STAT394 Notes

Eric Xia

Last Updated 29 March 2021

Contents

1 Introduction

1.1 Permutations and Combinations

Cardinality: the number of elements in a given set A, denoted by A

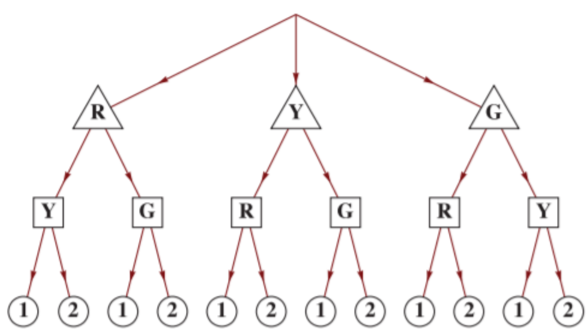
Property

Let A and B be finite sets and k a positive integer. Assume that there is a function f from A onto B so that each element of B is the image of exactly k elements of A. (Such a function is called k-to-one.) Then $A = k \cdot B$.

Below is a graphical illustration of this property. Set A has 12 elements represented by bullets, set B has 4 elements represented by squares, and set C has 4 elements represented by triangles. Each element of B is matched up with exactly three elements of A and each element of C is matched up with one element of B. Thus $A = 3 \cdot B$. The elements of B and C are matched up one-to-one, so $B = C$.

Example: To dress up for school in the morning, Joyce chooses from 3 dresses (red, yellow, or green), 3 blouses (also red, yellow, or green), and 2 pairs of shoes. She refuses to wear a dress and a blouse of matching colors. How many different outfits can she choose from?

We can imagine that Joyce goes through the choices one by one: first the dress, then the blouse, and finally the shoes. There are three choices for the dress. Once the dress is chosen, there are two choices for the blouse, and finally two choices for the shoes. By the multiplication principle, she has $3 \cdot 2 \cdot 2 = 12$ different outfits. The construction of the n-tuples in the general multiplication rule may be visualized by a **decision tree**, with arrows pointing to the available choices of the subsequent step. The dresses are the triangles and the blouses the squares, labeled by the colors. The shoes are the circles. The possible outfits resulting from the three choices can thus be read off from the directed paths through the tree along the arrows.



Corollary of General Multiplication Product

Let $A_1, A_2, \ldots, A_n$ be finite sets. Then

$$(A_1 \times A_2 \times \cdots \times A_n) = (A_1) \cdot (A_2) \ldots (A_n) = \Pi_{i=1}^{n}(A_i)$$

where the capital pi notation is shorthand for a product

Example 2: In a certain country license plates have three letters followed by three digits. How many different license plates can we construct if the country's alphabet contains 26 letters?

Each license plate can be described as an element of hte set $A \times A \times A \times B \times B \times B$ where A is the set of letters and B is the set of digits. Since $A = 26$ and $B = 10$, the answer is $26^3 \cdot 10^3 = 17,576,000$.

Example 3: We flip a coin three times and then roll a die twice. We record the resulting sequence of outcomes in order. How many different sequences are there?

The outcome of a coin flip is an element of the set $C = \{H, T\}$ and a die roll gives an element of the set $D = \{1, 2, 3, 4, 5, 6\}$. Each possible sequence is an element of the set $C \times C \times C \times D \times D$, which has $2^3 \cdot 6^2 = 288$ elements.

Example 4: How many distinct subsets does a set of size n have?

Each subset can be encoded by an n-tuple with entries 0 or 1, where the ith entry is 1 if the ith element of the set is in the subset (and 0 if it is not). Thus the number of subsets is the same as the number of elements in the Cartesian product

$$\{0, 1\} \times \cdot \times \{0, 1\} = \{0, 1\}^n,$$

which is 2^n. Note that the empty set $\emptyset$ and set itself are included in the count. These correspond to the n-tuples $(0, 0, \ldots, 0)$ and $(1, 1, \ldots, 1)$, respectively.

Permutations

Consider all k-tuples $(a_1, \ldots, a_k)$ that can be constructed from a set A of size $n (n \geq k)$ without repetition. So each $a_i \in A$ and $a_i \neq a_j$ if $i \neq j$. The total number of these k-tuples is

$$(n)_k = n \cdot (n-1) \ldots (n-k+1) = \frac{n!}{(n-k)!}$$

In particular, with $k = n$, each n-tuple is a **permutation** of the set A. So the total number of orderings of a set of n elements is $n! = n \cdot (n-1) \ldots 2 \cdot 1$.

Proof. Construct these k-tuples sequentially. Choose the first entry out of the full set A with n alternatives. If we have the first j entries chosen then the next one can be any one of the remaining $n - j$ elements of A. Thus the total number of k-tuples is the product $n \cdot (n-1) \ldots (n-k+1)$. □

$(n)_k$ is the **descending factorial**.

Example 5: I have a 5-day vacation from Monday to Friday in Santa Barbara. I have to specifically assign one day to exploring the town, one day to hiking the mountains, and one day for the beach. I can take up at most one activity per day. In how many ways can I schedule these activities?

The underlying set has 5 elements and we choose a 3-tuple. For example, (Mon, Fri, Tue) means Monday in town, Friday for hiking, and Tuesday on the beach. The number of choices is $5 \cdot 4 \cdot 3 = 60$.

Binomial Coefficient

Let n and k be nonnegative integers with $0 \leq k \leq n$. The number of distinct subsets of size k that a set of size n has is given by the **binomial coefficient**

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Conventionally, $0! = 1$ gives $\binom{n}{0} = 1$ which makes sense because the empty set $\emptyset$ is the unique subset with zero elements. The binomial coefficient $\binom{n}{k}$ is defined as 0 for integers $k < 0$ and $k > n$.

Example 6: In a class there are 12 boys and 14 girls. How many different teams of 7 pupils with 3 boys and 4 girls can be created?

Imagine that first we choose the 3 boys and then the 4 girls. We can choose the 3 boys in $\binom{12}{3}$ different ways and the 4 girls in $\binom{14}{4}$ different ways. Thus we can form the team in $\binom{12}{3} \cdot \binom{14}{4}$ different ways.

Example 7: Let $A = \{1, 2, 3, 4, 5, 6\}$.
(a) How many different two-element subsets of A are there that have one element from $\{1, 2, 3, 4\}$ and one element from $\{5, 6\}$?

The exhaustive list shows that the answer is 8:

$$\{1,5\}, \{1,6\}, \{2,5\}, \{2,6\}, \{3,5\}, \{3,6\}, \{4,5\}, \{4,6\}$$

We get the correct answer by first picking one element from $\{1, 2, 3, 4\}$ (four choices) and then one element from $\{5, 6\}$ (two choices) and multiplying: $4 \cdot 2 = 8$. Another way to arrive at the same answer is to take

$$\binom{6}{2} - \binom{4}{2} - \binom{2}{2} = 15 - 6 - 1 = 8$$

Multinomial Coefficient

Let n and r be positive integers and $k_1, \ldots, k_r$ nonnegative integers such that $k_1 + \cdots + k_r = n$. The number of ways of assigning labels $1, 2, \ldots, r$ to n items so that, for each $i = 1, 2, \ldots, r$ exactly k_i items receive label i, is the **multinomial coefficient**

$$\binom{n}{k_1, k_2, \ldots, k_r} = \frac{n!}{k_1! k_2! \ldots k_r!}$$

Example 8: 120 students signed up for a class. The class is divided into 4 sections numbered 1, 2, 3, and 4, which will have 25, 30, 31, 34 students, respectively. How many ways are there to divide up the students among the four sections?

The answer is

$$\binom{120}{25, 30, 31, 34} = \frac{120!}{25! \cdot 30! \cdot 31! \cdot 34!}$$

Example 9: If you flip a fair coin 4 times what is the probability that you will get exactly 2 tails?

We are trying to find the combination of 2 heads from 4 coins:

$$C(4,2) = \frac{4!}{2!(4-2)!} = 6$$

The total number of all combinations of 4 flips is:

$$C = 2^4 = 16$$

Thus the probability is

$$P(\text{Exactly 2H}) = \frac{6}{16}$$

Example 10: There are 9 students in a class: 5 boys and 4 girls. If the teacher picks a group of 4 at random,w hat is the probability that everyone in the group is a boy?

The number of ways to pick a group of 4 students out of 9 is

$$\frac{9!}{4!(9-4)!}$$

and the number of ways to pick a group of 4 students out of 5 is

$$\frac{5!}{4!(5-4)!}$$

Thus the probability is

$$\frac{\frac{5!}{4!(5-4)!}}{\frac{9!}{4!(9-4)!}} = \frac{5}{126}$$

1.2 Binomial Coefficients

A finite sum may be represented by

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n$$

A product may be represented similarly like so:

$$\Pi_{k=1}^{n} a_k = a_1 \cdot a_2 \cdot a_3 \ldots a_n$$

Exponentials and logarithms can be used to convert between sums and products:

$$\Pi_{k=1}^{n} a_k = \exp\left(\ln \Pi_{k=1}^{n} a_k\right) = e^{\sum_{k=1}^{n} \ln a_k}, a_k > 0$$

By convention, an empty sum has value zero while an empty product has value 1. This is because 0 is the additive identity and 1 is the multiplicative property. The $0! = 1$ is an instance of this convention.

Summation Identities

Let n be a positive integer. Then

$$\begin{aligned} 1 + 2 + \cdots + n &= \frac{n(n+1)}{2} \\ 1^2 + 2^2 + \cdots + n^2 &= \frac{n(n+1)(2n+1)}{6} \\ 1^3 + 2^3 + \cdots + n^3 &= \frac{n^2(n+1)^2}{4} \end{aligned}$$

This continues to hold for arbtirary upper and lower summation limits:

$$\sum_{k=a}^{b} k = a + a(a+1) + (a+2) + \cdots + b = (b-a+1)\frac{a+b}{2}$$

Binomial theorem

Let n be a positive integer. Then for any x, y,

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

Multinomial theorem

Let n and r be positive integers. Then for any $x_1, \ldots, x_r$,

$$(x_1 + x_2 + \cdots + x_r)^n = \sum_{\substack{k_1 \geq 0, k_2 \geq 0, \ldots, k_r \geq 0 \\ k_1 + k_2 + \cdots + k_r = n}} \binom{n}{k_1, k_2, \ldots, k_r} x_1^{k_1} x_2^{k_2} \ldots x_r^{k_r}$$

The sum runs over r-tuples $(k_1, k_2, \ldots, k_r)$ of nonnegative integers that add up to n.

2 Experiments with Random Outcomes

2.1 Sample Spaces and Probabilities

Sample space Ω: the set of all possible outcomes of the experiment.
• elements of Ω are called **sample points** and typically denoted by ω
• subsets of Ω are called **events**. The collection of events in Ω is denoted by $\mathcal{F}$.
a • the **probability measure** or **probability distribution** or **probability** P is a function from $\mathcal{F}$ into the real numbers. Each event A has a probability $P(A)$, and P satisfies the following axioms:
(a) $0 \leq P(A) \leq 1$ for each event A
(b) $P(\Omega) = 1$ and $P(\emptyset) = 0$.
(c) If $A_1, A_2, A_3, \ldots$ is a sequence of pairwise disjoint events then

$$P\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

The triple $(\Omega, \mathcal{F}, P)$ is called a **probability space**. Every mathematically precise model of a random experiment or collection of experiments must be of this kind.

The three axioms related to the probability measure P above are known as **Kolmogorov's Axioms**. **Pairwise disjoint** means that $A_i \cap A_j = \emptyset$ for each pair of indices $i \neq j$, i.e. the events A_i are **mutually exclusive**. Axiom (iii) says that the probability of the union of the mutually exclusive events is equal to the sum of their probabilities. Note that rule (iii) applies also to finitely many events.

If $A_1, A_2, \ldots, A_n$ are pairwise disjoint events then

$$P(A_1 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n)$$

Example: We flip a fair coin. The sample space is $\Omega = \{H, T\}$ (H for heads and T for tails). We take $\mathcal{F} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$, the collection of all subsets of Ω. The term "fair coin" means that the two outcomes are equally likely. So the probabilities of the singletons $\{H\}$ and $\{T\}$ are

$$P\{H\} = P\{T\} = \frac{1}{2}$$

By axiom (ii) we have $P(\emptyset) = 0$ and $P\{H, T\} = 1$.

Example 2: We roll a standard six-sided die. Then the sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$. Each sample point ω is an integer between 1 and 6. If the die is fair then each outcome is equally likely, i.e.

$$P\{1\} = P\{2\} = P\{3\} = P\{4\} = P\{5\} = P\{6\} = \frac{1}{6}$$

A possible event in this sample space is

$$A = \{\text{the outcome is even}\} = \{2, 4, 6\}$$

Then

$$P(A) = P\{2, 4, 6\} = P\{2\} + P\{4\} + P\{6\} = \frac{1}{2}$$

2.2 Random Sampling

If the sample space Ω has finitely many elements and each outcome is equally likely then for any event $A \subset \Omega$ we have

$$P(A) = \frac{A}{\Omega}.$$

Example: Suppose our urn contains 5 balls labeled 1, 2, 3, 4, 5. Sample 3 balls with replacement and produce an ordered list of the numbers drawn. At each step we have the same 5 choices. The sample space is

$$\Omega = \{1, 2, 3, 4, 5\}^3 = \{(s_1, s_2, s_3) : \text{ each } s_i \in \{1, 2, 3, 4, 5\}\}$$

and $\Omega = 5^3$. Since all outcomes are equally likely, we have for example

$$P\{\text{the sample is (2, 1, 5)}\} = P\{\text{the sample is (2, 2, 3)}\} = 5^{-3} = \frac{1}{125}$$

Example 2: Consider again the urn with 5 balls labeled 1, 2, 3, 4, 5. Sample 3 balls without replacement and produce an ordered list of the numbers drawn. Now the sample space is

$$\Omega = \{(2_1, s_2, s_3) : \text{ each } s_i \in \{1, 2, 3, 4, 5\} and s_1, s_2, s_3 \text{ are all distinct}$$

The first ball can be chosen in 5 ways, the second ball in 4 ways, and the third ball in 3 ways. So

$$P\{\text{the sample is (2,1,5)}\} = \frac{1}{5 \cdot 4 \cdot 3} = \frac{1}{60}$$

The outcome (2, 2, 3) is not possible because repetition is not allowed.

Example 3: Suppose our urn contains 5 balls labeled 1, 2, 3, 4, 5. Sample 3 balls without replacement and produce an unordered set of 3 numbers as the outcome. The sample space is

$$\Omega = \{\omega : \omega \text{ is a 3-element subset of } \{1,2,3,4,5\}\}$$

For example

$$P(\text{the sample is } \{1,2,5\}) = \frac{1}{\binom{5}{3}} = \frac{2! \cdot 3!}{5!} = \frac{1}{10}$$

The outcome $\{2,2,3\}$ does not make sense as a set of three numbers because of the repetition.

2.3 Infinitely Many Outcomes

Example: Flip a fair coin until the first tails comes up. Record the number of flips required as the outcome of the experiment. What is the space Ω of possible outcomes? The number of flips needed can be any positive integer, hence Ω must contain all positive integers. We can also imagine the scenario where tails never comes up. This outcome is represented by . Thus

$$\Omega = \{\infty, 1, 2, 3, \dots\}.$$

The outcome is k iff the first $k-1$ flips are heads and the kth flip is tails. This is one of the 2^k equally likely outcomes when we flip a coin k times, so the probability of this event is 2^{-k}. Thus

$$P\{k\} = 2^{-k} \text{ for each positive integer } k$$

This equation defines the **geometric probability distribution** with success parameter $\frac{1}{2}$ on the positive integers. The probability $P\{\infty\}$ can be derived from the axioms of probability:

$$1 = P(\Omega) = P\{\infty, 1, 2, 3 \dots\} = P\{\infty\} + \sum_{k=1}^{\infty} P\{k\}$$

Because $P\{k\} = 2^{-k}$ for each positive integer k, we have

$$\sum_{k=1}^{\infty} P\{k\} = \sum_{k=1}^{\infty} 2^{-k} = 1,$$

which implies that $P\{\infty\} = 0$.

Example 2: Consider a dartboard in the shape of a disk with a radius of 9 inches. The bullseye is a disk of diameter $\frac{1}{2}$ inch in the middle of the board. What is the probability that a dart randomly thrown on the board hits the bullseye? Let us assume that the dart hits the board at a uniformly chosen random location, that is, the dart is equally likely to hit anywhere on the board.

The sample space is a disk of radius 9. For simplicity take the center as the origin of our coordinate system, so

$$\Omega = \{(x,y) : x^2 + y^2 \leq 9^2\}$$

Let A be the event that represents hitting the bullseye. This is the disk or radius $\frac{1}{4}$: $A = (x,y) : x^2 + y^2 \leq \left(\frac{1}{4}\right)^2$. The probability should be uniform on the disk Ω, thus

$$P(A) = \frac{\text{area of } A}{\text{area of } \Omega} = \frac{\pi \left(\frac{1}{4}\right)^2}{\pi \cdot 9^2} = \frac{1}{36^2} \approx 0.00077$$

The set $\Omega = \{\infty, 1, 2, 3, \dots\}$ is **countably infinite**, meaning that its elements can be arranged in a sequence, or equivalently, labeled by positive integers. A countably infinite sample space works like a finite sample space. To specify a probability measure P, it is enough to specify the probabilities of the outcomes and then derive the probability of each event by additivity:

$$P(A) = \sum_{\omega : \omega \in A} P\{\omega\} \text{ for any event } A \subset \Omega$$

Finite and countably infinite sample spaces are both called **discrete sample spaces**.

2.4 Consequences of the Rules of Probabaility

Additivity property: if $A_1, A_2, A_3, \dots$ are pairwise disjoint events and A is their union, then $P(A) = P(A_1) + P(A_2) + P(A_3) + \dots$. Calculation of the probability of a complicated event A almost always involving decomposing A into smaller disjoint pieces whose probabilities are easier to find.

Example: An urn contains 30 red, 20 green, and 10 yellow balls. Draw two without replacement. What is the probability that the sample contains exactly one red or exactly one yellow?

We have

$$P(\text{exactly one red or exactly one yellow}) = P(\text{red and green}) + P(\text{yellow and green}) + P(\text{red and yellow})$$

Counting favorable arrangements for each of the simpler events gives:

$$
\begin{aligned}
P(\text{red and green}) &= \frac{30 \cdot 20}{\binom{60}{2}} = \frac{20}{59} \\
P(\text{yellow and green}) &= \frac{10 \cdot 20}{\binom{60}{2}} = \frac{20}{177} \\
P(\text{red and yellow}) &= \frac{30 \cdot 10}{\binom{60}{2}} = \frac{10}{59}
\end{aligned}
$$

which leads to

$$
P(\text{exactly one red or exactly one yellow}) = \frac{20}{59} + \frac{20}{177} + \frac{10}{59} = \frac{110}{177}
$$

We used unordered samples but we can get the answer also by using ordered samples.

Example 2: Peter and Mary take turns rolling a fair die. If Peter rolls 1 or 2 he wins and the game stops. If Mary rolls 3, 4, 5, or 6, she wins and the game stops. They keep rolling in turn until one of them wins. Suppose Peter rolls first.

(a) What is the probability that Peter wins and rolls at most 4 times?

For this event to occur, Peter must win on his first roll, win on his second roll, win on his third roll, or win on his fourth roll. These alternatives are mutually exclusive. So define events

$$
A = \{\text{Peter wins and rolls at most 4 times}\}
$$

and $A_k = \{\text{Peter wins on his } k\text{th roll}\}$. Then $A = \cup_{k=1}^{4} A_k$ and since the events A_k are mutually exclusive, $P(A) = \sum_{k=1}^{4} P(A_k)$. To find the probabilities $P(A_k)$ we need to think about the game and the fact that Peter rolls first. Peter wins on his kth roll if first both Peter and Mary fail $k-1$ times and then Peter succeeds. Each roll has 6 possible outcomes. Peter's roll fails in 4 different ways and Mary's roll fails in 2 different ways. Peter's kth roll succeeds in 2 different ways. Thus the ratio of the number of favorable alternatives over the total number of alternatives gives

$$
P(A_k) = \frac{(4 \cdot 2)^{k-1} \cdot 2}{(6 \cdot 6)^{k-1} \cdot 6} = \left(\frac{8}{36}\right)^{k-1} \frac{2}{6} = \left(\frac{2}{9}\right)^{k-1} \frac{1}{3}
$$

The probability asked is now obtained from a finite geometric sum:

$$
\begin{aligned}
P(A) &= \sum_{k=1}^{4} P(A_k) \\
&= \sum_{k=1}^{4} \left(\frac{2}{9}\right)^{k-1} \frac{1}{3} \\
&= \frac{1}{3} \sum_{j=0}^{3} \left(\frac{2}{9}\right)^{j} \\
&= \frac{1}{3} \cdot \frac{1 - \left(\frac{2}{9}\right)^4}{1 - \frac{2}{9}} \\
&= \frac{3}{7} \left(1 - \left(\frac{2}{9}\right)^4\right)
\end{aligned}
$$

Complements: Events A and A^C are disjoint and together make up Ω, no matter what the event A happens to be. Consequently

$$
P(A) + P(A^C) = 1
$$

A larger event must have larger probability:

$$
\text{if } A \subseteq B \text{ then } P(A) \le P(B)
$$

Inclusion-exclusion formulas for two and three events

$$
\begin{aligned}
P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\
P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)
\end{aligned}
$$

Example 1 (revisited): We solved this problem earlier by decomposing the event. Now apply first inclusion-exclusion and then count favorable arrangements using unordered samples:

$$
\begin{aligned}
P(\text{exactly one red or exactly one yellow}) &= P(\{\text{exactly one red}\}) \cup \{\text{exactly one yellow}\}) \\
&= P(\text{exactly one red}) + P(\text{exactly one yellow}) - P(\text{exactly one red and exactly one y} \\
&= \frac{30 \cdot 30}{\binom{60}{2}} + \frac{10 \cdot 50}{\binom{60}{2}} - \frac{30 \cdot 10}{\binom{60}{2}} \\
&= \frac{110}{177}
\end{aligned}
$$

General inclusion-exclusion formula

P(A$_1 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i_1 \le i_2 \le n} P(A_{i_1} \cap A_{i_2})$
$+ \sum_{1 \le i_1 \le i_2 \le i_3 \le n} P(A_{i_1} \cap A_{i_2} \cap A_{i_3})$
$- \sum_{1 \le i_1 \le i_2 \le i_3 \le i_4 \le n} P(A_{i_1} \cap A_{i_2} \cap A_{i_3} \cap A_{i_4})$
$+ \cdots + (-1)^{n+1} P(A_1 \cap \cdots \cap A_n)$
$= \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n} P(A_{i_1} \cap \cdots \cap A_{i_k})$

2.5 Random Variables: A First Look

2.6 Finer Points