Chapter 5: Probability

Yu Yang

School of Statistics
University of Minnesota

September 19, 2022

Randomness

Random Phenomenon

A phenomenon is *random* if individual outcomes are uncertain but there is a long-term regularity in the outcomes.

Random does not mean haphazard!

Probability

We use **probability** to quantify randomness.

Probability

The *probability* of an outcome of a random phenomenon is the proportion of times the outcome occurs in a very long series of repetitions (i.e., the "long-run proportion")

Example 5.1

Use coin flipping to illustrate the idea of probability as a long-run proportion. Suppose we keep flipping the **fair** coin and keep track of the proportion of rolls that have turned up heads.

Flip	Result	Cumulative Proportion of Heads
1	T	0
2	H	0.5
3	H	**0.67**
4	T	**0.50**
5	T	**0.40**
6	T	**0.33**
⋮	⋮	⋮

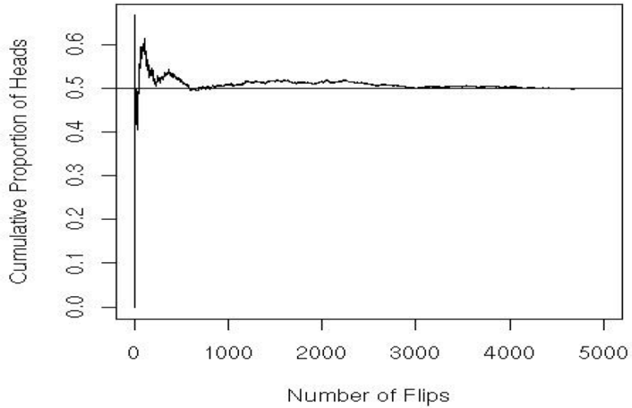
Example 5.1

And you can keep doing this for 5000 tosses, to get a table like:

Flip	Result	Cumulative Proportion of Heads
1	T	0.0000000
2	H	0.5000000
3	H	0.6666667
⋮	⋮	⋮
4998	H	0.5004002
4999	T	0.5005001
5000	H	0.5004000

We can simulate this in R and plot the cumulative proportion of heads.

Example 5.1 (cont.)



As the number of flips (x axis) increases, the proportion of heads converges to 0.5.

Probability Models

A probability model for a random phenomenon has two parts:

1. A list of all possible outcomes
2. The probability for each outcome

Sample Space and Event

Sample Space

A *sample space* is the collection of all possible outcomes of an experiment or random phenomenon. Usually denoted as S.

Event

An *event* is any **subset** of the sample space. Usually denoted as A, B, C.

Example 5.2

Sample Space (S): A sample space is the collection of all possible outcomes of an experiment or random phenomenon.

Suppose we randomly select a student from class and ask a question. Describe the sample space for these experiments.

1. How much time (in hours) did the student spend studying in the last 24 hours?

Example 5.2

Sample Space (S): A sample space is the collection of all possible outcomes of an experiment or random phenomenon.

Suppose we randomly select a student from class and ask a question. Describe the sample space for these experiments.

1. How much time (in hours) did the student spend studying in the last 24 hours?
   S = [0,24]. Event can be A = [0,2]

2. In what state was the student born if it is known that they were born in the US?

Example 5.2

Sample Space (S): A sample space is the collection of all possible outcomes of an experiment or random phenomenon.

Suppose we randomly select a student from class and ask a question. Describe the sample space for these experiments.

1. How much time (in hours) did the student spend studying in the last 24 hours?
   S = [0,24]. Event can be A = [0,2]

2. In what state was the student born if it is known that they were born in the US?
   S = {AL, AK, AZ, ..., WY}. Event can be B = {MN,NJ}

3. How many friends does the student have?

Example 5.2

Sample Space (S): A sample space is the collection of all possible outcomes of an experiment or random phenomenon.

Suppose we randomly select a student from class and ask a question. Describe the sample space for these experiments.

1. How much time (in hours) did the student spend studying in the last 24 hours?
   S = [0,24]. Event can be A = [0,2]

2. In what state was the student born if it is known that they were born in the US?
   S = {AL, AK, AZ, ..., WY}. Event can be B = {MN,NJ}

3. How many friends does the student have?
   S = { 0, 1, 2, 3, ...}. Event can be C = {2,3,4}

Example 5.3

Suppose we toss a fair coin 3 times.

1. What is the sample space, S?

Example 5.3

Suppose we toss a fair coin 3 times.

1. What is the sample space, S?
   S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

2. Let A be the event that we get **exactly** 2 heads on the 3 tosses.
   Write down A.

Example 5.3

Suppose we toss a fair coin 3 times.

1. What is the sample space, S?
   S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

2. Let A be the event that we get **exactly** 2 heads on the 3 tosses. Write down A.
   A = {HHT, HTH, THH}

3. Let B be the event that **at most** one tail is flipped. Write down B.

Example 5.3

Suppose we toss a fair coin 3 times.

1. What is the sample space, S?
   S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

2. Let A be the event that we get **exactly** 2 heads on the 3 tosses. Write down A.
   A = {HHT, HTH, THH}

3. Let B be the event that **at most** one tail is flipped. Write down B.
   B ={HHH, HHT, HTH, THH}

Special Events

- **Complimentary Event:** $A^c =$ "not A"
  $A^c =$ collection of all outcomes in S that are not in A.
  NOTE: $A \cup A^c = S$

- **Intersection:** $A \cap B =$"A and B"
  $A \cap B =$ all outcomes that are in both A and B

- **Union:** $A \cup B =$ "A or B or both"
  $A \cup B =$ all outcomes that are in either A or B (or both)

Event Relations

Disjoint

Two events A and B are *disjoint* if they have no outcomes in common.

Independent

Two events are *independent* if knowing one occurs does not change the probability that the other occurs. For example,

- The event that it rains today is independent of the event than 7 was one of the lottery numbers chosen last night.
- The event that it rains today is **not** independent of the event that it was cloudy this morning.

Example 5.4

Roll a fair die (one time)

1. Write down the sample space.

Example 5.4

Roll a fair die (one time)

1. Write down the sample space.
   S = {1, 2, 3, 4, 5, 6}

2. Let A be the event that you roll an even number and B be the event you roll a number bigger than 2. Write down the following events.
   (a) A, B:

Example 5.4

Roll a fair die (one time)

1. Write down the sample space.
   S = {1, 2, 3, 4, 5, 6}

2. Let A be the event that you roll an even number and B be the event you roll a number bigger than 2. Write down the following events.
   (a) A, B: A = {2, 4, 6}, B = {3, 4, 5, 6}
   (b) A^c:

Example 5.4

Roll a fair die (one time)

1. Write down the sample space.
   S = {1, 2, 3, 4, 5, 6}

2. Let A be the event that you roll an even number and B be the event you roll a number bigger than 2. Write down the following events.
   (a) A, B: A = {2, 4, 6}, B = {3, 4, 5, 6}
   (b) A^c: $A^c = \{1,3,5\}$
   (c) $A \cap B$:

Example 5.4

Roll a fair die (one time)

1. Write down the sample space.
   S = {1, 2, 3, 4, 5, 6}

2. Let A be the event that you roll an even number and B be the event you roll a number bigger than 2. Write down the following events.
   (a) A, B: A = {2, 4, 6}, B = {3, 4, 5, 6}
   (b) A^c: $A^c = \{1,3,5\}$
   (c) $A \cap B$: $A \cap B = \{4, 6\}$
   (d) $A \cup B$:

Example 5.4

Roll a fair die (one time)

1. Write down the sample space.
   S = {1, 2, 3, 4, 5, 6}

2. Let A be the event that you roll an even number and B be the event you roll a number bigger than 2. Write down the following events.
   (a) A, B: A = {2, 4, 6}, B = {3, 4, 5, 6}
   (b) A^c: $A^c = \{1,3,5\}$
   (c) $A \cap B$: $A \cap B = \{4, 6\}$
   (d) $A \cup B$: $A \cup B = \{2, 3, 4, 5, 6\}$
   (e) $(A \cup B)^c$:

Example 5.4

Roll a fair die (one time)

1. Write down the sample space.
   S = {1, 2, 3, 4, 5, 6}

2. Let A be the event that you roll an even number and B be the event you roll a number bigger than 2. Write down the following events.
   (a) A, B: A = {2, 4, 6}, B = {3, 4, 5, 6}
   (b) A^c: $A^c = \{1,3,5\}$
   (c) $A \cap B$: $A \cap B = \{4, 6\}$
   (d) $A \cup B$: $A \cup B = \{2, 3, 4, 5, 6\}$
   (e) $(A \cup B)^c$: $(A \cup B)^c = \{1\}$

Example 5.4 (cont.)

In the previous example, let C be the event that you roll a value less than 3 and let $A = \{2, 4, 6\}$ and $B = \{3, 4, 5, 6\}$.

- C = $\{1, 2\}$
- $A \cap B =$

Example 5.4 (cont.)

In the previous example, let C be the event that you roll a value less than 3 and let $A = \{2, 4, 6\}$ and $B = \{3, 4, 5, 6\}$.

- C = {1, 2}
- $A \cap B =$ {4, 6}
- $A \cap C =$

Example 5.4 (cont.)

In the previous example, let C be the event that you roll a value less than 3 and let $A = \{2, 4, 6\}$ and $B = \{3, 4, 5, 6\}$.

- C = {1, 2}
- $A \cap B =$ {4, 6}
- $A \cap C =$ {2}
- $B \cap C =$

Example 5.4 (cont.)

In the previous example, let C be the event that you roll a value less than 3 and let $A = \{2, 4, 6\}$ and $B = \{3, 4, 5, 6\}$.

- C = {1, 2}
- $A \cap B$ = {4, 6}
- $A \cap C$ = {2}
- $B \cap C$ = $\emptyset$ (empty set)
- Which ones are disjoint?

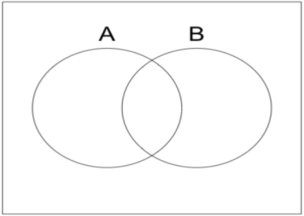
Example 5.4 (cont.)

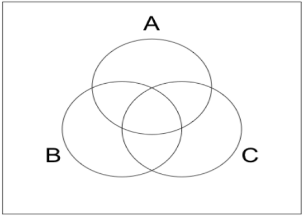
In the previous example, let C be the event that you roll a value less than 3 and let $A = \{2, 4, 6\}$ and $B = \{3, 4, 5, 6\}$.

- C = {1, 2}
- $A \cap B$ = {4, 6}
- $A \cap C$ = {2}
- $B \cap C$ = $\emptyset$ (empty set)
- Which ones are disjoint? B and C are disjoint, the rest are not.

Venn Diagrams

Venn diagrams allow us to visualize events, their intersections, and their unions.





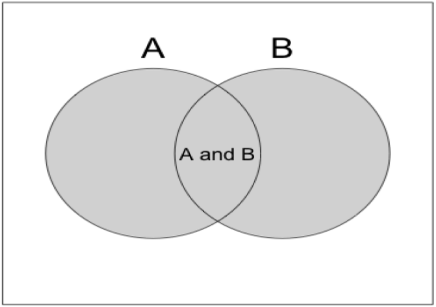
Probability Rules

Let A and B be events and let $P(A)$ and $P(B)$ denote the probabilities of these events occurring. The following statements are **always true:**

1. $0 \leq P(A) \leq 1$
   - $P(A) = 0 \implies A$ will never occur.
     Ex: Roll a die. $A =$ get a number larger than 7.
   - $P(A) = .0001 \implies A$ is very unlikely to occur but **will** occur in a long series of trials.
     Ex: Out of 10,000 tickets there is 1 golden ticket. $A =$ Charlie gets the golden ticket.
   - $P(A) = 0.6 \implies A$ will be observed more often.
     Ex: Record the temperature in the summer. $A =$ the highest temperature is above 80F.
   - $P(A) = 1 \implies A$ is certain to occur.
     Ex: Roll a die. $A =$ get a number smaller than 7.

Probability Rules

2. Law of Total Probability: $P(S) = 1$ where S is the sample space.
3. Complement Rule: $P(A^c) = 1 - P(A)$
4. General Addition Rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
5. Partitioning of Probability: $P(A) = P(A \cap B) + P(A \cap B^c)$



Probability Rules

6. If A and B are **disjoint**, then
   - $P(A \cap B) = 0$
   - $P(A \cup B) = P(A) + P(B)$

7. If A and B are **independent**, then $P(A \cap B) = P(A) \times P(B)$.

Applying Probability Rules - Example 1

A survey of students found that in the last month:

- 68% had gone to see a movie (A)
- 52% had attended a sporting event (B)
- 35% had done both ($A \cap B$)

(a) Draw a Venn diagram of these events.

(b) What is the probability that a randomly selected student has been to either a movie or a sporting event (or both) in the last month?
$P(A \cup B) = P(A) + P(B) - P(A \cap B) = .68 + .52 - .35 = .85$

(c) What is the probability that a randomly selected student has been to a movie but **not** a sporting event in the last month?
$P(A \cap B^c) = P(A) - P(A \cap B) = 0.68 - 0.35 = 0.33.$

(d) What is the probability that a randomly selected student has been to neither a movie nor a sporting event in the last month?
$P((A \cup B)^c) = 1 - P(A \cup B) = 1 - .85 = .15.$

Applying Probability Rules - Example 2

SurveyUSA polled 451 Americans regarding their opinion on federal gun control laws:

		Opinion				
		Too Restrictive	Not Restrictive Enough	About Right	Not Sure	Total
Age	18–34	31	67	49	6	153
	35–54	36	82	59	3	180
	55+	21	60	33	4	118
	Total	88	209	141	13	451

Applying Probability Rules - Example 2 (cont.)

Select one person at random from the sample and define events A and B as follows:

- $A =$ they think that federal gun control laws are too restrictive.
- $B =$ they are under the age of 55.

(a) Find $P(A)$ and $P(B)$.
$P(A) = 88/451 = 0.195$
$P(B) = (153 + 180)/451 = 0.738.$

(b) What is the probability that the person does **not** think federal gun control laws are too restrictive.
$P(A^c) = 1 - P(A) = 1 - 0.195 = 0.805.$

(c) What is the probability that the person is under the age of 55 **and** thinks federal gun control laws are too restrictive?
$P(A \cap B) = (31 + 36)/451 = 0.149.$

Applying Probability Rules - Example 3

According to an organization called Student Monitor, 83% of American college students own a laptop, 24% own a desktop, and 8% own neither a laptop nor a desktop.

(a) What is the probability that a randomly selected student owns either a laptop or a desktop or both?
$P(\text{laptop or desktop}) = 1 - P(\text{neither}) = 0.92$

(b) What is the probability that a randomly selected student owns both a desktop and a laptop?
$P(\text{laptop or desktop}) = P(\text{laptop}) + P(\text{desktop}) - P(\text{both})$.
$\Rightarrow P(\text{both}) = P(\text{laptop}) + P(\text{desktop}) - P(\text{laptop or desktop}) = 0.83 + 0.24 - 0.92 = 0.15$

(c) Are owning a desktop and owning a laptop independent for the population of American students?
$P(\text{laptop}) \times P(\text{desktop}) = 0.83 \times 0.24 = 0.1992 \neq 0.15 = P(\text{both})$.
Therefore they are not independent.

Conditional Probability: Motivating Example

Suppose I take a handful of M&Ms and get 5 red, 4 blue and 7 brown. I randomly pick one of the M&Ms and eat it. We can see that:

$$P(red) = 5/16, \; P(blue) = 4/16, \text{ and } P(brown) = 7/16$$

Now, suppose I tell you that the M&M was **not** brown. What is the probability that the M&M was blue *given that (or knowing that) it was not brown*.

Since it was not brown, that leaves 9 M&Ms: 4 blue and 5 red. I could have chosen any of the 4 blue out of the 9. So

$$P(\text{blue given not brown}) = 4/9$$

Conditional Probability

The *conditional probability* of event A *given* event B is the probability that A occurs given the knowledge that B has already occurred. When $P(B) > 0$, the conditional probability of A given B is:

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

Back to M&Ms

We can apply this formula to the M&M example. Let

- $A =$ the M&M is blue
- $B =$ the M&M is not brown.

We see that

$$P(\text{blue}|\text{not brown}) = \frac{P(\text{blue} \cap \text{not brown})}{P(\text{not brown})} = \frac{\frac{4}{16}}{1 - \frac{7}{16}} = \frac{4}{9}.$$

Another Multiplication Rule

By the definition of condition probability,

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

Consequently,

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

Conditional Probability and Independence

Recall that if A and B are independent, knowing that B has happened does **not** change the probability that A also occurs. Therefore, when A and B are independent,

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

Three Ways to Check Independence

A and B are independent if **any** of the following holds

1. $P(A \cap B) = P(A)P(B)$
2. $P(A|B) = P(A)$
3. $P(B|A) = P(B)$

These three statements are equivalent. If any of these conditions do not hold, then none of them hold and the events are **NOT** independent.

Example 5.5

According to an organization called Student Monitor, 83% of American college students own a laptop, 24% own a desktop, and 8% own neither a laptop nor a desktop. Define events as follows.

L = own a laptop

D = own a desktop

Example 5.5 (cont.)

(a) What is the probability that a student with a desktop also owns a laptop?
$P(L|D) = \frac{P(L \cap D)}{P(D)} = \frac{.15}{.24} = .625.$
Interpretation: 62.5% of desktop owners also owns a laptop.

(b) Find the conditional probability that a student owns desktop given that they own a laptop.
$P(D|L) = \frac{P(L \cap D)}{P(L)} = \frac{.15}{.83} = .18.$
Interpretation: 18% of students with a laptop also owns a desktop.

(c) Find the conditional probability that a student does NOT own a desktop given that the student owns a laptop.
$P(D^c|L) = 1 - P(D|L) = 1 - .18 = .82.$

(d) Are owning a desktop and owning a laptop independent?
$P(D|L) \neq P(D) \Rightarrow$ not independent.

Intersection vs. Conditional probability

In Example 5.5,

- what is the proportion of student own both a laptop **and** a desktop?
  Intersection: $P(L \cap D)$.

- what is the proportion of students who own a laptop also own a desktop?
  Conditional probability: $P(D|L)$.