6.1 Probability of an event

One of the primary applications of counting is to calculate probabilities of random events. Probability theory plays an important role in almost every area of science, economics, and business. In computer science, randomization and probability are ubiquitous. Here are a few examples:

- Algorithms can be designed to make random choices in order to avoid actions that correlate badly with input data. Probability theory is used to analyze the behavior and running time of randomized algorithms.
- Computer systems are often simulated and tested using randomized models for user behavior. For example, a network designer would use probability theory to design a network that avoids congestion under assumptions about the distribution of messages that need to be sent across the network.
- A computer chip manufacturer needs to predict the likelihood that a chip contains a defect depending on different features of the fabrication environment in which it is made. Probability theory would be used to analyze different environments and predict the likelihood that a chip has a defect.
- Probability is one of the primary tools in machine learning in which computers are used to discover patterns and rules in large sets of data. Bayes' Theorem about conditional probabilities provides a way to reason formally about the likelihood of different causes given a set of observed data.

A comprehensive introduction to probability theory and its role in computer science is beyond the scope of this material. Instead, the focus here is on how counting is used in probability.

Consider a simple example in which a red and a blue die are thrown. A player in a game of dice may be interested in different types of events that can occur as a result of the throw. For example, she may be interested in the likelihood that both dice come up with an even number or that the sum of the numbers on the dice is at least eight. The simple example of throwing a pair of dice contains many of the basic ingredients studied in probability.

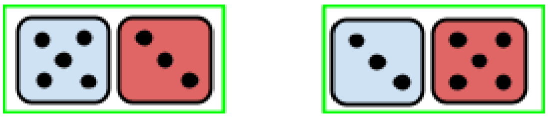
Definition 6.1.1: Experiments and outcomes.

An ***experiment*** is a procedure that results in one out of a number of possible ***outcomes***. The set of all possible outcomes is called the ***sample space*** of the experiment. A subset of the sample space is called an ***event***.

In the case of the red and blue die, the experiment is the process of throwing the dice. An outcome of a roll of the two dice is the number that shows up on the blue die and the number that shows up on the red die. The picture below shows two different outcomes of a roll of the dice:

Figure 6.1.1: Two possible outcomes of a red

and blue die.



For simplicity, we describe an outcome by an ordered pair where the first number denotes the outcome of the blue die and the second number denotes the outcome of the red die. The outcome on the left is denoted (5, 3) and the outcome on the right is denoted (3, 5). The set of outcomes for each die is {1, 2, 3, 4, 5, 6}, and the sample space for the two dice is

$$S = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$$

An event is a subset of S. For example, the event E that the sum of the dice is exactly 8 is the following set:

$$E = \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}$$

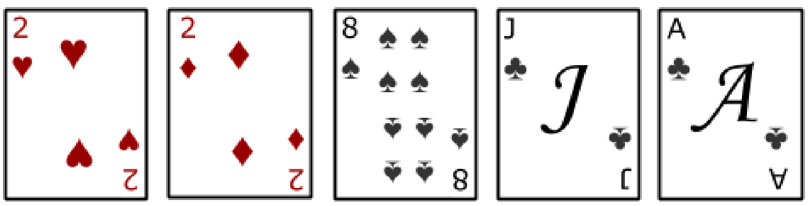
Standard playing cards

Another common example of an experiment is a card dealer who deals a hand with five cards to a player from a perfectly shuffled deck. A standard deck of playing cards consists of 52 cards. Each card has a rank and a suit. There are 13 possible ranks and 4 possible suits. Any rank-suit combination is possible, resulting in $13 \cdot 4 = 52$ different cards. The different ranks and suits are listed below:

- Ranks: 2, 3, 4, 5, 6, 7, 8, 9, 10, J (for jack), Q (for queen), K (for king), A (for ace).
- Suits: ♠ (spades), ♣ (clubs), ♥ (hearts), ♦ (diamonds).

The "8 of hearts", for example, is denoted 8♥. The sample space of 5-card hands consists of the set of all 5-subsets of the 52 cards. The order in which the cards are dealt is not important, just the set of cards in the hand after the cards have been dealt. For example, one possible outcome of the experiment of dealing a 5-card hand would be:

Figure 6.1.2: A 5-card hand from a standard card deck.



= {2♥, 2♦, 8♠, J♣, A♣}

The number of different outcomes from dealing a 5-card hand is the number of 5-subsets of the 52 cards: $\binom{52}{5}$. The event E_{jack} that the hand has four jacks would be the set of all hands of the form {J♠, J♣, J♥, J♦, *}, where the "*" could be any of the other 48 cards that are not jacks. The number of outcomes in E_{jack} is 48.

Discrete vs. continuous probability

Discrete probability is concerned with experiments in which the sample space is a finite or countably infinite set. Almost all of the experiments analyzed in this material have finite sample spaces. A set is ***countably infinite*** if there is a one-to-one correspondence between the elements of the set and the integers. A set that is not countably infinite is said to be ***uncountably infinite***. A formal discussion of the different kinds of infinity is beyond the scope of this material, but here are a couple of examples to give an intuitive feel for the difference between countably infinite and uncountably infinite sets.

Examples of countably infinite sets include the set of all binary strings (of any length), the set of ordered pairs of integers ($\mathbf{Z} \times \mathbf{Z}$), the set of all rational numbers.

The set of real numbers is an example of an uncountably infinite set. In fact the set of real numbers in a finite interval (for example the set of real numbers from 0 to 1) is also uncountably infinite. Here is an example of an experiment whose outcome is not countably infinite: a dart is thrown at a one meter square target. The point at which the dart lands on the target is the outcome of the experiment. The sample space of all possible outcomes is best modeled by pairs (x, y), where x and y are real numbers in the range from 0 to 1.

PARTICIPATION ACTIVITY | 6.1.1: Experiments, outcomes and events.

1) In the experiment of a 5-card hand, is the following an outcome or an event?

{4♣, 6♠, 7♥, Q♦, K♦}

- outcome
- event

2) In the experiment of a 5-card hand, is the following an outcome or an event?

{{4♣, 6♠, 7♥, Q♦, K♦}, {4♠, 9♥, 9♣, 9♦, K♦}}

- outcome
- event

3) In the example of a red and blue die that are thrown, define the event E to

be that the number on both dice are multiples of 3. Which set corresponds to E?

- (3, 6)
- {(3, 3), (6, 6)}
- {(3, 3), (3, 6),(6, 3), (6, 6)}

4) Suppose a coin is flipped three times. The outcome of the experiment is the sequence of outcomes from each flip. For example, HHH denotes the outcome in which the coin comes up heads in each flip. How many distinct outcomes are there?

- 8
- 2
- 4

5) In the experiment where the coin is flipped three times, which set corresponds to the event that at least two of the three flips come up heads?

- {HHH, HHT}
- {HHH, HHT, HTH, THH}
- {HHH, TTH, THT, HTT}

A natural question to ask about an experiment is: what is the likelihood (or probability) that a particular outcome occurs?

Definition 6.1.2: Probability distributions.

A ***probability distribution*** over the outcomes of an experiment with a countable sample space S is a function p from S to the set of real numbers in the interval from 0 to 1 with the property that

$$\sum_{s \in S} p(s) = 1.$$

If $E \subseteq S$ is an event, then the ***probability of event E*** is

$$p(E) = \sum_{s \in E} p(s).$$

Consider a dishonest dice player who shows up to a game with a loaded die. The player's die is biased so that an outcome of 6 is twice as likely to occur as the other numbers. Define an experiment which is a roll of the single die. The sample space is {1, 2, 3, 4, 5, 6}, and the probability distribution over the outcomes is defined by:

$$p(1) = p(2) = p(3) = p(4) = p(5) = \frac{1}{7}p(6) = \frac{2}{7}$$

The event that the die comes up an even number is {2, 4, 6}. The probability of the event that the die comes up even is:

p(2) + p(4) + p(6) = 1/7 + 1/7 + 2/7 = 4/7.

PARTICIPATION ACTIVITY | 6.1.2: Probability distributions - a loaded die.

1) In the loaded die example above, what is the probability of the event that the number on the die is 5 or 6?

Check Show answer

The uniform distribution

In many scenarios, the probability of every outcome in the sample space is the same. The probability distribution in which every outcome has the same probability is called the ***uniform distribution***. Since there are |S| outcomes in sample space S and their probabilities sum to 1, under the uniform distribution, for each s ∊ S, p(s) = 1/|S|. The uniform distribution reduces questions about probabilities to questions about counting because for every event E,

$$p(E) = \frac{|E|}{|S|}$$

Returning to the experiment with the red and blue die, if the dice are fair, then each of the 36 outcomes is equally likely. The event that the sum of the two numbers is 8 is the set

E = {(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)}.

The probability of the event E is |E|/|S| = 5/36.

Example 6.1.1: Three flips of a fair coin.

An experiment consists of three consecutive flips of a coin. The outcome is the sequence of outcomes of the three flips. There are eight possible outcomes in the sample space which are listed below. "H" stands for heads and "T" stands for tails:

HHH HHT HTH HTT THH THT TTH TTT

If the coin is a fair coin, then all eight outcomes are equally likely. The event that all three flips come out the same is E = {HHH, TTT}. The probability that all three flips come out the same is:

$$p(E) = \frac{|E|}{|S|} = \frac{2}{8} = \frac{1}{4}$$

PARTICIPATION ACTIVITY 6.1.3: Probability of events under a uniform distribution.

Give your answer as an integer or a fraction, simplified to its lowest terms, (e.g., 3/4 instead of 9/12).

1) In an experiment consisting of three flips of a fair coin, what is the probability that the first two flips are both heads?

Check Show answer

2) In an experiment consisting of three flips of a fair coin, what is the probability that the first two flips are the same?

Check Show answer

3) In an experiment consisting of a roll of a red and blue die, what is the probability that the red die is one more than the blue die?

Check Show answer

CHALLENGE ACTIVITY 6.1.1: Probability of an event.

Start

Two dice are rolled. Enter the size of the set that corresponds to the event that both dice are the same number.

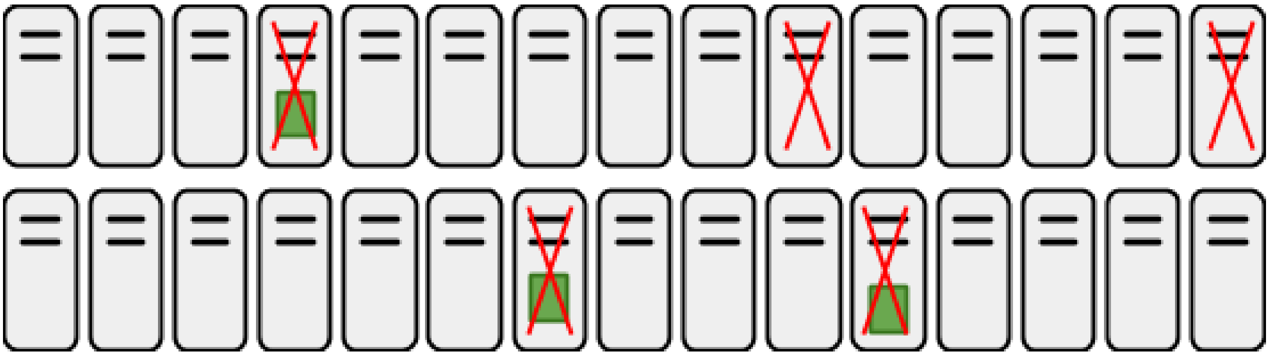
Ex: 8

1 2 3

Check Next

Example 6.1.2: Redundant data storage.

Consider a situation in which files are stored on a distributed network. Multiple copies of each file are stored around the network so that if one or more computers crash, the data is more likely to be available from at least one source. Suppose that three copies of a file are stored at different locations in a network of 30 computers and that at a particular moment, five random computers fail. Each subset of five computers are equally likely to be the five that have failed. What is the probability that there are no copies left of the file?

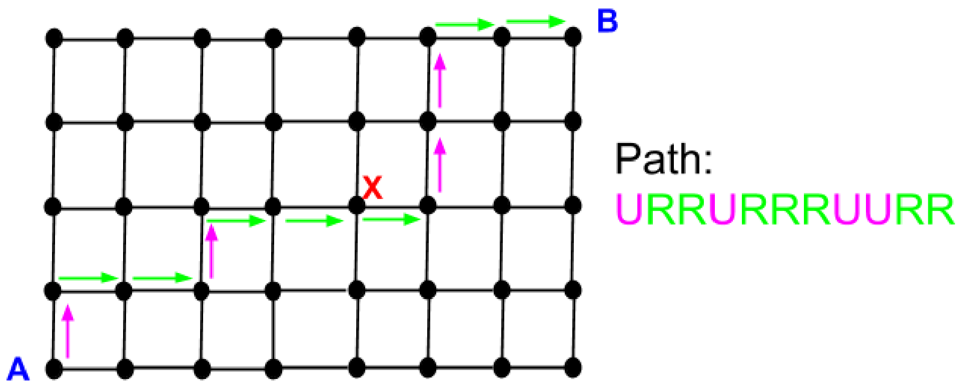


The experiment is that a set of five computers in the network fail. An outcome of the experiment is a particular set of five failing computers. Thus, the number of distinct outcomes is $\binom{30}{5}$. We assume the uniform distribution in which each of the possible 5-sets of computers is equally likely to fail. The file is stored at three particular computers in the network, and the event in question is that the set of five failed computers includes all three of the computers storing the file. How many outcomes are in the event? For every outcome in the event, the set of five failures includes the three that store the file and the remaining two failures are selected from the 27 computers that do not store the file. Thus, the number of outcomes in the event is $\binom{27}{2}$. The probability that there are no remaining computers storing the file after a random set of five computers fail is:

$$\frac{|E|}{|S|} = \frac{\binom{27}{2}}{\binom{30}{5}} \approx .002463$$

The next example explores a situation in which messages are routed in a grid network. Every vertex in the graph is a computer that routes messages through the network. To better understand traffic and congestion, we consider only the traffic generated by messages going from point A to point B. One reason to calculate congestion would be to put more powerful routers at network vertices that must handle more traffic. Thus, we would like to determine the probability that a randomly chosen path from A to B goes through the vertex marked X. We consider only the paths that go directly from A to B by a series of transitions upwards and to the right. The path never moves downward or to the left. The diagram below shows one such path.

Figure 6.1.3: Paths in a grid network.



A direct path from A to B crosses eleven edges, seven of them are moves to the right and four of them are moves upwards. The experiment is the process of selecting a direct path from A to B uniformly at random so that all such paths are equally likely. The number of direct paths from A to B is $\binom{11}{7}$. The outcome of the experiment is the path selected, so the size of the sample space is $\binom{11}{7}$. The animation below illustrates how to determine the number of outcomes in the event E that the chosen path passes through vertex X.

PARTICIPATION ACTIVITY 6.1.4: Calculating congestion in a grid network.

Animation captions:

1. Select a direct path from A to B at random. S is the set of directed paths from A to B. Each path in S has 11 edges: 4 up and 7 right. $|S| = \binom{11}{7}$.
2. E is the set of paths that go through a point X. $|E| = (\text{\# paths from A to X})(\text{\# paths from X to B})$.
3. Paths from A to X have 6 edges: 2 up and 4 right. There are $\binom{6}{4}$ paths from A to X.
4. Paths from X to B have 5 edges: 2 up and 3 right. There are $\binom{5}{3}$ paths from X to B.
5. The probability a random path goes through X = P(E) = $\binom{6}{4}\binom{5}{3}/\binom{11}{7} = \frac{5}{11}$.

Example 6.1.3: An example from a 5-card hand.

A dealer deals a 5-card hand. If the deck is perfectly shuffled, then each 5-card hand is equally likely and the distribution over the sample space of 5-card hands is uniform. What is the probability that the hand has two pairs? The definition of having two pairs is that there are two pairs of cards, each pair has the same rank. The pairs have different rank from each other and the fifth card (that is not part of a pair), has different rank from the pairs. The diagram below shows some examples that are and are not hands with two pairs:

{2♥, 2♦, 8♠, 8♣, A♣} A hand with 2 pairs

{2♥, 2♦, 2♠, 2♣, A♣} Not a hand with 2 pairs
(the pairs must have different rank)

{2♥, 2♦, 8♠, 8♣, 2♣} Not a hand with 2 pairs
(the last card must have different rank from the pairs)

Define P to be the set of 5-card hands that have two pairs. The probability of event P is:

$$p(P) = \frac{|P|}{\binom{52}{5}}$$

The animation below shows how to determine |P|.

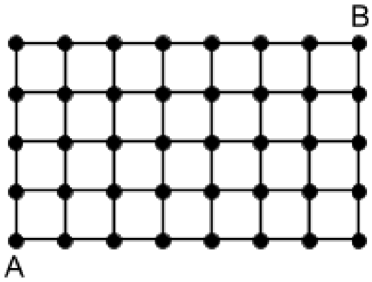
PARTICIPATION ACTIVITY 6.1.5: Number of 5-card hands with two pairs.

Animation captions:

1. Number of 5-card hands with two pairs. There are $\binom{13}{2}$ ways to select the ranks for the two pairs from {2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A}.
2. There are $\binom{4}{2}$ ways to select the suits for one pair from the set {Diamonds, Hearts, Clubs, Spades}.
3. There are $\binom{4}{2}$ ways to select the suits for the other pair.
4. The last card can be any card whose rank is not the same as the pairs. There are $52 - 4 - 4 = 44$ choices.
5. The probability that a random hand is a 2-pair hand is $\binom{13}{2}\binom{4}{2}\binom{4}{2}44/\binom{52}{5} \approx .04754$.

PARTICIPATION ACTIVITY 6.1.6: Using counting to determine the probability of events.

1) In the grid graph shown below, what's the probability that a random direct path from A to B goes through the vertex in the upper left corner?



- $\frac{11}{\binom{11}{7}}$
- $\frac{1}{\binom{7}{4}}$
- $\frac{1}{\binom{11}{7}}$

2) If a red and a blue die are thrown, then what is the probability that the numbers on the two dice are the same?

- $\frac{5}{36}$
- $\frac{1}{2}$
- $\frac{1}{6}$

3) What is the probability that a random 5-card hand has all four aces?

- $\frac{48}{\binom{52}{5}}$
- $\frac{4 \cdot 48}{\binom{52}{5}}$
- $\frac{1}{\binom{48}{5}}$

Additional exercises

Exercise 6.1.1: Coin flips and events.

About

A coin is flipped four times. For each of the events described below, express the event as a set in roster notation. Each outcome is written as a string of length 4 from {H, T}, such as HHTH. Assuming the coin is a fair coin, give the probability of each event.

(a) The first and last flips come up heads.

Solution

(b) There are at least two consecutive flips that come up heads.

(c) The first flip comes up tails and there are at least two consecutive flips that come up heads.

Exercise 6.1.2: The probability of an event under the uniform distribution - random permutations.

About

A class with n kids lines up for recess. The order in which the kids line up is random with each ordering being equally likely. There are two kids in the class named Celia and Felicity.

Give an expression for each of the probabilities below as a function of n. Simplify your final expression as much as possible so that your answer does not include any expressions of the form $\binom{a}{b}$.

(a) What is the probability that Celia is first in line?

Solution

(b) What is the probability that Celia is first in line and Felicity is last in line?

(c) What is the probability that Celia and Felicity are next to each other in the line?

Exercise 6.1.3: The probability of an event under the uniform distribution - selecting a pair of matching socks.

About

A drawer contains n white and n black socks. Each white sock has a unique design, and each black sock has a unique design. Two socks are selected at random from the drawer. Every way of selecting the two socks is equally likely, and the order in which the socks are selected does not matter.

Source: ADUni, modified by Sandy Irani.

(a) How many ways are there to select the two socks?

Solution

(b) How many ways of selecting the socks result in two socks of the same color being chosen?

Solution

(c) What is the probability that a randomly chosen pair of socks are the same color? Simplify your final expression as much as possible so that it does not include any expressions of the form $\binom{a}{b}$.

Solution

Exercise 6.1.4: The probability of an event under the uniform distribution - picking teams.

About

10 kids are randomly grouped into an A team with five kids and a B team with five kids. Each grouping is equally likely.

(a) What is the size of the sample space?

(b) There are two kids in the group, Alex and his best friend Jose. What is the probability that Alex and Jose end up on the same team?

(c) If the group consists of five girls and five boys, what is the probability that all the girls end up on the same team?

Exercise 6.1.5: The probability of an event under the uniform distribution - 5-card hands.

About

A 5-card hand is dealt from a perfectly shuffled deck of playing cards. What is the probability of each of the following events?

(a) What is the probability that the hand is a full house? A full house has three cards of the same rank and another pair of the same rank. For example, {4♠, 4♥, 4♦, J♠, J♣} is a full house.

Solution

(b) What is the probability that the hand is a three of a kind? A three of a kind has 3 cards of the same rank. The other two cards do not have the same rank as each other and do not have the same rank as the three with the same rank. For example, {4♠, 4♦, 4♣, J♠, 8♥} is a three of a kind.

(c) What is the probability that all 5 cards have the same suit?

(d) What is the probability that the hand is a two of a kind? A two of a kind has two cards of the same rank (called the pair). Among the remaining three cards, not in the pair, no two have the same rank and none of them have the same rank as the pair. For example, {4♠, 4♦, J♠, K♣, 8♥} is a two of a kind.

6.2 Unions and complements of events

Calculating probabilities for unions of events

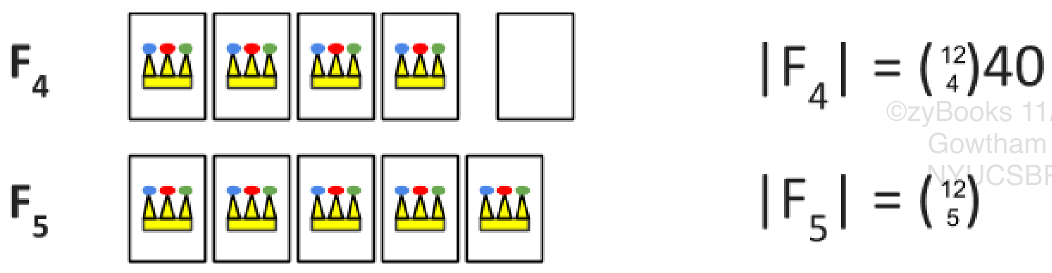
Sometimes it is easier to determine the probability of an event by defining the event in terms of other events. Suppose we would like to calculate the probability that a 5-card hand has at least four face cards (jack, queen, or king). It is easier to determine the probability that the 5-card hand has exactly four face cards or exactly five face cards. Adding the probabilities of the two events gives the correct answer because a hand can not have four face cards and five face cards at the same time.

Two events are ***mutually exclusive*** if the two events are disjoint (i.e., the intersection of the two events is empty). It follows from the definition of the probability of an event that if E_1 and E_2 are mutually exclusive, then:

$$p(E_1 \cup E_2) = p(E_1) + p(E_2)$$

If F_4 is the event that a random 5-card hand has four face cards and F_5 is the event that a random 5-card hand has five face cards, then $F_4 \cap F_5 = \emptyset$. In a deck of cards, there are twelve face cards (jack, queen, king in each of four suits) and therefore 40 non-face cards.

Figure 6.2.1: 5-card hands with at least four face cards.



The probability that a 5-card hand has at least four face cards is:

$$p(F_4 \cup F_5) = p(F_4) + p(F_5) = \frac{\binom{12}{4}40}{\binom{52}{5}} + \frac{\binom{12}{5}}{\binom{52}{5}}$$

PARTICIPATION ACTIVITY | 6.2.1: Identifying mutually exclusive events.

Are events A and B mutually exclusive?

1) A coin is tossed five times.

A: The first three flips come up heads
B: The last three flips come up tails

- Mutually exclusive
- Not mutually exclusive

2) A coin is tossed six times.

A: The first three flips come up heads
B: The last three flips come up tails

- Mutually exclusive
- Not mutually exclusive

3) A 5-card hand is dealt from a standard playing deck.

A: The hand has three aces
B: The hand has a pair of kings

- Mutually exclusive
- Not mutually exclusive

4) A 5-card hand is dealt from a standard playing deck.

A: The hand has three aces
B: The hand has three kings

- Mutually exclusive
- Not mutually exclusive

If two events are not mutually exclusive, the probability of the union of events can be determined by a version of the Inclusion-Exclusion principle:

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$$

The statement holds for non-uniform as well as uniform distributions. The following animation shows why the equality holds:

PARTICIPATION ACTIVITY 6.2.2: The probability of a union of events.

Animation captions:

1. Start with $p(E_1) = \sum_{s \in E_1} p(s)$. Every outcome in E_1 has been counted once.
2. Add in $p(E_2) = \sum_{s \in E_2} p(s)$. Every outcome in $E_1 - E_2$ and $E_2 - E_1$ has been counted once. Every outcome in $E_1 \cap E_2$ has been counted twice.
3. Subtract $p(E_1 \cap E_2) = \sum_{s \in E_1 \cap E_2} p(s)$. Every outcome in $E_1 \cup E_2$ has been counted once. Therefore $p(E_1) + p(E_2) - p(E_1 \cap E_2) = p(E_1 \cup E_2)$.

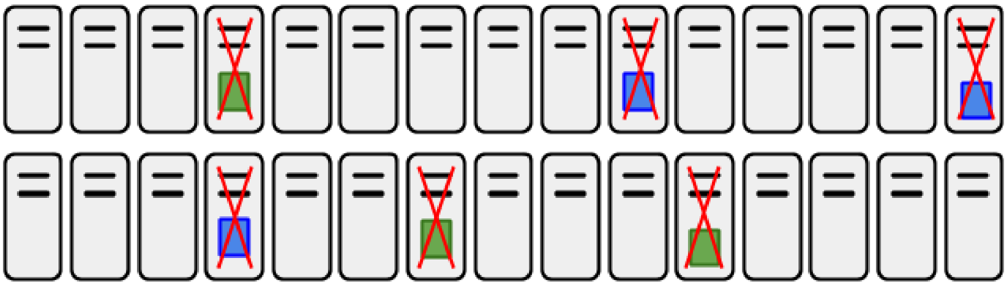
Example 6.2.1: Redundant data storage revisited.

Consider again the situation in which files are stored on a distributed network. Three copies of File 1 are stored at three distinct locations in the network, and three copies of File 2 are stored at three different locations in the network. Suppose that there are 6 random computers that have failed. What is the probability that either file has been wiped out? Let F_1 be the event that all three copies of File 1 are gone and F_2 the event that all three copies of File 2 are gone. What is $p(F_1 \cup F_2)$?

Using the same reasoning from the previous example:

$$\frac{|F_1|}{|S|} = \frac{|F_2|}{|S|} = \frac{\binom{27}{3}}{\binom{30}{6}}$$

However, F_1 and F_2 are not mutually exclusive. Since six computers have failed, it is possible that all three copies of File 1 and all three copies of File 2 were on failed computers. The event $F_1 \cap F_2$ happens when the six failed computers include the three with copies of File 1 and the three with copies of File 2. Since the copies of File 1 and File 2 are all stored on different computers, there is only one outcome in which both File 1 and File 2 are wiped out as pictured below:



Thus the probability that either File 1 or File 2 is gone is:

$$p(F_1 \cup F_2) = \frac{|F_1|}{|S|} + \frac{|F_2|}{|S|} - \frac{|F_1 \cap F_2|}{|S|} = \frac{2 \cdot \binom{27}{3} - 1}{\binom{30}{6}} \approx 0.0049$$

PARTICIPATION ACTIVITY | 6.2.3: Probabilities of unions of events.

The following questions pertain to an experiment in which a fair coin is flipped three times. Define the following three events:

- A: The first two flips come up heads.
- B: The last two flips come up heads.
- C: The last two flips come up tails.

Enter your answer as a simplified fraction. Ex: If your answer is 2/4, then enter: 1/2

1) Are the events A and C mutually exclusive?

Check Show answer

2) What is the probability that the first two flips come up heads and the last two flips come up tails? (What is p(A ∩ C)?)

Check Show answer

3) What is the probability that the first two flips are heads or the last two flips are tails. (What is p(A ∪ C)?)

Check Show answer

4) Are the events A and B mutually exclusive?

Check Show answer

5) What is the probability that the first two flips come up heads and the last two flips come up heads? (What is p(A ∩ B)?)

Check Show answer

6) What is the probability that the first two flips are heads or the last two flips are heads. (What is p(A ∪ B)?)

Check Show answer

The complement of an event

Sometimes it is easier to determine the probability that an event doesn't happen than to determine that the event does happen. Fortunately, there is an easy relationship between the two possibilities. The ***complement*** of an event E is S - E and is denoted by $\overline{E}$. Since $\overline{E}$ and E are disjoint events, p($\overline{E}$) + p(E) = 1. It follows that

$$p(\overline{E}) = 1 - p(E).$$

In the experiment in which the red and blue dice are thrown, define E to be the event that at least one of the dice comes up 6. The event $\overline{E}$ is the event that neither die comes up 6 which is equivalent to saying that each die has a number from the set {1, 2, 3, 4, 5}. Thus, $\overline{E}$ = {1, 2, 3, 4, 5} × {1, 2, 3, 4, 5}, and |$\overline{E}$| = 25.

$$p(E) = 1 - p(\overline{E}) = 1 - \frac{25}{36} = \frac{11}{36}.$$

PARTICIPATION ACTIVITY 6.2.4: Determining probabilities of events via the complement.

Express answer as an integer or a fraction in lowest terms (e.g., 3/4, not 9/12)

1) If a fair coin is flipped three times, what is the probability that there is at

least one head?

Check Show answer

2) If a fair coin is flipped ten times, what is the probability that there is at least one head? (Hint: The number of outcomes is $2^{10} = 1024$.)

Check Show answer

3) In a roll of a blue and a red die, what is the probability that the sum of the numbers is at most 10?

Check Show answer

Additional exercises

Exercise 6.2.1: The probability of an event - coin flips.

About

A fair coin is flipped n times. Give an expression for each of the probabilities below as a function of n. Simplify your final expression as much as possible.

(a) At least n - 1 flips come up heads.

Solution

(b) There are at least two consecutive flips that are the same.

(c) The number of heads is different from the number of tails. (Assume that n is even, otherwise the probability is one.)

Exercise 6.2.2: The probability of an event - random permutations.

About

A class with n kids lines up for recess. The order in which the kids line up is random with each ordering being equally likely. There are three kids in the class named Hubert, Celia and Felicity. The use of the word "or" in the description of the events, should be interpreted as the inclusive or. That is "A or B" means that A is true, B is true or both A and B are true.

Give an expression for each of the probabilities below as a function of n. Simplify your final expression as much as possible so that your answer does not include any expressions of the form $\binom{a}{b}$.

(a) What is the probability that Celia is first in line or Felicity is first in line?

Solution

(b) What is the probability that Celia is first in line or Felicity is last in line?

(c) What is the probability that Celia is not next to Felicity in line?

(d) What is the probability that Hubert is next to Celia or Felicity in line?

(e) What is the probability that Hubert is not next to either Celia or Felicity in line?

Exercise 6.2.3: The probability of an event - picking teams.

About

(a) 10 kids are randomly grouped into an A team with five kids and a B team with five kids. Each grouping is equally likely. There are three kids in the group, Alex and his two best friends Jose and Carl. What is the probability that Alex ends up on the same team with at least one of his two best friends?

Solution

Exercise 6.2.4: The probability of events - 5-card hands.

About

A 5-card hand is dealt from a perfectly shuffled deck of playing cards. What is the probability of each of the following events?

(a) The hand has at least one club.

(b) The hand has at least two cards with the same rank.

(c) The hand has exactly one club or exactly one spade.

(d) The hand has at least one club or at least one spade.

Exercise 6.2.5: Probability a random string is a valid password. About

(a) An online vendor requires that customers select a password that is a sequence of upper-case letters, lower-case letters and digits. A valid password must be at least 10 characters long, and it must contain at least one character from each of the three sets of characters. What is the probability that a randomly selected string with exactly ten characters results in a valid password? The alphabet for the strings in the sample space from which the string is drawn is the union of the three sets of characters.

Source: ADUni, modified by Sandy Irani.

6.3 Conditional probability and independence

Let's roll our blue and red dice again and look at the event E that the two numbers on the dice sum to at least 11. The event E is the set {(5,6),(6,5),(6,6)}, so the probability of E is |E|/|S| = 3/36 = 1/12. Now suppose we start the experiment by rolling the blue die first and the blue die comes up 5. How does the information that the blue die has come up 5 change the probability that the event E happens? In a sense, the sample space has shrunk from the set of all 36 outcomes to the set of outcomes (5, *) in which the blue die comes up 5. The probabilities need to be adjusted to reflect the smaller sample space.

There are two events needed in the analysis. The first is the original event E that the sum of the two numbers is at least 11. The second event F is that the blue die comes up 5. If the event F happens, the new probability of E is the ***conditional probability*** of E given F, denoted by p(E|F). The following definition gives a formula for calculating p(E|F):

Definition 6.3.1: Conditional probability.

The conditional probability of E given F is

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

If the distribution is uniform, then p(E) = |E|/|S| and the conditional probability becomes:

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)} = \frac{|E \cap F| \ / \ |S|}{|F| \ / \ |S|} = \frac{|E \cap F|}{|F|}$$

The animation below gives some intuition about the definition of conditional probability using a more abstract example with a uniform distribution:

PARTICIPATION ACTIVITY 6.3.1: Conditional probability.

Animation captions:

1. The sample space S = {a, b, c, d, e, f, g, h, i, j, k, l, m}. |S| = 13. The probability of each outcome is $\frac{1}{13}$.
2. Event F = {b, e, d, f, g}. If F happens, then p(x|F) = $\frac{1}{|F|}$ = $\frac{1}{5}$ for every outcome x in F and p(x|F) = 0 for every event x not in F.
3. Event E = {b, c, d, e}. p(E|F) = p(b|F) + p(d|F) + p(e|F) = $\frac{1}{5} + \frac{1}{5} + \frac{1}{5} = \frac{3}{5} = \frac{|E \cap F|}{|F|}$.

The quantities |E ∩ F| and |F| are needed to calculate the conditional probability of E given F under a uniform distribution. In the dice example:

- E ∩ F = {(5,6)}, since E ∩ F is the event that the blue die is 5 and the sum of the two dice is at least 11.
- F is all outcomes of the form (5, *). |F| = 6 since there are six remaining possibilities for the red die.

Therefore, the probability that the sum is at least 11, given that the blue die is 5, is:

$$p(E \mid F) = \frac{|E \cap F|}{|F|} = \frac{1}{6}$$

PARTICIPATION ACTIVITY 6.3.2: Calculating a conditional probability - three coin flips.

The following questions pertain to the experiment in which a fair coin is tossed 3 times.

1) Let W be the event that there are at least two heads. What is |W|?

Check Show answer

2) Let F be the event that the first flip comes up heads. What is |W ∩ F|?

Check Show answer

3) What is p(F|W)?

Check Show answer

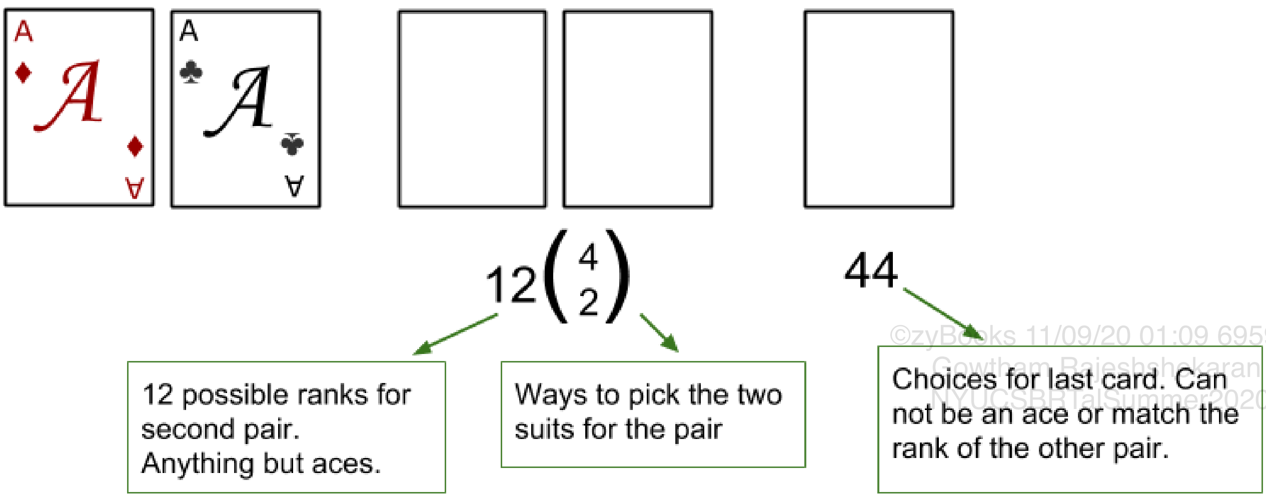
A 5-card hand example of conditional probability

We saw earlier that the probability that a 5-card hand has 2 pairs is about 0.0475. Now suppose that a player sees the first two cards he is dealt. Those two cards are the ace of clubs and the ace of diamonds. What is the likelihood after seeing those first two cards that the hand will have two pairs? Define A to be the event that the first two cards are A♣ and A♦. Let P be the event that the hand has two pairs. The goal is to determine p(P|A).

The number of outcomes in A is $\binom{50}{3}$ because the remaining three cards in the hand are chosen from the remaining 50 cards in the deck (i.e., the cards that are not A♣ or A♦).

The event A ∩ P is the event that the first two cards are A♣ or A♦ and the remaining three cards include a pair of non-aces and a card that is not an ace and does not match the rank of the second pair. The diagram below illustrates:

Figure 6.3.1: Number of hands in A ∩ P.



The probability of P given A is:

$$p(P\,|\,A) = \frac{|\,A \cap P\,|}{|\,A\,|} = \frac{12 \cdot \binom{4}{2} \cdot 44}{\binom{50}{3}} \approx 0.1616$$

The complement of an event and conditional probability

If E and F are both events in the same sample space S, then the probability of E and the probability of $\overline{E}$ still sum to 1, even when conditioned on the event F.

$$p(E|F) + p(\overline{E}|F) = 1$$

Proof.
The events E ∩ F and $\overline{E}$ ∩ F are mutually exclusive, so p((E ∩ F) ∪ ($\overline{E}$ ∩ F)) = p(E ∩ F) + p($\overline{E}$ ∩ F). Using the definition for conditional probability:

$$p(E \mid F) + p(\overline{E} \mid F) = \frac{p(E \cap F)}{p(F)} + \frac{p(\overline{E} \cap F)}{p(F)} = \frac{p(E \cap F) + p(\overline{E} \cap F)}{p(F)} = \frac{p((E \cap F) \cup (\overline{E} \cap F))}{p(F)}$$

We can use the set identities to establish the following set equality, with the role of the universe set played by the sample space S:

$$(E \cap F) \cup (\overline{E} \cap F) = (E \cup \overline{E}) \cap F = S \cap F = F$$

Therefore p((E ∩ F) ∪ ($\overline{E}$ ∩ F)) = p(F) and

$$p(E \mid F) + p(\overline{E} \mid F) = \frac{p((E \cap F) \cup (\overline{E} \cap F))}{p(F)} = \frac{p(F)}{p(F)} = 1 \blacksquare$$

PARTICIPATION ACTIVITY | 6.3.3: Calculating conditional probabilities: Two dice.

The following questions pertain to an experiment in which a blue and red die are tossed. Express probabilities as a simplified fraction (e.g., 0, 3/4 or 1/7, but not 4/8).

1) What is the probability that the sum is at least 10 given that the two die have the same number?

Check Show answer

2) What is the probability that the sum is less than 10 given that the two die have the same number?

Check Show answer

3) What is the probability that the sum is at least 10 given that the blue die comes up 2?

Check Show answer

4) What is the probability that the sum is less than 10 given that the blue die comes up 2?

Check Show answer

CHALLENGE ACTIVITY 6.3.1: Conditional Probability.

Start

A blue die and a red die are thrown.
B is the event that the blue comes up with a 3.
E is the event that the sum of the dice is 9.
Enter the sizes of the sets |E ∩ B| and |B|

|E ∩ B| = Ex: 5

|B| = Ex: 5

1 2

Check Next

Independent events

We calculated earlier that when the blue and red die are thrown, the event E that the two dice are the same has probability 1/6. Now suppose we consider the probability of E conditioned on the event F that the blue die comes up 5. What is p(E|F)? The event E ∩ F is the event that the first die comes up 5 and the two dice are the same. There is only one outcome then in E ∩ F: (5, 5).

$$p(E \mid F) = \frac{|E \cap F|}{|F|} = \frac{1}{6}$$

Conditioning on the event F does not change the probability of E. In other words, knowing that the first die came up 5 provides no additional information about the likelihood that the two dice will have the same number. Two events are ***independent*** if conditioning on one event does not change the probability of the other event. Here is a formal description in terms of probabilities:

Definition 6.3.2: Independent events.

Let E and F be two events in the same sample space. The following three conditions are equivalent:

1. $p(E \mid F) = \frac{p(E \cap F)}{p(F)} = p(E)$
2. $p(E \cap F) = p(E) \cdot p(F)$
3. $p(F \mid E) = \frac{p(E \cap F)}{p(E)} = p(F)$

If the three conditions hold, then events E and F are independent.

In the 5-card hand experiment, the event P that the hand has 2 pairs was shown to have probability around 0.04754. However, the probability of event P conditioned on the event A, that the first two cards are A♣ and A♦, is 0.1616. Thus, p(P|A) ≠ p(P), and therefore the events P and A are not independent.

PARTICIPATION ACTIVITY 6.3.4: Independent events.

Animation captions:

1. Experiment: pick a random permutation of a, b, c. There are 6 outcomes in the sample space. Event E: a comes before c = {abc, acb, bac}. $p(E) = \frac{3}{6} = \frac{1}{2}$.
2. Event F: b comes first = {bac, bca}. $p(F) = \frac{2}{6} = \frac{1}{3}$.
3. $E \cap F$: a comes before c and b comes first = {bac}. $p(E \cap F) = \frac{1}{6}$.
4. $p(E \cap F) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = p(E) \cdot p(F)$. Therefore events E and F are independent.
5. Event G: a comes before b = {abc, acb, cab}. $p(G) = \frac{3}{6} = \frac{1}{2}$. $E \cap G$: a comes before c and a comes before b = {abc, acb}.
6. $p(E \cap G) = \frac{2}{6} = \frac{1}{3} \neq \frac{1}{2} \cdot \frac{1}{2} = p(E) \cdot p(G)$. E and G are not independent.

PARTICIPATION ACTIVITY 6.3.5: Determining whether two events are independent - three coin flips.

A fair coin is flipped 3 times. Define the following events:

- F: the first flip comes up heads.
- E: the number of heads is even. (Zero is an even number.)
- W: at least two of the three flips come up heads.

1) What is p(F)?

Check Show answer

2) What is p(E)?

Check Show answer

3) What is the probability that the first flip comes up heads and there is an even number of heads? (What is p(F ∩ E)?)

Check Show answer

4) Are E and F independent?

Check Show answer

5) What is the probability that at least two flips come up heads? (What is p(W)?)

Check Show answer

6) What is the probability that the first flip is heads and there are at least two flips that are heads?

Check Show answer

7) Are F and W independent?

Check Show answer

Calculating the probabilities of two independent events

In some situations, it is reasonable to assume that two events are independent because it is unlikely that one event influences the other event. For example, if a single die is rolled twice, the event that the die comes up 5 on the first roll is unlikely to affect whether or not the die comes up 5 again on the second roll. If two events are independent, it is easier to calculate the probability that both events happen because the probability that both occur is the product of the probabilities that each event occurs. If X and Y are events in the same sample space, and X and Y are independent, then

$$p(X \cap Y) = p(X) \cdot p(Y)$$

PARTICIPATION ACTIVITY 6.3.6: Calculating probabilities of independent rolls of a die.

A loaded die is twice as likely to come up 6 than any of the other five possibilities. The probability distribution over the outcomes of a single roll of the die is defined by:

$$p(1) = p(2) = p(3) = p(4) = p(5) = \frac{1}{7} p(6) = \frac{2}{7}$$

The loaded die is rolled twice. The outcomes of the two rolls are independent, meaning that any event that only depends on the outcome of the first roll is independent of any event that only depends on the outcome of the second roll.

Express probabilities as a simplified fraction (e.g., 0, 3/4 or 1/7, but not 4/8).

1) What is the probability that the first roll does not come up 6?

Check Show answer

2) What is the probability that the first roll does not come up 6 and the second roll does come up 6?

Check Show answer

3) What is the probability that exactly one of the two rolls comes up 6?

Check Show answer

Mutual independence

Consider three events, A, B, and C in the same sample space. Even if every pair of the three events are independent, it is not necessarily true that $p(A \cap B \cap C) = p(A) \cdot p(B) \cdot p(C)$. There is a stronger version of independence that applies to more than two events.

Events $A_1, \ldots, A_n$ in sample space S are ***mutually independent*** if the probability of the intersection of any subset of the events is equal to the product of the probabilities of the events in the subset. In particular, if $A_1, \ldots, A_n$ are mutually independent, then

$$p(A_1 \cap A_2 \cap \ldots \cap A_n) = p(A_1) \cdot p(A_2) \ldots p(A_n).$$

Saying that n throws of a die are mutually independent means that any n events, $A_1, \ldots, A_n$, in which each event A_i only depends on the outcome of the i^{th} throw, are mutually independent events.

PARTICIPATION ACTIVITY | 6.3.7: Probabilities of mutually independent events.

A loaded die is twice as likely to come up 6 than any of the other five possibilities. The probability distribution over the outcomes of a single roll of the die is defined by:

$$p(1) = p(2) = p(3) = p(4) = p(5) = \frac{1}{7} p(6) = \frac{2}{7}$$

The die is rolled five times. The outcomes of the rolls of the die are mutually independent.

Let A_i be the event that the i^{th} roll comes up at least 5 (i.e. 5 or 6).

1) Are the events A_1, A_2, A_3, A_4 A_5 mutually independent?

- Yes
- No

2) What is the probability that on all of the five rolls, the die comes up at least 5?

- $(1/3)^5$
- $(3/7)^5$
- 3/7

3) Are the events $\overline{A}_1, A_2, A_3, A_4, A_5$ mutually independent?

- Yes
- No

4) What is the probability that the first roll comes up less than 5 and all of the other rolls come up at least 5?

- $(3/7)^4$

- $(4/7)^5$
- $(4/7) \cdot (3/7)^4$

Additional exercises

Exercise 6.3.1: Calculating conditional probabilities - two dice.

About

A red and a blue die are thrown. Both dice are fair. The events A, B, and C are defined as follows:

A: The sum on the two dice is even
B: The sum on the two dice is at least 10
C: The red die comes up 5

(a) Calculate the probability of each individual event. That is, calculate p(A), p(B), and p(C).

Solution

(b) What is p(A|C)?

Solution

(c) What is p(B|C)?

Solution

(d) What is p(A|B)?

Solution

(e) Which pairs of events among A, B, and C are independent?

Solution

Exercise 6.3.2: Calculating conditional probabilities - random permutations.

About

The letters {a, b, c, d, e, f, g} are put in a random order. Each permutation is equally likely. Define the following events:

A: The letter b falls in the middle (with three before it and three after it)
B: The letter c appears to the right of b, although c is not necessarily immediately to the right of b. For example, "agbdcef" would be an outcome in this event.
C: The letters "def" occur together in that order (e.g. "gdefbca")

(a)

Calculate the probability of each individual event. That is, calculate p(A), p(B), and

Calculate the probability of each individual event. That is, calculate p(A), p(B), and p(C).

(b)
What is p(A|C)?

(c)
What is p(B|C)?

(d)
What is p(A|B)?

(e)
Which pairs of events among A, B, and C are independent?

Exercise 6.3.3: Independence of events - random permutations.

About

A wedding party of eight people is lined up in a random order. Every way of lining up the people in the wedding party is equally likely.

(a)
What is the probability that the bride is next to the groom?

(b)
What is the probability that the maid of honor is in the leftmost position?

(c)
Determine whether the two events are independent. Prove your answer by showing that one of the conditions for independence is either true or false.

Exercise 6.3.4: The probability of an event - picking teams.

About

(a) 10 kids are randomly grouped into an A team with five kids and a B team with five kids. Each grouping is equally likely. There are three kids in the group, Alex and his two best friends, Jose and Carl. Define the events J and C as:

J: Alex ends up on the same team as Jose

C: Alex ends up on the same team as Carl

Are the events J and C independent? Prove your answer by showing that one of the conditions for independence is either true or false.

Solution

Exercise 6.3.5: Calculating conditional probabilities - a 5-card hand.

About

(a) A 5-card hand is dealt from a perfectly shuffled deck. Define the events:

A: the hand is a four of a kind (all four cards of one rank plus a 5^{th} card).

B: at least one of the cards in the hand is an ace

Are the events A and B independent? Prove your answer by showing that one of the conditions for independence is either true or false.

Exercise 6.3.6: Calculating probabilities of independent events.

About

A biased coin is flipped 10 times. In a single flip of the coin, the probability of heads is 1/3 and the probability of tails is 2/3. The outcomes of the coin flips are mutually independent. What is the probability of each event?

(a) Every flip comes up heads.

Solution

(b) The first 5 flips come up heads. The last 5 flips come up tails.

(c) The first flip comes up heads. The rest of the flips come up tails.

6.4 Bayes' Theorem

Suppose a gambler has two dice that look the same, except that one is a fair die and the other is loaded. When the fair die is rolled, each of the six outcomes is equally likely to occur. When the loaded die is rolled, a 6 is twice as likely to come up than the other outcomes. The probability distribution over the outcomes of the loaded die is:

$$p(1) = p(2) = p(3) = p(4) = p(5) = \frac{1}{7} p(6) = \frac{2}{7}$$

The gambler has both dice in her pocket and pulls out a random one for a game. Is it possible to reason about which die she picked based on the outcomes of the rolls? Presumably, if a 6 comes up more often than expected, it is more likely that the gambler is using the loaded die. Bayes' Theorem provides a way to reason quantitatively about the likelihood that the die is loaded based on the evidence of the outcomes. Bayes' Theorem is the cornerstone of many algorithms in machine learning in which the goal is to determine the likelihood of some event based on data obtained from observations.

In the loaded die problem, the event F is the event that the gambler selects the fair die and $\overline{F}$ is the event that she selects the loaded die. Since she picks a die at random, p(F) = p($\overline{F}$) = 1/2. Let X be the event that the die comes up 6. We know the probability of X conditioned on F and $\overline{F}$:

$$p(X \mid F) = \frac{1}{6} \qquad p(X \mid \overline{F}) = \frac{2}{7}$$

We would like to know the probability of F or $\overline{F}$ conditioned on observing a 6 (i.e. conditioned on the event X). Thus, we want to know p(F|X). Bayes' Theorem provides a way to determine p(F|X) from p(X|F), p(X|$\overline{F}$) and p(F):

Theorem 6.4.1: Bayes' Theorem.

Suppose that F and X are events from a common sample space and p(F) ≠ 0 and p(X) ≠ 0. Then

$$p(F \mid X) = \frac{p(X \mid F)p(F)}{p(X \mid F)p(F) + p(X \mid \overline{F})p(\overline{F})}$$

The animation below shows how to apply Bayes' Theorem to determine the likelihood that the die is loaded given that a roll comes up 6:

PARTICIPATION ACTIVITY 6.4.1: Application of Bayes' Theorem.

Animation content:

`undefined`

Animation captions:

1. F: fair die selected. $\overline{F}$: loaded die selected. X: the die comes up 6. $p(F) = p(\overline{F}) = \frac{1}{2}$. $p(X \mid F) = \frac{1}{6}$. $p(X \mid \overline{F}) = \frac{2}{7}$.
2. Bayes' Theorem: $\frac{p(X|F) \cdot p(F)}{p(X|F) \cdot p(F) + p(X|\overline{F}) \cdot p(\overline{F})}$. Replace each occurrence of $p(F)$ and $p(\overline{F})$ with $\frac{1}{2}$.
3. Replace both occurrences of $p(X \mid F)$ with $\frac{1}{6}$.

4. Replace the occurrence of $p(X|\overline{F})$ with $\frac{2}{7}$. The result is $\frac{(1/6)\cdot(1/2)}{(1/6)\cdot(1/2)+(2/7)\cdot(1/2)} \approx .37$.

The proof of Bayes' Theorem is given below. The proof makes use of the set identity X = ((X ∩ F) ∪ (X ∩ F̅)), which can be verified using the set identities given in another section.

Proof.

The numerator of the expression on the right is:

$$p(X \mid F)p(F) = p(X \cap F)$$

The first term in the denominator is also p(X ∩ F). The second term in the denominator is:

$$p(X \mid \overline{F})p(\overline{F}) = p(X \cap \overline{F}).$$

Adding the two terms in the denominator gives:

$$p(X \cap F) + p(X \cap \overline{F}) = p((X \cap F) \cup (X \cap \overline{F})) = p(X)$$

Plugging in the expressions for the numerator and the denominator:

$$\frac{p(X \mid F)p(F)}{p(X \mid F)p(F) + p(X|\overline{F})p(\overline{F})} = \frac{p(X \cap F)}{p(X)} = p(F \mid X) \blacksquare$$

PARTICIPATION ACTIVITY 6.4.2: Another Bayes problem with the fair and loaded dice.

Consider the example with the fair and loaded die. The gambler selects one of the two die at random and rolls the die twice. The outcomes of the two rolls are independent. The first outcome is not a 6 and the second outcome is a 6. Define Y to be the event that two rolls of the die result in the first roll not coming up 6 and the second roll coming up 6.

Express probabilities as a simplified fraction (e.g., 0, 3/4 or 1/7, but not 4/8).

1) What is p(Y|F)?

Check Show answer

2) What is p(Y|F̅)?

Check Show answer

3) Apply the results from the previous two questions to Bayes' Theorem to determine p(F|Y).

Check Show answer

Example 6.4.1: Testing for defects in computer chip manufacturing.

Suppose that 1 out of every 1000 computer chips produced by a chip manufacturer has a defect. The manufacturer has developed a method to test the chips, but the test is not perfect. If the chip has a defect, the test will discover the defect with probability 0.99. If the chip does not have a defect, the test will report that it does have a defect with probability 0.02. If the outcome of a particular test indicates that there is a defect, what is the likelihood that the chip is actually faulty?

Let D be the event that the chip has a defect. Let T be the event that the test outcome indicates a defect. Here is a summary of the probabilities given:

- p(D) = 0.001 because one out of every 1000 chips has a defect.
- p(T|D) = 0.99. If the chip is faulty, a test will indicate a defect with probability 0.99.
- $p(T|\overline{D})$ = 0.02. If a chip does not have a defect, the test reports that it has a defect with probability 0.02.

Applying Bayes' Theorem, we get:

$$p(D\,|\,T) = \frac{p(T\,|\,D)p(D)}{p(T\,|\,D)p(D) + p(T\,|\,\overline{D})p(\overline{D})} = \frac{0.99 \times 0.001}{(0.99 \times 0.001) + (0.02 \times 0.999)} \approx 0.047$$

It may seem counter-intuitive that the probability the chip has a defect is so low given that the test is quite accurate and predicted a defect. However, Bayes' Theorem also factors in the fact that the probability of a defect is very low in the absence of information from the test.

PARTICIPATION ACTIVITY 6.4.3: Bayes' Theorem.

One out of every 10,000 people has a particular genetic disease. A test has been developed for the disease that is very accurate but has some likelihood of error. If a person with the disease is tested, the test results say the person does not have the disease with probability 0.01. If the person does not have the disease, the test is incorrect with probability 0.005. Let D be the event that the person has the disease. Let T be the event that the person tests positive for the disease.

Note that a "positive" test result means that the results indicate that the person has the disease. A "negative" test result means that the results indicate that the person does not have the disease.

1) Select the mathematical equivalent of the following statement: "One out of every 10,000 people has a particular genetic disease."

- $p(D) = 0.0001$
- $p(T) = 0.0001$
- $p(T|D) = 0.0001$

2) Select the mathematical equivalent of the following statement: "If a person with the disease is tested, the test results say the person does not have the disease with probability 0.01."

- $p(\overline{D}|T) = 0.01$
- $p(T|D) = 0.01$
- $p(\overline{T}|D) = 0.01$

3) If $p(\overline{T}|D) = 0.01$, then what is $p(T|D)$?

- 1
- 0.99
- 0.01

4) Select the mathematical equivalent of the following statement: "If the person does not have the disease, the test is incorrect with probability 0.005."

- $p(T|\overline{D}) = 0.005$
- $p(\overline{T}|\overline{D}) = 0.005$

5) Use Bayes' Theorem to determine the expression for the probability that a person who tests positive actually has the disease.

- $$\frac{(0.99)(0.9999)}{(0.99)(0.9999) + (0.005)(0.0001)}$$
- $$\frac{(0.01)(0.0001)}{(0.01)(0.0001) + (0.005)(0.9999)}$$
- $$\frac{(0.99)(0.0001)}{(0.99)(0.0001) + (0.005)(0.9999)}$$

Additional exercises

Exercise 6.4.1: Bayes' Theorem - detecting a biased coin.

About

(a) Sally has two coins. The first coin is a fair coin and the second coin is biased. The biased coin comes up heads with probability .75 and tails with probability .25. She selects a coin at random and flips the coin ten times. The results of the coin flips are mutually independent. The result of the 10 flips is: T,T,H,T,H,T,T,T,H,T. What is the probability that she selected the biased coin?

Solution

Exercise 6.4.2: Bayes' Theorem - detecting a loaded die.

About

(a) Assume that you have two dice, one of which is fair, and the other is biased toward landing on six, so that 0.25 of the time it lands on six, and 0.15 of the time it lands on each of 1, 2, 3, 4 and 5. You choose a die at random, and roll it six times, getting the values 4, 3, 6, 6, 5, 5. What is the probability that the die you chose is the fair die? The outcomes of the rolls are mutually independent.

Exercise 6.4.3: Bayes' Theorem - drug screening.

About

(a) The national flufferball association decides to implement a drug screening procedure to test its athletes for illegal performance enhancing drugs. 3% of the professional flufferball players actually use performance enhancing drugs. A test for the drugs has a false positive rate of 2% and a false negative rate of 4%. In other words, a person who does not take the drugs will test positive with probability 0.02. A person who does take the drugs will test negative with probability 0.04. A randomly selected player is tested and tests positive. What is the probability that she really does take performance enhancing drugs?

Solution

Exercise 6.4.4: Bayes' Theorem - detecting a virus from test results.

About

(a) Assume one person out of 10,000 is infected with HIV, and there is a test in which 2.5% of all people test positive for the virus although they do not really have it. If you test negative on this test, then you definitely do not have HIV. What is the chance of having HIV, assuming you test positive for it?

Source: ADUni, modified by Sandy Irani.

6.5 Random variables

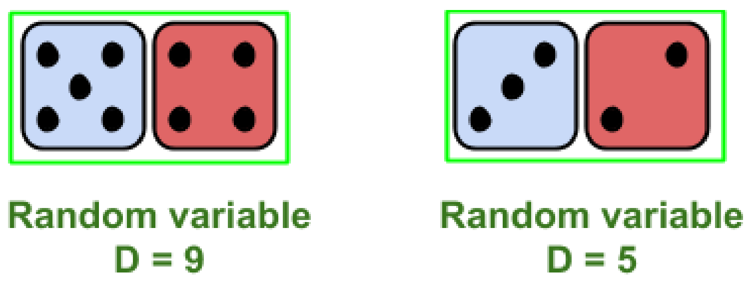
Many problems are concerned with a numerical value associated with the outcome of an experiment. For example, the rules of a game may depend on the sum of the values of a roll of two dice and not on the particular outcome of each individual die. In a random assignment of tasks to processors in a distributed network, the number of tasks assigned to a particular processor is likely to be an important factor in assessing the quality of the assignment. After a few days of observations of the stock market, the amount of money lost or gained is the quantity of interest. A random variable assigns a real number to every outcome of an experiment and therefore provides a way of targeting a particular quantity of interest.

Definition 6.5.1: Random variable.

A ***random variable*** X is a function from the sample space S of an experiment to the real numbers. X(S) denotes the range of the function X.

As an example, consider the experiment in which a red and blue die are rolled. Define the random variable D to be the sum of the two outcomes. The input to the function D is an outcome of the experiment which in the case of the dice is specified by a pair (x, y), where x is the number on the blue die and y is the number on the red die. Then D(x,y) = x + y. The diagram below shows two possible outcomes of the roll of the dice and the value of X for each outcome.

Figure 6.5.1: The random variable X for two outcomes of the dice.



The range of the random variable D is the set {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}. The dice can not sum to a number larger than 12 or less than 2. For every number in the range from 2 through 12, there is an outcome of the dice that sums to that value.

PARTICIPATION ACTIVITY | 6.5.1: Values and ranges of random variables.

1) Consider the sample space of all outcomes of a roll of a blue and a red die. Define M to be the value obtained by subtracting the number on the red die from the number on the blue die. What is the value of M(1, 5)?

Remember that in the ordered pair, the value of the blue die is first and the value of the red die is second.

- 4
- -4
- 6

2) Consider the sample space of all outcomes of a roll of a blue and a red die. Define M to be the value obtained by subtracting the number on the red die from the number on the blue die. What is the range of M?

- {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}
- {-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5}
- {-5, -4, -3, -2, -1, 0}

3) Consider an experiment in which a fair coin is tossed 3 times. Define the random variable A to be the number of flips that come up heads. What is the value of A(HTH)?

- 1
- 2
- 3

4) Consider an experiment in which a fair coin is tossed 3 times. Define the random variable A to be the number of flips that come up heads. What is the range of A?

- {0, 1, 2, 3}
- {1, 2, 3}
- {0, 1, 2}

If X is a random variable defined on the sample space S of an experiment and r is a real number, then X = r is an event. The event X = r consists of all outcomes s in the sample space such that X(s) = r. p(X = r) is the sum of p(s) for all s such that X(s) = r.

Consider again random variable X defined to be the sum of the numbers on a blue and red die after they are thrown. The event that X = 4 is the set {(1,3),(2,2),(3,1)}, and p(X = 4) = 3/36 = 1/12.

PARTICIPATION ACTIVITY | 6.5.2: Random variables and events.

Enter your answer as a simplified fraction. Ex: If your answer is 2/4, then enter: 1/2

1) Consider the sample space of all outcomes of a roll of a blue and a red die. Define M to be the value obtained by subtracting the number on the red die from the number on the blue die. What is the probability that M = 2?

   Remember that in the ordered pair, the value of the blue die is first and the value of the red die is second.

   Check Show answer

2) Consider an experiment in which a fair coin is tossed 3 times. Define the random variable A to be the number of flips that come up heads. What is the probability that A = 1?

   Check Show answer

Definition 6.5.2: Distribution over a random variable.

The ***distribution*** of a random variable is the set of all pairs (r, p(X = r)) such that r ∈ X(S).

The animation below shows the distribution over D, the sum of the numbers on the blue and red dice:

PARTICIPATION ACTIVITY | 6.5.3: Distribution over random variable - the sum of two dice.

Animation captions:

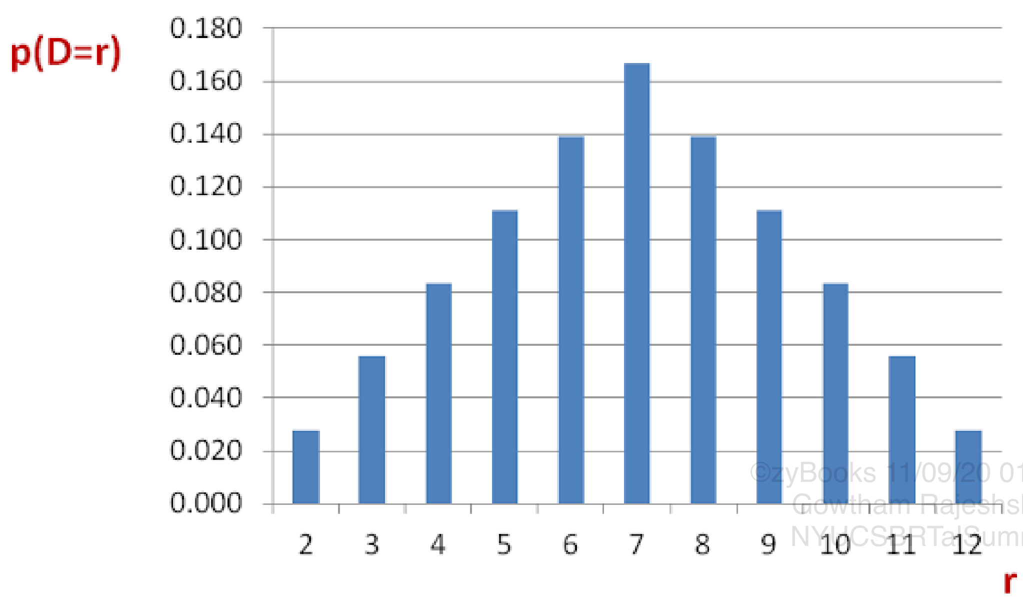
1. The outcomes of tossing a blue and red die are (1, 1), (1, 2), ... ,(6, 6). D = blue die + red die. There is one outcome for which $D = 2$, so $p(D = 2) = \frac{1}{36}$.
2. Add $(2, \frac{1}{36})$ to the distribution for D. Two outcomes have $D = 3$, so $p(D = 3) = \frac{2}{36} = \frac{1}{18}$. Add $(3, \frac{1}{18})$.
3. Three outcomes have $D = 4$, so $p(D = 4) = \frac{3}{36} = \frac{1}{12}$. Add $(4, \frac{1}{12})$.
4. Similarly, add $(5, \frac{1}{9}), (6, \frac{5}{36}), (7, \frac{1}{6}), (8, \frac{5}{36}), (9, \frac{1}{9}), (10, \frac{1}{12}), (11, \frac{1}{18}), (12, \frac{1}{36})$.
5. The resulting set of pairs is the distribution over D: the sum of the numbers on the blue and red dice.

Since a random variable has some value for every outcome in the sample space, the sum of the values p(X = r), over all r ∈ X(S), must equal 1. For example, the probabilities from the animation above all sum to 1:

$$\frac{1}{36} + \frac{1}{18} + \frac{1}{12} + \frac{1}{9} + \frac{5}{36} + \frac{1}{6} + \frac{5}{36} + \frac{1}{9} + \frac{1}{12} + \frac{1}{18} + \frac{1}{36} = 1$$

Sometimes it is helpful to visualize a distribution by plotting the values p(X = r) as a function of r. The diagram below shows the distribution of D for the dice example:

Figure 6.5.2: Histogram of the distribution over the sum of two dice.



PARTICIPATION ACTIVITY | 6.5.4: Distributions over random variables.

1) Consider an experiment in which a fair coin is flipped 3 times. Define the random variable A to be the number of flips that come up heads. Which choice corresponds to the distribution over A?

- (3,1/8), (2,3/8), (1, 3/8), (0,1/8)
- (3,1/4), (2,1/4), (1, 1/4), (0,1/4)
- (3,1/8), (2,3/8), (1, 3/8)

Additional exercises

Exercise 6.5.1: Probabilities and random variables - a roll of two dice.

About

Consider an experiment in which a red die and a blue die are thrown. Let X be the random variable whose value is the product of the numbers on the red and blue dice.

(a) What is the range of X?

(b) What is the probability that X = 6?

Exercise 6.5.2: Distribution over a random variable - aces in a 5-card hand.

About

A hand of 5 cards is dealt from a perfectly shuffled deck of playing cards. Let the random variable A denote the number of aces in the hand.

(a) What is the range of A?

Solution

(b) Give the distribution over the random variable A.

Solution

Exercise 6.5.3: Distribution over a random variable - selecting student council reps.

About

Two student council representatives are chosen at random from a group of 7 girls and 3 boys. Let G be the random variable denoting the number of girls chosen.

(a)
What is the range of G?

(b)
Give the distribution over the random variable G.

Exercise 6.5.4: Probabilities and random variables - redundant storage.

About

In a network of 40 computers, 5 hold a copy of a particular file. Suppose that 7 computers at random fail. Let F denote the number of computers that fail and have a copy of the file.

(a)
What is the range of F?

(b)
What is the probability that F = 2?

Exercise 6.5.5: Distribution over a random variable - 10 coin tosses.

About

A coin is tossed ten times. Let X be the random variable denoting the number of heads minus the number of tails.

(a)
What is the range of X?

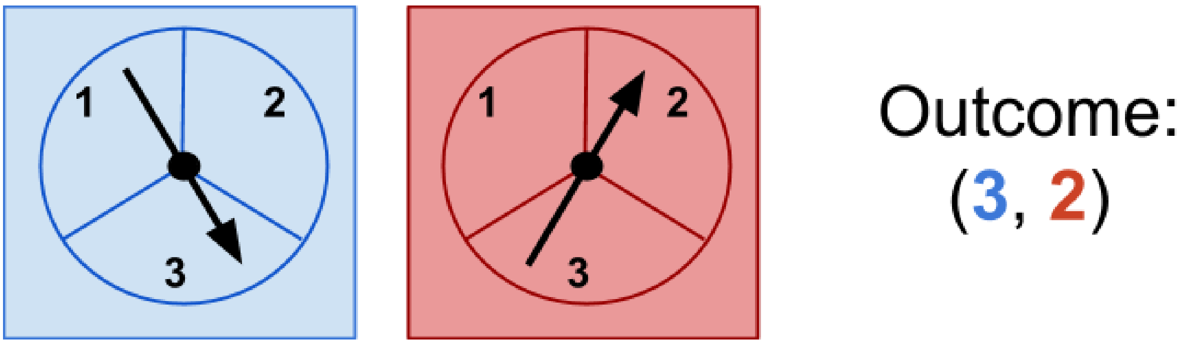
(b)
What is the distribution over X?

6.6 Expectation of a random variable

The distribution over a random variable provides a great deal of information about the outcome of the variable. The expected or average value of a random variable is a useful way to summarize the information in the distribution.

To illustrate expectations of a random variables, we consider an experiment in which a player in a game spins two spinners: a blue spinner and a red spinner. The outcome of each spinner is 1, 2, or 3. The outcome of the experiment is an ordered pair (x, y), where x is the outcome of the blue spinner and y is the outcome of the red spinner. Thus, the sample space of the experiment is {1, 2, 3} × {1, 2, 3}. The experiment is very similar to rolling two dice, but the sample space is smaller to keep the example simpler. Here is a picture of two outcomes of the spinner and the ordered pair representing the outcomes.

Figure 6.6.1: An outcome of the blue and red spinners.



Define random variable M to be the sum of the values of the two spinners. The table below summarizes the distribution over M:

Figure 6.6.2: Distribution over the sum of the spinners.

r	2	3	4	5	6
p(M=r)	1/9	2/9	1/3	2/9	1/9
The event M =r	{(1, 1)}	{(1, 2), (2, 1)}	{(1, 3), (2, 2), (3, 1)}	{(2, 3), (3, 2)}	{(3, 3)}

Here is the definition of the expectation of a random variable:

Definition 6.6.1: Expectation of a random variable.

The ***expected value*** of a random variable is denoted E[X] and is defined as

$$E[X] = \sum_{s \in S} X(s)p(s),$$

where p(s) is the probability of outcomes.

The animation below illustrates how to apply the definition to find E[M] for the two spinners:

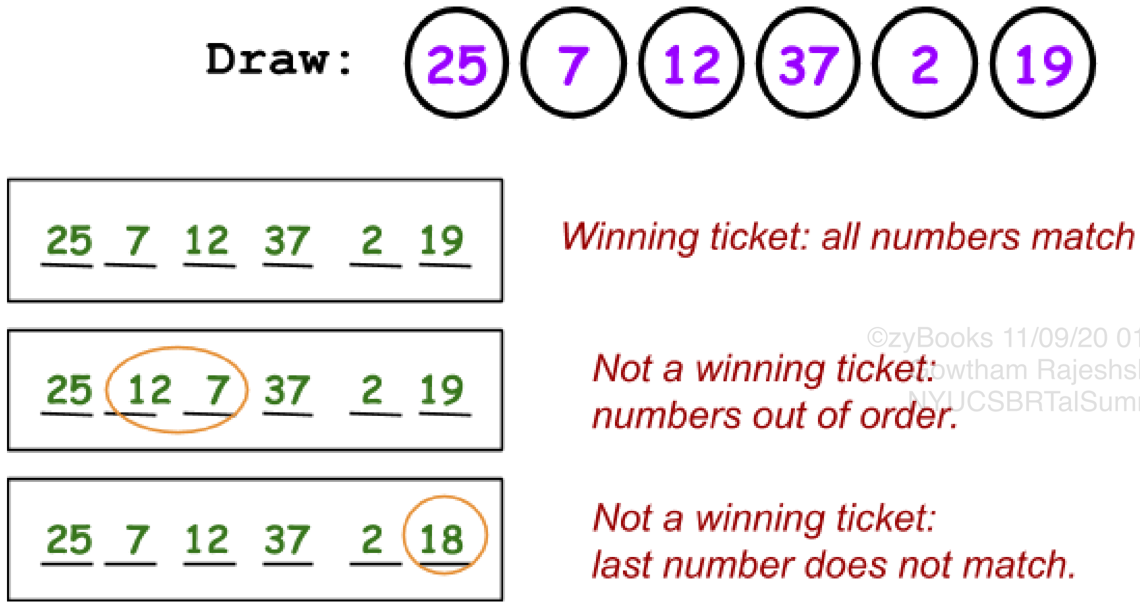
PARTICIPATION ACTIVITY 6.6.1: Calculating the expected value of the sum of the spinners.

Animation captions:

1. The set S of outcomes of spinning a blue and red spinner are {(1, 1), (1, 2), ... ,(3, 3)}. M = sum of red and blue spinner. For outcome $s = (1, 1)$, $M(s) = 1 + 1 = 2$.
2. The value for M can be determined for the rest of the outcomes.
3. Under the uniform distribution, the probability of each outcome is $\frac{1}{|S|} = \frac{1}{9}$. For outcome (1,1), the value of D is 2, and the probability is 1/9. To calculate E[M], start with $2 \cdot \frac{1}{9}$.
4. Add in $M(s) \cdot \frac{1}{9}$ for all the other outcomes s.
5. The result is $\frac{1}{9}(2 + 3 + 4 + 3 + 4 + 5 + 4 + 5 + 6) = \frac{36}{9} = 4$. The expected value of $M = 4$.

Example 6.6.1: Expected winnings from a lottery.

A lottery is run in which every ticket has six numbers. Each of the six numbers is in the range from 1 through 50. The person purchasing a lottery ticket can select the numbers on their ticket. On each Friday of the week, the state lottery commission holds a drawing in which six random numbers are generated. Each number in the range from 1 through 50 is equally likely. In order to win the jackpot, the ticket must match each number selected in order. The diagram below shows an example of the numbers selected in the drawing as well as the winning ticket and two losing tickets.



If a person has a winning ticket, they win \$100,000,000. What is the expected winning from a single ticket?

At the moment before the drawing, the numbers on the ticket are fixed. The experiment consists of the drawing in which the 6 numbers are selected. There are $(50)^6$ possible outcomes, one for each 6-tuple of numbers chosen. Since each number is equally likely, the probability of any single outcome of the whole drawing is $1/(50)^6$. The random variable W evaluates to 100 million for the single outcome that matches the tickets and is 0 for all other outcomes. The expected winning is E[W]:

$$E[W] = (100,000,000) \cdot \left(\frac{1}{50^6}\right) + 0 \cdot \left(1 - \frac{1}{50^6}\right) \approx \frac{10^8}{1.5626 \times 10^{10}} \approx .0064$$

The expected winnings is .0064 dollars or .64 cents. If a ticket costs a dollar, is it a good deal?

PARTICIPATION ACTIVITY 6.6.2: Calculating the expected value of a random variable.

Express the expectation as a simplified fraction (e.g., 0, 3/4 or 1/7, but not 4/8).

1) A fair coin is flipped. Define a random variable X that is 1 if the outcome of the flip is heads and 3 if the outcome of the flip is tails. X(H) = 1, X(T) = 3. What is E[X]?

Check Show answer

2) A fair die is tossed. Define random variable Y to be 1 if the outcome is a six and 0 otherwise. What is E[Y]?

Check Show answer

3) A single share of a stock is purchased for $200. With probability 0.3, the company will meet the deadline for their new product and the value of the stock will go to $300 in the next week. With probability 0.7, the company will miss the deadline and the value of the stock will go down to $100. What is the expected value of the stock next week?

Check Show answer

If the sample space is large, it can be tedious to add up the values of the random variable for each point in the sample space individually. The Theorem below gives a way to calculate the expectation of a random variable by grouping the outcomes according to the value of the random variable. There is a term in the sum for each value r in the range of X. The value of r is multiplied by the probability that the random variable X is equal to r.

Theorem 6.6.1: An alternative way to calculate the expectation of a random variable.

If X is a random variable defined over an experiment with sample space S,

$$E[X] = \sum_{r \in X(S)} r \cdot p(X = r),$$

where X(S) is the range of the function X.

The animation below uses the alternative way to calculate the expectation of a random variable to determine the value of E[M] for the spinner example:

PARTICIPATION ACTIVITY 6.6.3: Alternative way to calculate the expectation of a random variable.

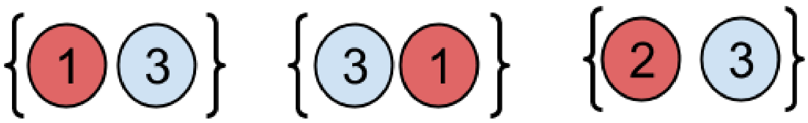
Animation captions:

1. Rearrange the outcomes of the red and blue spinners by value of M.
2. For $r = 2$, $p(M = 2) = \frac{1}{9}$, because there is only one outcome that has $M = 2$. Add in $2 \cdot \frac{1}{9}$.
3. For $r = 3$, $p(M = 3) = \frac{2}{9}$, because there are two outcomes that have $M = 3$. Add in $3 \cdot \frac{2}{9}$.
4. For $r = 4$, $p(M = 4) = \frac{3}{9}$, because there are three outcomes that have $M = 4$. Add in $4 \cdot \frac{3}{9}$.
5. For $r = 5$, $p(M = 5) = \frac{2}{9}$, because there are two outcomes that have $M = 5$. Add in $5 \cdot \frac{2}{9}$.
6. For $r = 6$, $p(M = 6) = \frac{1}{9}$, because there is one outcome that has $M = 6$. Add in $6 \cdot \frac{1}{9}$.
7. The result is $\frac{1}{9}(2 \cdot 1 + 3 \cdot 2 + 4 \cdot 3 + 5 \cdot 2 + 6 \cdot 1) = \frac{36}{9} = 4$. The expected value of $M = 4$.

PARTICIPATION ACTIVITY 6.6.4: Expected value: Balls drawn from a box.

A box contains a total of five balls. The balls are numbered so they are distinct. Two of the balls are colored red and three of the balls are colored blue. The set of balls in the box are pictured below:

2 balls are selected at random from the box. The order in which the balls are chosen does not matter but the numbers on the balls does matter. The diagram below shows three possible outcomes from the experiment. The outcome on the left and the outcome in the middle are considered the same. The outcome on the right is different than the other two.



Express your answer as an integer or simplified fraction (e.g., 0, 3/4 or 1/7, but not 4/8)

1) What is the size of the sample space?

Check Show answer

2) What is the probability that the two balls chose are both red?

Check Show answer

3) What is the probability that one ball is red and one ball is blue?

Check Show answer

4) What is the probability that both chosen balls are blue?

Check Show answer

5) What is the expected number of red balls chosen?

Check Show answer

Additional exercises

Exercise 6.6.1: Random variable expectations - expected number of girls chosen in a group.

About

(a) Two student council representatives are chosen at random from a group of 7 girls and 3 boys. Let G be the random variable denoting the number of girls chosen. What is E[G]?

Exercise 6.6.2: Random variable expectations - roll of a die.

About

(a) Consider a game in which a fair die is rolled. If the die comes up 1, the player wins \$2. If the die comes up 2, the player wins \$1. For all other outcomes, the player loses \$1. What is the expected amount that the player wins or loses? Round to the nearest cent.

Solution

Exercise 6.6.3: Expected winnings in a lottery.

About

(a) A lottery is run in which every ticket has six numbers in a particular order. Each of the six numbers is in the range from 1 through 50. The person purchasing a lottery ticket can select the numbers on their ticket. On each Friday of the week, the state lottery

commission holds a drawing in which six random numbers are generated. Each

commission holds a drawing in which six random numbers are generated. Each number in the range from 1 through 50 is equally likely.
The order of the numbers matters, so if the random numbers selected are (25, 7, 12, 37, 2, 19) then the ticket (7, 25, 12, 37, 20, 19) matches in location 3, 4, and 6. If a ticket matches in all six locations, then the ticket holder wins $100,000,000. If the ticket matches in five of the six locations, then the ticket holder wins $1,000,000. The ticket holder does not win any money for any of the other outcomes. What are the expected winnings?

Solution

Exercise 6.6.4: Expected values of squares.

About

(a) A fair die is rolled once. Let X be the random variable that denotes the square of the number that shows up on the die. For example, if the die comes up 5, then X = 25. What is E[X]?

(b) A fair coin is tossed three times. Let Y be the random variable that denotes the square of the number of heads. For example, in the outcome HTH, there are two heads and Y = 4. What is E[Y]?

6.7 Linearity of expectations

Linearity of expectations says that the expectation of the sum of two random variables is equal to the sum of the expectations. Also the expectation of a constant times a random variable is that same constant times the expectation of the random variable. When carefully applied, linearity of expectations can greatly simplify calculating expectations. Here is a formal statement of the theorem:

Theorem 6.7.1: Linearity of expectations.

If X and Y are two random variables defined on the same sample space S, and c is a real number,

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

The animation below applies linearity of expectations to determine the expectation of D, the sum of the outcomes of a blue and red die.

PARTICIPATION ACTIVITY 6.7.1: Expectation of the sum of two dice.

Animation captions:

1. D = sum of numbers on a blue and red die. B = # on blue die. Red = # on red die. For B = 5 and R = 3, D = B + R = 5 + 3 = 8.
2. $E[D] = E[B+R] = E[B] + E[R]$. Each outcome of a single die happens with probability $\frac{1}{6}$, so $E[B] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6}$.
3. $E[B] = \frac{1}{6}(21) = \frac{7}{2}$.
4. The red die is the same as the blue die, so $E[B] + E[R] = \frac{7}{2} + \frac{7}{2} = 7$.

The formal proof of linearity of expectations is given below:

Proof.
The proof involves plugging in the definition of the expectation and rearranging the sum:

$$E[X+Y] = \sum_{s \in S} p(s)[X(s) + Y(s)] = \sum_{s \in S} [p(s)X(s) + p(s)Y(s)]$$

$$= \sum_{s \in S} p(s)X(s) + \sum_{s \in S} p(s)Y(s) = E[X] + E[Y].$$

The first and last equalities are the definition of expectation. Here is the proof for E[cX]:

$$E[cX] = \sum_{s \in S} p(s)(cX(s)) = c \sum_{s \in S} p(s)X(s) = cE[X]. \blacksquare$$

Example 6.7.1: Expected profit from lemonade sales.

Two friends run a lemonade stand over the summer. It takes 10 lemons to make a batch of lemonade. Usually, lemons are 50 cents apiece, but with probability 1/4, they are on sale for 40 cents apiece. If there is a baseball game at the park where they put up their stand, they will sell $20 worth of lemonade. If there is no baseball game, they will sell $10 worth of lemonade. On a given day, there is a baseball game with probability 1/2. What is the expected profit?

Let R be the random variable denoting their earnings and C the random variable denoting their costs. The profit P = R - C. Using linearity of expectations we have:

$$E[P] = E[R-C] = E[R] + E[-C] = E[R] + (-1)E[C] = E[R] - E[C]$$

It remains to determine the expected earnings and the expected costs. The expected earnings are:

$$E[R] = \left(\frac{1}{2}\right)20 + \left(\frac{1}{2}\right)10 = 15$$

The cost to buy 10 lemons will be $5 if the lemons cost 50 cents apiece. The cost to buy 10 lemons will be $4 if the lemons cost 40 cents apiece. The expected costs are:

$$E[C] = \left(\frac{3}{4}\right)5 + \left(\frac{1}{4}\right)4 = 4.75$$

So the expected profit is $15 - $4.75 = $10.25.

PARTICIPATION ACTIVITY 6.7.2: Dice experiment.

Express the expectation as a simplified fraction (e.g., 0, 3/4 or 1/7, but not 4/8)

1) In the experiment in which a blue and red die are thrown, define M to be the value on the blue die minus the number on the red die. What is E[M]?

Check Show answer

2) Consider a blue spinner that has an outcome of 1, 2, 3, 4, or 5 and a red spinner that has an outcome of 1, 2, or 3. Each spinner is fair so that each outcome is equally likely. Let U be the sum of the outcomes of the two spinners. What is E[U]?

Check Show answer

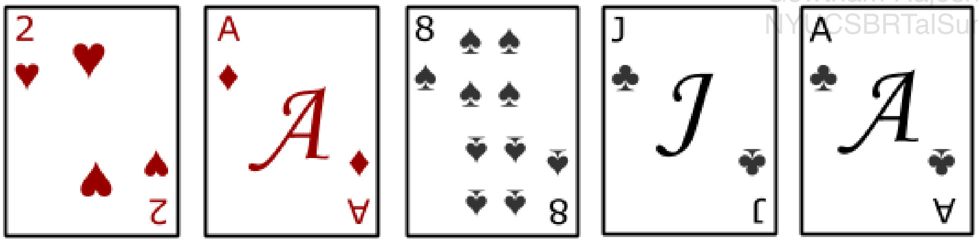
Linearity of expectations can be shown by induction to apply to more than two variables. If $X_1,\ldots,X_n$ are n variables defined on the same sample space, then

$$E\left[\sum_{j=1}^{n} X_j\right] = \sum_{j=1}^{n} E[X_j].$$

In particular, if $X = X_1 + \ldots + X_n$ and all the X_j have the same expectation, then $E[X] = nE[X_1]$.

Return to the 5-card hand dealt from a perfectly shuffled deck. What is the expected number of aces in the hand? One approach is to calculate the probability that the hand contains 0, 1, 2, 3, or 4 aces and apply the formula for the expectation of a random variable. An easier way to calculate the expected number of aces is to imagine that the player looks at his cards one at a time. Let A be the number of aces in the hand. Define random variable A_j to be 1 if the j^{th} card he looks at is an ace and 0 otherwise (for j = 1, 2, 3, 4, 5). In any particular hand, the random variable A is the sum of the A_j's. The diagram below gives an example of a particular hand and the corresponding random variables:

Figure 6.7.1: A 5-card hand and the number of aces.



Each A_j has the same distribution, so $E[A] = 5E[A_1]$. To determine A_1, we just need the probability that the first card is an ace. Since there are 4 aces and a total of 52 cards in the deck, $p(A_1 = 1) = 4/52 = 1/13$. Therefore:

$$E[A_1] = \left(\frac{1}{13}\right)1 + \left(1 - \frac{1}{13}\right)0 = \frac{1}{13}$$

The expected number of aces in a 5 card hand is $E[A] = 5 \cdot E[A_1] = 5/13$.

In the 5-card hand example, the random variables A_j are not independent. If you know that the first card was an ace, then it is less likely that the second card will be an ace. If you know that the first card was not an ace, then the likelihood that the second card will be an ace increases slightly. However, linearity of expectations does not require that the random variables be independent. The expectation of each A_j is determined without regard to the others.

Example 6.7.2: Load balancing.

In load balancing, tasks are assigned to processors in a network so as to even out the load among the processors. Consider a simple load balancing policy in which each incoming task is assigned to a random processor. If there m tasks and n processors, what is the expected load of (number of tasks assigned to) a particular processor P?

The sample space consists of all the possible assignment of jobs to processors. If the tasks are distinct, then an assignment is specified by the sequence of assignments for each task: $p_1,...,p_m$, where p_j is the processor to which task j is assigned. There are n possibilities for each p_j, so the sample space has n^m outcomes. Instead of analyzing the whole space at once, define m random variables $X_1,\ldots,X_m$. X_j is 1 if task j is assigned to processor P and otherwise $X_j = 0$.

$$X_j = \begin{cases} 1 \text{ if task j assigned to P} & p(X_j=1) = 1/n \\ 0 \text{ if task j } \underline{not} \text{ assigned to P} & p(X_j=0) = 1 - 1/n \end{cases}$$

If X is the random variable denoting the number of tasks assigned to processor P, then for each assignment, $X = X_1 + \ldots + X_m$. All the X_j's have the same expectation, so

$$E[X] = E[X_1 + \cdots + X_m] = E[X_1] + \cdots + E[X_m] = mE[X_1]$$

A particular task is assigned to P with probability 1/n since each processor is chosen uniformly at random. Therefore

$$E[X_1] = \left(\frac{1}{n}\right)1 + \left(1 - \frac{1}{n}\right)0 = \frac{1}{n}$$

$E[X] = m\,E[X_1] = m/n$.

PARTICIPATION ACTIVITY | 6.7.3: Linearity of expectations.

Express the expectation as an integer or simplified fraction (e.g., 0, 3/4 or 1/7, but not 4/8)

1) Suppose a 5-card hand is dealt to each of 2 players on a team. What is the expected total number of aces the team has?

Check   Show answer

2) A fair coin is flipped 9 times. What's the expected number of heads?

Check   Show answer

Additional exercises

Exercise 6.7.1: Expected number of hearts in a 5-card hand.

About

(a) A 5-card hand is dealt from a perfectly shuffled deck so that each 5-card hand is equally likely. What is the expected number of hearts in the hand?

Solution

Exercise 6.7.2: Expected number of copies of a file.

About

(a) In a network of 40 computers, 5 hold a copy of a particular file. Suppose that 7 computers at random fail. Let F denote the number of computers that fail and have a copy of the file. What is E[F]?

Exercise 6.7.3: Expected values - birthday matches.

About

(a) In this problem, assume that the probability that a person is born on a given day is 1/365. (For simplicity, ignore Feb 29.) In a group of 100, what is the expected number of pairs of people who have the same birthday?

Solution

Exercise 6.7.4: Expected values - matching coats.

About

(a) A class of 10 students hang up their coats when they arrive at school. Just before recess, the teacher hands one coat selected at random to each child. What is the expected number of children who get his or her own coat?

Exercise 6.7.5: Linearity of expectations - selecting teams.

About

(a) 10 kids are randomly grouped into an A team with five kids and a B team with five kids. Each grouping is equally likely. One kid in the group, Alex, has three friends who are also in the group whose names are Jose, Carl, and Rajiv. Let F be a random variable denoting the number of friends Alex has on his team. What is E[F]?

Exercise 6.7.6: Linearity of expectations - random permutations.

About

A family with a mother, father, two daughters and three sons lines up in a random order for a photo.

(a) Let D be a random variable denoting the number of daughters who are standing next to the mother. What is E[D]?

Solution

(b) Let N be a random variable denoting the number of sons who are standing next to the mother. What is E[N]?

6.8 Bernoulli trials and the binomial distribution

A ***Bernoulli trial*** is an experiment with two outcomes: ***success*** and ***failure***. In a sequence of independent Bernoulli trials, called a ***Bernoulli process***, the outcomes of the repeated experiments are assumed to be mutually independent and have the same probability of success and failure. Usually the probability of success is denoted by the variable p, and the probability of failure (1 - p) is denoted by the variable q.

Bernoulli processes model many important phenomena in computer science. For example, errors in transmitting bits over a communication channel are commonly modeled as a Bernoulli process. The event that a bit is transmitted correctly is a success and the event that the bit is flipped in transmission is a failure. The event that an individual bit is transmitted correctly is independent from all the other bits. An error-correcting code is a process of adding redundant data to a message so that transmission errors can be corrected. A larger number of errors can be corrected with more complex codes, so it is important to know the likelihood that a transmission will result in a given number of errors in order to know how complex a code needs to be. The diagram below gives some possible outcomes of transmitting the string 11111111 and their corresponding probabilities.

Figure 6.8.1: Outcomes and probabilities from transmitting the string 11111111.

10011110 *probability:* $pqqppppq = p^5q^3$

11111110 *probability:* $pppppppq = p^7q$

01111111 *probability:* $qppppppp = p^7q$

Usually the number of failures or successes in a sequence of trials is more important than the order in which they occur. For example, in transmitting a binary string, the important factor is how many errors occurred, not which particular bits were flipped. The number of successes in a sequence defines a random variable over the sample space of all outcomes. The animation below shows the probability that there are 3 successes in a sequence of 5 Bernoulli trials.

PARTICIPATION

ACTIVITY | 6.8.1: Probability of 3 successes in 5 Bernoulli trials.

Animation captions:

1. 5 Bernoulli trials. For each trial, s = success and f = failure. The sample space of 5 trials is $\{s,f\}^5$. Outcomes with 3 successes = {ffsss, fsfss, ... , sssff}.
2. Number of outcomes with 3 successes = $\binom{5}{3}$.
3. The probability of each outcomes with 3 successes is p^3q^2.
4. The probability of 3 successes = $\binom{5}{3}p^3q^2$.

The general theorem can be stated as follows:

Theorem 6.8.1: Bernoulli trial probabilities.

The probability of exactly k successes in a sequence of n independent Bernoulli trials, with probability of success p and probability of failure q = 1 - p is

$$\binom{n}{k}p^kq^{n-k}.$$

PARTICIPATION ACTIVITY | 6.8.2: Probability of k successes in a sequence of independent Bernoulli trials.

1) A biased coin is flipped 10 times. In a single flip, the probability of heads is 2/7 and the probability of tails is 5/7. What is the probability that 3 of the coin flips come up tails (and therefore 7 come up heads)?

- $\binom{10}{3}(5/7)^3(2/7)^7$
- $\binom{10}{3}(2/7)^3(5/7)^7$
- $\binom{7}{3}(5/7)^3(2/7)^7$

2) A single die is rolled 5 times. What is the probability that the die comes up 4 exactly 2 times?

- $\binom{5}{2}(1/3)^2(2/3)^3$
-

$\binom{5}{4}(1/6)^2(5/6)^3$

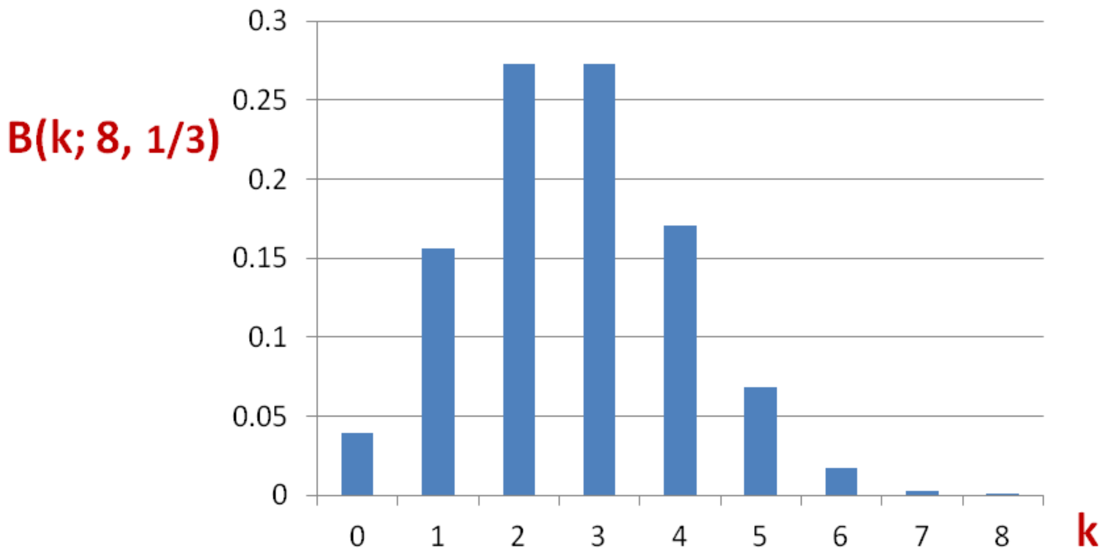
- $\binom{5}{2}(1/6)^2(5/6)^3$

The distribution over the random variable defined by the number of successes in a sequence of independent Bernoulli trials is called the ***binomial distribution***. The probability that the number of successes is k in a sequence of length n with probability of success p is denoted by b(k; n, p). By the theorem above:

$$b(k; n, p) = \binom{n}{k} p^k q^{n-k},$$

where q = 1 - p. The diagram below shows the distribution b(k; n, p) for n = 8 and p = 1/3.

Figure 6.8.2: The binomial distribution for n = 8 and p = 1/3.



The range of the random variable denoting the number of successes in a sequence of n Bernoulli trials is 0 through n. Since the values of b(k; n, p) are a probability distribution over the possible values for k, the probabilities should sum to 1 as k ranges from 0 through n:

$$\sum_{k=0}^{n} b(k; n, p) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$$

The second equality follows from the Binomial Theorem (covered on the material on the binomial coefficients) and the final equality follows from the fact that p + q = 1.

PARTICIPATION ACTIVITY 6.8.3: The binomial distribution.

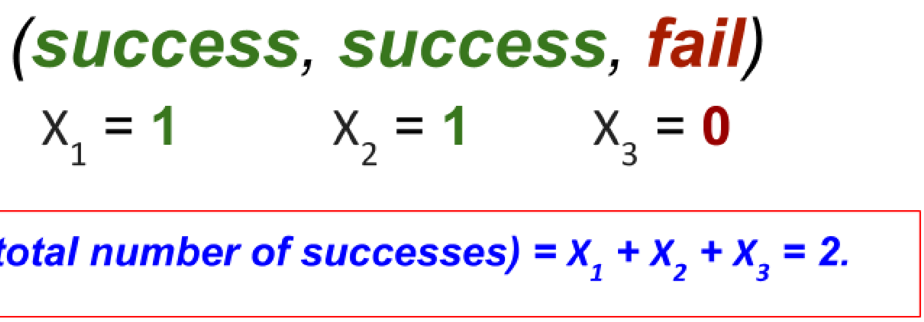
1) What is the probability of 6 successes in a sequence of 13 independent Bernoulli trials if the probability of failure is 1/3?

- b(6; 13, 1/3)
- b(6; 13, 2/3)
- b(13; 6, 1/3)

The expected number of successes

Linearity of expectations is the easiest way to determine the expected number of successes in a sequence of independent Bernoulli trials. Let K be the random variable denoting the number of successes in n Bernoulli trials with probability of success p. Define n additional random variables X_j for j = 1, ..., n. X_j = 1 if the j^{th} trial is a success and X_j = 0 if the j^{th} trial is a failure. The sum of the X_j for any sequence of trials is the number of successes. The diagram below shows a possible outcome for 3 Bernoulli trials and the values of the corresponding random variables.

Figure 6.8.3: An outcome of 3 Bernoulli trials.



The expectation of each X_j is

$$E[X_j] = p \cdot 1 + (1 - p) \cdot 0 = p$$

Thus the expected number of successes n Bernoulli trials with probability of success p is

$$E[K] = E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n] = np$$

In a sequence of 8 independent Bernoulli trials in which the probability of success is 2/3, the expected number of successes is 8(2/3) = 16/3.

PARTICIPATION ACTIVITY 6.8.4: Expected number of successes in a sequence of Bernoulli trials.

1) A biased coin comes up heads with probability 1/3. What is the expected number of heads in 100 tosses of the

coin? Give your answer as an integer or a fraction in lowest terms.

Check Show answer

Additional exercises

Exercise 6.8.1: Probability of manufacturing defects.

About

The probability that a circuit board produced by a particular manufacturer has a defect is 1%. You can assume that errors are independent, so the event that one circuit board has a defect is independent of whether a different circuit board has a defect.

(a) What is the probability that out of 100 circuit boards made *exactly* 2 have defects?

(b) What is the probability that out of 100 circuit boards made *at least* 2 have defects?

(c) What is the expected number of circuit boards with defects out of the 100 made?

(d) Now suppose that the circuit boards are made in batches of two. Either both circuit boards in a batch have a defect or they are both free of defects. The probability that a batch has a defect is 1%. What is the probability that out of 100 circuit boards (50 batches) at least 2 have defects? What is the expected number of circuit boards with defects out of the 100 made? How do your answers compared to the situation in which each circuit board is made separately?

Exercise 6.8.2: Error probability in a sequence of repeated tests.

About

(a) A test for a disease gives an answer "Positive" or "Negative". Because of experimental error, the test gives an incorrect answer with probability 1/4. That is, if the person has the disease, the outcome is negative with probability 1/4 and if the person does not have the disease, the outcome is positive with probability 1/4. You can assume that the probability an error is made in one test is independent of the outcomes of any other tests. In order to improve the probability of getting the correct answer, the test is run n times and the majority outcome is used as the final answer. If n = 3, what is the probability of getting the wrong answer? What about n = 5?

Solution

Exercise 6.8.3: Detecting a biased coin.

About

A gambler has a coin which is either fair (equal probability heads or tails) or is biased with a probability of heads equal to 0.3. Without knowing which coin he is using, you ask him to flip the coin 10 times. If the number of heads is at least 4, you conclude that the coin is fair. If the number of heads is less than 4, you conclude that the coin is biased.

(a) What is the probability you reach an incorrect conclusion if the coin is fair?

Solution

(b) What is the probability that you reach an incorrect conclusion if the coin is biased?

Exercise 6.8.4: Error-correcting codes and the probability of transmitting a message without errors.

About

(a) A communication channel flips each transmitted bit with probability 0.02. The event that one bit is flipped is independent of the event that any other subset of the bits is flipped.
A 100-bit message is sent across the communication channel and an error-correcting scheme is used that can correct up to three errors but expands the length of the message to 110 bits. What is the probability that there is an error in the message after the message has been transmitted?