Bayes’ Theorem is my favorite theorem. It enables you to start with a prior probability about a hypothesis and produced an updated (sometimes called posterior) probability, conditional on new evidence/data. Bayes’ Theorem applies to legal and medical cases, as well as many other areas of application. It even applies to philosophy, for some philosophers of science argue that Bayes’ Theorem provides a framework for analyzing how science proceeds. There’s a wonderful new book chronicling the many successes of Bayes’ Theorem, titled *The Theory That Would Not Die: How Bayes’ Rule Cracked the Enigma Code, Hunted Down Russian Submarines and Emerged Triumphant From Two Centuries of Controversy*. Moreover, an entire school of thought about how to conduct statistical inference is based on Bayes’ Theorem.

**Example 5-1: Document Errors**
Suppose that an office employs three associates who prepare a certain kind of document. Delia prepares 60% of these documents, Francis 30%, and Gino 10%. Delia makes an error in 10% of the documents that she prepares, Francis in 25%, and Gino in 5%.

a) Translate the given information into probability statements, using appropriate symbols.

b) Suppose that the given percentages hold exactly for a population of 1000 documents. Fill in the table below. [Hint: Start with the total number of documents prepared by each person. Then proceed to the error/not breakdown for each person.]

<table>
<thead>
<tr>
<th></th>
<th>Contains an error</th>
<th>Does not contain an error</th>
<th>Total documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prepared by Delia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepared by Francis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepared by Gino</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total documents</td>
<td></td>
<td></td>
<td>1000</td>
</tr>
</tbody>
</table>

c) Using the table, report the probability that a randomly selected document is both prepared by Delia and contains an error. Also use appropriate symbols to represent this probability, and report the name of the rule that you have used.

d) Use the table to determine the (unconditional) probability that a randomly selected document contains an error. Also use appropriate symbols to represent this probability, and report the name of the rule that you have used.
e) Now suppose that a randomly selected document is found to contain an error. Use the table to
determine the updated (conditional) probability that Delia prepared it. Also use appropriate
symbols to represent this probability.

f) Indicate how the probability in e) could have been found directly from the given probabilities.

- **Bayes’ Theorem**: If A is any event and $B_1, B_2, \ldots, B_k$ form a partition of the sample
  space $S$, then $\Pr(B_i \mid A) = \frac{\Pr(A \mid B_i) \Pr(B_i)}{\sum_{i=1}^{k} \Pr(A \mid B_i) \Pr(B_i)}$.
  - Notice that Bayes’ Theorem applies when you know, or can easily find, one
    conditional probability but what you want is the reverse conditional probability.
  - If the partition consists of just two events, then Bayes Theorem can be expressed
    as: $\Pr(H \mid E) = \frac{\Pr(E \mid H) \Pr(H)}{\Pr(E \mid H) \Pr(H) + \Pr(E \mid H^c) \Pr(H^c)}$.
    - Updated probability of a hypothesis (H) in light of new evidence (E).

g) Continue to suppose that a randomly selected document is found to contain an error.
Determine the updated (conditional) probability that Francis prepared it. Then determine the
updated (conditional) probability that Gino prepared it.

h) What do you notice about the three updated probabilities?

i) For each of the three associates, compare the (prior) probability that he/she prepared the
document to the updated (conditional) probability given that the document contains an error.

Delia:

Francis:

Gino:
j) Given that a randomly selected document is found to contain an error, who is most likely to have prepared it? Who is least likely? Explain why these make sense.

**Example 5-2: AIDS Testing**
The ELISA test for AIDS was widely used in the mid-1990’s for screening blood donations. As with most medical diagnostic tests, the ELISA test is not infallible. If a person actually carries the AIDS virus, experts estimate that this test gives a positive result 97.7% of the time. (This number is called the *sensitivity* of the test.) If a person does not carry the AIDS virus, ELISA gives a negative result 92.6% of the time (the *specificity* of the test). Estimates at the time were that 0.5% of the American public carried the AIDS virus (the *base rate* with the disease).

a) Suppose that a randomly selected person tests positive. Without doing any calculations, make a guess for the conditional probability, given this positive test result, that the person actually carries the AIDS virus.

b) Translate the given percentages into well-defined probabilities.

c) Express the probability of interest as a well-defined conditional probability.

d) Use Bayes’ Theorem to calculate the conditional probability of interest.

e) Is this conditional probability smaller than you guessed?

To gain insight into what’s happening here, imagine a hypothetical population of 1,000,000 people for whom these percentages hold exactly. Use the given percentages to fill in the following table, starting with the totals who do and do not carry the virus:

<table>
<thead>
<tr>
<th></th>
<th>Positive test</th>
<th>Negative test</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carries AIDS virus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does not carry AIDS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>1,000,000</td>
</tr>
</tbody>
</table>
f) From the filled-in table, determine the conditional probability that a randomly selected person carries the AIDS virus given that he/she tests positive. Does this agree with your answer to d)?

g) Use the table to explain why this probability turns out to be fairly small, compared to the sensitivity and specificity of the ELISA test.

h) Given that a randomly selected person tests negative on the ELISA test, how likely is it that he/she does not have the disease? Calculate an appropriate probability to address this question, both from the table and from the formula for Bayes’ Theorem.

Example 5-3: Bertrand’s Box Paradox
Suppose that three boxes have two coins each. Box 1 has two gold coins, box 2 has two silver coins, and box 3 has one of each. You first choose a box at random and then choose a coin at random from that box. Suppose the coin turns out to be gold. What is the probability that the other coin in the box is also gold?

a) What do you think most people guess this probability to be?

b) Define the relevant events in this example.

c) Use Bayes’ Theorem to determine the probability that the second coin is gold given that the first coin is gold.

d) Explain why it makes sense (in hindsight!) that this probability turns out differently from what most people expect. (Hint: You might want to fill in a hypothetical table with 3000 plays of this game to better understand what’s going on.)
Example 5-4: Forensic Evidence
Bayes’ Theorem was applied by expert witnesses testifying in a rape trial in Pittsburgh in the mid 1980’s. The defendant was accused of raping seven women in the Shadyside district of the city over a period from April 18, 1985, to January 30, 1986. By analyzing body secretion evidence taken from the scenes of the crimes, a forensic expert concluded that the assailant had the blood characteristics and genetic markers of type B, secretor, PGM 2+1-. She further testified that only 0.32% of the male population of Allegheny County had these blood characteristics and that the defendant himself was a type B, secretor, PGM 2+1-. The natural question to ask is how a juror should update his/her probability of the defendant’s guilt in light of this forensic evidence.

a) Let G represent the event that the defendant is guilty, and let E represent the forensic evidence that the criminal’s blood type was type B, secretor, PGM 2+1-. Let P(G) represent the prior probability that a juror assigns to the defendant’s guilt before hearing the forensic evidence. Rewrite Bayes’ Theorem in terms of G and E for expressing how to find the updated probability of guilt, conditional on the forensic evidence, from the prior probability of guilt.

b) What is P(E|G) in this situation?

c) What is P(E|Gc) in this situation? [Hint: Assume that if the defendant did not commit the crimes, then some other “random” male in Allegheny County did.]

d) Use your answers to the preceding questions to express the updated probability of guilt P(G|E) as a function of the prior probability of guilt P(G).

e) Construct a graph of P(G|E) as a function of P(G), for values of P(G) ranging from 0 to .5. [Hint: You could use R: > prior = (0:500)/100; updated = …] Comment on what the graph reveals.

f) Calculate the updated probabilities of guilt P(G|E) for the following prior probabilities P(G): .5, .1, .01, .001, and .00000278. [Hint: You could use R: > pr1 = c(.5, .1, .01, .001, .00000278); pr2 = …] Comment on what the calculations reveal.

The last entry in this list deserves special mention. The defense in this case argued that the prior probability of guilt should be 1 in 360,000, the estimated number of males in the appropriate age group in Allegheny County. The updated probability of guilt then becomes just 1 in 1150, the number of males with the same blood characteristics in the appropriate age group in Allegheny County.