

5.2: Patterns of Probable Inference

Some Patterns of Probable Inference

We are concerned with the likelihood of some hypothesized condition. In general, we have evidence for the condition which can never be absolutely certain. We are forced to assess probabilities (likelihoods) on the basis of the evidence. Some typical examples:

Table 5.3.

HYPOTHESIS	EVIDENCE
Job success	Personal traits
Presence of oil	Geological structures
Operation of a device	Physical condition
Market condition	Test market condition
Presence of a disease	Tests for symptoms

If H is the event the hypothetical condition exists and E is the event the evidence occurs, the probabilities available are usually $P(H)$ (or an odds value), $P(E|H)$, and $P(H^c|E)$. What is desired is $P(H|E)$ or, equivalently, the odds $P(H|E)/P(H^c|E)$. We simply use Bayes' rule to reverse the direction of conditioning.

$$\frac{P(H|E)}{P(H^c|E)} = \frac{P(E|H)}{P(E|H^c)} \cdot \frac{P(H)}{P(H^c)}$$

No conditional independence is involved in this case.

Independent evidence for the hypothesized condition

Suppose there are two “independent” bits of evidence. Now obtaining this evidence may be “operationally” independent, but if the items both relate to the hypothesized condition, then they cannot be really independent. The condition assumed is usually of the form $P(E_1|H) = P(E_1|HE_2)$ —if H occurs, then knowledge of E_2 does not affect the likelihood of E_1 . Similarly, we usually have $P(E_1|H^c) = P(E_1|H^cE_2)$. Thus $\{E_1, E_2\}$ ci $|H$ and $\{E_1, E_2\}$ ci $|H^c$.

Example 5.2.1 Independent medical tests

Suppose a doctor thinks the odds are 2/1 that a patient has a certain disease. She orders two independent tests. Let H be the event the patient has the disease and E_1 and E_2 be the events the tests are positive. Suppose the first test has probability 0.1 of a false positive and probability 0.05 of a false negative. The second test has probabilities 0.05 and 0.08 of false positive and false negative, respectively. If both tests are positive, what is the posterior probability the patient has the disease?

Solution

Assuming $\{E_1, E_2\}$ ci $|H$ and ci $|H^c$, we work first in terms of the odds, then convert to probability.

$$\frac{P(H|E_1E_2)}{P(H^c|E_1E_2)} = \frac{P(H)}{P(H^c)} \cdot \frac{P(E_1E_2|H)}{P(E_1E_2|H^c)} = \frac{P(H)}{P(H^c)} \cdot \frac{P(E_1|H)P(E_2|H)}{P(E_1|H^c)P(E_2|H^c)}$$

The data are

$$P(H)/P(H^c) = 2, P(E_1|H) = 0.95, P(E_1|H^c) = 0.1, P(E_2|H) = 0.92, P(E_2|H^c) = 0.05$$

Substituting values, we get

$$\frac{P(H|E_1E_2)}{P(H^c|E_1E_2)} = 2 \cdot \frac{0.95 \cdot 0.92}{0.10 \cdot 0.05} = \frac{1748}{5} \text{ so that } P(H|E_1E_2) = \frac{1748}{1753} = 1 - \frac{5}{1753} = 1 - 0.0029$$

Evidence for a symptom

Sometimes the evidence dealt with is not evidence for the hypothesized condition, but for some condition which is stochastically related. For purposes of exposition, we refer to this intermediary condition as a *symptom*. Consider again the examples above.

Table 5.4.

HYPOTHESIS	SYMPTOM	EVIDENCE
Job success	Personal traits	Diagnostic test results
Presence of oil	Geological structures	Geophysical survey results
Operation of a device	Physical condition	Monitoring report
Market condition	Test market condition	Market survey result
Presence of a disease	Physical symptom	Test for symptom

We let S be the event the symptom is present. The usual case is that the evidence is directly related to the symptom and not the hypothesized condition. The diagnostic test results can say something about an applicant's personal traits, but cannot deal directly with the hypothesized condition. The test results would be the same whether or not the candidate is successful in the job (he or she does not have the job yet). A geophysical survey deals with certain structural features beneath the surface. If a fault or a salt dome is present, the geophysical results are the same whether or not there is oil present. The physical monitoring report deals with certain physical characteristics. Its reading is the same whether or not the device will fail. A market survey treats only the condition in the test market. The results depend upon the test market, not the national market. A blood test may be for certain physical conditions which frequently are related (at least statistically) to the disease. But the result of the blood test for the physical condition is not directly affected by the presence or absence of the disease.

Under conditions of this type, we may assume

$$P(E|SH) = P(E|SH^c) \text{ and } P(E|S^cH) = P(E|S^cH^c)$$

These imply $\{E, H\}$ ci $|S$ and ci $|S^c$. Now

$$\begin{aligned} \frac{P(H|E)}{P(H^c|E)} &= \frac{P(HE)}{P(H^cE)} = \frac{P(HES) + P(HES^c)}{P(H^cES) + P(H^cES^c)} = \frac{P(HS)P(E|HS) + P(HS^c)P(E|HS^c)}{P(H^cS)P(E|H^cS) + P(H^cS^c)P(E|H^cS^c)} \\ &= \frac{P(HS)P(E|S)P(HS^c)P(E|S^c)}{P(H^cS)P(E|S) + P(H^cS^c)P(E|S^c)} \end{aligned}$$

It is worth noting that each term in the denominator differs from the corresponding term in the numerator by having H^c in place of H . Before completing the analysis, it is necessary to consider how H and S are related stochastically in the data. Four cases may be considered.

- Data are $P(S|H)$, $P(S|H^c)$, and $P(H)$.
- Data are $P(S|H)$, $P(S|H^c)$, and $P(S)$.
- Data are $P(H|S)$, $P(H|S^c)$, and $P(S)$.
- Data are $P(H|S)$, $P(H|S^c)$, and $P(H)$.

Case a:

$$\frac{P(H|S)}{P(H^c|S)} = \frac{P(H)P(S|H)P(E|S) + P(H^c)P(S|H^c)P(E|S)}{P(H)P(S^c|H)P(E|S^c) + P(H^c)P(S^c|H^c)P(E|S^c)}$$

Example 5.2.2 Geophysical survey

Let H be the event of a successful oil well, S be the event there is a geophysical structure favorable to the presence of oil, and E be the event the geophysical survey indicates a favorable structure. We suppose $\{H, E\}$ ci $|S$ and ci $|S^c$. Data are

$$P(H)/P(H^c) = 3, P(S|H) = 0.92, P(S|H^c) = 0.20, P(E|S) = 0.95, P(E|S^c) = 0.15$$

Then

$$\frac{P(H|E)}{P(H^c|E)} = 3 \cdot \frac{0.92 \cdot 0.95 + 0.08 \cdot 0.15}{0.20 \cdot 0.95 + 0.80 \cdot 0.15} = \frac{1329}{155} = 8.5742$$

$$\text{so that } P(H|E) = 1 - \frac{155}{1484} = 0.8956$$

The geophysical result moved the prior odds of 3/1 to posterior odds of 8.6/1, with a corresponding change of probabilities from 0.75 to 0.90.

Case b: Data are $P(S)P(S|H)$, $P(S|H^c)$, $P(E|S)$, and $P(E|S^c)$. If we can determine $P(H)$, we can proceed as in case a. Now by the law of total probability

$$P(S) = P(S|H)P(H) + P(S|H^c)[1 - P(H)]$$

which may be solved algebraically to give

$$P(H) = \frac{P(S) - P(S|H^c)}{P(S|H) - P(S|H^c)}$$

Example 5.2.3 Geophysical survey revisited

In many cases a better estimate of $P(S)$ or the odds $P(S)/P(S^c)$ can be made on the basis of previous geophysical data. Suppose the prior odds for S are 3/1, so that $P(S) = 0.75$. Using the other data in Example, we have

$$P(H) = \frac{P(S) - P(S|H^c)}{P(S|H) - P(S|H^c)} = \frac{0.75 - 0.20}{0.92 - 0.20} = 55/72, \text{ so that } \frac{P(H)}{P(H^c)} = 55/17$$

Using the pattern of case a, we have

$$\frac{P(H|E)}{P(H^c|E)} = \frac{55}{17} \cdot \frac{0.92 \cdot 0.95 + 0.08 \cdot 0.15}{0.20 \cdot 0.95 + 0.80 \cdot 0.15} = \frac{4873}{527} = 9.2467$$

$$\text{so that } P(H|E) = 1 - \frac{527}{5400} = 0.9024$$

Usually data relating test results to symptom are of the form $P(E|S)$ and $P(E|S^c)$, or equivalent. Data relating the symptom and the hypothesized condition may go either way. In cases a and b, the data are in the form $P(S|H)$ and $P(S|H^c)$, or equivalent, derived from data showing the fraction of times the symptom is noted when the hypothesized condition is identified. But these data may go in the opposite direction, yielding $P(H|S)$ and $P(H|S^c)$, or equivalent. This is the situation in cases c and d.

Data c: Data are $P(E|S)$, $P(E|S^c)$, $P(H|S)$, $P(H|S^c)$ and $P(S)$.

Example 5.2.4 Evidence for a disease symptom with prior P(S)

When a certain blood syndrome is observed, a given disease is indicated 93 percent of the time. The disease is found without this syndrome only three percent of the time. A test for the syndrome has probability 0.03 of a false positive and 0.05 of a false negative. A preliminary examination indicates a probability 0.30 that a patient has the syndrome. A test is performed; the result is negative. What is the probability the patient has the disease?

Solution

In terms of the notation above, the data are

$$P(S) = 0.30, P(E|S^c) = 0.03, P(E^c|S) = 0.05$$

$$P(H|S) = 0.93, \text{ and } P(H|S^c) = 0.03$$

We suppose $\{H, E\}$ ci $|S$ and ci $|S^c$.

$$\frac{P(H|E^c)}{P(H^c|E^c)} = \frac{P(S)P(H|S)P(E^c|S) + P(S^c)P(H|S^c)P(E^c|S^c)}{P(S)P(H^c|S)P(E^c|S) + P(S^c)P(H^c|S^c)P(E^c|S^c)}$$

$$= \frac{0.30 \cdot 0.93 \cdot 0.05 + 0.07 \cdot 0.03 \cdot 0.97}{0.30 \cdot 0.07 \cdot 0.05 + 0.70 \cdot 0.97 \cdot 0.97} = \frac{429}{8246}$$

which implies $P(H|E^c) = 429/8675 \approx 0.05$

Case d: This differs from case c only in the fact that a prior probability for H is assumed. In this case, we determine the corresponding probability for S by

$$P(S) = \frac{P(H) - P(H|S^c)}{P(H|S) - P(H|S^c)}$$

and use the pattern of case c.

Example 5.2.5 Evidence for a disease symptom with prior $P(h)$

Suppose for the patient in Example the physician estimates the odds favoring the presence of the disease are $1/3$, so that $P(H) = 0.25$. Again, the test result is negative. Determine the posterior odds, given E^c .

Solution

First we determine

$$P(S) = \frac{P(H) - P(H|S^c)}{P(H|S) - P(H|S^c)} = \frac{0.25 - 0.03}{0.93 - 0.03} = 11/45$$

Then

$$\frac{P(H|E^c)}{P(H^c|E^c)} = \frac{(11/45) \cdot 0.93 \cdot 0.05 + (34/45) \cdot 0.03 \cdot 0.97}{(11/45) \cdot 0.07 \cdot 0.05 + (34/45) \cdot 0.97 \cdot 0.97} = \frac{15009}{320291} = 0.047$$

The result of the test drops the prior odds of $1/3$ to approximately $1/21$.

Independent evidence for a symptom

In the previous cases, we consider only a single item of evidence for a symptom. But it may be desirable to have a "second opinion." We suppose the tests are for the symptom and are not directly related to the hypothetical condition. If the tests are operationally independent, we could reasonably assume

$$\begin{aligned} P(E_1|SE_2) &= P(E_1|SE_2^c) \{E_1, E_2\} \text{ ci } | S \\ P(E_1|SH) &= P(E_1|SH^c) \{E_1, H\} \text{ ci } | S \\ P(E_2|SH) &= P(E_2|SH^c) \{E_2, H\} \text{ ci } | S \\ P(E_1E_2|SH) &= P(E_1E_2|SH^c) \{E_1, E_2, H\} \text{ ci } | S \end{aligned}$$

This implies $\{E_1, E_2, H\} \text{ ci } | S$. A similar condition holds for S^c . As for a single test, there are four cases, depending on the tie between S and H . We consider a "case a" example.

Example 5.2.6 A market survey problem

A food company is planning to market nationally a new breakfast cereal. Its executives feel confident that the odds are at least 3 to 1 the product would be successful. Before launching the new product, the company decides to investigate a test market. Previous experience indicates that the reliability of the test market is such that if the national market is favorable, there is probability 0.9 that the test market is also. On the other hand, if the national market is unfavorable, there is a probability of only 0.2 that the test market will be favorable. These facts lead to the following analysis. Let

H be the event the national market is favorable (hypothesis)

S be the event the test market is favorable (symptom)

The *initial data* are the following probabilities, based on past experience:

- (a) Prior odds: $P(H)/P(H^c) = 3$
- (b) Reliability of the test market: $P(S|H) = 0.9$ $P(S|H^c) = 0.2$

If it were known that the test market is favorable, we should have

$$\frac{P(H|S)}{P(H^c|S)} = \frac{P(S|H)P(H)}{P(S|H^c)P(H^c)} = \frac{0.9}{0.2} \cdot 3 = 13.5$$

Unfortunately, it is not feasible to know with certainty the state of the test market. The company decision makers engage two market survey companies to make independent surveys of the test market. The reliability of the companies may be expressed as

follows. Let

E_1 be the event the first company reports a favorable test market.

E_2 be the event the second company reports a favorable test market.

On the basis of previous experience, the reliability of the evidence about the test market (the symptom) is expressed in the following conditional probabilities.

$$P(E_1|S) = 0.9 \quad P(E_1|S^c) = 0.3 \quad P(E_2|S) = 0.8 \quad P(E_2|S^c) = 0.2$$

Both survey companies report that the test market is favorable. What is the probability the national market is favorable, given this result?

Solution

The two survey firms work in an “operationally independent” manner. The report of either company is unaffected by the work of the other. Also, each report is affected only by the condition of the test market— regardless of what the national market may be. According to the discussion above, we should be able to assume

$$\{E_1, E_2, H\} \text{ ci } |S \text{ and } \{E_1, E_2, H\} \text{ ci } S^c$$

We may use a pattern similar to that in Example 2, as follows:

$$\begin{aligned} \frac{P(H|E_1E_2)P(H^c|E_1E_2)}{P(H^c)} &= \frac{P(H)}{P(H^c)} \cdot \frac{P(S|H)P(E_1|S)P(E_2|S) + P(S^c|H)P(E_1|S^c)P(E_2|S^c)}{P(S|H^c)P(E_1|S)P(E_2|S) + P(S^c|H^c)P(E_1|S^c)P(E_2|S^c)} \\ &= 3 \cdot \frac{0.9 \cdot 0.9 \cdot 0.8 + 0.1 \cdot 0.3 \cdot 0.2}{0.2 \cdot 0.9 \cdot 0.8 + 0.8 \cdot 0.3 \cdot 0.2} = \frac{327}{32} \approx 10.22 \end{aligned}$$

in terms of the posterior probability, we have

$$P(H|E_1E_2) = \frac{327/32}{1 + 327/32} = \frac{327}{359} = 1 - \frac{32}{359} \approx 0.91$$

We note that the odds favoring H , given positive indications from both survey companies, is 10.2 as compared with the odds favoring H , given a favorable test market, of 13.5. The difference reflects the residual uncertainty about the test market after the market surveys. Nevertheless, the results of the market surveys increase the odds favoring a satisfactory market from the prior 3 to 1 to a posterior 10.2 to 1. In terms of probabilities, the market surveys increase the likelihood of a favorable market from the original $P(H) = 0.75$ to the posterior $P(H|E_1E_2)$. The conditional independence of the results of the survey makes possible direct use of the data.

A classification problem

A population consists of members of two subgroups. It is desired to formulate a battery of questions to aid in identifying the subclass membership of randomly selected individuals in the population. The questions are designed so that for each individual the answers are independent, in the sense that the answers to any subset of these questions are not affected by and do not affect the answers to any other subset of the questions. The answers are, however, affected by the subgroup membership. Thus, our treatment of conditional independence suggests that it is reasonable to suppose the answers are conditionally independent, given the subgroup membership. Consider the following numerical example.

Example 5.2.7 A classification problem

A sample of 125 subjects is taken from a population which has two subgroups. The subgroup membership of each subject in the sample is known. Each individual is asked a battery of ten questions designed to be independent, in the sense that the answer to any one is not affected by the answer to any other. The subjects answer independently. Data on the results are summarized in the following table:

Table 5.5.

GROUP 1 (69 members)				GROUP 2 (56 members)		
Q	Yes	No	Unc.	Yes	No	Unc.
1	42	22	5	20	31	5

GROUP 1 (69 members)				GROUP 2 (56 members)		
2	34	27	8	16	37	3
3	15	45	9	33	19	4
4	19	44	6	31	18	7
5	22	43	4	23	28	5
6	41	13	15	14	37	5
7	9	52	8	31	17	8
8	40	26	3	13	38	5
9	48	12	9	27	24	5
10	20	37	12	35	16	5

Assume the data represent the general population consisting of these two groups, so that the data may be used to calculate probabilities and conditional probabilities.

Several persons are interviewed. The result of each interview is a “profile” of answers to the questions. The goal is to classify the person in one of the two subgroups on the basis of the profile of answers.

The following profiles were taken.

- Y, N, Y, N, Y, U, N, U, Y, U
- N, N, U, N, Y, Y, U, N, N, Y
- Y, Y, N, Y, U, U, N, N, Y, Y

Classify each individual in one of the subgroups.

Solution

Let G_1 = the event the person selected is from group 1, and $G_2 = G_1^c$ = the event the person selected is from group 2. Let

A_i = the event the answer to the i th question is “Yes”

B_i = the event the answer to the i th question is “No”

C_i = the event the answer to the i th question is “Uncertain”

The data are taken to mean $P(A_1|G_1) = 42/69$, $P(B_3|G_2) = 19/56$, etc. The profile

Y, N, Y, N, Y, U, N, U, Y, U corresponds to the event $E = A_1B_2A_3B_4A_5C_6B_7C_8A_9C_{10}$

We utilize the ratio form of Bayes' rule to calculate the posterior odds

$$\frac{P(G_1|E)}{P(G_2|E)} = \frac{P(E|G_1)}{P(E|G_2)} \cdot \frac{P(G_1)}{P(G_2)}$$

If the ratio is greater than one, classify in group 1; otherwise classify in group 2 (we assume that a ratio exactly one is so unlikely that we can neglect it). Because of conditional independence, we are able to determine the conditional probabilities

$$P(E|G_1) = \frac{42 \cdot 27 \cdot 15 \cdot 44 \cdot 22 \cdot 15 \cdot 52 \cdot 3 \cdot 48 \cdot 12}{69^{10}} \text{ and}$$

$$P(E|G_2) = \frac{29 \cdot 37 \cdot 33 \cdot 18 \cdot 23 \cdot 5 \cdot 17 \cdot 5 \cdot 24 \cdot 5}{56^{10}}$$

The odds $P(G_2)/P(G_1) = 69/56$. We find the posterior odds to be

$$\frac{P(G_1|E)}{P(G_2|E)} = \frac{42 \cdot 27 \cdot 15 \cdot 44 \cdot 22 \cdot 15 \cdot 52 \cdot 3 \cdot 48 \cdot 12}{29 \cdot 37 \cdot 33 \cdot 18 \cdot 23 \cdot 5 \cdot 17 \cdot 5 \cdot 24 \cdot 5} \cdot \frac{56^9}{69^9} = 5.85$$

The factor $56^9/69^9$ comes from multiplying $56^{10}/69^{10}$ by the odds $P(G_1)/P(G_2) = 69/56$. Since the resulting posterior odds favoring Group 1 is greater than one, we classify the respondent in group 1.

While the calculations are simple and straightforward, they are tedious and error prone. To make possible rapid and easy solution, say in a situation where successive interviews are underway, we have several m-procedures for performing the calculations. Answers to the questions would normally be designated by some such designation as Y for yes, N for no, and U for uncertain. In order for the m-procedure to work, these answers must be represented by numbers indicating the appropriate columns in matrices *A* and *B*. Thus, in the example under consideration, each Y must be translated into a 1, each N into a 2, and each U into a 3. The task is not particularly difficult, but it is much easier to have MATLAB make the translation as well as do the calculations. The following two-stage approach for solving the problem works well.

The first m-procedure *oddsdf* sets up the frequency information. The next m-procedure *odds* calculates the odds for a given profile. The advantage of splitting into two m-procedures is that we can set up the data once, then call repeatedly for the calculations for different profiles. As always, it is necessary to have the data in an appropriate form. The following is an example in which the data are entered in terms of actual frequencies of response.

```
% file oddsf4.m
% Frequency data for classification
A = [42 22 5; 34 27 8; 15 45 9; 19 44 6; 22 43 4;
     41 13 15; 9 52 8; 40 26 3; 48 12 9; 20 37 12];
B = [20 31 5; 16 37 3; 33 19 4; 31 18 7; 23 28 5;
     14 37 5; 31 17 8; 13 38 5; 27 24 5; 35 16 5];
disp('Call for oddsdf')
```

Example 5.2.8 Classification using frequency data

```
oddsf4           % Call for data in file oddsf4.m
Call for oddsdf  % Prompt built into data file
oddsdf          % Call for m-procedure oddsdf
Enter matrix A of frequencies for calibration group 1  A
Enter matrix B of frequencies for calibration group 2  B
Number of questions = 10
Answers per question = 3
  Enter code for answers and call for procedure "odds"
y = 1;           % Use of lower case for easier writing
n = 2;
u = 3;
odds             % Call for calculating procedure
Enter profile matrix E [y n y n y u n u y u] % First profile
Odds favoring Group 1:  5.845
Classify in Group 1
odds             % Second call for calculating procedure
Enter profile matrix E [n n u n y y u n n y] % Second profile
Odds favoring Group 1:  0.2383
Classify in Group 2
odds             % Third call for calculating procedure
Enter profile matrix E [y y n y u u n n y y] % Third profile
Odds favoring Group 1:  5.05
Classify in Group 1
```

The principal feature of the m-procedure odds is the scheme for selecting the numbers from the A and B matrices. If $E = [yynyuunnnny]$, then the coding translates this into the actual numerical matrix

[1 1 2 1 3 3 2 2 1 1] used internally. Then $A(:, E)$ is a matrix with columns corresponding to elements of E . Thus

```
e = A(:, E)
e = 42  42  22  42  5  5  22  22  42  42
    34  34  27  34  8  8  27  27  34  34
    15  15  45  15  9  9  45  45  15  15
    19  19  44  19  6  6  44  44  19  19
    22  22  43  22  4  4  43  43  22  22
    41  41  13  41  15 15  13  13  41  41
     9   9  52   9   8  8  52  52   9   9
    40  40  26  40   3  3  26  26  40  40
    48  48  12  48   9  9  12  12  48  48
    20  20  37  20  12 12  37  37  20  20
```

The i th entry on the i th column is the count corresponding to the answer to the i th question. For example, the answer to the third question is N (no), and the corresponding count is the third entry in the N (second) column of A . The element on the diagonal in the third column of $A(:, E)$ is the third element in that column, and hence the desired third entry of the N column. By picking out the elements on the diagonal by the command `diag(A(:,E))`, we have the desired set of counts corresponding to the profile. The same is true for `diag(B(:,E))`.

Sometimes the data are given in terms of conditional probabilities and probabilities. A slight modification of the procedure handles this case. For purposes of comparison, we convert the problem above to this form by converting the counts in matrices A and B to conditional probabilities. We do this by dividing by the total count in each group (69 and 56 in this case). Also, $P(G_1) = 69/125 = 0.552$ and $P(G_2) = 56/125 = 0.448$.

Table 5.6.

GROUP 1 $P(G_1) = 69/125$				GROUP 2 $P(G_2) = 56/125$		
Q	Yes	No	Unc.	Yes	No	Unc.
1	0.6087	0.3188	0.0725	0.3571	0.5536	0.0893
2	0.4928	0.3913	0.1159	0.2857	0.6607	0.0536
3	0.2174	0.6522	0.1304	0.5893	0.3393	0.0714
4	0.2754	0.6376	0.0870	0.5536	0.3214	0.1250
5	0.3188	0.6232	0.0580	0.4107	0.5000	0.0893
6	0.5942	0.1884	0.2174	0.2500	0.6607	0.0893
7	0.1304	0.7536	0.1160	0.5536	0.3036	0.1428
8	0.5797	0.3768	0.0435	0.2321	0.6786	0.0893
9	0.6957	0.1739	0.1304	0.4821	0.4286	0.0893
10	0.2899	0.5362	0.1739	0.6250	0.2857	0.0893

These data are in an m-file *oddsdp4.m*. The modified setup m-procedure *oddsdp* uses the conditional probabilities, then calls for the m-procedure odds.

Example 5.2.9 Calculation using conditional probability data

```
oddsp4           % Call for converted data (probabilities)
oddsdp           % Setup m-procedure for probabilities
Enter conditional probabilities for Group 1  A
Enter conditional probabilities for Group 2  B
Probability p1 individual is from Group 1  0.552
  Number of questions = 10
  Answers per question = 3
  Enter code for answers and call for procedure "odds"
y = 1;
n = 2;
u = 3;
odds
Enter profile matrix E  [y n y n y u n u y u]
Odds favoring Group 1:  5.845
Classify in Group 1
```

The slight discrepancy in the odds favoring Group 1 (5.8454 compared with 5.8452) can be attributed to rounding of the conditional probabilities to four places. The presentation above rounds the results to 5.845 in each case, so the discrepancy is not apparent. This is quite acceptable, since the discrepancy has no effect on the results.

This page titled [5.2: Patterns of Probable Inference](#) 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.