DOCUMENT RESUME

ED 173 096 SE 027 903

TITLE	School Mathematics Study Group, Unit Number Fourteen. Chapter 25 - Statistics. Chapter 26 - Systems of Sentences in Two Variables.
INSTITUTION	Stanford Univ., Calif. School Mathematics Study Group.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	68
NOTE	93p.; Not available in hard copy due to small, light and broken type
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Algebra; Curriculum; *Geometric Concepts; *Instruction; Mathematics Education; Secondary Education; *Secondary School Mathematics; *Statistics; *Textbooks
IDENTIFIERS	*School Mathematics Study Group

ABSTRACT

This is unit fourteen of a fifteen-unit SMSG secondary school text for high school students. The text is devoted almost entirely to mathematical concepts which all citizens should know in order to function satisfactorily in our society. Chapter topics include statistics and systems of sentences in two variables. (MP)

SCHOOL
MATHEMATICS
STUDY GROUP

UNIT NUMBER FOURTEEN

Chapter 25. Statistics

Chapter 26. Systems of Sentences in Two Variables

TABLE OF CONTENTS

Chapter 25

STATISTICS

25-1. Introduction

We often must deal with a list of numbers: scores in a bowling match, the number of inches annual rainfall in San Francisco, the wages paid the different employees of a company, the number of TV sets sold in a department store each month, and so on. What can we do with these numbers? It is hard to draw conclusions just by looking at the numbers. This is particularly true if the list is a long one. We need some way to see the situation as a whole.

It is often helpful to show the position of the numbers along the number line. When there are repetitions we can pile the points or dots above one another. The result is called a <u>dot frequency diagram</u>.

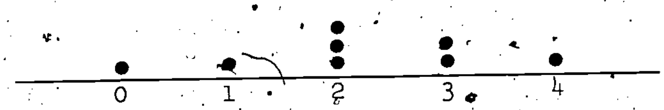


Figure 1.

Thus Figure 1 is a dot frequency diagram for 0, 1, 2, 2, 2, 3, 3, 4.

To illustrate the use of such a diagram, suppose that a manuscript has just been discovered. It is suspected that it was written by Sir Christopher Wren, the architect of St. Paul's Cathedral (1632-1723). In the attempt to decide whether this conjecture is correct, a careful study is made of the style of the manuscript so that it can be compared with the known writings of Wren. One obvious thing to study is the length of the words used. Some authors tend to use many long words. Others are more sparing in their use. To show how a study of word length can be carried out, we have chosen at random a page in a book and made the following tally of the number of letters (694) in the first 140 words.

Number of letters	Tally	Frequency	Totals
1	|	1	1
2	𝍸 𝍸 𝍸 𝍸	20	40
3	𝍸 𝍸 𝍸 𝍸 𝍸 ||	27	81
4	𝍸 𝍸 𝍸 𝍸 𝍸	25	100
5	𝍸 𝍸 𝍸 ||	17	85
6	𝍸 𝍸 𝍸	15	90
7	𝍸 𝍸 ||||	14	98
8	𝍸 𝍸	10	80
9	𝍸	5	45
10	|	1	10
11	|	1	11
12	|	1	12
13	||	2	26
14		0	0
15	|	1	15
		140	694

Table 1

The results of this tally are given in the frequency column of Table 1. The last column of this table gives the total number of letters in all the words of a given length. For example, for words of 5 letters the frequency is 17 and consequently there are $5 \times 17 = 85$ letters.

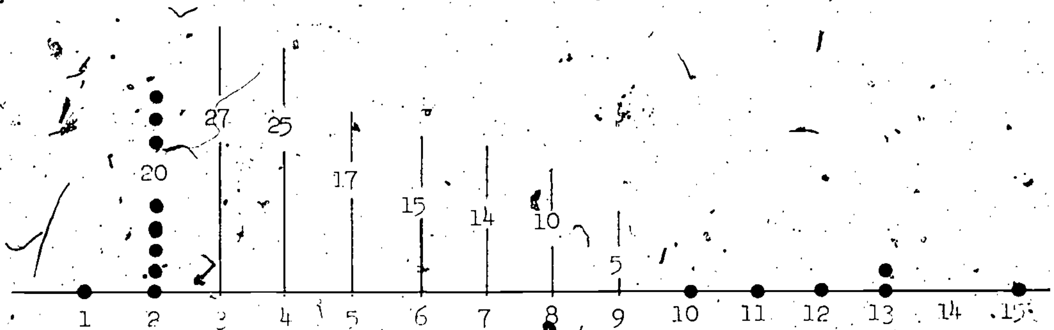


Figure 2

Figure 2 shows how these numbers appear on the number line. In cases where there would be many dots above a numeral, we have simplified the picture by running the dots together so that the number of occurrences (the frequency) becomes the ordinate of the topmost dot.

The numbers have a tendency to cluster about a central value. In practice there are three different ways that are used to determine the center.

(1) The mode is the most fashionable value, the one that has the highest frequency. In our example, the mode is 3 since its frequency (27) is the highest.

(2) The median is the middle number in a listing in order of size. If the number of items is even there are two candidates for the middle. In this case we choose the smaller. In our example, then, we locate the 70th number which is 4.

(3) The mean or average is the total number of letters divided by the number of words. In our example, 694 divided by 140 is 4.96. The total number of letters is obtained by adding the subtotals in the last column of the table.

We notice that the mode and the median are necessarily whole numbers, but that the mean need not be.

Each measure of central tendency is useful for certain purposes. However, we shall deal almost exclusively with the mean since the mathematics is so much more satisfactory. In the exercises you will discover certain properties of the mode, median, and mean.

Exercises 25-1

1. In a theme that you have written for English or History, make a tally of the number of words of different lengths from among the first 200 words. Construct a dot frequency diagram. Find the mode, median, and mean.

2. (a) Which of the measures of central tendency is certain to change if one of the numbers is changed?

   (b) Show that the other measures need not change.

3. Give an example to show that there could be two or more equally popular numbers in a list. In this case it is usually agreed that the set of numbers does not have a mode. Could a set of numbers have two medians? two means?

4. In a certain small town, there are 1001 persons with incomes: one with an income of $2,000,000 a year and 1000 wage earners, each with an annual income of $6,000. What is the mode? the median? the mean? If you were thinking of moving to this town, which of these three measures would you prefer to know?

5. Which of the three measures of central tendency can be found without arranging the numbers in order or finding the frequencies?

6. If you add a single number a to a list of n numbers, what is the greatest possible effect on the mode? on the median? on the mean? Illustrate your conclusions.

7. For each member of your class list the number of complete months that have passed since his or her last birthday. Make a dot frequency diagram. Find the mode, median, and mean for the numbers of months. Save your data for use in the next section.

8. List the weights of the members of your class to the nearest pound. Make a dot frequency diagram. Find the mode, median, and mean.

25-2. Cumulative Frequency

Let us turn back to the data on word length presented in the last section and consider another way of picturing these data. For convenience we copy the frequencies from the table obtained from the tally.

Number of Letters in Word	Frequency	Cumulative Frequency
1	1	1
2	20	21
3	27	48
4	25	73
5	17	90
6	15	105
7	14	119
8	10	129
9	5	134
10	1	135
11	1	136
12	1	137
13	2	139
15	1	140

We have added a third column headed Cumulative Frequency. To find an entry in this column we add to the entry above it the frequency listed on the same line. For example, on line 7, the entry is $105 + 14 = 119$.

Now for the new picture. We locate the points (x,y) where x is a number in the first column and y is the corresponding cumulative frequency listed in column three. (See Figure 3.)

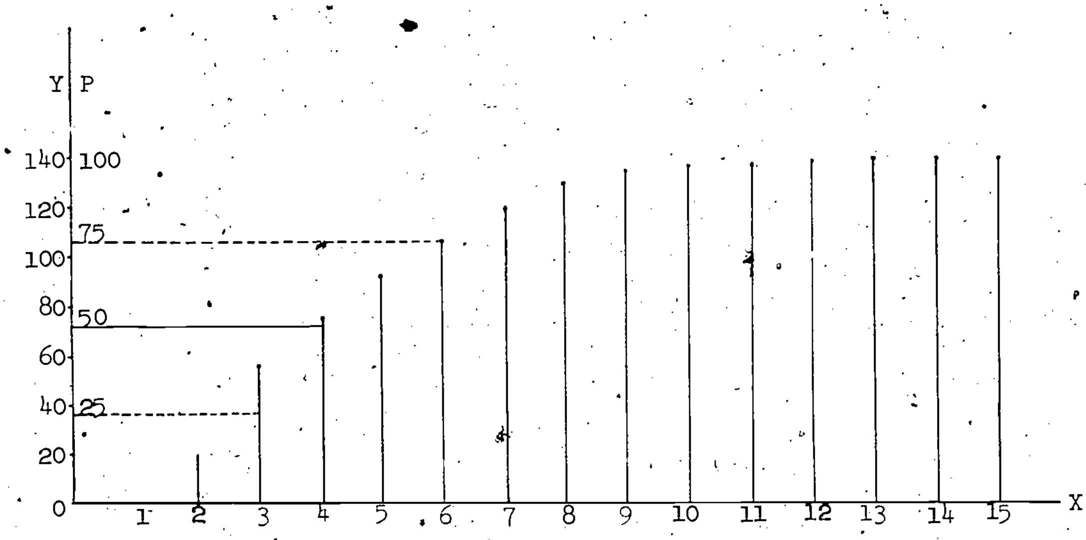


Figure 3

For greater clarity we draw vertical lines to represent the ordinates of these points.

Let us draw a horizontal line through the midpoint M of the interval [0,140] on the Y-axis and find its first intersection with an ordinate. This occurs at $x = 4$ which is the median of our set of numbers. The interval [0,4] on the X-axis contains all the points whose ordinate y is in the interval [0,70] on the Y-axis and the interval [4,15] on the X-axis contains all the points whose ordinate lies in the interval [70,140] on the Y-axis.

Often the entire cumulative frequency interval is divided into <u>four</u> equal subintervals. In our example, these intervals are [0,35], [35,70], [70,105], and [105,140]. The corresponding intervals on the X-axis are [0,3], [3,4], [4,6], and [6,15]. The number 3 is called the <u>lower quartile</u> and 6 is called the <u>upper quartile</u>.

Sometimes even finer divisions are used. This is done by dividing the total cumulative frequency interval into 100 equal parts. If we mark a new scale on the Y-axis with 100 opposite 140 we can read percents of the total frequency (140) directly.

Let a horizontal line segment be drawn on Figure 3 from the point marked P percent to its intersection with the first ordinate encountered. The corresponding abscissa x is called the Pth <u>percentile</u>. Thus the median is the 50th percentile and the lower and upper quartiles are the 25th and 75th percentiles respectively. The 10th percentile is 1. The 90th percentile corresponds to a cumulative frequency of $\frac{90}{100} \cdot 140 = 126$. It is therefore 8.

Exercises 25-2

1. For the example in the text, find the 30th percentile; the 80th percentile, and the 60th percentile.
2. In the example in the text, show that the 40th percentile and the 50th percentile are the same.
3. Use the data collected in problem 7 of Exercises 25-1 to construct a table of cumulative frequencies and to draw a cumulative frequency diagram.
4. From the results of Exercise 3, find the median, the lower quartile, the upper quartile, and the 60th percentile.
5. Find the lower quartile, median, and upper quartile for problem 4 of the previous section.

25-3. The Uses of Statistics

We started our discussion of word length with an imaginary discovery of a manuscript which was thought to be written by Sir Christopher Wren. The pattern of word length might be one piece of evidence. Of course, we would hope for something better, like a reference to some event that only he would have been likely to know.

However, sticking to the word length study, questions arise about the reliability of a sample of this sort. That is, if we assume that the document was written by Wren, how likely is it that this sample of 140 words would resemble similar samples taken from Wren's known writings? If it were <u>not</u> written by him, how much agreement could be expected? We would expect that the reliability of a sample would increase with its size so that if we knew the pattern for a 10,000 word sample we would have more confidence than if we had to reason from a rather small one. So sample size is surely important.

The problem of sampling is a central one in the study of statistics. Its importance is rather obvious. When we wish to predict the results of an election, we take a poll of the opinions of a relatively small sample of the population. To poll the whole population would be frightfully expensive. If we wish to test a new drug, we try it out on a sufficiently large sample of patients to draw what we hope are valid conclusions. What determines the reliability of a sample? We shall have something to say about this question toward the end of the chapter. Before we can do this we must learn to <u>describe</u> any set of numbers which might represent a sample. We shall learn to do this in the next few sections.

We shall first learn more about the average or mean of a set of numbers. Then we shall learn how to measure the extent to which a set of numbers is scattered about its mean as a center. It is very remarkable that a list of numbers, even a long list, can be rather well described by using only <u>two</u> numbers: one a measure of central tendency; the other a measure of scatter.

Hereafter we shall have nothing more to say about the mode or the median as a measure of central tendency. We shall use for this purpose only the mean (or average).

Exercises 25-3

(One third of the class should do Exercise 1, a second third Exercise 2, and the remaining third Exercise 3. Everyone should do Exercise 4.)

1. Select any book of a general or popular nature. Turn to a page at random and make a tally of the number of letters in the first 100 words. Find the average or mean word length. Repeat this experiment for four other pages in the same book. Compute the averages of your samples.
2. Select a child's book. Proceed as in Exercise 1.
3. Select a technical book. Proceed as in Exercise 1.
4. Discuss a number of practical situations in which it is important to use the results of a sample to draw conclusions about a population. The word "population" is used here to stand for any set of numbers from which the sample is chosen. The numbers need not be numbers of people.

25-4. Averages

Suppose that a student makes the following five bowling scores:

130, 120, 135, 145, 140.

What single number best respresents his performance? The usual answer is the average of these five numbers, that is

$$\frac{130 + 120 + 135 + 145 + 140}{5} = \frac{670}{5} = 134.$$

Instead of the word average we often use the term arithmetic mean or simply the word mean.

In general, if we have n numbers

$$x_1, x_2, x_3, \ldots, x_n,$$

their average $\bar{x}$, pronounced "x bar," is simply

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}.$$

The average gives equal importance to each of the numbers in the list. To find the average we merely add these numbers and divide by the number of items. This is a very simple process. Nevertheless a careful study of the properties of the average leads to remarkable results, as we shall see. But this is typical of mathematics where thinking so often pays very large dividends.

To begin with, we remark that it is not necessary that the numbers $x_1, x_2, \ldots, x_n$ be all different from each other. For example, the average of

80, 80, 90, 100

is

$$\frac{80 + 80 + 90 + 100}{4} = \frac{350}{4} = 87.5.$$

In the extreme case, all n numbers may be the same. The average of

15, 15, 15, 15, 15

is of course 15, and the average of any number n of equal items c is just

$$\frac{\overbrace{c + c + c + \ldots + c}^{n \text{ terms}}}{n} = \frac{nc}{n} = c.$$

Another obvious remark is that if any of the numbers is replaced by a larger one, the average $\bar{x}$ will be increased. Repeated use of this principle shows that $\bar{x}$ is less than the largest number in the list.

For example, in the list $(x_1, x_2, x_3, x_4) = (3, 1, 5, 2)$ if we replace 1 by 5, the largest number, we get the new list $(3, 5, 5, 2)$ whose average is larger than the original one. If we now replace 3 by 5 and 2 by 5, we obtain the list $(5, 5, 5, 5)$ whose average is certainly 5. We conclude that the original $\bar{x} < 5$.

More generally, let M (for maximum) be the largest number in the list. Let x be any smaller number. If we replace x by M, the average will be increased. If we repeat this process for <u>each</u> x less than M, we will finish with n equal numbers M. Their average is M. Since at each step we increased the average, the result M is bigger than the original $\bar{x}$. (Of course if all of the numbers are equal to begin with, any one of them can be called M and $\bar{x} = M$.)

It is equally clear that if any entry is replaced by a <u>smaller</u> number, the average will be <u>decreased</u>; and consequently $\bar{x}$ is greater than m (for <u>minimum</u>), the smallest number in the list. Thus

$$m < \bar{x} < M.$$

If we wish to include the possibility that all the given numbers are equal, we have

$$m \le \bar{x} \le M$$

where the equal signs can be used only when $m = M$.

Let us return to the problem with which we began. You will remember that the average of 130, 120, 135, 145, 140 turned out to be 134. We now introduce an idea which often simplifies the arithmetic. Let us consider the new list of numbers

$$10, \ 0, \ 15, \ 25, \ 20$$

obtained by subtracting the smallest one $(m = 120)$ from each of the original list. The new average is

$$\frac{10 + 0 + 15 + 25 + 20}{5} = \frac{70}{5} = 14.$$

If we add 14 to 120 we obtain 134, the average found before. Is this an accident?

Instead of subtracting the smallest number m, let us subtract 130 from each of the original numbers. We may think of 130 as a guess at the required average. The new list is

$$0,\ -10,\ 5,\ 15,\ 10$$

whose average is $\frac{20}{5} = 4$. Now $130 + 4 = 134$ is the average of the original list of numbers.

If we guess 135 for the average and subtract this from each number in the original list, we obtain the numbers

$$-5,\ -15,\ 0,\ 10,\ 5$$

whose average is $\frac{-5}{5} = -1$. Then $135 + (-1) = 134$ is the actual average.

Let c be the number which we subtract from each of those in the list. You will notice that the closer c is to $\bar{x}$, the smaller is the new average to be obtained. The arithmetic to be performed is usually easier. Of course, negative numbers may occur, but often they can be quickly combined with positive numbers.

Let us state the principle which seems to be behind these examples. Let c be any number. If we subtract c from a number x, we obtain $x - c$. The average of these diminished numbers can be written $\overline{(x - c)}$. In our examples we have seen that

$$\bar{x} = \overline{(x - c)} + c .$$

This equation is equivalent to

(1) $$\overline{(x - c)} = \bar{x} - c .$$

Can we prove that (1) is always true?

What is $\overline{(x - c)}$? It is

$$\frac{(x_1 - c) + (x_2 - c) + \dots + (x_n - c)}{n}$$

But this is

$$\frac{x_1 + x_2 + \dots + x_n}{n} - \frac{\overbrace{c + c + \dots + c}^{n\ c\text{'s}}}{n} = \bar{x} - c .$$

Exercises 25-4

1. We wish to average the following numbers

$$98, \ 93, \ 101, \ 110, \ 95.$$

(a) Do this directly from the definition.

(b) Now subtract a suitable number from each item of the list and apply equation (1). Compare the amount of arithmetic which must be performed in the two cases.

2. Use the method of this section to average

$$10.1, \ 10.3, \ 9.9, \ 10.8, \ 9.1.$$

3. In the example in the text, suppose that we choose $c = 134$ to subtract. What is the new list? What is the average of the new list?

4. In equation (1), let $c = \overline{x}$, the actual average of the x's. What is $\overline{(x - \overline{x})}$? Compare your result with Exercise 3.

5. Compare the average of $(2, 6, 16, 8, 10)$ with the average of $(1, 3, 8, 4, 5)$.

6. Show that $\overline{ax}$, that is,

$$\frac{ax_1 + ax_2 + \dots + ax_n}{n}$$

is equal to $a \cdot \overline{x}$.

7. What is the average of the number of spots on the faces of a die (die is the singular of "dice")?

What then is the average of 10, 20, 30, 40, 50, 60 ?

of 100, 200, 300, 400, 500, 600 ?

of .1, .2, .3, .4, .5, .6 ?

8. Find the average of

$$10.01, \ 10.02, \ 10.03, \ 10.04, \ 10.05, \ 10.06.$$

9. List the dates of birth for the members of your class. Find the average year of birth as easily as you can.

10. List the age in years for each member of your class on his or her last birthday. Find the average age. If you add this average age to

the answer of Exercise 9 will you get the present year? Explain any difference that you discover.

25-5. A Picture of $\bar{x}$

Let us represent the numbers x_1, x_2, ..., x_n and their average on the number line. This is done for 0, 1, 2, 3, 4, 5 in Figure 4. We have put an arrowhead at the point $\bar{x} = 2.5$ to suggest a fulcrum or a knife edge. In fact, if we imagine that equal weights are placed at the given points, there will be a balance about the point $\bar{x}$.

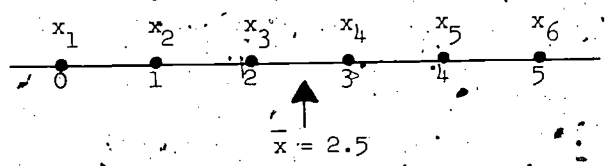


Figure 4

Another illustration is shown in Figure 5.

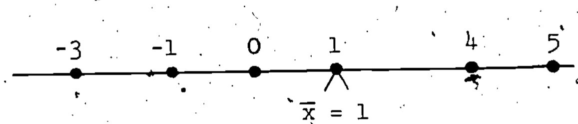


Figure 5

In every case, the fulcrum at $\bar{x}$ lies between the left-most point and the right-most point.

We can see very easily from the picture that the equation

$$(1) \qquad \overline{(x - c)} = \bar{x} - c$$

is true. Suppose that a set of points balances about a fulcrum at $\bar{x}$. Now imagine that the coordinate of each point is decreased by c but the points remain unchanged in position. The bar will balance at the same place as before. The bar neither knows nor cares how we name the points. Since the point previously called $\bar{x}$ now has the new coordinate $\bar{x} - c$, equation (1) makes sense.

For example, if we decrease each coordinate in Figure 5 by $c = 2$ we obtain the numbers shown on the lower scale in Figure 6 and

$$\overline{(x - 2)} = \bar{x} - 2 = 1 - 2 = -1.$$

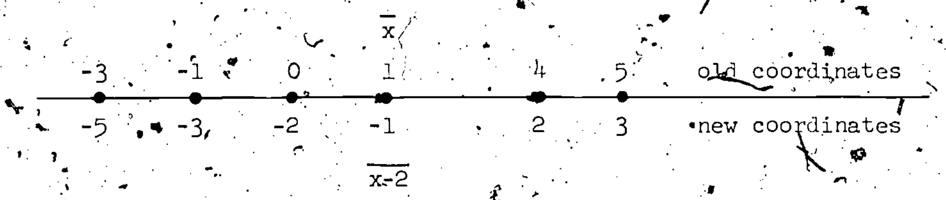


Figure 6

We saw in Exercise 6 that

$$\overline{ax} = a \cdot \overline{x}.$$

How can we picture this result? Simply in terms of a scale change! To be more definite, let $a = 3$. If we choose a new length unit which is $\frac{1}{3}$ of the original unit, each point will receive a new coordinate 3 times as large as before. This will not change the position of the fulcrum. The point that was called $\overline{x}$ is now called $3\overline{x}$. Consequently

$$\overline{3x} = 3\overline{x}.$$

Thus the average of $(\frac{3}{2}, 3, 6, 7\frac{1}{2})$ is three times the average of $(\frac{1}{2}, 1, 2, 2\frac{1}{2})$.

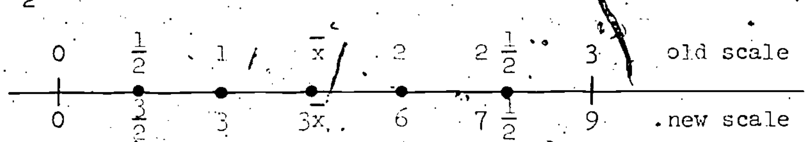


Figure 7

(A similar argument can be made for a different choice of a.)

Suppose that we have a long list of numbers to average, say the first 20 primes.

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29,$$
$$31, 37, 41, 43, 47, 53, 59, 61, 67, 71$$

It is easy to make a mistake in adding a long list of numbers. Instead of adding these numbers and dividing by 20, let us take them in groups of 10. The average of the first ten numbers is $\frac{129}{10} = 12.9$.

The remaining ten numbers average $\frac{510}{10} = 51$.

Suppose now that we average these two averages, obtaining

$$\frac{12.9 + 51}{2} = \frac{63.9}{2} = 31.95 .$$

Is this the same as the average of the 20 numbers? Yes, because

$$\frac{1}{2}\left(\frac{2 + 3 + \ldots + 29}{10} + \frac{31 + 37 + \ldots + 71}{10}\right)$$
$$= \frac{1}{2}\left(\frac{2 + 3 + \ldots + 29 + 31 + 37 + \ldots + 71}{10}\right)$$
$$= \frac{(2 + 3 + \ldots + 71)}{20}$$

The argument can be made more general. If

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

and

$$\overline{y} = \frac{y_1 + y_2 + \ldots + y_n}{n}$$

then

$$\frac{\overline{x} + \overline{y}}{2} = \frac{x_1 + x_2 + \ldots + x_n + y_1 + \ldots + y_n}{2n}$$

which is the average of the $2n$ numbers obtained by putting the x's and y's together in a single list.

It is important to remember that the number of x's is the same as the number of y's. If, for example, we take

$$1, 3, 5, 7$$

whose average is 4 and

$$2, 4$$

whose average is 3, the average of the combined list

$$1, 2, 3, 4, 5, 7$$

is not $\frac{4 + 3}{2} = \frac{7}{2}$ but $\frac{22}{6} = \frac{11}{3}$. The two lists do not contain the same number of items!

The principle that we have discovered may be given a slightly different twist. Instead of writing

$$(1) \qquad \frac{\overline{x} + \overline{y}}{2} = \frac{x_1 + x_2 + \ldots x_n + y_1 + \ldots + y_n}{2n}$$

we may write

$$\overline{x} + \overline{y} = \frac{x_1 + x_2 + \ldots + x_n + y_1 + \ldots + y_n}{n}$$

$$= \frac{(x_1 + y_1) + (x_2 + y_2) + \ldots + (x_n + y_n)}{n}$$

The right side is the average of the sums obtained by adding a number x to a number y. It can naturally be written as $\overline{(x + y)}$.

Hence

(2) $$\overline{x + y} = \overline{x} + \overline{y}$$

We can use this form (2) of the principle to find some other averages rather easily.

For example, suppose that we wish to average monthly automobile expenses of $53.20; $72.13; $38.42; $84.59; $27.62 over a five-month period. Let us think of 53, 72, 38, 84, 27 as the x's and .20, .13, .42, .59, .62 as the y's. Since $\overline{x} = 54.8$ and $\overline{y} = .39$, we obtain

$$\overline{x + y} = \overline{x} + \overline{y} = 54.8 + .39 = 55.19.$$

You should check this result by averaging the actual expenses.

However, if we do not require too great accuracy we can simply calculate $\overline{x}$, the average of the number of dollars and forget about $\overline{y}$. How much error could be committed if we did this? In the given problem, the average $\overline{y}$ cannot possibly exceed .62, the largest item. In no case could $\overline{y}$ exceed 1. Consequently we would never be more than $1 off.

Exercises 25-5

1. Illustrate on the number line that

$$\overline{(x - c)} = \overline{x} - c$$

with $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$, $x_5 = 5$ and $c = 1$.

2. Repeat Exercise 1 with $c = 2$ and $c = 3$.

3. Is it true that

$$\overline{(x + c)} = \overline{x} + c$$

where c is positive? Why? Interpret the result on the number line.

4. In showing on the number line that $\overline{ax} = a \cdot \overline{x}$, we chose the positive number $a = 3$. Is $\overline{(-1)\,x} = (-1)\,\overline{x}$? Interpret this equation on the number line.

5. Separate the list of 20 primes in this section into four lists of five primes each. Find the average for each set of five and the average of these averages. What conclusion do you draw?

6. Find the average of

$$50.03, \quad 64.01, \quad 39.06 \text{ and } 17.10$$

by using the method of this section. Show how this result could be estimated quickly with considerable accuracy.

7. In the city of Hioho there were two schools. In one school 80 percent of the students passed a nationally set test. In the other school 90 percent passed. It was argued that in the city 85 percent passed. Is the argument a valid one? Explain.

8. You have 10 numbers whose average is 35.5 and 4 numbers whose average is 43. You have lost the original lists so that this is all that you know. What is the average of the 14 items? (Hint: What was the sum of the 10 numbers? The sum of the 4 numbers? of the 14 numbers?)

9. If we have a list of n numbers with average $\overline{x}$ and another list of m numbers with average $\overline{y}$, what is the average $\overline{z}$ of the combined list of $n + m$ numbers? (Hint: Generalize the method of Exercise 8.)

Answer: $\overline{z} = \dfrac{n\overline{x} + m\overline{y}}{n + m}$.

What is $\overline{z}$ if $m = n$?

10. Show that if you have a set of n numbers with average $\overline{x}$, and add a number equal to $\overline{x}$, the new average remains unchanged. Use the result of Exercise 9.

25-6. Equally Spaced Numbers

If the numbers to be averaged are equally spaced when arranged in order of size, the averaging process becomes much simpler. Let us see if we can discover this simpler way of finding the average.

Exercises 25-6a

(Class Discussion)

Suppose that we wish to find the average or mean of the numbers

$$17,\ 21,\ 25,\ 29,\ 33,\ 37,\ 41.$$

1. What number do you guess the average to be?
2. Subtract your guess from each number in the list and average the new list.
3. If you guessed the average correctly what answer should you obtain for the new average found in Exercise 2? If you did not get this result, guess again.
4. Where is the correct average located in the given list?
5. If the given numbers are marked on the number line, where is the point $\bar{x}$ located? How can you find this point by using only the positions of the first and last point?
6. Let f be the first and ℓ the last of any set of numbers which are equally spaced when arranged in order. On the number line where is $\bar{x}$ located in relation to f and ℓ?
7. How does the distance between $\bar{x}$ and ℓ compare with the distance between f and $\bar{x}$?
8. Express $\bar{x}$ in terms of f and ℓ.

If all went well you should have found that $\bar{x} = \frac{f + \ell}{2}$ so that $\bar{x}$ may be obtained without using any of the numbers between the first and the last. How can we be sure that this conclusion is correct?

If we represent the equally spaced numbers by beads on the number line, it is clear that the beads balance about the midpoint of the interval $[f, \ell]$. Each bead on the left of the midpoint is balanced by a bead at the same distance on the right of the midpoint.

We conclude that

$$\bar{x} - f = \ell - \bar{x}$$

and therefore

$$\bar{x} = \frac{f + \ell}{2} .$$

We see that to average a set of equally spaced numbers it is sufficient to average the first and the last.

A particular case of this short cut concerns the average of the first n positive integers

$$1, 2, 3, \ldots, n .$$

The result is simply $\frac{n + 1}{2}$. We write

$$\bar{i} = \frac{n + 1}{2} .$$

(i stands for integer.)

Exercises 25-6b

1. Locate the following numbers on the number line

$$-5, -2, 1, 4, 7, 10 .$$

   (a) Find $\bar{x}$ by the straightforward method and by the short cut.

   (b) Show on the figure how the points balance in pairs about the point $\bar{x}$.

   (c) List the pairs of numbers which balance and show that their average is $\bar{x}$ in each case.

2. Repeat Exercise 1, using the numbers

$$7, 12, 17, 22, 27, 32, 37 .$$

3. If i takes on all values from 1 to n ; find $\overline{2i + 1}$ in two different ways: First, as a special case of an equally spaced set of numbers; secondly, by using the value of $\bar{i}$ and the fact that $\overline{2i + 1} = 2\bar{i} + 1$.

4. Show that if i takes on all values from 0 to $n - 1$, then

$$\bar{i} = \frac{n - 1}{2} .$$

5. Show that any list of n equally spaced numbers can be written

$$f, \ f + d, \ f + 2d, \ \ldots, \ f + id, \ \ldots, \ f + (n - 1)d,$$

   where d is the common difference and $f + (n - 1)d$ is the last number, and that their average is $\frac{f + \ell}{2}$.

6. When the famous mathematician Gauss (1777-1855) was only 10 years old, his arithmetic class was set the problem of adding something like 8129, 8149, ... , 10109. There were 100 equally spaced items. The pupils' slates were laid on the teacher's desk when they were done. The teacher had barely finished stating the problem when Gauss put his slate on the desk. When the others had finished at the end of the period, Gauss, who had simply written the answer, was the only one who was correct. How did he get the answer so quickly? What was his answer?

25-7. On Averaging Squares

Let us turn to a somewhat more difficult problem. What is the average of the first n squares

$$1,\ 4,\ 9,\ \ldots,\ n^2\ ?$$

Could the answer be $\frac{1 + n^2}{2}$? Let us test this guess with $n = 2$ and $n = 3$. The correct averages are

$$\frac{1 + 4}{2} = \frac{5}{2} \quad \text{and} \quad \frac{1 + 4 + 9}{3} = \frac{14}{3}\,.$$

With $n = 2$

$$\frac{1 + n^2}{2} = \frac{1 + 4}{2} = \frac{5}{2}\,, \text{ which is correct.}$$

However, with $n = 3$

$$\frac{1 + n^2}{2} = \frac{1 + 9}{2} = 5$$

gives an incorrect answer.

This result should not surprise us because the squares are not equally spaced. The short cut is based on the idea that the numbers <u>are</u> equally spaced.

Can we find a correct formula for

$$\frac{1 + 4 + 9 + \ldots + n^2}{n}\ ?$$

We can at least write down the correct averages for the first few values of n and hope to discover a pattern.

For $n = 1$, the average is 1. For $n = 2$ and $n = 3$, we have found the averages $\frac{5}{2}$ and $\frac{14}{3}$. For $n = 4$, we have $\frac{1 + 4 + 9 + 16}{4} = \frac{30}{4} = \frac{15}{2}$; and for $n = 5$, $\frac{1 + 4 + 9 + 16 + 25}{5} = 11$.

Let us make a little table. In the first column we give the value of n; in the second, the average of the squares of the first n integers which we write $\overline{i^2}$ (where i goes from 1 to n).

n	$\overline{i^2}$
1	1
2	$\frac{5}{2}$
3	$\frac{14}{3}$
4	$\frac{15}{2}$
5	11

The pattern is not obvious. However, let us write a third column which gives the averages of the first n integers (without squaring), that is,

$$\overline{i} = \frac{1 + 2 + \ldots + n}{n}.$$

n	$\overline{i^2}$	$\overline{i}$
1	1	1
2	$\frac{5}{2}$	$\frac{3}{2}$
3	$\frac{14}{3}$	2
4	$\frac{15}{2}$	$\frac{5}{2}$
5	11	3

Which set of numbers increases most rapidly, $\overline{i^2}$ or $\overline{i}$? Clearly $\overline{i^2}$. We can compare $\overline{i^2}$ with $\overline{i}$ by dividing. If we do this, and write the results in a fourth column, we can make a nice discovery.

n	$\overline{i^2}$	$\overline{i}$	$\overline{i^2}/\overline{i}$
1	1	1	1
2	$\frac{5}{2}$	$\frac{3}{2}$	$\frac{5}{3}$
3	$\frac{14}{3}$	2	$\frac{7}{3}$
4	$\frac{15}{2}$	$\frac{5}{2}$	3
5	11	3	$\frac{11}{3}$

Exercises 25-7a

(Class Discussion)

1. By what should we replace 1 in the last column to bring out the pattern? By what should we replace 3?

2. What general rule includes all the items in the last column? That is, how can you express $\overline{i^2}/\overline{i}$ in terms of n?

3. Test your rule by finding $\overline{i^2}$, $\overline{i}$ and $\overline{i^2}/\overline{i}$ for $n = 6$.

4. What is the general rule for $\overline{i}$?

5. Therefore, what is the apparently correct rule for $\overline{i^2}$? You should find $\overline{i^2} = \frac{(n+1)(2n+1)}{6}$.

Exercises 25-7b

1. Is $\overline{i^2} = (\overline{i})^2$?

2. Find in succession

   1^3; $1^3 + 2^3$; $1^3 + 2^3 + 3^3$; $1^3 + 2^3 + 3^3 + 4^3$;

   $1^3 + 2^3 + 3^3 + 4^3 + 5^3$ and discover a general rule for $\overline{i^3}$.

   Hint: Compare these sums with 1, $1 + 2$, $1 + 2 + 3$, $1 + 2 + 3 + 4$, $1 + 2 + 3 + 4 + 5$.

3. In the Class Discussion you discovered the apparently correct rule for $\overline{i^2} = \frac{1^2 + 2^2 + 3^2 + \ldots + n^2}{n}$. It is

$$\text{(1)} \qquad \overline{i^2} = \frac{(n+1)(2n+1)}{6} = \frac{2n^2 + 3n + 1}{6}.$$

You have verified the correctness of this rule as far as $n = 6$. This does not prove that the rule is correct for larger values of n. Prove that if the rule is correct for any particular integer n, then it is correct for the next larger integer $n + 1$.

Hints: Given (1), what ~~is~~ $1^2 + 2^2 + \ldots + n^2$?

Add $(n + 1)^2$ to this result.

Show that

$$\frac{1^2 + 2^2 + \ldots + n^2 + (n + 1)^2}{n + 1} = \frac{(n + 2)(2n + 3)}{6}$$

which means that Equation (1) is correct when n is replaced by $n + 1$.

25-8. Measures of Scatter

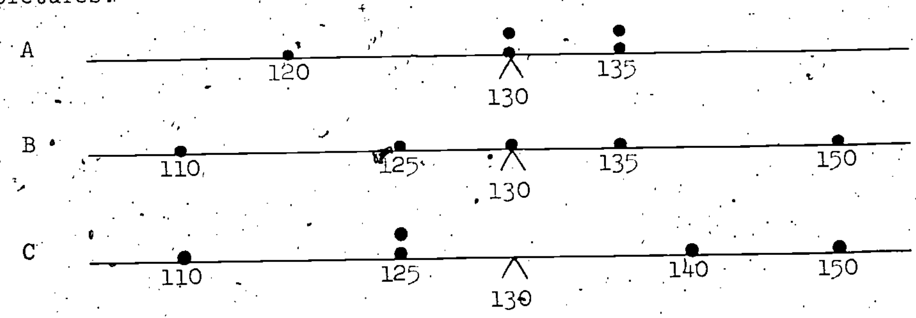
Two lists of numbers may have the same average but otherwise be very different. For example, suppose that bowling teams A, B and C have the following sets of scores:

A. 135, 130, 130, 135, 120

B. 110, 125, 130, 135, 150

C. 110, 125, 125, 140, 150

In each case, $\bar{x}$, the average, is 130. On the number line we have the following pictures:



It is clear that these pictures differ in the extent to which the points are spread out. In A the points are least spread out or scattered. In C the points are most spread out. It is useful to have a measure of the amount of this spread or <u>scatter</u> as it is called. We may also think of this as a measure of the degree of consistency of the scores.

There are several ways to measure scatter. The crudest measure is the <u>range</u>. This is the distance between the most widely separated points. In the examples, the ranges are:

A : 15

B : 40

C : 40

Both B and C have the same range. However, the scatter for C is greater than that for B. As a measure of scatter the range has the disadvantage that it <u>depends on only two of the numbers</u>.

If we subtract the mean value or average, 130, from each of the lists, we have the new lists:

A. 5, 0, 0, 5, -10

B. -20, -5, 0, 5, 20

C. -20, -5, -5, 10, 20

These numbers are called the <u>deviations</u> from the mean for A, B and C. The better measures of scatter use <u>all</u> of the deviations. If we average the deviations we shall of course get 0. For example, in A

$$\frac{5 + 0 + 0 + 5 + (-10)}{5} = \frac{0}{5} = 0.$$

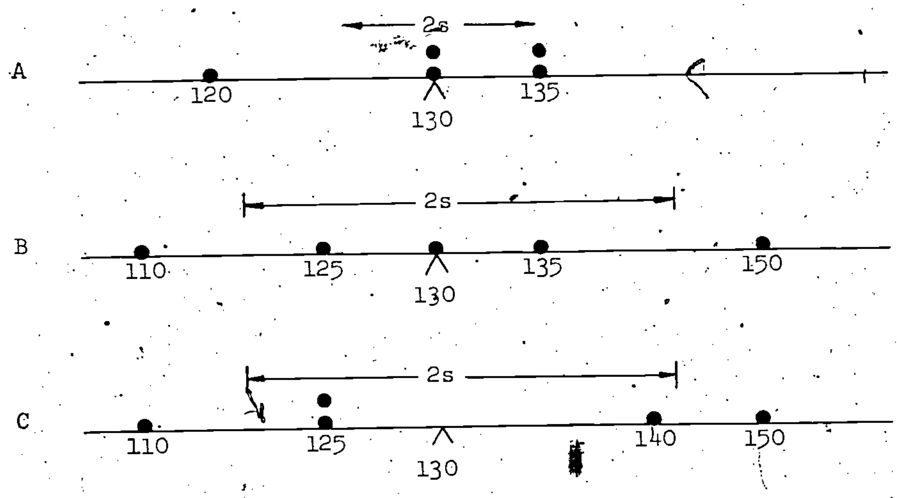
This average is therefore no measure of scatter. If, however, we average the <u>absolute values</u> of the deviations we <u>do</u> get a measure of scatter. We shall call the absolute value of a deviation its <u>distance</u>, since geometrically it represents the distance on the number line between a given point and the point $\overline{x}$. We shall use $\overline{d}$ to represent the average distance. The results for our examples are:

	distances (d)	av. dist. ($\overline{d}$)
A.	5, 0, 0, 5, 10	4
B.	20, 5, 0, 5, 20	10
C.	20, 5, 5, 10, 20	12

As you see, the values of $\bar{d}$ allow us to arrange our examples in what seems to be the right order of scatter.

Statisticians use a third measure of scatter called the <u>standard deviation</u> (s). This measure is formed as follows. The squares of the deviations (or, what is the same, the squares of the distances) are averaged. Then the square root of this average is taken to bring us back to a number comparable to the distances. For our examples we have

	Distances squared (d^2)	$\overline{d^2}$	$s = \sqrt{\overline{d^2}}$
A.	25, 0, 0, 25, 100	30	5.48
B.	400, 25, 0, 25, 400	170	13.04
C.	400, 25, 25, 100, 400	190	13.8



In each figure we have marked an interval of length 2s centered at $\bar{x} = 130$. The size of this interval gives some picture of the amount of scatter.

Why do statisticians prefer s to $\bar{d}$ as a measure of scatter? For one thing, because squares are easier to work with than absolute values. We also point out that the standard deviation s emphasizes large distances more strongly than the average distance $\bar{d}$ does. That is, the squares of the larger values of d have more effect than the squares of the smaller ones.

We can illustrate this remark by considering a list of 99 0's and one 100. $\bar{x} = 1$. The average distance

$$\bar{d} = \frac{\overbrace{1 + 1 + \dots + 1}^{99} + 99}{100} = 1.98$$

The square of the standard deviation

$$s^2 = \frac{\overbrace{1 + 1 + \dots + 1}^{99} + 9801}{100} = 99$$

so $s \approx 10$. As you see, s is about five times as large as $\bar{d}$, so that the larger number 100 produces a greater effect on s than on $\bar{d}$.

Exercises 25-8

1. Given the numbers

$$-4, -1, 0, 0, 5.$$

Find $\bar{x}$, $\bar{d}$ and s.

2. Find $\bar{d}$ and s for

(a) $-40, -10, 0, 0, 50$

(b) $-8, -2, 0, 0, 10$

Compare your results with those in Exercise 1.

3. Given any list of numbers

$$x_1, x_2, \dots, x_n \text{ with } \bar{x} = 0,$$

and standard deviation s, find the standard deviation for

$$ax_1, ax_2, \dots, ax_n$$

where a is any positive number. Check your result by referring to Exercise 2.

4. What are $\bar{d}$ and s for

$$4, 1, 0, 0, -5$$

and

$$8, 2, 0, 0, -10?$$

Compare with Exercise 1.

5. In Exercise 3 we took a to be positive. What are $\bar{d}$ and s if a is negative? Test your conclusion from Exercise 4.

6. Find the average or mean of

$$-4,\ 0,\ 2,\ 5,\ 7.$$

List the distances d of these points from the mean. Compare $\overline{d}$ and s.

25-9. Computing s^2

The numbers that can appear on the upper face of a die are

$$1,\ 2,\ 3,\ 4,\ 5,\ 6.$$

Their mean $m = 3.5$. What is the standard deviation s? If we subtract 3.5 from each of our numbers we have the deviations

$$-2.5,\ -1.5,\ -.5,\ .5,\ 1.5,\ 2.5.$$

The distances d from the mean are of course all positive

$$2.5,\ 1.5,\ .5,\ .5,\ 1.5,\ 2.5.$$

To find s, we square the deviations or the distances, average these squares and take the square root. The values of d^2 are

$$6.25,\ 2.25,\ .25,\ .25,\ 2.25,\ 6.25$$

whose average is $s^2 = \frac{17.50}{6} = 2\frac{11}{12} \approx 2.92$. Hence $s \approx \sqrt{2.92} \approx 1.70$.

If the mean m of a list of numbers is not an integer, the calculation of s^2 may involve a considerable amount of arithmetic. As we know,

$$s^2 = \overline{(x - m)^2}$$

where m is the mean. It often saves time to calculate $\overline{(x - a)^2}$ where a is an integer close to the mean. What is the connection between $\overline{(x - a)^2}$ and $\overline{(x - m)^2}$?

Since $(x - a)^2 = x^2 - 2ax + a^2$, $\overline{(x - a)^2} = \overline{x^2} - 2a\overline{x} + a^2$.

If we replace $\overline{x}$ by m, we have

(1) $$\overline{(x - a)^2} = \overline{x^2} - 2am + a^2.$$

This result is true for any number a. In particular, if $a = m$, the equation becomes $\overline{(x - m)^2} = \overline{x^2} - 2m^2 + m^2$

or

(2) $$\overline{(x - m)^2} = \overline{x^2} - m^2.$$

If we subtract (2) from (1), we obtain

$$\overline{(x-a)^2} - \overline{(x-m)^2} = m^2 - 2ma + a^2 = (m-a)^2.$$

This means that

(3) $$\overline{(x-m)^2} = \overline{(x-a)^2} - (m-a)^2.$$

That is, if we have calculated $\overline{(x-a)^2}$, we merely subtract $(m-a)^2$ from the result to obtain the desired $\overline{(x-m)^2} = s^2$.

Let us see how this works in practice.

<u>Example 1.</u> To find s^2 for

$$1,\ 2,\ 3,\ 4,\ 5,\ 6.$$

We have already found $s^2 = 2\,\frac{11}{12}$ by averaging the squares of the deviations from the mean, $m = 3\,\frac{1}{2}$.

Let us obtain this result by applying

(3) $$s^2 = \overline{(x-a)^2} - (m-a)^2 \quad \text{with} \quad a = 3$$

so that $s^2 = \overline{(x-3)^2} - (m-3)^2$.

$$\overline{(x-3)^2} = \frac{(-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 + 3^2}{6} = \frac{19}{6} = 3\,\frac{1}{6}.$$

$$(m-3)^2 = (3\,\tfrac{1}{2} - 3)^2 = (\tfrac{1}{2})^2 = \tfrac{1}{4}.$$

Therefore

$$s^2 = 3\,\tfrac{1}{6} - \tfrac{1}{4} = 2\,\tfrac{11}{12}.$$

<u>Example 2.</u> To find s^2 for the following 24 numbers:

16, 18, 19, 19, 20, 20, 22, 22, 22, 23, 24, 25,
25, 25, 26, 27, 28, 30, 30, 31, 32, 34, 37, 38.

The value of m is $25\,\frac{13}{24}$. Let us replace this by $a = 25$. The differences, $x - 25$, are the following:

-9, -7, -6, -6, -5, -5, -3, -3, -3, -2, -1, 0,
0, 0, 1, 2, 3, 5, 5, 6, 7, 9, 12, 13.

It is easy to find $\overline{(x-25)^2}$, the average of the squares of these numbers. It is $\frac{827}{24} = 34\,\frac{11}{24}$.

Since $m - 25 = 25\frac{13}{24} - 25 = \frac{13}{24}$, we have

$$s^2 = \overline{(x-25)^2} - \overline{(m-25)}^2 = 34\frac{11}{24} - (\frac{13}{24})^2.$$

Now $(\frac{13}{24})^2 = \frac{13}{24} \times \frac{13}{24} = \frac{169}{24} \times \frac{1}{24} \approx 7 \times \frac{1}{24} = \frac{7}{24}$.

Hence $s^2 \approx 34\frac{11}{24} - \frac{7}{24} = 34\frac{4}{24} = 34\frac{1}{6}$.

From this we can find $s \approx 5.85$.

Exercises 25-9

1. Find s^2 for 21, 25, 20, 30, 24 directly, that is, by averaging the squares of the deviations from the mean. Plot the numbers on the number line and show an interval $[m - s, m + s]$ of length $2s$ and center at m. What fraction of the points lie within this interval?

2. Find s^2 for 6, 8, 11, 12 in two ways:

   (a) by averaging the squares of the deviations from the mean.

   (b) by replacing the mean by the approximation $a = 9$ and applying equation (3).

   Compare the amount of arithmetic required in (a) and (b). Plot the numbers on the number line and show the interval $[m - s, m + s]$.

3. In Exercise 2, you found $m = \frac{37}{4}$ and $s^2 \approx 5.69$. Use these results to find m and s^2 for 24, 32, 44, 48. If necessary refer back to Exercises 25-8, Problem 3.

4. Given .28, .32, .35, .31. Find m and s^2 using any short cuts that you can.

5. Given 1, 2, 7, 11. Find m and s^2 directly and by using equation (3) with a suitable number $\underline{a}$.

6. In Example 2 in the text, show as simply as possible that $m = 25\frac{13}{24}$. Verify in detail the computations for s^2.

7. Draw a dot frequency diagram from Example 2 in the text and mark the interval $[m - s, m + s]$. What percent of the points lie within this interval?

8. List the ages of each of the members of your class in years at the last birthday. Find the mean m and the standard deviation s. Make a dot frequency diagram and show the interval from $m - s$ to $m + s$.

9. List the weights in pounds of the members of your class and find m and s. Make a dot frequency diagram and show the interval $[m - s, m + s]$. Comment on the difference between the results here and in Exercise 8.

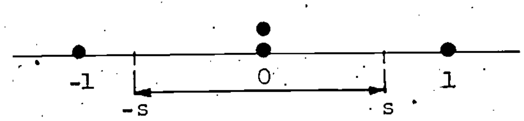
10. Find m and s^2 for $1, 2, 3, \ldots, 11$. Locate the points on the number line and show the interval $[m - s, m + s]$. What fraction of the points lie in this interval?

<u>Hint</u>: Remember that for $1, 2, 3, \ldots, n$, $m = \overline{i} = \frac{n + 1}{2}$ and $\overline{i^2} = \frac{(n + 1)(2n + 1)}{6}$, and use equation (2) in this section.

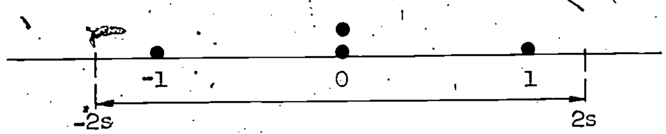
25-10. <u>What Does the Standard Deviation Tell Us?</u>

In the Exercises in Section 25-9, you have located the mean m on the number line and shown an interval which extends a standard deviation on each side of it. In some cases you were asked to find the fraction or percent of the points that lie inside this interval. Let us study a few more examples to see whether we can draw a general conclusion.

<u>Example 1.</u> Consider the list of numbers $-1, 0, 0, 1$ whose mean is $m = 0$. Here $s^2 = \frac{1 + 0 + 0 + 1}{4} = \frac{1}{2} = .5$ and $s \approx .7$. On the number line we have the following picture on which we have marked the interval $[m - s, m + s] = [-s, s]$.

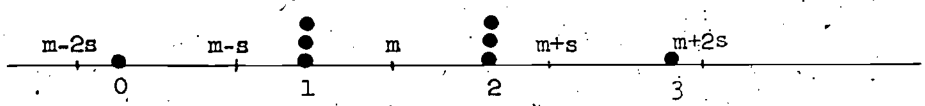


One half of the points lie in this interval. <u>All</u> four points lie in the interval $[-2s, 2s]$ which is approximately $[-1.4, 1.4]$.



<u>Example 2.</u> For the numbers $0, 1, 1, 1, 2, 2, 2, 3$ the mean $m = 1.5$ and the standard deviation $s = \frac{\sqrt{3}}{2} \approx \frac{1.73}{2} \approx .87$, as you should verify. A dot

frequency diagram shows the location of the numbers on the number line.



We have marked the intervals $[m - s, m + s]$ and $[m - 2s, m + 2s]$. As you see, 6 of the points are closer to m than 1 standard deviation and all of them are within $2s$ of m. That is, 75 percent are within the distance s and 100 percent are within the distance $2s$ of m.

The numbers that we have chosen can be imagined as winnings in a game in which 3 coins are tossed with 1¢ for a head H and 0 for a tail T. The possible outcomes are

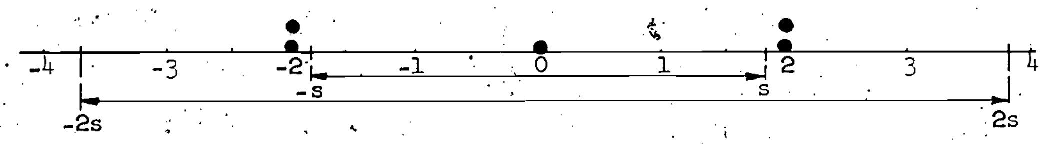
H H H	3
H H T	2
H T H	2
H T T	1
T H H	2
T H T	1
T T H	1
T T T	0

The second column shows the amount won for each of these outcomes. The different outcomes are equally likely if the coins and the method of tossing are honest. Therefore we expect that the average winning per throw is the mean (1.5) of the numbers

$$3, 2, 2, 1, 2, 1, 1, 0.$$

Also as we have found, 75 percent of the time we expect to win either 1 or 2 cents.

Example 3. We get a different picture with the numbers $-2, -2, -2, 0, 2, 2, 2$. You should verify that $m = 0$ and $s^2 = 3\frac{3}{7} \approx 3.42$ and therefore $s \approx 1.85$. The figure is a dot frequency diagram which shows the intervals $[-s, s]$ and $[-2s, 2s]$.



Only [illegible] of the points are within s of the mean. All of them are within $2s$ of the mean.

Most collections of numbers met in practice are more like Examples 1 and 2 than like Example 3. In what is called a <u>normal</u> <u>distribution</u>, about 68 <u>percent of the numbers are</u> within the interval $[m - s, m + s]$ and about 94 percent within the interval $[m - 2s, m + 2s]$.

In the next section we shall learn about an important statement which is true for <u>every</u> set of numbers. It concerns the number of points which lie within $2s$ of the mean. In the following Exercises, we shall see what we can discover for ourselves.

<u>Exercises 25-10</u>

For each of the following sets of numbers

(a) find m and s;

(b) show the numbers on the number line;

(c) mark the intervals $[m - s, m + s]$ and $[m - 2s, m + 2s]$;

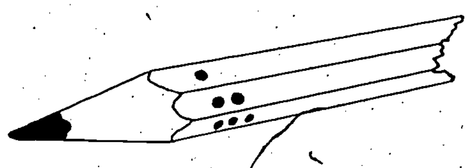
(d) find the percent of points inside each interval.

1. 1, 2, 2, 2, 3
2. 0, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4
3. -2, -1, 0, 1, 2
4. -2, -2, -2, 2, 2, 2
5. -3, 0, 0, 0, 0, 3
6. -1, 0, 0, 0, 0, 0, 0, 1
7. Is it possible to have no points inside the interval $[m - s, m + s]$? Is it possible to have all points inside this interval? Justify your answers.
8. What is s if all points are at the same distance (d) from the mean?
9. Is it possible that no points are within the interval $[m - 2s, m + 2s]$? Can all points be within this interval?
10. Suppose that $m = 0$ and $s = 1$. Is it possible that no more than half of the points lie in the interval $[-2, 2]$? If it helps to get started, begin with 8 points.

36

11. How large a fraction of a set of points must lie within $[m - 2s, m + 2s]$, do you guess?

12. Find a pencil that has six sides and mark it with dots as follows:



one dot on the first side, two dots on the second side, three dots on the third side, four dots on the fourth side, etc.

(a) Roll the pencil twenty times. Record the number of dots that come up on the top side of the pencil for each roll.

(b) Draw a dot frequency diagram of the data collected.

(c) Compute the mean and the standard deviation, using the method of this section.

(d) Show the mean and the 2s interval about the mean on the dot frequency diagram.

13. Each student should open his textbook, (any textbook), to a randomly selected page. Count the number of times the letter "e" occurs in the first complete sentence on the page. Continue to open the book and count until twenty numbers are recorded.

(a) Tally the frequencies for the data collected above.

(b) Draw a dot frequency diagram of the data collected.

(c) Compute the mean and the standard deviation, using the method of this section.

(d) Show the mean and the 2s interval about the mean on the dot frequency diagram.

14. Repeat the experiment described in Exercise 13 for the letter "a".

(a) How frequently did "a" occur exactly four times in a sentence? What was the relative frequency for "a" for a count of four? How does this compare with the relative frequency for the letter "e" for a count of four?

(b) What would you say is the most likely number of times the letter "e" would occur in a randomly selected English sentence?

(c) What would you say is the most likely number of times the letter "a" would occur in a randomly selected English sentence?

(d) Do you suppose that the relative frequencies are different for the letters "e" and "a" in a book written in a foreign language?

25-11. An Important Result

We found that it is possible to have no points inside the interval $[m - s, m + s]$. This comes about if half of the points are at $m - s$ and half at $m + s$.

Is it possible to have no points within $[m - 2s, m + 2s]$? Suppose that this were possible. Then for each of the n points the distance from m would be at least $2s$. That is, the square of each distance would be at least $4s^2$. The average square of the distance would then be $\overline{d^2}$

$$\overline{d^2} \geq 4s^2.$$

But $\overline{d^2} = s^2$ and we are led to the impossible conclusion that

$$s^2 \geq 4s^2$$

or

$$1 \geq 4!$$

There must therefore be at least one point within the double interval.

How many of the n points can fail to be inside the interval $[m - 2s, m + 2s]$? Let k be the unknown answer to this question. As in the argument just given, each of the k points would have a value of $d \geq 2s$ and therefore a value of $d^2 \geq 4s^2$. The sum of the squares for the k points therefore $\geq 4ks^2$. The sums of the squares of all the n distances can be no less. Therefore the average for all the points

$$(1) \qquad \overline{d^2} \geq \frac{k4s^2}{n}.$$

Since

$$\overline{d^2} = s^2, \quad s^2 \geq \frac{k4s^2}{n}.$$

Therefore

$$1 \geq \frac{4k}{n}$$

and so

(2) $$\frac{k}{n} \leq \frac{1}{4}.$$

The conclusion is that no more than a quarter of the points can be at a distance from the mean of $2s$ or more.

Is it possible for $\frac{k}{n}$ to be exactly <u>equal</u> to $\frac{1}{4}$? If you look at problem 6 of Exercises 25-9 you will see that the answer is "yes". In this problem $2s = 1$ and 6 of the 8 points are at 0 (inside the interval $[-1, 1]$), and 2 points are at the distance $2s = 1$ from the mean, $m = 0$.

If we examine the work above with the case $\frac{k}{n} = \frac{1}{4}$ in mind, we see that in this case sentence (1) becomes

$$\overline{d^2} \geq s^2.$$

Since $\overline{d^2} = s^2$, it is impossible for d to be $> 2s$ for any point. We must therefore have all k points at the distance $2s$. Moreover the remaining $n - k$ points must be at 0. So the equality in (2) occurs only in very rare circumstances. In any event, as we see, at least $\frac{3}{4}$ of the points must lie in the interval $[m - 2s, m + 2s]$.

This important theorem was discovered and proved by the Russian mathematician Chebychev (1821-1894). (The v is pronounced like f.) We shall soon see how important this theorem is.

Exercises 25-11

1. Test Chebychev's Theorem in Exercises 25-10, problems 12, 13, and 14.
2. In the argument given in the text, replace $2s$ by $3s$ and conclude that at most $\frac{1}{9}$ of the points lie at a distance from the mean of $3s$ or more.
3. In the argument given in the text, replace $2s$ by $4s$ and conclude that at most $\frac{1}{16}$ of the points lie at a distance from the mean of $4s$ or more.
4. From Exercise 3, prove that if you have 15 points, none of them can be as far as $4s$ away from the mean.

5. Generalize the previous three exercises to show that at most the fraction $\frac{1}{p^2}$ of a set of points are at the distance ps or more from the mean.

6. From Exercise 5 show that given n points with mean m and standard deviation s, at least

$$\left(1 - \frac{1}{p^2}\right) n$$

points lie in the interval $[m - ps, m + ps]$.

25-12. Using Statistics

Specialists on deciphering codes (cryptanalysts) have worked out tables to tell how often various letters of the alphabet occur in written English. The following table gives the percent for the different letters.

Letter	Percent	Letter	Percent	Letter	Percent
E	12.31	L	4.03	B	1.62
T	9.59	D	3.65	G	1.61
A	8.05	C	3.20	V	.93
O	7.94	U	3.10	K	.52
N	7.19	P	2.29	Q	.20
I	7.18	F	2.28	X	.20
S	6.59	M	2.25	J	.10
R	6.03	W	2.03	Z	.09
H	5.14	Y	1.88		

How could such a table possibly be constructed? Surely, no one has counted all the letters in all the books in the world written in English and determined what percent of these letters are E! Even if this were possible, what would this say about a page of English that you might write this evening?

Actually, of course, letter counts have been made for many samples of text. It is assumed that the results found for these samples are a reliable guide in estimating the occurrence of letters in a new sample of English of substantial size. Suppose that you had to decode the following message:

SZEV BLF VEVI DLIPVW LEG Z XLWV YVULIV

It is understood, of course, that the letters here correspond one-to-one to letters in the decoded message. We have made things easy for you by separating the message into words. If you count the number of times the various letters occur among these 30 letters you will get the following results:

B	1	K	1
C	1	L	5
D	1	T	1
E	2	U	1
F	2	V	7
G	1	W	2
I	3	Z	2

Therefore, the best guess is that V = E. With this possibility in mind look at the word

$$\text{VEVI} = E\ \_\ E\ \_$$

The letter E occurs twice in the coded message and I occurs three times. How about EVER or EVEN as a replacement for VEVI?

Now look at the last word.

$$\text{YVULIV} = \_\ E\ \_\ \_\ \left.\begin{matrix}N\\R\end{matrix}\right|E$$

The letter L occurs 5 times in the coded message. The most likely possibilities are T, A, or O. Hence we have

$$\text{YVULIV} = \_\ E\ \_\ \left.\begin{matrix}T\\A\\O\end{matrix}\right|\left.\begin{matrix}N\\R\end{matrix}\right]E$$

TNE looks impossible and TRE unlikely. Y and U occur only once. They probably represent consonants. How about BEFORE as a guess? We observe also that Z must be A or I since these are the only words with a single letter.

We shall leave the rest of the deciphering to you. When you have worked out the message you will find that the frequency of occurrence of the letters (after decoding) corresponds only roughly to the percents listed in the table. Is this surprising? What is the explanation? Clearly the sample was not large enough.

Exercises 25-12a

(Class Activity)

Each student should take a different page of a novel or magazine article and count the E's among the first 1,000 letters on the page. Keep a record of the count for each student.

Exercises 25-12b

1. Using the results for all the students in the class, what is the total number of letters counted and the total number of E's? Find the percent of E's and compare this with 12.31, as given in the table.

2. Find the average number of E's per hundred letters by averaging the percents found by each of the students. Does this result agree with that in Exercise 1?

3. Take the set of percents found by the different students as a set of numbers. Find the standard deviation of this set of numbers. If the mean found in Exercise 2 is not a very simple number you should use the short-cut of Section 25-9 to carry out the computation.

4. Read the following message written in code:

   VHH LI BRX FDQ JXHVV WKLV PHVVDJH

25-13. On the Behavior of Samples

In using statistics, it is important to know what can be concluded from a sample of a given population. If we choose 100 numbers at random from a population of 10,000 numbers, what does the sample tell us about the population? In particular, how close is the average of the 100 numbers in the sample to the average of the 10,000 numbers in the population?

To help us to answer such questions, let us start with a population that we know all about and investigate samples of different sizes which can be selected from it. To keep the arithmetic simple we shall take small populations.

Consider the population that consists of the following 6 numbers, which we have shown on the number line: 0, 4, 5, 8, 9, 10.

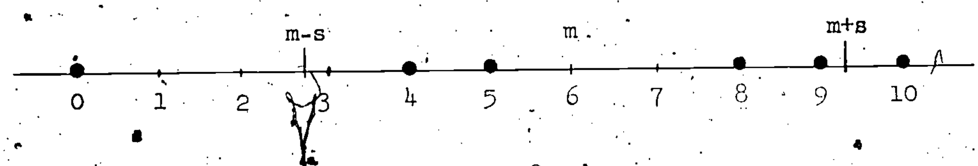


Figure 8

The mean $m = 6$ and the square of the standard deviation is $s^2 = \frac{35}{3}$ so that $s \approx 3.27$. The interval $[m - s, m + s]$ is shown in Figure 8.

From this given population, let us take all possible samples of size <u>two</u> and average the two numbers in each sample. For example, if we select the pair (4,9) we obtain the average $\frac{13}{2} = 6.5$. There are 15 different pairs which can be selected and therefore there are 15 averages. In Figure 9 we show these fifteen averages in a dot frequency diagram.

Figure 9

Let us call m_2 the mean of the averages of the fifteen samples of size <u>two</u>. You will find that $m_2 = 6$, the same as the population mean m. That is, $m_2 = m = 6$. We shall write the value of s^2 for the 15 sample means, $(s_2)^2 = s_2^2$. It turns out to be $5\frac{1}{15}$ so that $s_2 \approx 2.25$. The interval $[m_2 - s_2, m_2 + s_2]$ is shown on the figure. The points are somewhat less scattered than for the population. We see this by inspecting the figure and also by noting that $s_2 < s$.

Let us move on to samples of size 3, of which there are 20, and take the mean of the numbers in each sample. For example, the mean of the sample (4, 5, 10) is $\frac{19}{3} = 6\frac{1}{3}$. We show all of these means on the dot frequency diagram (Figure 10).

Figure 10

We shall call the average of these 20 means, m_3. You will find that m_3 is the same as the population average $m = 6$ and that for these points the square of the standard deviation is $s_3^2 = \frac{7}{3}$ so that $s_3 \approx 1.53$. Once again the interval $[m_3 - s_3, m_3 + s_3]$ is shown on the number line. Notice that $s_3 < s_2$.

When we find the mean of the numbers in a sample we replace these numbers by a single number at the center of the sample.

These center points form a tighter pattern than the set of points for the original population.

We could go on to find the average and standard deviation of all "sample means" of size 5. However, we shall be content to use the range which is a cruder measure of scatter than the standard deviation.

For samples of size 4, the least sample mean will come from the four left-most points 0, 4, 5, 8. This mean is $\frac{17}{4} = 4\frac{1}{4}$. The greatest sample mean of size 4 is the mean of 5, 8, 9, 10, the 4 right-most points. This mean is $\frac{32}{4} = 8$. The <u>range</u> of the sample means is therefore $8 - 4\frac{1}{4} = 3\frac{3}{4}$, as shown in Figure 11. The range, of course, is much easier to compute than s_4.

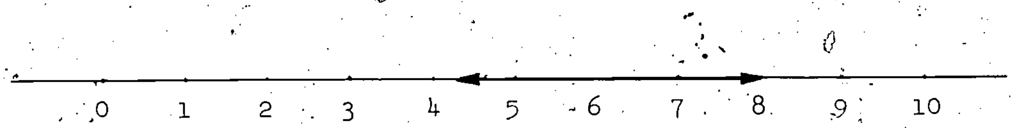


Figure 11

It can be proved that s_4 is less than one half of this range. Then $s_4 < 1\frac{7}{8}$.

For samples of size 5, the range of the sample means is only

$$\frac{36}{5} - \frac{26}{5} = \frac{10}{5} = 2$$

so that $s_5 < 1$.

These examples suggest two general conclusions about the mean values of all samples of a given size taken from a population:

(a) The average of all the sample means is equal to the population mean.

(b) The standard deviation of the set of all sample means is less than the standard deviation of the population. The greater the size of the sample, the smaller this standard deviation is.

Exercises 25-13 will bear out these conclusions. To <u>prove</u> that <u>these</u> two statements are always true is really a matter of careful bookkeeping.

This bookkeeping becomes rather complicated, particularly if the population is large so that there are many, many samples of a given size. Fortunately, statisticians have invented clever ways of doing the bookkeeping.

A very remarkable fact is that if the population (N) is large compared with the sample size (n), the value of s_n^2 for the sample means is almost exactly $\frac{1}{n}$ times the value of s^2 for the population.

(1) $$s_n^2 \approx \frac{1}{n} s^2.$$

To be exact

(2) $$s_n^2 = \frac{1}{n} \left(\frac{N - n}{N - 1}\right) s^2.$$

For example, if $n = 10$ and $N = 1000$,

$$s_{10}^2 = \frac{1}{10} \left(\frac{990}{999}\right) s^2 = \frac{11}{111} s^2$$

which is practically equal to $\frac{1}{10} s^2$.

In any case we see that $s_n^2 \le \frac{1}{n} s^2$.

<u>Exercises 25-13</u>

1. In the example given in the text, verify the values quoted for m, s^2, m_2, s_2^2, m_3, s_3^2.

2. Show that in the example given in the text the range of the sample means of size 5 is equal to 2, as stated.

3. Given the 15 numbers 0, 1, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 6, 6, 7.
   (a) Show these numbers on the number line.
   (b) Find the mean m and mark it on the number line.
   (c) What is the range?
   (d) What is the range of the means of all samples of size 2? of size 4? of size 8? of size 12?
   (e) Show the different range intervals on the number line.

4. In Exercise 3, replace the given list by the following:

0, 4, 5, 5, 6, 6, 6, 6, 6, 6, 6, 7, 7, 8, 12.

5. If we consider all samples of size 100 taken from a population of size 10,000, what does equation (2) give for s_{100}^2 ?

25-14. Estimating the Population Mean

In the previous section we began with a population and considered a number of samples chosen from it. Each of these samples contained the same number of items n.

We stated two facts:

(a) the average of the sample means is the same as the population mean m.

(b) the standard deviation (s_n) of the sample means of size n is smaller than the standard deviation (s) of the population.

Moreover

$$s_n^2 \approx \frac{s^2}{n} . \tag{1}$$

The result (1) is extremely important in the study of statistics. Note that $\frac{s^2}{n}$ does not depend upon the size of the population from which samples are chosen. It depends only on n, <u>the size of the sample</u>.

To see how we can use these results, let us suppose that we toss a coin 10,000 times and count the number of heads. In an actual experiment there were 54 heads in the first 100 tosses and 4979 heads in the 10,000 tosses. We show these results in a little table.

	Number of Tosses	Number of Heads
A	100	54
P	10,000	4,979

We may think of A as a sample of size 100 from a population P of 10,000. Let us assign the number 1 to each head and 0 to each tail in the population P. We now work out the population mean m and standard deviation s.

With 4979 heads (1's) and 5021 tails (0's) the mean is

$$m = \frac{5021 \times 0 + 4979 \times 1}{10,000} = .4979.$$

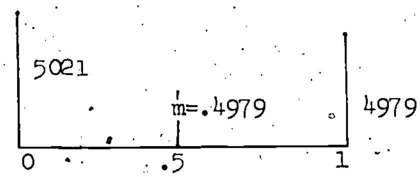


Figure 12

To compute s^2 we replace the mean .4979 by the approximation $.5 = \frac{1}{2}$. Then

$$s^2 = \overline{(x - \tfrac{1}{2})^2} - (.5 - .4979)^2.$$

But

$$\overline{(x - \tfrac{1}{2})^2} = \frac{(5021 \times \frac{1}{4}) + (4979 \times \frac{1}{4})}{10,000} = \frac{10,000 \times \frac{1}{4}}{10,000} = \frac{1}{4}.$$

Hence

$$s^2 = \frac{1}{4} - (.0021)^2.$$

Since $(.0021)^2 = .00000441$ is negligibly small, s^2 is almost exactly $\frac{1}{2}$.

Now let s_{100} be the standard deviation of all sample means of size 100 chosen from the population P.

According to (1),

$$s_{100}^2 \approx \frac{1}{100} s^2 \doteq \frac{1}{400}$$

so that

$$s_{100} \approx \frac{1}{20} = .05.$$

In Figure 13 we show the intervals $[m - s_{100}, m + s_{100}]$ and $[m - 2s_{100}, m + 2s_{100}]$.

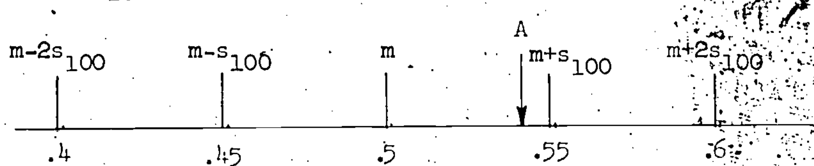


Figure 13

Where is the sample mean of A located in Figure 13? Since this mean is $\frac{46 \times 0 + 54 \times 1}{100} = .54$ we have the point shown by the arrow at A.

In the experiment reported, we now list the number of heads, not merely for the first set A of 100 tosses but also for the second, third, fourth and fifth sets of 100 tosses. The results are given in the table.

	Number of Heads	Sample Means
A	54	.54
B	46	.46
C	53	.53
D	55	.55
E	46	.46

In Figure 14, these five sample means are marked on the number line by the letters A, B, C, D, and E.

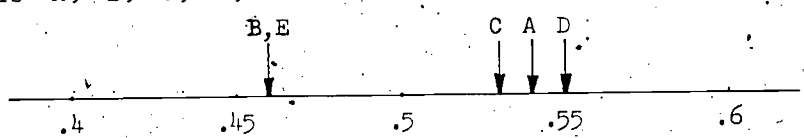


Figure 14

Exercises 25-14a

(Class Exercise)

A bucket contains 5,000 marbles, some black and some white. The number of black marbles is known to the teacher.

Each student shall draw a sample of size 20 and compute the sample mean. Black marbles are given the number 1 and white marbles the number 0. These results should be marked on the number line.

On the basis of the samples drawn, each student should guess the population mean. The teacher now gives the true value of the population mean. Each student now computes the standard deviation of the population and the standard deviation of the sample means and marks the intervals $[m - s_{20}, m + s_{20}]$ and $[m - 2s_{20}, m + 2s_{20}]$ on the number line.

The situation that we encounter in practice is very different from that of the previous section. There we knew the mean and the standard deviation of the population. In practice we have a sample chosen from the population. We wish to use this sample to obtain information about the entire population. Usually what we want is the population mean. Of course we have the sample mean. How close is this sample mean to the required population mean?

This question can be answered only with a certain probability. This is where Chebychev's Theorem comes to our aid. You will remember that this theorem says that at least $\frac{3}{4}$ of a set of numbers lies within 2 standard deviations of the mean. If we apply this result to the set of sample means of size n, we know that at least $\frac{3}{4}$ of the sample means are between $m - 2s_n$ and $m + 2s_n$. Therefore we are 75 percent certain that a given sample mean is within this interval.

More generally, Chebychev's Theorem says that not more than the fraction $\frac{1}{p^2}$ of the numbers lie more than p standard deviations away from the mean. Suppose that we guess that the population mean for a million coin tosses is $\frac{1}{2}$. Suppose, however, that we toss the coin 10,000 times and get 6000 heads. Then this sample mean is .60. What can we conclude? Is our guess about the population incorrect?

If the population mean is $\frac{1}{2}$ the standard deviation of all sample means for samples of 10,000 is

$$s_{10,000} = \frac{1}{\sqrt{10,000}} \times \frac{1}{2} = \frac{1}{100} \times \frac{1}{2} = \frac{1}{200} = .005 .$$

The interval $[m - s_{10,000}, m + s_{10,000}]$ is $[.495, .505]$. Now .60 is far outside this interval. In fact it is 20 standard deviations away from the mean $(\frac{1}{2})$. The chance that the mean of a sample of this size taken at random shall have a deviation as large as this cannot exceed $\frac{1}{20^2} = \frac{1}{400}$. With an honest coin and honest tosses there is only a $\frac{1}{4}$ of 1 percent chance that this would happen. We can almost certainly reject the idea that $m = \frac{1}{2}$.

Suppose that in a sample of 10,000 tosses we got 5200 heads. The sample mean .52 would be $4s_{10,000}$ away from the assumed mean $\frac{1}{2}$. The probability that this can occur if $m = \frac{1}{2}$ is only $\frac{1}{4^2} = \frac{1}{16}$. You would be in some doubt and might well decide to take a larger sample.

It should be said that much stronger statements can be made by using more powerful tools than Chebychev's Theorem. With these tools, the probability of a mean of .52 for a sample of 10,000, when the population mean is .50, can be shown to be very much less than $\frac{1}{16}$.

Exercises 25-14b

1. Draw a number line and mark the points corresponding to m, $m - s_{100}$, $m + s_{100}$, $m - 2s_{100}$, $m + 2s_{100}$ as in Figure 13. Locate the points that correspond to the means of the next five samples F, G, H, I, J of 100 in the coin tossing experiment given that the number of heads were as follows:

	Number of Heads
F	54
G	41
H	48
I	51
J	53

2. When taken together the samples A, B, C, D, E, F, G, H, I, J of 100 each make a sample of size 1000 from the population P.

   (a) What is the approximate value of s_{1000} according to the result (1)?

   (b) What is the more accurate value of s_{1000} given by

$$s_n^2 = \frac{1}{n}\left(\frac{N - n}{N - 1}\right) s^2 ?$$

3. Using the result of Exercise 2(b), draw a number line which shows the intervals

$$[m - s_{1000}, m + s_{1000}] \text{ and}$$

$$[m - 2s_{1000}, m + 2s_{1000}].$$

What is the mean of the sample of size 1000 which consists of A, B, C, D, E, F, G, H, I, and J? Locate this mean on the number line.

4. If the mean of a large population of coin tosses is actually .50, how large a sample is required to be 75 percent certain that a sample mean is between .49 and .51?

5. In a population of 1,000,000 letters, let us assume that 12 percent of the letters are E. Writing 1 for each E and 0 for any other letter, what is the mean m and standard deviation s of the population? In a sample of 900 letters it is found that only 10 percent are E's. Would this result cause you to reject the 12 percent assumption about the population?

25-15. Summary

We can represent a set of data (numbers) by a dot frequency diagram in which the frequency of occurrence of each number is shown by the number of dots above the corresponding point on the number line. The total number of points to the left of and including any marked position x on the number line is called the cumulative frequency at x. A diagram can be drawn to show the cumulative frequency at x as an ordinate.

If a line is drawn at a height which includes the lowest P percent of the total frequency, the largest number x that is included is called the Pth percentile. The lower quartile is the 25th percentile, the median is the 50th percentile and the upper quartile is the 75th percentile.

We can also describe a set of data briefly by two measures: (1) a measure of central tendency; (2) a measure of scatter.

Measuring Central Tendency

The most important measure of central tendency is the average or mean. The average of n numbers $x_1, x_2, \ldots, x_n$ is

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} .$$

The computation of the mean can often be simplified by using the following three properties:

(1) $\overline{x - c} = \overline{x} - c$;

(2) $\overline{ax} = a \cdot \overline{x}$;

(3) $\overline{x + y} = \overline{x} + \overline{y}$.

Two special cases are important:

(a) If the numbers are equally spaced when arranged in order, then $\overline{x} = \frac{f + \ell}{2}$, the average of the first and last numbers.

(b) $\overline{i^2} = \frac{1^2 + 2^2 + \ldots + n^2}{n} = \frac{(n + 1)(2n + 1)}{6}$.

Other measures of central tendency than the mean are the mode, the number that occurs most frequently (if there is one); and the median already mentioned.

Measuring Scatter

The most important measure of scatter is the standard deviation s. If m is the mean,

$$s^2 = \overline{(x - m)^2}.$$

The calculation of s^2 is often simplified by replacing m by an integer a, computing $\overline{(x - a)^2}$ and using the fact that

$$\overline{(x - m)^2} = \overline{(x - a)^2} - (m - a)^2.$$

It may be proved (Section 11) that at least $\frac{3}{4}$ of the numbers lie within the interval $[m - 2s, m + 2s]$ and more generally that at least $(1 - \frac{1}{p^2})\ n$ of the n numbers lie in the interval $[m - ps, m + ps]$. This is called Chebychev's Theorem.

Other less satisfactory measures of scatter are (1) the range, the difference between the largest and smallest numbers, and (2) the average distance, $d = \overline{|x - m|}$.

The Use of Statistical Measures

In using statistics, we often wish to use a sample chosen from a population to draw conclusions about the population itself. In particular, we can use the sample mean to test the accuracy of a guess about the population mean m. We use the fact that if s_n is the standard deviation of all samples of size n drawn from a much larger population with standard deviation s, then

$$s_n^2 \approx \frac{s^2}{n}.$$

We carry out this calculation for coin-tossing experiments and the occurrence of a given letter in English.

If we compare the difference between the sample mean and the assumed population mean (m) with s_n and if we apply Chebychev's Theorem, we can judge whether m is a reasonable value to assume for the population mean.

Chapter 26

SYSTEMS OF SENTENCES IN TWO VARIABLES

26-1. A Decision Problem

Suppose that you have just become president of a large corporation called "General Engines." Your division manufactures one make of car and one make of truck.

How many cars and trucks should be
scheduled for next year's production
to make as large a profit as possible?

Obviously, in order to make a decision, you need to have some detailed information. The following details might be available to a president of a corporation.

(1) The profit on each car sold is \$300, and the profit on each truck sold is \$400.

(2) It takes $1\frac{1}{2}$ tons of steel to build a car and 3 tons of steel to build a truck.

(3) The maximum number of cars and trucks that can be turned out in a year is 500,000.

(4) The amount of steel available to your division next year will be 960,000 tons. There will be 15,000 tons of steel left over from the current year, making 975,000 tons of steel on hand for next year.

Exercises 26-1

1. Suppose that you produced only cars. Keeping in mind the information given above:

(a) If you had unlimited plant capacity, how many cars could you build using the available steel?

(b) Can you actually produce as many cars as indicated in part (a)? Why or why not?

(c) How much profit would you make if you produced 500,000 cars?

(d) If you produce 500,000 cars, do you use all of the available steel? If not, how much steel is left over?

2. Suppose that you produce only trucks. Keeping in mind the information given above:

(a) If you had unlimited plant capacity, how many trucks could you make using the available steel?

(b) How much profit would you make if you produced the number of trucks indicated in part (a)?

(c) Do you use the full capacity of the plant when you produce the number of trucks found in part (a)? Do you use all of the available steel?

3. Suppose that we know that 90,000 trucks and 410,000 cars were produced and sold this year.

(a) How much profit was made on the cars? On the trucks?

(b) How much steel is needed for the above production plan?

(c) How many tons of steel will be surplus, at the end of next year, if you use the production plan indicated above?

4. Copy and complete the following table showing the information you have collected in previous exercises concerning possible production plans for next year.

	Number of cars produced	Number of trucks produced	Unused steel (in tons)	Unused plant capacity	Total Profit
(a)	500,000	0			$150,000,000
(b)		325,000			
(c)	410,000			0	

(d) What happens to the total profit when you decrease the number of cars produced from 500,000 to 410,000 and increase the number of trucks produced from 0 to 90,000?

(e) What happens to the amount of unused steel if the number of cars produced is decreased from 500,000 to 410,000 and if the number of trucks is increased from zero to 90,000?

(f) If you continue to decrease the number of cars produced from 410,000 to 0, and increase the number of trucks produced from 90,000 to 325,000, what happens to the total profit? What happens to the amount of unused steel?

5. Recommend a production plan so that the profit is maximum and the amount of steel left over is minimum. Support your plan with information about the expected profit, the amount of unused steel, and the plant use.

26-2. The Mathematical Model

One possible production plan would be a schedule calling for 400,000 cars and 100,000 trucks. A little arithmetic shows that for this plan

the number of tons of steel required is 900,000,

and the total profit is $160,000,000.

The increase in profit, over this year's profit, would be $1,000,000. The amount of steel left over would be only 75,000 tons. This appears to be a good plan, but is it the best possible plan? We will develop a mathematical model of this production problem to help us answer this question.

A great advantage of a mathematical model is that you can do "experiments" with it using just paper and pencil, or a computer. You can say, "What would happen if such and such were done?" Then you can carry out the mathematics and find out what the model predicts. You don't have to build something in a laboratory and test it, or wait until it happens in the real world. If no laboratory experiment is possible, then a mathematical model is often the only way one can get needed information. For example, no laboratory experiment is possible if you want to determine the route to be traveled to Mars by the first manned spaceship.

If our model is complete enough, then it will provide a good approximation to the real life situation; and we can rely on the answers it gives us. Of course the best test we have is to compare the predictions made using the model with the real situation and see how well they agree. Eventually this must always be done. If the agreement is poor we may have to add more features to the model. You can see that many different models can be made for the same real life situation, just as an artist can depict a scene in many different ways.

Now let us continue the analysis of our production problem and translate the statements into mathematical sentences.

Exercises 26-2a

(Class Discussion)

1. If x represents a number of cars produced, interpret the output of the function

$$f : x \to 300x.$$

2. If y represents a number of trucks produced, interpret the output of the function

$$g : y \to 400y.$$

3. Interpret the output of the function $h : (x, y) \to 300x + 400y$.

4. If P represents the total profit, then represent the output of the function h by an equation. (We will call this equation the <u>profit equation</u>.)

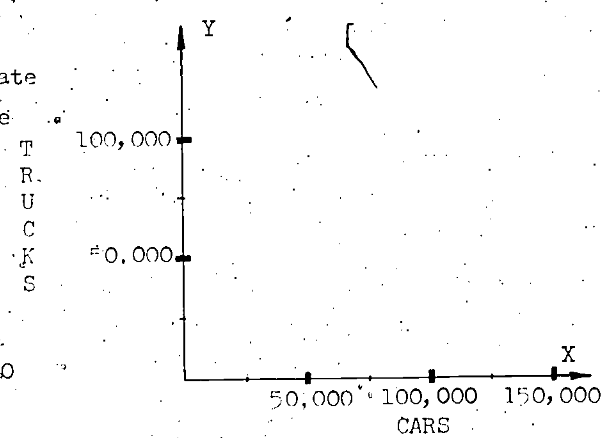
5. On a single pair of coordinate axes, draw the graphs of the profit equation if

   (a) $P = 150,000,000$.

   (b) $P = 130,000,000$.

   (c) $P = 180,000,000$.

   (<u>Note</u>: Use units of 25,000 on your graph.)



6. The graphs of the lines in Exercise 5 are ___?___ since the <u>slopes</u> are the same.

7. As P ___?___ (increases, decreases) the line moves away from the origin, and as P ___?___ (increases, decreases) the line moves in toward the origin.

8. From your graph estimate two different pairs of values of x and y which will yield the same profit of $130,000,000.

9. In this production problem do the values of the variables have any meaning when $x < 0$ or $y < 0$?

____

Now we need to incorporate the fact that the amount of steel used <u>cannot exceed 975,000 tons</u>. If x represents a number of cars produced, then the output of the function

$$s : x \to 1.5x$$

can be described as the number of tons of steel needed to produce x cars.

If y represents a number of trucks produced, then the output of the function

$$p : y \to 3y$$

can be described as the number of tons of steel needed to produce y trucks.

The output of the function $t : (x, y) \to 1.5x + 3y$ can be interpreted as the total number of tons of steel needed to produce x cars and y trucks. We can now write the <u>steel usage inequality</u>:

$$1.5x + 3y \leq 975,000.$$

Since we're interested only in non-negative values of x and y, we include the restrictions $x \geq 0$, $y \geq 0$. Why?

The plant capacity is also restricted. Therefore we need to write a sentence describing this restriction.

$$x + y \leq 500,000$$

is called the <u>plant capacity</u> inequality.

Now a mathematical statement of the problem is as follows:

We want the number pair (x, y) satisfying the conditions

$$x \geq 0, \quad y \geq 0,$$

and $1.5x + 3y \leq 975,000$; (Steel usage inequality)

and $x + y \leq 500,000$; (Plant capacity inequality)

for which the profit P,

$$P = 300x + 400y,$$

is a maximum.

Exercises 26-2b

1. Draw a graph of the region represented by

$$1.5x + 3y \leq 975,000 \text{ and } x \geq 0 \text{ and } y \geq 0.$$

2. On the same coordinate axes draw the graph of the region represented by

$$x + y \leq 500{,}000 \quad \text{and} \quad x \geq 0 \quad \text{and} \quad y \geq 0.$$

3. The points with integral coordinates in the intersection of the solution sets of $1.5x + 3y \leq 975{,}000$; $x \geq 0$; $y \geq 0$ and $x + y \leq 500{,}000$; $x \geq 0$; $y \geq 0$ represents the set of all permissible production plans.

   (a) Is there a production plan, represented by a point in this region, which will yield a profit of $P = 60{,}000{,}000$? Draw a graph to support your answer. (Hint: $(300)(?) = 60{,}000{,}000$, and $(400)(?) = 60{,}000{,}000$.)

   (b) Is there a production plan which will yield a profit of \$160,000,000? Draw a graph to support your answer. Is there more than one plan which will yield a profit of \$160,000,000?

   (c) Is there a production plan which will yield a profit of \$180,000,000? Draw a graph to support your answer.

4. The graph of the mathematical model suggests that the "profit line" should be moved "as far out as possible from the origin," since moving it out corresponds to increasing profit. However, the "profit line" must still intersect the solution set of the other given conditions in at least one point.

   (a) Find the coordinates of the point where the "profit line" intersects the solution set and where the profit line is as "far out" as possible from the origin.

   (b) Using the coordinates found in (a), calculate the profit for this production plan. Is this the maximum profit for the given conditions?

The method of solution corresponds to moving the profit line as "far out" as possible until it intersects only the boundary of the region that is described by the conditions of the problem and does not contain any interior points of the region. The coordinates of a point of intersection give the solution -- that is, the values of x and y. It can happen in a problem of this kind that instead of a single point the profit line will coincide with a segment of the boundary of the solution set. In this case, more than one solution to the problem exists. However, if the solution set is a <u>convex</u> region bounded by straight edges, a maximum can always be found at one of the vertices.

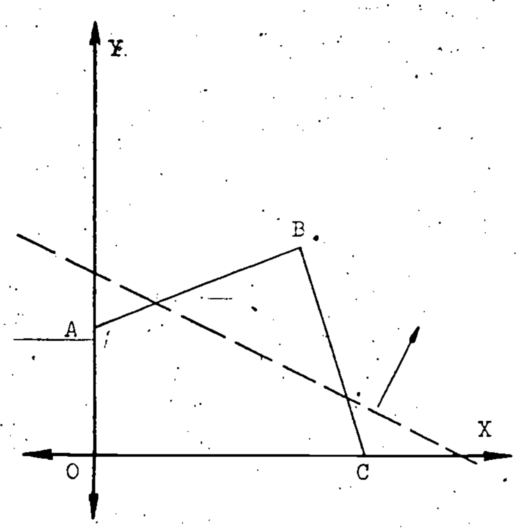
This means that in seeking a production plan we need only to check the profit at each vertex of the graph of the solution set. Since the number of vertices is finite, this is a rapid method of finding the best plan. In the study of these problems (called linear programming problems by mathematicians), there is a method of solution called the "simplex method" in which only the vertices are used.

Exercises 26-2c

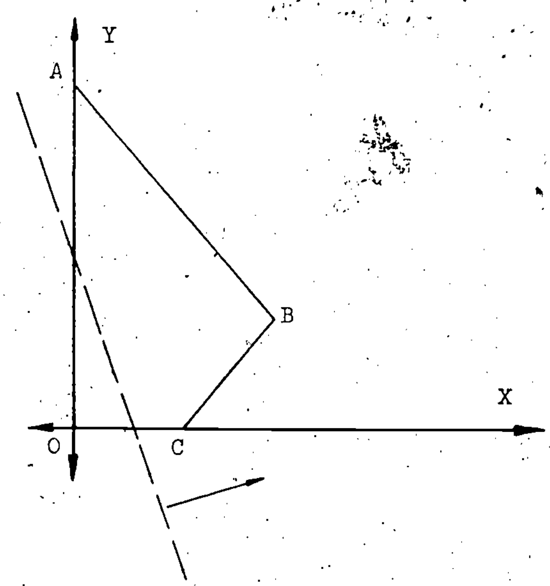
(Class Discussion)

The following drawings illustrate the mathematical model, in geometrical terms, corresponding to several different kinds of decision problems. The closed figure represents the set of possible solutions. The dashed line corresponds to one position of the profit line (or whatever quantity we wish to maximize). Move the profit line in the direction shown by the arrow. Tell in each case what intersection of the solution set with the profit line maximizes the profit.

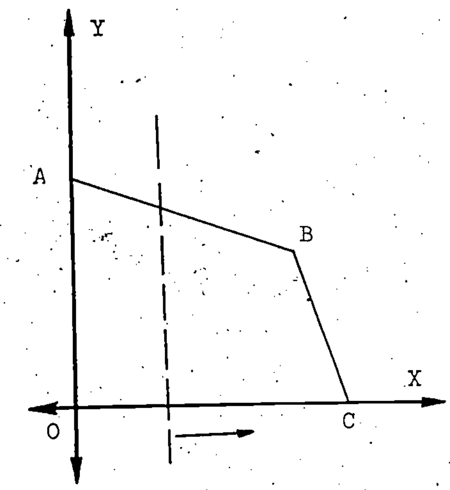
1.



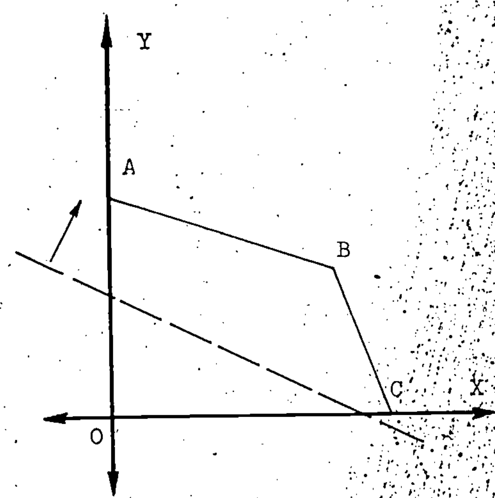
2.



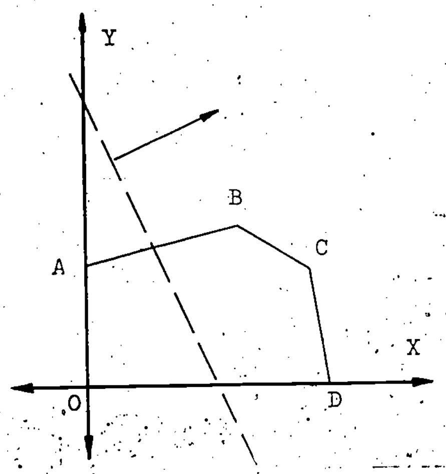
3.



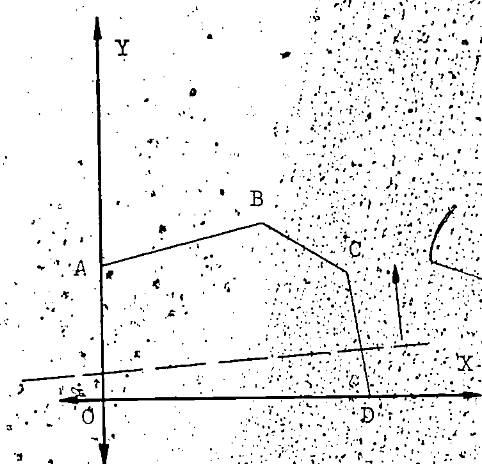
4.



5.



6.



26-3. Some Related Problems

You have just studied a particular linear programming problem concerning a decision about producing cars and trucks. People in government, business, and industry are often asked to make similar decisions. Usually these decisions have to do with money; a minimum cost or a maximum profit is desired. In this section we would like you to read about some different situations whose mathematical models are very similar to the one derived in the first two sections of this chapter. There are no solutions required in this section, but you could solve simple versions of most of these problems using the techniques illustrated in sections 26-1 and 26-2. We will develop some more techniques for solving these problems in later sections of the chapter.

a. The Diet Problem

A diet for losing or controlling weight is to be planned using a number of different foods. Each food contains certain amounts of different nutrients per ounce (such as protein, vitamin C, calcium, iron, etc.). Each food contains a certain number of calories per ounce. The diet requires certain minimum amounts of each nutrient per day. The problem is to find the amount of each food included in the diet which will give at least the minimum amount of nutrition, and will make the total number of calories as small as possible.

b. Transportation Problems

A company maintains a warehouse in each of a number of cities. Each warehouse holds a certain number of units of a given commodity (such as refrigerators). Orders come in from dealers in surrounding places. We are given the number of units required by each dealer, the distances from the dealers to the warehouses and the cost of shipping from each warehouse to each dealer. The problem is to decide how to fill the orders: how many units to ship to each dealer from each warehouse in order that the cost of shipping to meet all the orders be a minimum (as small as possible).

A variation of this problem (in wording alone) occurs in military operations planning. A nation maintains a number of naval bases. Each base is the home of a certain number of aircraft carriers, destroyers, etc. At a certain time it is necessary to assemble a task force at each of several locations for maneuvers. The number of ships of each type to rendezvous at these spots is assigned. The distance from each base to each rendezvous location is given. The problem is to decide on the orders for the ships: what destination should be assigned to the ships from each base in order to have the total travel time for all the ships a minimum. Notice that in this example and the warehouse problem, minimizing time is the same as minimizing distance. Since fuel consumption (for trucks or ships) is proportional to distance traveled, this is also the same as minimizing cost.

An interesting variation is the following: if the task forces must be assembled as fast as possible (say if there is an international emergency), then the problem is no longer the same. We would not be interested in adding the travel times of all the ships, but in the longest time taken by any ship to reach its destination. This leads to a different type of mathematical problem.

c. Blending Problems

An oil company refines crude oil to produce a certain number of barrels daily of several different chemical parts of the oil. These parts can be blended to make different marketable products such as grades of automotive gasolines and aviation gasolines. These products sell for different prices. We are given the number of barrels of each chemical part produced daily, the blending rules and the sale prices of the final products. The problem is how much of each product to produce daily to yield the maximum income.

A variation of this problem involves mixtures. As an example, suppose that two mixtures of nuts are to be offered for sale: a regular mix and a party mix. The proportions of the different kinds of nuts used for each mix are prescribed. Also given are the costs per pound of each kind of nut, the total supply of each kind of nut available and the selling price per pound of the two mixes. The problem is to decide how many pounds of each mix to produce out of the given supply so as to maximize the profit.

d. Network Problems

These are problems involving a network of telephone lines interconnecting cities, of roads and highway systems, of connections in an electronic circuit and so forth. As an example, suppose that special communications cables (say for TV) have to be laid to join a number of distant cities. It is not necessary to lay a cable directly between each pair of cities so long as some route can be found between them. For example, a cable need not join Chicago and Los Angeles directly if there is one from Chicago to San Francisco and one from San Francisco to Los Angeles since these can be joined at a switching station in San Francisco. Given the distance between each pair of cities, the problem is to determine which cities to join by cable in order that any city in the network can communicate with any other city and so that the total amount of cable to be laid is a minimum. This is called the shortest connecting network. The same problem frequently arises in the telephone business. If there are n points in the network to be connected, there are n^{n-2} possible networks. This number increases very rapidly as n increases, and it becomes impossible simply to measure all possible networks. As a linear programming problem the solution is easily found.

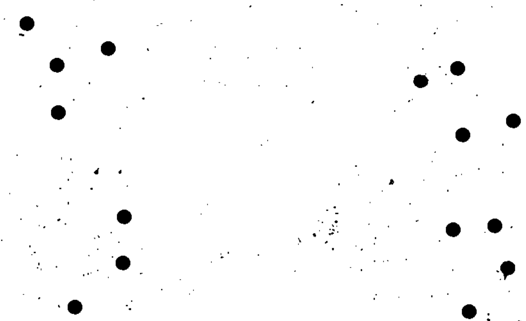


Figure 1

15 Points to be Joined by a Network

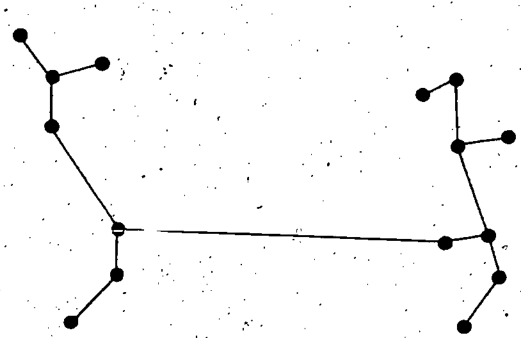


Figure 2

Shortest Connections Network

Figure 1 shows the location of a number of points to be joined by the shortest connecting network, and Figure 2 shows the solution.

A related problem concerns the maximum flow in a network. Suppose that the various cities in a network are joined by telephone "trunk lines," and that each trunk line can handle a certain number of calls. If a trunk is fully used, alternate routes can be found to place a call, using trunks to other intermediate cities. Given the location of all of the trunks and the maximum number of calls that each can handle, the problem is to find the maximum number of calls which can be made at one time from one particular city to another, say from New York to Los Angeles.

e. The Assignment Problem

Suppose that there are a number of jobs to be filled and a certain number of people available to carry out these assignments. Each person could be assigned to any one of the jobs, but he is better at some jobs than at others. Suppose that we are given a rating for each person for each of the jobs (say 10 if he is very good at it, down to 1 if he is very poor). The problem is to assign the people to the jobs so as to maximize the sum of all the ratings of the people in the jobs to which they are assigned.

f. The Trim Problem

Paper mills produce paper in large rolls in certain standard widths only. Customer orders are received specifying intermediate widths desired and the number of rolls of each width. These widths are obtained by cutting or trimming the existing rolls. Figure 3 shows the location of cuts which might occur along a roll.

A single roll might be cut in many ways to fill different orders. The problem is, given the widths of the standard rolls and given the customers' orders, how should the rolls be cut so that the amount of paper wasted (unused ends of rolls) is a minimum?

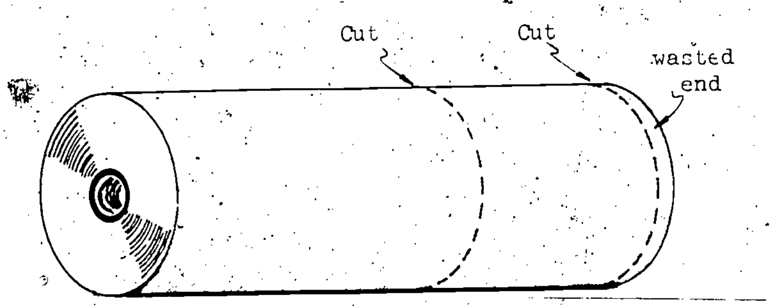


Figure 3

26-4. Solution Sets of Systems of Equations

As you recall, the solution set of a mathematical sentence is the set of real numbers which will make the sentence a true statement.

The production problem discussed in the first three sections had several mathematical sentences which represented restrictive conditions in the given situation. We now wish to consider a simpler system of linear equations like

$$2x + 3y - 9 = 0 \text{ and } 5x - 2y + 25 = 0,$$

and develop some methods for finding the solution set of this system. Such a system of equations can serve as a mathematical model for many different situations where the variables have two different conditions placed on them simultaneously. Our immediate problem is to find an ordered pair of numbers, (x, y), that will satisfy both clauses of the compound sentence, if such an ordered pair exists. In set-builder notation we are trying to find

$$\{(x, y) : 2x + 3y - 9 = 0\} \cap \{(x, y) : 5x - 2y + 25 = 0\}.$$

We will use the symbolism

$$\begin{cases} 2x + 3y - 9 = 0 \\ 5x - 2y + 25 = 0 \end{cases}$$

as another way to write the compound sentence

$$2x + 3y - 9 = 0 \text{ and } 5x - 2y + 25 = 0.$$

We will write only the equation portion of the set-builder statements, and carry the set implications mentally.

The solution set of each of the individual clauses of this compound sentence can be represented graphically as the set of all points on a line. The point of intersection of the two lines, if they do intersect, represents the solution set of the system of equations. The coordinates of the point of intersection are the values of x and y which satisfy both equations.

As you remember, it is convenient to write the equation of a line in y-form when graphing. Our system, then, is equivalent to

$$\begin{cases} y = -\frac{2}{3} x + 3 \\ y = \frac{5}{2} x + \frac{25}{2} . \end{cases}$$

The graph is as follows:

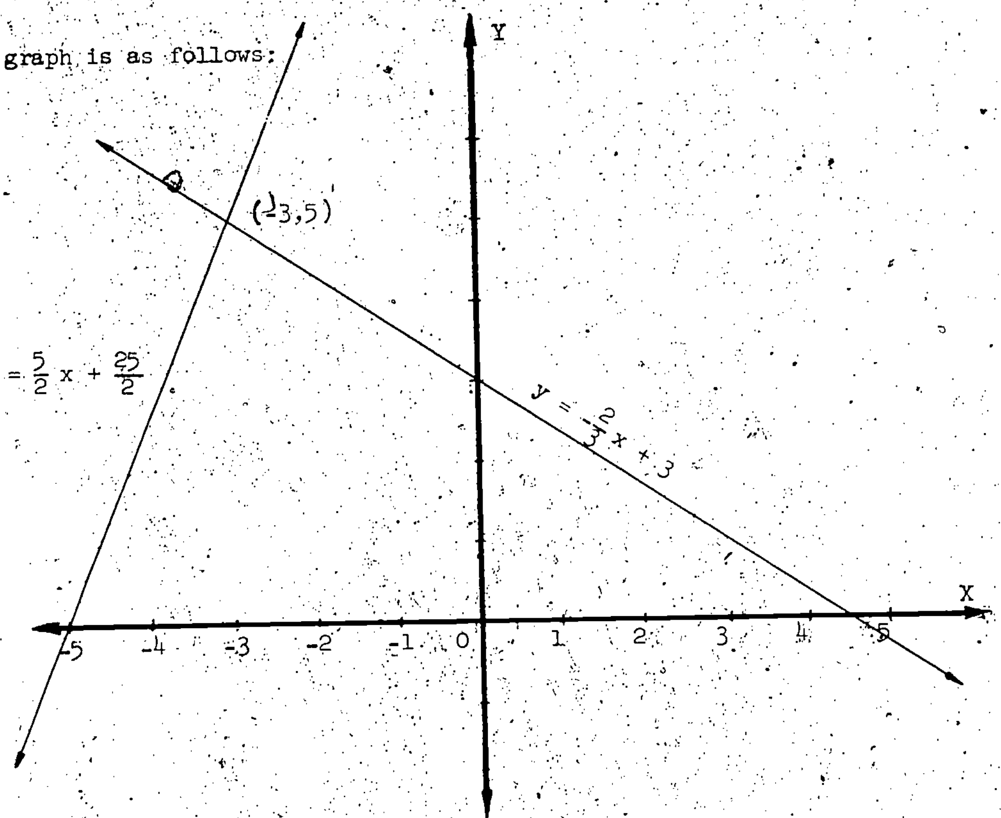


Figure 4

It is apparent that the solution set of this system is $\{(-3, 5)\}$.

Exercises 26-4a

(Class Discussion)

Simplify each of the sentences in the systems in Exercise 1. You will be given the coordinates of one point on each line. Put each equation in slope-intercept form and then draw the graph of each sentence on a single set of coordinate axes. The answers to Exercises 2 through 5 will be based on the eight lines drawn in Exercise 1.

Example:

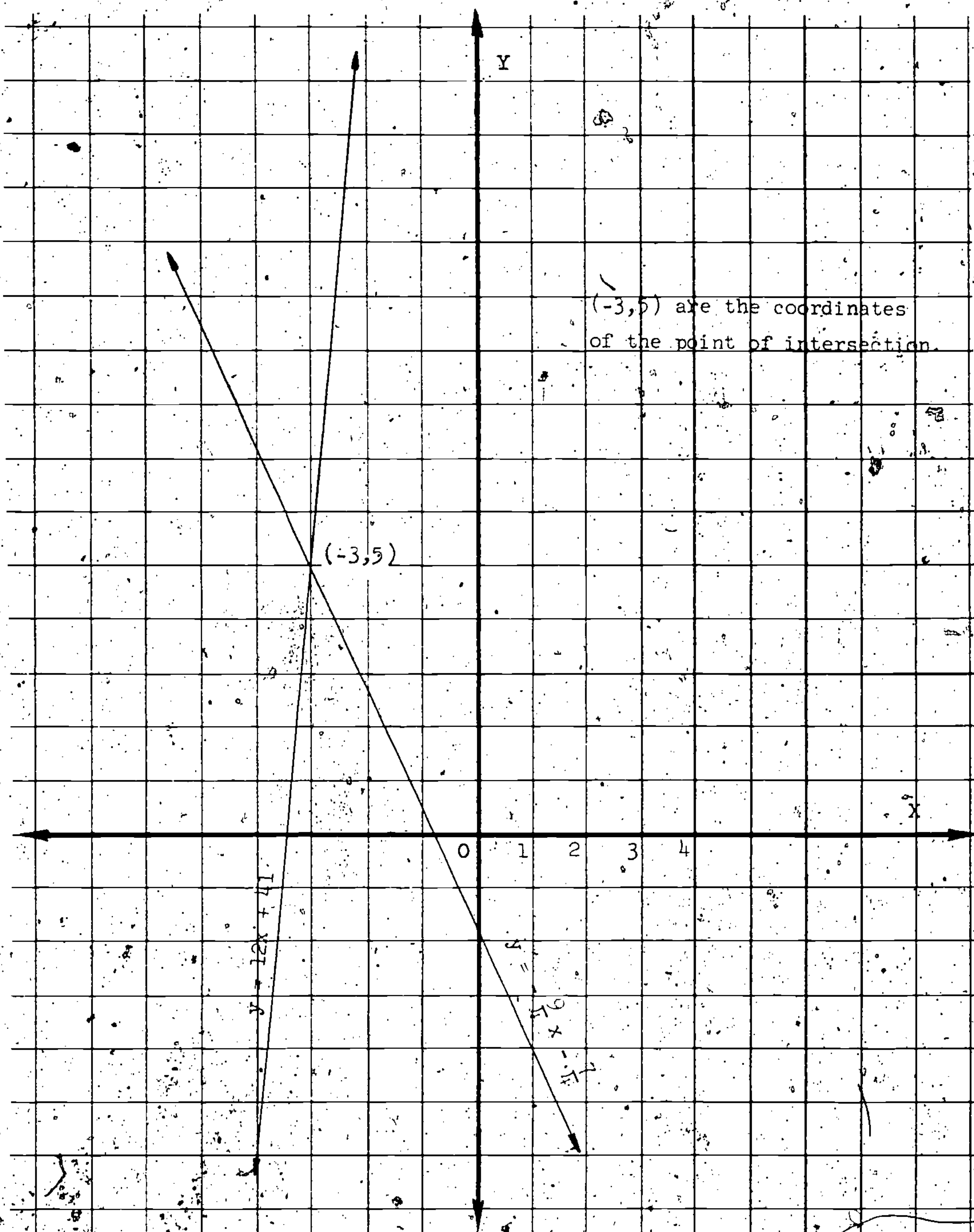
$$\begin{cases} (1)(2x + 3y - 9) + (2)(5x - 2y + 25) = 0 \\ (2)(2x + 3y - 9) + (1)(5x - 2y + 25) = 0 \end{cases}$$

is equivalent to

$$\begin{cases} 12x - y + 41 = 0 \\ 9x + 4y + 7 = 0, \end{cases}$$

which is equivalent to

$$\begin{cases} y = 12x + 41 \\ y = -\frac{9}{4}x - \frac{7}{4} \end{cases}$$



		Line Number	Coordinate of one point on line
1. (a)	$(1)(2x + 3y - 9) + (-1)(5x - 2y + 25) = 0$	#1	(-8, 2)
	$(1)(2x + 3y - 9) + (-2)(5x - 2y + 25) = 0$	#2	(-10, -3)
(b)	$(2)(2x + 3y - 9) + (1)(5x - 2y + 25) = 0$	#3	(1, -4)
	$(3)(2x + 3y - 9) + (1)(5x - 2y + 25) = 0$	#4	(4, -6)
(c)	$(-2)(2x + 3y - 9) + (1)(5x - 2y + 25) = 0$	#5	(5, 4)
	$(-3)(2x + 3y - 9) + (2)(5x - 2y + 25) = 0$	#6	(1, 10)
(d)	$(2)(2x + 3y - 9) + (3)(5x - 2y + 25) = 0$	#7	(-3, 0)
	$(-5)(2x + 3y - 9) + (2)(5x - 2y + 25) = 0$	#8	(0, 5)

2. What observation can you make about the coordinates of the points of intersection of the lines in Exercises 1(a) through 1(d)?

3. How would you describe the lines in Exercise 1(d)?

4. Do you think that it is possible to choose factors so that the lines are like the ones in Exercise 1(d)? Why or why not?

5. Did all of the systems of equations have the same solution set as the system $\begin{cases} 2x + 3y - 9 = 0 \\ 5x - 2y + 25 = 0 \end{cases}$? Why or why not?

It should be clear to you, at least geometrically, that no matter what factors, K and M, we use, if $K \cdot M \neq 0$, then the line

$$(K)(2x + 3y - 9) + (M)(5x - 2y + 25) = 0$$

will contain the point of intersection of the two lines in the original system $\begin{cases} 2x + 3y - 9 = 0 \\ 5x - 2y + 25 = 0 \end{cases}$. (This, of course, assumes that a point of intersection exists.) The expression $K(2x + 3y - 9) + M(5x - 2y + 25)$ is called a <u>linear combination</u> of the two polynomials $2x + 3y - 9$ and $5x - 2y + 25$.

Furthermore, if we choose K and M carefully, we can form two sentences that represent either a vertical or a horizontal line through the point of intersection of the original system of equations. That is, if the lines intersect we can find a system like

$$\begin{cases} x = a \\ y = b \end{cases}$$

which is equivalent to the original system of equations. From such a system it is possible to "read out" the solution of the original equivalent system. Hence, we need select only values of K and M which will result in a zero coefficient for y, and choose values of K and M which will result in a zero coefficient for x.

Exercises 26-4b

(Oral)

State one pair of values for K and M which will eliminate the term containing y and one pair of values which will eliminate the term containing x in the following sentences.

1. $(K)(x - 3y + 4) + (M)(x + 7y - 11) = 0$
2. $(K)(5x + 2y - 5) + (M)(x - 3y - 18) = 0$
3. $(K)(3x + y + 18) + (M)(2x - 7y - 34) = 0$
4. $(K)(x + 3y - 5) + (M)(2x - 3y - 2) = 0$
5. $(K)(x - 2y + \frac{1}{6}) + (M)(2x + y - \frac{4}{3}) = 0$
6. $(K)(3x + 2y - 1) + (M)(2x - \text{[illegible]}y - 18) = 0$

We are now in a position to solve our original system algebraically.

$$\begin{cases} 2x + 3y - 9 = 0 \\ 5x - 2y + 25 = 0 \end{cases}$$

$$\Longleftrightarrow \begin{cases} (2)(2x + 3y - 9) + (3)(5x - 2y + 25) = 0 \\ (5)(2x + 3y - 9) + (-2)(5x - 2y + 25) = 0 \end{cases}$$

$$\Longleftrightarrow \begin{cases} 19x + 0 + 57 = 0 \\ 0 + 19y - 95 = 0 \end{cases}$$

$$\Longleftrightarrow \begin{cases} x = -3 \\ y = 5 \end{cases}$$

The solution set is

$$\{(x, y) : x = -3\} \cap \{(x, y) : y = 5\}$$
$$= \{(-3, a)\} \cap \{(b, 5)\}$$
$$= \{(-3, 5)\}.$$

(Remember that the symbol "$\Longleftrightarrow$" means "is equivalent to".)

Exercises 26-4c

Find the solution sets of the following systems of equations by the method just developed. In Exercises 1 and 2 draw the graphs of each pair of equations.

1. $\begin{cases} 3x - 2y - 14 = 0 \\ 2x + 3y + 8 = 0 \end{cases}$

2. $\begin{cases} 5x + 2y = 4 \\ 3x - 2y = 12 \end{cases}$

3. $\begin{cases} 3 - 5x = 0 \\ 3y = x - 6 \end{cases}$

4. $\begin{cases} 5x - y = 32 \\ x - 2y - 19 = 0 \end{cases}$

5. $\begin{cases} 3x - 2y = 27 \\ 2x - 7y = -50 \end{cases}$

6. $\begin{cases} \frac{1}{2}x + y = 2 \\ y - \frac{1}{3}x = 1 \end{cases}$

7. $\begin{cases} x + y - 30 = 0 \\ x - y + 7 = 0 \end{cases}$

8. $\begin{cases} y = 7x + 5 \\ 4x = y - 3 \end{cases}$

9. $\begin{cases} 7x - 6y = 9 \\ 9x - 8y = 7 \end{cases}$

26-5. Parallel and Coincident Lines; Solution by Substitution

In the previous sections we considered only situations in which a system of equations had a solution set consisting of a single number pair. We interpret this situation by saying that the two lines that are graphs of the separate equations of each system have intersected in a single point.

Exercises 26-5a

(Class Discussion)

1. Consider the system $\begin{cases} 2x + y - 5 = 0 \\ 2x + y + 2 = 0 \end{cases}$

   (a) Write an equivalent system of equations with each equation in y-form, that is, in the form $y = mx + b$.

   (b) Draw the graph of each of the equations in (a) on the same coordinate axes.

   (c) Do the two lines intersect? Why or why not?

   (d) Forming an equivalent system of equations using the linear combination method and simplifying we get:

   $$\begin{cases} (1)(2x + y - 5) + (-1)(2x + y + 2) = 0 \\ (1)(2x + y - 5) + (-1)(2x + y + 2) = 0 \end{cases} \Leftrightarrow \begin{cases} 0 \cdot x + 0 \cdot y - 7 = 0 \\ 0 \cdot x + 0 \cdot y - 7 = 0 \end{cases}$$

   Are there any ordered pairs of real numbers (x, y) that will make the equations in this last system true statements?

   (e) The solution set of this system is the ________ set. Does this correspond to your graphical analysis in part (b)?

2. Consider the system $\begin{cases} 3x + y - 2 = 0 \\ 9x + 3y - 6 = 0 \end{cases}$

   (a) Write an equivalent system of equations with each equation in the form $y = mx + b$.

   (b) Draw the graphs of the two equations in part (a) on the same coordinate axes.

   (c) Do the two lines intersect? In how many points do the lines intersect?

   (d) Can you find K so that $(K)(3x + y - 2)$ is the same as $9x + 3y - 6$?

   (e) If you can find K in part (d), this means that every solution, (a, b), of $3x + y - 2 = 0$ is also a solution of $9x + 3y - 6 = 0$. Does this result correspond with your graphical interpretation of the situation in part (c)?

We have now seen that for the solution set of a system of equations

$$\begin{cases} ax + by - c = 0 \\ dx + ey - f = 0 \end{cases}$$

there are three possibilities:

(1) The solution set contains exactly one ordered pair of real numbers;

(2) The solution set is $\emptyset$;

(3) The solution set is an infinite set of ordered pairs of real numbers.

Graphically these possibilities can be interpreted as:

(1) Two lines that intersect in exactly one point;

(2) Two lines that are parallel; they do not intersect;

(3) A single line which is the graph of both equations.

The y-form of the linear equation, $y = mx + b$, is very helpful to us in recognizing which of the three possibilities exist.

(1) If the slopes are <u>not</u> equal, then the two lines intersect in exactly one point.

(2) If the slopes are equal, and the y-intercepts are different, then the lines are parallel; and the solution set is $\emptyset$.

(3) If the slopes are equal, and the y-intercepts are equal, then the two equations represent the same straight line; and the solution set is an infinite set of ordered pairs of real numbers.

If the two lines intersect in exactly one point, we notice in particular that at this point the value of y is the same for both sentences. This gives us another "method" for solving a system of linear equations.

<u>Example 1</u>: Solve $\begin{cases} 2x + y - 7 = 0 \\ x - y + 2 = 0 \end{cases}$

Put each equation in y-form.

$$\begin{cases} y = -2x + 7 \\ y = x + 2 \end{cases}$$

Since the slopes are different, the lines intersect in exactly one point.

Hence, for the point of intersection:

$$-2x + 7 = x + 2$$

$$\iff 5 = 3x$$

$$\iff \frac{5}{3} = x$$

The solution set of this equation is $\{\frac{5}{3}\}$.

If x is $\frac{5}{3}$ at the point of intersection, then using the second equation in y-form we find

$$y = \frac{5}{3} + 2. \quad \text{(We could also use the first equation.)}$$

Hence the solution set of the system is $\{(\frac{5}{3}, \frac{11}{3})\}$.

<u>Example 2</u>: Solve the system $\begin{cases} x + y = 7 \\ 2x - 3y = 4 \end{cases}$

It's easier to put the first equation in y-form than the second equation. That is,

$$y = -x + 7.$$

For "y" in the second equation we <u>substitute</u> "$-x + 7$".

$$\begin{aligned} & 2x - (3)(-x + 7) = 4 \\ \Longleftrightarrow\ & 2x + 3x - 21 = 4 \\ \Longleftrightarrow\ & 5x - 21 = 4 \\ \Longleftrightarrow\ & 5x = 25 \\ \Longleftrightarrow\ & x = 5 \end{aligned}$$

The solution set of this equation is $\{5\}$.

Since $y = -x + 7$ at the point of intersection, and we know that x is 5, we can write:

$$y = -5 + 7$$
$$y = 2$$

The solution set of the system is $\{(5, 2)\}$.

The method just described in the above examples is called the <u>substitution</u> method.

<u>Exercises 26-5b</u>

For each of the following systems of equations determine the nature of the solution set by putting the equations in y-form. If the solution set consists of exactly one ordered pair of real numbers, use the substitution method to solve the system.

1. $\begin{cases} 2x - y - 7 = 0 \\ 5x + 2y - 4 = 0 \end{cases}$

2. $\begin{cases} x - 2y - 5 = 0 \\ 3x - 6y - 12 = 0 \end{cases}$

3. $\begin{cases} 5x - 4y + 2 = 0 \\ 10x - 8y + 4 = 0 \end{cases}$

4. $\begin{cases} 2x - y + 13 = 0 \\ x + 4 = 0 \end{cases}$

5. $\begin{cases} x - 3y + 4 = 0 \\ x + 7y - 11 = 0 \end{cases}$

6. $\begin{cases} x + y = 56 \\ x - y = 18 \end{cases}$

7. $\begin{cases} y = 2x - 4 \\ x - \frac{1}{2} y - 2 = 0 \end{cases}$

8. $\begin{cases} x = 9y \\ \frac{1}{3} x = 3y + 2 \end{cases}$

9. $\begin{cases} \frac{1}{8} x - \frac{1}{3} y = 0 \\ \frac{1}{6} x - \frac{1}{9} y = 3 \end{cases}$

In the following exercises fill in the missing coefficients so that the graphs of the two equations are the same line.

10. $\begin{cases} 3x + (?)y + 6 = 0 \\ (?)x + 2y + 12 = 0 \end{cases}$

11. $\begin{cases} (?)x + \frac{1}{2} y + (?) = 0 \\ 2x + \frac{1}{3} y + 4 = 0 \end{cases}$

12. $\begin{cases} 5x + (?)y + (?) = 0 \\ 15x + (?)y + 3 = 0 \end{cases}$

13. What relation or relations must exist among the ratios $\frac{a}{d}$, $\frac{b}{e}$, and $\frac{c}{f}$ so that the graphs of the equations in the following system are the same line?

$$\begin{cases} ax + by + c = 0 \\ dx + ey + f = 0 \end{cases}$$

In the following exercises fill in the missing coefficients so that the graphs of the two equations are parallel lines.

14. $\begin{cases} x + 2y + 5 = 0 \\ 2x + (\ )y + (\ ) = 0 \end{cases}$

15. $\begin{cases} (\ )x + \frac{2}{3} y + (\ ) = 0 \\ (\ )x + \frac{1}{3} y + 6 = 0 \end{cases}$

16. $\begin{cases} 2x - 5y + 7 = 0 \\ 7x - (\ )y + (\ ) = 0 \end{cases}$

17. What relation or relations must exist among the ratios $\frac{a}{d}$, $\frac{b}{e}$, and $\frac{c}{f}$ so that the graphs of the equations in the following system are parallel lines?

$$\begin{cases} ax + by + c = 0 \\ dx + ey + f = 0 \end{cases}$$

26-6. Graphical Solution of Systems of Inequalities

As you remember, a line separates a plane into three non-intersecting sets of points; (1) the set of points on the line, (2) the set of points in the half plane on one side of the line, and (3) the set of points in the half plane on the opposite side of the line.

The simplest way to find the graph of a sentence like

$$3x - 2y + 6 < 0$$

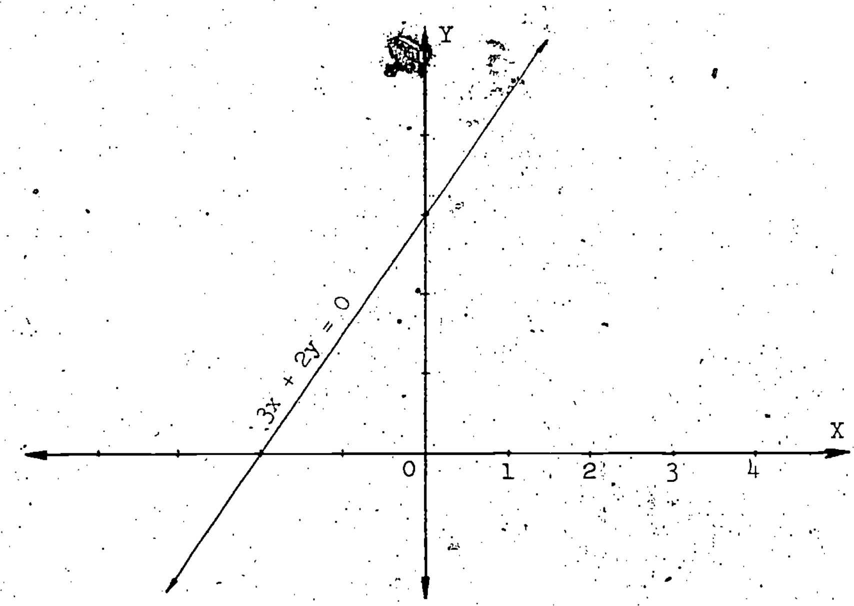
is to graph the equation

$$3x - 2y + 6 = 0.$$

In y-form an equivalent equation is

$$y = \frac{3}{2}x + 3.$$

The graph of this equation is the line with slope $\frac{3}{2}$ and y-intercept 3.



To find the half plane which is the solution set of $3x - 2y + 6 < 0$ we can use the equivalent sentence $y > \frac{3}{2}x + 3$. If $x = 0$, we see that $y > \frac{3}{2}(0) + 3$, or $y > 3$. These are the points on the part of the y-axis above the line $3x + 2y + 6 = 0$. Therefore the points in the half plane above the line, $3x + 2y + 6 = 0$, are in the solution set of the original sentence.

We can also determine which half plane is the graph of the inequality

$$3x - 2y : 6 < 0$$

by checking the coordinates of a single point in one of the half planes. Usually the easiest one to use is the origin, with coordinates $(0, 0)$. If the sentence, $3x + 2y + 6 < 0$, becomes a true statement when x and y are both replaced by 0, then the origin is a point in the graph of the solution set of the sentence. Furthermore, every point in the same half plane as the origin is on the graph of the solution set of the sentence. If the resulting statement is false, then the half plane on the opposite side of the line is the graph of the solution set of the sentence. If the line passes through the origin, it is helpful to choose a point on one of the coordinate axes so that one of the coordinates is zero.

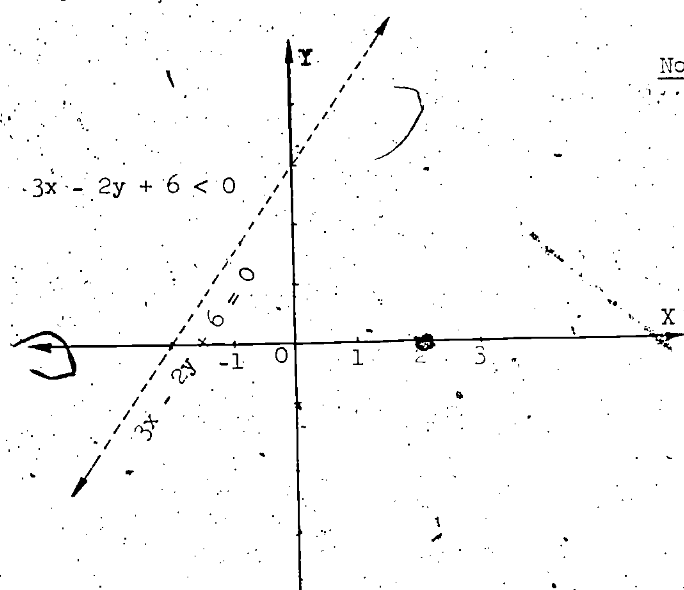


Figure 5

<u>Note</u>: The line is not solid because the points on the line are <u>not</u> included in the graph of the solution set. If the sentence had been

$$3x - 2y + 6 \leq 0,$$

then we would have included the points on the line.

The shaded region represents the solution set, namely

$$\{(x, y) : 3x - 2y + 6 < 0\}.$$

To find the solution set of a system of inequalities like

$$\begin{cases} x - 2y - 3 \geq 0 \\ 2x + y + 4 > 0 \end{cases}$$

we find the graphs of the functions $f : (x, y) \to x - 2y - 3$ and $g : (x, y) \to 2x + y + 4$, graph the solution set of each of the inequalities in the system, and find their intersection; that is

$$\{(x, y) : x - 2y - 3 \geq 0\} \cap \{(x, y) : 2x + y + 4 > 0\}.$$

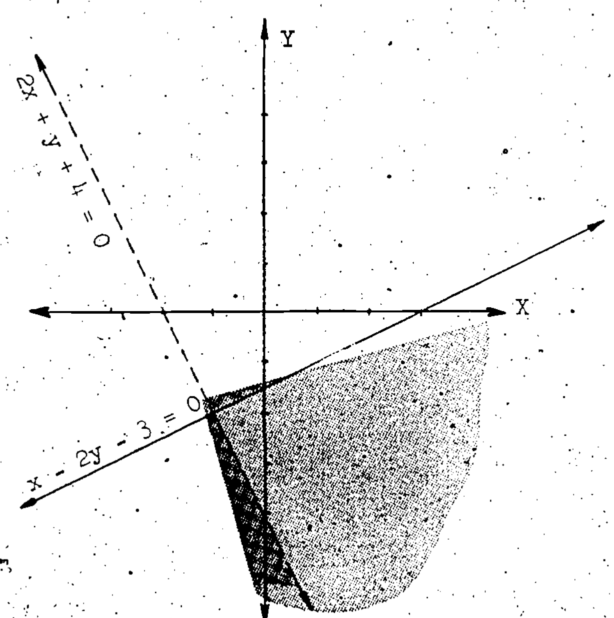


Figure 6

Note: The doubly shaded region, along with the solid half-line, represents the solution set of the system of inequalities.

Notice that the shaded region is unbounded; that is, it extends indefinitely to the right.

If the system had been written

$$x - 2y - 3 \geq 0 \text{ OR } 2x + y + 4 > 0,$$

then the graph of the solution set would be represented by all of the regions shaded in Figure 6, and the line $x - 2y - 3 = 0$. In set-builder notation the solution set is

$$\{(x, y) : x - 2y - 3 \geq 0\} \cup \{(x, y) : 2x + y + 4 > 0\}.$$

Graphs of solution sets of systems like

$$\begin{cases} 3x - 2y + 6 > 0 \\ x + 2y - 10 < 0 \\ x \geq 0 \\ y \geq 0 \end{cases}$$

can be found in a similar manner.

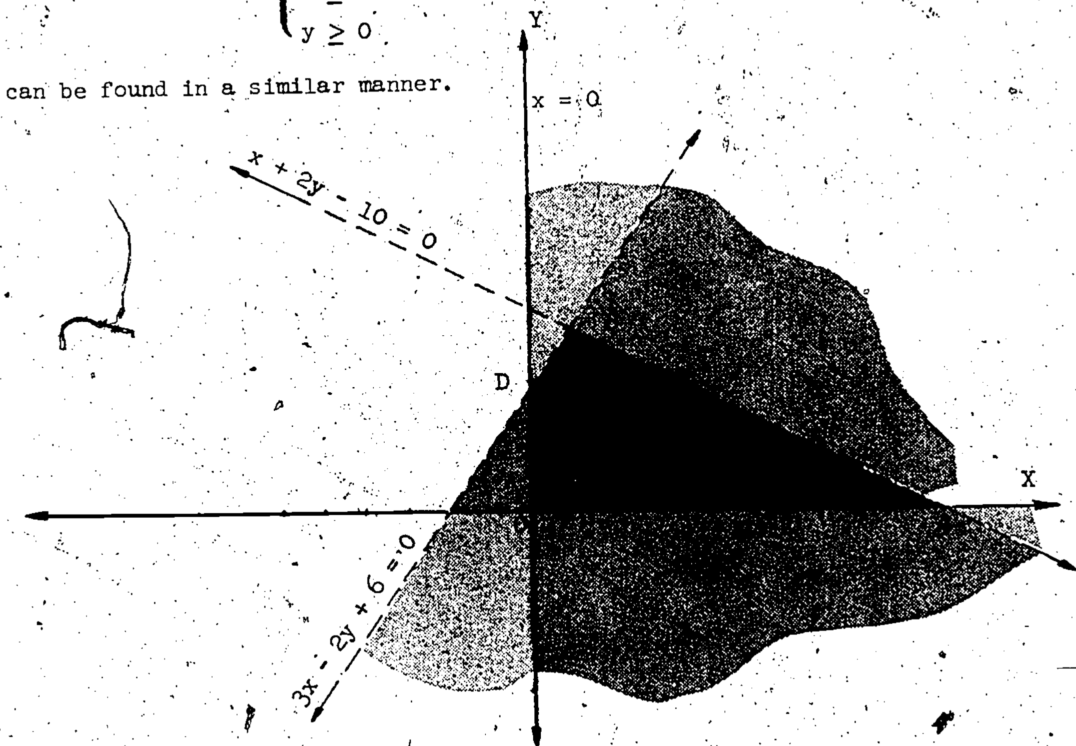


Figure 7

The solution set is represented by the region in the interior of ~~the~~ quadrilateral ODBA, and the segments $\overline{OD}$ and $\overline{OA}$ not including the end points A and D. Notice that the shaded region is <u>bounded</u>; that is, it is enclosed and a definite number will represent its area. In set-builder notation the solution set is

$$\{(x,y): 3x - 2y + 6 > 0\} \cap \{(x,y): x + 2y - 10 < 0\} \cap \{(x,y): x \geq 0\} \cap \{(x,y): y \geq 0\}.$$

In this kind of situation it is often simpler to shade the region that is <u>not</u> in the solution set of each of the sentences as in the following figure.

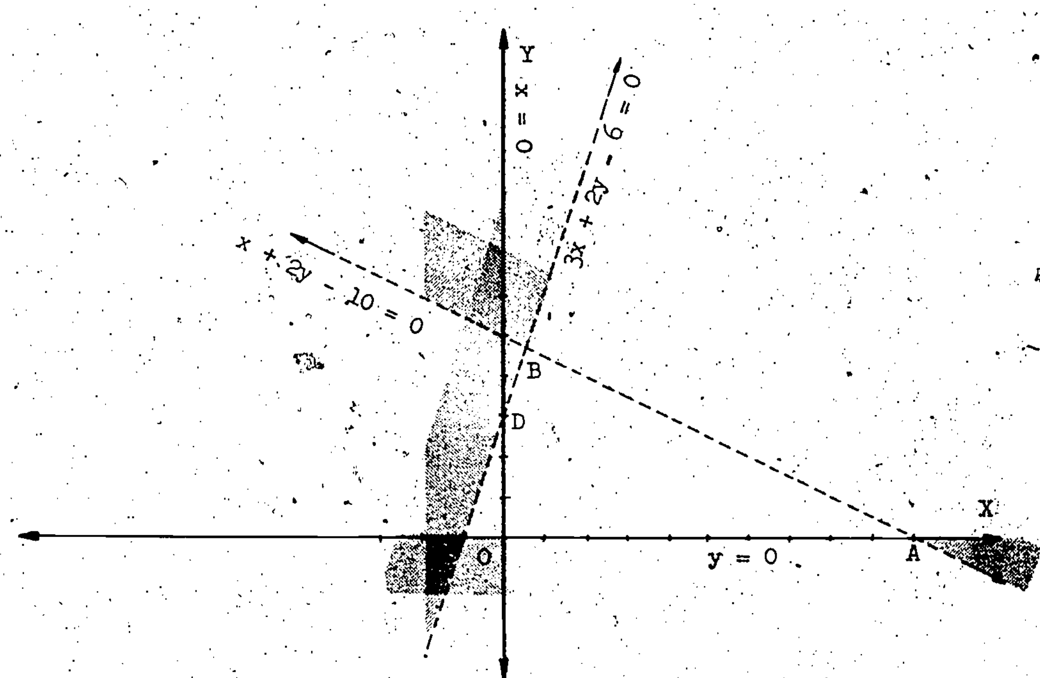


Figure 8

The solution set now is represented by the unshaded interior of the quadrilateral ODBA, and the line segments $\overline{OA}$ and $\overline{OD}$ not including the end points A and D.

Exercises 26-6

On separate coordinate axes graph the solution sets of the following systems of inequalities. State whether the solution set is bounded or unbounded.

1. $\begin{cases} x \geq 0 \\ y \geq 0 \\ x \leq 2 \\ y \leq 3 \end{cases}$ 2. $\begin{cases} y < x \\ x > 2 \end{cases}$ 3. $\begin{cases} 2x + 3y < 1 \\ x - y > 2 \end{cases}$ 4. $\begin{cases} x - y - 2 < 0 \\ x - y + 5 > 0 \end{cases}$

5. $\begin{cases} 2x + y > 2 \\ x + y < 1 \\ x \leq 3 \end{cases}$ 6. $\begin{cases} |y| < 2 \\ |x| < 1 \end{cases}$ 7. $\begin{cases} x + y \leq 5 \\ y \leq 3x + 4 \\ y \leq -x + 4 \end{cases}$

8. $\begin{cases} x + y \leq 2 \\ 2x + 2y \leq 4 \end{cases}$ 9. $\begin{cases} x + y \leq 2 \\ 2x + 2y \geq -5 \end{cases}$

10. $2x - y - 5 \leq 0$ or $x - 3y + 6 \geq 0$

11. $(2x - y - 5)(x - 3y + 6) < 0$ (Remember $ab < 0 \Leftrightarrow (a < 0$ and $b < 0)$ or $(a > 0$ and $b > 0)$)

12. $\begin{cases} -4 < x < 4 \\ -3 < y < 3 \end{cases}$

26-7. Applications

As indicated in the beginning sections, systems of sentences in two variables can serve as mathematical models for many different practical situations. The purpose of this section is to give you some practice in: (1) constructing systems of sentences in two variables that are good mathematical models, (2) finding the solution sets of the systems, and (3) interpreting the mathematical results in the language of the original situations. You should choose the method of finding the solution set which seems easiest to use. We will first consider some problems whose model is a system of linear equations.

Exercises 26-7a

(Class Discussion)

"Going with the tide a boat traveled 15 miles in 30 minutes. Returning halfway against the tide required 45 minutes. Find the speed of the boat in still water and find the speed of the tide."

1. Let x represent a speed of a boat in still water in miles per minute and y represent a speed of a tide in miles per minute. Then the output of the function

$$f : (x, y) \to (x + y)$$

is described as ______?______

2. The output of the function

$$g : (x, y) \to (x - y)$$

is described as ______?______

3. The output of the function

$$h : (x, y) \to (x + y)(30)$$

is described as _____?_____.

4. The output of the function

$$k : (x, y) \to (x - y)(45)$$

is described as _____?_____.

5. What do 7.5 and 15 represent in the above problem?

6. The system of equations representing this situation is $\begin{cases} ? \\ \end{cases}$

7. The solution set is $\{(\ ?\ )\}$.

8. The speed of the boat in still water is __?__ mpm, and the speed of the tide is __?__ mpm. (Note that the speed is in miles per minute.)

Exercises 26-7b

Analyze each of the following problem situations and write a system of equations which will serve as a mathematical model. Find the solution set of the system and answer the questions asked in each problem.

1. A bank teller has 154 bills of one-dollar and five-dollar denominations. He thinks his total is \$465. Has he counted his money correctly? Justify your answer.

2. Find an equation of the line which contains the intersection of the lines $5x - 7y - 3 = 0$ and $3x - 6y + 5 = 0$, and which passes through the origin. (Hint: What is the value of C so that $Ax + By + C = 0$ is a line through the origin?)

3. To chlorinate the water supply in Shady Hills, Horatio Algae must prepare each day exactly 20 gallons of a solution that is 35 percent chlorine. To do this, he mixes an 80 percent solution of chlorine with water that is 1 percent chlorine. How many gallons of the 80 percent solution should he use?

4. As a boy, Horatio Algae sold papers on a street corner. He received 1 cent for each paper sold on a week day, and 2 cents for each paper sold on Sunday. During one week he sold 1700 papers, including

Sunday sales. He received 22 dollars for the week. How many papers did he sell on Sunday?

5. Students in the mathematics club at Shady Hills School held a benefit Fiesta to raise money for their Spring Vacation trip to Acapulco. They sold enchilada dinners for 75 cents and taco dinners for 50 cents. The total number of dinners they prepared was 6000. The club needed $4000 to take the trip to Acapulco. How many dinners of each kind should they prepare?

6. Observations have shown that the Fahrenheit temperature is 39 more than the number of chirps made by the common black cricket in 15 seconds. What is the temperature when the number of chirps per 15 seconds is exactly two-fifths of the Fahrenheit temperature?

7. Two grades of gasoline are mixed, the one selling for 35 cents a gallon, and the other selling for 28 cents a gallon. How many gallons of each grade must be blended to obtain 500 gallons of the mixture to sell for 30 cents a gallon?

8. A man made two investments, the first at 4 percent, and the second at 6 percent. He received a yearly income from them of $400. If the total investment was $8000, how much did he invest at each rate?

9. An artist was called upon to design a mobile to decorate the entry to a large building. One of the parts consisted of two large, abstractly shaped pieces of cast iron. One piece weighed 50 pounds and the other 135 pounds. The pieces were to be suspended from the ends of a steel rod 20 ft. long. At what point should the attachment to the steel rod be made so that the mobile will balance and hang properly? (That is, where is the balance point?)

10. Foreign agents in a Mercedes SL 300, traveling 120 miles per hour, passed a group of hidden FBI agents on a highway. The FBI immediately radioed an alert to other agents and set out in pursuit in their supercharged Cobra. They had traveled only a short distance when the Mercedes hurtled past them going in the opposite direction traveling at their top speed of 120 mph. The Mercedes was finally stopped by a road block, but the secret documents they had stolen were not found. The FBI agents said that the Mercedes passed them the second time just 10 miles beyond the point where they were originally

hidden, and 17 minutes after they had passed them the first time. Assuming that the foreign agents took two minutes to make a stop, turn around, hide the documents, and reach top speed again, how far from the original observation point should the search start to recover the stolen documents?

____

Now let's consider some problems like the "decision problem" discussed at the beginning of the chapter. Most of the problems will be stated simply. We will consider only problems involving two variables so that we can illustrate our results graphically. However, the basic methods could be extended to solve more complicated problems.

Exercises 26-7c

(Class Discussion)

Transportation Problem: Suppose that a company maintains two warehouses in California containing transistor radios. One warehouse is in San Francisco and the other is in Los Angeles. Suppose that you want to transport by air 300 radios to stores in Chicago and 180 to New Orleans. Each store wants at least twice as many "Super Static" radios as "Wavering Sound" radios. "Super Static" radios are stored only in Los Angeles and "Wavering Sound" radios are stored only in San Francisco. If the following is a table of shipping costs per radio, what should the shipping order be to minimize the cost of transportation?

		(to) Chicago	New Orleans
(from)	S.F.	\$1.00	\$2.00
	L.A.	\$2.00	\$3.00

1. If x represents a number of radios shipped from Los Angeles to Chicago, then the output of the function

$$f : x \to 300 - x$$

can be described as _______________

If y represents a number of radios shipped from Los Angeles to New Orleans, then the output of the function

$$g : y \to 180 - y$$

can be described as ______________________.

2. Fill in the following shipping table showing the results of Exercise 1.

		(to) Chicago	New Orleans
(from)	S.F.		
	L.A.	x	y

3. Using the given table of shipping costs describe the outputs of the following functions:

(a) $h : x \to (2)\, x$

(b) $i : x \to (1)\,(300 - x)$

(c) $j : y \to (3)\, y$

(d) $k : y \to (2)\,(180 - y)$

4. $C = 2x + (300 - x) + 3y + 2(180 - y)$ represents the <u>total shipping cost</u>. Simplify this equation.

5. Each store wants at least twice as many "Super Static" radios as "Wavering Sound" radios. This means that $x \geq 2(300 - x)$ and $y \geq 2(180 - y)$. Simplify these sentences.

6. Since we know that a negative number of radios cannot be shipped, we want the number pair (x, y) satisfying the conditions:

$$x \geq 200$$
$$y \geq 120$$
$$300 - x \geq 0$$
$$180 - y \geq 0$$

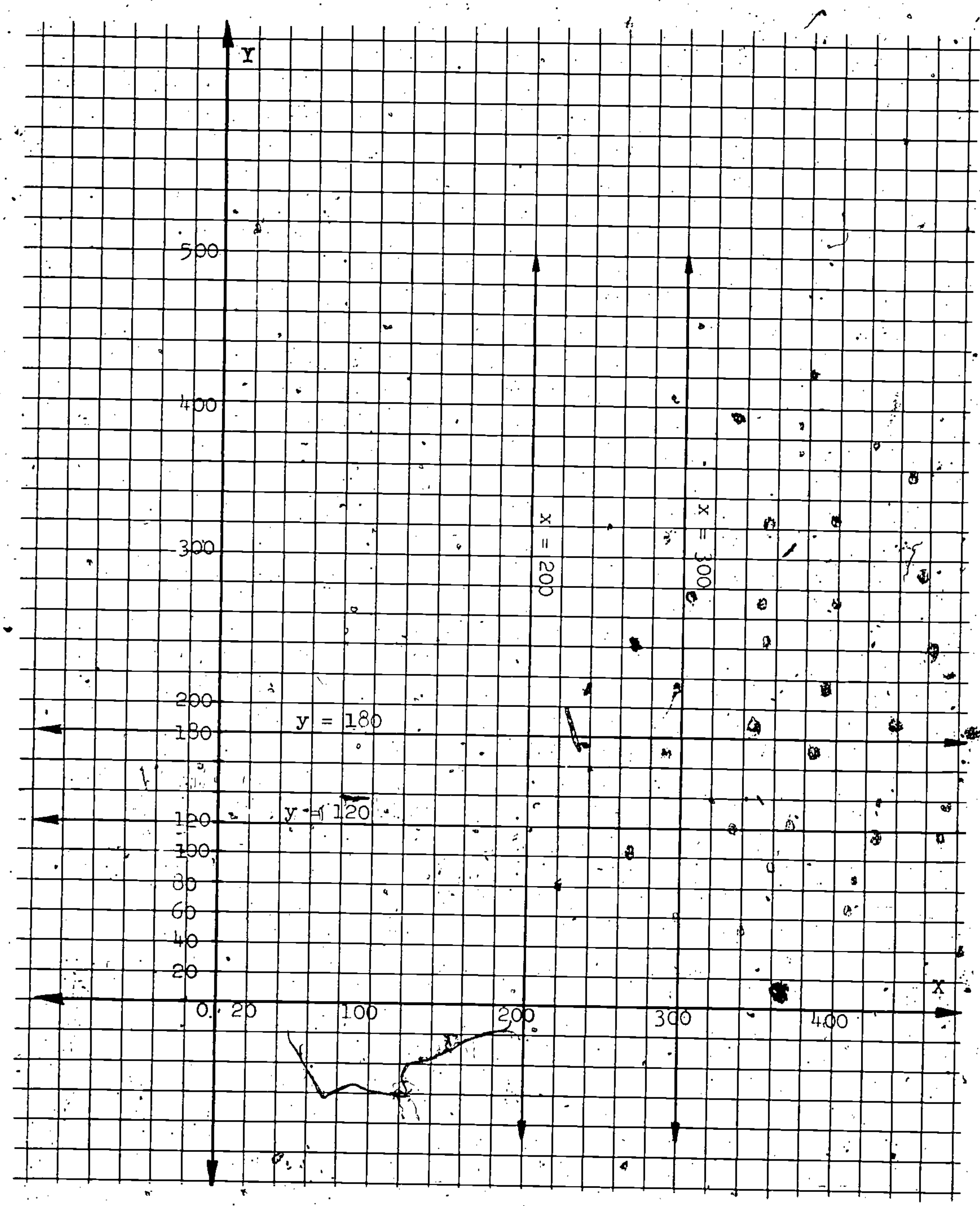
for which the cost C

$$C = x + y + 660$$

is a ______________ (minimum or maximum)

7. The shaded region is a graph of the solution set of the inequalities listed in Exercise 5. Copy the graph and draw the lines representing

the "Cost Equation" for $C = \$1160$, $C = \$860$, $C = \$1060$, and $C = \$980$.



8. What are the coordinates of the point where the cost will be a minimum?

9. Minimizing the cost, how many radios should be shipped from San Francisco? How many should be shipped from Los Angeles?

(to)

(from)		Chicago	New Orleans
	S.F.	?	?
	L.A.	200	?

Diet Problem: "Jack Spratt could eat no fat, his wife could eat no lean." Jack must eat at least 12 pounds of lean meat per week and his wife needs no more than 6 pounds of fat per week. The beef they buy is 10 percent fat and the pork they buy is 40 percent fat. If the cost of beef is \$1.00 a pound and the cost of pork is \$1.00 a pound, how much of each should they buy per week to satisfy their diet problems and minimize the cost, if they have to buy at least 10 pounds of beef and at least 5 pounds of pork, per week?

10. If x represents a number of pounds of beef purchased per week, then describe the output of the functions:

$$f : x \to .90x$$
$$\text{and } g : x \to .10x.$$

11. If y represents a number of pounds of pork purchased per week, then describe the output of the functions:

$$h : y \to .60y$$
$$\text{and } k : y \to .40y.$$

12. Describe the outputs of the following functions:

$$p : (x, y) \to .90x + .60y$$
$$q : (x, y) \to .10x + .40y.$$

13. $C = x + \underline{\ ?\ }$ represents the cost of x pounds of beef and y pounds of pork per week.

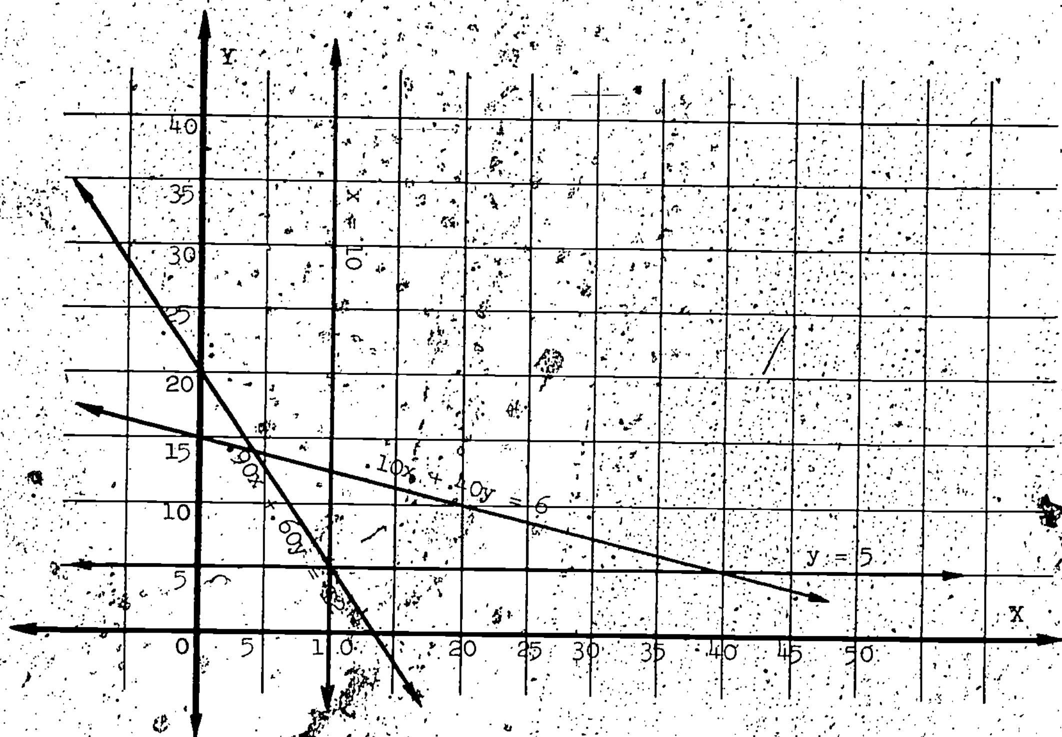
14. A mathematical model of the situation is

$$x \geq 10 \text{ and}$$
$$y \geq 5$$
$$.90x + .60y \geq \ ? \text{ and}$$

$.10x + .40y \le ?$,

for which the cost C, $C = x + y$ is a $\underset{\text{(minimum or maximum)}}{\underline{\quad ? \quad}}$.

15. The shaded region is a graph of the solution set of the inequalities listed in Exercise 14. Copy the graph and draw the lines representing the "Cost Equation" for $C = 10$, $C = 20$ and $C = 15$ where C is in dollars.



16. What are the coordinates of the point where the cost will be a minimum?

17. How many pounds of beef are bought per week?
How many pounds of pork are bought per week?

18. How many pounds of lean meat does Jack get?
How many pounds of fat does his wife get?

Exercises 26-7d

1. Blending Problem: A drugstore buys two different compounds, "X" and "Y", which contain the following quantities of Vitamins A and B:

	"X"	"Y"
Vitamin A	5 units per oz.	5 units per oz.
Vitamin B	10 units per oz.	5 units per oz.

"X" costs 5 cents per oz. and "Y" costs 10 cents per oz. The druggist wants a mixture of "X" and "Y" to contain at least 20 units of Vitamin A and no more than 60 units of Vitamin B. In order to satisfy company regulations the amount of Compound Y used must be at least three times the amount of Compound X. How many ounces of each does he use to make the mixture as cheaply as possible?

2. Horatio Alger designed two deluxe mousetraps. One model was called "Le Guillotine" and the other model was called "The Catapult." Since "Le Guillotine" is the most popular mousetrap, Horatio produces at least four "Le Guillotine" for every "Catapult" mousetrap produced. In his factory the mousetraps were built using two automated machines. To build each mousetrap, some work must be done on each machine. The time in minutes required on each machine for each mousetrap is given in the following table:

	"Le Guillotine"	"The Catapult"
Machine I	3	3
Machine II	[illegible]	5

Because of other important production requirements neither machine can build mousetraps for more than 150 minutes per day. Horatio makes a profit of $4 on "Le Guillotine" and a profit of $8 on "The Catapult." How many of each kind of mousetrap should Horatio build per day to get the maximum profit? (We assume that he can sell all that he can build.)

3. There are two student workers in the school cafeteria, Oliver and Shadow. Oliver earns \$2.00 an hour and Shadow earns \$1.50 an hour. Oliver can make 10 pies and 4 cakes in an hour and Shadow can make 6 pies and 4 cakes per hour.

	Pies	Cakes
Oliver	10	4
Shadow	6	4

No more than sixty pies and at least thirty-two cakes are needed. How many hours should each student work if it is desired to keep the labor cost at a minimum?

4. A store stocks two brands of transistor radios, "Super Static" and "Wavering Sound". The store owner has room for <u>no more</u> than sixty transistor radios in the store. He also knows that <u>at least</u> twice as many "Super Static" radios are sold as "Wavering Sound". He makes \$10 profit on "Super Static" radios and \$12 profit on "Wavering Sound" radios. How many of each kind should he stock to make the maximum profit?

5. A dairy farmer produces two grades of milk, Extra-Rich and Lo-Fat. For every gallon of Extra-Rich he produces from one to three gallons of Lo-Fat. He cannot produce more than ten gallons of Extra-Rich milk in one day. He sells the Extra-Rich milk to a commercial dairy 5 miles away at a profit of 20 cents a gallon. He sells the Lo-Fat milk to another dairy 3 miles away at a profit of 10 cents a gallon. He has only enough time and truck space to transport 60 gallon-miles per day. (This means that in one day 60 gallons can be moved one mile, or 30 gallons moved 2 miles, or 10 gallons moved 6 miles, etc.) How many gallons of each kind of milk should the farmer produce in order to maximize the profit?

26-8. <u>Summary</u>

<u>Section 26-2.</u>

A mathematical model of certain kinds of decision problems can be constructed and useful information can be derived from that model. The solution to such a problem is found on the boundary of the region described by the conditions stated in the mathematical model. We maximize (or minimize) a quantity by moving a line away (or toward) the origin so that it remains parallel to its original position.

<u>Section 26-3.</u>

A wide variety of problems in medicine, business, government, and industry can be described and solved using linear programming methods.

<u>Section 26-4.</u>

If a system of equations

$$\begin{cases} Ax + By + C = 0 \\ Dx + Ey + F = 0 \end{cases}$$

has a solution, then the coordinates of that point will also be a solution of

$$K(Ax + By + C) + M(Dx + Ey + F) = 0, \quad K \cdot M \neq 0.$$

We can derive an equivalent system of equations using this fact so that one equation represents a <u>vertical</u> line through the point of intersection and the other equation represents a <u>horizontal</u> line through the point of intersection.

$$\begin{cases} x = a \\ y = b \end{cases}$$

We achieve this result by choosing K and M so that in one linear combination the y-terms are eliminated, and in the other linear combination K and M are chosen so that the x-terms are eliminated.

Section 26-5.

The graph of a system of two linear equations;

(a) can intersect in one point,

(b) can be two parallel lines,

or (c) can be the same line.

Under these circumstances we say that the solution set of the system,

(a) contains exactly one ordered pair of real numbers,

(b) is the empty set,

or (c) is an infinite set of ordered pairs of real numbers.

We can use the slope-intercept form of the linear equation, $y = mx + b$, to recognize which situation exists. If the slopes are not equal the lines intersect in one point. If the slopes are equal but the y-intercepts are different, then the lines are parallel. If the slopes are equal and the y-intercepts are equal, then the lines are coincident.

If two lines intersect, then we note that the ordinate of the point of intersection is the same for both equations. We use this fact to develop another method of finding the solution of a system of linear equations. We call this method substitution and simply replace "y" in one equation by the expression in x found by putting the other equation into y-form.

In a system of linear equations $\begin{cases} ax + by + c = 0 \\ dx + ey + f = 0 \end{cases}$

if $\frac{a}{b} = \frac{b}{e} = \frac{c}{f}$, then the lines are coincident;

if $\frac{a}{b} = \frac{b}{e} \neq \frac{c}{f}$, then the lines are parallel;

if $\frac{a}{b} \neq \frac{b}{e}$, then the lines intersect in a point.

Section 26-6.

The solution sets of systems of inequalities can be found by graphing the related linear equations, and then determining which region is described by each inequality. Usually the region can be determined by using the coordinates of the origin or some point on the vertical axis to decide which half plane is in the solution set. If the coordinates of the point make the inequality a true statement, then all of the coordinates of points on that side of the line are in the solution set. If they do not make it a true statement, then the coordinates of points on the opposite side of the line are in the solution set. The inequality can also be put into y-form and the region that is in the solution set can be determined by seeing which part of the y-axes has points whose coordinates satisfy the inequality.

Section 26-7.

Systems of linear equations and systems of inequalities can serve as mathematical models of a large number of different situations. We use systems of sentences to derive important mathematical results, as well as to derive important results in various applications of mathematics.