

5.1: Conditional Independence

5.1. Conditional Independence*

The idea of stochastic (probabilistic) independence is explored in the unit [Independence of Events](#). The concept is approached as lack of conditioning: $P(A|B) = P(A)$. This is equivalent to the product rule $P(AB) = P(A)P(B)$. We consider an extension to conditional independence.

The concept

Examination of the independence concept reveals two important mathematical facts:

- Independence of a class of non mutually exclusive events depends upon the probability measure, and not on the relationship between the events. Independence cannot be displayed on a Venn diagram, unless probabilities are indicated. For one probability measure a pair may be independent while for another probability measure the pair may not be independent.
- Conditional probability is a probability measure, since it has the three defining properties and all those properties derived therefrom.

This raises the question: is there a useful conditional independence—i.e., independence with respect to a conditional probability measure? In this chapter we explore that question in a fruitful way.

Among the simple examples of “operational independence” in the unit on independence of events, which lead naturally to an assumption of “probabilistic independence” are the following:

- If customers come into a well stocked shop at different times, each unaware of the choice made by the other, the the item purchased by one should not be affected by the choice made by the other.
- If two students are taking exams in different courses, the grade one makes should not affect the grade made by the other.

Example 5.1.1 Buying umbrellas and the weather

A department store has a nice stock of umbrellas. Two customers come into the store “independently.” Let A be the event the first buys an umbrella and B the event the second buys an umbrella. Normally, we should think the events $\{A, B\}$ form an independent pair. But consider the effect of weather on the purchases. Let C be the event the weather is rainy (i.e., is raining or threatening to rain). Now we should think $P(A|C) > P(A|C^c)$ and $P(B|C) > P(B|C^c)$. The weather has a decided effect on the likelihood of buying an umbrella. But given the fact the weather is rainy (event C has occurred), it would seem reasonable that purchase of an umbrella by one should not affect the likelihood of such a purchase by the other. Thus, it may be reasonable to suppose

$$P(A|C) = P(A|BC) \text{ or, in another notation, } P_C(A) = P_C(A|B)$$

An examination of the sixteen equivalent conditions for independence, with probability measure P replaced by probability measure P_C , shows that we have independence of the pair $\{A, B\}$ with respect to the conditional probability measure $P_C(\cdot) = P(\cdot|C)$. Thus, $P(A|C^c) = P(A|BC^c)$. For this example, we should also expect that $P(A|C^c) = P(A|BC^c)$, so that there is independence with respect to the conditional probability measure $P(\cdot|C^c)$. Does this make the pair $\{A, B\}$ independent (with respect to the prior probability measure P)? Some numerical examples make it plain that only in the most unusual cases would the pair be independent. Without calculations, we can see why this should be so. If the first customer buys an umbrella, this indicates a higher than normal likelihood that the weather is rainy, in which case the second customer is likely to buy. The condition leads to $P(B|A) > P(B)$. Consider the following numerical case. Suppose $P(AB|C) = P(A|C)P(B|C)$ and $P(AB|C^c) = P(A|C^c)P(B|C^c)$ and

$$P(A|C) = 0.60, P(A|C^c) = 0.20, P(B|C) = 0.50, P(B|C^c) = 0.15, \text{ with } P(C) = 0.30.$$

Then

$$P(A) = P(A|C)P(C) + P(A|C^c)P(C^c) = 0.3200 \quad P(B) = P(B|C)P(C) + P(B|C^c)P(C^c) = 0.2550$$

$$P(AB) = P(AB|C)P(C) + P(AB|C^c)P(C^c) = P(A|C)P(B|C)P(C) + P(A|C^c)P(B|C^c)P(C^c) = 0.1110$$

As a result,

$$P(A)P(B) = 0.0816 \neq 0.1110 = P(AB)$$

The product rule fails, so that the pair is *not* independent. An examination of the pattern of computation shows that independence would require very special probabilities which are not likely to be encountered.

Example 5.1.2 Students and exams

Two students take exams in different courses. Under normal circumstances, one would suppose their performances form an independent pair. Let A be the event the first student makes grade 80 or better and B be the event the second has a grade of 80 or better. The exam is given on Monday morning. It is the fall semester. There is a probability 0.30 that there was a football game on Saturday, and both students are enthusiastic fans. Let C be the event of a game on the previous Saturday. Now it is reasonable to suppose

$$P(A|C) = P(A|BC) \text{ and } P(A|C^c) = P(A|BC^c)$$

If we know that there was a Saturday game, additional knowledge that B has occurred does not affect the likelihood that A occurs. Again, use of equivalent conditions shows that the situation may be expressed

$$P(AB|C) = P(A|C)P(B|C) \text{ and } P(AB|C^c) = P(A|C^c)P(B|C^c)$$

Under these conditions, we should suppose that $P(A|C) < P(A|C^c)$ and $P(B|C) < P(B|C^c)$. If we knew that one did poorly on the exam, this would increase the likelihood there was a Saturday game and hence increase the likelihood that the other did poorly. The failure to be independent arises from a common chance factor that affects both. Although their performances are “operationally” independent, they are not independent in the probability sense. As a numerical example, suppose

$$P(A|C) = 0.7 \quad P(A|C^c) = 0.9 \quad P(B|C) = 0.6 \quad P(B|C^c) = 0.8 \quad P(C) = 0.3$$

Straightforward calculations show $P(A) = 0.8400$, $P(B) = 0.7400$, $P(AB) = 0.6300$. Note that $P(A|B) = 0.8514 > P(A)$ as would be expected.

Sixteen equivalent conditions

Using the facts on [repeated conditioning](#) and the [equivalent conditions for independence](#), we may produce a similar table of equivalent conditions for conditional independence. In the hybrid notation we use for repeated conditioning, we write

$$P_C(A|B) = P_C(A) \text{ or } P_C(AB) = P_C(A)P_C(B)$$

This translates into

$$P(A|BC) = P(A|C) \text{ or } P(AB|C) = P(A|C)P(B|C)$$

If it is known that C has occurred, then additional knowledge of the occurrence of B does not change the likelihood of A .

If we write the sixteen equivalent conditions for independence in terms of the conditional probability measure $P_C(\cdot)$, then translate as above, we have the following equivalent conditions.

Table 5.1. Sixteen equivalent conditions

$P(A BC) = P(A C)$	$P(B AC) = P(B C)$	$P(AB C) = P(A C)P(B C)$
$P(A B^cC) = P(A C)$	$P(B^c AC) = P(B^c C)$	$P(AB^c C) = P(A C)P(B^c C)$
$P(A^c BC) = P(A^c C)$	$P(B A^cC) = P(B C)$	$P(A^cB C) = P(A^c C)P(B C)$
$P(A^c B^cC) = P(A^c C)$	$P(B^c A^cC) = P(B^c C)$	$P(A^cB^c C) = P(A^c C)P(B^c C)$

Table 5.2.

$P(A BC) = P(A B^cC)$	$P(A^c B^cC) = P(A^c B^cC)$	$P(B AC) = P(B A^cC)$	$P(B^c AC) = P(B^c A^cC)$
-----------------------	-----------------------------	-----------------------	---------------------------

The patterns of conditioning in the examples above belong to this set. In a given problem, one or the other of these conditions may seem a reasonable assumption. As soon as *one* of these patterns is recognized, then *all* are equally valid assumptions. Because of its simplicity and symmetry, we take as the defining condition the *product rule* $P(AB|C) = P(A|C)P(B|C)$.

Definition

A pair of events $\{A, B\}$ is said to be *conditionally independent, given C* , designated $\{A, B\}$ iff the following product rule holds: $P(AB|C) = P(A|C)P(B|C)$.

The equivalence of the four entries in the right hand column of the upper part of the table, establish

The replacement rule

If any of the pairs $\{A, B\}$, $\{A, B^c\}$, $\{A^c, B\}$ or $\{A^c, B^c\}$ is conditionally independent, given C , then so are the others.

—□

This may be expressed by saying that if a pair is conditionally independent, we may replace either or both by their complements and still have a conditionally independent pair.

To illustrate further the usefulness of this concept, we note some other common examples in which similar conditions hold: there is operational independence, but some chance factor which affects both.

- Two contractors work quite independently on jobs in the same city. The operational independence suggests probabilistic independence. However, both jobs are outside and subject to delays due to bad weather. Suppose A is the event the first contractor completes his job on time and B is the event the second completes on time. If C is the event of “good” weather, then arguments similar to those in Examples 1 and 2 make it seem reasonable to suppose $\{A, B\} \text{ ci } |C$ and $\{A, B\} \text{ ci } |C^c$. *Remark.* In formal probability theory, an event must be sharply defined: on any trial it occurs or it does not. The event of “good weather” is not so clearly defined. Did a trace of rain or thunder in the area constitute bad weather? Did rain delay on one day in a month long project constitute bad weather? Even with this ambiguity, the pattern of probabilistic analysis may be useful.
- A patient goes to a doctor. A preliminary examination leads the doctor to think there is a thirty percent chance the patient has a certain disease. The doctor orders two independent tests for conditions that indicate the disease. Are results of these tests really independent? There is certainly operational independence—the tests may be done by different laboratories, neither aware of the testing by the others. Yet, if the tests are meaningful, they must both be affected by the actual condition of the patient. Suppose D is the event the patient has the disease, A is the event the first test is positive (indicates the conditions associated with the disease) and B is the event the second test is positive. Then it would seem reasonable to suppose $\{A, B\} \text{ ci } |D$ and $\{A, B\} \text{ ci } |D^c$.

In the examples considered so far, it has been reasonable to assume conditional independence, given an event C , and conditional independence, given the complementary event. But there are cases in which the effect of the conditioning event is asymmetric. We consider several examples.

- Two students are working on a term paper. They work quite separately. They both need to borrow a certain book from the library. Let C be the event the library has two copies available. If A is the event the first completes on time and B the event the second is successful, then it seems reasonable to assume $\{A, B\} \text{ ci } |C$. However, if only one book is available, then the two conditions would *not* be conditionally independent. In general $P(B|AC^c) < P(B|C^c)$, since if the first student completes on time, then he or she must have been successful in getting the book, to the detriment of the second.
- If the two contractors of the example above both need material which may be in scarce supply, then successful completion would be conditionally independent, given an adequate supply, whereas they would not be conditionally independent, given a short supply.
- Two students in the same course take an exam. If they prepared separately, the event of both getting good grades should be conditionally independent. If they study together, then the likelihoods of good grades would not be independent. With neither cheating or collaborating on the test itself, if one does well, the other should also.

Since conditional independence is ordinary independence with respect to a conditional probability measure, it should be clear how to extend the concept to larger classes of sets.

Definition

A class $\{A_i : i \in J\}$, where J is an arbitrary index set, is *conditionally independent, given event C* , denoted $\{A_i : i \in J\} \text{ ci } |C$, iff the product rule holds for every finite subclass of two or more.

As in the case of simple independence, the replacement rule extends.

The replacement rule

If the class $\{A_i : i \in J\}$ ci $|C$, then any or all of the events A_i may be replaced by their complements and still have a conditionally independent class.

The use of independence techniques

Since conditional independence is independence, we may use independence techniques in the solution of problems. We consider two types of problems: an inference problem and a conditional Bernoulli sequence.

Example 5.1.3 Use of independence techniques

Sharon is investigating a business venture which she thinks has probability 0.7 of being successful. She checks with five “independent” advisers. If the prospects are sound, the probabilities are 0.8, 0.75, 0.6, 0.9, and 0.8 that the advisers will advise her to proceed; if the venture is not sound, the respective probabilities are 0.75, 0.85, 0.7, 0.9, and 0.7 that the advice will be negative. Given the quality of the project, the advisers are independent of one another in the sense that no one is affected by the others. Of course, they are not independent, for they are all related to the soundness of the venture. We may reasonably assume conditional independence of the advice, given that the venture is sound and also given that the venture is not sound. If Sharon goes with the majority of advisers, what is the probability she will make the right decision?

Solution

If the project is sound, Sharon makes the right choice if three or more of the five advisers are positive. If the venture is unsound, she makes the right choice if three or more of the five advisers are negative. Let H = the event the project is sound, F = the event three or more advisers are positive, $G = F^c$ = the event three or more are negative, and E = the event of the correct decision. Then

$$P(E) = P(FH) + P(GH^c) = P(F|H)P(H) + P(G|H^c)P(H^c)$$

Let E_i be the event the i th adviser is positive. Then $P(F|H)$ = the sum of probabilities of the form $P(M_k|H)$, where M_k are minterms generated by the class $\{E_i : 1 \leq i \leq 5\}$. Because of the assumed conditional independence,

$$P(E_1 E_2^c E_3^c E_4 E_5 | H) = P(E_1|H)P(E_2^c|H)P(E_3^c|H)P(E_4|H)P(E_5|H)$$

with similar expressions for each $P(M_k|H)$ and $P(M_k|H^c)$. This means that if we want the probability of three or more successes, given H , we can use ckn with the matrix of conditional probabilities. The following MATLAB solution of the investment problem is indicated.

```
P1 = 0.01*[80 75 60 90 80];
P2 = 0.01*[75 85 70 90 70];
PH = 0.7;
PE = ckn(P1,3)*PH + ckn(P2,3)*(1 - PH)
PE = 0.9255
```

Often a Bernoulli sequence is related to some conditioning event H . In this case it is reasonable to assume the sequence $\{E_i : 1 \leq i \leq n\}$ ci $|H$ and ci $|H^c$. We consider a simple example.

Example 5.1.4 Test of a claim

A race track regular claims he can pick the winning horse in any race 90 percent of the time. In order to test his claim, he picks a horse to win in each of ten races. There are five horses in each race. If he is simply guessing, the probability of success on each race is 0.2. Consider the trials to constitute a Bernoulli sequence. Let H be the event he is correct in his claim. If S is the number of successes in picking the winners in the ten races, determine $P(H|S = k)$ for various numbers k of correct picks. Suppose it is equally likely that his claim is valid or that he is merely guessing. We assume two conditional Bernoulli trials:

claim is valid: Ten trials, probability $p = P(E_i|H) = 0.9$.

Guessing at random: Ten trials, probability $p = P(E_i|H^c) = 0.2$.

Let S = number of correct picks in ten trials. Then

$$\frac{P(H|S=k)}{P(H^c|S=k)} = \frac{P(H)}{P(H^c)} \cdot \frac{P(S=k|H)}{P(S=k|H^c)}, 0 \leq k \leq 10$$

Giving him the benefit of the doubt, we suppose $P(H)/P(H^c) = 1$ and calculate the conditional odds.

```
k = 0:10;
Pk1 = ibinom(10,0.9,k);    % Probability of k successes, given H
Pk2 = ibinom(10,0.2,k);    % Probability of k successes, given H^c
OH  = Pk1./Pk2;           % Conditional odds-- Assumes P(H)/P(H^c) = 1
e   = OH > 1;             % Selects favorable odds
disp(round([k(e);OH(e)]'))

      6      2    % Needs at least six to have creditability
      7     73    % Seven would be creditable,
      8    2627   % even if P(H)/P(H^c) = 0.1
      9   94585
     10  3405063
```

Under these assumptions, he would have to pick at least seven correctly to give reasonable validation of his claim.

This page titled [5.1: Conditional Independence](#) 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.