MAT 142 College Mathematics

Probability

Terri L. Miller & Elizabeth E. K. Jones

Contents

Basic Probability

Vocabulary. In order to discuss probability we will need a fair bit of vocabulary.

Probability is a measurement of how likely it is for something to happen.

An ***experiment*** is a process that yields an observation.

Example 1. Experiment 1. The experiment is to toss a coin and observe whether it lands showing heads or tails.

Experiment 2. The experiment is to roll a die and observe the number of spots.

An ***outcome*** is the result of an experiment.

Example 2. Experiment 1. One possible outcome is heads, which we will designate H.

Experiment 2. One possible outcome is 5.

Sample space is the set of all possible outcomes, we will usually represent this set with S.

Example 3. Experiment 1. $S = \{H, T\}$

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

An ***event*** is any subset of the sample space. Events are frequently designated as E or E_1, E_2, etc. if there are more than one.

Example 4. Experiment 1. One possible event is landing heads up. $E = \{H\}$.

Experiment 2. One event might be getting an even number, a second event might be getting a number greater than or equal to 5. $E_1 = \{2, 4, 6\}, \quad E_2 = \{5, 6\}$.

A ***certain event*** is guaranteed to happen, a "sure thing". An ***impossible event*** is one that cannot happen.

Example 5. From our Experiment 1. If the event E_1 is landing heads up or landing tails up, then $E_1 = \{H, T\}$ and we can see that E_1 is a certain event. If the event E_2 is getting a 6, then $E_2 = \{\}$ and we can see than E_2 is impossible.

Outcomes in a sample space are said to be ***equally likely*** if every outcome in the sample space has the same chance of happening, *i.e.* one outcome is no more likely than any other.

Basic Computation of a Probability. If E is an event in an equally likely sample space, S, then the probability of E, denoted $p(E)$ is computed

$$\boxed{p(E) = \frac{n(E)}{n(S)}}$$

Example 6. Experiment 1. $E_1 = \{H, T\}$, so $p(E_1) = \dfrac{n(E_1)}{n(S)} = \dfrac{2}{2} = 1$.

Experiment 2. $E_2 = \{5, 6\}$, so $p(E_2) = \dfrac{n(E_2)}{n(S)} = \dfrac{2}{6} = \dfrac{1}{3}$.

You should note that since $n(\emptyset) = 0$, $p(\emptyset) = 0$ and since $\frac{n(S)}{n(S)} = 1$, $p(S) = 1$.

Relative frequency is the number of times a particular outcome occurs divided by the number of times the experiment is performed.

Example 7. Suppose you toss a fair coin 10 times and observe the results: heads occurred 6 times and tails occurred 4 times. In this case the relative frequency of heads would be $\frac{6}{10} = 0.6$.

Now suppose you toss a fair coin 1000 times and observe the results: heads occurred 528 times and tails occurred 472 times. Now the relative frequency of heads is $\frac{528}{1000} = 0.528$.

Something which can be observed from the previous two experiments is called the Law of Large Numbers. ***The Law of Large Numbers***: If an experiment is repeated a large number of times, the relative frequency tends to get closer to the probability.

Example 8. In the previous example, we observed the outcome heads. This was a fair coin so we know $p(H) = 0.5$. Notice that when we performed the experiment 1000 times, the relative frequency was closer to 0.5 than when we performed the experiment 10 times.

Some Applications. Probability is used in genetics. When Mendel experimented with pea plants he discovered that some genes were dominant and others were recessive. This means that given one gene from each parent, the dominant trait will show up unless a recessive gene is received from each parent. This is often demonstrated by using a "Punnett square". Here is a typical Punnett square:

	R	**R**
w	wR	wR
w	wR	wR

The letters along the top of the table represent the gene contribution from one parent and the letters down the left-hand side of the table represent the gene contribution from the other parent. Each cell of the table contains a genetic combination for a possible offspring. This particular Punnett square represent the crossing of a pure red flower pea with a pure white flower pea. The offspring will each inherit one of each gene; since red is dominant here, all offspring will be red.

Example 9. Suppose we cross a pure red flower pea plant with one of the offspring that has one of each gene. The resulting Punnett square would be

	R	**R**
w	wR	wR
R	RR	RR

From this we can see that the probability of producing a pure red flower pea is $\frac{2}{4} = \frac{1}{2}$, and we can see that each offspring would produce red flowers.

Example 10. This time we will cross two of the offspring which have one of each gene. The Punnett square is

	w	**R**
w	ww	wR
R	Rw	RR

Here we can find that the probability of an offspring producing white flowers is $\frac{1}{4}$.

Example 11. Sickle cell anemia is inherited. This is a co-dominant disease. A person with two sickle cell genes will have the disease while a person with one sickle cell gene will have a mild anemia called sickle cell trait. Suppose a healthy parent (no sickle cell gene) and a parent with sickle cell trait have children. Use a Punnett square to determine the following probabilities.

(1) the child has sickle cell
(2) the child has sickle cell trait
(3) the child is healthy

Solution: We will use S for the sickle cell gene and N for no sickle cell gene. Our Punnett square is

	N	**N**
S	SN	SN
N	NN	NN

(1) No offspring have two sickle cell genes so this probability is 0.
(2) We see that two of the offspring will have one of each gene, so this probability is $\frac{2}{4} = \frac{1}{2}$.
(3) Two of the children will have no sickle cell genes, this probability is $\frac{2}{4} = \frac{1}{2}$.

Basic Properties of Probability

We will need to start with one more new term. Two events are ***mutually exclusive*** if they cannot both happen at the same time, *i.e.* $E \cap F = \emptyset$. So, if $p(A \cap B) = 0$ then A and B are mutually exclusive,

We have the following rules for probability:

Basic Probability Rules
$p(\emptyset) = 0$,
$p(S) = 1$,
$0 \le p(E) \le 1$

Example 12. Roll a die and observe the number of dots; $S = \{1, 2, 3, 4, 5, 6\}$. Given E is the event that you roll a 15, F is the event that you roll a number between 1 and 6 inclusive, and G is the event that you roll a 3. Find $p(E)$, $p(F)$, and $p(G)$.

Solution: Since 15 is not in the sample space $E = \emptyset$ and $p(E) = 0$. $F = S$, so $p(F) = 1$. Lastly, $G = \{3\}$, so $p(G) = \frac{n(G)}{n(S)} = \frac{1}{6}$.

Example 13. Roll a pair of dice, one red and one white. Find the probabilities:

(1) the sum of the pair is 7
(2) the sum is greater than 9
(3) the sum is not greater than 9
(4) the sum is greater than 9 and even
(5) the sum is greater than 9 or even
(6) the difference is 3

Solution: This sample space is too large to list as there are $6 \times 6 = 36$ outcomes. Hence, we will just count up the number of ways to achieve each event and them divide by 36.

(1) The pairs that sum to 7 are $\{(6, 1), (1, 6), (5, 2), (2, 5), (4, 3), (3, 4)\}$, hence this probability is $\frac{6}{36} = \frac{1}{6}$.
(2) Sums greater than 9 are 10, 11 and 12, the pairs meeting this condition are $\{(6, 4), (4, 6), (5, 5), (6, 5), (5, 6), (6, 6)\}$, this gives the probability as $\frac{6}{36} = \frac{1}{6}$.
(3) The sum is not greater than 9 would be all of those not included in the previous part, so there must be $36 - 6 = 30$ outcomes in this set. Hence, this probability is $\frac{30}{36} = \frac{5}{6}$.
(4) Those with a sum greater than 9 and even would be those that sum to 10 or 12, $\{(6, 4), (4, 6), (5, 5), (6, 6)\}$. Since this set contains 4 outcomes, this probability is $\frac{4}{36} = \frac{1}{9}$.
(5) Half of the pairs would sum to an even number and part (3) gives us that there are 6 with sums greater than 9. Note that (4) gives us the number in the intersection. Using the counting formula $n(A \cup B) = n(a) + n(B) - n(A \cap B)$, this set contains $18 + 6 - 4 = 20$ and the probability is $\frac{20}{36} = \frac{5}{9}$.
(6) Those pairs whose difference is 3 are $\{(6, 3), (3, 6), (5, 2), (2, 5), (4, 1), (1, 4)\}$, hence the probability is $\frac{6}{36} = \frac{1}{6}$.

We have two more formulae for probability that will be useful.

Basic Probability Rules
$p(E) + p(E') = 1,$
$p(E \cup F) = p(E) + p(F) - p(E \cap F)$

Use of Venn Diagrams for Probability. It is often helpful to put the information given in a Venn Diagram to organize the information and answer questions. The following is an example of such a case.

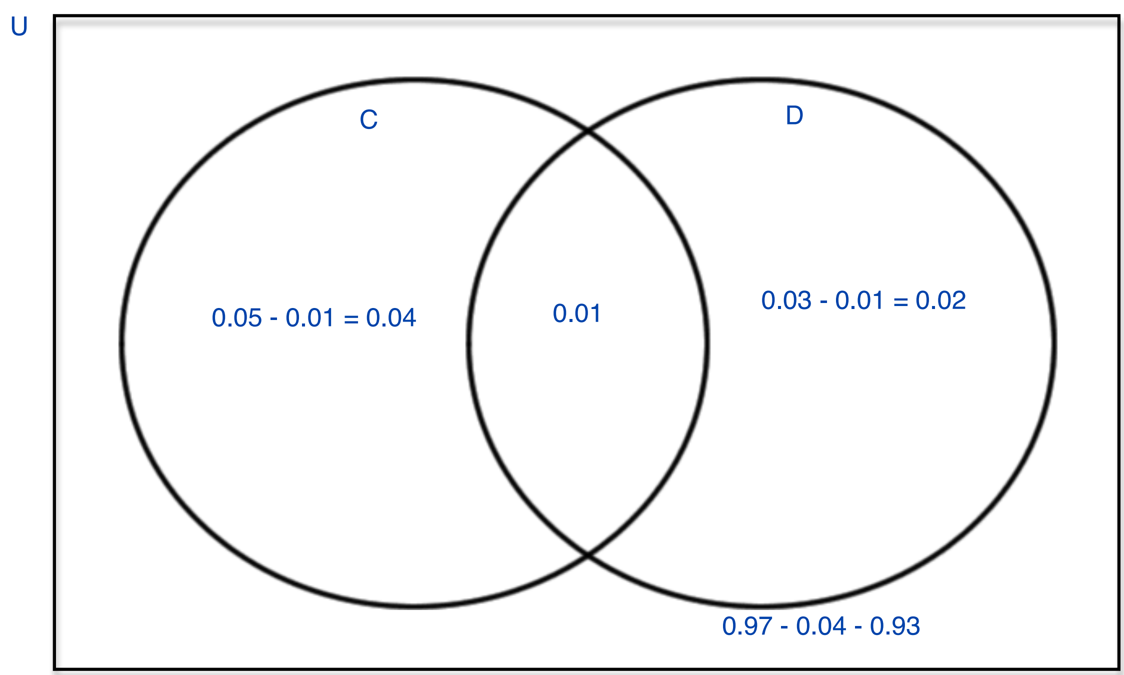
Example 14. Zaptronics makes CDs and their cases for several music labels. A recent sampling indicated that there is a 0.05 probability that a case is defective, a 0.97 probability that a CD is not defective, and a 0.07 probability that at lease one of them is defective.

(1) What is the probability that both are defective?
(2) What is the probability that neither is defective?

Solution: We start by writing down the information given. Let C represent a defective case and let D represent a defective CD. Then we are given:

$$p(C) = 0.05, \qquad p(D') = 0.97, \qquad p(C \cup D) = 0.07$$

Using the property $p(E) + p(E') = 1$, we get $p(D) = 1 - p(D') = 1 - 0.97 = 0.03$. Using the property $p(E \cup F) = p(E) + p(F) - p(E \cap F)$, we get $p(C \cap D) = p(C) + p(D) - p(C \cup D) = 0.05 + 0.03 - 0.07 = 0.01$. Put this together in a Venn diagram and recall that $p(U) = 1$.



From this we can see the answers we need.

(1) This is the middle football shape so, 0.01
(2) This is everything outside of the circles, so 0.93

Odds. The ***odds for*** an event, E, with equally likely outcomes are:

$$o(E) = n(E) : n(E')$$

and the ***odds against*** the event E are:

$$o(E') = n(E') : n(E).$$

Note, these are read as a to b.

Example 15. Toss a fair coin twice and observe the outcome. The sample space is

$$S = \{HH, HT, TH, TT\}.$$

If E is getting two tails, then

$$o(E) = 1 : 3,$$

the odds for E are 1 to 3. Also

$$o(E') = 3 : 1,$$

the odds against E are 3 to 1.

Example 16. The table below shows the percent of voters who voted for Clinton, Dole and Perot in the 1996 presidential election broken down by age.

Age	Clinton	Dole	Perot	Total
18 - 24	4.95	3.15	0.99	9
25 - 29	4.32	2.88	0.8	8
30 - 49	22	18.04	3.96	44
50 - 64	10.81	10.35	1.84	23
65 & over	8	7.04	0.96	16
Total	50.08	41.46	8.55	100

Table data from Survey by Voter News Service, a consortium ABC News, CBS News, CNN, FOX News, NBC News and the Associated Press (www.ropercenter.uconn.edu)

(1) Find the probability that a voter voted for Clinton and was between 30 and 49 years old?
(2) Find the probability that a voter voted for Clinton or was between 30 and 49 years old?
(3) Find the probability that a voter didn't support Clinton and is not between 30 and 49 years old?
(4) Find the probability that a voter did not support Dole or was not between 25 and 29 years old?

Solution:

(1) For this one, we want to look at the cell of the table where the column of voters for Clinton intersects with the row for voters between 30 and 49 years old. This percentage is 22. We would then divide this by the total to find the probability. This tells us that the probability a voter voted for Clinton and was between 30 and 49 years old is $p(E \cap F) = \frac{22}{100} = .22$.

(2) This question is an OR question. This means that we can use the formula $p(E \cup F) = p(E) + p(F) - p(E \cap F)$ If we let E be voting for Clinton and F be being between 30 and 49 years old, the $p(E) = \frac{50.08}{100}$ and the $p(F) = \frac{44}{100}$. We found $p(E \cap F)$ in part (1). When we plug all this into the formula, we get

$$p(E \cup F) = p(E) + p(F) - p(E \cap F) = \frac{50.08}{100} + \frac{44}{100} - \frac{22}{100} = 0.7208$$

Thus the probability that a voter voted for Clinton or was between 30 and 49 years old is 0.7208

(3) For this question, we can use the formula $p(E') = 1 - p(E)$. We remember from Sets and Counting de Morgan's laws, that $(A' \cap B') = (A \cup B)'$. It is easy for us to find the probability a voter voted for Clinton or is between 30 and 49 years old. We did this in part (2) of this problem. Thus the probability that a voter did not vote for Clinton and was not between 30 and 49 years old is $1 - 0.7208 = 0.2792$.

(4) For this problem, we will again use de Morgan's laws, that $(A' \cup B') = (A \cap B)'$. It is easy for us to find the probability a voter voted for Dole and is between 25 and 29 years old. This is the cell in the table where the Dole column intersects the 25 to 29 year old row. This probability is $\frac{2.88}{100} = 0.0288$. Thus the probability that a voter did not vote for Dole or was not between 25 and 29 years old is $1 - 0.0288 = 0.9712$.

Conditional Probability

We will introduce the idea of conditional probability with an example.

Example 17. Two coins are tossed. Event E is getting exactly one tail. Event F is getting at least one tail. The sample space for the experiment is

$$S = \{HH, HT, TH, TT\}.$$

The event

$$E - \{HT, TH\}$$

has probability $p(E) = 2/4$.

If I tell you that at least one of the coins is showing a tail, then you know that event F has occurred and that the outcome HH is no longer possible. Hence, we effectively reduce the size of the sample space to

$$F = \{HT, TH, TT\}.$$

With F as the sample space, the probability of event $E = 2/3$. Thus, the probability of one event occurring is changed by knowing that another event has already occured. This is conditional probability.

Conditional probability is a probability that is based on knowing that some event within a sample space has already occurred. The notation for the probability of the event E occurring when it is known that the event F has occurred is $p(E|F)$ and is read "the probability of E given F". The formula for computing the conditional probability is

$$\boxed{p(E|F) = \frac{p(E \cap F)}{p(F)}.}$$

Example 18. In our previous example we found $p(E|F) = 2/3$ by reducing the size of our sample space and using the basic probability formula. We could have found it using the formula.

$$E \cap F = \{HT, TH\} \quad \text{so} \quad p(E \cap F) = 1/2$$

Using the formula, we get

$$p(E|F) = \frac{1/2}{3/4} = \frac{1}{2} \times \frac{4}{3} = \frac{2}{3},$$

which is exactly the same value we got by counting.

Example 19. For this example, we will use the same table that shows the percent of voters who voted for Clinton, Dole and Perot in the 1996 presidential election broken down by age that we used in example 16.

Age	Clinton	Dole	Perot	Total
18 - 24	4.95	3.15	0.99	9
25 - 29	4.32	2.88	0.8	8
30 - 49	22	18.04	3.96	44
50 - 64	10.81	10.35	1.84	23
65 & over	8	7.04	0.96	16
Total	50.08	41.46	8.55	100

Table data from Survey by Voter News Service, a consortium ABC News, CBS News, CNN, FOX News, NBC News and the Associated Press (www.ropercenter.uconn.edu)

(1) Find the probability that a voter voted for Perot and was between 30 and 49 years old?
(2) Find the probability that a voter voted for Perot if they were between 30 and 49 years old?
(3) Find the probability that a voter who voted for Perot was between 30 and 49 years old?

Solution:

(1) For this one, we want to look at the cell of the table where the column of voters for Perot intersects with the row for voters between 30 and 49 years old. This percentage is 3.96. We would then divide this by the total to find the probability. This tells us that the probability a voter voted for Perot and was between 30 and 49 years old is $p(E \cap F) = \frac{3.96}{100} = 0.0396$.
(2) For this question, we know that the voter was between 30 and 49 years old. This means that the denominator in our probability will not be the total of 100 but instead will be the total for the 30 to 49 year old of 44. For the numerator, we are looking for the intersection of those voters who voted for Perot and were between the ages of 30 and 49 years old. We found this in part (1) of this example. Thus probability that a voter voted for Perot if they were between 30 and 49 years old is $\frac{3.96}{44} = 0.09$.
(3) For this question, we know that the voter voted for Perot. This mens that we will use the total percent of Perot voters as the denominator in our probability.For the numerator, we are looking for the intersection of those voters who voted for Perot and were between the ages of 30 and 49 years old. We found this in part (1) of this example. Thus probability that a voter voted for Perot if they were between 30 and 49 years old is $\frac{3.96}{8.55} = 0.4632$.

Notice that although parts (2) and (3) above were similar, the conditions in the questions were different and thus the probabilities were different.

Associated with conditional probability is the idea of independence of two events. Two events are ***independent events*** if knowing that one of them has occurred does not change the probability that the other occurs; $p(E|F) = p(E)$.

Example 20. You are rolling a pair of dice two times. What is the probability of the sum of the dice on the second roll is 6 given that the sum of the dice on the first roll was 7? Solution: Since what happens on the first roll does not affect what happens on the second roll. These two events are independent. Thus the probability of the sum of the dice on the second roll being a 6 is $\frac{5}{36}$

Example 21. There is a box of 24 donuts. There are 8 lemon cream, 10 custard filled, 3 chocolate cream, and 3 glazed donuts. You take two donuts from the box and eat them.

(1) What is the probability that the second donut eaten is custard filled given that the first donut was custard filled?
(2) What is the probability that the second donut eaten is custard filled given that the first donut was lemon cream?

Solution:

(1) In this question, the two event are not independent since once a donut is eaten, it is gone. For picking the second donut, there are only 9 custard filled donuts remaining in the box that could be picked and there are only 23 donuts remaining in the box. Thus the probability that the second donut eaten is custard filled given that the first donut was custard filled is $\frac{9}{23}$.
(2) In this question, the two event are again not independent. For picking the second donut, there are still 10 custard filled donuts remaining in the box that could be picked and there are only 23 donuts remaining in the box. Thus the probability that the second donut eaten is custard filled given that the first donut was lemon cream is $\frac{10}{23}$.

Expected Value

Once again, we will need some new vocabulary. The ***payoff*** is the value of an outcome.

The ***expected value*** of an event is a long term average of payoffs of an experiment. If the experiment were repeated often enough, the actual profit/loss will get close to the expected value. To compute the expected value, you first multiply each payoff by the probability of getting that payoff. Once you have done all of the multiplications, you add your results together. This sum is the expected value of an experiment.

Let $m_1, m_2, m_3, \ldots, m_k$ be the payoffs associated with the k outcomes of an experiment. Let $p_1, p_2, p_3, \ldots, p_k$, respectively, be the probabilities of those outcomes. Then the expected value is found as follows.

$$\boxed{\textbf{Expected Value:} \quad E = m_1p_1 + m_2p_2 + m_3p_3 + \cdots + m_kp_k}$$

In mathematics, the capital greek letter sigma, Σ, is used to tell us to add up the things that follow. Using this idea, an informal but perhaps more palatable form of the formula for expected value is

$$\boxed{E = \Sigma(\text{probability}) \times (\text{payoff}).}$$

Example 22. You toss a coin, if it comes up heads you win \$1, if it comes up tails, you pay me \$0.50. The outcome H, has the payoff 1. The outcome T, has the payoff -0.5. What is the expected value of tossing a coin in this instance?

Solution: I find it helpful to make a table for these kinds of problems. In the table, I list each outcome along with its probability, p and payoff, m.

Outcome	Head (H)	Tail (T)
Probability	$\frac{1}{2}$	$\frac{1}{2}$
Payoff (Value)	\$1.00	\$−0.50

Now that we have the probabilities and their payoffs lined up in columns, we can just multiply the values in each column and then add them together to find the expected value.

$$E = \frac{1}{2} \cdot \$1.00 + \frac{1}{2} \cdot \$(-0.50) = \$0.25$$

Thus the expect value of a toss of this coin is \$0.25.

Example 23. You are on a TV show. You have been asked to play a dice game in which you roll a single die. The dice game works as follows:

- If you roll a 1 or 2, you win \$50.
- If you roll a 3, you win \$20.
- If you roll a 4, 5, or 6, you lose \$30.

(1) What is the expected value of playing the game once?
(2) If you are given the opportunity to play the game ten times or get \$100, should you play to game or take the cash?

Solution:

(1) Once again we can make a table to solve this problem.

Outcome	1 or 2	3	4, 5, or 6
Probability	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{3}{6}$
Payoff (Value)	\$50	\$20	\$−30

We can now multiply each probability by its respective payoff to find the expected value of playing the game once.

$$E = \frac{2}{6} \cdot \$50 + \frac{1}{6} \cdot \$20 + \frac{3}{6} \cdot \$(-30) = \$5$$

Thus the expect value of a roll of this die is \$5.

(2) Since the expected value of playing the game once is \$5, then when you play it 10 times you should be expect to win \$50. Thus it would be in your best interest to take the \$100 and not play them game.

Example 24. You are taking a standardized multiple choice exam. Each questions has 5 possible answers. If you answer a question correctly, you will earn 5 points. If you answer a questions incorrectly, you lose $\frac{1}{4}$ point. When answering a question, you are certain that one of the answers is not possibly correct and thus you can eliminate it. Should you guess on such a question?

Solution: In order to be able to answer the question if we should guess, we need to find the expected value of guessing. We will once again set up a table.

Outcome	Right	Wrong
Probability	$\frac{1}{4}$	$\frac{3}{4}$
Payoff (Value)	5	$-\frac{1}{4}$

The probabilities are out of 4 since we were able to eliminate one of the 5 possible answer and are guessing from the remaining four answers.

We can now multiply each probability by its respective payoff to find the expected value of playing the game once.

$$E = \frac{1}{4} \cdot 5 + \frac{3}{4} \cdot \left(-\frac{1}{4}\right) = \frac{17}{16}$$

Thus the expect value of a toss of this coin is $\frac{17}{16} = 1.0625$. Since the expected value of guessing in this situation is positive, then it is worthwhile to guess when you can eliminate one of the answer choices.