5.9: Independence

Two events are considered **independent** if the probability of one event occurring is not changed by knowing if the other event occurred or not. Events that are not independent are called dependent.

Here are examples of independent (unrelated) events:

- A fair coin flip comes up heads; the coin is flipped again and comes up heads.
- A student is unable to attend a math class at De Anza College; it rains today in New York City.
- A house in San Francisco starts on fire; on the same day, a house in Dallas starts on fire.
- A patient is diagnosed with cancer; on the same day, another patient is diagnosed with pneumonia.

In these independent events, the probability of the second event occurring is not affected by whether the first event occurs.

Examples of dependent (related) events

- A student gets an A on the first exam; the same student gets an A on the second exam.
- A person has never smoked; the same person gets lung cancer.
- An earthquake destroys a home in San Francisco; on the same day, an earthquake destroys a home in Oakland.
- A student majors in Computer Science; the same student wants to work for Google.

In these dependent events, the probability of the second event occurring is affected by whether the first event occurs:

- A student who gets an A on an exam is more likely to get an A on another exam.
- A non-smoker is less likely to get lung cancer than is a smoker.
- A single strong earthquake will affect homes all over the Bay Area.
- A Computer Science major is more likely to work for a tech company, such as Google.

The mathematical definition of independent events means that the marginal probability of the first event occurring is the same as the conditional probability of the first occurring given the second event occurred. We can then adjust the Multiplicative Rule to get three formulas, any of which can be used to test for independence:

Independent events

If events A and B are **independent**, then the following statements are all true:

$$P(A) = P(A \mid B)$$

$$P(B) = P(B \mid A)$$

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

The last formula is particularly useful and can be easily generalized to find the joint probability of many independent events from looking at the simple marginal probabilities, making random sampling in statistical research so critical.

Example: Flip a coin ten times

A fair coin is flipped ten times. Find the probability of getting heads on all 10 tosses.

Solution

Because the coin tosses are independent, the multiplicative rule requires only marginal probabilities:

$P(\text{ all Heads }) = P(H)^{10} = 0.5^{10} = 0.0009766$

Example: Surprise quiz

On Monday, there is a 10% chance your history instructor will have a surprise quiz. On the same day, there is a 20% chance that your Math instructor will also have a surprise quiz. No other class you are taking has surprise quizzes. What is the probability that you will have a least one surprise quiz on Monday? Assume that all events are independent.

Solution

Let H be the event "Surprise quiz in History" and M be the event "Surprise quiz in Math." Then use both the Additive Rule and the Multiplicative Rule for independent events.

$P(H \text{ or } M) = P(H) + P(M) - P(M \text{ and } H)$

$P(H) = 0.10 \qquad P(M) = 0.20$

$P(H \text{ and } M) = P(H) \times P(M) = 0.10 \times 0.20 = 0.02$

$P(H \text{ or } M) = 0.10 + 0.20 - 0.02 = 0.28$

There is a 28% chance that there will be at least one surprise quiz on Monday.

Example: Accidents and DUI

1000 drivers were asked if they were involved in an accident in the last year. They were also asked if during this time, they were DUI, driving under the influence of alcohol or drugs. Are the events "Driver was DUI" and "Driver was involved in an accident" independent or dependent event

	Accident	No Accident	Total
DUI	70	130	200
Non-DUI	30	770	800
Total	100	900	1000

Solution

Let A be the event "the driver had an accident" and D be the event "the driver was DUI". We can use any of the rules for independence answer this question. Let's show all three possible methods here, but in practice choose the most convenient formula given the provided data.

Use Formula 1:

$P(A) = 100/1000 = 0.10$

$P(A|D) = 70/200 = 0.35$

$P(A) \neq P(A|D)$

Use Formula 2:

$P(D) = 200/1000 = 0.20$

$P(D|A) = 70/100 = 0.70$

$P(D) \neq P(D|A)$

Use Formula 3:

$P(A) = 100/1000 = 0.10$

$P(D) = 200/1000 = 0.20$

$P(A \text{ and } D) = 70/1000 = 0.07$

$P(A) \times P(D) = (0.10)(0.20) = 0.02$

$P(A \text{ and } D) \neq P(A) \times P(D)$

"Driver was DUI" and "Driver was involved in an accident" are dependent events.

Example: Accidents and origin of car

1000 drivers were asked if they were involved in an accident during the last year. They were also asked if during this time, if they were driving a domestic car or an imported car. Are the events "Driver drives a domestic car" and "Driver was involved in an accident" independent or dependent events?

	Accident	No Accident	Total
Domestic Car	40	540	600
Imported Car	60	360	400
Total	100	900	1000

Solution

Let A be the event “the driver had an accident” and D be the event “the driver drives a domestic car”. Let's again show all three possible methods here, but in practice choose the most convenient formula given the provided data.

Use Formula 1

$P(A)$ = 100/1000 = 0.10

$P(A|D)$ = 60/600 =0.10

$P(A) = P(A|D)$

Use Formula 2

$P(D)$ = 600/1000 = 0.60

$P(D|A)$ = 60/100 =0.60

$P(D) = P(D|A)$

Use Formula 3

$P(A)$ = 100/1000 = 0.10

$P(D)$ = 600/1000 = 0.60

$P(A \text{ and } D)$ = 60/1000 = 0.06

$P(A) \times P(D)$ = (0.10)(0.60) = 0.06

$P(A \text{ and } D) = P(A) \times P(D)$

"Driver has an accident" and "Driver drives a domestic car" are independent events.