Conditional Probability

Conditional probability: How prior probability on an event is updated?

- *When will you update you belief on an event?*

- *How do you update your belief?*

You are given a name and told that this person is living in USA.
What is the probability that this person is male? $P[A] =$? (A: a person is male)

Now, you are told that the person is a firefighter.
Will this extra information change the probability that this person is male?
What will be the new probability? $P[A|B]$ (B: a person is a firefighter)

Instead, you are told that it is raining today.
Will this extra information change the probability that this person is male?

Conditional Probability

Conditional probability: A sample space contains all the possible outcomes of an experiment. Sometimes, we obtain some additional information about an experiment that tells us that the outcome comes from a certain part of the sample space. In this case, the probability of an event is based on the outcomes in that part of the sample space. A probability that is based on a part of a sample space is called a *conditional probability.*

Example: Consider a population of 1000 aluminum rods. For each rod, the length is classified as too short, too long, or OK, and the diameter is classified as too thin, too thick, or OK. These 1000 rods form a sample space where each rod is equally likely to be sampled. The number of rods in each category is presented in the table.

If a rod is sampled, what is the probability that the diameter of the rod is OK?

$\mathrm{P(diameter = OK)} = \frac{928}{1000} = 0.928$ (unconditional probability calculated from the entire sample space)

Now, let's assume the length of the sampled rod is measured as OK. What is the probability that the diameter of the rod is OK?

$\mathrm{P(diameter = OK|length = OK)} = \frac{900}{942} = 0.955$ (conditional probability calculated from a part of the sample space)

	Diameter		
Length	*Too Thin*	*OK*	*Too Thick*
Too Short	10	3	5
OK	38	900	4
Too Long	2	25	13

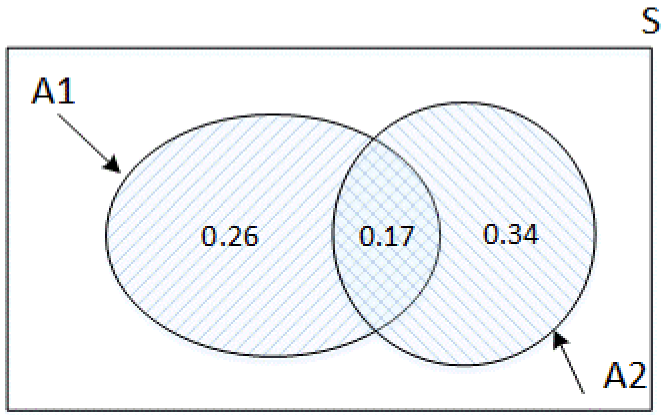
Conditional Probability

Example: A pregnancy test is used to determine the sex of a child. The test may reveal the presence of a protein zone. This zone is present in 43% of all pregnant woman. It is known that 51% of children born are male, 17% of all children born are male and the protein zone is present. Suppose that we are given the information that the protein zone is present in a test and asked: “what is the probability that the child is male?”

Let $A_1 = \{the\ protein\ zone\ is\ present\ in\ a\ pregnancy\}$
Let $A_2 = \{the\ born\ child\ is\ male\}$
It is known that, $P[A_1 \cap A_2] = 0.17, P[A_1] = 0.43, P[A_2] = 0.51$

Without the given information, the probability that a child born is male is 51%.
Will the information given change this probability? (increase? Decrease?)

Let’s denote this probability as: $P[A_2|A_1] = \frac{0.17}{0.43} = 0.395$
Receiving the information of presence of the protein zone reduces the probability of a male child from 51% to 39.5%

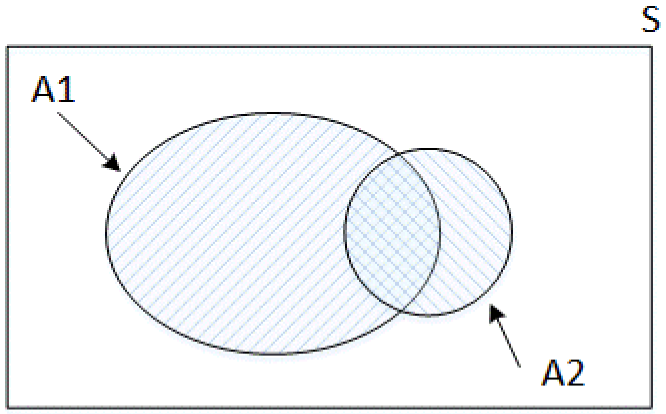


Venn Diagram

Conditional Probability

Definition: Let $A_1\ and A_2$ be events such that $P[A_1] \neq 0$. The conditional probability of A_2 given A_1, denoted by $P[A_2|A_1]$, is defined by:

$$P[A_2|A_1] = \frac{P[A_1 \cap A_2]}{P[A_1]}$$



Venn Diagram

Example: Consider a population of 1000 aluminum rods. For each rod, the length is classified as too short, too long, or OK, and the diameter is classified as too thin, too thick, or OK. These 1000 rods form a sample space where each rod is equally likely to be sampled. The number of rods in each category is presented in the table. Calculate $P(diameter = OK|length = OK)$ using the definition of conditional probability.

	Diameter		
Length	*Too Thin*	*OK*	*Too Thick*
Too Short	10	3	5
OK	38	900	4
Too Long	2	25	13

$$P(diameter = OK) = \frac{928}{1000} = 0.928$$

$$P(length = OK) = \frac{942}{1000} = 0.942$$

$$P(diameter = OK \text{ } and \text{ } length = OK) = \frac{900}{1000} = 0.9$$

Applying the definition of conditional probability gives:

$$P(diameter = OK|length = OK) = \frac{P(diameter = OK \text{ } and \text{ } length = OK)}{P(length = OK)}$$

$$= \frac{0.9}{0.942} = 0.955$$

Practice: calculate the following conditional probability

$P(diameter = OK|length = Too\ Long)$ 0.625

$P(length = OK|diameter = OK)$ 0.970

Is conditional probability a "probability"?

If conditional probability $P(A|B) = \frac{P(A \cap B)}{P(B)}$ is a probability, it has to satisfy the three axioms of probability.

For a fixed B with $P(B) \neq 0$,
(1) $0 \leq P(A|B) \leq 1$, for any event A in the sample space S.
(2) $P(S|B) = 1$ (why?)
(3) Given n mutually exclusive events $A_1, A_2, \dots A_n$,

$$P(A_1 \cup A_2 \cup \cdots \cup A_n|B) = P(A_1|B) + P(A_2|B) + \cdots + P(A_n|B)$$

From the definition of conditional probability, we have,

$$P(A_1 \cup A_2 \cup \cdots \cup A_n|B) = \frac{P[(A_1 \cup A_2 \cup \cdots \cup A_n) \cap B]}{P(B)}$$

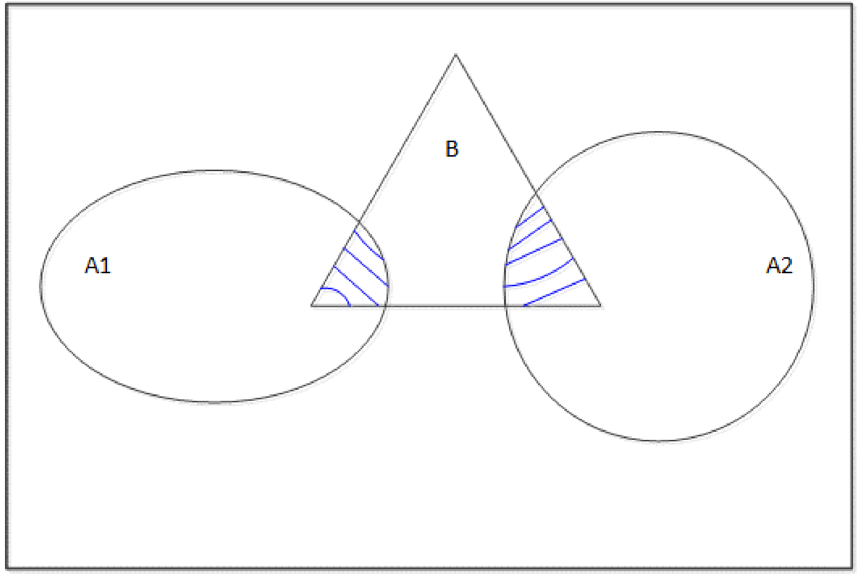
Since $A_1, A_2, \dots A_n$ are mutually exclusive, we have,

$$(A_1 \cup A_2 \cup \cdots \cup A_n) \cap B = (A_1 \cap B) \cup (A_2 \cap B) \dots \cup (A_n \cap B)$$

And, $(A_1 \cap B), (A_2 \cap B), \dots, (A_n \cap B)$ are mutually exclusive. Therefore,

$$P(A_1 \cup A_2 \cup \cdots \cup A_n|B) = \frac{P[(A_1 \cup A_2 \cup \cdots \cup A_n) \cap B]}{P(B)}$$

$$= \frac{P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B)}{P(B)}$$



$$= \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} + \cdots + \frac{P(A_n \cap B)}{P(B)} = P(A_1|B) + P(A_2|B) + \cdots + P(A_n|B)$$

Independent Events

Definition: Events $A_1\ and A_2$ are independent if and only if $P[A_1 \cap A_2] = P[A_1]P[A_2]$.

This definition can be used to test if two events are independent!

With this definition, let's revisit the definition of conditional probability: when $A_1\ and A_2$ are independent, we have

$$P[A_2|A_1] = \frac{P[A_1 \cap A_2]}{P[A_1]} = \frac{P[A_1]P[A_2]}{P[A_1]} = P[A_2]$$

The multiplication rule

let's revisit the definition of conditional probability.

$$P[A_2|A_1] = \frac{P[A_1 \cap A_2]}{P[A_1]} \Rightarrow P[A_1 \cap A_2] = P[A_2|A_1]P[A_1]$$

This is a way to calculate probability of joint events

Example: Recent research indicates that approximately 49% of all infections involve anaerobic bacteria. Furthermore, 70% of all anaerobic infections are polymicrobic. What is the probability that a given infection involve anaerobic bacteria and is polymicrobic?

Solution:

Let A_1 denote the event that the infection is anaerobic.
Let A_2 denote the event that the infection is polymicrobic.

We are given: $P[A_2|A_1] = 0.7; P[A_1] = 0.49$

Then, $P[A_1 \cap A_2] = P[A_2|A_1]P[A_1] = 0.7 \times 0.49 = 0.343$

The Law of Total Probability

If $A_1, A_2, \ \ldots, A_n$, are **mutually exclusive** and **exhaustive** events ($A_1 \cup A_2 \cup \cdots \cup A_n = S$), and B is any event in the same sample space.
Then,

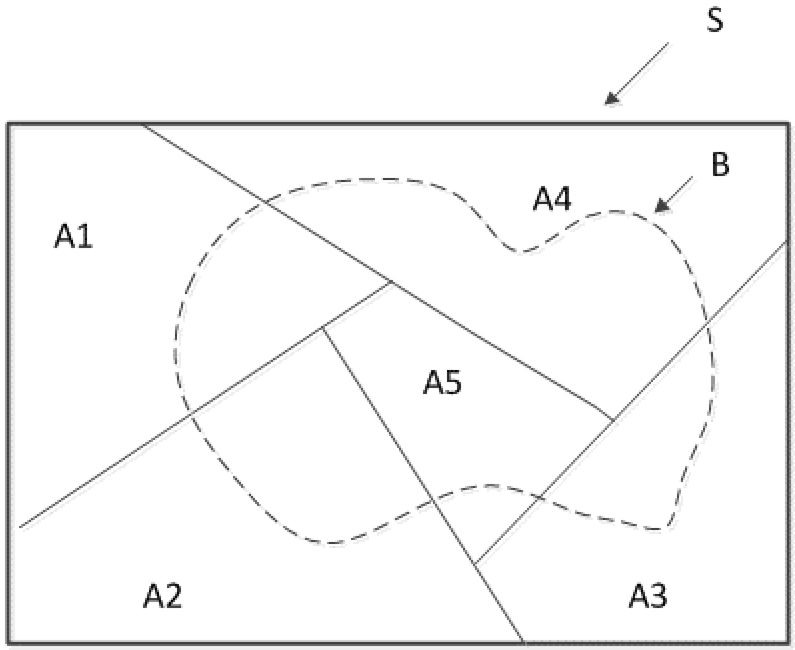
$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B) \\ &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n) \end{aligned}$$

if $P(A_i) \neq 0, for\ each\ A_i$.

This result follows from the fact that:

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B)$$

and $A_1 \cap B$, $A_2 \cap B$, $A_3 \cap B$, and $A_n \cap B$ are mutually exclusive.



Example on The Law of Total Probability

Example: Customers who purchase a certain make of car can order an engine in any of three sizes. Of all car sold, 45% have the smallest engine, 35% have the medium-sized one, and 20% have the largest. Of cars with the smallest engine, 10% fail an emissions test within two years of purchase, while 12% of those with the medium-sized engine and 15% of those with the largest engine fail. What is the probability that a randomly chosen car will fail an emissions test within two years?

Solution:
Let B denote the event that a car fails an emission test within two years.
Let A_1 denote the event that a car has a small engine, A_2 the event that a car has a medium-sized engine, A_3 the event that a car has a large engine. $A_1, A_2, and\ A_3$ are mutually exclusive and exhaustive. Then,

$$P(A_1) = 0.45 \quad P(A_2) = 0.35 \quad P(A_3) = 0.20$$
$$P(B|A_1) = 0.10 \quad P(B|A_2) = 0.12 \quad P(B|A_3) = 0.15$$
$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)$$
$$= 0.10 \times 0.45 + 0.12 \times 0.35 + 0.15 \times 0.20 = 0.117$$

Bayes’ Theorem

Bayes’ theorem is a formula that describes how to update the probabilities of hypotheses when given evidence (belief update).

Given a hypothesis (H) and a piece of evidence (E), Bayes’ theorem states that the relation between the probability of the hypothesis $P[H]$ before getting the evidence and the probability $P[H|E]$ of the hypothesis after getting the evidence is:

$$P[H|E] = \frac{P[E|H]P[H]}{P[E]}$$

$P[H|E]$ is called the **posterior probability**; $P[H]$ is the **prior probability**; $P[E|H]$ is called **likelihood**

The “diagnose problem” and the “inference problem”

Proof:

$$P[H|E] = \frac{P[H \cap E]}{P[E]} = \frac{P[E \cap H]}{P[E]} = \frac{P[E|H]P[H]}{P[E]}$$

Different forms of Bayes' Theorem

(1) $P[H|E] = \frac{P[E|H]P[H]}{P[E]}$

(2) $P[H|E] = \frac{P[E|H]P[H]}{P[E|H]P[H]+P[E|H']P[H']}$

(3) $P[H_j|E] = \frac{P[E|H_j]P[H_j]}{\sum_{i=1}^{n} P[E|H_i]P[H_i]}$, where $H_1, H_2, \dots, H_n$ are mutually exclusive and $H_1 \cup H_2 \cup \cdots \cup H_n = S$

Bayes' Theorem

Example: In St Louis county, 51% of the adults are males. Also, 9.5% of male smoke cigars, whereas, 1.7% of female smoke cigars. An adult is randomly selected for a survey.

(1) What is the probability that the selected person is a male?
(2) It is learned in the survey that the selected person smoke cigar. Use this additional information to update the probability that this selected person is a male.

Let's denote, M: a person is a male; M': a person is a female;
C: a person smokes cigar; C': a person does not smoke cigar

(1) $P[M] = 0.51$

(2) We know $P[M] = 0.51; P[M'] = 0.49; P[C|M] = 0.095; P[C|M'] = 0.017$

Applying the Bayes' Theorem:

$$P[M|C] = \frac{P[C|M]P[M]}{P[C|M]P[M] + P[C|M']P[M']} = \frac{0.095 \times 0.51}{0.095 \times 0.51 + 0.017 \times 0.49} \approx 0.853$$

This additional information increases the probability of the person to be male from 0.51 to 0.853

Practice Example: The proportion of people in a given community who have a certain disease is 0.005. A test is available to diagnose the disease. If a person has the disease, the probability that the test will produce a positive signal is 0.99 (detection rate). If a person does not have the disease, the probability that the test will produce a positive signal is 0.01 (false alarm rate). If a person tests positive, what is the probability that the person actually has the disease?

Solution:

Let D represent the event that the person has the disease, and let $+$ represent the event that the test result is positive. We want to find $P[D|+]$.

We are given the following probabilities: $P[D] = 0.005;\;\; P[+|D] = 0.99;\;\; P[+|D'] = 0.01$

By the Bayes' rule,

$$P\,[D|+] = \frac{P[+|D]P[D]}{P[+|D]P[D] + P[+|D']P[D']} = \frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.01 \times 0.995} = 0.332$$

The positive diagnose result increase one's probability of having the disease from 0.05 to 0.332. This diagnose method is useful.

If the false alarm rate of the method is 0.1 instead of 0.01,

$$P\,[D|+] = \frac{P[+|D]P[D]}{P[+|D]P[D] + P[+|D']P[D']} = \frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.1 \times 0.995} = 0.047$$

i.e., the positive diagnose result virtually does not change one's probability of having the disease. The test is not useful at all.