovered a special case of the binomial theorem!

Binomial Theorem

For all $n \in \mathbb{N}$ and $a, b \in \mathbb{R}$,

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Example: If $a = b = 1$, then $\sum_{k=0}^{n} \binom{n}{k} = 2^n$ (result from previous slide).

If $n = 2$, then $(a+b)^2 = \binom{2}{0} a^2 + \binom{2}{1} ab + \binom{2}{2} b^2 = a^2 + 2ab + b^2$.

Q: What is the coefficient of the terms with ab^3 and $a^2 b^3$ in $(a+b)^4$? Ans: $\binom{4}{1} = \binom{4}{3}$, 0.

Q: For $a, b > 0$, what is the coefficient of $a^{2n-7} b^7$ and $a^{2n-8} b^8$ in $(a+b)^{2n} + (a-b)^{2n}$?

Ans: $(a+b)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} a^{2n-k} b^k$,
$(a-b)^{2n} = -\sum_{k=0}^{2n} \binom{2n}{k} a^{2n-k} b^k \mathcal{I}\{\text{k is odd}\} + \sum_{k=0}^{2n} \binom{2n}{k} a^{2n-k} b^k \mathcal{I}\{\text{k is even}\}$.
$(a+b)^{2n} + (a-b)^{2n} = 2 \sum_{k=0}^{2n} \binom{2n}{k} a^{2n-k} b^k \mathcal{I}\{\text{k is even}\}$. Hence, coefficient of $a^{2n-7} b^7 = 0$, coefficient of $a^{2n-8} b^8 = 2\binom{2n}{8}$.

Questions?

Generalization to Multinomials

We saw how to split a set into two subsets - one that contains some elements, while the other does not. Can generalize the arguments to split a set into more than two subsets.

A $(k_1, k_2, \ldots, k_m)$-split of set A is a sequence of sets $(A_1, A_2, \ldots A_m)$ s.t. sets A_i form a partition ($A_1 \cup A_2 \cup \ldots = A$ and for $i \neq j$, $A_i \cap A_j = \emptyset$) and $|A_i| = k_i$.

An example of a $(2, 1, 3)$-split of $A = \{1, 2, 3, 4, 5, 6\}$ is $(\{2, 4\}, \{1\}, \{3, 5, 6\})$. Here, $m = 3$, $A_1 = \{2, 4\}$, $A_2 = \{1\}$, $A_3 = \{3, 5, 6\}$ s.t. $|A_1| = 2$, $|A_2| = 1$, $|A_3| = 3$, $A_1 \cup A_2 \cup A_3 = A$ and for $i \neq j$, $A_i \cap A_j = \emptyset$.

Example: Consider strings of length 6 of a's, b's and c's such that number of a's $= 2$; number of b's $= 1$ and number of c's $= 3$. Possible strings: abaccc, ccbaac, bacacc, cbacac.

Each possible string, e.g. bacacc can be written as a $(2, 1, 3)$-split of $A = \{1, 2, 3, 4, 5, 6\}$ as $(\{2, 4\}, \{1\}, \{3, 5, 6\})$ where A_1 records the positions of a, A_2 records the positions of b and A_3 records the positions of c.

Generalization to Multinomials

Q: Show that the number of ways to obtain an $(k_1, k_2, \ldots, k_m)$ split of A with $|A| = n$ is $\binom{n}{k_1, k_2, \ldots k_m} = \frac{n!}{k_1!\, k_2! \ldots k_m!}$ where $\sum_i k_i = n$.

Can map any permutation $(a_1, a_2, \ldots a_n)$ into a split by selecting the first k_1 elements to form set A_1, next k_2 to form set A_2 and so on. For the same split, the order of the elements in each subset does not matter. Hence f : number of permutations $\to$ number of splits is a $k_1!\, k_2!\, \ldots k_m!$-to-1 function.

Hence, $|\text{number of splits}| = \frac{|\text{number of permutations}|}{k_1!\, k_2! \ldots k_m!} = \frac{n!}{k_1!\, k_2! \ldots k_m!}$.

Generalization to Multinomials - Example

Q: Count the number of permutations of the letters in the word BOOKKEEPER.

We want to count sequences of the form $(1E, 1P, 2E, 1B, 1K, 1R, 2O, 1K) = EPEEBKROOK$. There is a bijection between such sequences and $(1, 2, 2, 3, 1, 1)$ split of $A = \{1, 2, \ldots, 10\}$ where A_1 is the set of positions of B's, A_2 is the set of positions of O's, A_3 is set of positions of K and so on.

For example, the above sequence maps to the following split:
({5}, {8,9}, { 6, 10}, { 1,3,4 }, { 2 }, { 7 })

Hence, the total number of sequences that can be formed from the letters in BOOKKEEPER = number of $(1, 2, 2, 3, 1, 1)$ splits of $A = [10] = \{1, 2, \ldots, 10\} = \frac{10!}{1!\,2!\,2!\,3!\,1!\,1!}$.

Q: Count the number of permutations of the letters in the word (i) ABBA (ii) A_1BBA_2 and (iii) $A_1B_1B_2A_2$? Ans: 6, 12, 24

Generalization to Multinomials - Example

Q: Suppose we are planning a 20 km walk, which should include 5 northward km, 5 eastward km, 5 southward km, and 5 westward km. We can move in steps of 1 km in any direction. For example, a valid walk is ($NENWSNSSENSWWESWEENW$) that corresponds to 1 km north followed by 1 km east and so on. How many different walks are possible?

Ans: The set $A = \{1, 2, \ldots, 20\}$ needs to be split into 4 subsets N, S, E, W s.t. $|N| = |S| = |E| = |W| = 5$. Counting the number of walks = counting the number of sequences of the form $(3N, 5W, 4S, 4E, 2N, 1E, 1S)$ = number of ways to obtain an $(5, 5, 5, 5)$-split of set $\{1, 2, 3, \ldots 20\}$. The total number of walks $= \frac{20!}{(5!)^4}$.

Multinomial Theorem

For all $m, n \in \mathbb{N}$ and $z_1, z_2, \dots z_m \in \mathbb{R}$,

$$(z_1 + z_2 + \dots + z_m)^n = \sum_{\substack{k_1, k_2, \dots, k_m \\ k_1 + k_2 + \dots k_m = n}} \binom{n}{k_1, k_2, \dots, k_m} z_1^{k_1} z_2^{k_2} \dots z_m^{k_m}$$

where $\binom{n}{k_1, k_2, \dots, k_m} = \frac{n!}{k_1! k_2! \dots k_m!}$.

Example 1: If $m = 2$, $k_1 = k$, $k_2 = n - k$ and $z_1 = a$, $z_2 = b$, recover the Binomial theorem.

Example 2: If $n = 4$, $m = 3$, then the coefficient of abc^2 in $(a + b + c)^4$ is $\binom{4}{1,1,2} = \frac{4!}{1!1!2!}$.

Questions?

Inclusion-Exclusion Principle

Recall that if A, B, C are disjoint subsets, then, $|A \cup B \cup C| = |A| + |B| + |C|$ (this is the Sum rule from Lecture 1).

For two general sets A, B, $|A \cup B| = |A| + |B| - |A \cap B|$. The last term fixes the "double counting".

Similarly, $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |A \cap C| + |A \cap B \cap C|$.

In general,

$$\begin{aligned} |\cup_{i=1,2,\ldots n} A_i| = \sum_i |A_i| - \sum_{i,j \text{ s.t. } 1 \le i < j \le n} |A_i \cap A_j| + \sum_{i,j,k \text{ s.t. } 1 \le i < j < k \le n} |A_i \cap A_j \cap A_k| \\ + \ldots + (-1)^{n-1} |\cap_{i=1,2,\ldots n} A_i| \end{aligned}$$

Inclusion-Exclusion Principle - Example

Q: Suppose there are 60 math majors, 200 EECS majors, and 40 physics majors. A student is allowed to double or even triple major. There are 4 math-EECS double majors, 3 math-physics double majors, 11 EECS-physics double majors and 2-triple majors. What is the total number of students across these three departments?

If M, E, P are the sets of Math, EECS and physics majors, then we wish to compute $|M \cup E \cup P| = |M| + |E| + |P| - |M \cap E| - |M \cap P| - |E \cap P| + |M \cap E \cap P| = 300$ - $|M \cap E| - |M \cap P| - |E \cap P| + |M \cap E \cap P|$.

$|M \cap E| = 4 + 2 = 6$, $|M \cap P| = 3 + 2 = 5$, $|P \cap E| = 11 + 2 = 13$. $|M \cap E \cap P| = 2$

$|M \cup E \cup P| = 300 - 6 - 5 - 13 + 2 = 278$.

Inclusion-Exclusion Principle - Example

Q: In how many permutations of the set $\{0, 1, 2, \ldots, 9\}$ do either 4 and 2, 0 and 4, or 6 and 0 appear consecutively? For example, in the following permutation $\underline{42}067891235$, 4 and 2 appear consecutively, but 6 and 0 do not (the order matters).

Let P_{42} be the set of sequences such that 4 and 2 appear consecutively. Similarly, we define P_{60} and P_{04}. So we want to compute
$|P_{42} \cup P_{60} \cup P_{04}| = |P_{42}| + |P_{60}| + |P_{04}| - |P_{42} \cap P_{60}| - |P_{42} \cap P_{04}| - |P_{60} \cap P_{04}| + |P_{42} \cap P_{60} \cap P_{04}|$.

Let us first compute $|P_{42}| = 9!$. Similarly, $|P_{60}| = |P_{04}| = 9!$.

What about intersections? $|P_{42} \cap P_{60}| =$ Number of sequences of the form $(42, 60, 1, 3, 5, 7, 8, 9) = 8!$. Similarly, $|P_{60} \cap P_{04}| = |P_{42} \cap P_{04}| = 8!$.

$|P_{42} \cap P_{60} \cap P_{04}| =$ Number of sequences of the form $(6042, 1, 3, 5, 7, 8, 9) = 7!$.

By the inclusion-exclusion principle, $|P_{42} \cup P_{60} \cup P_{04}| = 3 \times 9! - 3 \times 8! + 7!$.

Questions?