

1.3 Conditional Probability and Independence

All of the probabilities that we have dealt with thus far have been unconditional probabilities. A sample space was defined and all probabilities were calculated with respect to that sample space. In many instances, however, we are in a position to update the sample space based on new information. In such cases, we want to be able to update probability calculations or to calculate conditional probabilities.

Example 1.3.1 (Four aces) Four cards are dealt from the top of a well-shuffled deck. What is the probability that they are four aces?

The probability is $1/\binom{52}{4} = \frac{1}{270,725}$.

We can also calculate this probability by an “updating” argument, as follows. The probability that the first card is an ace is $4/52$. Given that the first card is an ace, the probability that the second card is an ace is $3/51$. Continuing this argument, we get the desired probability as

$$\frac{4}{52} \frac{3}{51} \frac{2}{50} \frac{1}{49} = \frac{1}{270,725}.$$

Definition of Independence

If A and B are events in S , and $P(B) > 0$, then the conditional probability of A given B , written $p(A|B)$, is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (1)$$

Note that what happens in the conditional probability calculation is that B becomes the sample space: $P(B|B) = 1$. The intuition is that our original sample space, S , has been updated to B . All further occurrence are then calibrated with respect to their relation to B .

Example 1.3.3 (Continuation of Example ??) Calculate the conditional probabilities given

that some aces have already been drawn.

$$\begin{aligned}
 & P(4 \text{ aces in 4 cards} | i \text{ aces in } i \text{ cards}) \\
 &= \frac{P(\{4 \text{ aces in 4 cards}\} \cap \{i \text{ aces in } i \text{ cards}\})}{P(i \text{ aces in } i \text{ cards})} \\
 &= \frac{P(4 \text{ aces in 4 cards})}{P(i \text{ aces in } i \text{ cards})} \\
 &= \frac{\binom{52}{i}}{\binom{52}{4} \binom{4}{i}} = \frac{(4-i)!48!}{(52-i)!} = \frac{1}{\binom{52-i}{4-i}}
 \end{aligned}$$

For $i = 1, 2$ and 3 , the conditional probabilities are .00005, .00082, and .02041, respectively.

(Three prisoners)

Three prisoners, A, B, and C, are on death row. The governor decides to pardon one of the three and chooses at random the prisoner to pardon. He informs the warden of his choice but requests that the name be kept secret for a few days. The next day, A tries to get the warden to tell him who has been pardoned. The warden refuses. A then asks which of B or C will be executed. The warden thinks for a while, then tells A that B is to be executed.

Warden's reasoning: Each prisoner has a $\frac{1}{3}$ chance of being pardoned. Clearly, either B or C must be executed, so I have given A no information about whether A will be pardoned.

A's reasoning: Given that B will be executed, then either A or C will be pardoned. My chance of being pardoned has risen to $\frac{1}{2}$.

Let A , B , and C denote the events that A, B, or C is pardoned, respectively. We know that $P(A) = P(B) = P(C) = 1/3$. Let W denote the event that the warden says B will die. Using (1), A can update his probability of being pardoned to

$$P(A|W) = \frac{P(A \cap W)}{P(W)}.$$

What is happening can be summarized in this table:

Prisoner Pardoned	Warden tells A	Probability
A	B dies	$\frac{1}{6}$
A	C dies	$\frac{1}{6}$
B	C dies	$\frac{1}{3}$
C	B dies	$\frac{1}{3}$

Using this table, we can calculate

$$\begin{aligned}
 P(W) &= P(\text{warden says B dies}) \\
 &= P(\text{warden says B dies and A pardoned}) \\
 &\quad + P(\text{warden says B dies and C pardoned}) \\
 &\quad + P(\text{warden says B dies and B pardoned}) \\
 &= \frac{1}{6} + \frac{1}{3} + 0 = \frac{1}{2}.
 \end{aligned}$$

Thus, using the warden's reasoning, we have

$$\begin{aligned}
 P(A|W) &= \frac{P(A \cap W)}{P(W)} \\
 &= \frac{P(\text{warden says B dies and A pardoned})}{P(\text{warden says B dies})} = \frac{1/6}{1/2} = \frac{1}{3}.
 \end{aligned}$$

However, A falsely interprets the event W as equal to the event B^c and calculates

$$P(A|B^c) = \frac{P(A \cap B^c)}{P(B^c)} = \frac{1/3}{2/3} = \frac{1}{2}.$$

We see that conditional probabilities can be quite slippery and require careful interpretation.

The following are several variations of (1):

$$P(A \cap B) = P(A|B)P(B)$$

$$P(A \cap B) = P(B|A)P(A)$$

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

(Bayes' Rule)

Let A_1, A_2, \dots be a partition of the sample space, and let B be any set.

Then, for each $i = 1, 2, \dots$,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}.$$

Example 1.3.6 (Codings)

When coded messages are sent, there are sometimes errors in transmission. In particular, Morse code uses “dots” and “dashes”, which are known to occur in the proportion of 3:4. This means that for any given symbol,

$$P(\text{dot sent}) = \frac{3}{7} \quad \text{and} \quad P(\text{dash sent}) = \frac{4}{7}.$$

Suppose there is interference on the transmission line, and with probability $\frac{1}{8}$ a dot is mistakenly received as a dash, and vice versa. If we receive a dot, can we be sure that a dot was sent?

Using Bayes’s Rule, we can write

$$P(\text{dot sent}|\text{dot received}) = P(\text{dot received}|\text{dot sent}) \frac{P(\text{dot sent})}{P(\text{dot received})}.$$

$$\begin{aligned} P(\text{dot received}) &= P(\text{dot received} \cap \text{dot sent}) \\ &\quad + P(\text{dot received} \cap \text{dash sent}) \\ &= P(\text{dot received}|\text{dot sent})P(\text{dot sent}) \\ &\quad + P(\text{dash received}|\text{dash sent})P(\text{dash sent}) \\ &= \frac{7}{8} \times \frac{3}{7} + \frac{1}{8} \times \frac{4}{7} = \frac{25}{56}. \end{aligned}$$

So we have

$$P(\text{dot sent}|\text{dot received}) = \frac{(7/8) \times (3/7)}{25/56} = \frac{21}{25}.$$

Example A (Incidence of Rare Disease)

In some cases it may happen that the occurrence of a particular event, B , has no effect on the probability of another event, A . For these cases, we have the following definition.

Definition of Independence

Two events, A and B , are statistically independent if

$$P(A \cap B) = P(A)P(B).$$

Example 1.3.8

The gambler introduced at the start of the chapter, Chevalier de Mere, was particularly interested in the event that he could throw at least 1 six in 4 rolls of a die. We have

$$\begin{aligned} P(\text{at least 1 six in 4 rolls}) &= 1 - P(\text{no six in 4 rolls}) \\ &= 1 - \prod_{i=1}^4 P(\text{no six on roll } i) \\ &= 1 - \left(\frac{5}{6}\right)^4 = 0.518. \end{aligned}$$

The equality of the second line follows by independence of the rolls.

Theorem 1.3.9

If A and B are independent events, then the following pairs are also independent:

- a. A and B^c .
- b. A^c and B .
- c. A^c and B^c .

PROOF: (a) can be proved as follows.

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) = P(A) - P(A)P(B) \quad (A \text{ and } B \text{ are independent}) \\ &= P(A)(1 - P(B)) = P(A)P(B^c) \end{aligned}$$

□

Example 1.3.10 Tossing two dice

Let an experiment consists of tossing two dice. For this experiment the sample space is

$$S = \{(1, 1), (1, 2), \dots, (1, 6), (2, 1), \dots, (2, 6), \dots, (6, 1), \dots, (6, 6)\}.$$

Define the following events:

$$A = \{\text{double appear}\} = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\},$$

$$B = \{\text{the sum is between 7 and 10}\}$$

$$C = \{\text{the sum is 2, 7 or 8}\}.$$

Thus, we have $P(A) = 1/6$, $P(B) = 1/2$ and $P(C) = 1/3$. Furthermore,

$$\begin{aligned} P(A \cap B \cap C) &= P(\text{the sum is 8, composed of double 4s}) \\ &= \frac{1}{36} = \frac{1}{6} \times \frac{1}{2} \times \frac{1}{3} \\ &= P(A)P(B)P(C) \end{aligned}$$

However,

$$P(B \cap C) = P(\text{sum equals 7 or 8}) = \frac{11}{36} \neq P(B)P(C).$$

Similarly, it can be shown that $P(A \cap B) \neq P(A)P(B)$. Therefore, the requirement $P(A \cap B \cap C) = P(A)P(B)P(C)$ is not a strong enough condition to guarantee pairwise independence.

Definition 1.3.12 (Mutual Independence)

A collection of events A_1, \dots, A_n are mutually independent if for any subcollection A_{i_1}, \dots, A_{i_k} , we have

$$P(\cap_{j=1}^k A_{i_j}) = \prod_{j=1}^k P(A_{i_j}).$$

Example 1.3.13 (Three coin tosses)

Consider the experiment of tossing a coin three times. The sample space is $\{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$. Let H_i , $i = 1, 2, 3$, denote the event that the i^{th} toss is a head. For example, $H_1 = \{HHH, HHT, HTH, HTT\}$. Assuming that the coin is fair and has an equal probability of landing heads or tails on each toss, the events H_1 , H_2 and H_3 are mutually independent.

To verify this we note that

$$P(H_1 \cap H_2 \cap H_3) = P(\{HHH\}) = \frac{1}{8} = P(H_1)P(H_2)P(H_3).$$

To verify the condition in Definition 1.3.12, we must check each pair. For example,

$$P(H_1 \cap H_2) = P(\{HHH, HHT\}) = \frac{2}{8} = P(H_1)P(H_2).$$

The equality is also true for the other two pairs. Thus, H_1 , H_2 and H_3 are mutually independent.