1 Probability of event A conditional on event B occurring

E.R. Morey: ConditionalProbability.tex September 8, 2010

Let $\Pr[A\,|B\,] \equiv$ Probability of event A conditional on event B occurring, assuming $\Pr[B] > 0$; if $\Pr[B] = 0$, $\Pr[A\,|B\,]$ is not defined.

Consider the following: $\Omega \equiv \{(x, y) : 0 \leq x \leq 100,\ 0 \leq y \leq 100\}$.
Now consider two sets: $A \equiv \{(x, y) : 20 \leq x \leq 50,\ 40 \leq y \leq 60\}$
and $B \equiv \{(x, y) : 0 \leq x \leq 40,\ 50 \leq y \leq 100\}$.

Note that A and B are each events. Note that in this example, these two event sets partially intersect.

We could ask the probability of a draw from A given that we are drawing from set B, this is a conditional probability.

Consider other examples where $A \subset B$ or $A = B$. Draw some Venn diagrams.

Definition 1 *of conditional probability (1)*

$$\Pr[A\,|B\,] = \frac{\Pr[AB]}{\Pr[B]}$$

if $\Pr[B] > 0$. *Read* $\Pr[AB] \equiv \Pr[A \cap B] \equiv \Pr[A \textit{ and } B]$.

Why can't the conditional probability be defined as $\Pr[A\,|B\,] = \Pr[AB]$?

Since there is nothing special about the names A and B, (2)

$$\Pr[B\,|A] = \frac{\Pr[AB]}{\Pr[A]}$$

if $\Pr[A] > 0$. One can rearrange (1) and (2) to obtain (1a) and (2a)

$$\Pr[AB] = \Pr[A\,|B\,]\Pr[B]$$

and

$$\Pr[BA] = \Pr[B\,|A]\Pr[A]$$

Note that $\Pr[AB] = \Pr[BA]$
Combining (1a) and (2a), one obtains (3)

$$\Pr[AB] = \Pr[A\,|B\,]\Pr[B] = \Pr[B\,|A]\Pr[A] = \Pr[BA]$$

What is the intuition behind defining conditional probability as we did above? In conditional probability, the sample space is effectively reduced to B; that is, the probability that A will happen given that ones lives in a world where B prevails.

So, $\Pr[A\,|B\,]$ is the probability that both A and B occur, $\Pr[AB]$, as a proportion of the probability of B, $\Pr[B]$.

Note the following, which following from (2) and (3)–divide (3) by $\Pr[A]$:

$$\begin{aligned} \Pr[B\,|A] &= \frac{\Pr[BA]}{\Pr[A]} = \frac{\Pr[AB]}{\Pr[A]} \\ &= \frac{\Pr[A\,|B]\,\Pr[B]}{\Pr[A]} \end{aligned}$$

This says that if one knows, $\Pr[A\,|B]$, $\Pr[B]$, and $\Pr[A]$, one can determine $\Pr[B\,|A]$. What is this result called? Bayes' theorem?

Bayes'theorem is quite useful. I have used it in my research on numerous occasions.

For example, assume that one has determined that there are 4 classes/types of individuals and that one has determined the probability that an individual will choose alternative 1 conditional on being a member of class III, $\Pr[1\,|III]$, determined the probability of answering 1, $\Pr[1]$, and determined the probability of being in class III, $\Pr[III]$.

One can use Bayes'theorem to determine the probability that an individual is in class III given that they chose alternative 1, $\Pr[III\,|1]$

Work it out: Applying Bayes' theorem

$$\Pr[III\,|1] = \frac{\Pr[1\,|III]\,\Pr[III]}{\Pr[1]}$$

There are many problems in the first review set about Bayes' theorem. Bayes' theorem is **amazing**. Look at the vegetarian-dog problem and the gubergomer problem

1.1 When are events A and B independent? Not independent

Note that, at this point, we are talking about *events* not *random variables*.

Start by considering a case where $\Pr[A] > 0$ and $\Pr[B] > 0$, but $A \cap B = \phi$, so $\Pr[AB] = 0$. If these assumptions hold, $\Pr[A\,|B\,] = \Pr[B\,|A\,] = 0$; that is, the two events are mutually exclusive. Draw a Venn diagram.

In this case, are A and B independent? NO. One event happening precludes the other event from happening: they are dependent.

When do we say A and B are independent?

Definition 2 *A and B are independent iff the following are true (each is a different way of saying the same thing):*

$\Pr[AB] \equiv \Pr[A \cap B] = \Pr[A]\Pr[B]$

$\Pr[A\,|B\,] = \Pr[A]$ *if* $\Pr[B] > 0$

$\Pr[B\,|A\,] = \Pr[B]$ *if* $\Pr[A] > 0$

Note that these are three equivalent statements: each implies the other two. Hopefully, you can show this? There is only one piece of information in the three equations, not three.

The following can be deduced from the definition of independence:

1. $\Pr[A] > 0$, $\Pr[B] > 0$ and $A \cap B = \phi \Rightarrow A$ and B not independent (this was my example above) - they are dependent

   Draw a Venn diagram to convince yourself this is true. I typically get confused on this issue: if two sets have no intersection, it is easy to wrongly conclude that they are independent–they don't tough each other.

   In the terminology of necessary and sufficient, 1. says ($\Pr[A] > 0$, $\Pr[B] > 0$ and $A \cap B = \phi$) is sufficient for (A and B not independent). In terms of the arrow, $\Longrightarrow$[1]

$$(\Pr[A] > 0, \Pr[B] > 0 \text{ and } A \cap B = \phi) \Longrightarrow (A \text{ and } B \text{ not independent})$$

Another way of expressing this is not(A and B not independent) $\Longrightarrow not(\Pr[A] > 0, \Pr[B] > 0 \text{ and } A \cap B = \phi)$.

Which is equivalent to (A and B are independent) $\Longrightarrow not(\Pr[A] > 0, \Pr[B] > 0 \text{ and } A \cap B = \phi)$.

[1] To say that x is sufficient for y is equivalent to saying x implies y (the existence of x guarentees y). It follows by the rules of logit that if x implies y, then $noty$ implies $notx$.

The following can also be deduced:

2. $\Pr[A] > 0$, $\Pr[B] > 0$ and A and B are independent $\Rightarrow A \cap B \neq \phi$

In the terminology of necessary and sufficient 2. says ($\Pr[A] > 0$, $\Pr[B] > 0$ and, A and B are independent) is sufficient for ($A \cap B \neq \phi$).

Another way of expressing this is $not(A \cap B \neq \phi) \Longrightarrow not(\Pr[A] > 0$, $\Pr[B] > 0$ and A and B are independent)

Which is equivalent to $(A \cap B = \phi) \Longrightarrow not(\Pr[A] > 0$, $\Pr[B] > 0$ and A and B are independent).

Think about how 1. and 2. are different? They are not equivalent statements.

Then think about the distinction between the following two statements

$\Pr[A] > 0$, $\Pr[B] > 0$ and $A \cap B = \phi \Rightarrow A$ and B are dependent (not independent) - this is 1. above

and

$A \cap B = \phi$ is consistent with A and B being independent

Both of these statements are correct statements. But, if $A \cap B = \phi$, A and B are independent only if one of the sets is empty ($P[A]P[B] = 0$).

Notice the first statement says that when $A \cap B = \phi$ and some other conditions hold, A and B are dependent, while the second statement notes that $A \cap B = \phi$ and A and B can be independent.

If two non-empty sets are independent, they must have common elements, a non-empty intersection. Does it go the other way? Does $A \cap B \neq \phi$ imply that A and B are independent? No.

Independence of n events is more complicated than independence of 2 events.

Definition 3 *A_1, A_2, A_3 and A_4 are independent iff* ***all*** *of the following are true:*

a) $\Pr[A_i A_j] = \Pr[A_i]\Pr[A_j]$ $\forall i$ *and* j, $i \neq j$

b) $\Pr[A_i A_j A_k] = \Pr[A_i]\Pr[A_j]P[A_k]$ $\forall i, j$ *and* k, $i \neq j \neq k$

c) $\Pr[A_i A_j A_k A_l] = \Pr[A_i]\Pr[A_j]P[A_k]P[A_l]$ $\forall i, j, k$ *and* l, $i \neq j \neq k \neq l$

Three things to note:

1. Independence of J sets requires pairwise independence between all the sets, three-wise independence between all the sets, four-wise independence ....J-wise independence. That is, all of the subsets of the J sets must exhibit independence.

2, c) does not imply b) and b) does not imply a)

MGB has an demonstration of this. For example, does $\Pr[A_1 A_2 A_3] = \Pr[A_1]\Pr[A_2]P[A_3]$ imply $\Pr[A_1 A_2] = \Pr[A_1]\Pr[A_2]$?

One can show that it does not with a counter-example. Consider $P[3] = 0$, in which case $\Pr[A_1 A_2 A_3] = \Pr[A_1]\Pr[A_2]P[A_3] = 0$, but this does not imply that $\Pr[A_1 A_2] = \Pr[A_1]\Pr[A_2]$. The later only hold if A_1 and A_2 are independent.

3. a) does not imply b) and b) does not imply c). MGB has a counter-example on page 42 of my version.

1.1.1 Independence, sets, and Venn diagrams,

1.1.2 Some student examples of independence and dependence

Example 1 Suppose each member of a population is classified as male or female, and as favoring or opposing the death penalty. Each individual belongs to one of four mutually exclusive groups: MD, MnD, FD, FnD

Anyone here from Texas?

One can view this as there being two events: the event male, M, and the event pro-death penalty, D. Female is the compliment of M. The question is whether the events M and D are independent.

Suppose, the proportions in each category were[2].

	Death	not Death
Male	.459	.441
Female	.051	.049

Where, for example, .051 is the probability, in this population, of being a female and favoring the death penalty - be aware that most of this population is male (maybe we are talking about the population of football players).

Note that $.459 + .441 + .051 + .049 = 1.0$ Events include : being male, being female, favoring death penalty, not favoring death penalty, being something, favoring something, etc.

The sum of the male row is the probability that an individual sampled from this population is a male; the sum of the female row is the probability that they are a female.

The sum of the first column is the probability that a sampled individual is for the death penalty.

[2]You can interpret this table as a Venn diagram for your population of people, where the fractions in each cell determine the total number of people that exhibit those two events.

In this example, the probability of an individual favoring the death penalty, conditional on being male is

$$\begin{aligned}\Pr(D\,|M\,) &= \frac{\Pr(MD)}{\Pr(M)} \\ &= \frac{.459}{.459 + .441} = .51\end{aligned}$$

But the $\Pr(D) = .459 + .051 = .51$, the same as $\Pr(D\,|M\,)$.

So the probability of favoring the death penalty equals the probability of favoring the death penalty conditional on one being male, so the two events (favoring death and being male) are independent.

As required for independence, the probability of an individual favoring the death penalty, conditional on being female (not a male) is

$$\begin{aligned}\Pr(D\,|F\,) &= \frac{\Pr(FD)}{\Pr(F)} \\ &= \frac{.051}{.051 + .049} = .51\end{aligned}$$

Note that we could have checked using a different approach.

For example if D and M are independent, the $\Pr[DM] = \Pr[D]\Pr[M]$

In this case, $\Pr[D] = .51$ and $\Pr[M] = .459 + .441 = 0.9$, so $\Pr[D]\Pr[M] = .9(.51) = 0.459 = \Pr(MD)$

However, if the proportions were

	Death	not Death
Male	.27	.21
Female	.24	.28

$$\begin{aligned}\Pr(D\,|M\,) &= \frac{\Pr(DM)}{\Pr(M)}\\ &= \frac{.27}{.27+.21} = .5625\end{aligned}$$

and $\Pr(D) = .27 + .24 = .51$, so the the probability of favoring the death penalty does not equal the probability of favoring the death penalty conditional on one being male, so the two events (favoring death and being male) are dependent. Males are more likely to prefer the death penalty.

We also see this by examining

$$\begin{aligned}\Pr(D\,|F\,) &= \frac{\Pr(FD)}{\Pr(F)}\\ &= \frac{.24}{.24+.28} = .46\end{aligned}$$

It also does not equal .51, as we know it cannot from the male result.

Can you draw some Venn diagrams representing the above examples that would be insightful?

Can you come up with some other examples of independence and dependence?

1.2 Cogitating on the existence god (God) in terms of conditional probabilities and Bayes' theorem

Does God exist. There are many "proofs" of the existence of god.

One is argued along the following lines: Miracles happen and this proves the existence of god.

Let's investigate in terms of Bayes' theorem:

Let $P[G\,|M\,]$ be the probability that god exists given that at least one miracle occurred, and, $\Pr[M\,|\overline{G}]$ be the probability of of at least one miracle in a universe with no god.

Use Bayes' Theorem to evaluate this argument for the existence of god.

Let me know if you disagree or agree with what follows.

I am going to use the following framework for thinking about this problem: Imagine that there is a Universe of universes, some have a god, some don't.[3]

So the Universal set, Ω, is the set of all possible universes. Consider our universe (our earth, solar system and beyond) a random draw from that set.

So, one event is the universe has a god, G. Another event is at least one miracle occurs in the universe, some universes have miracles, others don't. $P[G]$, the probability that there will be a god in a drawn universe; it is the proportion of universes that have a god, and $P[M]$ is the proportion of universes where at least one miracle occur. Consider some of the different statements of Bayes' theorem in this situation.

$$\Pr[G\,|M\,] = \frac{\Pr[M\,|G]\Pr[G]}{\Pr[M]}$$

$$\Pr[M\,|G] = \frac{\Pr[G\,|M\,]\Pr[M]}{\Pr[G]}$$

$$\Pr[M\,|\overline{G}] = \frac{\Pr[\overline{G}\,|M\,]\Pr[M]}{\Pr[\overline{G}]}$$

and

$$\Pr[G\,|\overline{M}\,] = \frac{\Pr[\overline{M}\,|G]\Pr[G]}{\Pr[\overline{M}]}$$

[3] We won't get into how the universes were created or how gods get associated with some universes and not with others.

Many advocates for the existence of god would argue that only god (or her helpers) can produce miracles, which means $\Pr[M\,|\overline{G}\,] = \Pr[\overline{G}\,|M\,] = 0$ - if there is no god there are no miracles and if there are miracles there is a god. But $\Pr[\overline{G}\,|M\,] = 0$ implies that $\Pr[G\,|M\,] = 1$ and $\Pr[M\,|\overline{G}\,] = 0$ implies that $[\Pr[\overline{M}\,|\overline{G}\,] = 1$.

Note that no one is arguing that god must do miracles.

Substituting these two assumptions in the above four expressions of Bayes' theorem, one gets

$$1 = \Pr[G\,|M\,] = \frac{\Pr[M\,|G\,]\Pr[G]}{\Pr[M]} \text{ which implies } \Pr[MG] = \Pr[M]$$

$$\Pr[M\,|G\,] = \frac{1\Pr[M]}{\Pr[G]} = \frac{\Pr[M]}{\Pr[G]}$$

$$0 = \Pr[M\,|\overline{G}\,] = \frac{(0)\Pr[M]}{\Pr[\overline{G}]} = 0 \text{ if } \Pr[\overline{G}] > 0$$

and

$$\Pr[G\,|\overline{M}\,] = \frac{\Pr[\overline{M}\,|G\,]\Pr[G]}{\Pr[\overline{M}]}$$

So, what does all of the above tell us.

The first equation simply says that if a miracle occurs in our universe then we have a god in our universe.

The practical question is how does one determine whether a miracle has occurred; this raises the issue of what is a miracle. Given the assumptions, one can prove the existence of god by proving at least one miracle has occurred.

The second line says that the probability that a miracle occurs given a god, is the probability of a miracle divided by the probability of a god.

What about the third line? It simply says if no god, no miracles.

The fourth line says there still might be a god even if one does not observe a miracle.

So what did I learn from writing all of the above down? A bunch about Bayes' theorem and that if one assumes only god can perform miracles and one want to prove god exists, go try an find an example of a miracle.

Did advocates of this theory prove there is a god? Yes, given there assumptions, if they proved there has been a miracle.

Come up with a numerical example where two events are both independent and disjoint.

For more on Baye's theorem see

our problem set on probability theory
Jeff Merrell's lecture note/example with fighter pilots

Wikipedia at
http://en.wikipedia.org/wiki/Bayes'_theorem

The Stanford Encyclopedia of Philosophy at

http://plato.stanford.edu/entries/bayes-theorem/