CMPT 210: Probability and Computing

Lecture 2

Sharan Vaswani

January 11, 2024

Functions

We can also define a function with a set as the argument. For a set $S \in D$, $f(S) := \{x | \forall s \in S, x = f(s)\}$.

$A = \{a, b, c, \dots z\}$, $B = \{1, 2, 3, \dots 26\}$. $f : A \to B$ such that $f(a) = 1$, $f(b) = 2, \dots$. $f(\{e, f, z\}) = \{5, 6, 26\}$.

If D is the domain of f, then range$(f) := f(D) = f(\text{domain}(f))$.

Q: If $f : \mathbb{N} \to \mathbb{R}$, and $f(x) = x^2$. What is the domain and codomain of f? What is the range?

Q: Consider $f : \{0, 1\}^5 \to \mathbb{N}$ s.t. $f(x)$ counts the length of a left to right search of the bits in the binary string x until a 1 appears. $f(01000) = 2$.

What is $f(00001)$, $f(00000)$? Is f a total function?

Surjective functions: $f : A \to B$ is a surjective function iff for every $b \in B$, there exists an $a \in A$ s.t. $f(a) = b$. $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x + 1$ is a surjective function.

For surjective functions, $|\#\text{arrows}| \geq |B|$.

Since each element of A is assigned at most one value, and some need not be assigned a value at all, $|\#\text{arrows}| \leq |A|$.

Hence, if f is a surjective function, then $|A| \geq |B|$.

$A = \{a, b, c, \ldots z, \alpha, \beta, \gamma, \ldots\}$, $B = \{1, 2, 3, \ldots 26\}$. $f : A \to B$ such that $f(a) = 1$, $f(b) = 2, \ldots$. f does not assign any value to the Greek letters. For every number in B, there is a letter in A. Hence, f is surjective, and $|A| > |B|$.

Injective & Bijective Functions

Injective functions: $f : A \to B$ is an injective function iff $\forall a \in A$, there is a *unique* $b \in B$ s.t. $f(a) = b$. If f is injective and $f(a) = f(b)$, then it implies that $a = b$.

Hence, $|\#\text{arrows}| = |A| \leq |B|$. Hence, if f is a injective function, then $|A| \leq |B|$.

$A = \{a, b, c, \ldots z\}$, $B = \{1, 2, 3, \ldots 26, 27, \ldots 100\}$. $f : A \to B$ such that $f(a) = 1$, $f(b) = 2, \ldots$. No element in A is assigned values $27, 28, \ldots$, and for every letter in A, there is a unique number in B. Hence, f is injective, and $|A| < |B|$.

Bijective functions: f is a bijective function iff it is both surjective and injective, implying that $|A| = |B|$.

$A = \{a, b, c, \ldots z\}$, $B = \{1, 2, 3, \ldots 26\}$. $f : A \to B$ such that $f(a) = 1$, $f(b) = 2, \ldots$. Every element in A is assigned a unique value in B and for every element in B, there is a value in A that is mapped to it. f is bijective, and $|A| = |B|$.

Functions

Converse of the previous statements is also true.

- If $|A| \geq |B|$, then it's always possible to define a surjective function $f : A \to B$.
- If $|A| \leq |B|$, then it's always possible to define a injective function $f : A \to B$.
- If $|A| = |B|$, then it's always possible to define a bijective function $f : A \to B$.

Q: Recall that the Cartesian product of two sets $S = \{s_1, s_2, \dots, s_m\}$, $T = \{t_1, t_2, \dots, t_n\}$ is $S \times T := \{(s, t) | s \in S, t \in T\}$. Construct a bijective function $f : (S \times T) \to \{1, \dots nm\}$, and prove that $|S \times T| = nm$.

Sequences

Examples: (a, b, a), (1,3,4), (4,3,1)

An element can appear twice. E.g. $(a, a, b) \neq (a, b)$.

The order of the elements does matter. E.g. $(a, b) \neq (b, a)$.

Q: What is the size of $(1, 2, 2, 3)$? What is the size of $\{1, 2, 2, 3\}$? .

Sets and Sequences: The Cartesian product of sets $S \times T \times U$ is a set consisting of all sequences where the first component is drawn from S, the second component is drawn from T and the third from U. $S \times T \times U = \{(s, t, u) | s \in S, t \in T, u \in U\}$.

Q: For set $S = \{0, 1\}$, $S^3 = S \times S \times S$. Enumerate S^3. What is $|S^3|$?

Questions?

Counting Sets – using a bijection

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?

Let A be the set of ways to select the 10 donuts. Each element of A is a potential selection. For example, 4 chocolate, 3 lemon, 0 sugar, 2 glazed and 1 plain.

Let's map each way to a string as follows: $\underbrace{0000}_{\text{chocolate}}\ \underbrace{000}_{\text{lemon}}\ \underbrace{\quad}_{\text{sugar}}\ \underbrace{00}_{\text{glazed}}\ \underbrace{0}_{\text{plain}}$.

Lets fix the ordering – chocolate, lemon, sugar, glazed and plain, and abstract this out further to get the sequence: 0000 1 000 1 1 00 1 0. Hence, each way of choosing donuts is mapped to a binary sequence of length 14 with exactly 4 ones. Now, let B be all 14-bit sequences with exactly 4 ones. An element of B is 11110000000000.

Q: The above sequence corresponds to what donut order?

For every way to select donuts, we have an equivalent sequence in B. And every sequence in B implies a unique way to select donuts. Hence, the mapping from $A \to B$ is a bijective function.

Counting Sets – using a bijection

Hence, $|A| = |B|$, meaning that we have reduced the problem of counting the number of ways to select donuts to counting the number of 14-bit sequences with exactly 4 ones.

General result: The number of ways to choose n elements with k available varieties is equal to the number of $n + k - 1$-bit binary sequences with exactly $k - 1$ ones.

Q: There are 2 donut varieties – chocolate and lemon-filled. Suppose we want to buy only 2 donuts. Use the above result to count the number of ways in which we can select the donuts? What are these ways?

Q: In the above example, I want at least one chocolate donut. What are the types of acceptable 3-bit sequences with this criterion? How many ways can we do this?

Counting Sets – using the sum rule

Q: Let R be the set of rainy days, S be the set of snowy days and H be the set of really hot days in 2023. A bad day can be either rainy, snowy or really hot. What is the number of good days?

Let B be the set of bad days. $B = R \cup S \cup H$, and we want to estimate $|\bar{B}|$. $|D| = 365$.
$|\bar{B}| = |D| - |B| = 365 - |B| = 365 - |R \cup S \cup H|$.

Since the sets R, S and H are disjoint, $|R \cup S \cup H| = |R| + |S| + |H|$, and hence the number of good days $= 365 - |R| - |S| - |H|$.

Sum rule: If $A_1, A_2 \dots A_m$ are disjoint sets, then, $|A_1 \cup A_2 \cup \dots \cup A_m| = \sum_{i=1}^{m} |A_i|$.

Counting Sequences – using the product rule

Q: Suppose the university offers Math courses (denoted by the set M), CS courses (set C) and Statistics courses (set S). We need to pick one course from each subject – Math, CS and Statistics. What is the number of ways we can select we can select the 3 courses?

The above problem is equivalent to counting the number of sequences of the form (m, c, s) that maps to choose the Math course m, CS course c and Stats course s.

Recall that the product of sets $M \times C \times S$ is a set consisting of all sequences where the first component is drawn from M, the second component is drawn from C and the third from S, i.e. $M \times C \times S = \{(m, c, s) | m \in M, c \in C, s \in S\}$. Hence, counting the number of sequences is equivalent to computing $|M \times C \times S|$.

Product Rule: $|M \times C \times S| = |M| \times |C| \times |S|$.

Using the above equivalence, the number of sequences and hence, the number of ways to select the 3 courses is $|M| \times |C| \times |S|$.

Counting – Example

Q: What is the number of length n-passwords that can be generated if each character in the password is allowed to be lower-case letter?

Counting – Example

Q: What is the number of passwords that can be generated if the (i) first character is only allowed to be a lower-case letter, (ii) each subsequent character in the password is allowed to be lower-case letter or digit $(0-9)$ and (iii) the length of the password is required to be between 6-8 characters?

Let $L = \{a, b, \ldots z\}$ and $D = \{0, 1, 2, \ldots\}$. Using the equivalence between sequences and products of sets, the set of passwords of length 6 is given by $P_6 = L \times (L \cup D)^5$. Using the product rule, $|P_6| = |L| \times (|L \cup D|)^5 = |L| \times (|L| + |D|)^5$.

Since the total set of passwords are $P = P_6 \cup P_7 \cup P_8$, and a password can be either of length 6, 7 or 8, sets P_6, P_7 and P_8 are disjoint. Using the sum rule, $|P| = |P_6| + |P_7| + |P_8| = |L| \times \left[(|L| + |D|)^5(1 + (|L| + |D|) + (|L| + |D|)^2)\right] = 26 \times 36^5 \times [1 + 36 + 1296]$.

Counting sequences – using the generalized product rule

Q: Suppose we have p prizes to be handed amongst the set A of n students. What are the number of ways in which we can distribute the prizes?

Q: Suppose we have p prizes to be handed amongst the set A of n students. What are the number of ways in which we can distribute the prizes such that each prize goes to a different student i.e. no student receives more than one prize?

Consider sequences of length p. The first entry can be chosen in n ways (the first prize can be given to one of the n students). After the first entry is chosen, since the same student cannot receive the prize, the second entry can be chosen in $n-1$ ways, and so on. Hence, the total number of ways to distribute the prizes $= n \times (n-1) \times \ldots \times (n-(p-1))$.

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 $|S| = n_1 \times n_2 \times \ldots n_k$. If $n_1 = n_2 = \ldots = n_k$, we recover the product rule.

Counting - Example

Q: A dollar bill is defective if some digit appears more than once in the 8-digit serial number. What is the fraction of non-defective bills?

In order to compute the fraction of non-defective bills, we need to compute the quantity $\frac{|\text{serial numbers with all different digits}|}{|\text{possible serial numbers}|}$.

For computing $|\text{possible serial numbers}|$, each digit can be one of 10 numbers. Hence, using the product rule, $|\text{possible serial numbers}| = 10 \times 10 \ldots = 10^8$.

For computing $|\text{serial numbers with all different digits}|$, the first digit can be one of 10 numbers. Once the first digit is chosen, the second one can be chosen in 9 ways, and so on. By the generalized product rule, $|\text{serial numbers with all different digits}| = 10 \times 9 \times \ldots 3 = 1,814,400$.

Fraction of non-defective bills $= \frac{1{,}814{,}400}{10^8} = 1.8144\%$.

Permutations

A permutation of a set S is a sequence of length $|S|$ that contains every element of S exactly once. Permutations of $\{a, b, c\}$ are $(a, b, c), (a, c, b), (b, c, a), (b, a, c), (c, a, b), (c, b, a)$.

Q: Given a set of size n, what is the total number of permutations?

Considering sequences of length n, the first entry can be chosen in n ways. Since each element can be chosen only once, the second entry can be chosen in $n - 1$ ways, and so on.

By the generalized product rule, the number of permutations $= n \times (n-1) \times \ldots \times 1$.

Factorial: $n! := n \times (n-1) \times \ldots \times 1$. By convention: $0! = 1$.

How big is $n!$? **Stirling approximation**: $n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$.

Q: Which is bigger? $n!$ vs $n(n-1)(n+2)(n-3)!$?

Q: In how many ways can we arrange n people in a line?

Counting – Division rule

k-**to**-1 **function**: Maps exactly k elements of the domain to every element of the codomain.

If $f : A \to B$ is a k-to-1 function, then, $|A| = k|B|$.

Example: E is the set of ears in this room, and P is the set of people. Then f mapping the ears to people is a 2-to-1 function. Hence, $|E| = 2|P|$.

Q: If $f : A \to B$ is a k-to-1 function, and $g : B \to C$ is a m-to-1 function, then what is $|A|/|C|$?

Q: If $f : A \to B$ is a k-to-1 function, and $g : C \to B$ is a m-to-1 function, then what is $|A|/|C|$?

Counting – Example

Q: In how many ways can we arrange n people around a round table? Two seatings are considered to be the same *arrangement* if each person sits with the same person on their left in both seatings.

Starting from the head of the table, and going clockwise, each seating has an equivalent sequence. $|\text{seatings}| = \text{number of permutations} = n!$.

n different seatings can result in the same arrangement (by clockwise rotation).

Hence, $f : \text{seatings} \to \text{arrangements}$ is an n-to-1 function. Hence, the $|\text{seatings}| = n\,|\text{arrangements}|$, meaning that the $|\text{arrangements}| = (n-1)!$.

Questions?

Counting subsets

Q: How many size-k subsets of a size-n set are there?
Example: How many ways can we select 5 books from 100?

Let us form a permutation of the n elements, and pick the first k elements to form the subset. Every size k subset can be generated this way. There are $n!$ total such permutations.

The order of the first k elements in the permutation does not matter in forming the subset, and neither does the order of the remaining $n - k$ elements.

The first k elements can be ordered in $k!$ ways and the remaining $n - k$ elements can be ordered in $(n - k)!$ ways. Using the product rule, $k! \times (n - k)!$ permutations map to the same size k subset.

Hence, the function f : permutations $\rightarrow$ size k subsets is a $k! \times (n - k)!$-to-1 function. By the division rule, $|\text{permutations}| = k! \times (n - k)! \, |\text{size } k \text{ subsets}|$. Hence, the total number of size k subsets $= \frac{n!}{k! \times (n-k)!}$.

n **choose** $k = \binom{n}{k} := \frac{n!}{k! \times (n-k)!}$.

Counting subsets

Q: Prove that $\binom{n}{k} = \binom{n}{n-k}$ - both algebraically (using the formula for $\binom{n}{k}$) and combinatorially (without using the formula)

Q: Which is bigger? $\binom{8}{4}$ vs $\binom{8}{5}$?

Counting subsets – Example

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

Consider set $A = \{1, \ldots, m\}$ and selecting S, a subset of size k. For example, say $m = 10, k = 3$ and $S = \{1, 7, 10\}$. 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 S of size k. Hence, there is a bijection:
f : m-bit sequence with exactly k ones $\rightarrow$ 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?

Counting subsets – Example

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

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

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

Questions?