Counting Theory, and Probability Theory

Composed by Nirattaya Khamsemanan, Ph.D

Disclaimer: this partial note is my attempt to help you learn the material in this course. Things in here might not be in the real exam and vice versa. Don't use this as your main study. There might be some typos and/or mistakes.

Be advised: these partial notes have been created for the sole purpose of aiding your studies in this class, and are *for personal use only*. They may not be duplicated, copied, modified or translated without my written consent.

1 Counting Theory

We are going to review a few basic knowledge of counting theory which is quite important in probability theory that we will learn in the next section. The definition of the probability of a certain type of outcome among equally likely events requires us to find two numbers: the total number n of outcomes possible and the number k of favorable outcomes. In many important types of probability problem, the counting problems are more complicated. To be sure that the numbers are correct, we need a systematic way to count. These counting procedures can turn the calculation of probability in even quite intricate problems into rather routine exercises.

Recall that,

Definition 1. An ordered k-tuple is a collection of k symbols in a fixed order.

For example, a polling organization divides the population into groups, call "strata" according to sex, race (white or nonwhite) and age (18-35, 36-50, over 50). For example, "white male between the ages of 18 and 5" is one stratum. We can represent a stratum as an ordered 3-tuple of the form

Theorem 1. If there are n_1 choices for the first member of an ordered k-tuple, n_2 for the second, regardless of the first choice, n_3 for the third, regardless of the first two choices, and so on, up to n_k forthe k th (independent of what the previous choices are), then there are

$$n_1 \cdot n_2 \cdot \cdot \cdot n_k$$

suck k-tuples in all.

Definition 2. The factorial of a non-negative integer n, denoted by n!, is the product of all positive integers less than or equal to n.

$$n! = 1 \cdot 2 \cdot 3 \cdot \cdots n$$

The zero factorial 0! = 1.

1.1 Permutations

Suppose we are given a set of n objects. Then a permutation of n objects taken k at a time is an arrange of k of the n objects in a specific order. Alternatively a permutation is an ordered k-tuple made up of k distinct elements from a set containing n elements. The total number of permutation s of k elements from a set of n elements is denoted by symbol P_k^n or P(n,k)

Permutation Formula

$$P_k^n = P(n,k) = n(n-1)(n-2)\dots(n-k+1) = \frac{n!}{(n-k)!}$$

1.2 Combination

In many counting problems, there is no sense of order. For example, how many different five-card poker hands can be deal from a standard 52-card deck? To a card player it makes no difference whether the card are picked up in the order. $(3\heartsuit, 5\clubsuit, 3\spadesuit, 3\diamondsuit, 5\diamondsuit)$ or $(5\diamondsuit, 3\spadesuit, 5\clubsuit, 3\heartsuit, 3\diamondsuit)$ or in some other order; the player will be delighted in any case because she has a very good hand—a "full house." If we thought of a poker hand in the terms of the las section, as an order 5-tuple of cards—that is, as a puermutation of five cars out of the deck of 52 then we would be looking at it unrealistically because both of these 5-tuples are the same hand.

For this very common kind of counting problem we need aformula for the number of distinct sets of k objects we can choose out of a collection of n objects, without imposing an ordering on the k objects. An unordered set of this kind is called a combination of k elements out of n, and the number of such combination is denoted by the symbol C_k^n .

Combination Formula

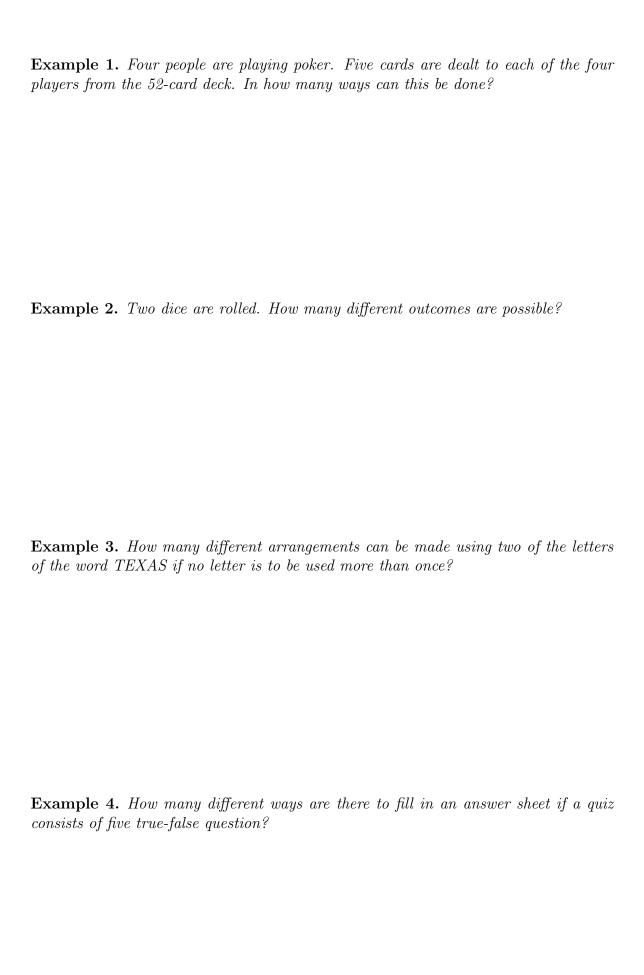
$$C_k^n = C(n, k) = \frac{n!}{k!(n-k)!}$$

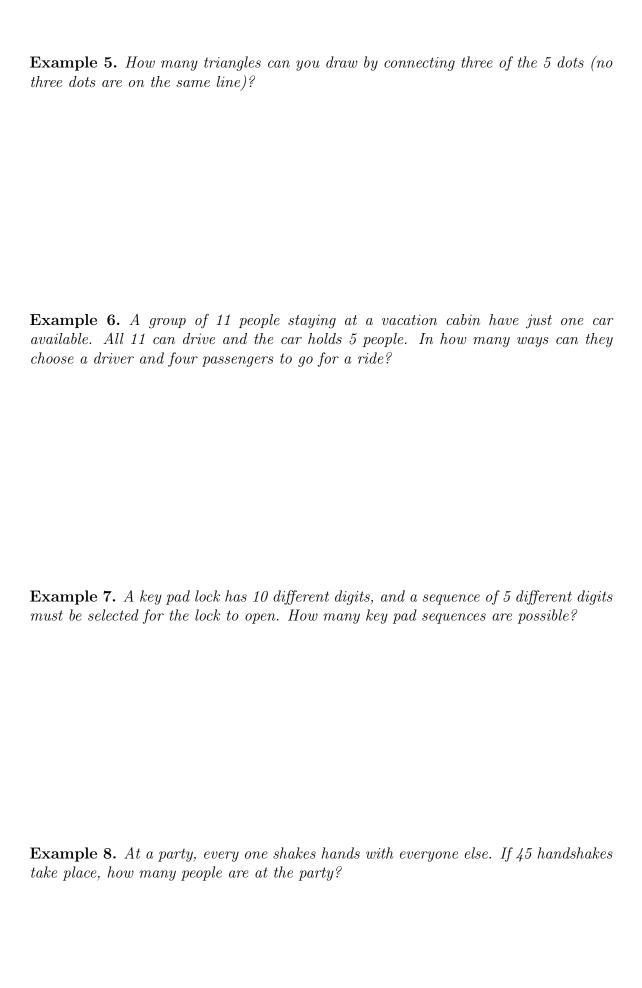
Remark 1. Note that

$$(1) C_k^n = \frac{P_k^n}{k!}$$

$$(2) C_{n-k}^n = C_k^n$$

$$(2) C_{n-k}^n = C_k^n$$





2 Basic Probability Theory

Definition 3. The probability of a favorable outcome A is the number of outcome A can occur divided by the total number of possible outcomes. If A happens k times and the total number of outcome is n, then we say that the probability of A is k/n or

$$P(A) = \frac{k}{n}$$

Example 9. An experiment consists of tossing a coin six times and observing the sequence of heads and tails.

1. How many different outcomes are possible?

2. What is the probability that the outcome has exactly three heads?

3. What is the probability that the outcome has more heads than tails?

4. What is the probability that the outcome at least 4 heads?

Example 10. Two dice are rolled. One of them is red and the other is blue.
1. What is the number of all possible outcomes?
2. Calculate the probability of rolling a sum of 10.
Example 11. Two dice are rolled. You can't distinguish the two of them.
1. What is the number of all possible outcomes?
1. What to the number of all possesse value mee.
2. Calculate the probability of rolling a sum of 10.

Example 12. A standard deck of playing cards consists of 52 different cards. Each deck consists of 4 suits:

$$clubs$$
 (\clubsuit), $diamonds$ (\diamondsuit), $hearts$ (\heartsuit), $spades$ (\spadesuit)

The club cards and the spade cards are black and the diamond cards and the heart cards are red.

Each suit consists of thirteen different values:

The values are ordered from left to right.

Each card is identified by its value and its suit.

The face cards are the Jacks, Queens, Kings. So there are 12 faces card in a deck.

1. What is the probability of being dealt a "royal flush" (Ace, King, Queen, Jack and ten, all of the same card suit) in poker?

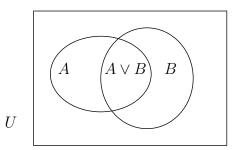
2. What is the probability of being dealt a "four of a kind" (all four cards of one value and any other card)?

2.1 Union of Events

The computation of the number of outcomes either in event A or in event B amounts to counting the number of points within at least one of the regions. If we count the number n(A) of points in event A and add to it the number n(B) of points in event B, we are counting those points that are in A and not in B once each and those points that are in B and not in A once each. But points in both A and B are counted twice, once as members of A and once as members of B. Thus, if we subtract from the sum n(A) + n(B) the number n(A and B) of points that are both in A and B, then we eliminate the double counting of these elements and we obtain a correct count of the number of successful outcomes. WE now have the counting rule:

$$n(A \lor B) = n(A) + n(B) - n(A \land B)$$

It is possible to visualize this geometrically by means of drawings known as Venn diagrams.



where U is set of all possible outcomes (or sample space). And $A \vee B$ is the set of outcomes that is A or B. $A \wedge B$ is the set of outcomes that is A and B.

Computing Probability by Counting If U is the nonempty set of all elements of the sample space, if A and B are two events and if all outcomes are equally likely, then according to the definition of the probability, we have

$$p(A \vee B) = \frac{n(A \vee B)}{n(U)}$$

In this equation, n(U) is the nonzero number of all possible outcomes. The counting rule tells us that

$$p(A \lor B) = \frac{n(A \lor B)}{n(U)}$$

$$= \frac{n(A) + n(B) - n(A \land B)}{n(U)}$$

$$= \frac{n(A)}{n(U)} + \frac{n(B)}{n(U)} - \frac{n(A \land B)}{n(U)}$$

$$= p(A) + p(B) - p(A \land B)$$

Thus,

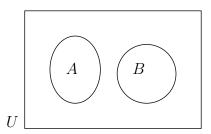
$$p(A \vee B) = p(A) + p(B) - p(A \wedge B)$$

2.2 Disjoint of Events

An important special case of the probability principle

$$p(A \lor B) = p(A) + p(B) - p(A \land B)$$

arises when no outcome is in both event A and even B.



In the language of the set theory, we said that the set A and set B are disjoint because they have no element in common i.e $n(A \wedge B) = 0$. This means that $p(A \wedge B) = 0$ also.

Thus: if event A and B are disjoint,

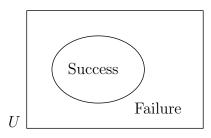
$$p(A \lor B) = p(A) + p(B)$$

2.3 Probability of Failure

No outcome can be both a success and failure (that is, not a success) so

$$p(\text{success} \lor \text{failure}) = p(\text{success}) + p(\text{failure})$$

On the other hand, every outcome is either a success or failure, so $p(\text{success} \lor \text{failure}) = 1$.



This means that

$$p(\text{success}) = 1 - p(\text{failure})$$

 $p(\text{failure}) = 1 - p(\text{success})$

Example 13. A campus radio station survey of music they liked. The survey results follows:	•
60 % like rock 75 % like pop 40 % like rock and pop	
1. How many students from the survey la How many like both?	ke pop music? How many like rock music?
2. Draw an appropriate Venn diagram fo	or these result?
3. What is the probability that a studen music but not pop music?	t who participated in this survey like rock
4. What is the probability that a student music but not rock music?	at who participated in this survey like pop
5. What is the probability that a student either pop nor rock music?	who participated in this survey doesn't like

Example 14. A banker rolls three dice (red, white and blue dice).

1. What is the probability he will get a 4-5-6 that is one of them is a 4, one of them is a 5 and one of them is a 6?

2. What is the probability that he will get any pair and a 6?

3. What is the probability that he will get three-of-a-kind (all three dice are the same)?

4. In a game call See-Low, the banker wins all bets if he throws three-of-a-kind, any pair and a 6, or 4-5-6. What is the probability that he will win?

3 Conditional Probability

In an experiment, the conditional probability that the outcome will come from a subset A, given that the possible outcomes are restricted to a subset B, is defined to be

$$p(A|B) = \frac{p(A \land B)}{p(B)}$$

provided that $p(B) \neq 0$. Or in the other words,

$$p(A \wedge B) = p(B) \cdot p(A|B) = p(A) \cdot p(B|A)$$

Let $A_1, A_2, ..., A_n$ be events in a sample space. The joint probability of $A_1, A_2, ..., A_n$ is given by

$$p(A_1, A_2, ..., A_n) = p(A_1 \land A_2 \land ... \land A_n) = p(A_1) \cdot p(A_2 | A_1) \cdot p(A_3 | A_1, A_2) \cdots p(A_n | A_1, ..., A_{n-1})$$

This is call the probability chain rule.

3.1 Independent of Events

Definition 4. If A and B are any events in a sample space, we say that A and B are independent if and only if

$$p(A \wedge B) = p(A)p(B)$$

otherwise, A and B are said to be dependent.

This means that two events A and B are independent if

$$p(A|B) = p(A)$$

Example 15. Suppose that we toss a coin twice and record the sequence of heads and tails. Let A and B be the events

A = "head on the second toss"

B = "head on the first toss"

Are these events independent?

Example 16. Suppose that we draw two consecutive cards from a 52-card deck. Let A and B be the events

A = "second card is black"

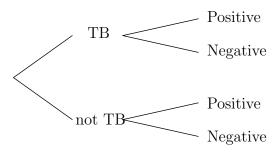
 $B = "first \ card \ is \ red."$

Are these events independent?

3.2 Tree Diagrams

In solving many probability problems, it is helpful to represent the various events and their associated probabilities by a tree diagram. To explain this useful notion, suppose that we wish to compute the probability of an event that results from performing a sequence of experiments. The various outcomes of each experiment are represented as branches emanating from a point.

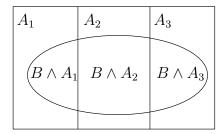
Example 17. Suppose that the reliability of a skin test for active pulmonary tuberculosis (TB) is specified as follows: Of people with TB, 98% have a positive reaction and 2% have a negative reaction; of people free of TB, 99% have a negative reaction and 1% have a positive reaction. From a large population of which 2 per 10,000 persons have TB, a person is selected at random and given a skin test, which turns out to be positive. What is the probability that the person has active pulmonary tuberculosis?



Example 18. Fifty percent of the students enrolled in a business statistics course had previously taken a finite mathematics course. Thirty percent of these students received an A for the statistics course, where as 20% of the other students received an A for the statistic course.
1. Draw a tree diagram and label it with the appropriate probabilities.
2. What is the probability that a student selected at random previously took a finite mathematics course and did not received an A in the statistics course?
3. What is the probability that a student selected at random received an A in the statistics course?
4. What is the conditional probability that a student previously took a finite mathematics course, given that he or she received and A in the statistics course?

3.3 Bayes'Theorem

If $A_1, A_2, ..., A_k$ is a partition of a sample space U of outcomes, and if B is viewed as the set of successful outcomes of the experiment, then we visualize the situation as the figure below.



This means that

$$p(B) = p(A_1)p(B|A_1) + p(A_2)p(B|A_2) + \dots + p(A_1)p(B|A_k)$$

When substitute this into the conditional probability, we get

$$p(A_i|B) = \frac{p(A_i)p(B|A_i)}{p(A_1)p(B|A_1) + p(A_2)p(B|A_2) + \dots + p(A_1)p(B|A_k)}$$

This formula for $p(B_i|A)$ is called **Bayes'Theorem**.

$$p(A_1) \quad A_1 \quad \frac{p(B|A_1)}{B} \quad p(A_1 \land B)$$

$$p(A_2) \quad A_2 \quad \frac{p(B|A_2)}{B} \quad p(A_2 \land B)$$

$$p(A_3) \quad A_3 \quad \frac{p(B|A_3)}{B} \quad p(A_3 \land B)$$

Example 19. Solve Example 17 using Bayes' Theorem.

Example 20. A population in which 3% are users of a drug is screened by a test taht is 80% accurate in the sense that the probability of a false-positive (result of a test means finding a drug that isn't there) is 0.20 and the probability of a false-negative (the failure of the test to find a drug that is there) is 0.15. If a person test positive, calculate the probability that he or she is, in fact, not a user of the drug?

Example 21. Ten percent of the pens made by Apex are defective. Only 5% of the pens made by its competitor, B-ink, are defective. Since Apex pens are cheaper than B-ink pens, an office order 70% of its stock from and Apex and 30% from B-ink. A pen is chosen at random and found to be defective. What is the probability that it was produced by Apex?