Quality Control Charts[1]

By W. A. SHEWHART

IRRESPECTIVE of the care taken in defining the production procedure, the manufacturer realizes that he cannot make all units of a given kind of product identical. This is equivalent to assuming the existence of non-assignable causes of variation in quality[2] of product. Of course, random fluctuations in such factors as humidity, temperature, wear and tear of machinery and the psychological and physiological conditions of those individuals engaged in carrying out the manufacturing procedure may give rise to some of these apparently uncontrollable variations. Knowing this, the manufacturer contents himself with trying to produce a product which is uniform and controlled—one which does not vary from one period to another by more than an amount which may be accounted for by a system of chance or non-assignable causes producing variations independent of time.

To make clear the significance of the terms "assignable causes" and "non-assignable causes," we may make use of the following illustration. Suppose a person were to fire one hundred rounds at a target. We know what probably would happen—the individual would not hit the bull's-eye every time. Possibly some of the shots would fall within the first ring, others within the second ring, and, in general, the shots would be distributed somewhat uniformly about the center of the target. We have a more or less definite picture of some of the possible reasons why the individual would not hit the bull's-eye every time, but we probably cannot assign the reasons or causes for his missing the bull's-eye in any particular instance—the causes of missing are non-assignable. Suppose, however, that the individual tended to shoot to the right of the bull's-eye. Naturally we would conclude that there was some discoverable cause for this general tendency, i.e., we would feel that the observed effect could be assigned to some particular cause.

The reason for trying to find assignable causes is obvious—it is only through the control of such factors that we are able to improve the product without changing the whole manufacturing process. But it would be a waste of time to try to ferret out or assign some cause for a

[1] A brief description of a newly developed form of control chart for detecting lack of control of manufactured product.

[2] Quality is some function of those characteristics X, Y, Z . . ., required to define a thing. For our present purpose we shall consider that quality is a function of a single characteristic X.

fluctuation in product which is no greater than that which could have resulted from the non-assignable causes as it would be to try to find the exact manner in which each of the causes contributed to missing the bull's-eye in the analogous case of target practice just considered.

Here then is the practical commercial problem—When do the observed differences between the product for one period and that for another indicate lack of control due to assignable causes, and when, on the other hand, do the differences in quality of manufactured product observed from one period to another indicate only fortuitous, chance or random effects which we cannot reasonably hope to control without radically changing the whole manufacturing process? We shall outline a typical example of the way this question arises, outline the basis for its solution and present the results in the form of a control chart.

Typical Example

Fig. 1 shows the frequency polygon for 15,050 instruments inspected for quality X. These instruments were selected at random throughout the year from a product manufactured in quantities of approximately

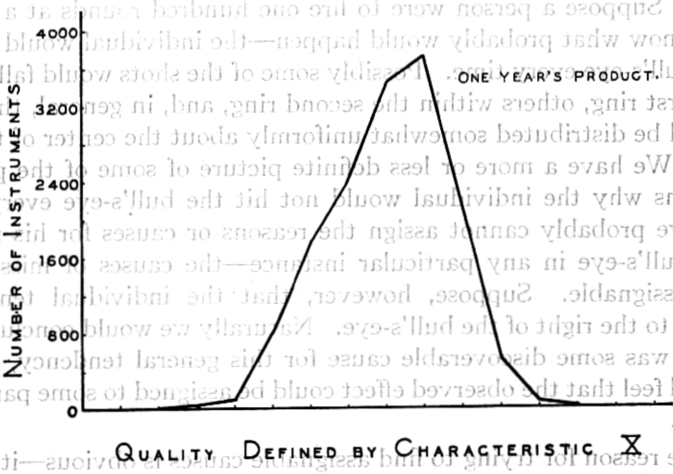

Fig. 1—Polygon showing distribution in quality for 15,050 units of product. Do these data present any evidence of lack of control?

2,000,000 per year. Is there any indication from these data that the product had not been uniform or controlled throughout the twelve month period in which the instruments had been selected?

Oftentimes we must decide from a study of a single frequency polygon of data such as given in Fig. 1, whether or not the product has been controlled during the period for which the data have been collected. In this instance, however, it was possible to group the 15,050 observations into twelve groups representing monthly samples of approximately 1250 instruments each. The data are presented in this form

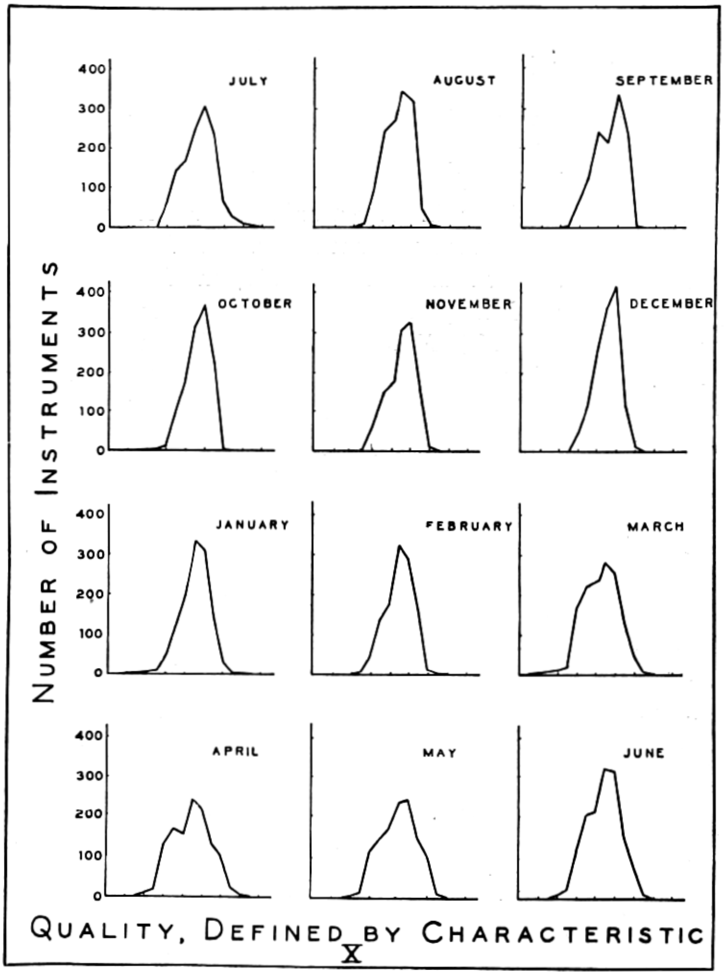

Fig. 2—Monthly polygons showing distribution in quality for samples of approximately 1250 units of product. Do these data present any evidence of lack of control?

in Fig. 2. Obviously no two polygons are the same in respect to average, dispersion and shape, but of course we would not expect them to be the same even though the product were uniform, any more than we would expect two targets to show the same distribution of shots even if the same individual had fired at both targets. In other words, non-assignable, fortuitous or chance causes introduce certain differences in the average, dispersion and shape of the observed polygons from one month to another, and we must set up some method of differentiating the effects of assignable from those of non-assignable causes.

Outline of Basis for Detecting Lack of Control

Uniform product was defined above as one for which the differences between the units or groups of units were controlled by a complex system of non-assignable chance causes producing results independent of time. Now, following a line of reasoning whose origin is attributed to Laplace, it may be shown that such a system of causes, in general, may be expected to give a unimodal distribution of product such that the probability $dy_{\lambda'}$ of the production of a unit having the quality X within the range X to $X+dX$ is independent of time, being a continuous function, f', of the quality X and certain parameters. We may represent the probability symbolically by the following equation

$$dy_{\lambda'}=f'(X,\ \lambda_1',\ \lambda_2' \ldots \lambda_{m'}')dX, \tag{1}$$

where the λ''s represent the m' parameters. Experimental evidence abounds in many fields of science to justify the adoption of Eq. 1 to represent the probability distribution of the effects of systems of chance causes. It is quite reasonable, therefore, to adopt this equation as a definition of uniform product and to use it as a basis for detecting lack of control.

Obviously, if we knew f' and the values of the m' parameters in Eq. 1, it would be comparatively easy to determine the limits within which the quality X or any estimate of a parameter derived from a sample of the product might be expected to vary because of chance causes. In practice, however, we know only the n observed values of quality obtained from inspecting a sample of as many units, and we do not know either the true functional relationship f' or any one of the m' parameters even though the product be uniform. We wish to find f' and each of the m' parameters, but, knowing that we cannot do this, we try to find some approximation f for the true function f' and some estimates $\Theta_1, \Theta_2 \ldots \Theta_m$ for the parameters $\lambda_1, \lambda_2 \ldots \lambda_m$ occurring in f. To do this we tentatively assume that the sample of n units has been drawn from a uniform product distributed in accord with the function f, and then use statistical theory to see if our assumption is justified.

Theoretically there are four fundamental steps in the procedure outlined above. They are:

1. *The Problem of Specification:* To find or specify a satisfactory form f of the distribution of the uniform product from which the sample of n pieces is assumed to have been drawn or to find the equation

$$dy_\lambda = f(X, \lambda_1, \lambda_2, \ldots \lambda_m)dX \tag{2}$$

where dy_λ is the assumed probability of a unit having a quality X within the interval X to $X+dX$.

For example we often assume the distribution to be normal so that Eq. 2 becomes

$$dy_\lambda = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-m_1)^2}{2\sigma^2}} dX, \tag{2'}$$

Here $m=2$, and λ_1 and λ_2 are respectively the arithmetic mean m_1 and the root mean square (or standard) deviation σ of X as defined by the normal curve Eq. 2′.

2. *The Problem of Estimation:* To find from the data given by the sample a suitable estimate for each of the m parameters in Eq. 2. These estimates of the parameters in terms of the data of the sample are often termed statistics. If we let Θ_i represent the chosen statistic for the parameter λ_i in Eq. 2, we may rewrite this equation as follows

$$dy_\Theta = f(X, \Theta_1, \Theta_2, \ldots \Theta_m)dX \tag{3}$$

as our theoretical approximation for the assumed true (Eq. 2) probability distribution.

An estimate of a given parameter may often be obtained in a number of ways by one or more methods.

In the above illustrative case of the normal law, we must estimate the two parameters m_1 and σ (Eq. 2′) from the n observed values of X in the sample. Now, it is well known[3] that σ may be expressed in an indefinitely large number of ways in terms of the arithmetic means of the absolute values of the integral powers of the deviations of X defined by Eq. 2′. Estimates of σ might be obtained in terms of the corresponding means calculated from the

[3] Whittaker and Robinson, Calculus of Observation, page 182.

sample. Two such estimates familiar to all are (letting Θ_2 stand in general for an estimate of σ, the second parameter of equation 2′)

$$\Theta_{21} = \sqrt{\frac{\pi}{2}} \Sigma \frac{|X - \overline{X}|}{n},$$

and

$$\Theta_{22} = \sqrt{\Sigma \frac{(X - \overline{X})^2}{n}},$$

where the summation extends over all the X's in the sample of n and $\overline{X}$ is the arithmetic mean of these values of X.

Thus for every λ occurring in Eq. 2, we may have many ways of securing an estimate from the sample. Of these ways, which one shall we choose? Obviously, as in the case of Θ_{21}, compared with Θ_{22}, one estimate may require less labor than another in its calculation. This, however, is not always the deciding factor, because one estimate may have a larger error than another. This leads us to the third problem.

3. *The Problem of Distribution:* To determine how each of the proposed estimates of a parameter might be distributed in a sequence of samples so that we may obtain some measure of its error.

In general we desire that estimate of a given parameter which has the smallest error or highest precision. Thus, in the case of Θ_{21}, it requires a sample of $1.14n$ to give as high a precision as the estimate Θ_{22} has for a sample of size n because the ratio of the error of Θ_{21} to Θ_{22} is $\sqrt{1.14}$. Hence the economic savings effected by using the better of two estimates may be very appreciable.

Furthermore the errors of the statistics are used in establishing the limits within which observed values of the statistics calculated from different samples may be expected to lie as will be illustrated below in discussing the data of Fig. 2. Naturally such errors are used in preparing the control chart Fig. 4.

Suppose now that we have taken the three steps outlined above and found the calculated or theoretical distribution in the form of Eq. 3. What assurance have we that the observed sample could have come from such a distribution? This question leads us to the fourth problem.

4. *The Problem of Fit:* To calculate the probability of fit between the observed and theoretical distributions.

Thus, if the n observed values of X are grouped into $m+1$ cells having frequencies $n_o, n_1, \ldots n_{m+1}$ and if the calculated or theoretical frequencies in these same cells as determined from Eq. 3 are $n_{o\Theta}, n_{1\Theta}, \ldots n_{m\Theta}$ where $\Sigma n_i = \Sigma n_{i\Theta} = n$, we may calculate by Pearson's method the probability P of random samples exhibiting as large or larger values of X'' than that observed in our sample where $\chi^2 = \Sigma \frac{(n_{i\Theta} - n_i)^2}{n_{i\Theta}}$. If the value of probability P thus found is small, we may conclude that it is highly improbable that the sample of n units of product came from uniform product of the form assumed. Of course, this theoretically does not settle the question as to whether the sample might have come from a uniform product other than that assumed, because, as we see, f is only an assumed form for f'. Practically, however, we seem justified in concluding that it is unlikely that the product is uniform if P is small, particularly since the choice of f is customarily made upon the basis of large samples. The application of this test is illustrated in connection with the discussion of the data in Fig. 3.

Practical Application of Theory

The application of the steps just outlined will be illustrated by an analysis of the data in Figs. 1 and 2 to show that the product had not been controlled for the period therein indicated. Carrying out steps 1 and 2 we conclude that the best theoretical equation representing the data in Fig. 1 is either[4] the Gram-Charlier series (two terms) or the Pearson curve of type IV for both of which the estimates of the parameters may be expressed in terms of the first four moments μ_1, μ_2, μ_3 and μ_4 of Fig. 3. These two distributions are shown in columns 10 and 14 respectively.[5] Pearson's test for goodness of fit (step 4) gives negligible results[6] (the probabilities of fit as measured by P on the chart are for practical purposes zero) in both instances, and this was taken as indicating that assignable causes of variation had entered the product. Further investigation of an engineering nature justified this conclusion.

We should not fail to note as suggested above, however, that a small value of fit technically indicates only that the chance is small that a random sample drawn from the theoretical universe (either the two-

[4] Equations for these curves may be found in Bowley's Elements of Statistics, pages 267 and 345 respectively.

[5] Bowley's table, page 303 in his "Elements of Statistics," was used in the calculation of the Gram-Charlier graduation.

[6] Corrections were applied to take account of the number of degrees of freedom, etc., in the calculation of goodness of fit.

OBSERVATIONS

CHARACTERISTIC INSPECTED (a)	QUALITY X
UNITS (b)	
TO NEAREST	
SOURCE OF DATA	

INSPECTION ENGINEERING ANALYSIS SHEET

SUBJECT

	TYPE A-INSTRUMENT
1	Quality from
2	July 1923 -
3	June 1924 incl.

CALC. BY MBC	REPORT NO.
CHK. BY MSH	DATE 1-14-26
APP'D.	
SHEETS	SHEET

$f_{(x)}$ Gram-Charlier (2 terms) — $f_{(x)}$ Pearson Type IV

0 Cell No.	1 Quality (a) in X (b)	2 Cell Boundaries Lower	3 Cell Boundaries Upper	4 Dev. in cells from origin at $\bar{o}$ -5.5 X	5 Obs. Freq. y	6 yX	7 yX^2	8 yX^3	9 yX^4	10 Theor. Freq. y_1	11 $y-y_1$	12 $(y-y_1)^2$	13 $\frac{(y-y_1)^2}{y_1}$	14 Theor. Freq. y_2	15 $y-y_2$	16 $(y-y_2)^2$	17 $\frac{(y-y_2)^2}{y_2}$	Cell No.
0	-5.5	-5.75	-5.26	0	13	0	0	0	0									0
1	-5.0	-5.25	-4.76	1	10	10	10	10	10									1
2	-4.5	-4.75	-4.26	2	8 [74]	16	32	64	128	86	- 12	144	1.674	103	- 29	841	8.165	2
3	-4.0	-4.25	-3.76	3	43	129	387	1161	3483									3
4	-3.5	-3.75	-3.26	4	100	400	1600	6400	25600	253	-153	23409	92.526	229	-129	16641	72.668	4
5	-3.0	-3.25	-2.76	5	815	4075	20375	101875	509375	675	140	19600	29.037	631	184	33856	53.655	5
6	-2.5	-2.75	-2.26	6	1761	10566	63396	380376	2282256	1466	295	87025	59.362	1463	298	88804	60.700	6
7	-2.0	-2.25	-1.76	7	2397	16779	117453	822171	5755197	2662	-265	70225	26.381	2708	-311	96721	35.717	7
8	-1.5	-1.75	-1.26	8	3431	27448	219584	1756672	14053376	3679	-248	61504	16.718	3716	-285	81225	21.858	8
9	-1.0	-1.25	- .76	9	3703	33327	299943	2699487	24295383	3441	262	68644	19.949	3477	226	51076	14.690	9
10	- .5	- .75	- .26	10	2165	21650	216500	2165000	21650000	2024	141	19881	9.823	997	168	28224	14.133	10
11	0	- .25	.24	11	510	5610	61710	678810	7466910	683	-173	29929	43.820	626	-116	13456	21.495	11
12	.5	.25	.74	12	77	924	11088	133056	1596672									12
13	1.0	.75	1.24	13	15 [94]	195	2535	32955	428415	81	13	169	2.086	100	- 6	36	.360	13
14	1.5	1.25	1.74	14	2	28	392	5488	76832									14
15																		15
	Σ				15050	121157	1015005	8783525	78143637	15050			301.376	15050			303.441	
										NO. OF CELLS = 10		P = .00000		NO. OF CELLS = 10		P = .00000		

MOMENTS ABOUT ORIGIN $\bar{o}$

$$_1\mu_1 = \frac{\Sigma yX}{\Sigma y} = \frac{121157}{15050} = 8.050299 \qquad m = \text{UNITS (b) PER CELL} = .5$$

$$_1\mu_2 = \frac{\Sigma yX^2}{\Sigma y} = \frac{1015005}{15050} = 67.442193 \qquad \bar{X} = \bar{o} + m\, {}_1\mu_1 = -5.500000 + 4.025150 = -1.474850$$

$$_1\mu_3 = \frac{\Sigma yX^3}{\Sigma y} = \frac{8783525}{15050} = 583.622924$$

$$_1\mu_4 = \frac{\Sigma yX^4}{\Sigma y} = \frac{78143637}{15050} = 5192.268239$$

NOTE: THE ORIGIN $\bar{o}$ IS THE MID-VALUE BETWEEN THE BOUNDARIES OF THE CELL CHOSEN.

CORRECTED MOMENTS

$$\sigma = m\,\mu_2^{1/2} = .5 \times 1.597356 = .798678$$

$$k = \frac{\mu_3}{\mu_2^{3/2}} = \frac{-1.730023}{4.075727} = -.424470$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{22.144232}{6.510387} = 3.401370$$

$$\beta_1 = k^2 = (-.424470)^2 = .180175$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}} = \frac{.798678}{\sqrt{15050}} = .00651034$$

$$\sigma_\sigma = \frac{\sigma}{\sqrt{2N}} = \frac{.798678}{\sqrt{30100}} = .00460350$$

$$\sigma_k = \sqrt{\frac{6}{N}} = \sqrt{.000398671} = .0199667$$

$$\sigma_{\beta_2} = \sqrt{\frac{24}{N}} = \sqrt{.00159468} = .0399334$$

UNCORRECTED MOMENTS ABOUT ARITH. MEAN

$$\mu_1 = 0$$

$$\mu_2 = {}_1\mu_2 - {}_1\mu_1^2 = 67.442193 - 64.807314 = 2.634879$$

$$\mu_3 = {}_1\mu_3 - 3\,{}_1\mu_1\,{}_1\mu_2 + 2\,{}_1\mu_1^3 = 583.622924 - 1628.789457 + 1043.436510 = -1.730023$$

$$\mu_4 = {}_1\mu_4 - 4\,{}_1\mu_1\,{}_1\mu_3 + 6\,{}_1\mu_1^2\,{}_1\mu_2 - 3\,{}_1\mu_1^4 = 5192.268239 - 18793.356164 + 26224.484274 - 12599.963844 = 23.432505$$

$$3\sigma_{\bar{X}} = .0195310$$

$$3\sigma_\sigma = .0138105$$

$$3\sigma_k = .0599001$$

$$3\sigma_{\beta_2} = .119800$$

CORRECTED MOMENTS ABOUT ARITH. MEAN (SHEPPARD'S CORRECTIONS)

$$\mu_2(\text{cor.}) = \mu_2 - \frac{h^2}{12} = 2.634879 - .08333 = 2.551546$$

$$\mu_4(\text{cor.}) = \mu_4 - \frac{1}{2}h^2\mu_2 + \frac{7}{240}h^4 = 23.432505 - 1.317440 + .029167 = 22.144232$$

$$h = 1$$

Fig. 3

term Gram-Charlier series or Pearson IV type in this case) would give as large or larger value of χ^2 than that observed. Therefore the basis for the conclusion at the end of the previous paragraph is that we have faith[7] that the customary method of taking theoretical steps 1 and 2 gives a close approximation to the true distribution of the product when it is uniform or controlled.

Turning to a study of the data grouped into monthly distributions (Fig. 2), we find additional evidence of lack of control. Naturally the monthly observed values of the four statistics, average $\overline{X}$, standard deviation σ, skewness $k=\sqrt{\beta_1}$, and kurtosis β_2 should lie within well-defined limits established by sampling theory (step 3) and shown in Fig. 4, if the product had been controlled. Furthermore, the observed values of percentage defective p (percentage of instruments having quality less than some value X) from month to month also should fall within well-defined limits. Using the grand average[8] of a statistic as the basis for establishing limits, the first five sections of the control chart in Fig. 4 were constructed. The dotted lines calculated upon the basis of a uniform sample of 1250 indicate the limits within which the different statistics should lie, if the product had been controlled. The chart shows that observed values of these statistics often fall outside their respective limits indicating, subject to limitations imposed by the method of calculation, lack of control of product.

We may go still further and, without carrying out the analysis of Fig. 3, make use of Pearson's test of goodness of fit to calculate the probability that the first two months' samples could have been drawn from the same universe (the same uniform product), then that the third month's sample could have come from the same universe as the combined samples for the first and second months, etc.[9] Obviously the values of χ^2 used as a basis for this calculation of the goodness of fit

[7] Such faith may be based upon the a priori conception that an observed difference in two values of X is the resultant effect of a large number of causes (following in the steps of Laplace, Charlier, Edgeworth, Gram, Thiele and others) and upon the experience that observed homogeneous distributions always have been fitted by some one of the well-known forms of probability curves (following in the steps of Pearson and others).

[8] Some objection may be raised to the use of the observed average as a basis for establishing the limits of a given statistic, because this observed average almost certainly would not be the true value even though the product had been uniform. In the present case, however, we are probably justified in using the observed average because previous experience based upon thousands of observations has given approximately the same values for these quantities. Rigorously, of course, we should find the standard deviations of monthly differences from the grand average and set up limits on this basis. Wherever necessary this method is followed and in fact has been carried out for the case in hand where it gives results similar to those indicated in Fig. 4.

[9] Pearson, K.P., Biometrika, vol. viii, 1911, p. 250 and vol. x, 1914, p. 85.
Rhodes, E.C., Biometrika, vol. xvi, 1924, p. 239.

should fall within well-defined limits such as indicated on the chart. Reference to the χ^2-part of the control chart, Fig. 4, shows that this test gives more conclusive evidence than any other for deciding that the product had not been controlled. As previously noted, further investigation revealed the assignable causes of lack of control. This is a

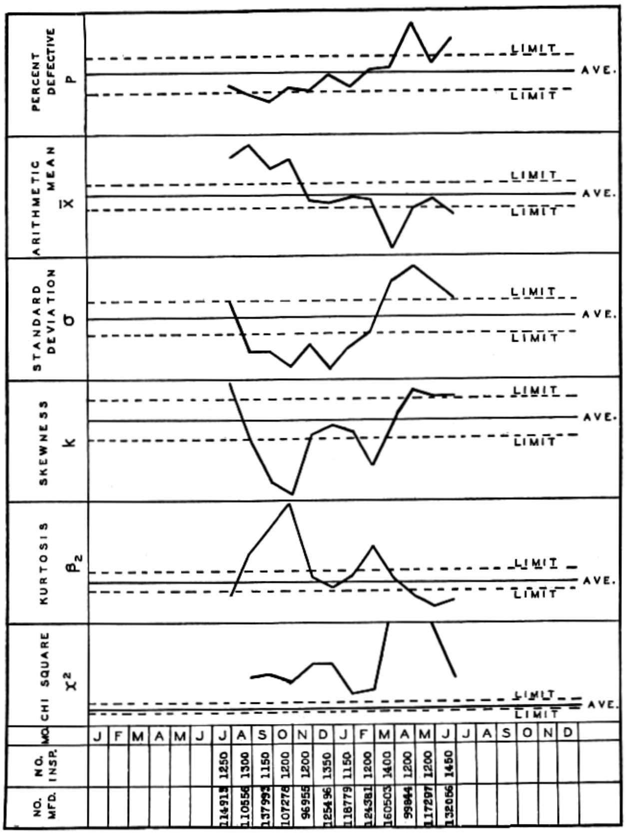

Fig. 4

common experience under such circumstances. Furthermore, it is of interest to note that the preparation of such a chart requires but a small amount of labor on the part of a computer.

Discussion and Conclusion

This paper shows how statistical methods may be used to detect lack of control of product. It describes a recently developed form of manufacturing control chart which helps in the use of inspection and production data by applying some of the modern tools of the statistician. The chart tells the manufacturer at a glance whether or not the product has been controlled. Evidence of lack of control calls for immediate attention, but there need be no time lost in looking for causes of variation in product when these variations are not large enough to indicate lack of control.

There is an obvious advantage in using all parts of the chart wherever possible, because, as the illustration shows, one part may reveal trouble even though some other parts do not. However, when the inspection is made on the basis of attributes, the data will be available for the first or percentage defective part of the chart only.