ECE 315 - BASIC PROBABILITY - INVESTIGATION 4
BAYES' RULE FOR DISCRETE EVENTS

WINTER 2004 **A.P. FELZER**

To do "well" on this investigation you must not only get the right answers but must also do neat, complete and concise writeups that make obvious what each problem is, how you're solving the problem and what your answer is. You also need to include drawings of all circuits as well as appropriate graphs and tables.

In the last two Investigations we defined and worked with conditional probabilities. The main objective of this Investigation is to continue this work with the development of Bayes' Rule for finding conditional probabilities in situations where we know the outcome of a test - like the result of a blood test - and we want to know the probability that we actually have a disease.

1. The objective of this first problem is to introduce the basic idea of **Bayes' Rule** with the following observation

$$\text{P(Have Disease | Test Positive)} = \frac{\text{P(Have Disease and Test Positive)}}{\text{P(Test Positive)}}$$

$$= \frac{\text{P(Test Positive and Have Disease)}}{\text{P(Test Positive)}}$$

$$= \frac{\text{P(Test Positive | Have Disease) P(Have Disease)}}{\text{P(Test Positive)}}$$

   From this result we see that we can find the probability of having a disease when we test positive if we have sufficient information about the disease and the test to know the probabilities in the final expression.

   Suppose, in particular, that one out of a thousand people has a particular disease. And suppose there is a test for this disease that will be positive if you have the disease but is also positive for 5% of the people who take the test but don't have the disease. What's the probability you have the disease if you test positive

2. The objective of this problem is to apply Bayes' Rule in Problem (1) to a communication channel with outcomes as follows

   T_0 = 0 is transmitted $\qquad$ R_0 = 0 is received
   T_1= 1 is transmitted $\qquad$ R_1 = 1 is received

   that has the following probabilities

   $P(T_0) = 0.6 \qquad P(T_1) = 0.4 \qquad$ P(Error at the Receiver) = 0.1

   Note that an error occurs whenever the received signal is different from the transmitted signal

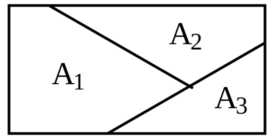
   a. Find $P(T_0 | R_0)$
   b. Find $P(T_1 | R_0)$

3. The objective of this and the next five problems is to develop a general equation for Bayes' Rule. We begin with partitions. A set of events $A_1, A_2, \ldots, A_n$ of a sample space S form a **partition** if they satisfy the following conditions

(1) They're mutually exclusive

(2) $A_1 \quad A_2 \quad \cdot \cdot \cdot \quad A_n = S$

a. Explain why A_1, A_2 and A_3 in the following Venn diagram form a partition



b. Draw a Venn diagram of a partition of four events

c. Verify that the events A_1 = {At least one head} and A_2 = {No heads} are a partition of the sample space when we flip a coin twice

d. Come up with another partition for the sample space when we flip a coin twice

e. Come up with a partition of three events A_1, A_2 and A_3 for the sample space when we flip a coin four times

4. Suppose we flip a coin twice and form partitions from the outcomes like in Problem (3)

   a. What is $P = P(A_1) + P(A_2)$ if the partition consists of the two events

$$A_1 = \{\text{At least one head}\} \quad \text{and} \quad A_2 = \{\text{No heads}\}$$

   b. What would $P = P(A_1) + P(A_2) + P(A_3)$ equal if we had formed a partition of three events A_1, A_2, A_3

   c. What would $P = P(A_1) + \cdot \cdot \cdot + P(A_n)$ equal if we had formed a partition of n events $A_1, A_2, \ldots, A_n$

5. Given a partition $A_1, A_2, \ldots, A_n$ of a sample space S

   a. Draw a Venn diagram to illustrate the fact that

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \cdot \cdot \cdot + P(B \cap A_n)$$

   b. Verify that this relation is true for an example that you make up

6. Given the following result from Problem (5)

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_n)$$

for partitions $A_1, A_2, \ldots, A_n$

   a. Show that

$$P(B) = P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \cdots + P(B \mid A_n)P(A_n) = \sum_{k=1}^{n} P(B \mid A_k)P(A_k)$$

   b. We call the result in part (a) the **theorem on total probability**. Illustrate the theorem on total probability for $n = 3$

7. Make use of the theorem on total probability to show that if $A_1, A_2, \ldots, A_n$ form a partition of a sample space S then

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B \cap A_j)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{k=1}^{n} P(B|A_k)P(A_k)}$$

We call this result **Bayes' Rule**.

8. Write out the equation for Bayes' Rule for $P(A_1 | B)$ for a partiton consisting of A_1 and A_2

9. The objective of this problem is to apply our general expression for Bayes' Rule to the disease problem of Problem (1) above. Suppose one out of a thousand people has a particular disease. And suppose there is a test for this disease that will be positive if you have the disease but is also positive for 5% of the people who take the test but don't have the disease.
   a. Define a partition on the sample space
   b. What's the probability you have the disease if you test positive

Note that P(Have Disease) is called an **a priori probability**. It's our probability of having the disease before we know anything about the test result. P(Have Disease | Test Positive) , on the other hand, is referred to as an **a posteriori probability** because it depends on the test result. **Memorize** these terms.

10. The objective of this problem is to apply our general Bayes' Rule expression to the communications problem in Problem (2) above with a communication channel with outcomes

T_0 = 0 is transmitted $\qquad$ R_0 = 0 is received
T_1= 1 is transmitted $\qquad$ R_1 = 1 is received

and probabilities

$P(T_0) = 0.6 \qquad P(T_1) = 0.4 \qquad$ P(Error at the Receiver) = 0.1

where an error occurs whenever the received signal is different from the signal that was sent
   a. Verify that T_0 and T_1 form a paritition
   b. Make use of Bayes' Rule to find $P(T_0 | R_0)$
   c. Make use of Bayes' Rule to find $P(T_1 | R_0)$
   d. What are the a priori and a posteriori probabilities that a zero was transmitted

11. Now for the "famous" Monty Hall problem. A popular TV game show had three closed boxes - two of which were empty and one of which had $10,000. The contestant got to choose one of the three boxes. But before the chosen one was opened, Monty Hall would always open one of the two remaining boxes and show that it was empty and then give the contestant the opportunity to stick with the box he chose originally or change to the other remaining unopened box. What should he do. Hint - make use of the theorem on total probability as follows

$$P(B) = P(B | A_1)P(A_1) + P(B | A_2)P(A_2) + \cdots + P(B | A_n)P(A_n)$$

with B = Winning after switching
A_1 = Initially choosing an empty box
A_2 = Initially choosing the box with the prize

You may also find it useful to do a coin flipping simulation