Chapter 7

Probability

7.1 Experiments, Sample Spaces and Events

Start with some definitions we will need in our study of probability.

An EXPERIMENT is an activity with an observable result. Tossing coins, rolling dice and choosing cards are all probability experiments.

The result of the experiment is called the OUTCOME or SAMPLE POINT. So the two possible outcomes from tossing a coin are H (heads) and T (tails).

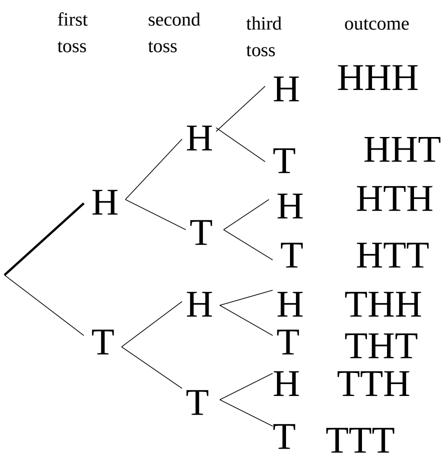
The set of all outcomes or sample points is called the SAMPLE SPACE of the experiment.

An EVENT is a subset of a sample space. That is, an event can contain one or more outcomes that are in the sample space.

Consider tossing a coin. The sample space is $S = \{H, T\}$. The events that are possible in this experiment are $\emptyset, \{H\}, \{T\}, S$. So, while there are 2 outcomes in the sample space, there are 4 different events.

If a 6-sided die is rolled, the sample space is $S = \{1, 2, 3, 4, 5, 6\}$.

Sometimes we use a tree diagram to find all the possible outcomes of an experiment. Consider tossing a coin 3 times and noting the result of each toss.



So we find there are 8 elements in the sample space. We would show this as

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

[0] ©Janice Epstein 1998, 1999, 2000. These notes may not be distributed for profit

Find the event E where $E = \{x|x \text{ has exactly one head}\}$ $\quad E = \{HTT, THT, TTH\}$

Find the event E where $E = \{x|x \text{ has two or more heads}\}$ $\quad E = \{HHT, HTH, THH, HHH\}$

Find the event E where $E = \{x|x \text{ has more than 3 heads}\}$ $\quad E = \emptyset$

Another common sample space you must be familiar with is that from rolling two dice. We put this in a chart. You can think of the dice as being two different colors to avoid missing any outcomes.

1-1	2-1	3-1	4-1	5-1	6-1
1-2	2-2	3-2	4-2	5-2	6-2
1-3	2-3	3-3	4-3	5-3	6-3
1-4	2-4	3-4	4-4	5-4	6-4
1-5	2-5	3-5	4-5	5-5	6-5
1-6	2-6	3-6	4-6	5-6	6-6

These sample spaces have all been finite. That is, we can list all the elements. An infinite sample space has to be described, you can't list all the elements:

What is the sample space for the time spent working on a homework set?

$S = \{t|t \geq 0, t\ in\ minutes\}$

Describe the event of spending between one and two hours on a homework set.

$E = \{t|60 < t < 120\}$

7.2 Definition of Probability

When we toss a fair coin the two outcomes in the sample space $S = \{H, T\}$ are equally likely, so the probability of each outcome is 1/2. We would write this as $P(\{H\}) = 1/2$ and $P(\{T\}) = 1/2$. This is a THEORETICAL PROBABILITY based on the sample space having equally likely outcomes. In general, this is the way we will find probability, by using a sample space of EQUALLY LIKELY OUTCOMES. The probability of an event, $P(E)$ is a number between 0 and 1, $0 \leq P(E) \leq 1$. An event with a probability of 0 is IMPOSSIBLE and an event with a probability of 1 is CERTAIN. The closer $P(E)$ is to 1, the more likely the event is to happen.

We can also calculate the EMPIRICAL PROBABILITY of an event by doing an experiment many times. For example, you could toss a coin and note how many times it comes up heads (shown in book) or you could roll a die and count how many times a 1 is rolled.

number of tosses (m)	number of 1's rolled (n)	relative frequency (n/m)
20	3	3/20 = .15
100	18	18/100 = .18
1000	160	160/1000 = .16
10000	1662	1662/10000 = .1662

If the die is fair, the theoretical probability is 1/6 = .16666666..... We would expect that if we did the experiment enough times we would approach the theoretical value rather closely if the die were fair. In fact, the empirical probability that an event occurs is the relative frequency that the event occurs as the number of experiments get very large ($n \to \infty$). So, if the die (or coin) was not fair, we could still find the probability of each outcome by doing lots of experiments. The balls used in the Lotto game are checked this way - they do many many draws and make sure each ball has a 1/50 chance of being drawn.

A third way of "calculating" probability is called SUBJECTIVE PROBABILITY. That is when an expert estimates the probability of something happening based on their opinion. Betting lines, economic predictions and weather forcasting are all based on this kind of probability.

When finding probability theoretically we will always start by finding the sample space. You must be certain that it is a UNIFORM SAMPLE SPACE - each of the outcomes are equally likely. We write $S = \{s_1, s_2, ..., s_n\}$, a sample space with n equally likely outcome.

The events $\{s_1\}$, $\{s_2\}$ ... $\{s_n\}$ are called SIMPLE events because they consist of exactly one outcome. Notice these simple events are

MUTUALLY EXCLUSIVE as only one can occur. The probability of each simple event occuring is the same,

$P(s_1) = P(s_2) = ...P(s_n) = 1/n$

We can put the probability of each event in a table called a PROBABILITY DISTRIBUTION TABLE:

outcome	probability
$\{s_1\}$	$P(s_1) = 1/n = P_1$
$\{s_2\}$	$P(s_2) = 1/n = P_2$
$\vdots$	
$\{s_n\}$	$P(s_n) = 1/n = P_n$

Notice the following properties of probability distributions:

1. $0 \leq P(s_i) \leq 1$
2. $P_1 + P_2 + ...P_n = 1$
3. $P(\{s_i\} \cup \{s_j\}) = P_i + P_j, i \neq j$

So we can go back to our sample spaces from section 7.1 and write the probability distribution table:

Toss a coin three times. There are 8 equally likely outcomes:

simple event	Probability
$\{HHH\}$	1/8
$\{HHT\}$	1/8
$\{HTH\}$	1/8
$\{HTT\}$	1/8
$\{THH\}$	1/8
$\{THT\}$	1/8
$\{TTH\}$	1/8
$\{TTT\}$	1/8

Notice each event is mutually exclusive. this must be true of the list in your probability distribution table or it won't work.

We can also use our relative frequency interpretation of empirical probabilities for doing an experiment and assigning probability.

Suppose the instructor of a class polled the students about the number of hours spent per week studying math during the previous week. The results were

time studying (hours)	number of students	Probability (relative frequency)
$0 \leq x \leq 2$	69	69/309 = .2233
$2 < x \leq 4$	128	128/309 = .4142
$4 < x \leq 6$	68	68/309 = .2201
$6 < x \leq 8$	30	30/309 = .0971
$x > 8$	14	14/309 = .0453
total	309	1

Since the catagories are mutually exclusive, we can find the probability that a student studies more than 4 hours per week as

$P(x > 4) = P(4 < x \leq 6) + P(6 < x \leq 8) + P(x > 8)$

$= .2201 + .0971 + .0453 = .3625$

What is the probability of rolling a sum 2 or a sum of 12 using two fair die?

Remember the sample space has 36 outcomes. Look at the sample space and find that the sum of 2 happens once and the sum of 12 happens once and they are mutually exclusive.

$P = 1/36 + 1/36 = 2/36 = 1/18$

What is the probability of rolling a sum of 7? Look at the sample space:

$P = 1/36 + 1/36 + 1/36 + 1/36 + 1/36 + 1/36 = 6/36 = 1/6$

7.3 Rules of Probability

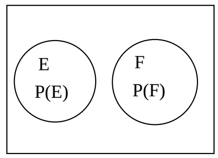
Consider a sample space S with events E and F. So far we have learned the following:

$P(E) \geq 0$ for any E

$P(S) = 1$

If E and F are mutually exclusive (that is, one one of them can occur, $E \cap F = \emptyset$) then

$P(E \cup F) = P(E) + P(F)$



In general, E and F have some outcomes in common so we have the union rule for probability: $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

Consider our two dice. Let

$E = \{x | x$ is a sum of 7 $\}$

$F = \{x | x$ is a 6 on the green die $\}$

Find these events in our sample space:

$E = \{1-6, 2-5, 3-4, 4-3, 5-2, 6-1\}$

$F = \{6-1, 6-2, 6-3, 6-4, 6-5, 6-6\}$

What is the probability that you have a sum of 7 OR a 6 on the green die?

We can find that $P(E) = 6/36$ and $P(F) = 6/36$ but what is $P(E \cap F)$? That is the probability that we have a sum of 7 AND a 6 on the green die. That is going to be the outcome that the two events have in common. The two conditions are met with the roll $6-1$, so $P(E \cap F) = 1/36$,

$P(E \cup F) = P(E) + P(F) - P(E \cap F) = 6/36 + 6/36 - 1/36 = 11/36$

an alternate way to do this with a dice problem is to write out the sample space and count how many outcomes fulfill the two requirements.

Example - If a single card is drawn from a standard deck of cards, what are the probabilities of

a) a 9 or a 10? Let

$E = \{9\heartsuit, 9\diamondsuit, 9\clubsuit, 9\spadesuit\}, P(E) = 4/52$

$F = \{10\heartsuit, 10\diamondsuit, 10\clubsuit, 10\spadesuit\}, P(F) = 4/52$

$E \cap F = \emptyset$, so $P(E \cap F) = 0$

$P(E \cup F) = 4/52 + 4/52 - 0/52 = 8/52 = 2/13$

b) a black card or a 3?

$E = \{x | x$ is a black card $\}, n(E) = 26$

$F = \{3\heartsuit, 3\spadesuit, 3\diamondsuit, 3\clubsuit\}, n(F) = 4$

$E \cap F = \{3\spadesuit, 3\clubsuit\}, n(E \cap F) = 2.$

$P(E \cup F) = \frac{26}{52} + \frac{4}{52} - \frac{2}{52} = \frac{28}{52}$

Sometimes you need to use a Venn diagram to see what the probability is:

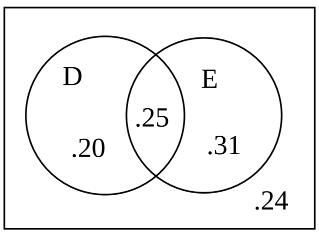
Example - a survey gave the following results: 45% of the people surveyed drank diet drinks (D) and 25% drank diet drinks and exercised ($D \cap E$) and 24% did not exercise and did not drink diet drinks ($D^c \cap E^c$). Find the probability that:

a) a person does not drink diet drinks (D^c).

b) does not exercise and drinks diet drinks ($E^c \cap D$).

c) excercises and does not drink diet drinks ($E \cap D^c$).

Draw the diagram. Fill in the numbers. We must have a total probability of 1 in the whole diagram, so we can find the last number by subtraction from 1.



So we find that 55% do not drink diet drinks
20% do not exercise and drink diet drinks
and 31% exercise and do not drink diet drinks.

7.4 Use of Counting Techniques in Probability

We have casually be using the following theoretical technique for find the probability of an event E:

Let S be a uniform sample space and E be any event in S. Then

$$P(E) = \frac{number\ of\ outcomes\ in\ E}{number\ of\ outcomes\ in\ S} = \frac{n(E)}{n(S)}$$

In this section the sample spaces and events will be larger, so we will use our counting techniques from chapter 6 to find the number of outcomes.

Example - suppose we have a jar with 8 blue and 6 green marbles. What is the probability that in a sample of 2, both will be blue?

$E = \{x|x$ is two blue marbles $\}$, $n(E) = \binom{8}{2} \cdot \binom{6}{0} = 28$

$S = \{x|x$ is a sample of two marbles $\}$, $n(S) = \binom{14}{2} = 91$

$P(E) = \frac{n(E)}{n(S)} = \frac{28}{91}$

You do not need to write out all that work, if you wish you can just do

$$P(2\ blues) = \frac{\binom{8}{2} \cdot \binom{6}{0}}{\binom{14}{2}} = \frac{28}{91}$$

You should show at least that much work for partial credit.

What is the probability there is at least one blue marble?

We can have $2B0G + 1B1G$:

$$P(\geq 1Blue) = \frac{\binom{8}{2}\cdot\binom{6}{0}+\binom{8}{1}\cdot\binom{6}{1}}{\binom{14}{2}} = \frac{28+48}{91} = \frac{76}{91}$$

Example - a stack of 100 copies has 3 defective papers. What is the probability that in a sample of 10 there will be no defective papers?

The number of ways to choose a sample of 10 from 100 is

$$n(S) = \binom{100}{10}$$

The number of ways to have no defective papers in the sample is

$$n(E) = \binom{3}{0}\cdot\binom{97}{10}$$

So we find the probability is

$$P(E) = \frac{n(E)}{n(S)} \approx .73$$

Example - A student takes a true/false test with 5 questions by guessing (choose answer at random). Write a probability distribution table for the number of correct answers.

Start by finding the sample space. Try a tree diagram to see how many ways we can answer the questions. See that we need the multiplication principle to find

$$n(S) = 2\cdot 2\cdot 2\cdot 2\cdot 2 = 2^5 = 32$$

The possible results are $0W5R, 1W4R, 2W3R, 3W2R, 4W1R, 5W0R$. We need to see how many outcomes are in each event.

There is only one way to get all of them wrong, or do

$$n(5W0R) = \binom{5}{0} = 1$$

So $P(5W0R) = 1/32$

To get four wrong and one right we can get any one of the 5 questions for our correct one, or

$$n(4W1R) = \binom{5}{1} = 5$$

So $P(4W1R) = 5/32$

To get three wrong and two right we can get choose any two of the 5 questions for our correct ones, or

$$n(3W2R) = \binom{5}{2} = 10$$

So $P(3W2R) = 10/32$

To get two wrong and three right we can get choose any 3 of the 5 questions for our correct ones, or

$$n(2W3R) = \binom{5}{3} = 10$$

So $P(2W3R) = 10/32$

To get one wrong and four right we can get choose any 4 of the 5 questions for our correct ones, or

$$n(1W4R) = \binom{5}{4} = 5$$

So $P(1W4R) = 5/32$

To get zero wrong and five right we must choose all 5 of the 5 questions for our correct ones, or

$$n(0W5R) = \binom{5}{5} = 1$$

So $P(0W5R) = 1/32$

Put this into the table:

number wrong	probability
0	1/32
1	5/32
2	10/32
3	10/32
4	5/32
5	1/32
total	1

You should be able to do these probability distribution tables for things like the number of blue jelly beans in the sample or the number of hearts in the hand of cards and so on.

7.5 Conditional Probability and Independent Events

What is conditional probability? It is where you know some information but not enough to get a complete answer. For example,

A survey is done of people making purchases at a gas station. Most people buy gas or a drink. The results are shown in the table:

	buy drink (B)	no drink (B^c)	total
buy gas (A)	20	15	35
no gas (A^c)	10	5	15
total	30	20	50

We can find the following probabilities from the table:

$P(A) = 35/50 = .7$ and $P(A^c) = 15/50 = .3$

$P(B) = 30/50 = .6$ and $P(B^c) = 20/50 = .4$

But what if we wanted to know the probability that a person who buys a drink also buys gas? Or, in other words, given that a person bought a drink (B), what is the probability that they bought gas (A)? We write this as

$P(A|B) =$ the probability of A happening given that B has happened = PROBABILITY OF A GIVEN B

In this case, since we know that the person bought a drink, our sample space is not all 50 people, it is just the 30 people who bought drinks. Of these 30 people who bought drinks, 20 of them bought gas, so

$$P(A|B) = \frac{20}{30} = \frac{2}{3} \approx .67$$

The knowledge that we wanted to only consider the people who bought drinks will shrink our sample space. This is called conditional probability as this probability we found was based on the condition of having bought a drink. In general,

The CONDITIONAL PROBABILITY of event E given event F is

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

So using the formula here,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{20/50}{30/50} = \frac{2}{3}$$

What is the probability that a person who buys gas also buys a drink? Reword this as given the person has bought gas, what is the probability they bought a drink? Or, $P(B|A)$=?

We can look only at those 35 people who bought gas and find that 20 of them bought drinks, so $P(B|A) = 20/35 = 4/7$ or use the formula,

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{20/50}{35/50} = \frac{4}{7} \approx .57$$

What is the probability that a person who doesn't buy a drink buys gas?

$$P(A|B^c) = \frac{P(A \cap B^c)}{P(B^c)} = \frac{15/50}{20/50} = \frac{15}{20} = .75$$

We can rearrange our formula for conditional probability to find the PRODUCT RULE:

$P(E|F) = \frac{P(E \cap F)}{P(F)}$ and $P(F|E) = \frac{P(F \cap E)}{P(E)}$

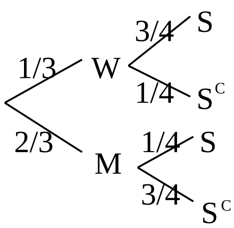
But $P(E \cap F) = P(F \cap E)$, so

$$P(E \cap F) = P(F)P(E|F) = P(F \cap E) = P(E)P(F|E)$$

We can use tree diagrams to show these conditional processes:

At a party, 1/3 of the guest are women. 75% of the women wore sandals and 25% of the men wore sandals. What is the probability that a person chosen at random at the party is a man wearing sandals?

We can show these probabilities in a tree diagram:



We find the probability of each outcome by multiplying along the branches. This is the product rule. The probability of being a woman and wearing sandals, $P(W \cap S)$, is the probabililty of being a woman, $P(W) = 1/3$, times the probability of wearing sandals given that you are a woman, $P(S|W) = 3/4$:

$$P(W \cap S) = P(W)P(S|W) = 1/3 \cdot 3/4 = 3/12 = 1/4$$

We can find all of the probabilities by multiplying along the branches.

$$P(W \cap S^c) = P(W)P(S^c|W) = 1/3 \cdot 1/4 = 1/12$$

$$P(M \cap S) = P(M)P(S|M) = 2/3 \cdot 1/4 = 2/12 = 1/6$$

$$P(M \cap S^c) = P(M)P(S^c|M) = 2/3 \cdot 3/4 = 6/12 = 1/2$$

Notice that the probability adds up to be 1 (3/12+1/12+2/12+6/12). Everyone has to be in one of the catagories, so we must have a total probability of 1.

To find the probability that a person chosen at random is wearing sandals will come from adding up all the ways that you could be wearing sandals:

$$P(S) = P(W \cap S) + P(M \cap S) = 3/12 + 2/12 = 5/12$$

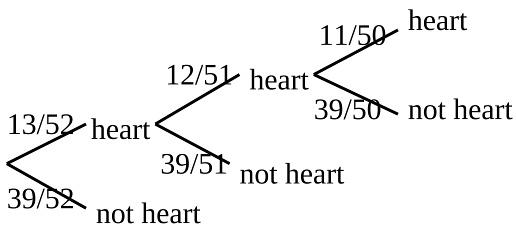
To find the probability that a person chosen at random is a woman or is wearing sandals you add up the probabilities of the cases that fit our statement:

$P(W \cup S) = P(W \cap S) + P(W \cap S^c) + P(M \cap S)$

$= 3/12 + 1/12 + 2/12 = 6/12 = 1/2$

We can have three branches if our problem has three steps. When what you do depends on what went before, it is called a stochastic process. Here we had to know if it was a man or a woman before we could say what the probability is of having sandals on. Another stochastic process is drawing cards from a deck without replacement. As the cards are removed, the probabilities change...

Consider drawing 3 cards from a standard deck of 52 cards without replacement. What is the probability that the three cards are hearts? what is the probability that the third card drawn is a heart given the first two cards are hearts? Use a tree diagram as this is a stochastic process. We do not need the whole tree diagram, the do the part we need:



Find the probability of 3 hearts by multiplying along the branch that does what we want:

$$P(3hearts) = \frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} = \frac{1716}{132600} = \frac{\binom{13}{3} \cdot \binom{39}{0}}{\binom{52}{3}} \approx .0129$$

To find the probability that the third card is a heart given the first two are hearts we just use the probability on the branch as it it GIVEN that we are already that far:

$$P(H_3|H_1 \cap H_2) = \frac{11}{50}$$

Events don't have to depend on each other. We call these kinds of events INDEPENDENT. For example, when you roll two dice what number each shows is independent - the dice are not influencing each other in any way. If you know that the green die shows a 6, do you know anything about the red die! NO - the dice are independent of each other. The drawing of cards without replacement is not independent, because when you know what card is drawn out, you know something about what is left. The definition of independce is that E and F are independent if

$P(E|F) = P(E)$ or $P(F|E) = P(F)$

We say that the "given" doesn't matter. Looking at how this changes the product rule,

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = P(E) \rightarrow P(E) \cdot P(F) = P(E \cap F)$$

So we can determine that E and F are independent if and only if $P(E) \cdot P(F) = P(E \cap F)$

You will need to use this rule to find if events are independent and if you are given that the events are independent, you will use the rule to find the probabilities.

Example - A certain part on the space shuttle has a 1% probability of failure during the mission. If they carry two back-ups to this part, what is the probability that they will all fail?

$$P(all\ fail) = .01 \cdot .01 \cdot .01 = .000001 = .0001\%$$

Example - A medical experiment showed the probability that a new medicine was effective was .75, the probability of a certain side effect was .4 and the probability for both occuring is .3. Are these events independent?

If E is the event that the medicine is effective and F is the event that there is a side effect, then $P(E \cap F) = .3$. Use our rule to check

$$P(E) \cdot P(F)(? =?)P(E \cap F) \rightarrow .75 \cdot .4(? =?).3 \rightarrow YES$$

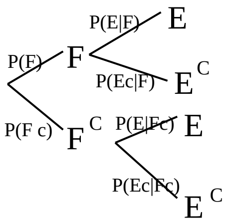
Since the rule held, the two events are independent.

You want to use the rule to determine if the events are indendent unless you are sure from the data given that they are independent. In those cases, you use the product to find the probability of both events happening.

7.6 Baye's Theorem

Say we know $P(E|F)$, can we find $P(F|E)$? YES - use the definition, $P(F|E) = \dfrac{P(F \cap E)}{P(E)}$

Now do a tree diagram to find $P(E)$

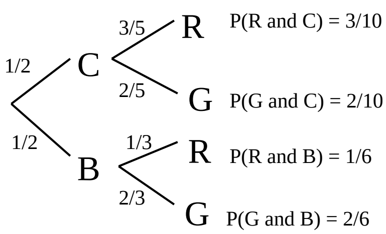


$$P(F|E) = \frac{P(F \cap E)}{P(F)P(E|F) + P(F^c)P(E|F^c)}$$

Your book uses the product rule to write this in a more confusing way. You should simply remember the definition of conditional probability and use a tree diagram or venn diagram to find the probability called for in the denominator. This formula with the probability in the denominator being a sum is Baye's theorem.

Example - We are to choose a marble from a cup or a bowl. We have flip a coin to decide to choose from the cup or the bowl. The bowl contains 1 red and 2 green marbles. The cup contains 3 red and 2 green marbles. What is the probability that a marble came from the bowl given that it is red?

Make a tree diagram. The choice of cup or bowl has to come first as we can't pick until we know where we are picking from.

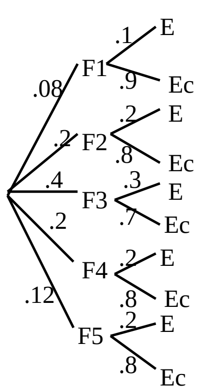


$$P(B|R) = \frac{P(B \cap R)}{P(R)} = \frac{P(B \cap R)}{P(R \cap B) + P(R \cap C)} = \frac{1/6}{1/6 + 3/10} = \frac{5}{14}$$

Example - Given the following table (from awhile back), what is the probability that a family of two wage earners selected at random will have a family income of less than \$10,000?

family income	proportion of families	proportion with 2 wage earners
$< 10,000 (F_1)$	.08	.1 (E)
$10,000 - 20,000 (F_2)$	.20	.2 (E)
$20,001 - 30,000 (F_3)$	.40	.3 (E)
$30,001 - 40,000 (F_4)$	.20	.2 (E)
$> 40,001 (F_5)$	.12	.2 (E)

We can show this in a tree diagram. Look carefully at what data we are given. The different income groups divide up all the families. Then when you know which income group you are in, then you can find the proportion with two wage earners. So we must have the income groups as our first branch.



$$P(F_1|E) = \frac{P(F_1 \cap E)}{P(E)}$$

$$= \frac{P(F_1 \cap E)}{P(F_1 \cap E) + P(F_2 \cap E) + P(F_3 \cap E) + P(F_4 \cap E) + P(F_5 \cap E)}$$

$$= \frac{.08 \cdot .1}{.08 \cdot .1 + .2 \cdot .2 + .4 \cdot .3 + .2 \cdot .2 + .12 \cdot .2} = .034$$