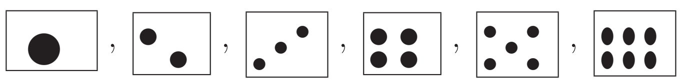
2 Probability

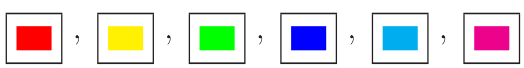
2.1 Sample Spaces and Events

Definition: The **sample space** of an experiment, denoted by $\mathcal{S}$, is the set of all possible outcomes of that experiment.

Example: Roll a (not necessarily fair) six-sided die once. The possible outcomes are the **faces** (i.e., dots) on the die.

$\mathcal{S} = \{$  $\}$

A different die might have six colors for the six sides (like a Rubik's cube).

$\mathcal{S} = \{$  $\}$

□

Example: Toss a (not necessarily fair) coin once.

Example: Number of people living in a household.

Definition: An **event** is a subset of the sample space.

Example: For a (six-sided, numeric) die roll, name an **event**.

Additional definitions for events A and B:

(a) **union:** A or B, $A \bigcup B$

(b) **intersection:** A and B, $A \bigcap B$, AB

(c) **complement:** A'

(d) **mutually exclusive** or **disjoint:** $A \bigcap B = \emptyset$

A **Venn diagram** is a graphical representation of events.

Example: Let $\mathcal{S}$ be the sample space of all Harrisonburg residents. let A ={Person is female}, and let B ={Person has high school diploma}.

Example: Let $\mathcal{S}$ be the sample space of all JMU students; let A ={Person is a freshman}, and let B ={Person is a sophomore}.

Exercise 2.9, p. 58:

2.2 Axioms, Interpretations, and Properties of Probabilities

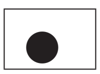
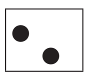
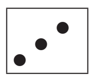
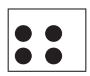
Axiom 1: For any event A, $P(A) \geq 0$.

Axiom 2: $P(\mathcal{S}) = 1$.

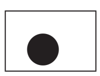
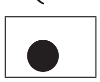
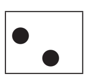
Axiom 3: For A_1, A_2, A_3, $\ldots$ is an infinite collection of **disjoint** events, then

$$P(A_1 \bigcup A_2 \bigcup A_3 \bigcup \cdots) = \sum_{i=1}^{\infty} P(A_i).$$

Example: Consider **one** roll of any (unfair) **four**-sided die.

$\mathcal{S} = \Big\{$  ,  ,  ,  $\Big\}$

Remark: $0 \leq P(A) \leq 1$, and $P(A) \leq P(B)$ if $A \subseteq B$.

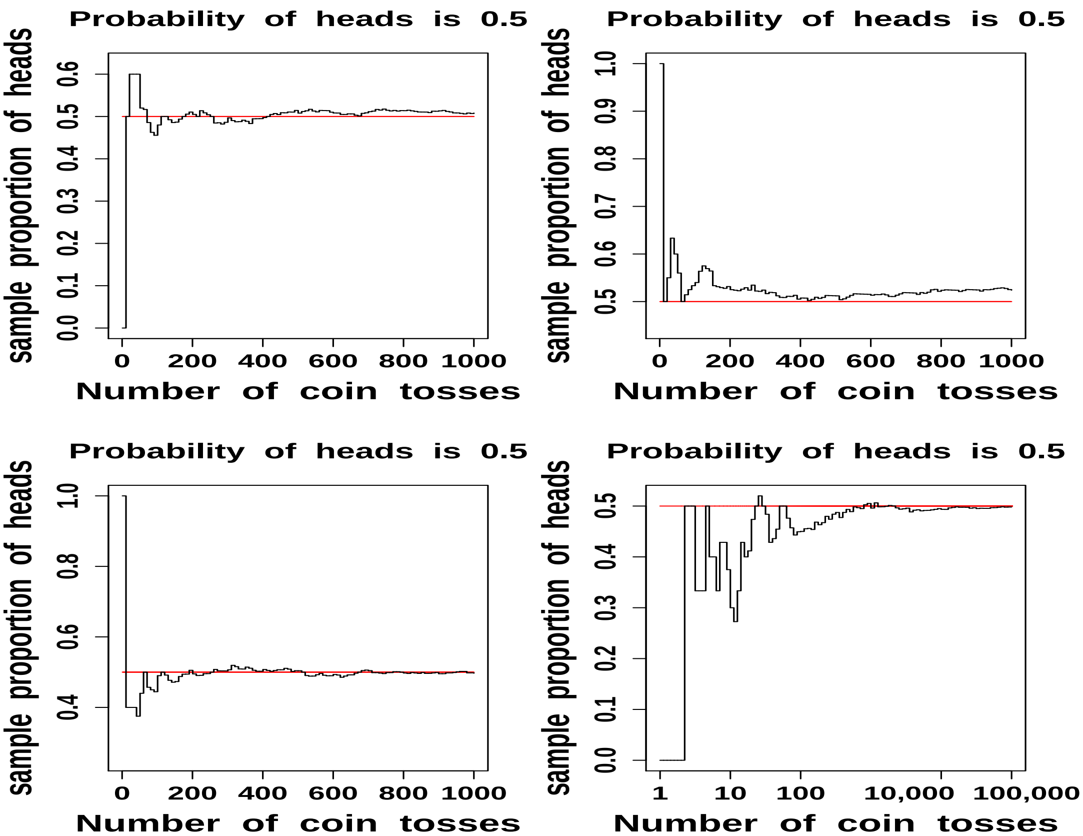
Example: Consider **one** roll of any (unfair) **four**-sided die.
Let $A = \Big\{$  $\Big\}$, and let
$B = \Big\{$  $\bigcup$  $\Big\}$.

□

Interpreting Probability

What does it mean when we say that P(heads)=0.5 for a coin? *Alternative question:* How do we find P(heads) for a coin?

The graphs below represent the sample proportion of heads, in tosses of a fair coin, for a large number of tosses.



More rigorous definition of *probability*, based on a *long run proportion*: A **probability** of an outcome is the proportion of times that the outcome is expected to occur when the experiment is repeated many times under identical conditions.

A **law of large numbers** states that a sample proportion, $\hat{p}$, "gets close to" the population proportion or probability, over a long run.

Example: What happens to the sample proportion of heads after many tosses of the coin?

Example: Consider a club consisting of 6 females and 4 males, and selections of club members are made WITH replacement. What happens to the sample proportion of females after many selections (WITH replacement)?

More Probability Properties

Same notation: Consider the sample space $\mathcal{S}$, and events A, B, A_1, A_2, $\ldots$.

Property: $P(A') = 1 - P(A)$

Proof:

Recall: If A and B are mutually exclusive, then $P(A \bigcap B) = 0$.

Example: Consider **one** roll of any (unfair, numeric) **four**-sided die.

Property: *Addition rule (in general).* $P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$

Example: Suppose that in a city, 72% of the population got the flu shot, 33% of the population got the flu, and 18% of the population got the flu shot and the flu. Determine P(flu shot **or** flu).

$\square$

Exercise 2.12, p. 65: *Slightly modified.* Consider randomly selecting a student at a certain university, and let A denote the event that the selected individual has a Visa credit card and B be the analogous event for a MasterCard.

Suppose that $P(A) = 0.5$, $P(B) = 0.4$, and $P(A \bigcap B) = 0.22$.

The textbook uses 0.25 instead of 0.22.

(a) Compute the probability that the selected individual has at least one of the two types of cards.

(b) Compute the probability that the selected individual has neither type of card.

(c) Describe, in terms of A and B, the event that the selected student has a Visa card but not a MasterCard, and then calculate the probability of this event.

□

Equally Likely Outcomes

If a sample space consists of N elementary outcomes such that each outcome is equally likely to be selected, then the probability of any particular outcome occurring is ____________________.

Example: Fair (*unbiased*) six-sided die.

Example: Fair (*unbiased*) coin.

Example: When rolling a fair die once, determine $P\big($ ⚀ $\cup$ ⚁ $\big)$.

□

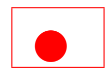
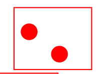
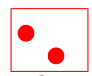
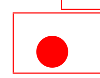
2.3 Counting Techniques

Example: Suppose we roll a 3-sided red die once, and a 4-sided blue die once.

(a) List all possible pairs of outcomes.

(b) Determine the total number of pairs of outcomes.

(c) Assuming that the dice are fair, determine $P(\text{sum} = 4)$.

(d) Determine the total number of possible outcomes (where order matters) if we roll the 3-sided red die **five** times and the 4-sided blue die **two** times.
"Where order matters" implies that rolling a  *and a*  *on the first and second rolls differs from rolling a*  *and a*  *on the first and second rolls of the red die, respectively.*

Example: How many different types of pizzas can be made if **7** toppings are available?

Example: Suppose a club consists of **10** members. How many ways can we select one president, one chef, and one custodian, such that individuals **may** have more than one title?

Permutations

Example: Suppose a club consists of **10** members, and members may **NOT** hold more than one office.

(a) How many ways can we select one president, one vice-president, and one secretary?

(b) If Larry, Moe, and Curly are in this club, and officers are selected at random, what is the probability that Larry is president, Moe is vice-president, and Curly is secretary?

□

Determine 4!, 5!, 1!, and 0!.

In general, $P_{k,n} = \frac{n!}{(n-k)!}$

Note: With permutations, order **is** relevant.

Combinations

Example: Suppose a club consists of **10** members. How many ways can we select **3** members to serve on the charity committee?

Note: A committee consisting of persons (A, B, C) is identical to (A, C, B), (B, A, C), (B, C, A), (C, A, B), and (C, B, A).

In general, $\binom{n}{k} = \frac{n!}{k!\ (n-k)!}$

How many ways can we select **7** members **NOT** to serve on the charity committee?

Exercise 2.44, p. 75: Show that $\binom{n}{k} = \binom{n}{n-k}$. Give an interpretation involving subsets.

□

Example: Prove that **combinations** can be generated from Pascal's triangle. For example, $\binom{5}{0} = 1$, $\binom{5}{1} = 5$, $\binom{5}{2} = 10$, $\binom{5}{3} = 10$, $\binom{5}{4} = 5$, and $\binom{5}{5} = 1$.

$$
\begin{array}{ccccccccccccccc}
 & & & & & & & 1 & & & & & & & \\
 & & & & & & 1 & & 1 & & & & & & \\
 & & & & & 1 & & 2 & & 1 & & & & & \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots
\end{array}
$$

Proof (by induction):

□

Example: ♠♡♣♢ Assume a shuffled standard deck of 52 cards.

(a) Determine the number of ways **5** cards can be drawn **WITH** replacement.

(b) Determine the number of ways **5** cards can be drawn **withOUT** replacement, where order is **irrelevant**.

(c) Determine the probability of drawing **all diamonds** in a 5-card draw.

(d) Determine the probability of drawing a **flush** in a 5-card draw.

(e) Determine the probability of drawing a **royal flush** in a 5-card draw.

□

2.4 Conditional Probability

Example: (fictitious) Suppose that in Rhode Island, there are 100,000 college students. Among these 100,000 Rhode Island students, 10,000 attend (fictitious) Laplace-Fisher University (which has no out-of-state students) and exactly half of the Rhode Island students are female. Among the 10,000 Laplace-Fisher students, 6,000 are female. Define the events F and L by $F = \{\text{Student is female}\}$. and $L = \{\text{Student attends Laplace-Fisher University}\}$

(a) Determine the probability that a randomly selected Rhode Island college student attends Laplace-Fisher University. In other words, what proportion of Rhode Island college students attend Laplace-Fisher University?

(b) Determine the probability that a randomly selected Rhode Island college student attends Laplace-Fisher University AND is female.
In other words, what proportion of Rhode Island college students attend Laplace-Fisher University AND are female?

(c) Determine the probability that a randomly selected Laplace-Fisher University student is female.
In other words, determine the probability that a randomly selected Rhode Island college student is female, **given** that the student attends Laplace-Fisher University.
In other words, determine the probability that a randomly selected Rhode Island college student is female, **conditional** that the student attends Laplace-Fisher University.
In other words, what proportion of Laplace-Fisher students are female?

□

Definition: For events A and B, the **conditional probability** of event A, given that event B has occurred, is

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

This above definition is *always true*, regardless of whether or not A and B are independent (to be defined in section 2.5).

Multiplication rule for conditional probabilities (*always true*):
$P(A \cap B) = P(A|B)P(B)$ if $P(B) > 0$. Similarly, $P(A \cap B) = P(B|A)P(A)$ if $P(A) > 0$.

Example: ♠♡♣♢ In a standard deck of 52 shuffled cards, determine the probability that the top card and bottom card are both diamonds. *Notation:* $T = \{\text{Top card is a diamond}\}$ and $B = \{\text{Bottom card is a diamond}\}$.

The Law of Total Probability

$P(B) = P(B \bigcap A) + P(B \bigcap A') = P(B|A)P(A) + P(B|A')P(A')$

Example: Define the events
$G = \{\text{Individual has GENE linked to cancer}\}$, and
$C = \{\text{Individual gets CANCER}\}$. In a certain community, suppose that
20% of the people have this gene linked to cancer,
70% of the people WITH this gene will get cancer, and
10% of the people withOUT this gene will get cancer.
Determine the probability that a randomly selected individual will get cancer.

□

In general, we have equation (2.5): Let $A_1, A_2, \ldots, A_k$ be mutually exclusive such that $\cup_{i=1}^{k} A_i = \mathcal{S}$, and let $P(B) > 0$. Then,

$$P(B) = \sum_{i=1}^{k} P(B \cap A_i) = \sum_{i=1}^{k} P(B|A_i)P(A_i).$$

Bayes' Theorem

Example: Define the events

$D = \{\text{Athlete uses DRUGS}\}$, and

$T = \{\text{Athlete TESTS positive for drugs}\}$.

For the athletes in a particular country, suppose that **10%** of the athletes use drugs, **60%** of the drug-users test positive for drugs, and **1%** of the non-drug-users test positive for drugs.

(a) State the probabilities in terms of the notation D and T.

(b) What proportion of the athletes would test **positive** for drugs? Alternatively, what is the probability that a randomly selected athlete would test **positive** for drugs?

(c) What proportion of the athletes would test **negative** for drugs? Alternatively, what is the probability that a randomly selected athlete would test **negative** for drugs?

(d) What proportion of the **drug-users** would test **negative** for drugs? Alternatively, what is the probability that a randomly selected **drug-user** would test **negative** for drugs?

(e) Given that an athlete tested **positive** for drugs, what is the likelihood that the athlete actually **uses** drugs?

(f) Using a *tree diagram*, repeat part (e).

(g) Given that an athlete tested **positive** for drugs, what is the likelihood that the athlete does **NOT use** drugs?

In general, **Bayes' theorem** is the following:

Let $A_1, A_2, \ldots, A_k$ be mutually exclusive such that $\cup_{i=1}^k A_i = \mathcal{S}$. Then,

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^k P(B|A_i)P(A_i)},$$

for $j = 1, \ldots, k$.

Example: *Let's Make a Deal!* In this game, the contestant selects **one** door from among **three** doors. Behind one of the three doors is a car, and behind the other two doors are goats. Suppose the contestant selects door **#1**. Next, Monty Hall (the game-show host) opens a **different** door to **intentionally show a goat,** and then offers the contestant the choice to switch doors.

Hence, if Monty Hall opens, say, door **#2**, should the contestant switch to door **#3** or keep door **#1**?

Define the events

$C_i = \{\text{Car is behind door } \#i\}$, and

$D_i = \{\text{Monty opens door } \#i\}$, for $i = 1, 2, 3$.

(a) Determine the **prior** probabilities.

(b) Determine the probability that Monty Hall opens door **#2**.

(c) Determine the **posterior** probability that the car is behind door **#1**, given that Monty Hall opens door **#2**.

(d) Determine the **posterior** probability that the car is behind door **#2**, given that Monty Hall opens door **#2**.

(e) Determine the **posterior** probability that the car is behind door **#3**, given that Monty Hall opens door **#2**.

(f) Recall that the contestant originally selected door **#1**. If Monty Hall opens door **#2** (to intentionally show a goat), should the contestant switch to door **#3** or keep door **#1**?

(g) Rework parts (b), (c), (d), and (e) using a tree diagram.

(h) Now suppose there are 100 doors (containing 1 car and 99 goats). You select door #1, and Monty Hall opens 98 of the remaining 99 doors intentionally showing only goats.

What is the likelihood that the car is behind door #1?

What is the likelihood that the car is behind the only other closed door?

Should you switch doors?

□

On exams you may use the tree diagram, or you may use the formulas directly. It is your choice!

2.5 Independence

Definition: Different trials of a random phenomenon are **independent** if the outcome of any one trial is not affected by the outcome of any other trial.

Example:

Definition: Trials which are not independent are called **dependent**.

Example:

Example: **Tree diagram** – Use a tree diagram to show the resulting **pairs** when a fair coin is tossed once and a fair **four**-sided die is rolled once.

Multiplication rule for independent events: If outcomes A and B are **independent**, then $P(A \cap B) = P(A)P(B)$.

Example: Suppose that in a city, 72% of the population got the flu shot, 33% of the population got the flu, and 18% of the population got the flu shot and the flu. Are the events {Person gets the flu shot} and {Person gets the flu} independent?

Example: *(Know this example.)*

(a) What is the likelihood of any particular airplane engine failing during flight?

(b) If failure status of an engine is *independent* of all other engines, what is the likelihood that all three engines fail in a three-engine plane?

(c) What happened to a three-engine jet from Miami headed to Nassau, Bahamas (Eastern Air Lines, Flight 855, May 5, 1983)?

□

Definition: (Independence in terms of conditional probabilities.) The following four statements are equivalent.

(a) Events A and B are **independent**.

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

(c) $P(A|B) = P(A)$ if $P(B) > 0$

(d) $P(B|A) = P(B)$ if $P(A) > 0$

Example: According to the American Red Cross, Greensboro Chapter, 42% of Americans have type A blood, 85% of Americans have the Rh factor (positive), and 35.7% of Americans have type $A+$ blood. Are the events {Person has type A blood} and {Person has Rh factor} independent?

□

Example: *Revisit earlier example.* ♠♡♣♢ In a standard deck of 52 shuffled cards, let $T =$ {Top card is a diamond} and $B =$ {Bottom card is a diamond}. Determine the following probabilities and discuss whether or not T and B are independent.

(a) $P(B)$

(b) $P(B|T)$

(c) $P(T)$

(d) $P(T|B)$

(e) $P(T \cap B)$

(f) $P(T)P(B)$

□

Example: ♠♡♣♢ In a standard deck of 52 shuffled cards, let
$F =$ {FIRST card drawn is a diamond} and let

$S = \{\text{SECOND card drawn is a diamond}\}$.

(a) Determine the probability that the first **two** cards drawn are diamonds, if the cards are drawn **WITH** replacement.

(b) Determine the probability that the first **two** cards drawn are diamonds, if the cards are drawn **withOUT** replacement.

(c) Determine the probability that the **second** card drawn is a diamond, if the cards are drawn **WITH** replacement.

(d) Determine the probability that the **second** card drawn is a diamond, if the cards are drawn **withOUT** replacement.

$\square$

Example: A fair coin is tossed twice. Let $H_1 = \{\text{First toss is heads}\}$, and $H_2 = \{\text{Second toss is heads}\}$. Determine $P(H_2|H_1)$.

$\square$

Remark: Events A_1, A_2, …, A_n can be **mutually independent** (p. 79).