

Chapter 2: Probability

Shiwen Shen

University of South Carolina

2017 Summer I

Overview of Probability

- ▶ Probability is a measure of one's belief in the occurrence of a future *event*.
- ▶ The subject of probability theory is the foundation upon which all of statistics is built, providing a means for modeling populations, experiments, or almost anything else that could be considered a random phenomenon.

Some Interesting Problem

- ▶ **(Birthday Problem)** What is the probability that at least 2 people in a class with 48 students have the same birthday (Month and day)?
 - ▶ Year has 365 days (forget leap year).
 - ▶ Equal likelihood for each day.
- ▶ You should be able to solve this problem by the end of this chapter.

Elements of Probability: Random Experiment

Prior to performing the experiment or activity, you do not know for certain which particular outcome will occur. For example,

- ▶ Tossing a coin once.
- ▶ Tossing a coin three times.
- ▶ Observing the lifetime of an electronic component.
- ▶ Drawing three cards (with replacement) from a standard deck of cards.

Elements of Probability: Sample Space

The set of *all possible* outcomes for a given random experiment is called the *sample space*, denoted by S . (discrete vs. continuous)

- ▶ Tossing a coin once. $S = \{\text{Head}, \text{Tail}\}$.
- ▶ Tossing a coin three times.
 $S = \{\text{HHH}, \text{HHT}, \dots, \text{HTT}, \text{TTT}\}$.
- ▶ Drawing three cards (with replacement) from a standard deck of cards. $S = \{?\}$.
- ▶ Observing the lifetime of an electronic component.
 $S = \{\omega : \omega > 0\}$.

The element of the sample space is called *outcome*, denoted by ω . We use n_S to denote the number of elements in the sample space (discrete).

Elements of Probability: Event

- ▶ An *event* is a set of possible outcomes that is of interest. Using mathematical notation, suppose that S is a sample space for a random experiment. We say that A is an event in S if the outcome ω satisfies $\{\omega \in S : \omega \in A\}$.
- ▶ We would like to develop a mathematical framework so that we can **assign** probability to an event A . This will quantify how likely the event is. The probability that the event A occurs is denoted by $P(A)$.
- ▶ More precisely, probability measure is a real-valued function defined on a set of events in a probability space that satisfies measure properties.

Equiprobability Model

- ▶ Suppose that a sample space S contains $n_S < \infty$ outcomes, each of which is **equally likely**. If the event A contains n_A outcomes, then

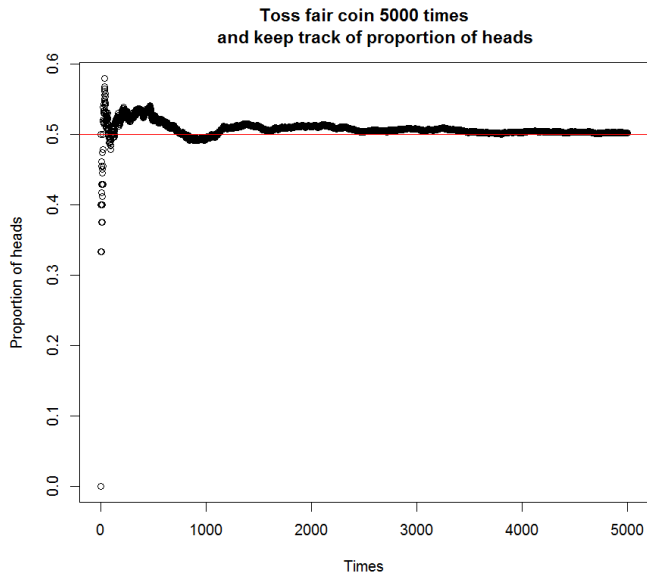
$$P(A) = \frac{n_A}{n_S}.$$

- ▶ Tossing a fair coin three times. Let event $A = \{\text{exactly two heads}\} = \{HHT, HTH, THH\}$. Then, $P(A) = \frac{n_A}{n_S} = \frac{3}{8}$.
- ▶ If I randomly pull a card out of a pack of 52 cards, what is the chance it's a black card? (___) Not a diamond? (___)
- ▶ **Warning:** If the outcomes in S are not equally likely, then this result is not applicable.

Interpretation of Probability

- ▶ One particular interpretation of probability is view probability as a **limiting proportion**.
- ▶ If the experiment is repeatable, then $P(A)$ can be interpreted as the percentage of times that A will occur “over the long run” This is called the **relative frequency** interpretation.

Simulation: Toss A Fair Coin 5000 Times



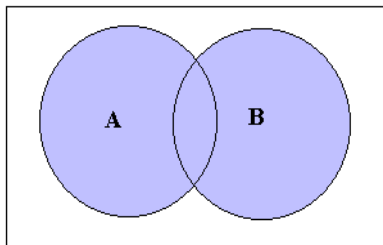
```
x <- rbinom(n=5000,size=1,prob=0.5)
y <- rep(0,5000)
y[1] <- x[1]
for(i in 2:5000){
  y[i] <- y[i-1] + x[i]
}
prop <- y / seq(from=1, to=5000, by=1)
plot(prop, xlab="Times", ylab="Proportion of heads",
      main="Toss fair coin 5000 times
      and keep track of proportion of heads")
abline(h = 0.5, col="red")
```

Union and Intersection of Events

- ▶ The **union** of two events A and B is the set of all possible outcomes ω in either event or both. By notation,

$$A \cup B = \{\omega : \omega \in A \text{ or } \omega \in B\}.$$

- ▶ Using Venn diagram, the union can be expressed as

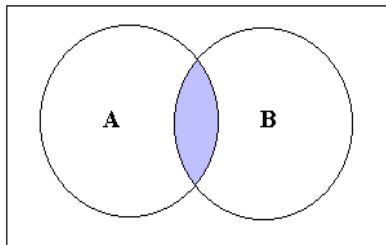


Union and Intersection of Events (cont.)

- ▶ The **intersection** of two events A and B is the set of all possible outcomes ω in the both events. By notation,

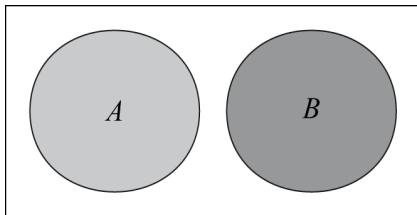
$$A \cap B = \{\omega : \omega \in A \text{ and } \omega \in B\}.$$

- ▶ Using Venn diagram, the intersection can be expressed as



Mutually Exclusive Events

- ▶ **Mutually exclusive** events **can not** occur at the same time. If events A and B are mutually exclusive, then $A \cap B = \emptyset$, where \emptyset is the null event.
- ▶ Using Venn diagram, the mutually exclusive events can be expressed as



Example: Roommate Profile

Suppose your old roommate moves out, and you want to pick up a new roommate, the following table summarizes the roommate profile

| | Snores | Doesn't Snore | Total |
|---------------|--------|---------------|-------|
| Parties | 150 | 100 | 250 |
| Doesn't Party | 200 | 550 | 750 |
| Total | 350 | 650 | 1000 |

Table : Roommate Profile (Frequency - Counts)

Example: Roommate Profile (cont.)

| | Snores | Doesn't Snore | Total |
|---------------|--------|---------------|-------|
| Parties | 0.15 | 0.1 | 0.25 |
| Doesn't Party | 0.2 | 0.55 | 0.75 |
| Total | 0.350 | 0.65 | 1 |

Table : Roommate Profile (Relative Frequency - Probability)

- ▶ What is the probability that a randomly chosen roommate will snore? _____
- ▶ What is the probability that a randomly chosen roommate will like to party? _____
- ▶ What is the probability that a randomly chosen roommate will snore **or** like to party? _____

Union of Two Events

- ▶ If events A and B intersect, you have to subtract out the “double count”. From the Venn diagram, it is easy to see

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

which is called **additive law**.

- ▶ If A and B are **mutually exclusive**, then $A \cap B = \emptyset$ and $P(\emptyset) = 0$, it follows that

$$P(A \cup B) = P(A) + P(B)$$

For any sample space S , a probability P must satisfy

1. $0 \leq P(A) \leq 1$, for any event A
2. $P(S) = 1$
3. If A_1, A_2, \dots, A_n are **pairwise mutually exclusive** events, then

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

Note that “pairwise mutually exclusive” means that $A_i \cap A_j = \emptyset$ for all $i \neq j$, $i, j = 1, 2, \dots, n$.

These three properties are usually referred to as the **Axioms of Probability**. Any function P satisfies the Axioms of Probability is called a probability function. Other nice properties can be derived simply by using them. For example,

- ▶ $P(\emptyset) = 0$
- ▶ If $A \subseteq B$ then $P(A) \leq P(B)$

Question: can you derive these properties using the axioms?

Conditional Probability

- ▶ A conditional probability is the probability that an event will occur, **given or when** another event is known to occur or to have occurred
- ▶ **Definition:** Let A and B be events in a sample space S with $P(B) > 0$. The **conditional probability** of A , given that B has occurred, is

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

- ▶ Solving the conditional probability formula for the probability of the intersection of A and B yields

$$P(A \cap B) = P(A|B) \times P(B).$$

Conditional Probability (cont.)

- ▶ If $B = S$, then the conditional probability of A , given B has occurred, is equal to unconditional probability of A :

$$P(A|B) = P(A|S) = \frac{P(A \cap S)}{P(S)} = \frac{P(A)}{P(S)} = \frac{P(A)}{1} = P(A)$$

- ▶ If $B \neq S$, then we can understand the conditional probability in this “informal” way: we are only interested in the probability of event A when event B has occurred. So, we manually “shrink” the sample space S to be B so that we can calculate the probability of A in the restricted space with the scale of B .

Example: Conditional Probability

In a company, 36 percent of the employees have a degree from a SEC university, 22 percent of the employees that have a degree from the SEC are also engineers, and 30 percent of the employees are engineers. An employee is selected at random.

- ▶ Compute the probability that the employee is an engineer and is from the SEC.
- ▶ Compute the conditional probability that the employee is from the SEC, given that s/he is an engineer.

Solution: Let

$A = \{\text{The employee is an engineer}\}$

$B = \{\text{The employee is from SEC}\}$

Example: Conditional Probability (cont.)

We want to calculate $P(A \cap B)$. From the information in the problem, we have $P(A) = 0.3$, $P(B) = 0.36$, and $P(A|B) = 0.22$. Therefore,

$$P(A \cap B) = P(A|B)P(B) = 0.22 \times 0.36 = 0.0792.$$

For the second part, we want to calculate $P(B|A)$. From the previous calculation

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = 0.0792/0.3 = 0.264.$$

Example: Roommate Profile Revisit

- ▶ Given that a randomly chosen roommate snores, what is the probability that he/she likes to party?

| | Snores | Doesn't Snore | Total |
|---------------|--------|---------------|-------|
| Parties | 150 | 100 | 250 |
| Doesn't Party | 200 | 550 | 750 |
| Total | 350 | 650 | 1000 |

Table : Roommate Profile (Frequency - Counts)

Solution:

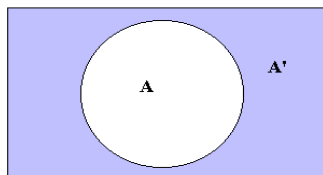
Example: Conditional Probability (cont.)

- ▶ Note that, in both examples, the conditional probability $P(B|A)$ and the unconditional probability $P(B)$ are **NOT** equal. In other words, knowledge that A “has occurred” has changed the likelihood that B occurs.
- ▶ In other situations, it might be that the occurrence (or non-occurrence) of a companion event has no effect on the probability of the event of interest. This leads us to the definition of **independence**.

- ▶ **Definition:** When $P(A|B) = P(A)$, we say that events A and B are **independent**. If A and B are not independent, they are said to be **dependent** events.
- ▶ This definition is symmetric,
 $P(A|B) = P(A) \iff P(B|A) = P(B)$.
- ▶ Under independence assumption,
 $P(A \cap B) = P(A|B)P(B) = P(A)P(B)$

Complementary Events

- ▶ The **complement** of an event is the set of the outcomes not included in the event, but still part of the sample space. The complement of A is denoted by \bar{A} or A^c or A' .
- ▶ The Venn diagram is



Complementary Events (cont.)

- ▶ By additive law $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, which implies that $1 = P(S) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$, so that

$$P(\bar{A}) = 1 - P(A)$$

- ▶ **DeMorgan's Law:** it is useful when the calculation involves two events and their complements. Let A and B be two events

$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$

- ▶ Question: Can you prove the DeMorgan's Law (formally or informally) ?

Example: Complementary Events

The probability that Tom will be alive in 20 years is 0.75 (A).
The probability that Nancy will be alive in 20 years is 0.85 (B).

- ▶ Assuming independence, what is the probability that neither will be alive 20 years from now?
- ▶ Still assuming independence, what is the probability that only one of the two, Tom or Nancy, will be alive in twenty years?
- ▶ What is the probability of the complement of both Nancy and Tom not being alive in 20 years?

All mutually exclusive events are complementary.

- ▶ Yes.
- ▶ No.

If events A and B are mutually exclusive, then A and B are independent.

- ▶ Yes.
- ▶ No.

1. **Complement rule:** Suppose that A is an event.

$$P(\bar{A}) = 1 - P(A).$$

2. **Additive law:** Suppose that A and B are two events.

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

3. **Multiplicative law:**

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A).$$

4. **Law of Total Probability:**

$$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B}).$$

5. **Bayes' rule:**

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}.$$

Applications of Probability Rules

- ▶ We purchase 30% of our parts from Vendor A. Vendor A's defective rate is 5%. What is the probability that a randomly chosen part is a defective part from Vendor A?
- ▶ We are manufacturing widgets. 50% are red and 30% are white. What is the probability that a randomly chosen widget will not be white?

Solution:

Applications of Probability Rules (cont.)

When a computer goes down, there is a 75% chance that it is due to overload and a 15% chance that it is due to a software problem. There is an 85% chance that it is due to an overload or a software problem. What is the probability that both of these problems are at fault?

Solution:

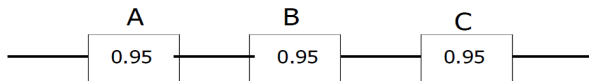
Applications of Probability Rules (cont.)

It has been found that 80% of all accidents at foundries involve human error and 40% involve equipment malfunction. 35% involve both problems. If an accident involves an equipment malfunction, what is the probability that there was also human error?

Solution:

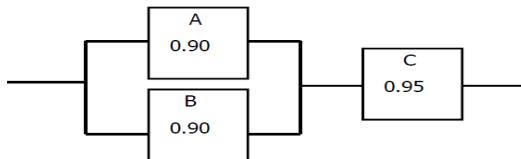
Applications of Probability Rules (cont.)

Consider the following electrical circuit: The probability on the components is their reliability (probability that they will operate). Components are independent of each other. What is the probability that the circuit **will not** operate when the switch is thrown?



Applications of Probability Rules (cont.)

Consider the electrical circuit below. Probabilities on the components are reliabilities and all components are independent. What is the probability that the circuit will work when the switch is thrown?



Applications of Probability Rules (cont.)

An insurance company classifies people as “accident-prone” and “nonaccident-prone”. For a fixed year, the probability that an accident-prone person has an accident is 0.4, and the probability that a non-accident-prone person has an accident is 0.2. The population is estimated to be 30 percent accident-prone. Define the events

$$A = \{\text{policy holder has an accident}\}$$

$$B = \{\text{policy holder is accident-prone}\}$$

We are given that $P(B) = 0.3$, $P(A|B) = 0.4$, and $P(A|\bar{B}) = 0.2$.

Applications of Probability Rules (cont.)

- ▶ What is the probability that a new policy-holder will have an accident?

- ▶ Suppose that the policy-holder does have an accident. What is the probability that s/he is “accident-prone” ?

- ▶ For equal likelihood probability model, the probability of event A depends on the size of A , i.e., the number of elements in A , which is denoted by n_A .
- ▶ Recall that if an experiment can result in any one of N different, but equally likely, outcomes and if exactly n of these outcomes corresponds to event A , then the probability of event A is

$$P(A) = \frac{n_A}{n_S} = \frac{n}{N}.$$

- ▶ In the next several slides, several counting methods will be presented, including Fundamental Theorem of Counting, Factorial, Combination, and Permutations.

Fundamental Theorem of Counting

- ▶ **Theorem:** Stated simply, it is the idea that if there are a ways of doing something and b ways of doing another thing, then there are $a \times b$ ways of performing both actions.
- ▶ **Example:** When you decide to order pizza, you must first choose the type of crust: thin or deep dish (2 choices). Next, you choose one topping: cheese, pepperoni, or sausage (3 choices). Using the fundamental theorem of counting, you know that there are $2 \times 3 = 6$ possible combinations of ordering a pizza.

Example: Fundamental Theorem of Counting

There are 5 processes needed to manufacture the side panel for a car: clean, press, cut, paint, polish. Our plant has 6 cleaning stations, 3 pressing stations, 8 cutting stations, 5 painting stations, and 8 polishing stations.

- ▶ How many possible different “pathways” through the manufacturing exist? (_____)
- ▶ What is the number of “pathways” that include a particular pressing station? (_____)
- ▶ Assuming stations are chosen randomly, what is the probability that a panel follows any particular path? (_____)
- ▶ Assuming stations are chosen randomly, what is the probability that a panel goes through pressing station 1? (_____)

Permutation: Ordered Samples

- ▶ **Permutation:** ways to arrange distinct objects into a sequence. For example, the number of **permutations** of n different elements is

$$n! \equiv n(n-1)(n-2)\dots(2)(1).$$

- ▶ $n!$ is read as **n factorial**, and $0! = 1$.
- ▶ **Theorem:** *Let there be n distinct objects. The number of possible ordered samples of size r from these n objects is*

$$P_r^n \equiv n \times (n-1) \times (n-2) \times \dots \times (n-r+1) = \frac{n!}{(n-r)!}.$$

- ▶ **Example:** Let there be $N = 10$ people and suppose that we are to sit $r = 6$ of them into 6 chairs, assume the chairs are numbered. Then the number of ways of doing this is _____

Combinations: Unordered Samples

- ▶ **Theorem:** *Let there be n distinct objects and consider taking an unordered sample of size r from these objects. Then the number of possible samples is*

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

- ▶ **Example:** Let there be 10 people and suppose we want to obtain a sample of size 6 from this group without considering the order of people, then the number of possible samples is _____.

Example: Counting Methods

9 out of 100 computer chips are defective. Chips are different. We choose a random sample of $n = 3$.

- ▶ How many different samples of 3 are possible? _____
- ▶ How many of the samples of 3 contain exactly 1 defective chip? _____
- ▶ What is the probability of choosing exactly 1 defective chip in a random sample of 3? (This is called **hypergeometric distribution** later). _____
- ▶ What is the probability of choosing at least 1 defective chip in a random sample of 3? (Hint: it is easier to calculate the prob of choosing no defective chip.) _____

Example: Coin Tossing

What is the probability of getting 4 or 5 heads when tossing a fair coin 10 times?

Solution: Let us define events

$$A = \{\text{getting 4 heads}\},$$

$$B = \{\text{getting 5 heads}\}.$$

$P(\text{getting 4 or 5 heads}) = P(A \cup B) = P(A) + P(B)$, since $A \cap B = \emptyset$.

The event A “looks” like

$$A = \{HHHHTTTTTT, HTTHTHHTTT, \dots, TTTTTHHHH\}.$$

It is easy to see that each element in A is assigned probability $(0.5^4)(0.5^6) = 0.5^{10}$. By combination formula, these are total $\binom{10}{4}$ elements in A .

Example: Coin Tossing (cont.)

It follows that $P(A) = \binom{10}{4}0.5^{10}$. By a similar argument, one can show that $P(B) = \binom{10}{5}0.5^{10}$.

Therefore,

$$P(A \cup B) = P(A) + P(B) = \binom{10}{4}0.5^{10} + \binom{10}{5}0.5^{10} = 0.4512.$$

Later, we will solve this problem using **binomial distribution**.

Birthday Problem

What is the probability that at least 2 people in this class ($n=48$) have the same birthday (Month and day)?

- ▶ Year has 365 days (forget leap year).
- ▶ Equal likelihood for each day.

Solution: This is a classical model. Let event

$$A = \{\text{at least 2 students share the same birthday.}\}$$

Then the complement of event A is

$$\bar{A} = \{\text{no student shares the same birthday.}\}$$

We know $n_S = 365^{48}$, and

$$n_{\bar{A}} = (365)(364)\dots(365 - 48 + 1) = \frac{365!}{(365 - 48)!} (= P_{48}^{365}).$$

Birthday Problem (Cont.)

- ▶ Therefore,

$$P(\bar{A}) = \frac{n_{\bar{A}}}{n_S} = \frac{365! / (365 - 48)!}{365^{48}} = \frac{365!}{(365 - 48)! 365^{48}} \approx 0.0394.$$

- ▶ $P(A) = 1 - P(\bar{A}) \approx 0.9606$. We have a surprisingly high chance!
- ▶ How to we get the 0.0394 in R? One possible code is

```
factorial(365) / ( factorial(365-48) * 365^48 )
```

- ▶ R returns an error message from the above code because 365! is way to large.

Birthday Problem (Cont.)

- ▶ A smart way calculate this probability is to use “combination” to avoid $365!$ as following:

$$\begin{aligned}P(\bar{A}) &= \frac{365!}{(365 - 48)!365^{48}} \\&= \left[\frac{365!}{(365 - 48)!48!} \right] \left(\frac{48!}{365^{48}} \right) \\&= C_{48}^{365} \left(\frac{48!}{365^{48}} \right)\end{aligned}$$

- ▶ In R, it is

```
choose(365, 48) * factorial(48) / 365^48
```