

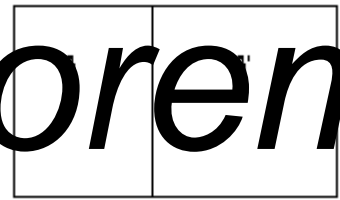
$$P(A \cap B) = P(A | B) * P(B)$$

Explaining

Given A subset B then $P(A | B) = \frac{\text{card}(A)}{\text{card}(B)}$

Bayes Theorem

$$P(S) = P(A) + P(A')$$



$$A \cap A' = \emptyset$$

and $A \cup A' = S$

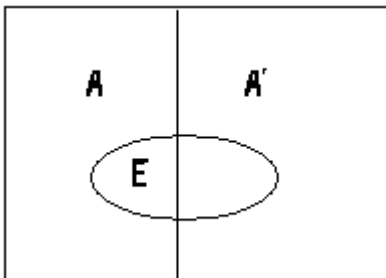
A Progression of

$$P(A \cap B) = P(B \cap A)$$

Ideas Illuminating an

Important Mathematical Concept

$$P(E) = P(E \cap A) + P(E \cap A')$$



$$P(A | B) = \frac{P(B | A) * P(A)}{P(B)}$$

By Dan Umbarger
www.mathlogarithms.com

Foreword

When I did my undergraduate work it was “de rigueur” for computer science majors to also study calculus in order to complement the training we were taking in the Computer Science. Additionally, since all our professors were also mathematics teachers, we usually had assignments requiring knowledge of math. As time went on Computer Science at the post high school level has pretty well separated from mathematics and, at least for the non-engineer majors, requires less function based mathematics and more discrete math and statistics. I do not wish to debate whether or not a high school student contemplating a university computer science major should take calculus or statistics or discrete math. It is quite clear to me that they should take as much mathematics as their high school offers.

One area of non-calculus based mathematics that has been increasingly used in artificial intelligence applications of computer science is “probability based decision making.” High school teachers have always taught decision making (if, if-else, switch-case) based upon randomly generated numerical values. At the university and applied level we see decisions being made based upon more sophisticated criteria...conditional probability. A major part of this work is predicated upon knowledge of Bayes’ Theorem. Bayes’ Theorem is named after the Reverend Thomas Bayes of England, 1702-1761 (who reportedly developed it while trying to prove the existence of God) although there is some evidence that a Nicholas Saunderson, the blind Lucasian professor at Cambridge (think Sir Issac Newton and Stephen Hawkings!!) actually discovered the principles that are generally credited to Rev. Bayes.
http://en.wikipedia.org/wiki/Nicholas_Saunderson

In Computer Science Bayes Theorem is used in enhancing low resolution imaging and in filtering situations such as spam and noise filters...all situations involving “probability based decision making.”

As a long time high school computer science teacher it occurs to me that such a powerful tool should be made available to interested high school students. We all have students who are pushing beyond the AP Computer Science curriculum and applications of Bayes Theorem would be a good topic for such a student to investigate. It could possibly benefit them greatly after high school. As I was not able to locate any high school age appropriate materials explaining Bayes Theorem I have determined to try to fill the void. When I looked back at my university business stat book I found the standard Bayes Theorem formulas given with no discussion or motivation or progression of ideas anywhere... a great example of “Athena jumping out of the head of Zeus.” No wonder I never understood Bayes Theorem while in university! My math stat book was too advanced for me and my business stat book was too trivial and vague to understand...just magic formulas to manipulate and quickly forget after the unit test Hopefully I have sailed between Scylla and Charybdis here.

By traditional standards I am not particularly qualified to undertake this task but in a strange way that makes me even more suited to such a task. I noticed some time ago that many (most) of those people for whom abstract ideas come easily cannot communicate well with others who do not have similar background and/or ability. Many (most) of my high school computer geniuses can jibber-jabber up a storm with their intellectual equals but are totally ineffective in trying to communicate with less capable classmates. My struggle to master concepts usually results in developing visuals. Those visuals achieve two goals. They allow my brighter students to master the material even more quicker than ever while still providing access of ideas to those less academically gifted students who still need or could benefit from the material.

Table of Contents

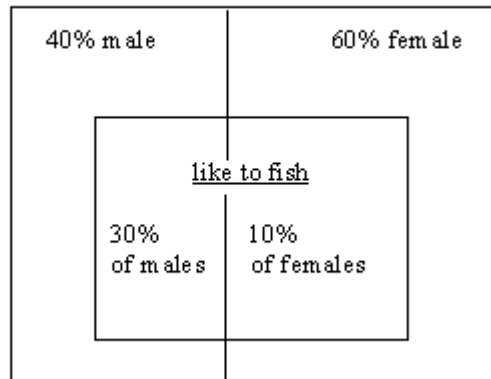
	Page
Chapter 1, Bayes Theorem, An Anticipatory Set.....	4
Chapter 2, Bayes Theorem, An Informal Derrivation.....	6
Chapter 3, Bayes Theorem, Applied, No Grid.....	13
Chapter 4, Extended Bayes Theorem.....	15

Chapter 1, Bayes Theorem, An Anticipatory Set

In the educational world an “anticipatory set” is a preliminary discussion of a topic that introduces ideas and vocabulary and “sets the stage” so to speak for what is to come. It gives the student the “big picture” before proceeding to work with the details.

Problem: A given population is 40% male and 60% women. 30% of the males like to fish but only 10% of the females like to fish. A visual can make all this data immediately available to our brains to process.

Population P



List all the probabilities this data suggests. Several are obvious from the figure.

Probability a person chosen at random from set P is male, $P(\text{male}) = 40\%$ of entire population
 Probability a person chosen at random from set P is female, $P(\text{female}) = 60\%$ of entire population

Probability of picking at random from set of males a subject who likes to fish:
 $P(\text{fish} | \text{male}) = 30\%$ of males... given... duh! (but $30\% * 40\% = 12\%$ of entire population)

Probability of picking at random from the set of females a subject who likes to fish:
 $P(\text{fish} | \text{female}) = 10\%$ of females... given ... duh! (but $10\% * 60\% = 6\%$ of entire population)

Notice the symbolism $P(\text{fish} | \text{male})$ and $P(\text{fish} | \text{female})$.

$P(\text{fish} | \text{male})$ is read as “the probability of choosing a person who likes to fish from the set of males.”

$P(\text{fish} | \text{female})$ is read as “the probability of choosing a person who likes to fish from the set of females.”

Probabilities of the form $P(\text{set 1} | \text{set 2})$ are called “conditional probabilities.”

Question: What would be the probability of randomly choosing a male from the set of people who like to fish? Symbolically this would be $P(\text{male} | \text{fish}) = ?$

It was pretty easy to ascertain $P(\text{fish} | \text{male})$. (30% of males but $30\% * 40\% = 12\%$ of entire population)
 It is not so easy to ascertain $P(\text{male} | \text{fish})$. That is what learning Bayes Theorem will enable you to do.

$$P(\text{male} | \text{fish}) = \frac{P(\text{fish} | \text{male}) * P(\text{male})}{P(\text{fish})}$$

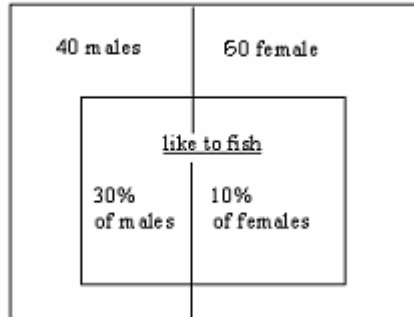
Bayes Theorem, A Sneak Peak

Notice that the values $P(\text{fish} | \text{male})$ and $P(\text{male})$ are readily available from the given data but value $P(\text{male} | \text{fish})$ and $P(\text{fish})$ are not!

Suppose the original problem was changed as follows:

Problem: In a given population are 40 males and 60 female. 30% of the males like to fish but only 10% of the females like to fish. A visual can make all this data immediately available to our brains to process.

Population P

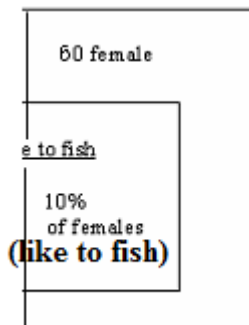
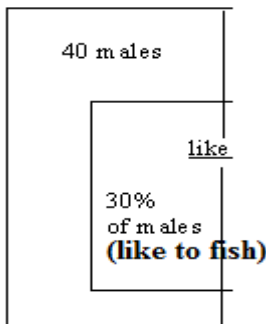


Here we have cardinal numbers for our gender data rather than percentages. Again list all the information that this data implies.

What % of the males like to fish? ... 30% What % of the females like to fish?... 10%
 What % of the population is male?...40/100. What % of the population is female?...60/100.

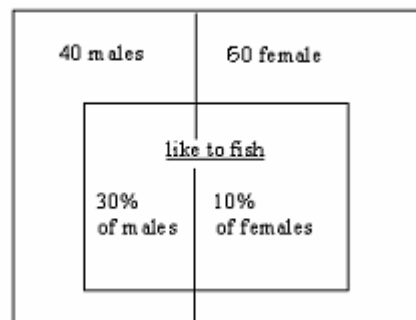
How many males are there in the population? ... 40. How many females are in the population?...60
 How many males like to fish? $30\% * 40 = 12$ How many females like to fish? $10\% * 60 = 6$

This new data can be calculated from the given data.



Given that a person is chosen at random from the set of people who like to fish what is the probability that that person is a male? $P(\text{male} | \text{fish})$.
 Although slightly disguised this is same situation as posed on the previous page and is the kind of problem that Bayes theorem will allow us to solve. (Note that if $\# \text{male} + \# \text{female} \neq 100$ we could use the given counts and calculate %.)

Population P



Chapter 2, Bayes Theorem, An Informal Derrivation

Where in the world does Bayes Theorem come from? How do we know that it works? Smart university math teachers are more than eager to derrive this theorem. Even if I knew how to derrive this theorem it would not be appropriate to do so in materials which are being targeted to high school students. What follows is considered an informal, non-rigorous, intuitive, demonstration which will appropriately serve to create a familiarity with and a confidence in Bayes Theorem.

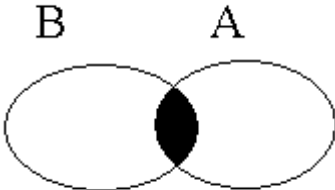
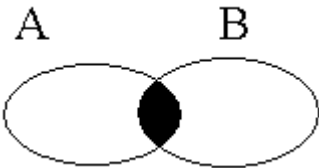
$$P(male | fish) = \frac{P(fish | male) * P(male)}{P(fish)}$$

We start by learning (or reviewing) a basic rule in set theory...

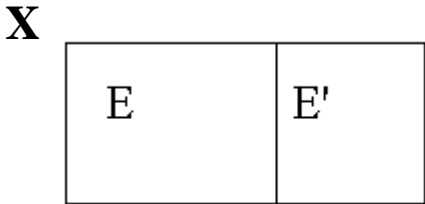
$(A \cap B) = (B \cap A)$

The Commutative Property of Interesection

The figures below seem to support this idea. We could use a truth table but it seems unnecessary.

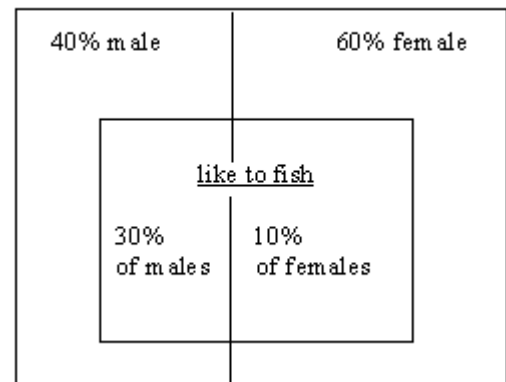


Another basic rule from set theory which we will need and you should recognize from Geometry is that the “whole is equal to the sum of its parts.” In the figure shown below set $X = \text{set } E \cup \text{set } E'$. Notice that this rule only works if the parts are mutually exclusive and entirely partition set X.

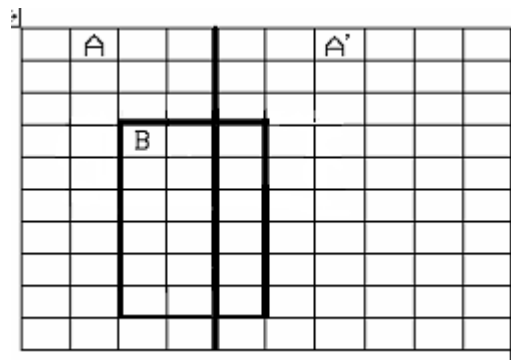


We now move back to the example shown before (at right) but this time will work with a more generic situation. The population P will be broken into mutually exclusive subspaces A and A' with subset B straddling spaces A & A' (below right). The spaces A and A' will not represent people but will merely represent generic sets. Subset B will not represent people who like to fish but merely a set which straddles two other sets. Furthermore each of the spaces will be gridded so that we can count units in each of the respective areas. This latter trick will help us to make the necessary deductive leap allowing us to “derive” Bayes Theorem.

Population P



Sample space - S



Paralleling the previous example we divide The 100 space population grid into two subsets, A and A'. Set A will have 40 tiles (40% of the region) while Set A' will have 60 tiles (60% of the region.) Now we need to draw subset B so that it covers 30% of set A while at the same time covering 10% of set A'.

30% * 40 tiles = 12 tiles while at the same time
10% * 60 tiles = 6 tiles.

So we draw set B into our grid with 18 tiles,
12 of them in set A and 6 of them in set A'.

We see now that the original figure (above) was not draw proportional to the given information.

Note 1:) $S = A \cup A'$. 2.) $Set B \cap Set A$ also $Set B \cap Set A'$ Finally 3.) $A \cap A' = \emptyset$

By counting grid tiles in each of the regions above answer the following questions.

Question: If you were blindfolded and randomly threw a dart so that you hit the dart board every time...

What % of the time would your dart hit region A? $P(A) = 40\%$
 What % of the time would your dart hit region A'? $P(A') = 60\%$
 What % of the time would your dart hit region B? $P(B) = 18\%$ (count the tiles)

Given your dart hits region A what % of the time will it also hit region B?

$$P(B | A) = 12 / 40 = 6 / 20 = 30\% \text{ (of region A)}$$

Given your dart hits region A' what % of the time will it also hit region B?

$$P(B | A') = 6 / 60 = 10\% \text{ (of region A')}$$

What % of the time would your dart hit region B and also be in region A?

$$P(A | B) = (12 / 18 = 2 / 3) = 66 \& 2/3 \%$$

(Information available by counting grid tiles.)

What % of the time would your dart hit region B and also be in region A'?

$$P(A' | B) = (6 / 18 = 1 / 3) = 33 \& 1/3 \%$$

(Information available by counting grid tiles.)

Gathering together all the observations from the bottom of pg. 7

$$P(A) = 40\%, \quad P(A') = 60\%$$

$$P(B) = 18\%$$

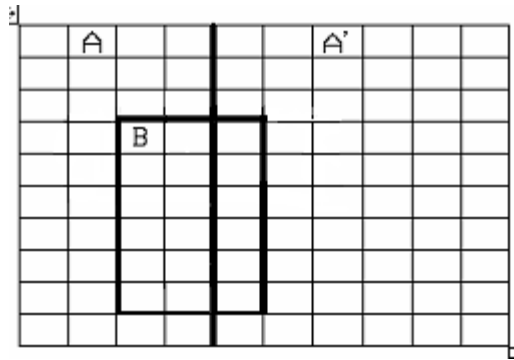
$$P(B|A) = 30\%, \quad P(B|A') = 10\%$$

$$P(A|B) = 2/3 \quad (2/3 \text{ of } B \text{ grids are in } A)$$

$$P(A'|B) = 1/3 \quad (1/3 \text{ of } B \text{ grids are in } A')$$

$$A \cap B = 12 \text{ tiles so } P(A \cap B) = 12/100 = 12\%$$

$$A' \cap B = 6 \text{ tiles so } P(A' \cap B) = 6/100 = 6\%$$



The next part is guaranteed to cause a university mathematics professor to cringe. There is a relation among the different probability values shown above.

$$P(A \cap B) = 12\%, \quad P(A|B) = 2/3, \text{ and } P(B) = 18\%$$

$$12\% = (2/3) * 18\%$$

$$P(A \cap B) = P(A|B) * P(B) \quad *** \text{ special equation \#1}$$

Also there is a relation among the following probability values

$$P(B \cap A) = 12\%, \quad P(B|A) = 30\%, \text{ and } P(A) = 40\%$$

$$12\% = 30\% * 40\%$$

$$P(B \cap A) = P(B|A) * P(A) \quad *** \text{ special equation \#2}$$

All the pieces are now in place....

$$P(A \cap B) = P(B \cap A) \quad \text{commutative property of intersection}$$

$$P(A|B) * P(B) = P(A \cap B) = P(B \cap A) = P(B|A) * P(A) \quad ** \text{ special eqs. \#1 \& 2}$$

$$\text{Thus by transitive} \quad P(A|B) * P(B) = P(B|A) * P(A)$$

Hence

Ta da !

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

Recopying from previous page.

$$P(A | B) = \frac{P(B | A) * P(A)}{P(B)}$$

This is what we wanted... an expression for what we do not know in terms of what we do know (or can find out). In the derivation of this formula we knew $P(A | B)$ to be $2 / 3$ because we had access to grid tiles and could count them to come up with $12 / 18$ which is the probability that a value is in set A given that it is in set B. Likewise $P(B) = 18\%$... we counted Set B grid tiles to get the 18 (12 + 6). However this example is a contrived one utilizing a grid with countable squares which is not usually going to be the case. In general you will only be able to access $P(B | A)$ and $P(A)$. You will have to use a formula to access $P(A | B)$ and $P(B)$.

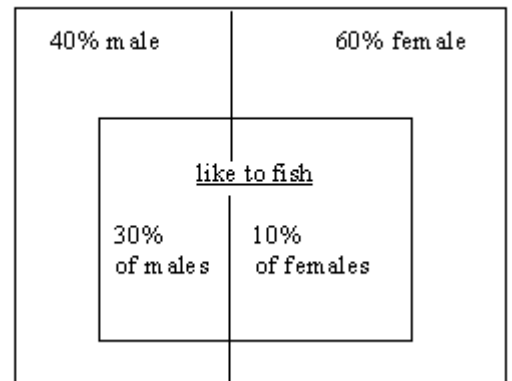
If this paragraph did not make sense take a moment to review the data at right.

$P(\text{like to fish given you are male}) = P(\text{fish} | \text{male}) = 30\%$ of male
 (given) or $(30\% * 40\%) = 12\%$ of entire population
 $P(\text{male}) = 40 / 100$ (given)

$P(\text{likes to fish}) = ???$... cannot do without a grid or a formula.
 [It is **not** $30\% + 10\%$, but rather $\frac{\text{cardinality}(30\%) + \text{cardinality}(10\%)}{\text{cardinality}(\text{population})}$]

$P(\text{someone chosen from the set of fishophiles is a male}) = ???$
 [It is not $(30\% + 10\%) / 40\%$, stay tuned... this is coming.]

Population P



Lets compare this new formula with the one given several pages back on pages 4 & 6.

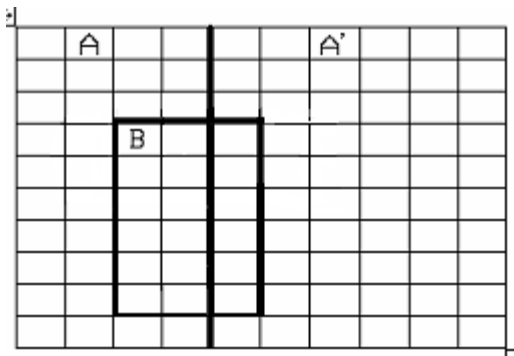
$$P(A | B) = \frac{P(B | A) * P(A)}{P(B)}$$

$$P(\text{male} | \text{fish}) = \frac{P(\text{fish} | \text{male}) * P(\text{male})}{P(\text{fish})}$$

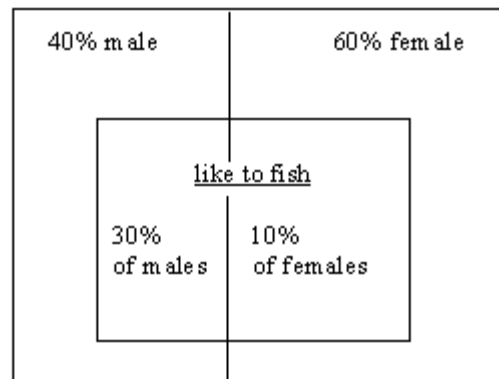
Now if only we could determine $P(B)$ (or $P(\text{fish})$) without a grided figure we would have everything necessary to calculate $P(A | B)$.

Clearly $P(B)$ below is $18 / 100 = 18\%$ by counting grid tiles. (or $(8 + 12) / 100$)

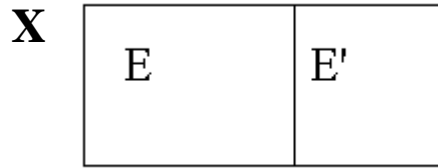
However $P(\text{like to fish})$ is not clear. It is not $30\% + 10\%$!!!



Population P



A basic rule from set theory which you should recognize from Geometry is that the whole is equal to the sum of its parts. In the figure shown below set $X = \text{set } E \cup \text{set } E'$. Notice that this rule only works if the parts are mutually exclusive and entirely partition set X .



Therefore $P(X) = P(E) + P(E')$

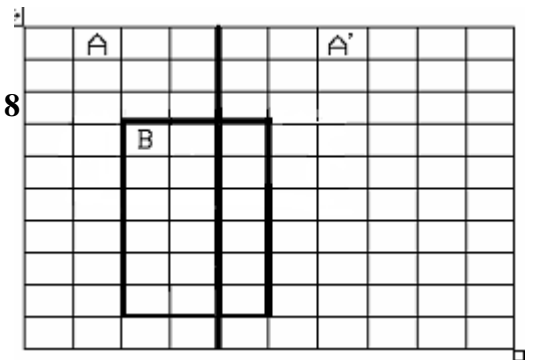
Previously $P(B) = 12 / 100 + 6 / 100$
 $= 18 / 100$

Without the grid tiles in the figure at right, but using the special equation 2 (pg 8)

$P(B \cap A) = P(B | A) * P(A)$ *** special equation #2, pg. 8

we can obtain $P(B)$ as follows.

$$\begin{aligned}
 P(B) &= [(A \cap B) \cup (A' \cap B)] \\
 &= [(B \cap A) \cup (B \cap A')] \\
 &= \{ [P(B | A) * P(A)] \cup [P(B | A') * P(A')] \} \quad \dots \text{special equation \#2, applied twice} \\
 &= [(0.30 * 0.40) + [0.10 * 0.60]] \\
 &= (0.12 + 0.06) \\
 &= 18\% \text{ This works as expected from the grid count} \\
 &\quad \text{because we are adding cardinalities here not \%s (30\% + 10\% \neq 18\%)}
 \end{aligned}$$

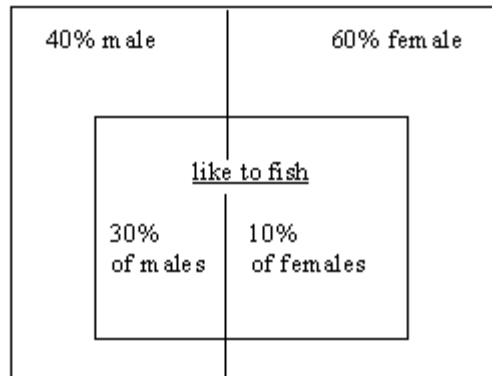


For the figure at right

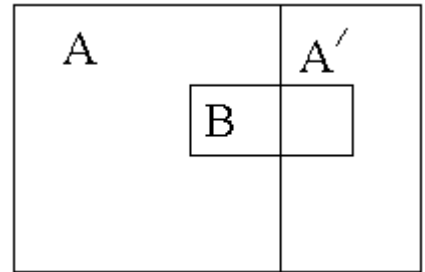
Probability that someone chosen at random from the entire set of data who likes fish is:

$$\begin{aligned}
 P(\text{like to fish}) &= P(\text{male and like to fish}) \\
 &\quad + \\
 &\quad P(\text{female and like to fish}) \\
 &= 30\% * 40\% + 10\% * 60\% \\
 &= 12\% + 6\% \\
 &= 18\% \quad (\text{no counting of grid tiles})
 \end{aligned}$$

Population P



Thus for a sample space subdivided into two mutually exclusive partitions with $A \cup A'$ completely filling the sample space



Bayes Theorem

$$P(A | B) = \frac{P(B | A) * P(A)}{P(B)}$$

However since

$$P(B) = (B \cap A) \cup (B \cap A') \text{ or using special equation 2 from pg. 8 } P(B \cap A) = P(B | A) * P(A)$$

$$P(B) = [P(B | A) * P(A)] \cup [P(B | A') * P(A')] \text{ *** spec. eq. \#2 applied twice on both } A \text{ and } A'$$

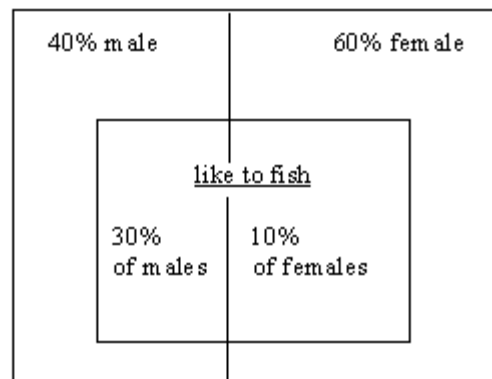
The later P(B) form indicates that the probability of B happening is equal to the sum of the probability of B happening given A + the probability of B happening given A'

Substituting for P(B) in the denominator Bayes Theorem can now be rewritten as

$$P(A | B) = \frac{P(B | A) * P(A)}{P(B | A) * P(A) + P(B | A') * P(A')} \quad \frac{\text{(probability of interest)}}{\text{(probability of either event)}}$$

Population P

or for the fish problem on page 10



$$P(\text{male} | \text{fish}) = \frac{P(\text{fish} | \text{male}) * P(\text{male})}{P(\text{fish} | \text{male}) * P(\text{male}) + P(\text{fish} | \text{female}) * P(\text{female})}$$

$$= \frac{30\% * 40\%}{30\% * 40\% + 10\% * 60\%} = \frac{12\%}{12\% + 6\%} = \frac{12\%}{18\%} = \frac{2}{3}$$

= which we got previously by counting grid tiles (pg. 10)

i.e. $P(\text{male} | \text{fish}) = \frac{\text{probability a male who likes to fish is chosen from the set of males}}{\text{probability of choosing at random from all those who like to fish}}$

A caveat. The reason that college math professors would object to this treatment of Bayes Theorem is that I have taken a sample of size one and used it to deduce a general rule. Induction is actually a commonly used and accepted technique by mathematicians but not in its limited form as shown here. Specifically I have inductively concluded that

$$\begin{aligned}P(A \cap B) &= P(A | B) * P(B) \text{ and} \\P(B \cap A) &= P(B | A) * P(A)\end{aligned}$$

based upon a sample of size one. In the world of mathematics that is not enough evidence to make a generalized conclusion. This paper, however, is not written for mathematicians but for high school students. Unlike the business and psych stat books I had in college it at least tries to motivate and explain where Bayes Theorem comes from.

Chapter 3, Bayes Theorem, Applied...No Grid

When doing an experiment there are four things that can happen. The experiment can:

- 1.) correctly identify something as true (correct true test)
- 2.) incorrectly identify something as true (false positive)
- 3.) correctly identify something as false (correct false test)
- 4.) incorrectly identify something as false (false negative)

For example. Suppose there was a drug test that was developed to test athletes for use of steroids. It is not a 100% accurate test. Sometimes:

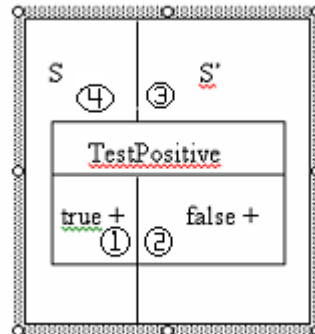
- 1.) it correctly identifies an athlete as being a drug user (true positive)
- 2.) it incorrectly identifies an athlete as being a drug user (false positive)
- 3.) it correctly concludes that an athlete does not use a drug (true negative)
- 4.) it incorrectly concludes that an athlete does not use a drug (false negative)

A manufacturer claims that its test for steroids results in 95% true positives and 15% false positives. 10% of the team uses steroids. Your teammate just tested positive. What is the probability that your friend uses steroids? Let S represent those athletes on your team that use steroids. Hence S' will be those athletes that do not use steroids. Let TP represent all the people who test positive for steroids... $TestPositive = true\ positive + false\ positive$

Your teammate just tested positive. What is the probability that your friend is a user?

You are to calculate $P(S | TestPositive)$

Let $TP = TestPositive$ (true positive + false positive)



The generic Bayes theorems are:

$$P(A | B) = \frac{P(B | A) * P(A)}{P(B)}$$

$$P(A | B) = \frac{P(B | A) * P(A)}{P(B | A)*P(A) + P(B | A')*P(A')}$$

For this specific problem the formulas would be:

$$P(S | TP) = \frac{P(TP | S) * P(S)}{P(TP)}$$

$$P(S | TP) = \frac{P(TP | S) * P(S)}{P(TP | S)*P(S) + P(TP | S')*P(S')}$$

Putting everything together now

A manufacturer claims that its test for steroids results in 95% true positives and 15% false positives.

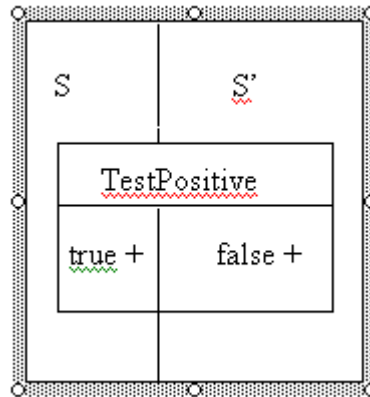
10% of the team uses steroids. Your teammate just tested positive. What is the probability that your he/she uses steroids?

Let S represent those athletes on your team that use steroids. Hence S' will be those athletes that do not use steroids. Let TP represent all the people who test positive for steroids...
 $TP = \text{true positive} + \text{false positive}$

Your friend just tested positive. What is the probability that your friend is a user?

Given: $P(S) = 10\%$
 $P(TP | S) = 95\%$
 $P(TP | S') = 15\%$

Implied: $P(S') = 90\%$



The generic Bayes Theorems are:

$$P(A | B) = \frac{P(B | A) * P(A)}{P(B)}$$

$$P(A | B) = \frac{P(B | A) * P(A)}{P(B | A) * P(A) + P(B | A') * P(A')}$$

For this specific problem the formulas would be:

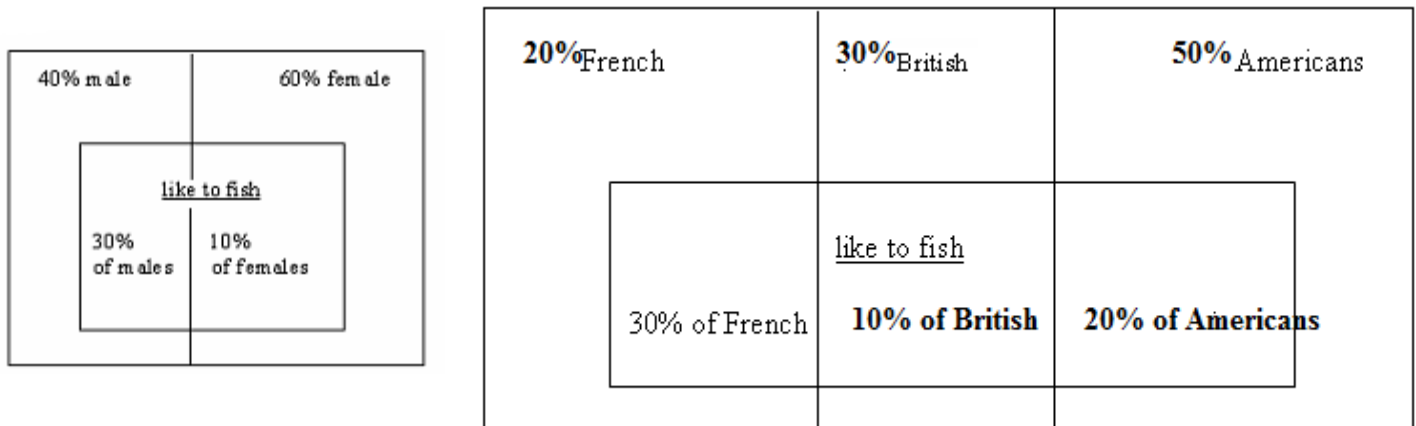
$$P(S | TP) = \frac{P(TP | S) * P(S)}{P(TP)}$$

$$P(S | TP) = \frac{P(TP | S) * P(S)}{P(TP | S) * P(S) + P(TP | S') * P(S')}$$

$$\begin{aligned}
 P(S | TP) &= \frac{95\% * 10\%}{(95\% * 10\%) + (15\% * 90\%)} \\
 &= \frac{0.095}{0.23} \\
 &= 0.413 \\
 &= 41\% \text{ probability that friend uses steroids}
 \end{aligned}$$

Chapter 4, Extended Bayes Theorem

Recall from the original example that the sample space was split into two mutually exclusive subspaces. (See fig. below left.) Bayes Theorem can be extended to apply to those situations where the subspace is split into more than two subspaces (See fig. below right) as long as the subspaces completely partition the original space and the subspaces are mutually exclusive. Notice in the fig. at right that % French + % British + % American = 100%. Assume that there are no subjects with dual citizenship.



What is the probability of picking a person at random from the population of French, British, and American suggested above and

- picking a French citizen? Answer 20% (given)
- picking a British citizen? Answer 30% (given)
- picking an American? Answer 50% (given)

For the new problem condition what is the probability of French who like to fish?

Duh! Could it be 30%? $P(\text{fish} | \text{French}) = 30\%$ (it was given)
(Alternatively $30\% * 20\% = 6\%$ of entire population are French fishermen)

For the new problem condition what is the probability of British who like to fish?

Duh! Could it be 10%? $P(\text{fish} | \text{British}) = 10\%$ (it was given)
(alternatively $10\% * 30\% = 3\%$ of entire population are British fishermen)

For the new problem condition what is the probability of Americans who like to fish?

Duh! Could it be 20%? $P(\text{fish} | \text{Americans}) = 20\%$ (it was given)
(alternatively $20\% * 50\% = 10\%$ of entire population are American fishermen)

Can you anticipate the next question?

Given that all the people who like to fish are standing in a room all by themselves (discussing fish stuff) and that one of those fishophiles were chosen at random what is the probability that person will be French? (be British?) (be American?). Intuitively, for the French, that would be:

$$\frac{\text{Number of French who like to fish}}{\text{Total number of people in entire population}}$$

We can extend the Chapters 1 and 4 Bayes Theorems

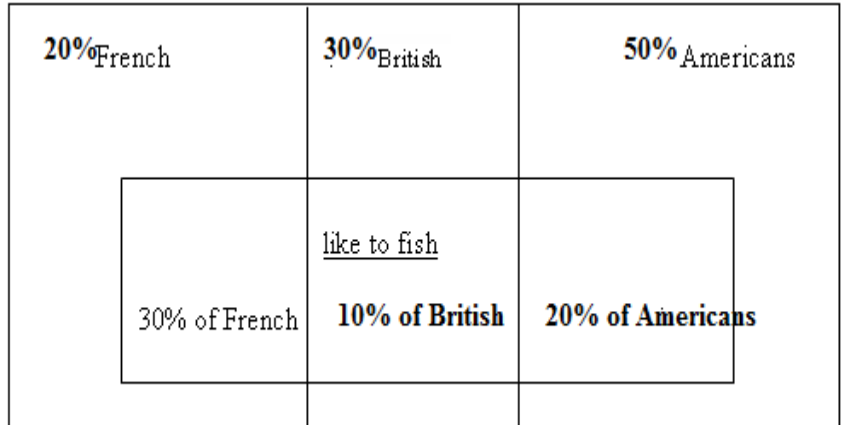
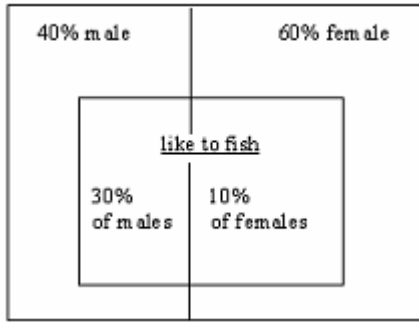
Chap. 1 Bayes Theorem

Chap. 4 Bayes Theorem

$$P(\text{male} | \text{fish}) = \frac{P(\text{fish} | \text{male}) * P(\text{male})}{P(\text{fish})}$$

$$P(\text{French} | \text{fish}) = \frac{P(\text{fish} | \text{French}) * P(\text{French})}{P(\text{fish})}$$

to apply to 3 subspaces



Two subspaces...male & female

Three subspaces...French, British, American

but we must adjust the formula now to reflect that the sample space has three mutually exclusive subspaces and not two.

$$P(\text{fish}) = P(\text{fish} \cap \text{male}) + P(\text{fish} \cap \text{female})$$

$$P(\text{fish}) = P(\text{fish} \cap \text{French}) + P(\text{fish} \cap \text{British}) + P(\text{fish} \cap \text{American})$$

So the two Bayes Theorems would be

$$P(\text{male} | \text{fish}) = \frac{P(\text{fish} | \text{male}) * P(\text{male})}{P(\text{fish} | \text{male})P(\text{male}) + P(\text{fish} | \text{female})P(\text{female})} \quad \text{For 2 subspaces}$$

$$P(\text{French} | \text{fish}) = \frac{P(\text{fish} | \text{French}) * P(\text{French})}{P(\text{fish} | \text{French}) * P(\text{French}) + P(\text{fish} | \text{British}) * P(\text{British}) + P(\text{fish} | \text{American}) * P(\text{American})} \quad \text{For 3 subspaces}$$

What is the probability that a person chosen at random from the group of fishermen is French?

$$\begin{aligned} P(\text{French} | \text{fish}) &= \frac{(30% * 20\%)}{(30% * 20\%) + (10% * 30\%) + (20% * 50\%)} \\ &= \frac{0.06}{0.06 + 0.03 + 0.1} \\ &= \frac{0.06}{0.19} = 0.3947 = 31.58\% \end{aligned}$$

20% _{French}	30% _{British}	50% _{Americans}
30% of French	<u>like to fish</u> 10% of British	20% of Americans

Using the following extended Bayes Theorem

$$P(\mathbf{X} | fish) = \frac{P(fish | \mathbf{X}) * P(\mathbf{X})}{P(fish | French) * P(French) + P(fish | British) * P(British) + P(fish | American) * P(American)}$$

Calculate $P(\text{British} | \text{fish})$ and $P(\text{American} | \text{fish})$. If you are clever you could reuse the denominator for $P(\text{French} | \text{fish})$ at the bottom of the previous page to save yourself a lot of time and effort.

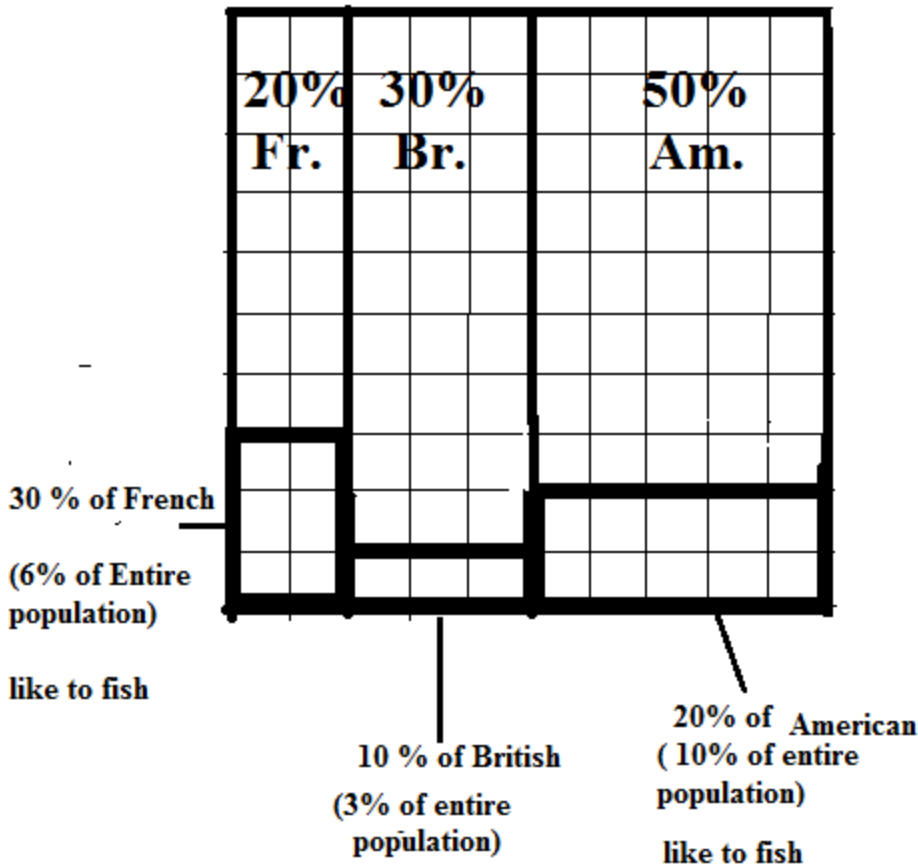
$$P(\text{French} | \text{fish}) = \frac{0.3 * 0.2}{0.06 + 0.03 + 0.1} = \frac{0.06}{0.19} = 31.58\%$$

$$P(\text{British} | \text{fish}) = \frac{.10 * .30}{0.19} = \frac{0.03}{0.19} = 15.79\%$$

$$P(\text{American} | \text{fish}) = \frac{.2 * .5}{0.19} = \frac{.1}{0.19} = 52.63\%$$

Since all the three partitions (French who like to fish, British who like to fish and Americans who like to fish) are mutually exclusive and since that set of people (fishermen) is completely partitioned what would be the probability choosing from the group of fisherman someone who is either French, British, or American? You could find the two unknown answers from above and add up all three probabilities (31.58% + 15.79% + 52.63%) or you could just use logic. Or better yet get the answer both mathematically and using logic and compare the two.

Another way to check our work is to go back to the idea of “gridded spaces.” Notice (below right) that the original figure was not proportional to the respective space sizes.



20% _{French}	30% _{British}	50% _{Americans}
	like to fish	
30% of French	10% of British	20% of Americans

Notice that the total percentage of people who like to fish make up 19% of the entire population
 $P(Fr|fish) + P(Br|fish) + P(Am|fish)$

Probability (X | chosen from the people who like to fish) =

Probability (X chosen from [French fishermen + British fishermen + American Fishermen])

$$P(X | fish) = \frac{P(fish | X) * P(X)}{P(fish | French) * P(French) + P(fish | British) * P(British) + P(fish | American) * P(American)}$$

$$P(\text{French} | \text{fish}) = \frac{.30 * .20}{0.06 + 0.03 + 0.1} = \frac{0.06}{0.19} = 31.58\% \text{ (6/19 by counting grids)}$$

$$P(\text{British} | \text{fish}) = \frac{.10 * .30}{0.19} = \frac{0.03}{0.19} = 15.79\% \text{ (3/19 by counting grids)}$$

$$P(\text{American} | \text{fish}) = \frac{.2 * .5}{0.19} = \frac{.10}{0.19} = 52.63\% \text{ (10/19 by counting grids)}$$