IV. CONDITIONAL PROBABILITY ; INTERSECTION; INDEPENDENCE

Conditional Probability:

Let A and B be events in a sample space S for a given experiment. Suppose we are told that the experiment has been performed and that the event A has occurred. What effect does this additional knowledge have on the occurrence of event B? Does the probability that B will also occur increase? Decrease? Stay the same? The probability that B will occur given that A has already occurred is called a *conditional probability* and is denoted by $P(B|A)$ (read "the probability of B given A").

Example 4.1:

An experiment consists of tossing a penny, a nickel and a dime. Events A, B, C, and D are defined as follows:

A = "the penny comes up heads",

B = "at least two heads come up",

C = "exactly two tails come up",

D = "the dime comes up tails" .

(a) Calculate $P(B|A)$ and compare it with $P(B)$.

(b) Calculate $P(C|A)$ and compare it with $P(C)$.

(c) Calculate $P(D|A)$ and compare it with $P(D)$.

Solution:

We know from previous examples that the sample space is:

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

where, for example, HTH means that the penny came up heads, the nickel came up tails and the dime came up heads. Also, the outcomes are equally likely and so the probability of each of the simple events is 1/8.

(a) Event $B = \{HHH, HHT, HTH, THH\}$ $P(B) = 4/8 = 1/2$.

Now suppose that we are told that event A *has occurred*, that is the penny came up heads. Since we *know* that the penny came up heads, the sample space has to be changed. It is now

$$S' = \{HHH, HHT, HTH, HTT\}$$

These are equally likely events and the probability of getting at least two heads is now $3/4$; $P(B|A) = 3/4$. Knowing that A has occurred has *increased* the probability B will occur.

(b) Event $C = \{HTT, THT, TTH\}$. In absence of any information, the probability that C occurs is $P(C) = 3/8$.

Now assume that event A has occurred . Then S' is the new sample space of equally likely events and the probability of getting exactly two tails is $1/4$; $P(C|A) = 1/4$. Knowing that A has occurred *decreased* the probability of C occurring.

(c) Event $D = \{HHT, HTT, THT, TTT\}$ and $P(D) = 4/8 = 1/2$

If we assume that event A has occurred, then the sample space is S' and $P(D|A) = 2/4 = 1/2$. In this case, knowing that A has occurred *did not change* the probability of D occurring.

Example 4.1 illustrates that in a given experiment, the occurrence of some event A may or may not influence the occurrence of some other event B.

Examples 4.2:

1. A box contains 10 disks numbered 1, 2, 3, …, 10. A person draws a disk from the box.
   (a) What is the probability that the number is a prime?

   (b) Suppose the person states that the number on the disk is odd. Now what is the probability that the number is a prime?

2. One hundred people responded to a survey that asked whether they thought women in the armed forces should be permitted to participate in combat. The results of the survey were:

Gender	Yes	No	Totals
Male	18	32	50
Female	8	42	50

Totals	26	74	100

(a) What is the probability that a respondent answered “yes” given the respondent was a female?

(b) What is the probability that a respondent was a male, given that the respondent answered “no”?

Solutions:

1. The sample space $S = \{1, 2, 3, ..., 10\}$ consists of 10 equally likely outcomes.
Let A = “the number on the disk is odd” and let B = “the number on the disk is a prime”

(a) $B = \{2, 3, 5, 7\}$ and

$$P(B) = \frac{n(A)}{n(S)} = \frac{4}{10} = \frac{2}{5}.$$

(b) Given that the disk drawn has an odd number on it, the sample space becomes $S' = A = \{1, 3, 5, 7, 9\}$. Now $A \cap B = \{3, 5, 7\}$ and

$$P(B|A) = \frac{n(A \cap B)}{n(A)} = \frac{3}{5}.$$

Knowing that the number on the disk is odd improved the probability that the number is a prime.

2. (a) Given that the respondent was a female, the sample space becomes the set of 50 female respondents. The probability that she answered “yes” (Y) is

$$P(Y) = \frac{8}{50} = \frac{4}{25}.$$

(b) Given that the respondent answered “no”, the sample space becomes the set of 74 respondents who answered “no”. The probability that the respondent was a male (M) is

$$P(M) = \frac{32}{74} = \frac{16}{27}.$$

We want to obtain a formula for the conditional probability of B given A in terms of the probabilities of the events A and B themselves. We will use a sample space S of equally likely outcomes to motivate the formula. Recall that in the equally likely case, the probability of an event E is given by

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

Knowing that the event A has occurred means that we must change the sample space from S to the subset $S' = A$. Now, to determine the probability that B will occur, given that A has already occurred, we need the number of elements in B that are also in A; that is, the number of elements in $B \cap A$. Thus,

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

Dividing numerator and denominator by $n(S)$, we obtain

$$P(B|A) = \frac{\frac{n(B \cap A)}{n(S)}}{\frac{n(A)}{n(S)}} = \frac{P(B \cap A)}{P(A)}.$$

This formula holds in general. For events A and B in an arbitrary sample space S, the conditional probability of B given A is given by

(4) $$P(B|A) = \frac{P(B \cap A)}{P(A)}, \quad \text{provided } P(A) \neq 0.$$

Examples 4.3:

1. A box contains black chips and white chips. A person draws two chips in succession without replacement. If the probability of selecting a black chip and a white chip is 15/56 and the probability of selecting a black chip on the first draw is 3/8, find the probability of selecting a white chip on the second draw given that first chip selected was a black chip.

2. In a certain housing subdivision, 35% of the homes have a family room and a fireplace. If 70% of the homes have a family room, find the probability that a home has a fireplace given that it has a family room.

Solutions:

1. Let B = "black chip" and W = "white chip" . Then

$$P(W|B)=\frac{P(B\cap W)}{P(B)}=\frac{\frac{15}{56}}{\frac{3}{8}}=\frac{15}{56}\cdot\frac{8}{3}=\frac{5}{7}.$$

2. Let R = "family room" and let F = "fireplace" . Then

$$P(F|R)=\frac{P(R\cap F)}{P(R)}=\frac{0.35}{0.70}=\frac{1}{2}.$$

Intersection of Events; the Product Rule:

Let A and B be two events in a sample space S, and assume $P(A)\neq 0$ and $P(B)\neq 0$. Then, by the formula (4),

$$P(B|A)=\frac{P(B\cap A)}{P(A)} \qquad \text{and} \qquad P(A|B)=\frac{P(A\cap B)}{P(B)}.$$

Solving the first equation for $P(B\cap A)$ and the second equation for $P(A\cap B)$, we get

$$P(B\cap A)=P(A)P(B|A) \qquad \text{and} \qquad P(A\cap B)=P(B)P(A|B).$$

Since $A\cap B=B\cap A$, we have the *product rule*:

(5) $$P(A\cap B)=P(A)P(B|A)=P(B)P(A|B).$$

We can use either of these equations to calculate $P(A\cap B)$.

Examples 4.4:

1. 70% of the shoppers at a certain grocery store are female and 80% of the female shoppers have a charge account at the store. Find the probability that a shopper selected at random is a female and has a charge account.

2. Suppose 60% of the male shoppers at the same grocery store have a charge account at the store. What is the probability that a shopper selected at random is a male who does not have a charge account.

Solutions:

Let C = “shoppers with charge account”, F = “female shoppers” , and M = “male shoppers” .

1. Since 70% of the shoppers are female, the probability that a shopper selected at random is a female is

$$P(F) = 0.70 .$$

Since 80% of the female shoppers have charge accounts at the store, the probability that a shopper has a charge account, given that the shopper is a female, is

$$P(C|F) = 0.80 .$$

Therefore, the probability that a shopper selected at random is a female and has a charge account is

$$P(C \cap F) = P(F)P(C|F) = 0.70 \cdot 0.80 = 0.56 \ .$$

2. We conclude that 30% of the shoppers at the grocery store are male. Since 60% of the males have charge accounts, 40% do not. Therefore, $P(M) = 0.30$ and $P(C^c|M) = 0.40$. The probability that a shopper selected at random is a male and does not have a charge account is

$$P(C^c \cap M) = P(M)P(C^c|M) = 0.30 \cdot 0.40 = 0.12 \ .$$

Probability Trees:

We used tree diagrams in the section on Counting to help us count the number of outcomes in an experiment that involved several steps. Here we use tree diagrams, called *probability trees*, to help us calculate the probabilities of events for experiments that consist of a sequence of steps.

Example 4.5.

A box contains 9 balls; 6 of them are red and 3 are white. A ball is drawn from the box and, without replacing it, a second ball is drawn.

(*a*) Use a tree diagram to determine the sample space *S*.

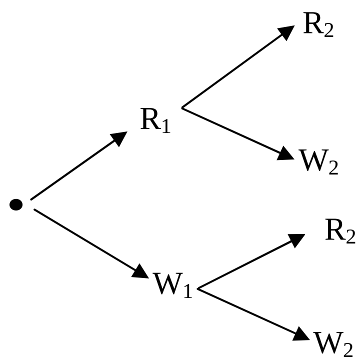
(*b*) Assign probabilities to the simple events.

(*c*) Let *A* be the event "the second ball is red". Calculate *P*(*A*).

(*d*) What is the probability that the first ball is red, given that the second ball is white?

Solutions:

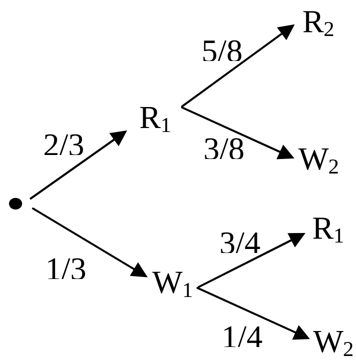
(*a*)



The sample space is $S = \{R_1 \cap R_2, R_1 \cap W_2, W_1 \cap R_2, W_1 \cap W_2\}$ where $R_1 \cap R_2$ represents the outcome "red on the first ball and red on the second ball", $R_1 \cap W_2$ represents red on the first ball and white on the second, and so on.

(*b*) The probability of drawing a red ball on the first draw is $6/9 = 2/3$ so we assign the probability $2/3$ to the branch SR_1. The probability that the first ball drawn is white is $3/9 = 1/3$ so we assign that value to the branch SW_1.

Now to the next level: What probability should we assign to the branch R_1R_2? This is the conditional probability $P(R_2|R_1)$, that is, the probability that the second ball is red given that the first ball was red. Since the box now contains 8 balls and 5 of them are red, $P(R_2|R_1) = 5/8$. Continuing in this manner, we assign probabilities to the other branches of the tree to obtain the result in Figure A.



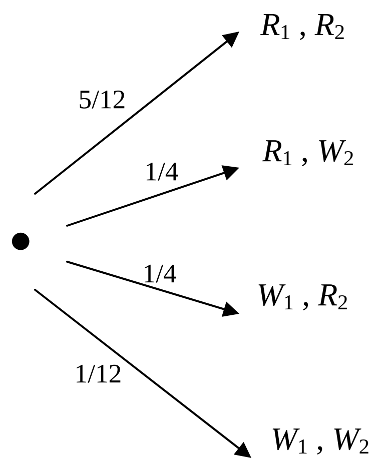


Figure A

Now, $P(R_1 \cap R_2) = P(R_1)P(R_2 | R_1) = \frac{2}{3} \cdot \frac{5}{8} = \frac{5}{12}$,
$P(R_1 \cap W_2) = P(R_1)P(W_2 | R_1) = \frac{2}{3} \cdot \frac{3}{8} = \frac{1}{4}$, and so on. The results are shown in Figure B.

(c) $A = \{R_1 \cap R_2, W_1 \cap R_2\}$ and $P(A) = P(R_1 \cap R_2) + P(W_1 \cap R_2) = \frac{5}{12} + \frac{1}{4} = \frac{2}{3}$.

(d) By the conditional probability formula, $P(R_1 | W_2) = \dfrac{P(R_1 \cap W_2)}{P(W_2)}$. We get $P(R_1 \cap W_2) = \frac{1}{4}$ from the tree diagram. Now, the event W_2 = "the second ball is white" is given by $W_2 = \{R_1 \cap W_2, W_1 \cap W_2\}$ and so $P(W_2) = \frac{1}{4} + \frac{1}{12} = \frac{1}{3}$. Therefore,

$$P(R_1 | W_2) = \frac{P(R_1 \cap W_2)}{P(W_2)} = \frac{\frac{1}{4}}{\frac{1}{3}} = \frac{3}{4}.$$

Independent/Dependent Events:

As we discovered in Example 4.1, the occurrence of some event A may or may not change the probability that some other event B occurs. Events A and B are *independent* if the occurrence of one does not change the probability of the occurrence of the other. In symbols, A and B are independent if

$$P(A|B) = P(A) \quad \text{or if} \quad P(B|A) = P(B).$$

To be precise, we should say that if $P(A|B) = P(A)$, then A is independent of B, or if $P(B|A) = P(B)$, then B is independent of A. However, it can be shown that if A and B are events with nonzero probabilities, then A is independent of B if and only if B is independent of A. This is one of the problems in the Exercise Set.

If $P(A|B) \neq P(A)$ or if $P(B|A) \neq P(B)$, then we say that A and B are *dependent* events.

Formula (5) provides another way to view independence. In particular, if A and B are independent in the sense that $P(A|B) = P(A)$ or $P(B|A) = P(B)$, then by formula (5),

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

Moreover, if A and B are events with nonzero probabilities, then (6) implies that A and B are independent.

There are some obvious examples of the concept of independence/dependence. If we toss a fair coin twice, the outcome on the second toss is independent of the outcome on the first toss, a coin has no memory. If we draw a card from a standard deck, put it back, and then draw a second card, the outcome on the second card is independent of the outcome on the first card. On the other hand, if we draw a card from the deck and do not replace it, the outcome on the second draw *does* depend on the outcome of the first draw. Here are some less obvious examples.

Examples 4.6.

1. Continuing with Example 4.5, let D = ‘the balls have different colors” and let W = “second ball is white”.

(a) Determine $P(D)$.

(b) Test D and R_1 (“red ball on the first draw”) for independence.

(c) Test W and D for independence.

2. Let E be the event that all of a family’s children are the same sex, and let F be the event that the family has at most 1 boy. Assuming that having a girl is as likely as having a boy, test events E and F for independence if:

(a) The family has 2 children.

(b) The family has 3 children.

Solutions:

1. See the solution for Example 4.5.1. Write the sample space as $S=\{R_1R_2, R_1W_2, W_1R_2, W_1W_2\}$. The events D and W are given by $D=\{R_1W_2, W_1R_2\}$ and $W=\{R_1W_2, W_1W_2\}$.

(a) $P(D)=P(R_1W_2)+P(W_1R_2)=\frac{6}{24}+\frac{6}{24}=\frac{12}{24}=\frac{1}{2}$.

(b) We'll compare $P(D|R_1)$ with $P(D)$:

$$P(D|R_1)=\frac{P(D\cap R_1)}{P(R_1)}=\frac{6/24}{2/3}=\frac{6}{24}\cdot\frac{3}{2}=\frac{3}{8}\neq P(D).$$

Alternatively, we can compare $P(D\cap R_1)$ with $P(D)P(R_1)$:

$$P(D\cap R_1)=\frac{6}{24}=\frac{1}{4}\quad\text{and}\quad P(D)P(R_1)=\frac{1}{2}\cdot\frac{2}{3}=\frac{1}{3}.$$

Either way, we conclude that D and R_1 are dependent.

(c) $P(W)=P(R_1W_2)+P(W_1W_2)=\frac{6}{24}+\frac{2}{24}=\frac{8}{24}=\frac{1}{3}$;

$$P(W|D)=\frac{P(W\cap D)}{P(D)}=\frac{10/24}{1/2}=\frac{5}{12}\cdot 2=\frac{5}{6}\neq P(W).$$

Thus, W and D are dependent.

2. (a) The sample space is $S=\{BB, BG, GB, GG\}$ and the outcomes are equally likely; $E=\{BB, GG\}$ and $P(E)=\frac{1}{2}$; $F=\{BG, GB, GG\}$, and $P(F)=\frac{3}{4}$.

Testing for independence:

$$P(E|F)=\frac{P(E\cap F)}{P(F)}=\frac{1/4}{3/4}=\frac{1}{3}\neq P(E).$$

Thus, E and F are dependent events.

(b) The sample space is $S = \{BBB, BBG, BGB, BGG, GBB, GBG, GGB, GGG\}$ and the outcomes are equally likely; $E = \{BBB, GGG\}$ and $P(E) = \frac{1}{4}$;

$F = \{BGG, GBG, GGB, GGG\}$, and $P(F) = \frac{1}{2}$.

Testing for independence:

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{1/8}{1/2} = \frac{1}{4} = P(E).$$

Thus, in this case, *E* and *F* are independent events.

Bayes' Formula:

In experiments that consist of several stages, we typically move from stage to stage in order. That is, we compute the probabilities of the outcomes of the first stage, then, given the outcome at the first stage, we compute the probabilities at the second stage, and so on. See Example 4.5, parts (b) and (c). On the other hand, if you look at part (d), we took an outcome at the second stage and asked for the conditional probability of an outcome at the first stage. Here are some more examples of this type.

Examples 4.7.

1. Box 1 contains 4 red balls and 3 white balls, Box 2 has 2 red balls and 5 white balls. A card is drawn from a standard deck. If the card is a heart, a ball is drawn from the Box 1; if the card is not a heart, a ball is drawn from Box 2.

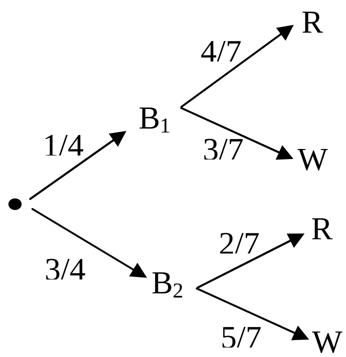
> (*a*) If the ball is white, what is the probability that it came from **B**ox 1? What is the probability that it came from Box 2?
>
> (*b*) If the ball is red, what is the probability that it came from Box 1? From Box 2?

2. A company produces 100 DVD's per day at three plants. Plant A produces 400 DVD's per day; plant B produces 250 per day; and plant C produces 350 per day. Production records show that 4% of the DVD's produced at plant A will be defective, 3% at plant B will be defective, and 5% at plant C will be defective. At the end of each day, the DVD's are shipped to a distribution center. Suppose a DVD at the center is selected at random and found to be defective.

> (*a*) What is the probability that it came from plant A?
>
> (*b*) What is the probability that it came from either B or C?

Solutions:

1. Since the probability of drawing a heart is $\frac{1}{4}$ and the probability of not drawing a heart is $\frac{3}{4}$, we choose Box 1 with probability $\frac{1}{4}$ and we choose Box 2 with probability $\frac{3}{4}$. Let *R* = "red ball" and *W* = "white ball". Now, tree diagram for this experiment is:



From this diagram, we calculate that $P(W)=\frac{3}{28}+\frac{15}{28}=\frac{18}{28}=\frac{9}{14}$ and $P(R)=\frac{5}{14}$.

(*a*) $$P(B_1|W)=\frac{P(B_1\cap W)}{P(W)}=\frac{3/28}{9/14}=\frac{3}{28}\cdot\frac{14}{9}=\frac{1}{6}$$

$$P(B_2|W)=\frac{P(B_2\cap W)}{P(W)}=\frac{15/28}{9/14}=\frac{15}{28}\cdot\frac{14}{9}=\frac{5}{6}$$

Note that since "$B_2|W$" is the "complement" of "$B_1|W$", we could have used

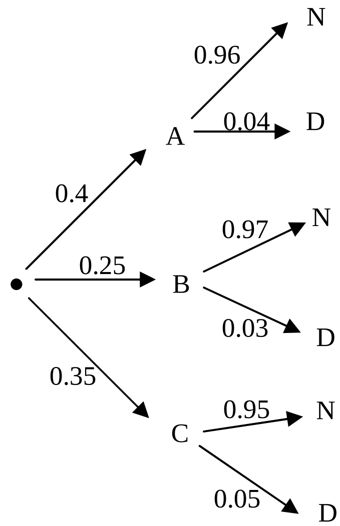
$$P(B_2|W)=1-P(B_1|W)=1-\frac{1}{6}=\frac{5}{6}$$

to calculate $P(B_2|W)$.

(*b*) $P(B_1|R)=\frac{P(B_1\cap R)}{P(R)}=\frac{4/28}{5/14}=\frac{4}{28}\cdot\frac{14}{5}=\frac{2}{5}$ and $P(B_2|R)=\frac{3}{5}$.

2. Let D = "defective DVD" and N = "not defective DVD". The probability that a DVD selected at random comes from plant A is $P(A)=\frac{400}{1000}=0.4$, the probability that it comes

from plant B is $P(B)=\frac{250}{1000}=0.25$, and the probability that it comes from plant C is $P(C)=\frac{350}{1000}=0.35$. This problem represented by the tree diagram:



Now, $P(D)=0.4(0.04)+0.25(0.03)+0.35(0.05)=0.0410$

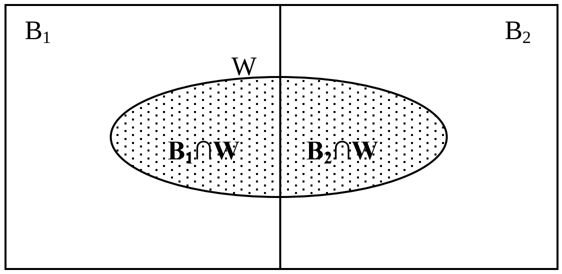
(*a*) $P(A|D)=\dfrac{P(A\cap D)}{P(D)}=\frac{0.4(0.04)}{0.0410}\approx 0.3902$.

The probability that the defective DVD came from plant A is 0.3902.

(*b*) If the defective DVD did not come from plant A, then it came from either plant B or plant C. Therefore,

$$P(B\cup C)=1-P(A)=1-0.3902=0.6098.$$

Let's analyze Parts 1 and 2 of Examples 4.7 using Venn diagrams. In Part 1, we have two events B_1 = "Box 1 is chosen", and B_2 = "Box 2 is chosen" , and we have the event W = "the ball is white". Note that B_1 and B_2 are mutually exclusive (disjoint) sets and $S=B_1\cup B_2$. The Venn diagram shows these events:



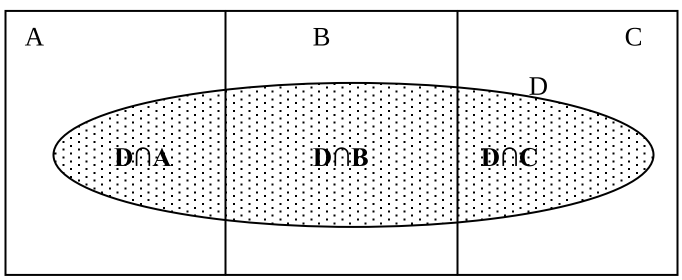
Now, from the diagram, $W = (B_1 \cap W) \cup (B_2 \cap W)$ and

$$P(W) = P(B_1 \cap W) + P(B_2 \cap W) \text{ since } (B_1 \cap W) \cap (B_2 \cap W) = \varnothing .$$

Thus,

$$P(B_1 | W) = \frac{P(B_1 \cap W)}{P(W)} = \frac{P(B_1 \cap W)}{P(B_1 \cap W) + P(B_2 \cap W)} .$$

In Part 2, we have the three events A = "it came from plant A", B = "it came from plant B", C = "it came from plant C", and D = "defective DVD". Events A, B, and C are mutually exclusive and their union is $S = A \cup B \cup C$. A Venn diagram for the events is



From the diagram, $D = (A \cap D) \cup (B \cap D) \cup (C \cap D)$ and

$$P(D) = P(A \cap D) + P(B \cap D) + P(C \cap D) .$$

Therefore,

$$P(A | D) = \frac{P(A \cap D)}{P(D)} = \frac{P(A \cap D)}{P(A \cap D) + P(B \cap D) + P(C \cap D)} .$$

These examples are special cases of a general result known as Bayes' formula. [Thomas Bayes (1702-1763) was a Presbyterian Minister.]

Bayes' Formula: Let $U_1, U_2, ..., U_n$ be n mutually exclusive events whose union is the sample space S. Let E be an arbitrary event in S such that $P(E) \neq 0$. Then

$$P(U_1|E) = \frac{P(U_1 \cap E)}{P(U_1 \cap E) + P(U_2 \cap E) + ... + P(U_n \cap E)}.$$

Corresponding results hold for $U_2, U_3, ..., U_n$.

Exercise Set 3.4:

1. Two cards are drawn in succession, without replacement, from a standard 52 card deck. What is the probability that the second card is a heart, given that the first card is a heart? Are these two events dependent or independent?

2. What is the probability that the second card is a heart, given that the first card is red? Dependent or independent?

3. What is the probability that the second card is a spade, given that the first card is a Jack? Dependent or independent?

4. What is the probability that the second card is red, given that the first card is a heart? Dependent or independent?

5. One of 3 prisoners is randomly selected to be shot in the morning. They are not told who it will be. The guard offers to tell Prisoner A the name of one of Prisoners B and C who is not going to be shot. The guard is truthful. What is the probability that A will be shot, given that C will not be shot? What is the probability that A will be shot, given that the guard says that C will not be shot? To answer, you need to make an assumption about the guard's behavior. State that assumption.

6. A certain genetic defect occurs in only 5% of the population. Among patients with the defect, a test gives the correct diagnosis 90% of the time and the wrong diagnosis 10% of the time. Among patients without the defect, the test is right 80% of the time and wrong 20% of the time. Given that the test is positive (says that the defect is present), what is the probability that the patient actually has the defect?

7. The table below gives the result of a survey of employees about their supervisor/employee gender relationships (supervisor listed first) and their job satisfaction. Assume that one respondent is randomly selected from the set of all

respondents to this survey.

	Female/Male	Female/Female	Male/Male	Male/Female
Very Satisfied	21	25	54	71
Barely Satisfied	39	49	50	38
Not Satisfied	31	48	10	11

What is the probability that the respondent is a male? Given that the respondent is a male, what is the probability that he is very satisfied with his job? Given that the respondent is barely satisfied, what is the probability that she is a female?

8. Players A and B take turns flipping a fair coin. A goes first. The first player to toss a head is the winner. What is the probability that B wins? (Hint: Let p denote the probability that B wins. Then 1-p is the probability that A wins.