Conditional Probability

Colby Community College

Example 1

The `photo_classify` data set represents a machine learning algorithm classifying a sample of 1822 photos as either about fashion or not.

Example 1

The `photo_classify` data set represents a machine learning algorithm classifying a sample of 1822 photos as either about fashion or not.

		truth		
		fashion	*not*	Total
mach_learn	*pred_fashion*	197	22	219
	pred_not	112	1491	1603
	Total	309	1513	1822

Example 1

The `photo_classify` data set represents a machine learning algorithm classifying a sample of 1822 photos as either about fashion or not.

		truth		
		fashion	*not*	Total
mach_learn	*pred_fashion*	197	22	219
	pred_not	112	1491	1603
	Total	309	1513	1822

If a photo is actually about fashion, what is the chance the algorithm will correctly identify the photo as being about fashion?

Example 1

The photo_classify data set represents a machine learning algorithm classifying a sample of 1822 photos as either about fashion or not.

		truth		
		fashion	*not*	Total
mach_learn	*pred_fashion*	197	22	219
	pred_not	112	1491	1603
	Total	309	1513	1822

If a photo is actually about fashion, what is the chance the algorithm will correctly identify the photo as being about fashion?

Of the 309 fashion photos, the algorithm correctly classifies 197 of them.

$$P\left(\texttt{mach\_learn} \text{ is } \textit{pred\_fashion} \text{ given } \texttt{truth} \text{ is } \textit{fashion}\right) = \frac{197}{309} = 0.638$$

Example 2

Using the same data set as in Example 1.

		truth		
		fashion	*not*	Total
mach_learn	*pred_fashion*	197	22	219
	pred_not	112	1491	1603
	Total	309	1513	1822

If the algorithm predicts the photo as being about fashion, what is the probability is actually is?

Example 2

Using the same data set as in Example 1.

		truth		
		fashion	*not*	Total
mach_learn	*pred_fashion*	197	22	219
	pred_not	112	1491	1603
	Total	309	1513	1822

If the algorithm predicts the photo as being about fashion, what is the probability is actually is?

Of the 1603 photos predicted to be about fashion, 112 we actually about fashion.

$$P\left(\texttt{truth}\text{ is }\mathit{fashion}\text{ given }\texttt{mach\_learn}\text{ is }\mathit{pred\_fashion}\right)=\frac{197}{219}=0.900$$

Note

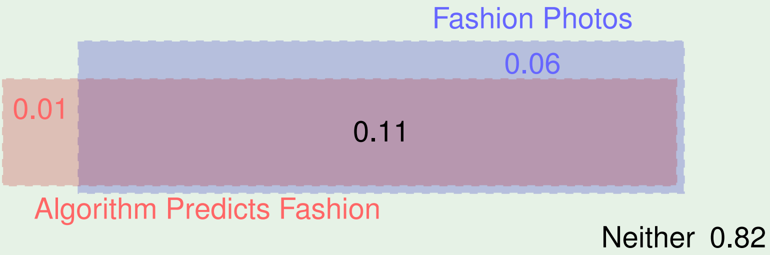
It can be helpful to draw Venn Diagrams of these contingency tables using rectangles.

Note

It can be helpful to draw Venn Diagrams of these contingency tables using rectangles.

Example 3

The Venn Diagram for Example 1 is:



Definition

A **marginal probability** is a probability based on a single variable without regard to other variables.

Definition

A **marginal probability** is a probability based on a single variable without regard to other variables.

Example 4

$$P\left(\texttt{mach\_learn}\text{ is } \textit{pred\_fashion}\right) = \frac{219}{1822} = 0.12$$

Definition

A **marginal probability** is a probability based on a single variable without regard to other variables.

Example 4

$$P\left(\texttt{mach\_learn} \text{ is } \mathit{pred\_fashion}\right) = \frac{219}{1822} = 0.12$$

Definition

A probability of outcomes for two or more variables is called a **joint probability**.

Definition

A **marginal probability** is a probability based on a single variable without regard to other variables.

Example 4

$$P\left(\texttt{mach\_learn} \text{ is } \textit{pred\_fashion}\right) = \frac{219}{1822} = 0.12$$

Definition

A probability of outcomes for two or more variables is called a **joint probability**.

Example 5

$$P\left(\texttt{mach\_learn} \text{ is } \textit{pred\_fashion} \text{ and } \texttt{truth} \text{ is } \textit{fashion}\right) = \frac{197}{1822} = 0.11$$

Note

Sometimes a comma is substituted for "and" in a joint probability.

$$P\left(\texttt{mach\_learn} \text{ is } \textit{pred\_fashion}, \texttt{truth} \text{ is } \textit{fashion}\right)$$

means the same thing as

$$P\left(\texttt{mach\_learn} \text{ is } \textit{pred\_fashion} \text{ and } \texttt{truth} \text{ is } \textit{fashion}\right)$$

Definition

A **table proportions** is a table that summarizes joint probabilities. The proportions are computed by dividing each count by table's total.

Definition

A **table proportions** is a table that summarizes joint probabilities. The proportions are computed by dividing each count by table's total.

Example 6

The table proportions for `photo_classify` are:

	truth: *fashion*	truth: *not*	Total
mach_learn: *pred_fashion*	$\frac{197}{1822}$		
mach_learn: *pred_not*			
Total			

$\downarrow\downarrow\downarrow$

	truth: *fashion*	truth: *not*	Total
mach_learn: *pred_fashion*	0.1081		
mach_learn: *pred_not*			
Total			

Definition

A **table proportions** is a table that summarizes joint probabilities. The proportions are computed by dividing each count by table's total.

Example 6

The table proportions for `photo_classify` are:

	`truth`: *fashion*	`truth`: *not*	Total
`mach_learn`: *pred_fashion*	$\frac{197}{1822}$	$\frac{22}{1822}$	
`mach_learn`: *pred_not*			
Total			

$\downarrow\downarrow\downarrow$

	`truth`: *fashion*	`truth`: *not*	Total
`mach_learn`: *pred_fashion*	0.1081	0.0121	
`mach_learn`: *pred_not*			
Total			

Definition

A **table proportions** is a table that summarizes joint probabilities. The proportions are computed by dividing each count by table's total.

Example 6

The table proportions for `photo_classify` are:

	`truth`: *`fashion`*	`truth`: *`not`*	Total
`mach_learn`: *`pred_fashion`*	$\frac{197}{1822}$	$\frac{22}{1822}$	$\frac{219}{1822}$
`mach_learn`: *`pred_not`*			
Total			

$$\downarrow\downarrow\downarrow$$

	`truth`: *`fashion`*	`truth`: *`not`*	Total
`mach_learn`: *`pred_fashion`*	0.1081	0.0121	0.1202
`mach_learn`: *`pred_not`*			
Total			

Definition

A **table proportions** is a table that summarizes joint probabilities. The proportions are computed by dividing each count by table's total.

Example 6

The table proportions for `photo_classify` are:

	`truth`: *`fashion`*	`truth`: *`not`*	Total
`mach_learn`: *`pred_fashion`*	$\frac{197}{1822}$	$\frac{22}{1822}$	$\frac{219}{1822}$
`mach_learn`: *`pred_not`*	$\frac{112}{1822}$		
Total			

$\downarrow\downarrow\downarrow$

	`truth`: *`fashion`*	`truth`: *`not`*	Total
`mach_learn`: *`pred_fashion`*	0.1081	0.0121	0.1202
`mach_learn`: *`pred_not`*	0.0615		
Total			

Definition

A **table proportions** is a table that summarizes joint probabilities. The proportions are computed by dividing each count by table's total.

Example 6

The table proportions for `photo_classify` are:

	truth: *fashion*	truth: *not*	Total
mach_learn: *pred_fashion*	$\frac{197}{1822}$	$\frac{22}{1822}$	$\frac{219}{1822}$
mach_learn: *pred_not*	$\frac{112}{1822}$	$\frac{1491}{1822}$	
Total			

$\downarrow\downarrow\downarrow$

	truth: *fashion*	truth: *not*	Total
mach_learn: *pred_fashion*	0.1081	0.0121	0.1202
mach_learn: *pred_not*	0.0615	0.8183	
Total			

Definition

A **table proportions** is a table that summarizes joint probabilities. The proportions are computed by dividing each count by table's total.

Example 6

The table proportions for `photo_classify` are:

	truth: *fashion*	truth: *not*	Total
mach_learn: *pred_fashion*	$\frac{197}{1822}$	$\frac{22}{1822}$	$\frac{219}{1822}$
mach_learn: *pred_not*	$\frac{112}{1822}$	$\frac{1491}{1822}$	$\frac{1603}{1822}$
Total			

$\downarrow\downarrow\downarrow$

	truth: *fashion*	truth: *not*	Total
mach_learn: *pred_fashion*	0.1081	0.0121	0.1202
mach_learn: *pred_not*	0.0615	0.8183	0.8798
Total			

Definition

A **table proportions** is a table that summarizes joint probabilities. The proportions are computed by dividing each count by table's total.

Example 6

The table proportions for `photo_classify` are:

	`truth`: *`fashion`*	`truth`: *`not`*	Total
`mach_learn`: *`pred_fashion`*	$\frac{197}{1822}$	$\frac{22}{1822}$	$\frac{219}{1822}$
`mach_learn`: *`pred_not`*	$\frac{112}{1822}$	$\frac{1491}{1822}$	$\frac{1603}{1822}$
Total	$\frac{309}{1822}$		

$$\downarrow\downarrow\downarrow$$

	`truth`: *`fashion`*	`truth`: *`not`*	Total
`mach_learn`: *`pred_fashion`*	0.1081	0.0121	0.1202
`mach_learn`: *`pred_not`*	0.0615	0.8183	0.8798
Total	0.1696		

Definition

A **table proportions** is a table that summarizes joint probabilities. The proportions are computed by dividing each count by table's total.

Example 6

The table proportions for `photo_classify` are:

	`truth`: *fashion*	`truth`: *not*	Total
`mach_learn`: *pred_fashion*	$\frac{197}{1822}$	$\frac{22}{1822}$	$\frac{219}{1822}$
`mach_learn`: *pred_not*	$\frac{112}{1822}$	$\frac{1491}{1822}$	$\frac{1603}{1822}$
Total	$\frac{309}{1822}$	$\frac{1513}{1822}$	

$\downarrow\downarrow\downarrow$

	`truth`: *fashion*	`truth`: *not*	Total
`mach_learn`: *pred_fashion*	0.1081	0.0121	0.1202
`mach_learn`: *pred_not*	0.0615	0.8183	0.8798
Total	0.1696	0.8304	

Definition

A **table proportions** is a table that summarizes joint probabilities. The proportions are computed by dividing each count by table's total.

Example 6

The table proportions for `photo_classify` are:

	`truth`: *fashion*	`truth`: *not*	Total
`mach_learn`: *pred_fashion*	$\frac{197}{1822}$	$\frac{22}{1822}$	$\frac{219}{1822}$
`mach_learn`: *pred_not*	$\frac{112}{1822}$	$\frac{1491}{1822}$	$\frac{1603}{1822}$
Total	$\frac{309}{1822}$	$\frac{1513}{1822}$	$\frac{1822}{1822}$

$$\downarrow\downarrow\downarrow$$

	`truth`: *fashion*	`truth`: *not*	Total
`mach_learn`: *pred_fashion*	0.1081	0.0121	0.1202
`mach_learn`: *pred_not*	0.0615	0.8183	0.8798
Total	0.1696	0.8304	1.0

Example 7

The table proportions from Example 6 make a probability distribution.

Joint Outcome	Probability
`mach_learn` is *pred_fashion* and `truth` is *fashion*	0.1081
`mach_learn` is *pred_fashion* and `truth` is *not*	0.0121
`mach_learn` is *pred_not* and `truth` is *fashion*	0.0615
`mach_learn` is *pred_not* and `truth` is *not*	0.8182

Example 7

The table proportions from Example 6 make a probability distribution.

Joint Outcome	Probability
`mach_learn` is *pred_fashion* and `truth` is *fashion*	0.1081
`mach_learn` is *pred_fashion* and `truth` is *not*	0.0121
`mach_learn` is *pred_not* and `truth` is *fashion*	0.0615
`mach_learn` is *pred_not* and `truth` is *not*	0.8182

Note

Joint probabilities can be used to calculate marginal probabilities in simple cases.

Example 7

The table proportions from Example 6 make a probability distribution.

Joint Outcome	Probability
`mach_learn` is *pred_fashion* and `truth` is *fashion*	0.1081
`mach_learn` is *pred_fashion* and `truth` is *not*	0.0121
`mach_learn` is *pred_not* and `truth` is *fashion*	0.0615
`mach_learn` is *pred_not* and `truth` is *not*	0.8182

Note

Joint probabilities can be used to calculate marginal probabilities in simple cases.

Example 8

$$\begin{aligned} P\left(\texttt{truth}\text{ is }\textit{fashion}\right) &= P\left(\texttt{mac\_learn}\text{ is }\textit{pred\_fashion}\text{ and }\texttt{truth}\text{ is }\textit{fashion}\right) \\ &\quad + P\left(\texttt{mac\_learn}\text{ is }\textit{pred\_not}\text{ and }\texttt{truth}\text{ is }\textit{fashion}\right) \end{aligned}$$

Example 7

The table proportions from Example 6 make a probability distribution.

Joint Outcome	Probability
`mach_learn` is *pred_fashion* and `truth` is *fashion*	0.1081
`mach_learn` is *pred_fashion* and `truth` is *not*	0.0121
`mach_learn` is *pred_not* and `truth` is *fashion*	0.0615
`mach_learn` is *pred_not* and `truth` is *not*	0.8182

Note

Joint probabilities can be used to calculate marginal probabilities in simple cases.

Example 8

$$\begin{aligned} P\left(\texttt{truth} \text{ is } \textit{fashion}\right) &= P\left(\texttt{mac\_learn} \text{ is } \textit{pred\_fashion} \text{ and } \texttt{truth} \text{ is } \textit{fashion}\right) \\ &\quad + P\left(\texttt{mac\_learn} \text{ is } \textit{pred\_not} \text{ and } \texttt{truth} \text{ is } \textit{fashion}\right) \\ &= 0.1081 + 0.0615 \end{aligned}$$

Example 7

The table proportions from Example 6 make a probability distribution.

Joint Outcome	Probability
`mach_learn` is *pred_fashion* and `truth` is *fashion*	0.1081
`mach_learn` is *pred_fashion* and `truth` is *not*	0.0121
`mach_learn` is *pred_not* and `truth` is *fashion*	0.0615
`mach_learn` is *pred_not* and `truth` is *not*	0.8182

Note

Joint probabilities can be used to calculate marginal probabilities in simple cases.

Example 8

$$\begin{aligned} P(\texttt{truth}\text{ is }\textit{fashion}) &= P(\texttt{mac\_learn}\text{ is }\textit{pred\_fashion}\text{ and }\texttt{truth}\text{ is }\textit{fashion}) \\ &\quad + P(\texttt{mac\_learn}\text{ is }\textit{pred\_not}\text{ and }\texttt{truth}\text{ is }\textit{fashion}) \\ &= 0.1081 + 0.0615 = 0.1696 \end{aligned}$$

Definition

A **conditional probability** is a probability computed under a condition.

Definition

A **conditional probability** is a probability computed under a condition.

Example 9

$$P\left(\texttt{truth}\text{ is }\textit{fashion}\text{ given }\texttt{mach\_learn}\text{ is }\textit{pred\_fashion}\right) = \frac{197}{219} = 0.900$$

Definition

A **conditional probability** is a probability computed under a condition.

Example 9

$$P\left(\texttt{truth}\text{ is } \mathit{fashion} \text{ given } \texttt{mach\_learn} \text{ is } \mathit{pred\_fashion}\right) = \frac{197}{219} = 0.900$$

Definition

There are two parts to a conditional probability, the **outcome of interest** and the **condition**.

$$P\left(\text{outcome of interest given condition}\right)$$

is the same as

$$P\left(\text{outcome of interest} \mid \text{condition}\right)$$

Definition

A **conditional probability** is a probability computed under a condition.

Example 9

$$P\left(\texttt{truth}\text{ is }\mathtt{\textit{fashion}}\text{ given }\texttt{mach\_learn}\text{ is }\mathtt{\textit{pred\_fashion}}\right) = \frac{197}{219} = 0.900$$

Definition

There are two parts to a conditional probability, the **outcome of interest** and the **condition.**

$$P\left(\text{outcome of interest given condition}\right)$$

is the same as

$$P\left(\text{outcome of interest} \mid \text{condition}\right)$$

Example 10

$$P\left(\texttt{truth}\text{ is }\mathtt{\textit{fashion}} \mid \texttt{mach\_learn}\text{ is }\mathtt{\textit{pred\_fashion}}\right) = \frac{197}{219} = 0.900$$

Note

Conditional probabilities are computed as:

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

Note

Conditional probabilities are computed as:

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

Example 11

$$\begin{aligned}
&P(\texttt{truth}\text{ is }\mathit{fashion} \mid \texttt{mach_learn}\text{ is }\mathit{pred_fashion}) \\
&\quad = \frac{P(\texttt{truth}\text{ is }\mathit{fashion}\text{ and }\texttt{mach_learn}\text{ is }\mathit{pred_fashion})}{P(\texttt{mach_learn}\text{ is }\mathit{pred_fashion})} \\
&\quad = \frac{0.0615}{0.1696} \\
&\quad = 0.3626
\end{aligned}$$

Example 12

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated			inoculated		
		yes	*no*	Total	*yes*	*no*	Total
result	*lived*	238	5136	5374	0.0382	0.8252	0.8634
	died	6	844	850	0.0010	0.1356	0.1366
	Total	244	5980	6224	0.0392	0.9608	1.0000

Example 12

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated		
		yes	*no*	Total
result	*lived*	238	5136	5374
	died	6	844	850
	Total	244	5980	6224

inoculated		
yes	*no*	Total
0.0382	0.8252	0.8634
0.0010	0.1356	0.1366
0.0392	0.9608	1.0000

What is the probability that a randomly selected person who was not inoculated died from smallpox?

Example 12

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated: *yes*	inoculated: *no*	Total
result	*lived*	238	5136	5374
	died	6	844	850
	Total	244	5980	6224

		inoculated: *yes*	inoculated: *no*	Total
result	*lived*	0.0382	0.8252	0.8634
	died	0.0010	0.1356	0.1366
	Total	0.0392	0.9608	1.0000

What is the probability that a randomly selected person who was not inoculated died from smallpox?

$$P\left(\texttt{result}\text{ is } \textit{died} \,\middle|\, \texttt{inoculated}\text{ is } \textit{no}\right)$$

Example 12

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated		
		yes	*no*	Total
result	*lived*	238	5136	5374
	died	6	844	850
	Total	244	5980	6224

inoculated		
yes	*no*	Total
0.0382	0.8252	0.8634
0.0010	0.1356	0.1366
0.0392	0.9608	1.0000

What is the probability that a randomly selected person who was not inoculated died from smallpox?

$$P\left(\texttt{result} \text{ is } \textit{died} \,\middle|\, \texttt{inoculated} \text{ is } \textit{no}\right)$$

$$= \frac{P\left(\texttt{result} \text{ is } \textit{died} \text{ and } \texttt{inoculated} \text{ is } \textit{no}\right)}{P\left(\texttt{inoculated} \text{ is } \textit{no}\right)}$$

Example 12

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated		
		yes	*no*	Total
result	*lived*	238	5136	5374
	died	6	844	850
	Total	244	5980	6224

inoculated		
yes	*no*	Total
0.0382	0.8252	0.8634
0.0010	0.1356	0.1366
0.0392	0.9608	1.0000

What is the probability that a randomly selected person who was not inoculated died from smallpox?

$$P\left(\texttt{result}\text{ is }\textit{died} \;\middle|\; \texttt{inoculated}\text{ is }\textit{no}\right)$$

$$= \frac{P\left(\texttt{result}\text{ is }\textit{died}\text{ and }\texttt{inoculated}\text{ is }\textit{no}\right)}{P\left(\texttt{inoculated}\text{ is }\textit{no}\right)}$$

$$= \frac{0.1356}{0.9608}$$

Example 12

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated			inoculated		
		yes	*no*	Total	*yes*	*no*	Total
result	*lived*	238	5136	5374	0.0382	0.8252	0.8634
	died	6	844	850	0.0010	0.1356	0.1366
	Total	244	5980	6224	0.0392	0.9608	1.0000

What is the probability that a randomly selected person who was not inoculated died from smallpox?

$$P\left(\texttt{result}\text{ is }\textit{died} \mid \texttt{inoculated}\text{ is }\textit{no}\right)$$

$$= \frac{P\left(\texttt{result}\text{ is }\textit{died}\text{ and }\texttt{inoculated}\text{ is }\textit{no}\right)}{P\left(\texttt{inoculated}\text{ is }\textit{no}\right)}$$

$$= \frac{0.1356}{0.9608}$$

$$= 0.1411 \approx 14.11\%$$

Example 13

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated			inoculated		
		yes	*no*	Total	*yes*	*no*	Total
result	*lived*	238	5136	5374	0.0382	0.8252	0.8634
	died	6	844	850	0.0010	0.1356	0.1366
	Total	244	5980	6224	0.0392	0.9608	1.0000

Example 13

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated		
		yes	*no*	Total
result	*lived*	238	5136	5374
	died	6	844	850
	Total	244	5980	6224

inoculated		
yes	*no*	Total
0.0382	0.8252	0.8634
0.0010	0.1356	0.1366
0.0392	0.9608	1.0000

What is the probability that a randomly selected inoculated person died from smallpox?

Example 13

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated			inoculated		
		yes	*no*	Total	*yes*	*no*	Total
result	*lived*	238	5136	5374	0.0382	0.8252	0.8634
	died	6	844	850	0.0010	0.1356	0.1366
	Total	244	5980	6224	0.0392	0.9608	1.0000

What is the probability that a randomly selected inoculated person died from smallpox?

$$P\left(\texttt{result}\text{ is }\textit{died} \,\middle|\, \texttt{inoculated}\text{ is }\textit{yes}\right)$$

Example 13

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated			inoculated		
		yes	*no*	Total	*yes*	*no*	Total
result	*lived*	238	5136	5374	0.0382	0.8252	0.8634
	died	6	844	850	0.0010	0.1356	0.1366
	Total	244	5980	6224	0.0392	0.9608	1.0000

What is the probability that a randomly selected inoculated person died from smallpox?

$$P\left(\texttt{result}\text{ is }\textit{died} \,\middle|\, \texttt{inoculated}\text{ is }\textit{yes}\right)$$

$$= \frac{P\left(\texttt{result}\text{ is }\textit{died}\text{ and }\texttt{inoculated}\text{ is }\textit{yes}\right)}{P\left(\texttt{inoculated}\text{ is }\textit{yes}\right)}$$

Example 13

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated		
		yes	*no*	Total
result	*lived*	238	5136	5374
	died	6	844	850
	Total	244	5980	6224

inoculated		
yes	*no*	Total
0.0382	0.8252	0.8634
0.0010	0.1356	0.1366
0.0392	0.9608	1.0000

What is the probability that a randomly selected inoculated person died from smallpox?

$$P\left(\texttt{result}\text{ is }\textit{died} \;\middle|\; \texttt{inoculated}\text{ is }\textit{yes}\right)$$

$$= \frac{P\left(\texttt{result}\text{ is }\textit{died}\text{ and }\texttt{inoculated}\text{ is }\textit{yes}\right)}{P\left(\texttt{inoculated}\text{ is }\textit{yes}\right)}$$

$$= \frac{0.0010}{0.0392}$$

Example 13

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

		inoculated		
		yes	*no*	Total
result	*lived*	238	5136	5374
	died	6	844	850
	Total	244	5980	6224

inoculated		
yes	*no*	Total
0.0382	0.8252	0.8634
0.0010	0.1356	0.1366
0.0392	0.9608	1.0000

What is the probability that a randomly selected inoculated person died from smallpox?

$$P\left(\texttt{result}\text{ is }\textit{died} \mid \texttt{inoculated}\text{ is }\textit{yes}\right)$$

$$= \frac{P\left(\texttt{result}\text{ is }\textit{died}\text{ and }\texttt{inoculated}\text{ is }\textit{yes}\right)}{P\left(\texttt{inoculated}\text{ is }\textit{yes}\right)}$$

$$= \frac{0.0010}{0.0392}$$

$$= 0.0255 \approx 2.55\%$$

Example 14

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

The residents of Boston self-selected whether or not to be inoculated.

Is this study observational or experimental?

Example 14

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

The residents of Boston self-selected whether or not to be inoculated.

Is this study observational or experimental? Observational

Example 14

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

The residents of Boston self-selected whether or not to be inoculated.

Is this study observational or experimental? Observational

Can we infer any causal connection using this data?

Example 14

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

The residents of Boston self-selected whether or not to be inoculated.

Is this study observational or experimental? Observational

Can we infer any causal connection using this data?

No, the fact this is an observational study combined with the self-selecting bias means we cannot.

Example 14

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

The residents of Boston self-selected whether or not to be inoculated.

Is this study observational or experimental? Observational

Can we infer any causal connection using this data?

No, the fact this is an observational study combined with the self-selecting bias means we cannot.

What are some potential confounding variables that might influence whether someone `lived` *or* `died`*?*

Example 14

The `smallpox` data set provides a sample of 6,224 individuals from the year 1721 who were exposed to smallpox in Boston.

The residents of Boston self-selected whether or not to be inoculated.

Is this study observational or experimental? Observational

Can we infer any causal connection using this data?

No, the fact this is an observational study combined with the self-selecting bias means we cannot.

What are some potential confounding variables that might influence whether someone `lived` *or* `died`*?*

People die for many reasons, wealth determines level of medical care available, etc...

General Multiplication Rule

If A and B represent two outcomes or events, then

$$P(A \text{ and } B) = P(A \mid B) \cdot P(B)$$

General Multiplication Rule

If A and B represent two outcomes or events, then

$$P(A \text{ and } B) = P(A \mid B) \cdot P(B)$$

Example 15

Suppose we are only given two pieces of information:

- 96.08% of Boston residents were not inoculated.
- 85.88% of Boston residents who were not inoculated ended up surviving.

Let us find the probability that a resident who was no inoculated and lived.

General Multiplication Rule

If A and B represent two outcomes or events, then

$$P(A \text{ and } B) = P(A \mid B) \cdot P(B)$$

Example 15

Suppose we are only given two pieces of information:

- 96.08% of Boston residents were not inoculated.
- 85.88% of Boston residents who were not inoculated ended up surviving.

Let us find the probability that a resident who was no inoculated and lived.

$$P(\texttt{result} \text{ is } \textit{\texttt{lived}} \text{ and } \texttt{inoculated} \text{ is } \textit{\texttt{no}})$$

General Multiplication Rule

If A and B represent two outcomes or events, then

$$P(A \text{ and } B) = P(A \mid B) \cdot P(B)$$

Example 15

Suppose we are only given two pieces of information:

- 96.08% of Boston residents were not inoculated.
- 85.88% of Boston residents who were not inoculated ended up surviving.

Let us find the probability that a resident who was no inoculated and lived.

$$\begin{aligned} &P(\texttt{result} \text{ is } \textit{lived} \text{ and } \texttt{inoculated} \text{ is } \textit{no}) \\ &\quad = P(\texttt{result} \text{ is } \textit{lived} \mid \texttt{inoculated} \text{ is } \textit{no}) \cdot P(\texttt{inoculated} \text{ is } \textit{no}) \end{aligned}$$

General Multiplication Rule

If A and B represent two outcomes or events, then

$$P(A \text{ and } B) = P(A \mid B) \cdot P(B)$$

Example 15

Suppose we are only given two pieces of information:

- 96.08% of Boston residents were not inoculated.
- 85.88% of Boston residents who were not inoculated ended up surviving.

Let us find the probability that a resident who was no inoculated and lived.

$$\begin{aligned} &P(\texttt{result} \text{ is } \texttt{\textit{lived}} \text{ and } \texttt{inoculated} \text{ is } \texttt{\textit{no}}) \\ &\qquad = P(\texttt{result} \text{ is } \texttt{\textit{lived}} \mid \texttt{inoculated} \text{ is } \texttt{\textit{no}}) \cdot P(\texttt{inoculated} \text{ is } \texttt{\textit{no}}) \\ &\qquad = 0.8588 \cdot 0.9698 \end{aligned}$$

General Multiplication Rule

If A and B represent two outcomes or events, then

$$P(A \text{ and } B) = P(A \mid B) \cdot P(B)$$

Example 15

Suppose we are only given two pieces of information:

- 96.08% of Boston residents were not inoculated.
- 85.88% of Boston residents who were not inoculated ended up surviving.

Let us find the probability that a resident who was no inoculated and lived.

$$\begin{aligned} &P(\texttt{result} \text{ is } \texttt{\textit{lived}} \text{ and } \texttt{inoculated} \text{ is } \texttt{\textit{no}}) \\ &\qquad = P(\texttt{result} \text{ is } \texttt{\textit{lived}} \mid \texttt{inoculated} \text{ is } \texttt{\textit{no}}) \cdot P(\texttt{inoculated} \text{ is } \texttt{\textit{no}}) \\ &\qquad = 0.8588 \cdot 0.9698 \\ &\qquad = 0.8251 \approx 82.51\% \end{aligned}$$

Example 16

Let's use

- $P\left(\texttt{inoculated}\text{ is }\mathit{yes}\right) = 0.0392$
- $P\left(\texttt{result}\text{ is }\mathit{lived} \mid \texttt{inoculated}\text{ is }\mathit{yes}\right) = 0.9754$

to find the probability that a person was both inoculated and lived.

Example 16

Let's use

- $P(\texttt{inoculated}\text{ is }\textit{yes}) = 0.0392$
- $P(\texttt{result}\text{ is }\textit{lived} \mid \texttt{inoculated}\text{ is }\textit{yes}) = 0.9754$

to find the probability that a person was both inoculated and lived.

$$P(\texttt{result}\text{ is }\textit{lived}\text{ and }\texttt{inoculated}\text{ is }\textit{yes})$$

Example 16

Let's use

- $P\left(\texttt{inoculated}\text{ is }\textit{yes}\right) = 0.0392$
- $P\left(\texttt{result}\text{ is }\textit{lived} \mid \texttt{inoculated}\text{ is }\textit{yes}\right) = 0.9754$

to find the probability that a person was both inoculated and lived.

$$
\begin{aligned}
&P\left(\texttt{result}\text{ is }\textit{lived}\text{ and }\texttt{inoculated}\text{ is }\textit{yes}\right) \\
&\qquad = P\left(\texttt{result}\text{ is }\textit{lived} \mid \texttt{inoculated}\text{ is }\textit{yes}\right) \cdot P\left(\texttt{inoculated}\text{ is }\textit{yes}\right)
\end{aligned}
$$

Example 16

Let's use

- $P\left(\texttt{inoculated is } \textit{yes}\right) = 0.0392$
- $P\left(\texttt{result is } \textit{lived} \mid \texttt{inoculated is } \textit{yes}\right) = 0.9754$

to find the probability that a person was both inoculated and lived.

$$
\begin{aligned}
&P\left(\texttt{result is } \textit{lived} \text{ and } \texttt{inoculated is } \textit{yes}\right)\\
&\qquad = P\left(\texttt{result is } \textit{lived} \mid \texttt{inoculated is } \textit{yes}\right) \cdot P\left(\texttt{inoculated is } \textit{yes}\right)\\
&\qquad = 0.0382 \cdot 0.9754
\end{aligned}
$$

Example 16

Let's use

- $P\,(\texttt{inoculated is } \mathit{yes}) = 0.0392$
- $P\,(\texttt{result is } \mathit{lived} \mid \texttt{inoculated is } \mathit{yes}) = 0.9754$

to find the probability that a person was both inoculated and lived.

$$\begin{aligned}
&P\,(\texttt{result is } \mathit{lived} \text{ and } \texttt{inoculated is } \mathit{yes}) \\
&\qquad = P\,(\texttt{result is } \mathit{lived} \mid \texttt{inoculated is } \mathit{yes}) \cdot P\,(\texttt{inoculated is } \mathit{yes}) \\
&\qquad = 0.0382 \cdot 0.9754 \\
&\qquad = 0.0382 \approx 3.82\%
\end{aligned}$$

Sum of Conditional Probabilities

Let $A_1, A_2, \ldots, A_k$ represent all the disjoint outcomes for a variable. Then if B is an event, possibly for another variable, we have:

$$P(A_1 \mid B) + P(A_2 \mid B) + \cdots + P(A_k \mid B) = 1$$

Sum of Conditional Probabilities

Let $A_1, A_2, \ldots, A_k$ represent all the disjoint outcomes for a variable. Then if B is an event, possibly for another variable, we have:

$$P(A_1 \mid B) + P(A_2 \mid B) + \cdots + P(A_k \mid B) = 1$$

Note

If an event and it's complement are conditioned on the same information, then:

$$P(A \mid B) = 1 - P\left(A^C \mid B\right)$$

Sum of Conditional Probabilities

Let $A_1, A_2, \ldots, A_k$ represent all the disjoint outcomes for a variable. Then if B is an event, possibly for another variable, we have:

$$P(A_1 \mid B) + P(A_2 \mid B) + \cdots + P(A_k \mid B) = 1$$

Note

If an event and it's complement are conditioned on the same information, then:

$$P(A \mid B) = 1 - P\left(A^C \mid B\right)$$

Example 17

If 97.54% of the inoculated people lived, what is the proportion of people that must have died?

Sum of Conditional Probabilities

Let $A_1, A_2, \ldots, A_k$ represent all the disjoint outcomes for a variable. Then if B is an event, possibly for another variable, we have:

$$P(A_1 \mid B) + P(A_2 \mid B) + \cdots + P(A_k \mid B) = 1$$

Note

If an event and it's complement are conditioned on the same information, then:

$$P(A \mid B) = 1 - P\left(A^C \mid B\right)$$

Example 17

If 97.54% of the inoculated people lived, what is the proportion of people that must have died?

There are only two outcomes: `lived` and `died`. Which means that $100\% - 97.54\% = 2.46\%$ people who were inoculated died.

Example 18

Let X and Y represent the outcomes of rolling two dice.

Let us compute the following:

$$P\,(Y = 1 \mid X = 1)$$

Example 18

Let X and Y represent the outcomes of rolling two dice.

Let us compute the following:

$$P(Y = 1 \mid X = 1) \quad = \frac{P(Y = 1 \text{ and } X = 1)}{P(X = 1)}$$

Example 18

Let X and Y represent the outcomes of rolling two dice.

Let us compute the following:

$$\begin{aligned} P(Y=1 \mid X=1) &= \frac{P(Y=1 \text{ and } X=1)}{P(X=1)} \\ &= \frac{P(Y=1)\cdot P(X=1)}{P(X=1)} \end{aligned}$$

Example 18

Let X and Y represent the outcomes of rolling two dice.

Let us compute the following:

$$\begin{aligned} P(Y = 1 \mid X = 1) &= \frac{P(Y = 1 \text{ and } X = 1)}{P(X = 1)} \\ &= \frac{P(Y = 1) \cdot P(X = 1)}{P(X = 1)} \\ &= \frac{P(Y = 1) \cdot \cancel{P(X = 1)}}{\cancel{P(X = 1)}} \end{aligned}$$

Example 18

Let X and Y represent the outcomes of rolling two dice.

Let us compute the following:

$$
\begin{aligned}
P(Y = 1 \mid X = 1) &= \frac{P(Y = 1 \text{ and } X = 1)}{P(X = 1)} \\
&= \frac{P(Y = 1) \cdot P(X = 1)}{P(X = 1)} \\
&= \frac{P(Y = 1) \cdot \cancel{P(X = 1)}}{\cancel{P(X = 1)}} \\
&= P(Y = 1)
\end{aligned}
$$

Example 18

Let X and Y represent the outcomes of rolling two dice.

Let us compute the following:

$$\begin{aligned} P(Y = 1 \mid X = 1) &= \frac{P(Y = 1 \text{ and } X = 1)}{P(X = 1)} \\ &= \frac{P(Y = 1) \cdot P(X = 1)}{P(X = 1)} \\ &= \frac{P(Y = 1) \cdot \cancel{P(X = 1)}}{\cancel{P(X = 1)}} \\ &= P(Y = 1) \end{aligned}$$

Note

We have shown that if two events are independent, then knowing the outcome of one should provide no information about the other.

Example 19

Ron is watching a roulette table in a casino and notices that the last five outcomes were `black`. He figures that the chances of getting `black` six times in a row is very small and puts his paycheck on red.

Why is this a really bad idea?

Example 19

Ron is watching a roulette table in a casino and notices that the last five outcomes were `black`. He figures that the chances of getting `black` six times in a row is very small and puts his paycheck on red.

Why is this a really bad idea?

Each spin of a roulette wheel is independent of any of the previous spins. The next spin has the exact same probability of getting black and any other.

Example 19

Ron is watching a roulette table in a casino and notices that the last five outcomes were `black`. He figures that the chances of getting `black` six times in a row is very small and puts his paycheck on red.

Why is this a really bad idea?

Each spin of a roulette wheel is independent of any of the previous spins. The next spin has the exact same probability of getting black and any other.

Note

Posting the last several outcomes of a betting game is a real practice casinos use to trick people into believing the odds are in their favor. It's known as the **gambler's fallacy**.

Definition

A **tree diagram** is a tool to organize outcomes and probabilities around the structure of the data.

They are most useful when two or more processes occur in a sequence and each process is conditioned on its predecessor.

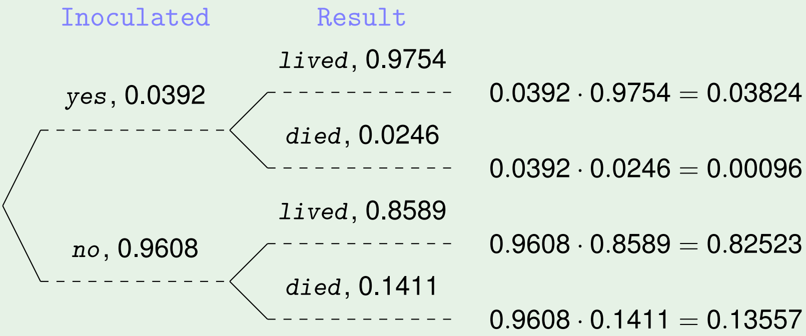
Definition

A **tree diagram** is a tool to organize outcomes and probabilities around the structure of the data.

They are most useful when two or more processes occur in a sequence and each process is conditioned on its predecessor.

Example 20

Here is the tree diagram for `smallpox` dataset.

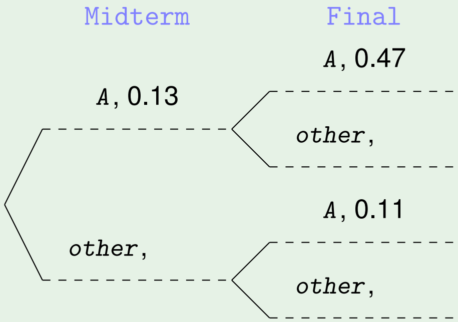


Example 21

Suppose that 13% of students earned an A on the midterm. Of those students that earned an A, 47% received an A on the final and 11% of the student who earned a lower grade than an A on the midterm received an A on the final.

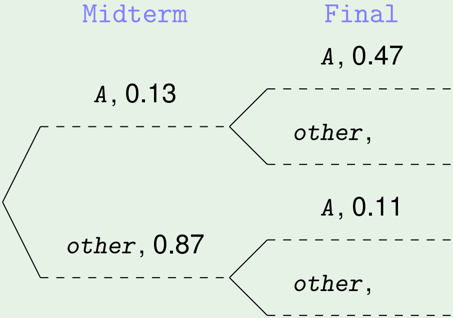
Example 21

Suppose that 13% of students earned an A on the midterm. Of those students that earned an A, 47% received an A on the final and 11% of the student who earned a lower grade than an A on the midterm received an A on the final.



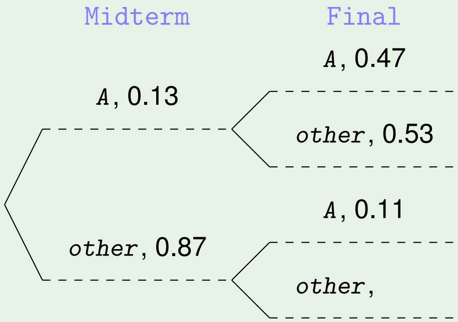
Example 21

Suppose that 13% of students earned an A on the midterm. Of those students that earned an A, 47% received an A on the final and 11% of the student who earned a lower grade than an A on the midterm received an A on the final.



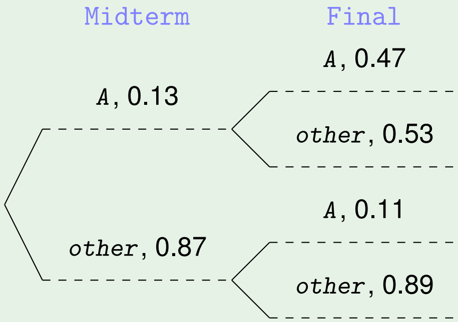
Example 21

Suppose that 13% of students earned an A on the midterm. Of those students that earned an A, 47% received an A on the final and 11% of the student who earned a lower grade than an A on the midterm received an A on the final.



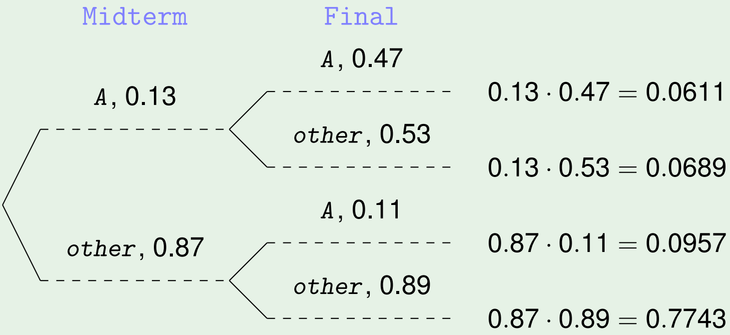
Example 21

Suppose that 13% of students earned an A on the midterm. Of those students that earned an A, 47% received an A on the final and 11% of the student who earned a lower grade than an A on the midterm received an A on the final.



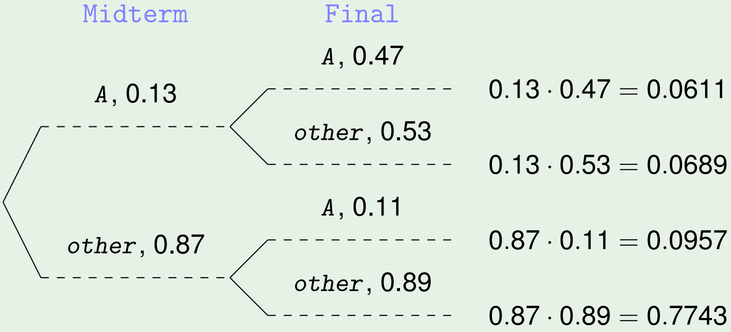
Example 21

Suppose that 13% of students earned an A on the midterm. Of those students that earned an A, 47% received an A on the final and 11% of the student who earned a lower grade than an A on the midterm received an A on the final.



Example 21

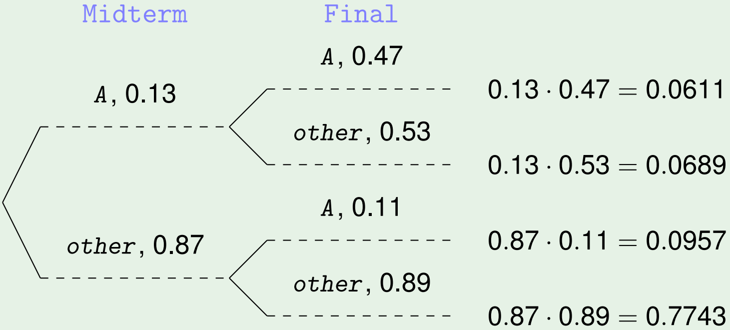
Suppose that 13% of students earned an A on the midterm. Of those students that earned an A, 47% received an A on the final and 11% of the student who earned a lower grade than an A on the midterm received an A on the final.



$$P\left(\texttt{midterm}\text{ is }A \mid \texttt{final}\text{ is }A\right)$$

Example 21

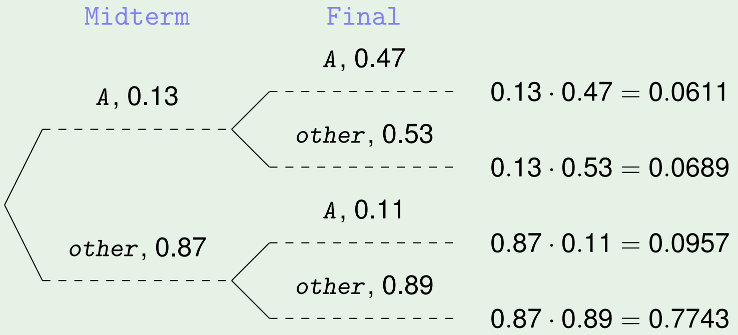
Suppose that 13% of students earned an A on the midterm. Of those students that earned an A, 47% received an A on the final and 11% of the student who earned a lower grade than an A on the midterm received an A on the final.



$$P\left(\texttt{midterm is } A \mid \texttt{final is } A\right) = \frac{P\left(\texttt{midterm is } A \text{ and } \texttt{final is } A\right)}{P\left(\texttt{final is } A\right)}$$

Example 21

Suppose that 13% of students earned an A on the midterm. Of those students that earned an A, 47% received an A on the final and 11% of the student who earned a lower grade than an A on the midterm received an A on the final.

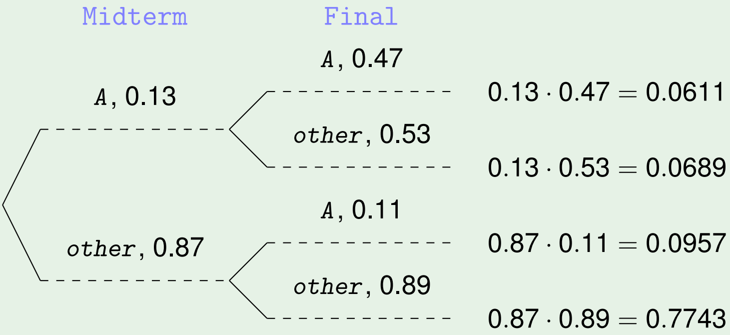


$$P\left(\texttt{midterm is } A \mid \texttt{final is } A\right) = \frac{P\left(\texttt{midterm is } A \text{ and } \texttt{final is } A\right)}{P\left(\texttt{final is } A\right)}$$

$$= \frac{0.0611}{0.1568}$$

Example 21

Suppose that 13% of students earned an A on the midterm. Of those students that earned an A, 47% received an A on the final and 11% of the student who earned a lower grade than an A on the midterm received an A on the final.



$$P\left(\texttt{midterm is } A \mid \texttt{final is } A\right) = \frac{P\left(\texttt{midterm is } A \text{ and } \texttt{final is } A\right)}{P\left(\texttt{final is } A\right)}$$

$$= \frac{0.0611}{0.1568} = 0.3897 \approx 38.97\%$$