MATHEMATICS 191, FALL 2004
MATHEMATICAL PROBABILITY
Outline #3 (Conditional Probability)

Last modified: October 5, 2004
References:

- PRP, sections 1.4, 1.5 and 1.7
- EP, Chapter 2

1. Define conditional probability (PRP, p.9) and show examples of how to calculate it.

2. (PRP, section 1.4, EP, section 2.1) State and prove Lemma 4 on p. 10 of PRP,
$$\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)$$
Use this lemma and the formula for conditional probability to analyze the "bearded man" problem in the notes, the "lemons problem" on p. 53 of EP, or a similar example of your own invention.

3. Use conditional probability to analyze example 3 on page 9 of PRP. This is equivalent to solution I to the problem discussed in section 2.13, part A of EP.

4. (EP, Section 2.13, part B) Describe the "Monty Hall problem" and analyze it in terms of conditional probability. This is equivalent to example 7 on page 9 of PRP.

5. Analyze a "division into sets" problem in terms of conditional probability. Here is one that is based on the example from outline 2.

   Lisa purchases six Dunkin Munchkins, four plain and two chocolate. She chooses three at random and puts them in a bag for her son Thomas. The other three go into a bag for her daughter Catherine.

   Catherine reaches into her bag and extracts a Munchkin at random. It is a plain one (event B). Show that the conditional probability, given event B, that Thomas has both chocolate Munchkins, precisely one chocolate Munchkin, or no chocolate Munchkins are $P(A_2) = 0.3, P(A_1) = 0.6$, and $P(A_0) = 0.1$ respectively.

6. Define independent events (PRP, p. 13, EP, section 2.2) and illustrate the concept using sets.exe.

7. Give an example of three events that are independent in pairs but not independent.

   The following example is based on one presented in Math S-152 by Harvard physics concentrator Marc Parris.

   George Steinbrenner, owner of the Yankees, wants to strengthen his pitching staff for the World Series. He secretly negotiates deals with Roger Clemens, Randy Johnson, and Barry Zito and takes them to the commissioner for approval. The commissioner's office looks at the deals and announces that the commissioner will make his decision within 24 hours by rolling a fair tetrahedral die.

   - if the die comes up 1, the Yankees get Clemens only.
   - if the die comes up 2, the Yankees get Johnson only.
   - if the die comes up 3, the Yankees get Zito only.
   - if the die comes up 4, the Yankees get Clemens, Johnson, and Zito.

   Let C, J, and Z be the events the the Yankees get Clemens, Johnson, and Zito respectively. Show that these events are independent in pairs, e.g. $P(Z|J) = P(Z)$ but that they are not independent, because

   $$P(Z|J \cap C) \neq P(Z).$$

8. Explain how to create independent events by means of a "compound experiment," for example a die roll followed by a coin flip. Use sets.exe to illustrate events that are independent whenever "Random 4×4" is used.

9. Give an example of Simpson's paradox and state the paradox in general terms (PRP, p. 19).

10. Describe and explain "Eddington's controversy (EP page 76), but do not address the issue of why Eddington thought it was controversial.

11. As an example of a generalized Monty Hall problem, solve the following "Christine's random dessert" problem.

Shea and Marty walk into Christine's Random Desserts, Harvard Square's trendiest after-dinner spot, and ask for a table for two. They are escorted to a round table with five places, in front of each of which is a metal dome with a handle, of the sort associated with upscale French restaurants. They choose two of the five places at random and sit down.

Christine herself arrives to take their order. "We only need two places, really," says Marty. Christine then explains that under two of the domes are complimentary servings of her famous cookies, while under the other three are tiny vegetarian snacks. "Let me show you a randomly-chosen one at an empty place," she says, and lifts a dome to reveal a scrawny carrot.

"Would either or both of you like to change your seat?" asks Christine.

(a) What is the probability, when Shea and Marty first sit down, that they both have cookies in front of them? That neither has cookies?

(b) If Shea and Marty both take a new seat, what is the probability that they now both have cookies in front of them? That neither has cookies?

(c) If Shea takes a new seat but Marty does not, what is the probability that they now both have cookies in front of them? That neither has cookies?

It would be tedious to enumerate what happens for each of the $5! = 120$ possible arrangements of the cookies and carrots. Here is a shorter but equivalent analysis.

Events A_2, A_1, A_0 are "Shea and Marty have 2, 1, or 0 cookies respectively."

This is a Munchkin problem with two unequal bags. There are $\binom{5}{2} = 10$ ways to choose two items. There is 1 way to get 2 cookies, there are $3 \cdot 2 = 6$ ways to get 1 carrot and 1 cookie, and $\binom{5}{2} = 3$ ways to get 0 cookies and 2 carrots.

So

$$\mathbb{P}(A_2) = 0.1, \mathbb{P}(A_1) = 0.6, \mathbb{P}(A_0) = 0.3.$$

Assume that Shea and Marty are at places 1 and 2. Event C is that Christine reveals a carrot at place 3. This place was chosen at random from among the subset of places 3, 4, and 5 that have carrots, but any arrangement of carrots and cookies at these places is equally probable. Because places 3, 4, and 5 are equivalent, C is independent of A_2, A_1 and A_0.

If they have 1 cookie, there is 1 cookie at places 4 and 5 and if they both switch it changes nothing.

If they have 2 cookie, there are 0 cookies at places 4 and 5 and switching loses cookies.

If they have 0 cookies, there are 2 cookies at places 4 and 5 and switching gains cookies.

So if they both switch, the probabilities for 0 and 2 cookies are simply interchanged.

Events B_2, B_1, B_0 are "Shea and Marty have 2, 1, or 0 cookies respectively after just one of them switches."

If they start with 1 cookie and make 1 switch, the probability that the one with no cookie switches to the empty place with the cookie is $\frac{1}{4}$. The probability for ending up with no cookies is likewise $\frac{1}{4}$, and so the probability of still having 1 cookie is $\frac{1}{2}$.

$$\mathbb{P}(B_2) = \mathbb{P}(B_2|A_2)\mathbb{P}(A_2) + \mathbb{P}(B_2|A_1)\mathbb{P}(A_1) + \mathbb{P}(B_2|A_0)\mathbb{P}(A_0)$$

$$\mathbb{P}(B_2) = 0 + \frac{1}{4} \cdot 0.6 + 0 = 0.15$$

$$\mathbb{P}(B_1) = \mathbb{P}(B_1|A_2)\mathbb{P}(A_2) + \mathbb{P}(B_1|A_1)\mathbb{P}(A_1) + \mathbb{P}(B_1|A_0)\mathbb{P}(A_0)$$

$$\mathbb{P}(B_1) = 1 \cdot 0.1 + \frac{1}{2} \cdot 0.6 + 1 \cdot 0.3 = 0.7$$

$$\mathbb{P}(B_0) = \mathbb{P}(B_0|A_2)\mathbb{P}(A_2) + \mathbb{P}(B_0|A_1)\mathbb{P}(A_1) + \mathbb{P}(B_0|A_0)\mathbb{P}(A_0)$$

$$\mathbb{P}(B_1) = 0 + \frac{1}{4} \cdot 0.6 + 0 = 0.15$$

Interestingly, this is a higher probability for 1 cookie, and a correspondingly smaller probability for 0 or 2 cookies, than if they had originally chosen from among 4 places, 2 with cookies and 2 with carrots.

Notes:

- Conditional probability: Suppose that A and B are events (subsets) of a sample space Ω. If we know that event B has occurred, what is then the probability that A has also occurred?

  Example: I roll two dice and note that the sum is six. What is the probability that at least one 2 is showing?
  Answer: there are 5 pairs in set B: (1,5), (2,4), (3,3), (4,2), and (5,1)
  Two of them belong to set A so the probability is $\frac{2}{5}$

  at least one 3 is showing?
  Answer: there are 5 pairs in set B: (1,5), (2,4), (3,3), (4,2), and (5,1)
  Only one of them belongs to set A so the probability is $\frac{1}{5}$

  Definition of conditional probability: Suppose that $P(B) > 0$.
  Then $P(A|B)$ "the probability of A, given B," is

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

  Many simple problems in conditional probability can be analyzed by arranging the probabilities into a rectangular array where the rows and columns correspond to mutually exclusive events and each cell gives the probability for the intersection of the "row" and "column" events.

- Example: the "bearded man" problem:

  At an airport with severe terrorism problems, security personnel have established the following:
  Of male passengers with explosives in their shoes, 60% have beards.
  Of male passengers with no explosives in their shoes, 5% have beards.
  20% of male passengers have explosives in their shoes.
  What is the probability that a bearded male passenger has explosives in his shoes?

  Event A is "explosive shoes" while event A^c is "non-explosive shoes"
  Event B is "bearded" while event B^c is "clean-shaven"

  From the given information we can make a table

	B	B^c
A	.12	.08
A^c	.04	.76
Sum	.16	84

  Calculate $P(B)$ by using the lemma on page 10 of PRP.

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c) = .05 \times .8 + .60 \times .2 = .16$$

  So $P(A \cap B) = .12, P(B) = .16$, and

$$P(A|B) = P(A \cap B)/P(B) = .12/.16 = 3/4.$$

- The most famous probability problem (the Monty Hall problem)

  In a computer game show, there are three doors. Behind one door, chosen at random, is a car C. Behind each of the other two doors is a goat. There is a small goat $G1$ and a large goat $G2$. You select a door. Before revealing whether you have won the car, host Monty Hall, who knows where the car is located, opens a door that you have not chosen and that does not have the car behind it, revealing a goat. If he has a choice, he reveals the smaller goat. He then asks if you would like to switch your choice to the remaining unopened door.

  Suppose (without loss of generality) you pick door 1. There are six ways to arrange the car and the two goats behind the three doors

  Here are the six 6 equally likely possibilities, along with Monty Hall's action in each case.

Case	Door 1	Door 2	Door 3	Monty opens
1	C	G1	G2	door 2
2	C	G2	G1	door 3
3	G1	C	G2	door 3
4	G2	C	G1	door 3
5	G1	G2	C	door 2
6	G2	G1	C	door 2

  Event A is "the car is behind door 1"
  Event B is "Monty Hall opened door 2" So $P(A \cap B) = 1/6$ (case 1 only)
  $P(B) = 1/2$ (cases 1, 5, and 6)

  So $P(A|B) = P(A \cap B)/P(B) = 1/3, P(C|B) = 2/3$, and you double your chances of winning by switching to door 3!

  A subtle variant: Monty flips a coin to choose a door (not yours) at random and open it. This means that you may discover that you have lost. Now the 6 possibilities are

  Car behind 1, heads, opens 2
  Car behind 1, tails, opens 3
  Car behind 2, heads, opens 2
  Car behind 2, tails, opens 3
  Car behind 3, heads, opens 2
  Car behind 3, tails, opens 3

  Suppose that event B is "he opens 2, and there is no car behind it."

  So $P(A \cap B)$ is 1/6 from Car behind 1, heads, opens 2

  $P(B)$ is 2/6 from
  Car behind 1, heads, opens 2
  Car behind 3, heads, opens 2

Now $P(A|B)$ is 1/2 and there is no advantage to switching.

- Independence

  Two events A and B are called independent if

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

  Alternative view: In this case

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

  so that the conditional probability of A is unaffected by knowledge of B.

  Simple example:
  A = "Drawing a heart"
  B = "Drawing an ace"
  Independent - easy argument on PRP p. 13.

  A = "Drawing two hearts"
  B = "Drawing two aces"
  Not independent since $P(A \cap B) = 0, P(A) > 0$

  Slightly more complicated:
  A = two fair dice are rolled and the numbers are the same.
  $P(A) = 1/6$
  B = "sum of the numbers on the 2 dice is 8"

  There are 5 possibilities: (2,6) (3,5) (4,4) (5,3) (6,2) So $P(A \cap B) = 1/36$ and $P(B) = 5/36$. Thus $P(A|B) = 1/5$ and the events are not independent.

  For three events to be independent it is required that $P(A \cap B \cap C) = P(A)P(B)P(C)$

- Compound Experiments
  A compound experiment might consist of a die roll followed by a coin flip. In such a case we can construct independent events in a very general manner.

  We assume that we know probabilities $P_1(x)$ for the first experiment (the die roll) even if the die is loaded. We assume that we know probabilities $P_2(y)$ for the second experiment (the coin flip) even if the coin is not true.

  The outcome of the compound experiment is a pair (x, y) e.g. (4, heads) It is reasonable to define $P(x, y) = P_1(x)P_2(y)$, e.g. 1/12 for each outcome if the die and coin are both fair. With this formula, $P(x, y)$ is positive, and the sum over all possible outcomes is 1.

  Now we have to show that the condition for independence is satisfied:

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

This is tricky, since events A, B, and $A \cap B$ all have to be subsets of the same set.

Event A: result of first experiment is in C_1, result of experiment 2 is anything. $P(A) = P_1(C_1)$

Event B: result of second experiment is in C_2, result of experiment 1 is anything. $P(B) = P_2(C_2)$

Event $A \cap B$: result of first experiment is in C_1, result of second experiment is in C_2.

Then from the definition $P(x, y) = P_1(x)P_2(y)$ it follows that $P(A \cap B) = P(A)P(B)$.

An illustration is in sets.exe, the "Random 4 x 4" probabilities. It works this way:
Generate 4 random numbers to get $P_1(q)$ for the quotient q. (values 0 - 3)
Generate 4 more random numbers to get $P_2(r)$ for the quotient r. (values 0 - 3)

Assign to each number $0 \leq x \leq 15$
$P(x) = P_1(x/4)P_2(x \mod 4)$ Then, no matter what the random numbers, a "remainder event" like $x \mod 4 = 1$ is independent of a "quotient event" like $x \leq 7$.

- Simpson's paradox

Simpson's paradox requires an event A to be conditioned on a pair of events B and C. The "paradox" can arise only if events B and C are not independent.

Event A is the "success" event. In the classic version, a drug cures a disease. When A occurs, people think that it is good.

Event B is the "my product" event, while B^c is the "competitor's product" event.

Event C is the "test group" event. Generally both products are good for one test group, bad for the other.

Event A is conditioned on the pair of events B and C, but the public are willing to view A as conditioned on B only. Here is a concrete example of how to exploit the paradox to make an inferior product look good.

The experiment is "Choose a clinical trial at random. The subject is a man (or woman) who uses my product (or my competitor's) and is or is not cured." There are 8 possible outcomes, and their probabilities sum to 1.

Let's build things up from scratch.

My product is worthless for men, good for women. My competitor's product is bad but not worthless for men and always works for women. To be precise, if A is "cured," B is "used my product," and C is "male"

- $\mathbb{P}(A|(B \cap C)) = 0$.
- $\mathbb{P}(A|(B \cap C^c)) = 3/4$.
- $\mathbb{P}(A|(B^c \cap C)) = 1/10$.
- $\mathbb{P}(A|(B^c \cap C^c)) = 1$.

I hire a testing firm to compare the products. They put my product into a face cream and advertise it on Home and Garden TV. They put the competing product into a pill that must be taken with beer and advertise in at halftime on football games. The result, not surprisingly, is that the people who test my product are women, while those who test my competitor's are men. To be precise,

- $\mathbb{P}(C|B) = \frac{1}{9}$. (people who test my product are rarely men)
- $\mathbb{P}(C|B^c) = \frac{10}{11}$. (people who test the competing product are usually men)
- $\mathbb{P}(C^c|B) = \frac{8}{9}$. (people who test my product are usually women)
- $\mathbb{P}(C^c|B^c) = \frac{1}{11}$. (people who test the competing product are rarely women)

Since both products work much better on women, this is unfair, but that's what marketing is all about.

We want roughly the same number of people trying each product, so it is reasonable to have $\mathbb{P}(B) = 0.45$.

Now, of course,

- $\mathbb{P}(C \cap B) = \mathbb{P}(C|B)\mathbb{P}(B) = \frac{1}{9}(0.45) = 0.05$.
- $\mathbb{P}(C^c \cap B) = \mathbb{P}(C^c|B)\mathbb{P}(B) = \frac{8}{9}(0.45) = 0.4$.
- $\mathbb{P}(C \cap B^c) = \mathbb{P}(C|B^c)\mathbb{P}(B^c) = \frac{10}{11}(0.55) = 0.5$.
- $\mathbb{P}(C^c \cap B^c) = \mathbb{P}(C^c|B^c)\mathbb{P}(B^c) = \frac{1}{11}(0.5) = 0.05$.

The probabilities for these disjoint events sum to 1.

Now it is easy to calculate the probability of a cure in each case by using the conditional probabilities and present the result in a table.

	Men(C) Mine(B)	Men(C) His(B^c)	Women(C^c) Mine(B)	Women(C^c) His(B^c)
Cured (A)	0	.05	.3	.05
Not cured (A^c)	.05	.45	.1	0

On looking at this table, it is apparent that either for men or for women, the competitor's product works better.

No problem: just ignore event C and present the data for the four cases involving just events A and B.

	Mine(B)	His(B^c)
Cured (A)	.3	.1
Not cured (A^c)	.15	.45

Now it is clear that my product worked $\frac{2}{3}$ of the time, while the competing product worked $\frac{2}{11}$ of the time. It is true, but misleading, that

$\mathbb{P}(A|B) = \frac{2}{3}$ while $\mathbb{P}(A|B^c) = \frac{2}{11}$. This is the supposed "paradox."

What happened? I rigged the clinical trials so that women (for whom both products work) would try my product while men (for whom neither product works) would try my competitor's.

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

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B \cap C) + \mathbb{P}(A \cap B \cap C^c)}{\mathbb{P}(B)}$$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A|(B \cap C))\mathbb{P}(B \cap C) + \mathbb{P}(A|(B \cap C^c))\mathbb{P}(B \cap C^c)}{\mathbb{P}(B)}$$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A|(B \cap C))\mathbb{P}(C|B)\mathbb{P}(B)}{\mathbb{P}(B)} + \frac{\mathbb{P}(A|(B \cap C))\mathbb{P}(C^c|B)\mathbb{P}(B)}{\mathbb{P}(B)}$$

$$\mathbb{P}(A|B) = \mathbb{P}(A|(B \cap C))\mathbb{P}(C|B) + \mathbb{P}(A|(B \cap C^c))\mathbb{P}(C^c|B)$$

Checking the numbers:

For my product,

$$\frac{2}{3} = 0 \cdot \frac{1}{9} + \frac{3}{4} \cdot \frac{8}{9}$$

For the competing product,

$$\frac{2}{11} = \frac{1}{10} \cdot \frac{10}{11} + 1 \cdot \frac{1}{11}$$