Conditional Probability

Example: The following table shows the probabilities of a randomly selected student from our Finite Math class of falling into one of the four categories.

	Athlete	**Non-Athlete**
Freshman	0.45	0.16
Sophomore	0.26	0.13

I randomly select a student from our class. What is the probability that the student is an athlete? 0.71 or 71%

I randomly select a student from our class, but I tell you that person is a sophomore. How does your answer change?

The probability now becomes 0.26 or 26%

Conditional Probability

Having additional information/knowledge about a random process event can sometimes change the probability!

Example: A baseball player has a 0.292 batting average for 2016, so we would expect the player to have a hit with probability 29% during the 2017 season.

Q: What if at the beginning of the 2017 season, our player makes 5 successful hits in a row! Do we still think that his batting probability is 29%?

Q: What if (close to opening day) we learn that the pitcher our player will be facing is left-handed. We should use this new information to re-assign the probability.

Conditional Probability

This new probability is called **conditional probability**.

We have some **prior information** about the outcome, or about the conditions under which the experiment will be performed.

This **additional information can change the sample space**, and it also changes the subset corresponding to our event!

Conditional Probability

Example: A single card is drawn at random from a standard deck of cards.

- Let H be the event that *a heart is drawn*
- Let R be the event that *a red card is drawn*

a) What is $P(H)$?

$$\frac{13}{52} = 25\%$$

b) I draw a card at random, and (without showing you the card) I tell you that it is red. What is the probability that it is a Heart?

$$\frac{13}{26} = 50\%$$

Conditional Probability

We calculated $P(H)$ **given that the card is red**.

The notation is $P(H|R)$ (read "the conditional probability of H given R"). In our example, the events were

- H = the event that *a heart is drawn* (13 outcomes)
- R = the event that *a red card is drawn* (26 outcomes)

and the conditional probability was $P(H|R) = \frac{13}{26} = 0.50$.

Note: The knowledge that the card drawn is red changed the sample space from 52 to 26 possible outcomes. These are exactly the outcomes from the event R.

The original sample space (entire deck) is reduced to R!

Conditional Probability

Definition: If A and B are events in a sample space S and $P(B) \neq 0$, the **conditional probability** that A will occur, given that B has occurred, is

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

For equally likely outcomes, this is simply

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

In some cases, $P(A|B) = P(A)$ (the occurrence of event B doesn't change the probability of A). In this case, A and B are independent events. We will study these next week.

Calculating Conditional Probabilities

Example: I recorded the following data over a period of 30 days. I kept track of my mood, and I kept track of the weather on that day. Pick a day at random out of the 30.

	Sunny (S)	Cloudy (NS)
Good (G)	9	6
Bad (NG)	1	14

a) What is $P(G)$? $P(G) = \frac{9+6}{30} = \frac{15}{30} = 0.5.$

b) What is $P(S)$? $P(S) = \frac{9+1}{30} = \frac{10}{30} \approx 0.333.$

Calculating Conditional Probabilities

	Sunny (S)	Cloudy (NS)
Good (G)	9	6
Bad (NG)	1	14

c) What is $P(G|S)$? $P(G|S) = \frac{P(G \cap S)}{P(S)} = \frac{9/30}{10/30} = \frac{9}{10}$

d) What is $P(S|G)$? $P(S|G) = \frac{P(G \cap S)}{P(G)} = \frac{9/30}{15/30} = \frac{9}{15}$

In this example, $P(S) \neq P(S|G)$ (50% versus 60%).

Note: This does NOT imply a cause-effect relationship! The weather might have an effect on my mood, however it is unlikely that my mood would have any effect on the weather.

Visualizing Probabilities

Example: Of the students at a university, 50% regularly attend the football games, 30% are first-year students and 40% are upper-class students who do not regularly attend football games.

a) What is the probability that a randomly chosen student is both is a first-year student and regularly attends football games?

We could use an algebra, but it is easier represent the probabilities visually. We have the following information:

G : students who regularly attend games (50%);
F : freshman students (30%);
F' : sophomores and beyond;

We are given $P(F' \cap G') = 40\%$. We want $P(F \cap G)$.

Visualizing Probabilities

Relevant events:

G : students who regularly attend games (50%);

F : freshman students (30%);

F' : sophomores and beyond;

We are given $P(F' \cap G') = 40\%$. We want $P(F \cap G)$.

	G	**G'**	Row Totals
F	0.2	0.1	0.3
F'	0.3	0.4	0.7
Col Totals	0.5	0.5	1

$P(F \cap G) = 0.2$

Visualizing Probabilities

b) What is the (conditional) probability that a randomly chosen student attends the games, given that he/she is a first year student?

$$P(G|F) = \frac{P(G \cap F)}{P(F)} = \frac{0.2}{0.3} \approx 66.7\%$$

c) What is the conditional probability that the person is a first year student given that he/she regularly attends football games?

$$P(F|G) = \frac{P(F \cap G)}{P(G)} = \frac{0.2}{0.5} = 0.4 = 40\%$$

Calculating Conditional Probabilities

Example: E and F are events in a sample space with

$$P(E) = 0.50 \qquad P(F) = 0.40 \qquad P(E \cap F) = 0.30$$

a) What is $P(E|F)$?

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{0.3}{0.4} = 75\%$$

b) What is $P(F|E)$?

$$P(F|E) = \frac{P(F \cap E)}{P(E)} = \frac{0.3}{0.5} = 60\%$$

A Formula For $P(E \cap F)$

We can rearrange the formula for conditional probability

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

to get

$$\boxed{P(E \cap F) = P(F) \cdot P(E|F)}$$

Also from $P(F|E)$, we have the alternate form:

$$\boxed{P(E \cap F) = P(E) \cdot P(F|E)}$$

This is very useful in calculating probabilities for **sequential events** (events that happen one after the other).

A Formula For $P(E \cap F)$

Example: If $P(E|F) = .2$ and $P(F) = .3$, find $P(E \cap F)$.

$$P(E) = P(E|F) \cdot P(F) = 0.2 \cdot 0.3 = 0.06 = 6\%$$

Example: A bag of marbles contains 6 red and 4 blue. Select a marble at random from the bag, and then (without replacing the first marble) select a second marble. What is the probability that both marbles are red?

Consider the following events:

$A =$*the first marble is red*

$B =$*the second marble is red*

We are interested in $P(A \cap B)$, which is

$$P(A \cap B) = P(A) \cdot P(B|A) = \frac{6}{10} \cdot \frac{5}{9} \approx 0.333$$

A Formula For $P(E \cap F)$

Example: A bag of marbles contains 6 red and 4 blue. Select a marble at random from the bag, and then (without replacing the first marble) select a second marble.

What is the probability that the second marble is blue?

This can happen in two different (disjoint) ways. The first marble is red (and the second blue), or the first marble is blue (and the second blue again).

We add the probabilities for each of the two disjoint scenarios:

$$\frac{6}{10} \cdot \frac{4}{9} + \frac{4}{10} \cdot \frac{3}{9} = \frac{24}{90} + \frac{12}{90} = \frac{36}{90} = 40\%$$

Note: This is similar to the addition and multiplication principles we were using for counting elements in a set!

Quality Control

Example: A box of 20 apples is ready for shipment. Four of the apples are spoiled. An inspector picks **at most four** apples for testing. He selects one apple at a time, inspects it, and if it is not spoiled, sets it aside. The moment he selects a defective apple, the box fails the quality control test. If all four apples selected are good, puts them back and ships the box.

a) What is P(box passes)?

He must pick 4 good apples in a row:

$$\frac{16}{20} \cdot \frac{15}{19} \cdot \frac{14}{18} \cdot \frac{13}{17} \approx 0.376$$

b) What about P(box fails)? $1 - 0.376 = 0.624$

Quality Control

Example: 20 apples, 4 are spoiled, select at most 4 (one at a time) for testing.

c) What is the probability that either the first or second apple selected is bad?

First apple is bad with probability 4/20.
For the second apple to be bad, the first must be good (otherwise the testing process stops early). So in total, the probability we want is

$$\frac{4}{20} + \frac{16}{20} \cdot \frac{4}{19} \approx 0.368$$

Conditional Probability

Example: A library has 40% Fiction and 60% Non-Fiction books. Of the Fiction books, 20% are Word and need replacement. Of the Non-Fiction, 10% are Worn and need replacement. What is the probability that a randomly chosen book needs to be replaced?

We are given $P(F) = 0.4$, $P(NF) = 0.6$, and the rest are conditional: $P(W|F) = 0.2$ and $P(W|NF) = 0.1$. We are looking to find $P(W)$.

The book can either be F and need replacement, or NF and need replacement (mutually exclusive), so we have:

$$\begin{aligned} P(W) &= P(W \cap F) + P(W \cap NF) \\ &= 0.4 \cdot 0.2 + 0.6 \cdot 0.1 \\ &= 0.14 \end{aligned}$$

Genetics

For a certain type of pea plant, the color of its flower (either Red or White) is determined by a pair of genes.

Each gene is of one of the types C (dominant, red) or c (recessive, white).

When two plants are crossed, the offspring receives one gene from each parent. If the parent is of type Cc, both genes are equally likely to be passed on.

Plants for which the genotype cc produce white flowers, while plants of genotypes CC or Cc produce red flowers.

Genetics

Example: You cross two pea plants of genotype Cc.

a) What is the probability that their offspring produces white flowers?

Each parent must pass on the recessive gene. Since both C and c are equally likely, $P(c) = 0.5$ for each parent.

$$(0.5)(0.5) = 0.25 = 25\%$$

b) What is the probability that their offspring produces red flowers?

$$1 - 0.25 = 0.75$$

Note: we can alternately compute the disjoint scenarios $P(CC)$, $P(Cc)$, or $P(cC)$, and add.

Genetics

Example: You have a batch of red flowering pea plants, of which 60% have genotype Cc and 40% have genotype CC.

You select one of these plants at random and cross it with a white flowering pea plant (cc). What is the probability that the offspring produces red flowers?

The white flowering plant can only contribute with the recessive gene, so the entire outcome is determined by the randomly chosen parent.

If the randomly chosen parent is CC, we always get red flowers, and this happens with probability $(0.4)(1)$.

If the randomly chosen parent is Cc, it must pass on the C gene in order for the offspring to produce red flowers.

$$(0.4)(1) + (0.6)(0.5) = 0.7$$