



Conditional Probability & Independence

The general formula for determining the probability of an event is:

$$P(\text{event}) = \frac{\text{the total number of successes}}{\text{the total number of outcomes}}$$

This is still true even if I tell you some information about the outcome before you calculate the probability. These sorts of problems involve **conditional probability**. The given information in a problem reduces the sample space and changes the value of the denominator of the fraction in the general formula above, and sometimes the numerator's value as well. We might describe a conditional probability as "the chances that A happens, given that B happens", and the notation for that is $P(A|B)$.

Example 1: I draw a red card from a well-shuffled deck of cards. What is the probability that the card I drew was (a) a heart, (b) a king, and (c) a club?

Solution: There are 52 cards in a deck of cards, but only 26 of them are red. Since we know that the card that was drawn is a red card, the total number of outcomes gets reduced from 52 to 26.

(a) There are 13 hearts in the deck, all of them red, so the number of successes doesn't change. $P(\text{heart} | \text{red card}) = \frac{13}{26} = \frac{1}{2}$.

(b) There are 4 kings in the deck, but two of them are not in the sample space anymore because they are black cards, so there are only two successes in this case. $P(\text{king} | \text{red card}) = \frac{2}{26} = \frac{1}{13}$. (Notice that this is the same probability as drawing a king without being told the colour of the card.)

(c) Since clubs are black cards, not only is the denominator of the fraction changed, the numerator is too. All of the successes have been removed from the sample space, so $P(\text{club} | \text{red card}) = \frac{0}{26} = 0$.

In situations where the outcomes are not countable, we can calculate a conditional probability from other probabilities.

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

This formula works on countable probabilities as well; for the card example above, we would have had $\frac{13/52}{26/52}$, $\frac{2/52}{26/52}$ and $\frac{0/52}{26/52}$, which would reduce to the probabilities we calculated anyway since all those "over 52"s would have cancelled out.



Example 2: At the Learning Centre, we have English and Math/Science desks at both campuses. The table at the right lists the percentages of the people who visited the Learning Centre during the month of September for each situation.

	Broadway	Downtown
Math/Sci	27%	19%
English	31%	23%

What is the probability that in the month of September that (a) a student who comes to the Broadway campus is looking for math or science help, and (b) a student who needs math or science help comes to the Broadway campus?

Solution: These two questions are different! In each one, the given information is different, so the probability is calculated from a different sample space.

(a) In this question, the location is the given information, and we are asked for the probability of the subject matter:

$$\begin{aligned}
 P(\text{math/sci and Broadway}) &= .27 \\
 P(\text{Broadway}) &= .27 + .31 = .58 \\
 \therefore P(\text{math/sci} \mid \text{Broadway}) &= \frac{.27}{.58} = .466\dots
 \end{aligned}$$

(b) In this question, the subject is the given information, and we are asked about the probability of the location:

$$\begin{aligned}
 P(\text{math/sci and Broadway}) &= .27 \\
 P(\text{math/sci}) &= .27 + .19 = .46 \\
 \therefore P(\text{Broadway} \mid \text{math/sci}) &= \frac{.27}{.46} = .587\dots
 \end{aligned}$$

INDEPENDENT VARIABLES

Two probability variables or events, A and B, are **independent** when telling you about one does not affect the probability of the other. Mathematically, that means that $P(A) = P(A|B)$ and $P(B) = P(B|A)$. In the first example in this worksheet, telling you that a red card was drawn didn't affect the probability that the card was a king.

Example 3: Show that the events "drawing a red card" and "drawing a king" from the experiment in Example 1 are independent.

Solution: We've already seen that $P(\text{king})$ and $P(\text{king} \mid \text{red card})$ are equal; they're both $\frac{1}{13}$. Now we should show that $P(\text{red card}) = P(\text{red card} \mid \text{king})$. $P(\text{red card}) = \frac{26}{52} = \frac{1}{2}$.
 $P(\text{red card} \mid \text{king}) = \frac{P(\text{red card and king})}{P(\text{king})} = \frac{\frac{2}{52}}{\frac{4}{52}} = \frac{2}{4} = \frac{1}{2}$. Since these probabilities are the same, the two events are independent.



EXERCISES

A. A box of crayons has these colours: red, orange, yellow, green, blue, violet, brown, black, pink, white, gold, turquoise. You take a crayon at random from the box with your eyes closed.

- 1) How many outcomes are there?
- 2) What is the probability that you drew the green crayon?
- 3) While your eyes are still closed, I tell you that the colour of the crayon you picked has exactly five letters in its name. How many outcomes are there now?
- 4) What is the conditional probability that you drew the green crayon, given that the colour has five letters in its name?
- 5) What is the conditional probability that you drew the blue crayon, given that the colour has five letters in its name?

B. Two fair, six-sided dice are rolled, one red and one blue. Create a grid displaying the outcomes of rolling these dice. Shade (using a different colour each time) the boxes corresponding to the events:

- 1) doubles were rolled.
- 2) the total on the dice was 8.
- 3) doubles were rolled and the total on the dice was 8.

Find the probability that:

- 4) doubles were rolled, given that the total on the dice was 8.
- 5) the total on the dice was 8, given that doubles were rolled.

C. A study examined residents of a mining town to see if the mining operation had adverse health effects. The table at the right shows the percentages of respondents who had either of two health factors related to lung cancer, and lung cancer itself.

- 1) Finish the table by adding totals to the rows and columns.

	Lung cancer	No lung cancer
Smoker	12.9	6.05
Smoker & Family History	8.95	1.15
Family History	5.25	3.9
Neither factor	4.45	57.35

What is the probability that a respondent picked at random:

- 2) both had cancer in their family history and smoked?
- 3) had lung cancer, given that they were a smoker with no family history of cancer?
- 4) had both health factors, given that they had lung cancer?
- 5) had neither health factor, given that they had no lung cancer?
- 6) had lung cancer, given that they had at least one health factor?

100%



D. The table at the right shows the ordering habits of customers at a fast-food restaurant. The number of orders for each situation in a given month are shown. What is the probability that, during this month:

1) a kids' meal customer ordered a chicken burger?

2) a customer who bought a beef burger by itself ordered a MegaBurger?

3) a customer ordered a chicken nugget combo if they ordered chicken?

4) a beef purchase was for a cheeseburger or a combo?

5) a kids' meal with chicken was for chicken nuggets?

6) an order was a combo if it wasn't an order for something by itself?

		By itself	Combo	Kids' meal	
Beef	Hamburger	26	253	121	400
	Cheeseburger	34	389	184	607
	MegaBurger	56	407	0	463
Chicken	Chicken nuggets	152	61	209	422
	Chicken wrap	283	129	7	419
	Chicken burger	12	274	171	457
		563	1513	692	2768

E. In a certain factory, 64% of the employees are men. 19% of the men have taken more than 1 week of leave during a single calendar year while working there. 16% of the employees are women who have taken more than 1 week off during a year.

1) Are extended leave (i.e., more than 1 week of leave) and gender independent variables?

2) If happiness and gender are independent, and 54% of the employees are men who are happy with working conditions, what percent are women who are unhappy?

SOLUTIONS

A. (1) 12 (2) $\frac{1}{12}$ (3) 4 (4) $\frac{1}{4}$ (5) 0

B. (1) 6 boxes shaded (2) 5 boxes shaded (3) 1 box shaded (4) $\frac{1}{5}$ (5) $\frac{1}{6}$

C. (1) LC: 31.55, NLC: 68.45; S: 18.95, S&FH: 10.1, FH: 9.15, Nf: 61.8 (2) 10.1%
 (3) $\frac{12.9}{18.95} \approx 68.07\%$ (4) $\frac{8.95}{31.55} \approx 28.37\%$ (5) $\frac{57.35}{68.45} \approx 83.78\%$ (6) $\frac{27.1}{38.2} \approx 70.94\%$

D. (1) $\frac{171}{692} \approx 24.71\%$ (2) $\frac{56}{116} \approx 48.28\%$ (3) $\frac{61}{1298} \approx 4.70\%$ (4) $\frac{1267}{1470} \approx 86.19\%$
 (5) $\frac{209}{387} \approx 54.01\%$ (6) $\frac{1513}{2205} \approx 68.6\%$

E. (1) No. $P(\text{man}) = .64$ and $P(\text{man} \mid \text{extended leave}) = \frac{.19 \times .64}{.19 \times .64 + .16} \approx .43$

(2) Since $P(\text{happy}) = P(\text{happy} \mid \text{man})$, $P(\text{happy woman}) = 0.30375$, and by elimination 5.63% of the employees are unhappy women.

