Lecture Note: Week 1

MATH 461: Probability Theory, Spring 2021
Daesung Kim

Lecture 1. Basic Combinatorics I (Sec 1.1-3)

We start with a simple question. How many two-letter words are there (using 26 alphabet)?

> **The Basic Principle of Counting**
> Suppose that two experiments are to be performed. Then if experiment 1 can result in any one of m possible outcomes and if, for each outcome of experiment 1, there are n possible outcomes of experiment 2, then together there are mn possible outcomes of the two experiments.

There are two spots for letters and there are 26 possibilities for each spot. Thus, there are $26 \times 26 = 676$ possible two-letter words. How many two-letter words without repetition? There are 26 repeated cases so that the answer is $676 - 26 = 650$.

> **The Generalized Basic Principle of Counting**
> If r experiments that are to be performed are such that the first one may result in any of n_1 possible outcomes; and if, for each of these n_1 possible outcomes, there are n_2 possible outcomes of the second experiment; and if, for each of the possible outcomes of the first two experiments, there are n_3 possible outcomes of the third experiment; and if ..., then there is a total of $n_1 n_2 \cdots n_r$ possible outcomes of the r experiments.

Example 1. How many different 7-place license plates are possible if the first 3 places are to be occupied by letters and the final 4 by numbers? How many license plates would be possible if repetition among letters or numbers were prohibited?

Example 2. Consider a set S of n elements, say $S = \{1, 2, \cdots, n\}$. How many different subsets of S are there? If $n = 2$, then $S = \{1, 2\}$ and all possible subsets are $\varnothing, \{1\}, \{2\}, \{1, 2\}$. In general, for each element in the set S, we have two choices: included or not. Since there are n elements and 2 choices each, the number of possible subsets is $2 \times \cdots \times 2 (n \text{ times}) = 2^n$.

> **Permutation**
> Each ordered arrangement of n distinct objects is called a permutation. The number of all possible permuations is $n! = n \cdot (n-1) \cdots 2 \cdot 1$.

For example, all possible permutation of three letters *a, b, c* are *abc, acb, bac, bca, cab, cba.*

Example 3. Ms. Jones has 10 books that she is going to put on her bookshelf. Of these, 4 are mathematics books, 3 are chemistry books, 2 are history books, and 1 is a language book. Ms. Jones wants to arrange her books so that all the books dealing with the same subject are together on the shelf. How many different arrangements are possible?

Example 4. How many different letter arrangements can be formed from the letters *arrange*?

Lecture 2. Basic Combinatorics II (Sec 1.4-6)

Combination
Consider n distinct objects. How many different groups, called combinations, of size r $(1 \le r \le n)$ of these objects can be formed? The number of combinations is

$$\binom{n}{r} = \frac{n \cdot (n-1) \cdots (n-r+1)}{r \cdot (r-1) \cdots 1} = \frac{n!}{(n-r)!r!}.$$

For example, for $n = 3$ objects, say $a,\ b,\ c$ and $r = 2$, there are three possible combinations: $ab,\ bc,\ ac$.

Example 5. A committee of 3 is to be formed from a group of 20 people. How many different committees are possible?

Example 6. Consider a set of 8 antennas of which 5 are defective and 3 are functional and assume that all of the defectives and all of the functionals are considered indistinguishable. How many linear orderings are there in which no two defectives are consecutive?

The Binomial Theorem
For real numbers x, y and a natural number n, we have

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Multinomial Coefficients
For integer $n \ge 1$ and integers $n_1, \cdots, n_r \ge 0$ such that $n_1 + \cdots + n_r = n$,

$$\binom{n}{n_1, \cdots, n_r} = \frac{n!}{n_1! \cdots n_r!}.$$

How to interpret the multinomial coefficients?

The first interpretation of multinomial coefficients
The number of ways to divide n distinct objects into r *distinct groups* of size $n_1, \cdots, n_r$ with $n_1 + \cdots + n_r = n$ is

$$\binom{n}{n_1, \cdots, n_r}.$$

Example 7. Ten children are to be divided into an A team and a B team of 5 each. The A team will play in one league and the B team in another. How many different divisions are possible? What if we just divide them into two teams?

The second interpretation of multinomial coefficients
The number of ordered arrangements of n objects of which n_1 are alike, $\cdots$, n_r are alike with $n_1, \cdots, n_r$ with $n_1 + \cdots + n_r = n$ is

$$\binom{n}{n_1, \cdots, n_r}.$$

Example 8. How many different letter arrangements can be formed from the letters *PEPPER*?

The Multinomial Theorem
For real numbers $x_1, \cdots, x_r$ and an integer $n \geq 1$, we have

$$(x_1 + \cdots + x_r)^n = \sum \binom{n}{n_1, \cdots, n_r} x_1^{n_1} \cdots x_r^{n_r}$$

where the sum is over all integers $n_1, \cdots, n_r \geq 0$ such that $n_1 + \cdots + n_r = n$.

For example,

$$(x + y + z)^2 = x^2 + y^2 + z^2 + 2xy + 2yz + 2xz.$$

Example 9. What is the coefficient of the term x^2y^3z in the expansion of $(x + y + z)^6$?

Example 10. Consider n distinct balls and r distinct urns. How many different ways are there to distribute balls into urns? (An urn can contain any number of balls, including zero.) What if the balls are indistinguishable?

Lecture 3. Axioms of Probability I (Sec 2.2-3)

Sample Spaces and Events

We consider an experiment whose outcome is unpredictable. A sample space is the set of all possible outomes, which is usually denoted by S. For example,

(i) If the outcome of an experiment consists of the determination of the sex of a newborn child, then $S = \{g, b\}$.

(ii) If the experiment consists of flipping two coins, then $S = \{(H, H), (H, T), (T, H), (T, T)\}$.

(iii) If the experiment consists of tossing two dice, then $S = \{(i, j) : 1 \leq i, j \leq 6\}$.

Let S be a sample space. An event is a subset of the sample space S. For example, $\varnothing$ and S itself are also events.

(i) If $S = \{g, b\}$, then the events are $\varnothing$, $\{g\}$, $\{b\}$, and $\{g, b\}$.

(ii) If $S = \{(H, H), (H, T), (T, H), (T, T)\}$, then $\{(H, H), (H, T)\}$ (the first coin is head) is an event.

(iii) If $S = \{(i, j) : 1 \leq i, j \leq 6\}$, then $\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$ is an event.

Operations on Events
Let S be a sample space and E, F events.

(i) The union of E and F is an event, denoted by $E \cup F$, that either E or F occurs.

(ii) The intersection of E and F is an event, denoted by EF or $E \cap F$, that both E and F occur.

(iii) The complement of E is an event, denoted by E^c, that E does not occur.

(iv) If all outcomes in E are also in F, we say that E is contained in F, denoted by $E \subset F$.

(v) Let $E_1, E_2, \cdots$ be countably many events. The union of these events is an event, denoted by $\cup_{i=1}^{\infty} E_i$, that at least one of these event occurs. The intersection of these events is an event, denoted by $\cap_{i=1}^{\infty} E_i$, that all the events occur.

Example 11. Let $S = \{(H, H), (H, T), (T, H), (T, T)\}$. Let $E = \{\text{the first coin is head.}\} = \{(H, H), (H, T)\}$ and $F = \{\text{the first is different from the second.}\} = \{(H, T), (T, H)\}$. Then, $E \cup F = \{(H, H), (H, T), (T, H)\}$, $E \cap F = \{(H, T)\}$, and $E^c = \{(T, T), (T, H)\} = \{\text{the first is tail}\}$.

Example 12. We have $S^c = \varnothing$, $\varnothing^c = S$, and $\varnothing \subset E \cap F \subset E, F \subset E \cup F \subset S$.

Laws for operations
Let S be a sample space and E, F, G events.

(i) (Commutativity) $E \cup F = F \cup E$ and $EF = FE$.

(ii) (Associativity) $(E \cup F) \cup G = E \cup (F \cup G)$ and $(E \cap F) \cap G = E \cap (F \cap G)$.

(iii) (Distribution laws) $(E \cup F) \cap G = (E \cap G) \cup (F \cup G)$ and $(E \cap F) \cup G = (E \cup G) \cap (F \cup G)$.

(iv) (De Morgan's laws) For events $E_1, \cdots, E_n$,

$$\left(\bigcup_{k=1}^{n} E_k\right)^c = \bigcap_{k=1}^{n} E_k^c, \qquad \left(\bigcap_{k=1}^{n} E_k\right)^c = \bigcup_{k=1}^{n} E_k^c.$$

Example 13. Let E, F, G be events in a sample space S. Find expressions of the following events.

(i) Both E and G, but not F, occur: $E \cap G \cap F^c$.

(ii) At least two of the events occur: $(E \cap F) \cup (F \cap G) \cup (G \cap E)$.

(iii) At most one of them occur: the complement of the event in (ii). That is,

$$\begin{aligned}((E \cap F) \cup (F \cap G) \cup (G \cap E))^c &= (E \cap F)^c \cap (F \cap G)^c \cap (G \cap E)^c \\ &= (E^c \cup F^c) \cap (F^c \cup G^c) \cap (G^c \cup E^c).\end{aligned}$$

Axioms of Probability

An intuitive way of defining the probability of an event E is to consider its relative frequency. Suppose we perform the same experiments n times and count the number of occurance of the event E, denotes by $\phi_n(E)$. Then, we consider

$$\lim_{n\to\infty} \frac{\phi_n(E)}{n}$$

as a definition of probability of E. However, this definition leads to several delicate questions: how do we know that the limit exists? Even if so, how do we know the limit is consistent? And so on.

Modern probability theory starts from the axioms of probability and show that the relative frequency converges to its probability.

Axioms of Probability
Let S be a sample space. For each event E, the probability $\mathbb{P}(E)$ is an assignment so that the following axioms are satisfied:

(i) For all events, $0 \le \mathbb{P}(E) \le 1$.

(ii) $\mathbb{P}(S) = 1$.

(iii) For any sequence of mutually exclusive events $E_1, E_2, \cdots$, (meaning tthat $E_iE_j = \varnothing$ for $i \neq j$),

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i).$$

Direct consequences are the following:

(i) $\mathbb{P}(\varnothing) = 0$.

(ii) If $E_1, E_2, \cdots, E_n$ are mutually exclusive events, then

$$\mathbb{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mathbb{P}(E_i).$$

In particular if $A \cap B = \varnothing$, then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Example 14. Let $S = \{H, T\}$, then all the events are $\varnothing, \{H\}, \{T\}, S$. We assign the probabilities $\mathbb{P}(\varnothing) = 0$, $\mathbb{P}(\{H\}) = p$, $\mathbb{P}(\{T\}) = 1 - p$, and $\mathbb{P}(S) = 1$ for some $0 \leq p \leq 1$. One can check that the three axioms are satisfied.

The pair of a sample space and probability $(S, \mathbb{P})$ satisfying the three axioms is called a probabilty space.

References

[SR] Sheldon Ross, *A First Course in Probability*, 9th Edition, Pearson

Department of Mathematics, University of Illinois at Urbana-Champaign
E-mail address: daesungk@illinois.edu