

### 3.1: Conditional Probability

The *original* or *prior* probability measure utilizes all available information to make probability assignments  $P(A)$ ,  $P(B)$ , etc., subject to the defining conditions (P1), (P2), and (P3). The probability  $P(A)$  indicates the *likelihood* that event  $A$  will occur on any trial.

Frequently, *new information* is received which leads to a reassessment of the likelihood of event  $A$ . For example

- An applicant for a job as a manager of a service department is being interviewed. His résumé shows adequate experience and other qualifications. He conducts himself with ease and is quite articulate in his interview. He is considered a prospect highly likely to succeed. The interview is followed by an extensive background check. His credit rating, because of bad debts, is found to be quite low. With this information, the likelihood that he is a satisfactory candidate changes radically.
- A young woman is seeking to purchase a used car. She finds one that appears to be an excellent buy. It looks “clean,” has reasonable mileage, and is a dependable model of a well known make. Before buying, she has a mechanic friend look at it. He finds evidence that the car has been wrecked with possible frame damage that has been repaired. The likelihood the car will be satisfactory is thus reduced considerably.
- A physician is conducting a routine physical examination on a patient in her seventies. She is somewhat overweight. He suspects that she may be prone to heart problems. Then he discovers that she exercises regularly, eats a low fat, high fiber, variagated diet, and comes from a family in which survival well into their nineties is common. On the basis of this new information, he reassesses the likelihood of heart problems.

New, but partial, information determines a *conditioning event*  $C$ , which may call for reassessing the likelihood of event  $A$ . For one thing, this means that  $A$  occurs iff the event  $AC$  occurs. Effectively, this makes  $C$  a new basic space. The new unit of probability mass is  $P(C)$ . How should the new probability assignments be made? One possibility is to make the new assignment to  $A$  proportional to the probability  $P(AC)$ . These considerations and experience with the classical case suggests the following procedure for reassignment. Although such a reassignment is not logically necessary, subsequent developments give substantial evidence that this is the appropriate procedure.

#### Definition

If  $C$  is an even having positive probabilty, the *conditional probability of  $A$ , given  $C$*  is

$$P(A|C) = \frac{P(AC)}{P(C)}$$

For a fixed conditioning event  $C$ , we have a new likelihood assignment to the event  $A$ . Now

$$P(A|C) \geq 0, P(\Omega|C) = 1, \text{ and } P(\bigvee_j A_j|C) = \frac{P(\bigvee_j A_j C)}{P(C)} = \sum_j P(A_j C)/P(C) = \sum_j P(A_j|C)$$

Thus, the new function  $P(\cdot|C)$  satisfies the three defining properties (P1), (P2), and (P3) for probability, so that for fixed  $C$ , we have a *new probability measure*, with all the properties of an ordinary probability measure.

*Remark.* When we write  $P(A|C)$  we are evaluating the likelihood of event  $A$  when it is known that event  $C$  has occurred. This is *not* the probability of a conditional event  $A|C$ . Conditional events have no meaning in the model we are developing.

#### Example 3.1.1 Conditional probabilities from joint frequency data

A survey of student opinion on a proposed national health care program included 250 students, of whom 150 were undergraduates and 100 were graduate students. Their responses were categorized Y (affirmative), N (negative), and D (uncertain or no opinion). Results are tabulated below.

	Y	N	D
U	60	40	50
G	70	20	10

Suppose the sample is representative, so the results can be taken as typical of the student body. A student is picked at random. Let  $Y$  be the event he or she is favorable to the plan,  $N$  be the event he or she is unfavorable, and  $D$  is the event of no opinion (or uncertain). Let  $U$  be the event the student is an undergraduate and  $G$  be the event he or she is a graduate student. The data may reasonably be interpreted

$$P(G) = 100/250, P(U) = 150/250, P(Y) = (60 + 70)/250, P(YU) = 60/250, \text{ etc.}$$

Then

$$P(Y|U) = \frac{P(YU)}{P(U)} = \frac{60/250}{150/250} = \frac{60}{150}$$

Similarly, we can calculate

$$P(N|U) = 40/150, P(D|U) = 50/150, P(Y|G) = 70/100, P(N|G) = 20/100, P(D|G) = 10/100$$

We may also calculate directly

$$P(U|Y) = 60/130, P(G|N) = 20/60, \text{ etc.}$$

Conditional probability often provides a natural way to deal with compound trials carried out in several steps.

### Example 3.1.2 Jet aircraft with two engines

An aircraft has two jet engines. It will fly with only one engine operating. Let  $F_1$  be the event one engine fails on a long distance flight, and  $F_2$  the event the second fails. Experience indicates that  $P(F_1) = 0.0003$ . Once the first engine fails, added load is placed on the second, so that  $P(F_2|F_1) = 0.001$ . Now the second engine can fail only if the other has already failed. Thus  $F_2 \subset F_1$  so that

$$P(F_2) = P(F_1 F_2) = P(F_1)P(F_2|F_1) = 3 \times 10^{-7}$$

Thus reliability of any one engine may be less than satisfactory, yet the overall reliability may be quite high.

The following example is taken from the UMAP Module 576, by Paul Mullenix, reprinted in UMAP Journal, vol 2, no. 4. More extensive treatment of the problem is given there.

### Example 3.1.3 Responses to a sensitive question on a survey

In a survey, if answering “yes” to a question may tend to incriminate or otherwise embarrass the subject, the response given may be incorrect or misleading. Nonetheless, it may be desirable to obtain correct responses for purposes of social analysis. The following device for dealing with this problem is attributed to B. G. Greenberg. By a chance process, each subject is instructed to do one of three things:

1. Respond with an honest answer to the question.
2. Respond “yes” to the question, regardless of the truth in the matter.
3. Respond “no” regardless of the true answer.

Let  $A$  be the event the subject is told to reply honestly,  $B$  be the event the subject is instructed to reply “yes,” and  $C$  be the event the answer is to be “no.” The probabilities  $P(A)$ ,  $P(B)$ , and  $P(C)$  are determined by a chance mechanism (i.e., a fraction  $P(A)$  selected randomly are told to answer honestly, etc.). Let  $E$  be the event the reply is “yes.” We wish to calculate  $P(E|A)$ , the probability the answer is “yes” given the response is honest.

#### Solution

Since  $E = EA \vee B$ , we have

$$P(E) = P(EA) + P(B) = P(E|A)P(A) + P(B)$$

which may be solved algebraically to give

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

Suppose there are 250 subjects. The chance mechanism is such that  $P(A) = 0.7$ ,  $P(B) = 0.4$  and  $P(C) = 0.16$ . There are 62 responses “yes,” which we take to mean  $P(E) = 62/250$ . According to the pattern above

$$P(E|A) = \frac{62/250 - 14/100}{70/100} = \frac{27}{175} \approx 0.154$$

The formulation of conditional probability assumes the conditioning event  $C$  is well defined. Sometimes there are subtle difficulties. It may not be entirely clear from the problem description what the conditioning event is. This is usually due to some ambiguity or misunderstanding of the information provided.

#### Example 3.1.4 What is the conditioning event?

Five equally qualified candidates for a job, Jim, Paul, Richard, Barry, and Evan, are identified on the basis of interviews and told that they are finalists. Three of these are to be selected at random, with results to be posted the next day. One of them, Jim, has a friend in the personnel office. Jim asks the friend to tell him the name of one of those selected (other than himself). The friend tells Jim that Richard has been selected. Jim analyzes the problem as follows.

##### Analysis

Let  $A_i$ ,  $1 \leq i \leq 5$  be the event the  $i$ th of these is hired ( $A_1$  is the event Jim is hired,  $A_3$  is the event Richard is hired, etc.). Now  $P(A_i)$  (for each  $i$ ) is the probability that finalist  $i$  is in one of the combinations of three from five. Thus, Jim's probability of being hired, before receiving the information about Richard, is

$$P(A_1) = \frac{1 \times C(4, 2)}{C(5, 3)} = \frac{6}{10} = P(A_i), 1 \leq i \leq 5$$

The information that Richard is one of those hired is information that the event  $A_3$  has occurred. Also, for any pair  $i \neq j$  the number of combinations of three from five including these two is just the number of ways of picking one from the remaining three. Hence,

$$P(A_1 A_3) = \frac{C(3, 1)}{C(5, 3)} = \frac{3}{10} = P(A_i A_j), i \neq j$$

The conditional probability

$$P(A_1|A_3) = \frac{P(A_1 A_3)}{P(A_3)} = \frac{3/10}{6/10} = 1/2$$

This is consistent with the fact that if Jim knows that Richard is hired, then there are two to be selected from the four remaining finalists, so that

$$P(A_1|A_3) = \frac{1 \times C(3, 1)}{C(4, 2)} = \frac{3}{6} = 1/2$$

##### Discussion

Although this solution seems straightforward, it has been challenged as being incomplete. Many feel that there must be information about *how* the friend chose to name Richard. Many would make an assumption somewhat as follows. The friend took the three names selected: if Jim was one of them, Jim's name was removed and an equally likely choice among the other two was made; otherwise, the friend selected on an equally likely basis one of the three to be hired. Under this assumption, the information assumed is an event  $B_3$  which is *not* the same as  $A_3$ . In fact, computation (see Example 5, below) shows

$$P(A_1|B_3) = \frac{6}{10} = P(A_1) \neq P(A_1|A_3)$$

Both results are mathematically correct. The difference is in the conditioning event, which corresponds to the difference in the information given (or assumed).

#### Some properties

In addition to its properties as a probability measure, conditional probability has *special properties* which are consequences of the way it is related to the original probability measure  $P(\cdot)$ . The following are easily derived from the definition of conditional probability and basic properties of the prior probability measure, and prove useful in a variety of problem situations.

**(CP1) Product rule** If  $P(ABCD) > 0$ , then  $P(ABCD) = P(A)P(B|A)P(C|AB)P(D|ABC)$ .

Derivation

The defining expression may be written in product form:  $P(AB) = P(A)P(B|A)$ . Likewise

$$P(ABC) = P(A) \frac{P(AB)}{P(A)} \cdot \frac{P(ABC)}{P(AB)} = P(A)P(B|A)P(C|AB)$$

and

$$P(ABCD) = P(A) \frac{P(AB)}{P(A)} \cdot \frac{P(ABC)}{P(AB)} \cdot \frac{P(ABCD)}{P(ABC)} = P(A)P(B|A)P(C|AB)P(D|ABC)$$

This pattern may be extended to the intersection of any finite number of events. Also, the events may be taken in any order.

—□

### Example 3.1.5 Selection of items from a lot

An electronics store has ten items of a given type in stock. One is defective. Four successive customers purchase one of the items. Each time, the selection is on an equally likely basis from those remaining. What is the probability that all four customers get good items?

#### Solution

Let  $E_i$  be the event the  $i$ th customer receives a good item. Then the first chooses one of the nine out of ten good ones, the second chooses one of the eight out of nine good ones, etc., so that

$$P(E_1 E_2 E_3 E_4) = P(E_1)P(E_2|E_1)P(E_3|E_1 E_2)P(E_4|E_1 E_2 E_3) = \frac{9}{10} \cdot \frac{8}{9} \cdot \frac{7}{8} \cdot \frac{6}{7} = \frac{6}{10}$$

Note that this result could be determined by a combinatorial argument: under the assumptions, each combination of four of ten is equally likely; the number of combinations of four good ones is the number of combinations of four of the nine. Hence

$$P(E_1 E_2 E_3 E_4) = \frac{C(9, 4)}{C(10, 4)} = \frac{126}{210} = 3/5$$

### Example 3.1.6 A selection problem

Three items are to be selected (on an equally likely basis at each step) from ten, two of which are defective. Determine the probability that the first and third selected are good.

#### Solution

Let  $G_i, 1 \leq i \leq 3$  be the event the  $i$ th unit selected is good. Then  $G_1 G_3 = G_1 G_2 G_3 \vee G_1 G_2^c G_3$ . By the product rule

$$P(G_1 G_3) = P(G_1)P(G_2|G_1)P(G_3|G_1 G_2) + P(G_1)P(G_2^c|G_1)P(G_3|G_1 G_2^c) = \frac{8}{10} \cdot \frac{7}{9} \cdot \frac{6}{8} + \frac{8}{10} \cdot \frac{2}{9} \cdot \frac{7}{8} = \frac{28}{45} \approx 0.6$$

**(CP2) Law of total probability** Suppose the class  $\{A_i : 1 \leq i \leq n\}$  of events is mutually exclusive and every outcome in  $E$  is in one of these events. Thus,  $E = A_1 E \vee A_2 E \vee \cdots \vee A_n E$ , a disjoint union. Then

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

### Example 3.1.7 a compound experiment

Five cards are numbered one through five. A two-step selection procedure is carried out as follows.

1. Three cards are selected without replacement, on an equally likely basis.
  - If card 1 is drawn, the other two are put in a box
  - If card 1 is not drawn, all three are put in a box
2. One of cards in the box is drawn on an equally likely basis (from either two or three)

Let  $A_i$  be the event the  $i$ th card is drawn on the first selection and let  $B_i$  be the event the card numbered  $i$  is drawn on the second selection (from the box). Determine  $P(B_5)$ ,  $P(A_1 B_5)$ , and  $P(A_1 | B_5)$ .

### Solution

From Example 3.1.4, we have  $P(A_i) = 6/10$  and  $P(A_i A_j) = 3/10$ . This implies

$$P(A_i A_j^c) = P(A_i) - P(A_i A_j) = 3/10$$

Now we can draw card five on the second selection only if it is selected on the first drawing, so that  $B_5 \subset A_5$ . Also  $A_5 = A_1 A_5 \cup A_1^c A_5$ . We therefore have  $B_5 = B_5 A_5 = B_5 A_1 A_5 \cup B_5 A_1^c A_5$ . By the law of total probability (CP2),

$$P(B_5) = P(B_5 | A_1 A_5)P(A_1 A_5) + P(B_5 | A_1^c A_5)P(A_1^c A_5) = \frac{1}{2} \cdot \frac{3}{10} + \frac{1}{3} \cdot \frac{3}{10} = \frac{1}{4}$$

Also, since  $A_1 B_5 = A_1 A_5 B_5$ ,

$$P(A_1 B_5) = P(A_1 A_5 B_5) = P(A_1 A_5)P(B_5 | A_1 A_5) = \frac{3}{10} \cdot \frac{1}{2} = \frac{3}{20}$$

We thus have

$$P(A_1 | B_5) = \frac{3/20}{1/4} = \frac{6}{10} = P(A_1)$$

Occurrence of event  $B_1$  has no effect on the likelihood of the occurrence of  $A_1$ . This condition is examined more thoroughly in the chapter on "Independence of Events".

Often in applications data lead to conditioning with respect to an event but the problem calls for "conditioning in the opposite direction."

### Example 3.1.8 Reversal of conditioning

Students in a freshman mathematics class come from three different high schools. Their mathematical preparation varies. In order to group them appropriately in class sections, they are given a diagnostic test. Let  $H_i$  be the event that a student tested is from high school  $i$ ,  $1 \leq i \leq 3$ . Let  $F$  be the event the student fails the test. Suppose data indicate

$$P(H_1) = 0.2, P(H_2) = 0.5, P(H_3) = 0.3, P(F | H_1) = 0.10, P(F | H_2) = 0.02, P(F | H_3) = 0.06$$

A student passes the exam. Determine for each  $i$  the conditional probability  $P(H_i | F^c)$  that the student is from high school  $i$ .

### Solution

$$P(F^c) = P(F^c | H_1)P(H_1) + P(F^c | H_2)P(H_2) + P(F^c | H_3)P(H_3) = 0.90 \cdot 0.2 + 0.98 \cdot 0.5 + 0.94 \cdot 0.3 = 0.952$$

Then

$$P(H_1 | F^c) = \frac{P(F^c | H_1)P(H_1)}{P(F^c)} = \frac{0.10 \cdot 0.2}{0.952} = \frac{180}{952} = 0.1891$$

Similarly,

$$P(H_2 | F^c) = \frac{P(F^c | H_2)P(H_2)}{P(F^c)} = \frac{0.02 \cdot 0.5}{0.952} = 0.5147 \text{ and } P(H_3 | F^c) = \frac{P(F^c | H_3)P(H_3)}{P(F^c)} = \frac{0.06 \cdot 0.3}{0.952} = 0.2962$$

The basic pattern utilized in the reversal is the following.

**(CP3) Bayes' rule** If  $E \subset \bigcup_{i=1}^n A_i$  (as in the law of total probability), then

$$P(A_i | E) = \frac{P(A_i E)}{P(E)} = \frac{P(E | A_i)P(A_i)}{P(E)} \quad 1 \leq i \leq n$$

The law of total probability yields  $P(E)$

Such reversals are desirable in a variety of practical situations.

### Example 3.1.9 A compound selection and reversal

Begin with items in two lots:

1. Three items, one defective.
2. Four items, one defective.

One item is selected from lot 1 (on an equally likely basis); this item is added to lot 2; a selection is then made from lot 2 (also on an equally likely basis). This second item is good. What is the probability the item selected from lot 1 was good?

#### Solution

Let  $G_1$  be the event the first item (from lot 1) was good, and  $G_2$  be the event the second item (from the augmented lot 2) is good. We want to determine  $P(G_1|G_2)$ . Now the data are interpreted as

$$P(G_1) = 2/3, P(G_2|G_1) = 4/5, P(G_2|G_1^c) = 3/5$$

By the law of total probability (CP2),

$$P(G_2) = P(G_1)P(G_2|G_1) + P(G_1^c)P(G_2|G_1^c) = \frac{2}{3} \cdot \frac{4}{5} + \frac{1}{3} \cdot \frac{3}{5} = \frac{11}{15}$$

By Bayes' rule (CP3),

$$P(G_1|G_2) = \frac{P(G_2|G_1)P(G_1)}{P(G_2)} = \frac{4/5 \times 2/3}{11/15} = \frac{8}{11} \approx 0.73$$

### Example 3.1.10 Additional problems requiring reversals

- *Medical tests.* Suppose  $D$  is the event a patient has a certain disease and  $T$  is the event a test for the disease is positive. Data are usually of the form: prior probability  $P(D)$  (or prior odds  $P(D)/P(D^c)$ ), probability  $P(T|D^c)$  of a *false positive*, and probability  $P(T^c|D)$  of a *false negative*. The desired probabilities are  $P(D|T)$  and  $P(D^c|T^c)$ .
- *Safety alarm.* If  $D$  is the event a dangerous condition exists (say a steam pressure is too high) and  $T$  is the event the safety alarm operates, then data are usually of the form  $P(D)$ ,  $P(T|D^c)$ , and  $P(T^c|D)$ , or equivalently (e.g.,  $P(T^c|D^c)$  and  $P(T|D)$ ). Again, the desired probabilities are that the safety alarms signals correctly,  $P(D|T)$  and  $P(D^c|T^c)$ .
- *Job success.* If  $H$  is the event of success on a job, and  $E$  is the event that an individual interviewed has certain desirable characteristics, the data are usually prior  $P(H)$  and reliability of the characteristics as predictors in the form  $P(H)$  and  $P(E|H^c)$ . The desired probability is  $P(H|E)$ .
- *Presence of oil.* If  $H$  is the event of the presence of oil at a proposed well site, and  $E$  is the event of certain geological structure (salt dome or fault), the data are usually  $P(H)$  (or the odds),  $P(E|H)$ , and  $P(E|H^c)$ . The desired probability is  $P(H|E)$ .
- *Market condition.* Before launching a new product on the national market, a firm usually examines the condition of a test market as an indicator of the national market. If  $H$  is the event the national market is favorable and  $E$  is the event the test market is favorable, data are a prior estimate  $P(H)$  of the likelihood the national market is sound, and data  $P(E|H)$  and  $P(E|H^c)$  indicating the reliability of the test market. What is desired is  $P(H|E)$ , the likelihood the national market is favorable, given the test market is favorable.

The calculations, as in Example 3.8, are simple but can be tedious. We have an m-procedure called *bayes* to perform the calculations easily. The probabilities  $P(A_i)$  are put into a matrix PA and the conditional probabilities  $P(E|A_i)$  are put into matrix PEA. The desired probabilities  $P(A_i|E)$  and  $PA_i|E^c$  are calculated and displayed

### Example 3.1.11 matlab calculations for

```
>> PEA = [0.10 0.02 0.06];
>> PA = [0.2 0.5 0.3];
>> bayes
Requires input PEA = [P(E|A1) P(E|A2) ... P(E|An)]
and PA = [P(A1) P(A2) ... P(An)]
```

Determines PAE = [P(A1|E) P(A2|E) ... P(An|E)]  
 and PAEc = [P(A1|Ec) P(A2|Ec) ... P(An|Ec)]  
 Enter matrix PEA of conditional probabilities PEA  
 Enter matrix PA of probabilities PA

P(E) = 0.048

P(E Ai)	P(Ai)	P(Ai E)	P(Ai Ec)
0.1000	0.2000	0.4167	0.1891
0.0200	0.5000	0.2083	0.5147
0.0600	0.3000	0.3750	0.2962

Various quantities are in the matrices PEA, PA, PAE, PAEc, named above

The procedure displays the results in tabular form, as shown. In addition, the various quantities are in the workspace in the matrices named, so that they may be used in further calculations without recopying.

The following variation of Bayes' rule is applicable in many practical situations.

**(CP3\*) Ratio form of Bayes' rule**  $\frac{P(A|C)}{P(B|C)} = \frac{P(AC)}{P(BC)} = \frac{P(C|A)}{P(C|B)} \cdot \frac{P(A)}{P(B)}$

The left hand member is called the *posterior odds*, which is the odds *after* knowledge of the occurrence of the conditioning event. The second fraction in the right hand member is the *prior odds*, which is the odds *before* knowledge of the occurrence of the conditioning event  $C$ . The first fraction in the right hand member is known as the *likelihood ratio*. It is the ratio of the probabilities (or likelihoods) of  $C$  for the two different probability measures  $P(\cdot|A)$  and  $P(\cdot|B)$ .

#### Example 3.1.12 A performance test

As a part of a routine maintenance procedure, a computer is given a performance test. The machine seems to be operating so well that the prior odds it is satisfactory are taken to be ten to one. The test has probability 0.05 of a false positive and 0.01 of a false negative. A test is performed. The result is positive. What are the posterior odds the device is operating properly?

#### Solution

Let  $S$  be the event the computer is operating satisfactorily and let  $T$  be the event the test is favorable. The data are  $P(S)/P(S^c) = 10$ ,  $P(T|S^c) = 0.05$ , and  $P(T^c|S) = 0.01$ . Then by the ratio form of Bayes' rule

$$\frac{P(S|T)}{P(S^c|T)} = \frac{P(T|S)}{P(T|S^c)} \cdot \frac{P(S)}{P(S^c)} = \frac{0.99}{0.05} \cdot 10 = 198 \quad \text{so that } P(S|T) = \frac{198}{199} = 0.9950$$

The following property serves to establish in the chapters on "Independence of Events" and "Conditional Independence" a number of important properties for the concept of independence and of conditional independence of events.

**(CP4) Some equivalent conditions** If  $0 < P(A) < 1$  and  $0 < P(B) < 1$ , then

$$P(A|B) * P(A) \text{ iff } P(B|A) * P(B) \text{ iff } P(AB) * P(A)P(B) \text{ and} \\ P(AB) * P(A)P(B) \text{ iff } P(A^c B^c) * P(A^c)P(B^c) \text{ iff } P(AB^c) \diamond P(A)P(B^c)$$

where  $*$  is  $<$ ,  $\leq$ ,  $=$ ,  $\geq$ , or  $>$  and  $\diamond$  is  $>$ ,  $\geq$ ,  $=$ ,  $\leq$ , or  $<$ , respectively.

Because of the role of this property in the theory of independence and conditional independence, we examine the derivation of these results.

VERIFICATION of (CP4)

$$\begin{aligned} P(AB) * P(A)P(B) &\text{ iff } P(A|B) * P(A) \text{ (divide by } P(B) \text{ - may exchange } A \text{ and } A^c) \\ P(AB) * P(A)P(B) &\text{ iff } P(B|A) * P(B) \text{ (divide by } P(A) \text{ - may exchange } B \text{ and } B^c) \\ P(AB) * P(A)P(B) &\text{ iff } [P(A) - P(AB^c)] * P(A)[1 - P(B^c)] \text{ iff } -P(AB^c) * -P(A)P(B^c) \text{ iff} \\ P(AB^c) \diamond P(A)P(B^c) & \\ \text{we may use c to get } P(AB) * P(A)P(B) &\text{ iff } P(AB^c) \diamond P(A)P(B^c) \text{ iff } P(A^c B^c) * P(A^c)P(B^c) \end{aligned}$$

—□

A number of important and useful propositions may be derived from these.

$$P(A|B) + P(A^c|B) = 1, \text{ but, in general, } P(A|B) + P(A|B^c) \neq 1.$$

$$P(A|B) > P(A) \text{ iff } P(A|B^c) < P(A).$$

$$P(A^c|B) > P(A^c) \text{ iff } P(A|B) < P(A).$$

$$P(A|B) > P(A) \text{ iff } P(A^c|B^c) > P(A^c).$$

VERIFICATION — Exercises (see problem set)

—□

## Repeated conditioning

Suppose conditioning by the event  $C$  has occurred. Additional information is then received that event  $D$  has occurred. We have a new conditioning event  $CD$ . There are two possibilities:

Reassign the conditional probabilities.  $P_C(A)$  becomes

$$P_C(A|D) = \frac{P_C(AD)}{P_C(D)} = \frac{P(ACD)}{P(CD)}$$

Reassign the total probabilities:  $P(A)$  becomes

$$P_{CD}(A) = P(A|CD) = \frac{P(ACD)}{P(CD)}$$

*Basic result:*  $P_C(A|D) = P(A|CD) = P_D(A|C)$ . Thus repeated conditioning by two events may be done in any order, or may be done in one step. This result extends easily to repeated conditioning by any finite number of events. This result is important in extending the concept of "[Independence of Events](#)" to "[Conditional Independence](#)". These conditions are important for many problems of probable inference.

---

This page titled [3.1: Conditional Probability](#) is shared under a [CC BY 3.0](#) license and was authored, remixed, and/or curated by [Paul Pfeiffer](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.