Lecture 2

More Counting and Probability

Review Problems

1. How many different strings of length 12 can be made from the letters in “AABBBCCDDDE”?

2. A student must answer 7 out of 10 questions on an exam.
   (a) How many ways are there to take the exam?
   (b) If the student also must answer at least 3 out of the first 5 questions, how many ways are there?
   (c) Which is larger?

3. Ten people are split into 2 teams of 5 that will compete. How many possible match-ups are there?

4. A standard deck of 52 playing cards has 12 picture cards. How many ways are there to deal the cards to 4 players (each getting 13 cards) so that each player receives exactly 3 picture cards. Assume the order in which a given player receives the cards doesn’t matter.

Review Solutions

1. \( \frac{12!}{2!3!3!3!1!} = \binom{12}{2,3,3,3,1} \)

2. (a) \( \binom{10}{7} \)
   (b) \( \binom{5}{3}\binom{5}{4} + \binom{5}{4}\binom{5}{3} + \binom{5}{3}\binom{5}{2} \)
   (c) The first is bigger (less restrictive).

3. \( \binom{10}{5}/2 \) (each pair of teams is counted twice)

4. \( \binom{12}{3,3,3,3}\binom{40}{10,10,10,10} \)
Multinomial Theorem

Consider the expressions \((x + y)^n\) and \((x + y + z)^n\). By looking at \((x + y)^2\) and \((x + y)^3\) we can see that they are sums of all possible strings of length \(n\) using \(x, y,\) and \(x, y, z\) respectively. Using the counting method above we can group the strings by their counts of \(x\)'s, \(y\)'s and \(z\)'s to obtain

**Theorem 1** (Multinomial Theorem). For any \(x, y, z \in \mathbb{R}\), and \(n \geq 0\) we have

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

and

\[
(x + y + z)^n = \sum_{0 \leq i,j,k \leq n \atop i+j+k=n} \binom{n}{i,j,k} x^i y^j z^k
\]

where

\[
\binom{n}{i,j,k} = \frac{n!}{i!j!k!}.
\]

This can be extended to any number of variables.

Counting Proofs

One type of proof technique the occurs frequently in combinatorics is counting the same quantity using two different methods. Here is an example.

**Theorem 2.**

\[
\binom{n}{k} = \frac{n(n-1) \cdots (n-k+1)}{k!}
\]

**Proof.** Consider all ordered sequences of length \(k\) taken from a set of size \(n\). By the multiplication principle there are \(n(n-1) \cdots (n-k+1)\) of these. Counted differently, first select one of the \(\binom{n}{k}\) subsets of size \(k\) to use. Then choose one of the \(k!\) orderings for these \(k\) elements. \(\square\)

**Theorem 3.**

\[
\frac{(2n)!}{2^n n!} = (2n-1)(2n-3) \cdots 1.
\]

**Proof.** Suppose we want to pair up \(2n\) people to form \(n\) teams of two. Give each person a number from \(1, \ldots, 2n\). If at each step we choose the lowest numbered unpaired player to pair up, we get the formula on the right. Now that we know the number of ways to pair up people, we can count the number of ways to order \(2n\) people. First choose the \(n\) pairs. Then choose one of the \(n!\) orderings of the pairs. Then choose one of the \(2\) orderings within each pair. This shows

\[
(2n)! = (2n-1)(2n-3) \cdots 1 \cdot n! \cdot 2^n.
\]

\(\square\)
Counting Proof Exercises

1. What is $(3x + 4y)^3$?
2. Prove that $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ using a counting proof.
3. Prove that $\binom{n+m}{k} = \sum_{j=0}^{k} \binom{n}{j}\binom{m}{k-j}$ using a counting proof.
4. What is $\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots + (-1)^n\binom{n}{n}$?

Solutions

1. $\binom{3}{0}(3x)^3 + \binom{3}{1}(3x)^2(4y) + \binom{3}{2}(3x)(4y)^2 + \binom{3}{3}(4y)^3$.

2. Consider the set $\{1, \ldots, n\}$. If we want to choose a $k$ element subset we can either include $n$ or not. If we include $n$, we must then choose $k-1$ elements from the remaining $n-1$. If we don’t include $n$ we must choose $k$ elements from the remaining $n-1$.

3. Consider $n$ red numbers and $m$ black numbers. Suppose we want to choose $k$ total numbers. Then we can sum over the total quantity of red numbers we will choose.

4. This is $(1 + (-1))^n = 0$.

Axioms of Probability

Suppose we are modeling a random world or experiment. Let $S$ (called the sample space) denote the set of all possible outcomes of our experiment, or possible states of our world. Subsets of $S$ are called events. In a graduate course, we would get very precise on which subsets of $S$ are allowed to be events. In this course, we won’t worry about this point. Define a function $P$ (called a probability measure) whose domain is the set of events (remember events are subsets of $S$) and whose codomain is a real number such that

1. $0 \leq P(E) \leq 1$,
2. $P(S) = 1$,
3. (Countable additivity)

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

where $E_i$ are pairwise disjoint. That is, $E_i \cap E_j = \emptyset$ for $i \neq j$. 
A few notes on the axioms:

1. This is it. All of the probability theory in this course will come from the 3 rules above.
2. The probability function (measure) $P$ takes subsets of $S$ as input, not elements of $S$.
3. Sometimes people include $P(\emptyset) = 0$ in the axioms, but we can just derive it. Let $E_1 = S$ and $E_k = \emptyset$ for $k \geq 2$. Then by countable additivity

$$1 = P(S) = P\left( \bigcup_{k=1}^{\infty} E_k \right) = P(E_1) + \sum_{k=2}^{\infty} P(E_k) = 1 + \sum_{k=2}^{\infty} P(\emptyset).$$

Thus we must have $P(\emptyset) = 0$.

We can derive various properties of probability from these axioms. For example:

**Theorem 4.** If $A, B$ are events with $A \subset B$ then $P(A) \leq P(B)$.

**Proof.** We can write $B = A \cup (B \cap A^c)$ so that

$$P(B) = P(A) + P(B \cap A^c) \geq P(A).$$

The above is valid, assuming we have finite additivity (which isn’t one of the axioms). To complete the proof we will establish finite additivity in the exercises below.

**Axiom Exercises**

1. For any $A \subset S$, derive an expression for $P(A^c)$ in terms of $P(A)$.
2. (Finite Additivity) For any $n$ events $A_1, \ldots, A_n \subset S$ with $A_i \cap A_j = \emptyset$ for $i \neq j$ prove that

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

3. For any events $A, B$ prove $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
4. (⋆⋆) (Countable Subadditivity, or Boole’s Inequality) For any events $A_1, \ldots$ (not necessarily disjoint) we have

$$P\left( \bigcup_{i=1}^{\infty} A_i \right) \leq \sum_{i=1}^{\infty} P(A_i).$$
Solutions

1. Note that \( P(A^c) + P(A) = P(S) = 1 \) so \( P(A^c) = 1 - P(A) \).

2. Define \( A_{n+1} = \emptyset, A_{n+2} = \emptyset, \ldots \) and apply countable additivity.

3. Note that \( P(A \cup B) = P(A) + P(B \cap A^c) \) (can help to look at a Venn diagram). Then note that \( P(B \cap A^c) + P(B \cap A) = P(B) \).

4. Let \( B_1 = A_1, B_2 = A_2 \cap A_1^c, B_3 = A_3 \cap A_2^c \cap A_1^c, \) etc. In general \( B_n = A_n \cap \bigcap_{k=1}^{n-1} A_k^c \). Then the \( B_i \) are disjoint sets with the same union. Hence

\[
P \left( \bigcup_{i=1}^{\infty} A_i \right) = P \left( \bigcup_{i=1}^{\infty} B_i \right) = \sum_{i=1}^{\infty} P(B_i) \leq \sum_{i=1}^{\infty} P(A_i).
\]

Inclusion-Exclusion for Probabilities

We have already seen that

\[
P(A \cup B) = P(A) + P(B) - P(A \cap B).
\]

Using induction, you can extend this property \( n \) sets.

\[
P \left( \bigcup_{k=1}^{n} A_k \right) = \sum_{k=1}^{n} P(A_k) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) + \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n).
\]

You can easily see this formula at work for \( n = 3 \) using a Venn diagram. The idea of the general proof is to write

\[
\bigcup_{k=1}^{n+1} A_k = A_{n+1} \cup \bigcup_{k=1}^{n} A_k
\]

and apply the case for 2 sets followed by the case for \( n \) sets. It can also be shown that each time we subtract or add (before the last) we are overcorrecting. More precisely,

\[
P \left( \bigcup_{k=1}^{n} A_k \right) \leq \sum_{k=1}^{n} P(A_k)
\]

\[
P \left( \bigcup_{k=1}^{n} A_k \right) \geq \sum_{k=1}^{n} P(A_k) - \sum_{i<j} P(A_i \cap A_j)
\]

\[
P \left( \bigcup_{k=1}^{n} A_k \right) \leq \sum_{k=1}^{n} P(A_k) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)
\]

\[\vdots\]
Examples of Probabilities

**Example 5** (Equally Likely Outcomes). Let $S$ be finite, and define $P(A)$ for $A \subset S$ to be

$$P(A) = \frac{|A|}{|S|}.$$

We note that $0 \leq P(A) \leq 1$, $P(S) = 1$, and if $A_1, A_2, \ldots$ are disjoint then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^{k} A_i\right) \quad (S \text{ is finite})$$

$$= \frac{\left|\bigcup_{i=1}^{k} A_i\right|}{|S|} \quad \text{(Defn of } P)$$

$$= \sum_{i=1}^{k} \frac{|A_i|}{|S|} \quad \text{(Addition Principle)}$$

$$= \sum_{i=1}^{k} P(A_i) \quad \text{(Defn of } P).$$

Above we use the fact that since $S$ is finite, there could not be infinitely many disjoint nonempty sets in our union. Thus there must be a $k$ so that $A_j = \emptyset$ for all $j > k$.

**Example 6** (Arbitrary Coin Flip). Fix $p \in [0, 1]$, let $S = \{H, T\}$, and define

$$P(\{H\}) = p, \quad P(\{T\}) = 1 - p, \quad P(\emptyset) = 0, \quad P(S) = 1.$$

Then all properties hold. We will often use the letter $q$ to denote $1 - p$.

**Example 7** (General Finite Space). Let $S$ be a finite set with distinct elements $s_1, \ldots, s_n$. Define, for $i = 1, \ldots, n$,

$$P(\{s_i\}) = p_i$$

where each $p_i \in [0, 1]$, and $\sum_{i=1}^{n} p_i = 1$. For more general events $A \subset S$, define

$$P(A) = \sum_{s_i \in A} P(\{s_i\}) = \sum_{s_i \in A} p_i.$$

This can be shown to have all the required properties.

This example includes both previous examples.

**Example 8** (General Countably Infinite Space). Let $S$ be an infinite set with distinct elements $s_1, s_2, \ldots$. Define, for $i = 1, 2, \ldots$,

$$P(\{s_i\}) = p_i$$

where each $p_i \in [0, 1]$, and $\sum_{i=1}^{\infty} p_i = 1$. For more general events $A \subset S$, define

$$P(A) = \sum_{s_i \in A} P(\{s_i\}) = \sum_{s_i \in A} p_i.$$

This can be shown to have all the required properties.