From: Downie and Heath (1975). "Basic Statistical Methods." Harper and Rowe.

Standard Scores and the Normal Curve

Chapter 6

STANDARD SCORES

The formula for the standard score is

$$z = \frac{X - \bar{X}}{s} = \frac{x}{s} \qquad \textbf{(6.1)}$$

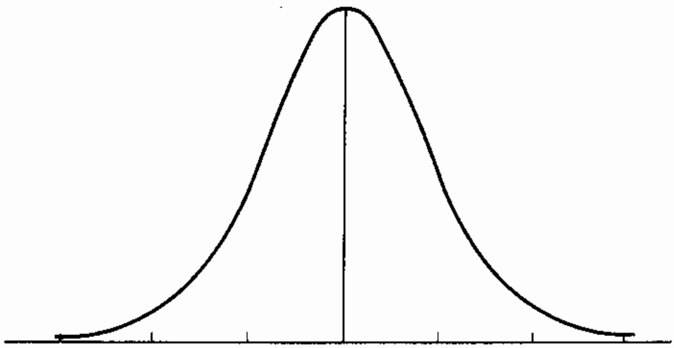
where X = any raw score or unit of measurement
$\bar{X}, s$ = mean and standard deviation of the distribution of scores

To illustrate the computation and nature of standard scores, let us take the following scores, which are a part of a distribution with a mean of 60 and a standard deviation of 10:

X	x	z
70	10	1.00
60	0	.00
50	−10	−1.00
54	−6	−.60
46	−14	−1.40

In the first column we have the raw scores (X). The mean is subtracted from each of these, and then this deviation from the mean, or x, is divided by the standard deviation to change the deviation values into standard score values. The raw score of 60 is at the mean. There is no deviation; hence the standard score is zero. A raw score of 70 is 1 standard deviation above the mean. This results in a z score of 1. When we change raw scores to standard scores, we are expressing them in standard deviation units. These standard scores tell us how many standard deviation units any given raw score deviates from the mean.

Since 3 standard deviations on either side of the mean include practically



z-Scores	−3	−2	−1	0	+1	+2	+3
Navy Scores	20	30	40	50	60	70	80
AGCT	40	60	80	100	120	140	160
CEEB	200	300	400	500	600	700	800

Figure 6.1 Distribution of the various types of standard scores.

all of the cases, it follows that the highest z score usually encountered is +3 and the lowest is −3. We can describe the distribution of z scores by saying that they have a mean of zero and a standard deviation of 1. This is shown in Figure 6.1. Thus any time we see a standard score, we should be able to place exactly where an individual falls in a distribution. A student with a z score of 2.5 is 2.5 standard deviations above the mean on that test distribution and has a very good score. These standard scores are equal units of measurement and hence can be manipulated mathematically. It should be noted here also that changing a distribution of scores to z scores does not change the shape of the original distribution of scores. If the distribution was positively skewed to begin with, z scores made from such a distribution would be positively skewed.

Since z scores are expressed in decimals and since about half of them are negative, they are rather cumbersome to handle. Many times so-called linear transformations are made. Such transformations consist of making the scale larger, so that negative scores are eliminated, and of using a larger standard deviation, so that decimals are done away with. Transformed scores can be obtained from the following equation:

$$\text{Standard score} = z(\text{new standard deviation}) + \text{the new mean}$$

A common form for these transformations is based upon a mean of 50 and a standard deviation of 10. In equation form this becomes

$$\text{Standard score} = z(10) + 50$$

or, starting with the raw scores, we have

$$\text{Standard score} = \frac{(X - \bar{X})}{s}(10) + 50$$

This system with a mean of 50 and a standard deviation of 10 is very popular. It has been widely used by branches of the Armed Forces for many years. With this system we have a mean of 50 and a range of between 20 and 80. All scores are positive and all can be rounded to two-place numbers. Figure 6.1 shows this distribution of scores and also several others. The third row in Figure 6.1 shows the scores used on the American College Test (ACT), a test battery used to select college freshmen. Here the mean is 15 and the standard deviation is 5. The last row of these figures shows the type used by the College Entrance Examination Board and the Graduate Record Examination. Here we note a mean of 500 and a standard deviation of 100. Any system can be set up, but those just noted are most frequently encountered.

Some standard scores have been normalized. By this we mean that the distributions of these scores has been made to conform to that of the normal curve. We shall consider the normalizing of scores later. Here we shall briefly note some of these normalized standard scores. A very common standard score is the *T* score, which has a mean of 50 and a standard deviation of 10. *T* scores are frequently normalized. Then there are stanines (standard nines), which have a mean of 5 and a standard deviation of approximately 2. Many of the standard scores used in nationwide testing programs have been normalized.

Uses of Standard Scores

Since standard scores are equal units of measurement and since their size is the same from distribution to distribution, they become a very useful tool both in the reporting of test scores and in doing research using test results. When the results of different tests taken by the same individuals are to be compared, this is best done by the use of standard scores. This process is illustrated in Table 6.1.

In part A of Table 6.1 the scores of students on three elementary school tests are presented. Only the scores of three students are listed. At the bottom are shown the mean and standard deviation of each test. Looking at the data in part A of Table 6.1, it becomes apparent that, as they stand, they convey little meaning. Which student had the best overall performance? On which test did the different students do best? or worst? None of these questions can be answered as the scores are shown.

Suppose that we now change these scores to standard scores. For this we shall use the system with a mean of 50 and a standard deviation of 10. We shall start with the geography test, and for all of these we shall use this transformation equation:

$$\text{Standard score} = \frac{X - \bar{X}}{s}(10) + 50$$

The standard score in geography for student A in this test is

$$SS = \frac{60 - 60}{10}(10) + 50$$
$$= 0 + 50$$
$$= 50$$

The geography score for student B becomes

$$SS = \frac{72 - 60}{10}(10) + 50$$
$$= \frac{120}{10} + 50$$
$$= 12 + 50$$
$$= 62$$

This is continued until all scores are transformed. The results of transforming the scores in part A of Table 6.1 are shown in part B. This may seem to be a laborious and time-consuming process. However, if there are many scores to be transformed, it is most convenient to arrange the scores from high to low and find the deviation of each from the mean (since these are in order, the rest of the process becomes very easy). Scores with a negative deviation are the same distance below the new mean as the positive scores with the same deviations are above it.

Table 6.1 Comparing and Combining Scores Made on Different Tests by the Use of Standard Scores

	Part A: Raw Scores		
Student	Geography	Spelling	Arithmetic
A	60	140	40
B	72	100	36
C	46	110	24
etc.			
Mean	60	100	22
Standard Deviation	10	20	6

	Part B: Standard Scores			
Student	Geography	Spelling	Arithmetic	Average
A	50	70	80	67
B	62	50	73	62
C	36	55	53	48
etc.				

Let us now examine part B of Table 6.1. Note that student A is at the mean in geography, 2 standard deviations above the mean in spelling, and 3 standard deviations above the mean in arithmetic. His average performance on these three tests was 67, 1.7 standard deviations above the mean. For student A we can then say that his performance is average in geography, excellent in spelling, and superior in arithmetic. In this manner we can consider the achievement of each of the students on each of the three tests, and we can get an average measure of his performance on all three tests. It should be noted that the only justifiable manner to compare scores is to first change the scores to standard scores, and then do the comparing. Standard scores change the raw scores to equal and comparable units.

Another use of standard scores is in determining final grades for a course. Let us take a course which has three examinations of 1 hour each during the semester and a final examination 2 hours long. At the end of the semester the instructor averages the three 1-hour examinations and then combines this average with the final examination in some manner or other. This method is not correct. Suppose that the means and standard deviations of the three 1-hour examinations were as follows:

	First Test	Second Test	Third Test	Final Exam
$\bar{X}$	52	87	62	124
s	8	17	11	21

If a student's scores on these three 1-hour examinations are averaged, each will not contribute equally to the average. The second test with the largest standard deviation would contribute more to the final average than the other two, and the first examination would contribute least. The proposed method tends to equalize contributions of 1-hour examinations to final scores.

In grading situations, if the teacher wants each of a series of tests to contribute equally to the final grade, it follows that the scores of individuals on each test should be changed to standard scores and that these should be averaged. Then each 1-hour test is contributing more equally to the final grade. Suppose now that the instructor wants the final examination to have twice as much weight as each of the other tests in the final grade. These final grades should also be changed to standard scores. Then for any student his overall average is determined by taking his standard score on each of the first three tests, adding these, then adding to this sum 2 times his standard score on the final examination and dividing the total by 5. In symbols:

$$\frac{z_1 + z_2 + z_3 + z_f(2)}{5}$$

where z_1, z_2, z_3 = the standard scores on the three 1-hour examinations

zf = standard score on the final examination

2 = weight final examination is to have

This method gives each examination the weight that the instructor desires it to have. Some teachers convert all scores to standard scores before entering them into their class books. This is strongly recommended.

Science teachers and some others have special problems with laboratory grades, project grades, and the like. It is desired that the laboratory grades, for example, contribute a certain amount to the final grade. This can only be done with any accuracy by changing these to standard scores and giving them the desired weight in the final average as shown above for the combination of 1-hour examinations and finals.

THE NORMAL CURVE

General Nature of the Normal Curve

For some time we have been talking about the normal curve. Now we shall examine it in detail, discuss its characteristics, its properties, and some of the basic ways in which it can be used. In the eighteenth century gamblers were interested in the chances of beating various gambling games and they asked mathematicians to help them out. DeMoivre (1733) was the first to develop the mathematical equation of the normal curve. In the early nineteenth century Gauss and LaPlace further developed the concept of the curve and probability. It was at about the same time that errors of observation made by astronomers were represented by a curve of this type. Today the normal curve is referred to as the curve of error, the bell-shaped curve, the Gaussian curve, or DeMoivre's curve.

By now the shape of this curve is familiar to you. Its maximum height is at the mean. In the language of the curve we say that the maximum ordinate (ordinates are given the symbol y) is at the mean. All other ordinates are shorter than this one. The normal curve is also said to be asymptotic. By this we mean that theoretically the tails never touch the base line but extend to infinity in either direction. In actual practice, however, 3 standard deviations on either side of the curve will include practically all of the cases. As mentioned previously, the skewness of the normal curve is zero and its peakedness is described as being mesokurtic.

In our educational and psychological work, we assume that certain traits are normally distributed. In actuality, probably no distribution ever takes on the absolute form of the normal distribution. Many of our frequency distributions are very close to the normal one, and we assume that they have a normal distribution. To the extent that our distributions differ from normal, error enters into our work. The normal curve is important not primarily because *scores* are assumed to be normally distributed, but because the *sampling* distributions of various statistics are known or assumed to be normal. Hence the normal curve's importance is primarily in sampling statistics. (This will be discussed in Chapters 11 and 12.)

Appendix B is based upon a normal curve with a mean of zero, a standard deviation of 1, and an area reduced to unity. It is referred to as the unity normal curve or the normal curve in standard score units. In solving problems we assume that our cases are spread evenly over this area. Suppose that we have a problem with 500 cases. In solving the problem we find that .262 of the area of the curve falls above a certain point. To determine the number of cases above this point, we merely multiply our proportion by the number of cases, in this example .262(500). We also use this area of the curve in talking about probability. Suppose that as an illustration we take a raw score which has an equivalent standard score of +1.88. By referring to the normal probability table, Appendix B, we note that approximately .03 of the area of the curve falls above this point. We then say that the probability of obtaining a score equal to or higher than this one of +1.88 is .03 or 3 in 100.

Areas Under the Normal Curve

In the chapter in which we discussed standard deviations, we mentioned some of the relationships between standard deviation units and the normal curve. The most important of these relationships are again repeated in Figure 6.2. To summarize, 1 standard deviation taken on each side of the mean includes a total area of 68.26 percent of the curve, or approximately two-thirds of the cases. In terms of probability we can state that the chances are about two out of three of a score in any normally distributed sample falling within the area of 1 standard deviation on each side of the mean. A second standard deviation measured beyond the first cuts off 13.59 percent of the area. By adding all of the area included by 2 standard deviation units on both sides of the mean, we have accounted for more than 95 percent of the area or cases. If we continue and measure off another, or third, standard deviation on each side of the mean, we cut off another piece equal to 2.15 percent of the area. The sum of all of the areas included by these 6 standard deviation units is equal to 99.74 percent

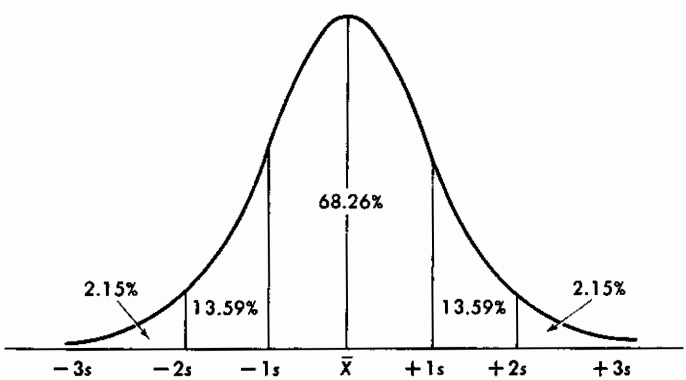


Figure 6.2 Percentages under the normal curve at various standard deviation units from the mean.

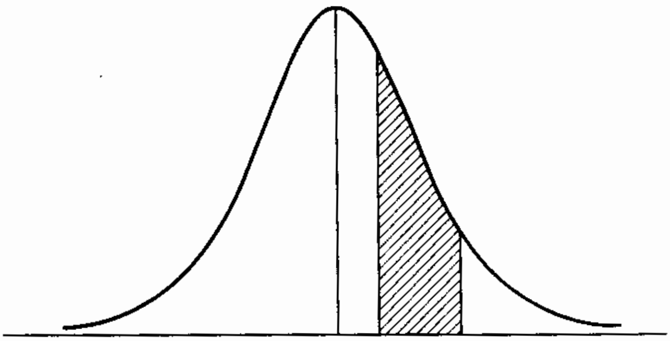


Figure 6.3 Percentage of the area of the normal curve between two points on the same side of the mean.

of the total. From this it follows that .26 percent of the cases are beyond three standard deviation units from the mean. This means 26 in 10,000. Dividing this by 2 to distribute these equally on both sides of the mean, we see that on each side we can expect 13 cases in 10,000 to fall beyond 3 standard deviation units from the mean.

AREAS CUT OFF BETWEEN DIFFERENT POINTS. Suppose, to begin with, we take the problem illustrated in Figure 6.3. We wish to find the proportion of the area or of the number of cases included between the two points $+.5s$ and $+1.76s$ units above the mean. The problems are all worked using the normal probability table, Appendix B. We begin by going to the table and finding the area of the curve cut off between the mean and a point equivalent to a standard score of .5 above the mean. This value appears in column 2 of the table and is found to be .1915. Next we continue down column 1 of the table until we come to a standard score of 1.76. By looking in column 2, we find that .4608 of the area is included between the mean and this point. Then the area of the curve between these two points is the difference between the two points, $.4608 - .1915$, which equals .2693. We can then state that approximately 27 percent of the cases fall between these two points, or that the probability of a score's falling between these two points is .27.

In the next illustration, we shall take two points on different sides of the mean. This time we wish to determine what proportion of the normal curve falls between a standard score of $-.48$ and one of $+1.5$ (Figure 6.4). There are no values for negative standard scores in Appendix B. As far as areas are concerned, equal standard scores, whether positive or negative, include equal areas when taken from the mean. From the table we find that a standard score of $-.48$ cuts off an area of .1844 between it and the mean. A standard score of $+1.5$ likewise includes .4332 of the area of the curve between it and the mean. The area included between both points is then equal to the sum of

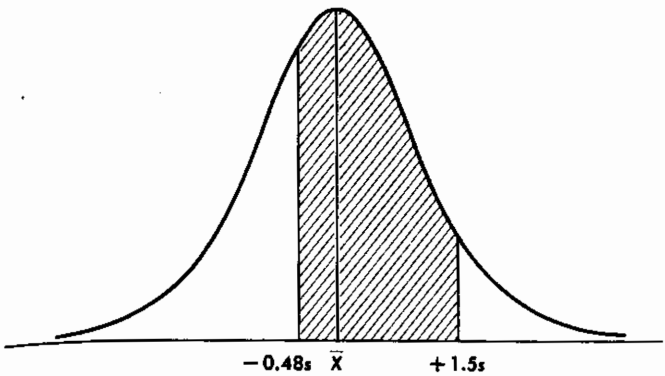


Figure 6.4 Percentage of the area of the normal curve between two points on different sides of the mean.

these two areas, .1844 + .4332, which is equal to .6176 or approximately 62 percent of the area.

Before we go any further, we might examine the other columns in Appendix B. Column 3 is labeled *Area in Larger Portion,* and column 4, *Area in Smaller Portion.* Any time we take a point on the base line of a curve and erect a perpendicular at this point, we divide the curve into two areas, a larger and a smaller area. For any given standard score, the sum of these two areas is equal to unity. Column 5 of the table, *Ordinate,* gives the size of the ordinates for the various standard scores.

Here is another type of problem that can be solved using these tables. Suppose that we have a distribution of test scores for which the following statistics have been computed: $\bar{X} = 80$, $s = 16$, $N = 510$. We wish to know what percentage of the scores in this distribution fall above a raw score of 110, assuming a normal distribution (Figure 6.5). This can of course be solved for

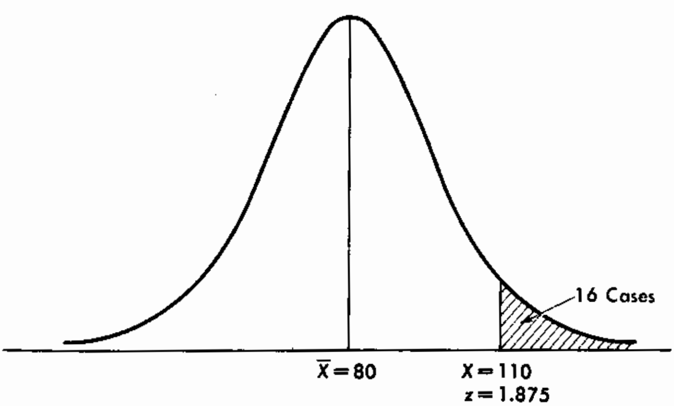


Figure 6.5 Percentage of cases in a normal distribution falling above a certain score.

any score in this distribution. Since the table of the normal curve is in standard units, the first step is to change the raw score of 110 to a standard score as follows:

$$z = \frac{X - \bar{X}}{s}$$
$$= \frac{110 - 80}{16}$$
$$= \frac{30}{16}$$
$$= 1.875$$

Inasmuch as this score is halfway between 1.87 and 1.88, we must interpolate. This time we are interested in the area in the smaller portion of the curve at a given standard score point. The value from the table is half the distance between .0307 and .0301, which equals .0304. Next we multiply our N by this proportion to find the number of cases above this raw score of 110. We find that $510(.0304) = 15.504$ cases. This in round numbers is approximately 16 cases.

POINTS ABOVE OR BELOW WHICH CERTAIN PROPORTIONS OF THE AREA FALL. Now we are going to reverse the process we have been using. Suppose that we wish to find that point in the distribution which has 10 percent of the cases below it and 90 percent of the cases above it (Figure 6.6). This point would be the tenth centile. Again we enter the normal probability table and using column 4, the area in the smaller portion, we locate the value closest to .10.

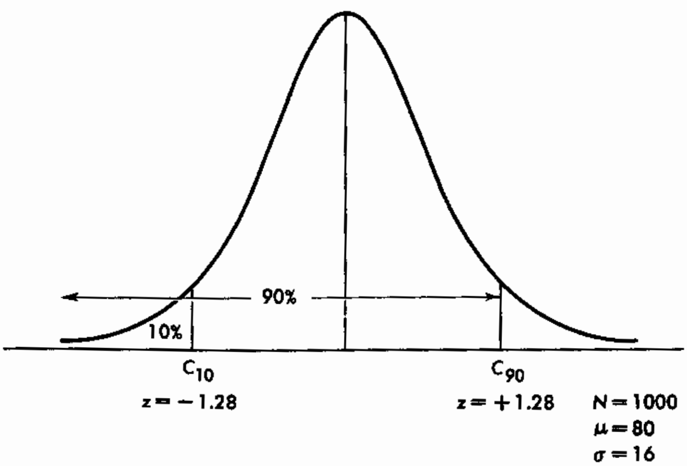


Figure 6.6 Points in the normal curve above and below which different percentages of the curve lie.

In the table we find this to be .1003, and reading to the left we find that a standard score of 1.28 divides the area of the curve into the two proportions as desired. Since this standard score is to the left of the mean, it has to have a minus sign in front of it and is correctly written as −1.28. We could have solved this problem just as well by using column 3 of the table, the area in the larger portion. We would have gone down this column until we came as close to .90 as possible. This value is .8997, which puts us in the same row as before. The procedure from this point on is identical. If we wanted to find the point in the distribution with 10 percent of the cases above it (C_{90}), the work would be identical to that just completed, except that in the end the value of the standard score would be +1.28 because this time our point is to the right of the mean.

Suppose that we now want to know above which score in the distribution 90 percent of the cases will fall. In the last paragraph we found that a z of 1.28 is the point above which 90 percent of the cases fall. To find the raw score corresponding to this z score of 1.28 we solve the equation of z for a different unknown.

$$z = \frac{X - \mu}{\sigma}$$

$$1.28 = \frac{X - 80}{16}$$

$$X - 80 = 16(1.28)$$

$$X = 80 + 20.48$$

$$X = 100.5$$

Similarly, the raw score equivalent of the point below which 10 percent of the cases fall is

$$-1.28 = \frac{X - 80}{16}$$

$$X - 80 = -20.48$$

$$X = 80 - 20.48$$

$$X = 59.5$$

NORMALIZING A DISTRIBUTION OF SCORES

In the paragraphs that follow, the process of normalizing a distribution of scores will be demonstrated. From the results, the normal curve for the data will be plotted. This normal curve for any set of data is referred to as the *curve of best fit* for that set of data. The best-fitting curve for any set of data

Table 6.2 Normalizing a Distribution of Scores

(1)	(2) f_0	(3) Upper Limit	(4) x	(5) z	(6) Proportion Below	(7) Proportion Within	(8) f_e	(9) f_e
90–94	1	94.5	30.6	2.51	.9940	.0119	1.785	1.8
85–89	3	89.5	25.6	2.10	.9821	.0276	4.140	4.1
80–84	8	84.5	20.6	1.69	.9545	.0548	8.220	8.2
75–79	12	79.5	15.6	1.28	.8997	.0919	13.875	13.8
70–74	28	74.5	10.6	.87	.8078	.1306	19.590	19.6
65–69	36	69.5	5.6	.46	.6772	.1573	23.595	23.6
60–64	12	64.5	.6	.05	.5199	.1605	24.075	24.1
55–59	18	59.5	−4.4	−.36	.3594	.1388	20.820	20.8
50–54	10	54.5	−9.4	−.77	.2206	.1016	15.240	15.2
45–49	8	49.5	−14.4	−1.18	.1190	.0631	9.465	9.5
40–44	8	44.5	−19.4	−1.59	.0559	.0331	4.965	5.0
35–39	5	39.5	−24.4	−2.00	.0228	.0148	2.220	2.2
30–34	1	34.5	−29.4	−2.41	.0080	.0080	1.200	1.2
$N = 150$						$\Sigma = .9940$		$\Sigma f_e = 149.1$
$\bar{X} = 63.9$								
$s = 12.2$								

has the same mean and standard deviation and is based upon the same number of cases as the original data.

This process will be demonstrated by using the data in Table 6.2. These data are based on 150 cases, with a mean of 63.9 and a standard deviation of 12.2. Column 1 lists the intervals and column 2 shows the observed frequencies, which are labeled this time f_0. After setting up these two columns, we proceed as follows:

1. Determine the upper limit of each interval and record these in column 3.

2. Determine the x values of column 4 by subtracting the mean, 63.9, from each of the upper limits in column 3.

3. Change each of these x values of column 4 to a standard score, z, by dividing by the standard deviation, 12.2.

4. From the normal probability table, Appendix B, determine the proportion of the area in the normal curve below each of these standard scores. For the top score, 2.51, we see in the table that .9940 of the area is below this point. These proportions are recorded in column 6.

5. The values in column 7 are determined as follows. The bottom value in column 7 has the same value as that at the bottom of column 6, for the proportion within and below this interval is the same as the proportion below the upper limit of the interval 30–34. The proportion within the second interval is obtained by subtracting .0080, the proportion within the bottom interval, from .0228, the proportion below the upper limit of the second interval. This results in a value of .0148. For the proportion within the

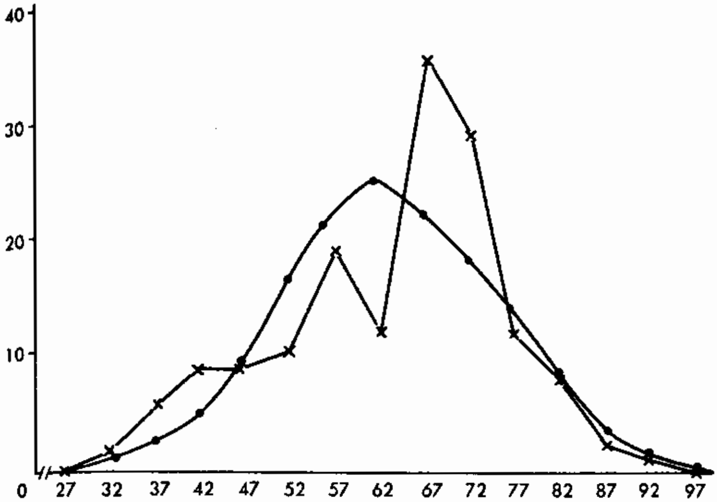


Figure 6.7 Frequency polygon and normalized curve for the same data.

third interval, we subtract .0228 from .0559. This process is continued until all the values in column 7 are determined. The sum of this column should be close to 1, but it is usually actually a little less than this, for there are always cases in the extremes of the tails of the normal curve that are not taken into account in this process.

6. The values in column 8 are obtained by multiplying the number of cases, 150, by each of the proportions in column 7. Again these add up to slightly less than 150.

7. In column 9 these expected frequencies (f_e) are rounded to the nearest tenth.

In Figure 6.7, the axes have been set up in the usual manner for constructing a frequency polygon. First the f_0's are plotted and these points connected with a rule as in the usual method for plotting a frequency polygon. Then the values in column 9 are plotted and these are connected by means of a smooth curve. In Figure 6.7 we have the curve of best fit for these data superimposed upon the frequency polygon for the original data.

The student may now be wondering when he should normalize a curve or even if there is any justification for performing the act. Later, in Chapter 14 on chi-square, we shall see that one of the uses of chi-square is to see if a distribution of measures departs from normal. To make a chi-square test, we must be able to set up the expected frequencies for any set of data. Also there are times when a distribution of a sample of scores is not normal; yet the research worker has a hunch or knows that the distribution of the trait with which he is concerned is normally distributed in the population. Since he can justify this, he may find it useful to normalize the data in his sample.