Probability Rules: Probability Tables

Learning Objective: Apply probability rules in order to find the likelihood of an event.

Learning Objective: When appropriate, use tools such as Venn diagrams or probability tables as aids for finding probabilities.

In our delivery example, there are two categorical variables of interest in the background:

- On-time delivery by service A (yes/no)
- On-time delivery by service B (yes/no)

Since each of the two has two possible values (yes/no), there are four possible combinations altogether, which correspond to the four possible outcomes of using the two services.

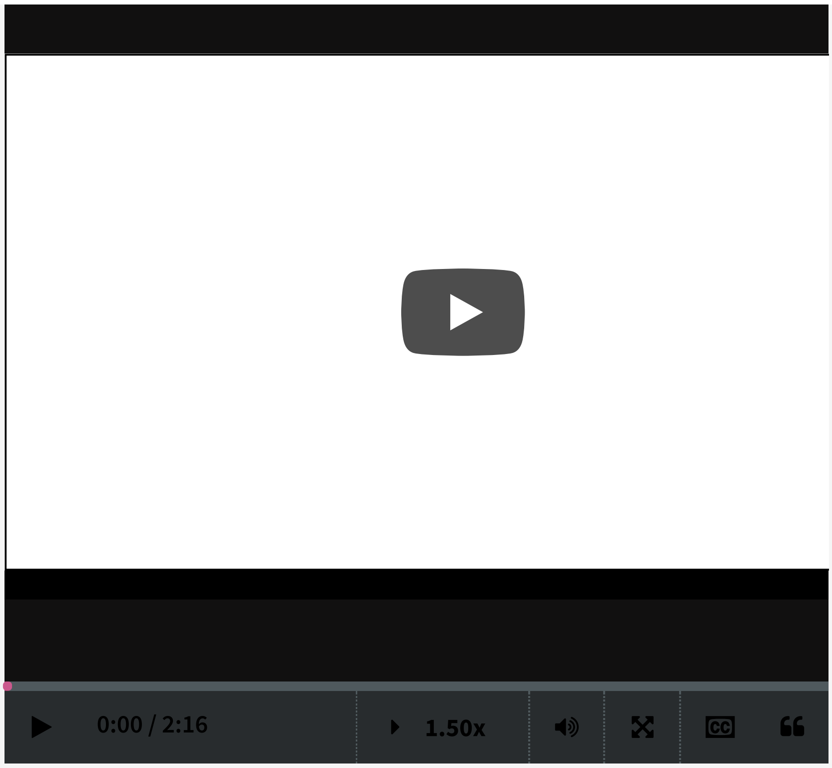
While the Venn diagrams were great to visualize the General Addition Rule, in cases like these it is much easier to display the information in and work with a two-way table of probabilities, much as we examined the relationship between two categorical variables in the Exploratory Data Analysis section.

How do we build a two-way table of probabilities?

Let's use our delivery example to illustrate this simple process:

Probability Rules

Start of transcript. Skip to the end.



Here again is the probability table. Let's fill in the table with information that

the problem provides. The probability of on-time delivery by service A is given to

be .90. The probability of on-time delivery by service B is given to be

.80, and the probability of on-time delivery by both service A

Video

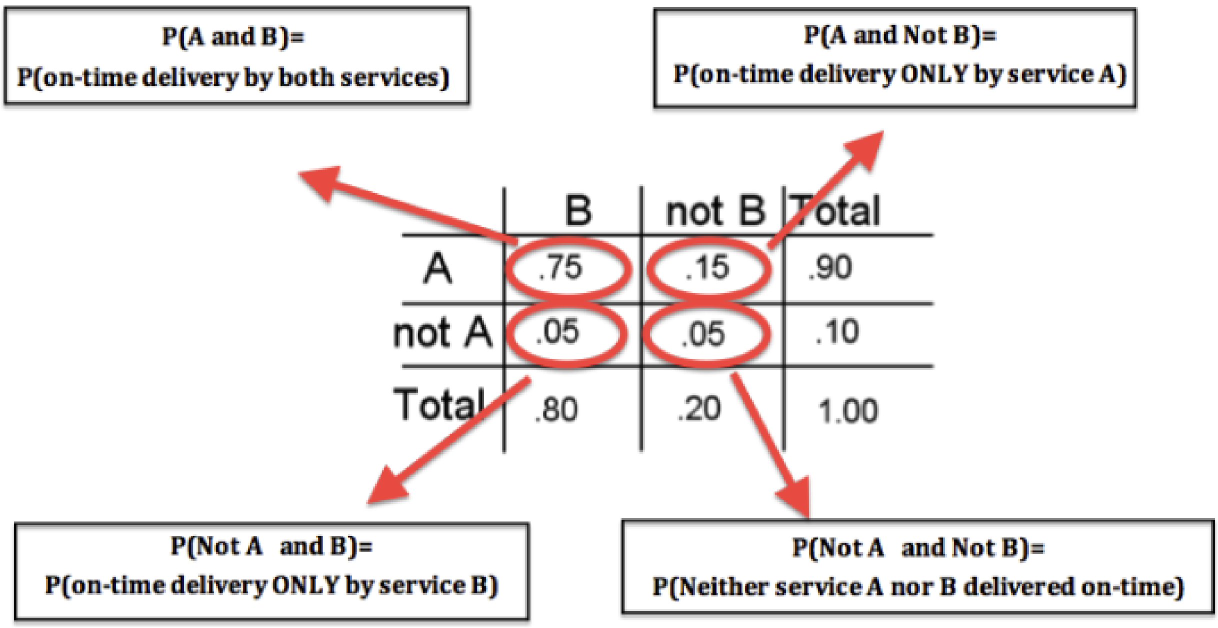
Download video file

Transcripts

Download SubRip (.srt) file

Download Text (.txt) file

Now that we’ve completed the table, it is important to understand what each of the table’s entries mean in context.



Comment

A common mistake is to confuse between: P(A)=P(event A occurs) and P(A and Not B)=P(ONLY event A occurs)[and similarly, between P(B)=P(event B occurs) and P(Not A and B)=P(only event B occurs)].

Looking at the probability table is a great way to clear-up this confusion:

	B	not B	Total
A	.75	.15	.90
not A	.05	.05	.10
Total	.80	.20	1.00

P(A) = 0.90 means that in 90% of the cases when service A is used, it delivers the document on time.

These cases of on-time delivery by service A can be decomposed into two sub-cases:

- P(A and B) =0.75 → 75% of the time the document is delivered on time also by service B (i.e., the document is delivered on time by both services)
- P(A and Not B) = 0.15 → 15% of the time the document is not delivered on time by service B (i.e., delivered on time only by service A).

Similarly,

	B	not B	Total
A	.75	.15	.90
not A	.05	.05	.10
Total	.80	.20	1.00

P(B) = 0.80 means that in 80% of the cases when service B is used, it delivers the document on time.

These cases of on-time delivery by service B can be decomposed into two sub-cases:

- P(A and B) = 0.75 → 75% of the time the document is delivered on time also by service A (i.e., the document is delivered on time by both services)
- P(Not A and B) = 0.05 → 5% of the time the document is not delivered on time by service A (i.e., delivered on time **only by service B**).

Example

Recall the smoke detector example from the last activity. Here is a quick recap:

D—the dining room alarm is set off by smoke in the kitchen

B—the bedroom alarm is set off by smoke in the kitchen

P(D) = 0.95

P(B) = 0.40

D and B are independent → P(D and B) = 0.38

Complete the table below. Start with the information that is given and go from there.

Complete the following table - what is the probability that goes in each cell?

	B	not B	Total
D			
not D			
Total			

Reset this activity

0.02 0.03 0.05 0.38 0.40 0.57 0.60 0.95 1.00

Learn By Doing

1/1 point (graded)

What is the probability that only the dining room alarm rings? Pick the correct symbolic representation.

- ○ P(D)
- ○ P(not B)

- (●) P(D and ‘not B’) ✔

Answer
Correct: If only the dining room alarm rings, then D and ‘not B’ both occur.

Submit

Learn By Doing

1/1 point (graded)
What event is described by the probability statement P(‘not D’ and ‘not B’)?

- (●) neither alarm rang ✔
- ( ) at least one alarm rang
- ( ) only one alarm rang
- ( ) both alarms rang

Answer
Correct:
‘Not D’ means the dining room alarm did not ring. ‘Not B’ means the bedroom alarm did not ring. ‘Not D’ and ‘not B’ means these events occurred together, so neither alarm rang.

Submit

Did I Get This

1/1 point (graded)
Recall the example concerning the delivery of an important document. To maximize the chances of on-time delivery, two copies of the document are sent using two services, service A and service B. It is known that the probabilities of on-time delivery are:

0.90 for service A: P(A) = 0.90

0.80 for service B: P(B) = 0.80

0.75 for both services being on time: P(A and B) = 0.75

What is the probability that only service B delivers the document on time?

- P(B) = 0.80
- P(not A) = 0.10
- P('not A' and B) = 0.05 ✔

Answer

Correct: If only service B delivers on time, then B and 'not A' both occur.

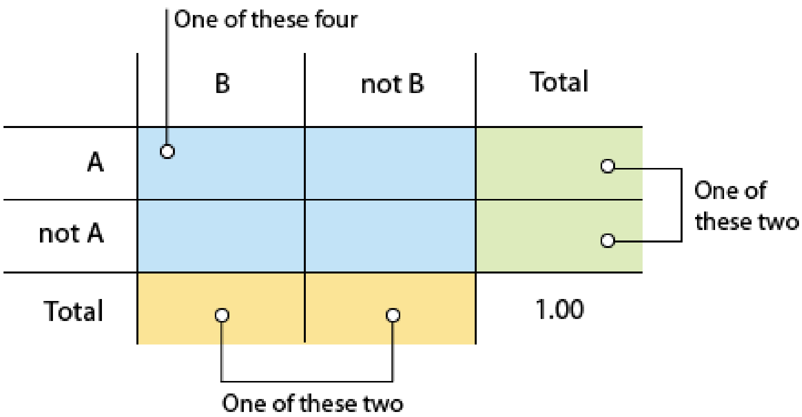
Submit

Comment

In both the delivery problem and the smoke detector problem, we knew P(A), P(B) and P(A and B). (In the smoke detector problem, we actually needed to work a bit to get P(A and B), but it wasn't too bad.) Visually, we had the probability for the three shaded cells below, which was enough information to complete the table.

	B	not B	Total
A			
not A			
Total			1.00

This, however, is not the only combination of three cells that would provide sufficient information to complete the table. Essentially, as long as we are given (or can calculate) one cell in each of the margins (the total row and column), and one of the four cells in the body of the table, we'll be able to complete the entire table. Visually, we need:



Scenario: Jury and Judge Decisions

Researchers studied thousands of court cases. For each case, they recorded the jury's decision. In addition, they asked the judge in each case how he or she would have decided the same case if there were no jury. In 67% of the cases the jury voted to convict, in 83% of the cases the judge would have convicted, and in 19% of the cases only the judge would have convicted.

Let A be the event "jury convicts".

Let B be the event "judge convicts".

Complete the following table - what is the probability that goes in each cell?

	B	not B	Total
A			
not A			
Total			

Reset this activity

0.03 0.14 0.17 0.19 0.33 0.64 0.67 0.83 1.00

Scenario: Suicide and Gender

According to www.jointogether.com, in 2000, 87% of all suicides were committed by males, 56% of all suicides were committed using a gun, and 10% of all suicides were committed by women not using a gun.

(We'll use M for suicide committed by a male, F (= not M) for suicide committed by female, and G for a suicide committed using a gun.)

Did I Get This

1/1 point (graded)

Which of the following is a correct representation of the given information in a two-way probability table?

Table A

	G	not G	Total
M	.10		.87
F			
Total	.56		

Table B

	G	not G	Total
M			.87
F		.10	
Total	.56		

Table C

	G	not G	Total
M			.87
F			.10
Total	.56		

Table D

	G	not G	Total
M	.56		.87
F		.10	
Total			

- Table A
- Table B
- Table C
- Table D

Answer

Correct: The given information was P(M) = 0.87, P(G) = 0.56, and P(F and not G) = 0.10.

Submit

Complete the following table - what is the probability that goes in each cell?

	G	not G	Total
M			
F			
Total			

Reset this activity

0.03 0.10 0.13 0.34 0.44 0.53 0.56 0.87 1.00

Comment

When we used two-way tables in the Exploratory Data Analysis (EDA) section, it was to record values of two categorical variables for a concrete **sample** of individuals. In contrast, the information in a probability two-way table is for an entire **population**, and the values are rather abstract. If we had treated something like the delivery example in the EDA section, we would have recorded the actual numbers of on-time (and not-on-time) deliveries for samples of documents mailed with service A or B. In this section, the long-term probabilities are presented as being known. Presumably, those probabilities were based on relative frequencies recorded over many repetitions.