CS1231 Part 10 - Counting and Probability, Continued

Based on lectures by Terence Sim and Aaron Tan
Notes taken by Andrew Tan
AY18/19 Semester 1

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

1 Combinations

Combinations refer to the various ways in which objects may be selected in a set, ignoring order. Each subset of size r of a given set is called an **r-combination** of the set.

More formally, let n and r be non-negative integers with $r \leq n$. An **r-combination** of a set of n elements is a subset of r of the n elements.

$\binom{n}{r}$, read as "n choose r", denotes the number of subsets of size r that can be chosen from a set of n elements. We can also denote it by $C(n, r)$, ${}^nC_r$

1.1 Relationship between permutations and combinations

Recall that an r-permutation of a set of n elements is an ordered selection of r elements from the set. We can think of constructing an r-permutation of a set of n elements in 2 steps:

1. Choose a subset of r elements from the set.
2. Choose an ordering for the r-element subset.

The number of ways to perform step 1 is $\binom{n}{r}$, and the number of ways to perform step 2 is $r!$. Hence, we obtain the equation:

$$P(n, r) = \binom{n}{r} \times r!$$

Thus, the number of subsets of size r that can be chosen from a set of n elements, $\binom{n}{r}$, is given by the formula

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

Note that this implies that

$$\binom{n}{r} = \binom{n}{n-r}$$

We can interpret this by saying that a set A with n elements has exactly as many subsets of size r as it has subsets of size $n - r$.

1.2 Permutations of a set with repeated elements

Suppose a collection consists of n objects of which

n_1 are of type 1 and are indistinguishable from each other

n_2 are of type 2 and are indistinguishable from each other

...

n_k are of type k and are indistinguishable from each other and suppose that $n_1 + n_2 + \cdots + n_k = n$. Then the number of distinguishable permutations of the n objects is

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-n_2-\cdots-n_{k-1}}{n_k} = \frac{n!}{n_1!n_2!n_3!\ldots n_k!}$$

Essentially, for each type i, we are choosing a subset of n_i positions from the number of remaining positions available.

2 Multisets

Multisets are combinations with repetition allowed. More formally, a multiset of size r, chosen from a set X of n elements, is an unordered selection of elements taken from X with repetition allowed.

If $X = \{x_1, x_2, \ldots, x_n\}$, we write a multiset as $[x_{i_1}, x_{i_2}, \ldots, x_{i_r}]$, where each x_{i_j} is in X and some of the x_{i_j} may equal each other.

The number of multisets of size r that can be selected from a set of n elements is:

$$\binom{r+n-1}{r}$$

To understand this, consider the numbers $1, 2, 3, 4$ in $\{1, 2, 3, 4\}$ as categories, and imagine choosing a total of 3 numbers from the categories with multiple selections from any category allowed.

	Category 1	Category 2	Category 3	Category 4
$[1, 1, 1]$:	xxx	|	|	|
$[1, 3, 4]$:	x	|	| x	| x
$[2, 4, 4]$:		| x	|	| xx

Hence, we may write $[1, 1, 1]$ as "$xxx|||$, $[1, 3, 4]$ as "$x||x|x$, and $[2, 4, 4]$ as "$|x||xx$. Thus, finding the number of multisets of size 3 is the same as $\binom{4+3-1}{3} = \binom{6}{3}$.

2.1 Using the appropriate formula

As we have discussed permutations and combinations, we shall summarize the different cases and the corresponding formulas as shown

	order Matters	order does not matter
Repetition is allowed	n^k	$\binom{r+n-1}{r}$
Repetition is not allowed	$P(n, k)$	$\binom{n}{k}$

3 Pascal's Formula and the Binomial Theorem

Suppose n and r are positive integers with $r \leq n$. Then Pascal's formula denotes the following:

$$\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}$$

Pascal's triangle is a geometric version of Pascal's forumla:

n							
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10	10	5	1	
6	1	6	15	20	15	6	1
	0	1	2	3	4	5	6
				k			

Notice how Pascal's formula is easily visualized through the triangle.

Furthermore, we can use Pascal's formula to continuously derive a formula for other values by substitution. For example,

$$\binom{n+2}{r} = \binom{n+1}{r-1} + \binom{n+1}{r} = [\binom{n}{r-2} + \binom{n}{r-1}] + [\binom{n}{r-1} + \binom{n}{r}] = \binom{n}{r-2} + 2\binom{n}{r-1} + \binom{n}{r}$$

3.1 Binomial theorem

In algebra, the sum of two terms, such as $a + b$, is called a binomial. The binomial theorem gives an expression for the powers of a binomial $(a+b)^n$, for each positive integer n and all real numbers a and b.

More formally, the binomial theorem states that given any real numbers a and b ad any non-negative integer n,

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = a^n + \binom{n}{1} a^{n-1} b^1 + \binom{n}{2} a^{n-2} b^2 + \cdots + \binom{n}{n-1} a^1 b^{n-1} + b^n$$

4 Probability Axioms and Expected Value

Recall that a sample space is a set of all outcomes of a random process or experiment, and that an event is a subset of a sample space.

Now let S be a sample space. A **probability function** P from the set of all event in S to the set of real numbers satisfies the following **probability axioms**: For all events A and B in S,

1. $0 \leq P(A) \leq 1$
2. $P(\emptyset) = 0$ and $P(S) = 1$
3. If A and B are disjoint (i.e. $A \cap B = \emptyset$), then $P(A \cup B) = P(A) + P(B)$.
4. $P(A^c) = 1 - P(A)$
5. If A and B are any events in a sample space S, then $P(A \cup B) = P(A) + P(B) + P(A \cap B)$.
6. If $A \subseteq B$, then $P(B - A) = P(B) - P(A)$.

4.1 Expected Value

Suppose the possible outcomes of an experiment, or random process, are real numbers $a_1, a_2, a_3, \ldots, a_n$ which occur with probabilities $p_1, p_2, p_3, \ldots, p_n$. The **expected value** of the process is

$$\sum_{k=1}^{n} a_k p_k = a_1 p_1 + a_2 p_2 + a_3 p_3 + \cdots + a_n p_n$$

4.1.1 Linearity of Expectation

The expected value of the sum of random variables is equal to the sum of their individual expected values, regardless of whether they are independent. For random variables X and Y (which may be dependent),

$$E[X + Y] = E[X] + E[Y]$$

More generally, for random variables $X_1, X_2, \ldots, X_n$ and constants $c_1, c_2, \ldots, c_n$,

$$E[\sum_{i=1}^{n} c_i X_i] = \sum_{i=1}^{n} (c_i E[X_i])$$

In essence, we can think of expectation as a linear function.

5 Conditional Probability

Let A and B be events in a sample space S. If $P(A) \neq 0$, then the **conditional probability of B given A**, denoted $P(B|A)$, is

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

5.1 Bayes' Theorem

Suppose that a sample space S is a union of mutually disjoint events $B_1, B_2, B_3, \ldots, B_n$. Suppose A is an event in S, and suppose A and all the B_i have non-zero probabilities. If k is an integer with $1 \leq k \leq n$, then

$$P(B_k|A) = \frac{P(A|B_k) \cdot P(B_k)}{P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + \cdots + P(A|B_n) \cdot P(B_n)}$$

5.2 Independent Events

If two events A and B are **independent**, then the probability of one occuring shouldn't affect the probability of the other occuring. More succinctly,

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

Furthermore, A and B are independent if and only if

$$P(A \cap B) = P(A) \cdot P(B)$$

5.3 Pairwise independence/Mutually independent

We say that three events A, B, and C are **pairwise independent** if and only if

$$P(A \cap B) = P(A) \cdot P(B)$$

$$P(A \cap C) = P(A) \cdot P(C)$$

$$P(B \cap C) = P(B) \cdot P(C)$$

Events can be pairwise independent without satisfying the condition $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$. Conversely, they can satisfy the condition $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ without being pairwise independent.

We extend the definition to include **mutual independence** as follows:
Let A, B, and C be events in a sample space S. A, B, and C are **pairwise independent** if and only if they satisfy conditions $1 - 3$ below. They are **mutually independent** if and only if they satisfy all four conditions below.

$$P(A \cap B) = P(A) \cdot P(B)$$

$$P(A \cap C) = P(A) \cdot P(C)$$

$$P(B \cap C) = P(B) \cdot P(C)$$

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$

We can also generalize the definition of mutual independence as follows:
Events $A_1, A_2, \ldots, A_n$ in a sample space S are **mutually independent** if and only if the probability of the intersection of any subset of the events is the product of the probabilities of the events in the subset.