STATISTICAL QUALITY CONTROL APPLIED TO A TELEMETRY SYSTEM ACCEPTANCE PROCEDURE

N66 34638

FACILITY FORM 602

(ACCESSION NUMBER) N66 34638 (THRU)

(PAGES) 117 (CODE) 1

(NASA CR OR TMX OR AD NUMBER) CR 77305 (CATEGORY) 07

By

MARVIN A. GRIFFIN

RICHARD S. SIMPSON

Project Directors

H. PAUL HASSELL, JR.

Research Associate

WILLIAM S. SPIVEY

Graduate Assistant

June, 1966

GPO PRICE $ ______

CFSTI PRICE(S) $ ______

Hard copy (HC) $4,00

Microfiche (MF) 1,00

ff 653 July 65

TECHNICAL REPORT NUMBER 13

SYSTEMS ENGINEERING GROUP

BUREAU OF ENGINEERING RESEARCH

UNIVERSITY OF ALABAMA UNIVERSITY, ALABAMA

STATISTICAL QUALITY CONTROL

APPLIED TO A TELEMETRY SYSTEM

ACCEPTANCE PROCEDURE

by

Marvin A. Griffin
Richard S. Simpson
Project Directors

H. Paul Hassell, Jr.
Research Associate

William S. Spivey
Graduate Assistant

June, 1966

TECHNICAL REPORT NUMBER 13

Prepared for

National Aeronautics and Space Administration
Marshall Space Flight Center
Huntsville, Alabama

Under

CONTRACT NUMBER NAS8-20172

Systems Engineering Group
Bureau of Engineering Research
University of Alabama

ABSTRACT

This is the thirteenth of a series of technical reports concerned with the Telemetry Systems on the Saturn vehicle.

The purpose of this report is to develop a methodology for implementing statistical control charts as a basis for telemetry package acceptance procedures. The methodology is developed for both the control chart for mean values and for the control chart for standard deviations.

A total expected cost model which relates alpha and beta errors as well as the sample size is developed. This model is used to establish optimum upper and lower control limits for the chart for mean values. Control limits for the chart for standard deviations are then established based on this model.

An experiment that was designed and conducted for the purpose of testing the feasibility of the optimum control limits is reported. Results of this experiment confirm the reasonableness of the assumptions made in the cost model. Subcarrier oscillators in the experimental telemetry package that were intentionally maladjusted are detected by the control charts.

Estimates of the accuracy and precision of the telemetry package are obtained and ninety-nine per cent confidence limits are established for these limits.

Standards for future control chart analysis are established for both charts. These standards may be used for future package checkout procedures.

34638

TABLE OF CONTENTS

	Page
ABSTRACT	ii
LIST OF TABLES	v
LIST OF ILLUSTRATIONS	vi
Chapter	
I. INTRODUCTION	1
Statement of the Problem	
The Proposed Methodology	
II. THE GENERAL THEORY OF CONTROL CHARTS	8
III. THE DEVELOPMENT OF THE METHODOLOGY	14
The $\overline{X}$-Chart	
The σ-Chart	
Estimation of σ'	
Effects of Non-Normality	
Summary of the Methodology	
IV. THE DEVELOPMENT OF THE PROBABILITY FUNCTIONS FOR THE TYPE I AND TYPE II ERRORS	37
Operating Characteristic Function for the $\overline{X}$-Chart	
Operating Characteristic Function for the σ-Chart	
Probability Function for α for the $\overline{X}$-Chart	
Probability Function for α for the σ-Chart	
V. THE DETERMINATION OF THE PROPER LIMIT CONSTANT	47
VI. APPLICATION OF THE METHODOLOGY	55
Description of the Experimental Output	
Determination of the Optimum K and n	
Analysis by Control Charts	
VII. CONCLUSION	79

Chapter | Page

Summary
Sources of Possible Error
Recommendations for Future Research and Application

APPENDICES 85

A. Miscellaneous Proofs, Theorems, and Sample Calculations 86

B. Glossary of Symbols 93

LIST OF REFERENCES 97

BIBLIOGRAPHY 100

LIST OF TABLES

Table | Page

1. Values of $\bar{\bar{X}}$, $\hat{\sigma}_{\bar{x}}$, $\bar{\sigma}$, and σ_{σ} for 0% Input Level 56

2. Means and Standard Deviations for 0% Input Level 57

3. Optimum Values of K 58

4. Values of Total Cost ($TC = c_1\beta + c_2\alpha + c_3 n$) 59

5. Means and Standard Deviations for 25%, 50%, 75%, and 100% Input Levels 60

6. Values of $\hat{\sigma}_i$, $\bar{\bar{X}}_i$, and $\hat{\sigma}_{\bar{\bar{X}}_i}$ and Channels Requiring Investigation for 5 Input Levels 64

LIST OF ILLUSTRATIONS

Figure		Page
1.	Distribution of Chance Variations in a Sample Measure of Quality	9
2.	Illustration of the Theoretical Basis for a Control Chart	10
3.	Control Chart for $\overline{X}$ for Shoulder Depth of Fragmentation Bombs	12
4.	Control Chart for σ for Shoulder Depth of Fragmentation Bombs	12
5.	Block Diagram of the FM/FM Experimental Telemetry System	16
6.	Graph of $f(\overline{X}_{ij})$ Depicting Areas Under the Curve	22
7.	Control Chart for Mean Values of SCO's	24
8.	Control Chart for Standard Deviation of SCO's	29
9.	Possible States of the Process	38
10.	Graphic Representation of a Shift in the Process Mean from $\overline{\overline{X}}'$ to $\overline{\overline{X}}'_{\delta}$	41
11.	Control Charts for σ, 0% Input	63
12.	Control Charts for $\overline{X}$, 0% Input	65
13.	Control Charts for σ, 25% Input	66
14.	Control Charts for $\overline{X}$, 25% Input	67
15.	Control Charts for σ, 50% Input	68
16.	Control Charts for $\overline{X}$, 50% Input	69
17.	Control Charts for σ, 75% Input	70
18.	Control Charts for $\overline{X}$, 75% Input	71

Figure		Page

19. Control Charts for σ, 100% Input 72

20. Control Charts for $\overline{X}$, 100% Input 73

21. OC Curve for $\overline{X}$-Chart Based on Standard Values, 0% Input 78

22. OC Curve for σ-Chart Based on Standard Values 78

CHAPTER I

INTRODUCTION

While the attempt to control the quality of a manufactured product is as old as industry itself, the concept of statistical quality control is relatively new. The greatest development in statistics has occurred in the last sixty years, and it was not until the 1920's that statistical theory began to be applied effectively to quality control (4).[1] In recent years great progress has been made in applying statistical methods to problems of research and development. At the same time the application of statistical quality control to the manufacturing process has become recognized as a major factor in the reduction of the costs associated with improved quality and in the improvement of product quality. The integration of statistical quality control into the area of research and development has also been a tremendous aid in attaining both process and product control.

Statement of the Problem

A particular quality control problem has recently become

---

[1]Numbers is parentheses throughout the thesis indicate the references as listed in the LIST OF REFERENCES.

evident in the field of aerospace telemetry. Aerospace telemetry is the "science of transmission of information from air and space vehicles to accessible locations" (15).

The advent of the missile age has brought about a phenomenal increase in the usage of telemetry equipment. With the evolution of new techniques and equipment for telemetering in-flight space vehicle data, the need for increased accuracy and precision of the transmitting equipment is obvious. An airborne telemetry package[2] is placed on the spacecraft for the purpose of transmitting the most critical measurements to a telemetry ground station. At the ground station personnel continuously monitor and analyze these measurement data to determine the effect of flight conditions at the vehicle. Therefore it is of prime necessity for the telemetry package to be of sufficient quality to assure that the transmitted data is actually that measured at the vehicle.

During the time required for each telemetry package to be sent from the manufacturer to the space vehicle, there are several places where the control of the quality of the package needs to be established. The first of these is at the manufacturing plant immediately before shipping the package to the telemetry personnel. Another is at the test laboratory immediately after the telemetry

---

[2] A telemetry package is an electrical system consisting of a set of subcarrier oscillators for converting measured voltage into frequency, a mixer amplifier, a transmitter, and a power amplifier used for transmitting signals from space vehicle to a ground receiving station.

personnel have received the equipment. A third place for a quality control program is at the test laboratory immediately before sending the package to the vehicle. A fourth area for controlling the quality of the equipment is at the vehicle prior to launch time. A final area at which the control of the telemetry package performance is a necessity is in the spacecraft during flight. The monitoring of actual flight calibration data and subsequent analysis by statistical methods would indicate whether or not the package was performing adequately during this critical phase.

It is believed that the methodology presented in this thesis could be applied to any of these areas. However, the research will be conducted at, and applied to, the third of these areas; the NASA telemetry test laboratory immediately prior to sending the package to the vehicle.

The present program for determining if a telemetry package is operating in a satisfactory manner and is ready to be sent to the spacecraft consists of a series of rigorous electrical tests. After these tests have been conducted and adjustments made on the components, a five point calibration sequence[3] is fed through the package and transmitted over a cable to the ground station. Here it is received and sent through a bank of discriminators[4] and then

---

[3]A five point calibration sequence consists of supplying voltage in five distinct steps to a telemetry package. The steps represent 0, 25, 50, 75, and 100% of 5 volts.

[4]A discriminator is an instrument for separating a mixed frequency signal into various frequency bands corresponding to those produced by a related subcarrier oscillator.

recorded by an oscillograph.[5] This record made by the oscillograph is at present the only means for analyzing the quality of the assembled package. It is thought that this method furnishes inconclusive answers to the questions "what is the accuracy and precision of the package?" and "are there any assignable causes of variation within the package?" Accuracy is a measure of systematic errors while precision is a measure of random (chance) errors (5,6).

The researcher believes that by establishing an effective program of quality control utilizing statistical methods, these questions can be answered and ultimately a decision to reject or accept the package as satisfactory can be made. Also, through the establishment of a quality control program a quantitative history of the performance of telemetry packages can be assembled.

The Proposed Methodology

The methodology employed in the establishment of this quality control program will assure that the basic components of the telemetry package, the subcarrier oscillators, are in statistical control. The methodology is based on the statistical control of product quality by Shewhart control charts. These control charts will provide a graphical method for comparing with

---

[5]An oscillograph is a device for producing a written curve representing variable voltage.

an average value the output of the different subcarrier oscillators over several levels of input voltage. Therefore subcarrier oscillators (henceforth referred to as SCO's) which differ significantly from the overall mean value may be investigated for assignable causes of variation and subsequently replaced if abnormalities exist.

There are two distinct but related phases of control chart analysis (10). In the first phase, which is often termed "Control with No Standard Given," the control chart is used as a device for specifying a state of statistical control and judging whether the state of control has been attained based on past data. The purpose in this phase is thus to discover whether measurements from samples vary among themselves by an amount greater than should be attributed to chance. The second phase, usually termed "Control with Respect to a Given Standard," is used to discover whether measurements obtained in current production depart significantly from "standard values" which may have been established by experience based on prior data, by economic considerations, or by reference to the desired state of the process as designated by the specifications. The methodology developed in this thesis will be applied to the first of these phases. The result of the analysis will therefore be to judge whether a state of control has been attained for the SCO's and then to specify a goal of statistical control for future action.

Following a brief discussion of the general theory of quality control in CHAPTER II, a theoretical development of the methodology consists of two types of control charts. The control chart for standard deviations (σ - chart) is used for analyzing the variability of the SCO's. The control chart for mean values ($\overline{X}$ - chart) supplies a basis for judging whether the various $\overline{X}$ values, the mean values for each SCO-input level grouping, are in statistical control. Therefore SCO's which are not functioning properly with respect to either their mean values or their variability can be quickly detected and investigated. After these assignable causes have been removed, the performance of the SCO's can be predicted and standards set for their variability and mean values. These standards can then be used for the evaluation of future telemetry package performance.

A second major area of investigation is the development of the operating characteristic (OC) function for the chart for means and the chart for standard deviations for analysis based on past data. The OC functions are developed in CHAPTER IV, and a general procedure is given for obtaining the OC curve for both charts.

A third major problem to be considered in this thesis is the determination of the proper limit constant, K, to be used in calculating the control chart limits. The selection of the proper K factor requires considerable analysis, and the decision must be based on an economic evaluation of the risks involved in making incorrect decisions. This problem is solved in CHAPTER V.

If the thesis is to have any practical significance, the methodology must be applied to an actual situation and meaningful results obtained. Therefore, a final major thesis contribution is an investigation of the methodology applied to a real world environment in CHAPTER VI. An experiment was performed in the telemetry ground station at the Marshall Space Flight Center in Huntsville, Alabama. The application of the methodology is tested by observing the ability of the control charts to detect certain malfunctioning components in a telemetry package set-up. Finally when the control charts indicate that the experimental package is in statistical control, the precision and accuracy of this package is estimated.

CHAPTER II

THE GENERAL THEORY OF CONTROL CHARTS

"A control chart is a statistical device principally used for the study and control of repetitive processes" (4). The discovery and development of control charts were made in 1924 and the following years by a young physicist of the Bell Telephone Laboratories, Walter A. Shewhart. In trying to solve a problem which was complicated by the presence of random variation, he decided that the problem was statistical in nature. Some of the observed variation in performance was inherent in the process and could be explained as being the result of chance causes. This type variation was unavoidable. However, from time to time variations occurred which could not be explained by chance alone, but were the result of some change in the process. The differentiation of these two causes of quality variation is the basis for the theory of control charts.

If a group of data is studied and it is found that their variation conforms to a statistical pattern that might reasonably have been produced by chance causes, then it is assumed that there has been no change in the process; i.e., there are no assignable causes of variation present, and the process is said to be in "statistical control." If, however, the variations in the data do not conform to a pattern that might be expected by chance causes, then it is concluded that one or more assignable causes are at work, and the process is said to be "out of control." The nature of this statistical pattern of variation is described as follows.

Suppose samples of a given size are taken at regular intervals and suppose that for each sample some statistic X (sample mean, sample standard deviation, etc.) is computed. Since this statistic X is a sample result it will be subject to sampling fluctuations. If there are no assignable causes of variation present, these sampling fluctuations will take the form of some definite statistical distribution. Suppose, for example, theory suggests that the sampling distribution of X is normal in form, as in Figure 1. This distribution will have a mean which can be computed from the sample means, and a standard deviation which can be computed from the

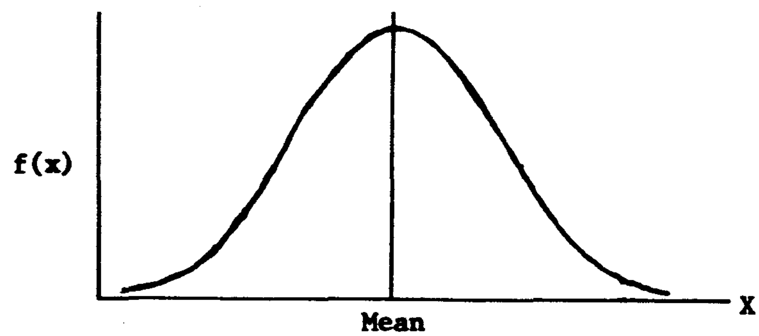

FIGURE 1. DISTRIBUTION OF CHANCE VARIATIONS IN A SAMPLE MEASURE OF QUALITY.

within-sample variation for the various samples. From this mean and standard deviation certain probability points can be calculated. If the vertical scale of a chart is calibrated in units of X and the horizontal scale marked with respect to some rational basis for ordering X and if horizontal lines are drawn through the mean of X, and through an extreme value representing a certain probability point on the upper and lower tail of the distribution of X, the result is a control chart for X, as in Figure 2.

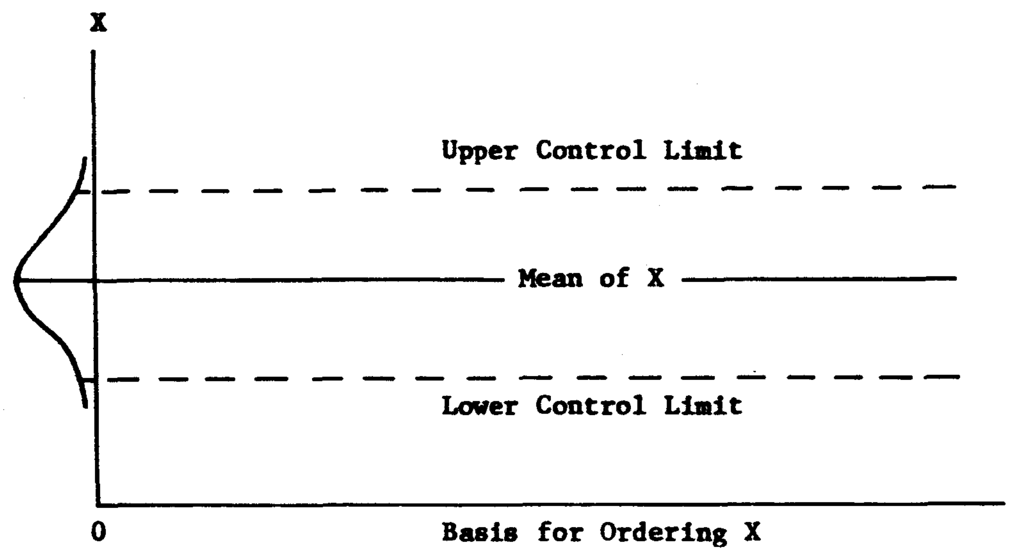

FIGURE 2. ILLUSTRATION OF THE THEORETICAL BASIS FOR A CONTROL CHART.

Since the control chart is constructed in conformity with statistical theory it can consequently be used to test the hypothesis of control. To see whether a process is in control sample values of X from data pertaining to the process are plotted on the control chart. If these values all fall within the control limits without varying in a nonrandom manner within the limits, then the process can be said to be in control at the level indicated by the chart. If the data do not conform to this pattern then departures from the pattern are investigated and assignable causes are tracked down. After a condition of control has been satisfactorily established, departure from the condition may be quickly detected by maintaining a control chart on current output (3).

Figures 3 and 4 illustrate the use of two type control charts. Figure 3 gives a chart for the averages of samples of shoulder depth measurement of fragmentation bombs. The vertical scale is calibrated in units of the sample means, $\overline{X}$, and the horizontal scale is marked for 10 samples taken from two days of production. As can be seen, this chart shows that the sample averages are not in a state of control since the means for samples 3 and 8 on May 5 are above the upper control limit. The process should be investigated to determine why these two sample values are significantly higher than the others. Figure 4 gives a chart for controlling the variability in the measurement of shoulder depth of fragmentation bombs. The vertical scale for this chart is calibrated in units of the sample standard deviation, σ, and the horizontal scale is marked as in Figure 3. This chart shows that the product variability is in a state of statistical control since all points fall within the control limits and the data varies in a random manner within these limits.

It is important to note that the samples on a control chart should represent subgroups of output that are as homogeneous as possible. In other words, the subgroups should be such that if

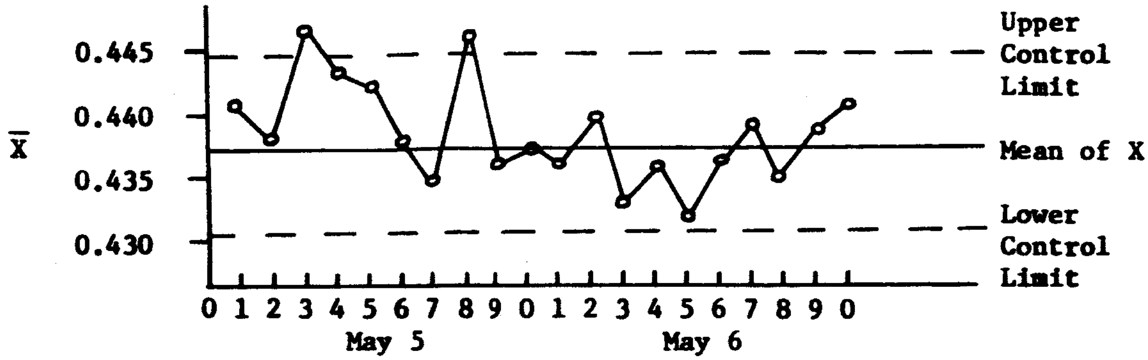

FIGURE 3. CONTROL CHART FOR $\bar{X}$ FOR SHOULDER DEPTH OF FRAGMENTATION BOMBS (1).

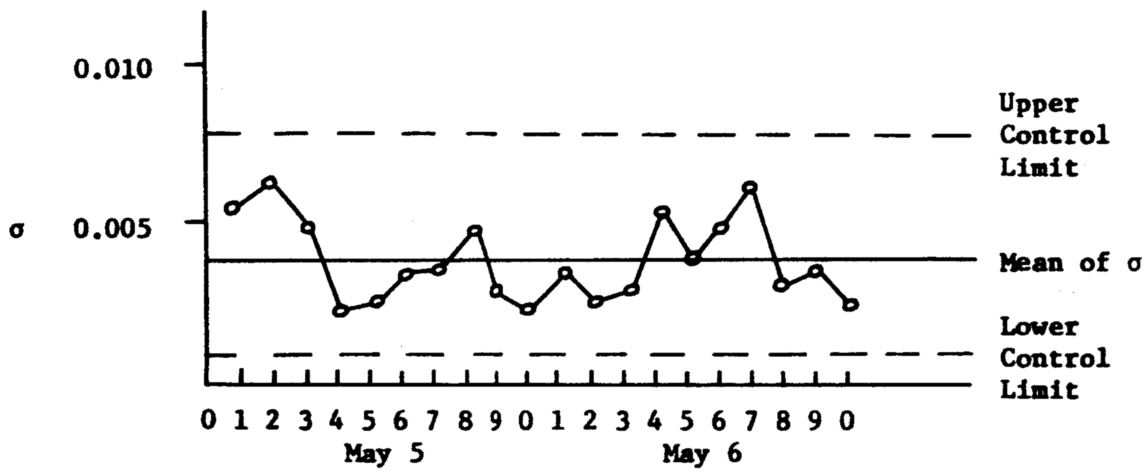

FIGURE 4. CONTROL CHART FOR σ FOR SHOULDER DEPTH OF FRAGMENTATION BOMBS (1).

assignable causes are present, they will show up in differences between the subgroups rather than in differences between the members of a subgroup (13). A natural subgroup, for example, would be the output of a particular subcarrier oscillator. If a system consisted of say 10 subcarrier oscillators, it would be better to take a separate sample from the output of each SCO than to have each sample made up of items from all 10 SCO's. For differences between the subcarrier oscillators may be an assignable cause that is the object of the control chart analysis to detect.

A control chart, then, provides a reasonable test for determining when a process can be considered to be in control. Some of the advantages which may accrue when a process is brought into good control by control chart analysis are (8):

1. The act of getting a process into good statistical control ordinarily involves the identification and removal of undesirable assignable causes. Hence, quality performance has been much improved.

2. A process in good statistical control is predictable.

3. If our process is in good statistical control, we can more safely guarantee our product.

4. A chart in control in experimentation enables us to determine soundly the experimental error.

5. The sound way to cut inspection is through getting the process in control.

CHAPTER III

THE DEVELOPMENT OF THE METHODOLOGY

It has been stated that the major problem to be investigated in this thesis is the use of control charts for determining whether a telemetry package is in a state of statistical control. In the two previous chapters the problem was discussed in general. The purpose in this chapter is to develop a theoretical foundation for the control chart models.

Let us consider the telemetry system as a type of industrial process. The basic component of the telemetry package, the SCO, may also be thought of as a sub-process. An analogous situation would be a large factory within which there are several manufacturing divisions. We might be interested in comparing these manufacturing divisions to determine whether they are producing essentially the same output. Control chart analysis would certainly not seem applicable if we were only interested in comparing the output of one entire factory with perhaps three other such factories. However, the individual divisions within each factory would produce a repetitive output which could be analyzed by control chart techniques. Therefore, we shall view each SCO as a manufacturing division and as a distinct process, capable of producing output which is of a repetitive nature. The input to the process is variable voltage capable of being applied in discrete steps of 0, 1.25, 2.50, 3.75, and 5.00 volts to the telemetry package.[6] The package converts these voltages to frequencies (each SCO represents a different frequency range) and transmits these frequencies to a receiving station. Here a set of discriminators decodes these mixed frequency signals and separates them into frequency ranges corresponding to a particular SCO. These frequencies are then sent to data reduction equipment and converted into digital units. These units may then be printed with each column of the printout corresponding to a particular SCO. This system is represented by Figure 5. Now consider these printed values as a sequence of random variables from the process. A particular random variable can be represented by the symbol X_{ijk} where i corresponds to the different levels of input voltage ($i = 1, 2, \ldots h$), j corresponds to the different SCO's ($j = 1, 2, \ldots m$), and k corresponds to an individual value in a particular SCO-level combination ($k = 1, 2, \ldots n$). Any random variable X_{ijk} can therefore be thought of as being the result of a population average value, where the population is composed of all possible values for a particular SCO at a particular input level, plus some random error. Let us assume that this random error is drawn from a

[6]These five voltage levels represent a synthetic calibration of the telemetry package. Actual in-flight calibration is performed somewhat differently as the output of the mixed signal from the package is calibrated rather than each SCO individually.

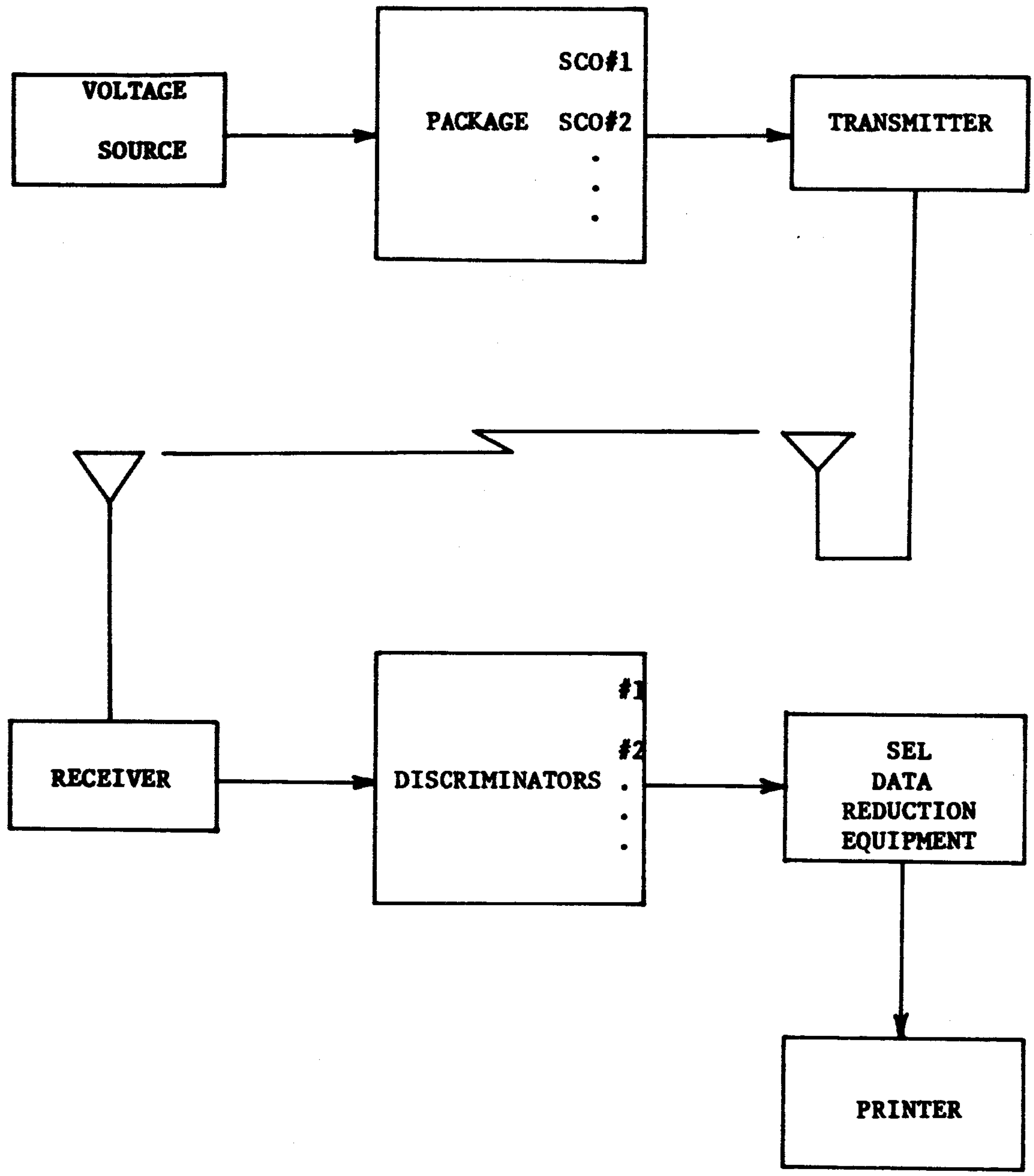

FIGURE 5. BLOCK DIAGRAM OF THE FM/FM EXPERIMENTAL TELEMETRY SYSTEM.

universe of errors that are statistically independent (the value of any one error does not depend on the value of any other errors) and are distributed in a normal distribution with a mean value of zero and with some amount of variation. We shall further assume that the several possible population average values that we are considering are drawn from some large universe of average values that are normally distributed. The mean value of this universe is the mean for all possible SCO's at a particular level of input, and there may be some amount of variability of the population averages about this mean. Let us also assume that the population average values are statistically independent of the random errors and that the universe mean value and variation are fixed quantities, but are unknown to us and must be estimated from available data. We may now express the random variable X_{ijk} by a linear mathematical model and list several statements clarifying the terms in the model.

$$X_{ijk} = \bar{X}'_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \ldots h; \; j = 1, 2, \ldots m; \; k = 1. 2, \ldots n,$$

where,

1. the ε_{ijk} are statistically independent and distributed according to $N(0, \sigma_i'^2)$;[7]
2. the $\bar{X}'_{ij}$ are statistically independent and distributed according to $N(\bar{\bar{X}}'_i, \sigma_{\bar{x}}'^2)$, where $\sigma_{\bar{x}}'^2 = \theta^2\sigma_i'^2$;
3. the $\bar{X}'_{ij}$ are statistically independent of the ε_{ijk};
4. the $\bar{\bar{X}}'_i$ and $\sigma_i'^2$ are fixed but unknown.

[7]This notation indicates that the errors are distributed according to a normal probability distribution (see p. 21) with mean of zero and variance $\sigma_i'^2$.

With this model in mind we shall now turn to the problem of developing a set of decision rules for determining whether a group of SCO's are in statistical control with respect to the average values of the random variables associated with the different SCO's. The testing of the hypothesis of statistical control in this manner is termed "control through the use of the $\overline{X}$-chart."

The $\overline{X}$-Chart

We shall first formulate the hypothesis of control of the mean values and the alternative hypothesis and then develop the $\overline{X}$-Chart test for these hypotheses.

If the SCO's are in a state of statistical control one consequence is that the population average values for the various SCO's are equivalent. In other words, if we apply a certain voltage to the telemetry package, this same voltage will be applied to all the SCO's simultaneously. Therefore, if we digitize the output of these SCO's for a very long period of time, and if the SCO's are known to be properly adjusted and functioning correctly, the average of all possible digitized values for any one SCO will be exactly the same as for all the other SCO's. Thus $\overline{X}'_{i1} = \overline{X}'_{i2} = \ldots = \overline{X}'_{im}$. This is the hypothesis of control and is often termed the null hypothesis, symbolized by H_0. On the other hand, if for some reason any of the SCO's have not been properly adjusted or are not functioning correctly, then all of the $\overline{X}'_{ij}$ values will not be equal. Thus $\overline{X}'_{i1} \neq \overline{X}'_{i2} \neq \ldots \neq \overline{X}'_{im}$. This is called the alternative hypothesis and is symbolized by H_1. Let us examine these hypotheses from a slightly different viewpoint. Statement 2. listed on page 17 states that the $\overline{X}'_{ij}$ are normally distributed with mean $\overline{\overline{X}}'_i$ and variance $\theta^2\sigma_i'^2$. Now if a state of control exists, these $\overline{X}'_{ij}$ will be equal and there will be, of course, no variation among them. Therefore θ will be zero. However, if any of the SCO's are not functioning correctly, then these $\overline{X}'_{ij}$ will be unequal and there will be some variation among them. Thus θ will be greater than zero. Our two hypotheses may now be stated as:

$$H_0: \quad \theta = 0;$$

$$H_1: \quad \theta > 0.$$

By treating the $\overline{X}'_{ij}$ as random variables we are taking into account the average effect of m independent assignable causes of varying magnitude. Although we are unable to specify the size of any particular shift in the process averages, $\overline{X}'_{ij}$, a measure of the size of the $\overline{X}'_{ij}$ as a group is given by the parameter θ.

Thus a test for the hypothesis of control, H_0: $\theta = 0$, would be to compute each population average $\overline{X}'_{ij}$ and compare these values with the universe mean $\overline{\overline{X}}'_i$. If each $\overline{X}'_{ij}$ is exactly equal to $\overline{\overline{X}}'_i$, then $\theta = 0$ and the process is in control. However, this method is impossible since each $\overline{X}'_{ij}$ is an average of <u>all possible</u> digitized values for a particular SCO, and the universe mean $\overline{\overline{X}}'_i$ is an average value for <u>all possible</u> digitized values for all of the SCO's at a given voltage level. We would have to let the process run for an infinite period of time to assemble these values. Therefore, we must find some method for testing H_0 that will be based on sample estimates of these population and universe parameters.

Let us take a sample of n values of X_{ijk} for a particular SCO. The mean value of this sample is given by

$$\overline{X}_{ij} = \frac{\sum_{k=1}^{n} X_{ijk}}{n} \quad . \qquad [1]$$

We must show that the expected value of this sample mean for an infinite number of samples of size n is $\overline{X}'_{ij}$. In other words, we must show that

$$E(\overline{X}_{ij}) = \overline{X}'_{ij} \quad .$$

This can be done by using elementary theorems of expectation (see Apprendix A, p. 86) as follows:

$$\begin{aligned} E(\overline{X}_{ij}) &= E\left[\frac{1}{n}\sum_{k=1}^{n} X_{ijk}\right] \\ &= \frac{1}{n} E\left[\sum_{k=1}^{n} X_{ijk}\right] \\ &= \frac{1}{n}\sum_{k=1}^{n} E\,(X_{ijk}) \\ &= \frac{1}{n}\sum_{k=1}^{n} \overline{X}'_{ij} \\ &= \overline{X}'_{ij} \quad . \end{aligned}$$

It can also be shown in the same manner that the expected value of $\overline{X}'_{ij}$ is $\overline{\overline{X}}'_{i}$. Thus,

$$\overline{\overline{X}}_{i} = \frac{\sum_{j=1}^{m} \overline{X}_{ij}}{m} \quad . \qquad [2]$$

Now since these $\overline{X}_{ij}$ values are sample rather than population means there will be some amount of variation among them that can be attributed to chance causes. A measure of this amount of variation is given by what is termed the standard error of the mean, symbolized by $\sigma_{\overline{x}_{ij}}$. The standard error of the mean is merely a standard deviation computed from the various $\overline{X}_{ij}$ about $\overline{\overline{X}}_i$. This term is defined by the formula

$$\sigma_{\overline{x}_{ij}} = \sqrt{\frac{\sum_{j=1}^{m} (\overline{X}_{ij} - \overline{\overline{X}}_i)^2}{m}} .$$

The $\overline{X}_{ij}$ values will therefore form some statistical distribution, of which the mean is $\overline{\overline{X}}_i$ and the variance is $\sigma^2_{\overline{x}_{ij}}$. This distribution of mean values will approach a normal (Gaussian) distribution regardless of the distribution of individual X_{ijk} values. The frequency function of this distribution is given by

$$f(\overline{X}) = \frac{1}{\sigma_{\overline{x}} \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\overline{X} - \overline{\overline{X}}}{\sigma_{\overline{x}}}\right)^2} \qquad [3]$$

This fact is attributed to the central limit theorem of statistics which states, "The form of the distribution of sample means approaches the form of a normal probability distribution as the size of the sample is increased" (9).

Thus, the probability for any $\overline{X}_{ij}$ to lie between any two standard values $\overline{\overline{X}}_i + K\sigma_{\overline{x}_{ij}}$ and $\overline{\overline{X}}_i - K\sigma_{\overline{x}_{ij}}$ can be found by integrating equation [3] for these limits. Suppose, for example, that $K = 3$. The probability that $\bar{X}_{ij}$ will lie between $\bar{\bar{X}}_i \pm 3\sigma_{\bar{X}_{ij}}$ is given by

$$\int_{\bar{\bar{X}}-3\sigma_{\bar{X}}}^{\bar{\bar{X}}+3\sigma_{\bar{X}}} f(\bar{X}) = \int_{\bar{\bar{X}}-3\sigma_{\bar{X}}}^{\bar{\bar{X}}+3\sigma_{\bar{X}}} \frac{1}{\sigma_{\bar{X}}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\bar{X}-\bar{\bar{X}}}{\sigma_{\bar{X}}}\right)^2} d\bar{X}$$

The value of this integral can be shown to be 0.9973 (see Appendix A, p.87). Figure 6 shows a graph of $f(\bar{X}_{ij})$ with probabilities (areas under the curve) given for two value of K.

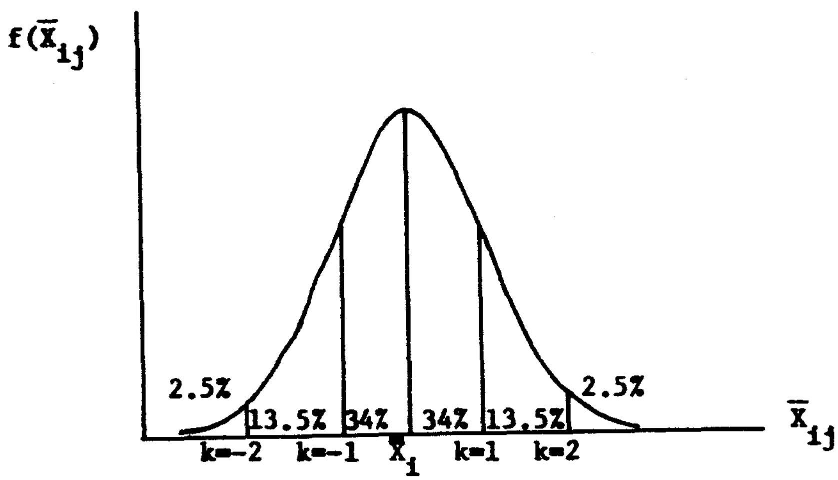

FIGURE 6. GRAPH OF $f(\bar{X}_{ij})$ DEPICTING AREAS UNDER THE CURVE.

Since the total area under the frequency function curve is one, that is,

$$\int_{-\infty}^{\infty} f(\bar{X}_{ij})d\bar{X}_{ij} = 1,$$

the area outside limits of $\bar{\bar{X}}_i \pm 3\sigma_{\bar{x}_{ij}}$ is 1 - 0.9973 = 0.0027. Therefore, the probability that $\bar{X}_{ij}$ will fall outside these limits due to chance alone is 0.0027. This probability could easily be computed for other values of K.

If was previously indicated that the expected value of $\bar{X}_{ij}$ is $\bar{\bar{X}}'$. However, the expected value of $\sigma_{\bar{x}_{ij}}$ is not $\sigma'_{\bar{x}_{ij}}$. Sample variances used to estimate population variances tend to be consistently small by a factor of (n-1)/n (see p.27). Therefore, to correct for this "bias," $\sigma_{\bar{x}_{ij}}$ must be multiplied by $\sqrt{m/(m-1)}$ to give an unbiased estimate of $\sigma'_{\bar{x}_{ij}}$. In computational form this unbiased estimate is given by

$$\hat{\sigma}_{\bar{x}_{ij}} = \sigma_{\bar{x}_{ij}} \sqrt{m/(m-1)}$$

$$= \sqrt{\frac{\sum_{j=1}^{m} (\bar{X}_{ij} - \bar{\bar{X}}_i)^2}{m-1}} \quad . \qquad [4]$$

Thus, to test the hypothesis of statistical control, H_0: $\theta = 0$, we will set the $\bar{\bar{X}}_i \pm K\hat{\sigma}_{\bar{x}_{ij}}$ limits for a particular value of K and then observe whether the computed $\bar{X}_{ij}$ values fall within these limits. If all $\bar{X}_{ij}$ are within the limits, we will assume that the SCO's are in statistical control. However, if any $\bar{X}_{ij}$ falls outside the limits, we will assume that an assignable cause of variation has occurred and this particular SCO must be investigated.

We must realize that there is a certain risk involved in making an incorrect decision. A point might fall outside the limits due merely to chance. In this case we would be investigating an SCO for no reason. This type of decision error is called a Type I error and the associated risk; i.e., the probability of committing this error, is termed the α risk. On the other hand, we might conclude that an SCO which has an assignable cause of variation would be in control due to its $\bar{X}_{ij}$ falling within the limits. This type of decision error is called a Type II error and the associated risk is termed the β risk. An analysis of these risks will be given in subsequent chapters. It will suffice at this point to realize that they exist.

We now have enough information to justify the establishment of the $\bar{X}$-Chart. This chart will be a graphical method for interpreting the decision rules we have developed.

The chart is established by drawing a center line at $\bar{\bar{X}}_i$ and placing control limits at $\bar{\bar{X}}_i \pm K\hat{\sigma}_{\bar{x}_{ij}}$. Values of $\bar{X}_{ij}$ are then plotted on the chart for the m SCO's and the hypothesis of control tested by observing whether these plotted $\bar{X}_{ij}$ fall within the limits. An example of this chart is shown in Figure 7.

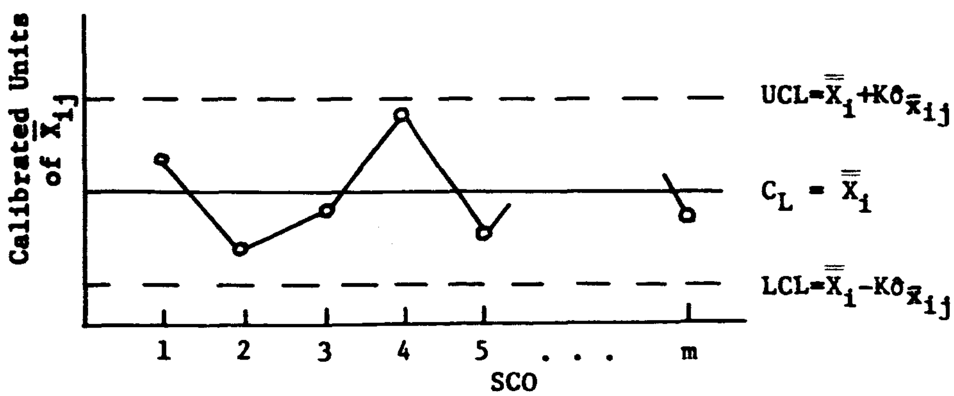

FIGURE 7. CONTROL CHART FOR MEAN VALUES OF SCO'S.

We shall now turn to the problem of developing a control chart for testing the variability within the different SCO's.

The σ-Chart

In this section the problem that we are involved with is to develop a set of decision rules for determining whether a group of SCO's are in statistical control with respect to a measure of variability of the random variables associated with the different SCO's. One such measure of the variability is the standard deviation, σ'_{ij}.

We shall once again formulate the hypothesis of control of the standard deviations and the alternative hypothesis and then develop the σ-Chart to test these hypotheses.

First, let us make the assumption that the various X_{ijk} values are normally distributed. This is perhaps a rather restrictive assumption. However, theoretically we know nothing about the frequency function of the standard deviation of samples from a non-normal universe. Therefore, in order to develop the methodology we will assume normality and then later analyze the effects of non-normality.

This hypothesis of control for variability is that the population standard deviations for the SCO's are equal. Thus, $\sigma'_{i1} = \sigma'_{i2} = \dots = \sigma'_{im}$. Now suppose that these σ'_{ij} values come from a universe of standard deviations whose average value is σ'_i

and whose variance is $\sigma_{\sigma}'^2 = \theta^2\sigma_i'^2$. If a state of control exists then θ will be zero. If the process is not in control, then the population standard deviations will be unequal and θ will be greater than zero. Once again the two hypotheses may be stated as:

$$H_0: \theta = 0;$$

$$H_1: \theta > 0.$$

As with the testing of the hypothesis for the mean values, we know that it is impossible to calculate a true population standard deviation since this would indicate a deviation of <u>all possible</u> X_{ijk} values about their mean. Therefore, we must once again rely on sample estimates for our test.

Suppose that in addition to calculating sample averages, $\bar{X}_{ij}$, we calculate sample standard deviations from the formula

$$\sigma_{ij} = \sqrt{\frac{\sum_{k=1}^{n} (X_{ijk} - \bar{X}_{ij})^2}{n}} \qquad [5]$$

The frequency function of the distribution of variances, σ_{ij}^2, for samples of size n from a normal universe and its differential is given by, momentarily dropping the ij subscript, (1)

$$f(\sigma^2)d(\sigma^2) = \left(\frac{n}{2\sigma'^2}\right)^{(n-1)/2} e^{-(n\sigma^2)/2} \quad (\sigma^2)^{(n-3)/2} \quad \frac{1}{\Gamma((n-1)/2)} d(\sigma^2), \quad [6]$$

where $\Gamma((n-1)/2)$ is the gamma or factorial function (see Appendix A, p. 88). Now changing the variable to σ_{ij} in equation [6] gives

$$f(\sigma)d\sigma = \left(\frac{n}{2\sigma'^2}\right)^{(n-1)/2} e^{-(n\sigma^2)/2\sigma'^2} (\sigma)^{n-3} \frac{2\sigma d\sigma}{\Gamma((n-1)/2)}$$

$$= \frac{n^{(n-1)/2}}{2^{(n-3)/2}\,\sigma'^{(n-1)}} \; \frac{e^{-n\sigma^2/2\sigma'^2}\,\sigma^{n-2}\,d\sigma}{\Gamma((n-1)/2)} \; . \qquad [7]$$

The expected value of σ_{ij} in an infinite number of samples of size n can be evaluated by

$$E(\sigma) = \int_0^\infty \sigma\, f(\sigma)\, d\sigma.$$

Substituting $v = n\sigma^2/2\sigma'^2$, $\sigma = \sqrt{2\sigma'^2 v/n}$, and $2\sigma d\sigma = 2\sigma'^2/n\, dv$ in the middle of expression [7] and in the integral, we obtain

$$E(\sigma) = \int_0^\infty \left(\frac{2\sigma'^2 v}{n}\right)^{(n-2)/2} \left(\frac{n}{2\sigma'^2}\right)^{(n-1)/2} \frac{e^{-v}\, 2\sigma'^2 dv}{\Gamma((n-1)/2)n}$$

$$= \left(\frac{2\sigma'^2}{n}\right)^{1/2} \frac{1}{\Gamma((n-1)/2)} \int_0^\infty v^{(n/2)-1} e^{-v}\, dv$$

$$= \left(\frac{2\sigma'^2}{n}\right)^{1/2} \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}$$

$$= \sqrt{\frac{2}{n}} \cdot \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}\, \sigma' \qquad , \text{ or}$$

$$E(\sigma_{ij}) = C_2 \sigma'_{ij} \; ,$$

where

$$C_2 = \sqrt{\frac{2}{n}} \frac{\Gamma(n/2)}{\Gamma((n-1)/2)} \; . \qquad [8]$$

In a similar manner it may be shown that $E(\sigma^2) = ((n-1)/n)\sigma'^2$.

We may calculate $\bar{\sigma}_i$ by averaging the various σ_{ij}, thus

$$\bar{\sigma}_i = \frac{\sum_{j=1}^{m} \sigma_{ij}}{m} . \qquad [9]$$

Now since we have shown that $E(\bar{X}'_{ij}) = \bar{\bar{X}}'$,

then by a similar argument $E(\bar{\sigma}_i) = C_2\sigma'_i$.

As was stated previously, if a state of control exists, then

$C_2\sigma'_i = C_2\sigma'_{ij}$, or $\sigma'_i = \sigma'_{ij}$. Since the σ_{ij} are sample rather than population values there will again be some variation among them that can be attributed to chance. A measure of this variation is given by the standard deviation of the distribution of sample standard deviations, $\sigma_{\sigma_{ij}}$, defined by

$$\sigma_{\sigma_{ij}} = \sqrt{\frac{\sum_{j=1}^{m} (\sigma_{ij} - \bar{\sigma}_i)^2}{m}} . \qquad [10]$$

Thus, as in the case for testing mean values, we can test the hypothesis of control of variability by setting limits of $\bar{\sigma}_i \pm K\sigma_{\sigma_{ij}}$ for a particular value of K and then observe whether computed values of σ_{ij} fall within these limits. If a point falls outside these limits we will assume that an assignable cause of variation has occurred and this particular SCO must be investigated.

The probability of a point falling outside limits of $\bar{\sigma}_i \pm K\sigma_{\sigma_{ij}}$ can be found by integrating [7] for these limits. The evaluation of this integral is rather complicated, and depends on the sample size n and the universe standard deviation, σ'_i . Previous work (1, 2, 13) has shown that even for samples of n = 5 the form of the distribution of $f(\sigma_{ij})$ roughly resembles the normal distribution. As n increases this resemblance becomes greater. Therefore, although the probability of a sample σ_{ij} falling outside the control limits is not exactly the same as for a sample $\overline{X}_{ij}$, let us assume that these probabilities are roughly equal. Thus, the same K factor can be used for determining the limits for both charts.

We can now justify the establishment of the σ-Chart. The chart is established by drawing a center line at $\bar{\sigma}_i$ and placing control limits at $\bar{\sigma}_i \pm K\sigma_{\sigma_{ij}}$. Values of σ_{ij} are then plotted on the chart for the m SCO's, and the hypothesis of control tested by observing whether these plotted σ_{ij} fall within the limits. An example of this chart is shown in Figure 8.

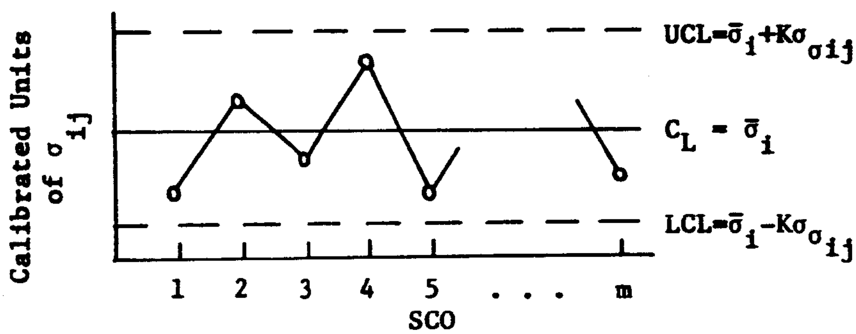

FIGURE 8. CONTROL CHART FOR STANDARD DEVIATION OF SCO'S.

Since σ_{ij} can never be negative (see equation [5]) the smallest possible value for the lower control limit for the σ-Chart is zero. Therefore, if $\bar{\sigma}_{ij} - K\sigma_{\sigma_{ij}}$ yields a negative number the lower limit will be set at zero.

Estimation of σ'

Although the methodology for the $\overline{X}$-Chart was developed before the σ-Chart, in actual application of the control charts the σ-Chart should be established first. The reason for this is that when the process variability is in statistical control the process standard deviation, σ', may be estimated from the collected data. If the variability of the process is not in approximate control there is little basis for estimating σ', and therefore little basis for an $\overline{X}$-Chart.

Recall from the previous section that $E(\bar{\sigma}_i) = C_2 \sigma'_i$. Therefore, if the SCO's are in control the universe standard deviation for each level of input voltage may be estimated as

$$\hat{\sigma}_i = \frac{\bar{\sigma}_i}{C_2} \quad , \qquad [11]$$

where C_2 is defined in equation [8].

Thus, an estimate of the variability of the process may be obtained by averaging the different $\hat{\sigma}_i$ values for the five levels of input voltage, and

$$\hat{\sigma}' = \frac{\sum_{i=1}^{h} \hat{\sigma}_i}{h} \quad . \qquad [12]$$

It should be noted that there are other methods available for estimating σ'. For instance, an estimate of the standard deviation for each level is given by

$$\hat{\sigma}_i = \sqrt{\frac{\sum_{j=1}^{m} \sum_{k=1}^{n} (X_{ijk} - \bar{\bar{X}}_i)^2}{mn}} . \qquad [13]$$

Once this estimate is obtained equation [12] could be applied to give an estimate of σ'.

The major advantages in the first method given over this method are:

1. The estimate of σ' by the first method is on the average less than the corresponding estimate by the above method. Therefore, criteria involving the use of the first method will in the long run detect trouble more often than similar criteria involving the second method (13).

2. The estimate of σ' by the first method involves the use of the $\bar{\sigma}_i$ values which have already been calculated. The second method requires considerable additional calculation. Therefore, a savings in computational labor is brought about by the first method.

A third method for estimating σ' would be to calculate $\hat{\sigma}_i$ from the formula

$$\hat{\sigma}_i = \frac{\sigma_{\sigma_{ij}} \sqrt{2n}}{[2(n-1) - 2nC_2^2]^{1/2}}, \qquad [14]$$

where $\sigma_{\sigma_{ij}}$ is defined by [10] and C_2 is defined in [8], and then apply [12] to give $\hat{\sigma}'$. This method of calculating $\hat{\sigma}'$ can be shown to be valid by computing

$$E(\sigma^2) = \int_0^{\infty} \sigma^2 f(\sigma) d\sigma,$$

where $f(\sigma)d\sigma$ is given in [7], and then computing

$$\sigma_\sigma^{\;2} = [E(\sigma^2) - E(\sigma)]^2 ,$$

$$\sigma_\sigma = \sqrt{\sigma_\sigma^{\;2}} .$$

This method gives essentially the same result as does the first method. However, once again a savings in computational labor can be achieved by using the first method.

There are other methods of estimating σ' from the data. However, the methods presented here are considered to be the most appropriate and therefore are the only ones mentioned. Consequently, we shall use the first method presented; i.e., the application of [11] and [12], to estimate σ'.

Effects of Non-Normality

It was assumed for the development of the σ-Chart that the distribution of individual X_{ijk} values follows the normal distribution. This assumption may not always be valid; therefore, we shall discuss briefly some of the effects of a non-normal population.

The primary limitation caused by a non-normal population is on the statement that the probability that a sample σ_{ij} will fall outside the σ-Chart limits is roughly the same as for an $\overline{X}_{ij}$ value falling outside the $\overline{X}$-Chart limits. If the population is not normal this statement is not necessarily true since we know nothing about the frequency function of a sample standard deviations from a non-normal universe.

Previous work with distributions of telemetry data have indicated that although quite often these distributions are not normal, they are usually unimodal[8] (7). The Camp-Meidel theorem

[8] A unimodal distribution is one which is monotonically decreasing on both sides of its one modal value, or value which occurs the most frequently.

states that if the distribution of the random variable X is unimodal, the probability that X should deviate from its mean by more than K times its standard deviation is equal to or less than $1/2.25K^2$ (4). It can be shown (13) that, although the distributions of σ_{ij} of samples of n are not known for other than the normal universe, nevertheless the moments of the distributions of σ_{ij} are known in terms of the moments of the universe. Hence, we can always establish limits

$$\bar{\sigma}_i \pm K\sigma_{\sigma_{ij}}$$

within which the observed standard deviation should fall more than $100(1 - 1/2.25K^2)$ per cent of the total number of times a sample of n is chosen, so long as the quality of product is controlled. Now if $K = 3$, $100(1 - 1/2.25K^2) = 95.1$ per cent. This is compared with a value of approximately 99 per cent if the normality assumption holds. It is further believed that the main cause of non-normality in telemetry data is peakedness rather than asymmetry. This would possibly indicate that even a larger percentage of the sample σ_{ij} values would fall within the limits than could be predicted by the Camp-Meidel theorem. We shall therefore feel justified in selecting the same K factor for both charts regardless of the form of the distribution of individual values.

The limitation imposed by non-normality would become even more evident in the discussion of the operating characteristic function in CHAPTER IV. This function could not be evaluated without the assumption of normality. Therefore, we will recognize the fact that the data may not be perfectly normally distributed, but will make the assumption of normality realizing that perhaps we have introduced some error into our analysis. It is believed that this error will not seriously affect the methodology and thus may be tolerated.

Summary of the Methodology

The control chart methodology presented in this chapter may be summarized in the following step-by-step procedure:

1. Obtain means and standard deviations for samples of size n for each of the j SCO's, and for each input level, from the formulas

$$\bar{X}_{ij} = \frac{\sum_{k=1}^{n} X_{ijk}}{n}$$

$$\sigma_{ij} = \sqrt{\frac{\sum_{k=1}^{n} (X_{ijk} - \bar{X}_{ij})^2}{n}}$$

$$= \sqrt{\frac{\sum_{k=1}^{n} X_{ijk}^2}{n} - \bar{X}_{ij}^2}$$

2. Compute an average mean, average standard deviation, and standard error of the mean for each level of input from the formulas

$$\bar{\bar{X}}_i = \frac{\sum_{j=1}^{m} \bar{X}_{ij}}{m}$$

$$\bar{\sigma}_i = \frac{\sum_{j=1}^{m} \sigma_{ij}}{m}$$

$$\hat{\sigma}_{\bar{X}_{ij}} = \sqrt{\frac{\sum_{j=1}^{m} (\bar{X}_{ij} - \bar{\bar{X}}_i)^2}{m-1}}$$

3. Determine the center line and control limits for the σ-Chart for each level of input. (The method for choosing the proper K is given in CHAPTER V.)

$$C_L = \bar{\sigma}_i$$
$$UCL = \bar{\sigma}_i + K\sigma_{\sigma_{ij}}$$
$$LCL = \bar{\sigma}_i - K\sigma_{\sigma_{ij}}$$

4. Plot the σ_{ij} values. If these values fall within the limits, the package variability is in control. Any σ_{ij} that falls outside the limits represents an SCO whose variability may not be within proper specifications. These SCO's must then be examined. If an assignable cause is found, the corresponding σ_{ij} must be eliminated (and also the $\bar{X}_{ij}$ value) and the center line and control limits recomputed.

5. Repeat step 4. until all σ_{ij} values are within the limits.

6. Determine the center line and control limits for the $\bar{X}$-Chart for each level of input. (The method for choosing the proper K is given in CHAPTER V.)

$$C_L = \bar{\bar{X}}_i$$
$$UCL = \bar{\bar{X}}_i + K\hat{\sigma}_{\bar{X}_{ij}}$$
$$LCL = \bar{\bar{X}}_i - K\hat{\sigma}_{\bar{X}_{ij}}$$

7. Plot the $\bar{X}_{ij}$ values. If these values fall within the control limits, then the package is in control. Any $\bar{X}_{ij}$ value that falls

outside the limits calls for an examination of that particular SCO. If an assignable cause of variation is found, the corresponding $\overline{X}_{ij}$ value must be eliminated and the center line and limits adjusted.

8. Repeat step 7. until all $\overline{X}_{ij}$ values are within the limits. When this has been accomplished the telemetry package will be in statistical control.

9. Estimate the standard deviation of the package, σ', as

$$\hat{\sigma}' = \frac{\sum_{i=1}^{h} \hat{\sigma}_i}{h} ,$$

where, $\hat{\sigma}_i = \bar{\sigma}_i / C_2$.

CHAPTER IV

THE DEVELOPMENT OF THE PROBABILITY FUNCTIONS FOR THE TYPE I AND TYPE II ERRORS

In the first part of CHAPTER III, we stated in terms of a mathematical model, that there are two possibilities for the state of control of the telemetry process.

(a) Process in Control

Here we assume that the distribution of the process (i.e., the distribution of individual items of product) is normal with a fixed mean $\bar{\bar{X}}_i'$, and fixed standard deviation σ_i', both unknown.

(b) Process out of Control

In this case we again assume that the distribution of the process (at any particular time) is normal with fixed but unknown standard deviation σ_i', but now the process mean is regarded as a chance quantity itself, having a normal distribution with unknown mean $\bar{\bar{X}}_i'$ and standard deviation $\theta\sigma_i'$ where θ is a positive constant.

State (b) was later given in terms of the standard deviation as a chance quantity having mean value σ_i' and standard deviation $\theta\sigma_i'$.

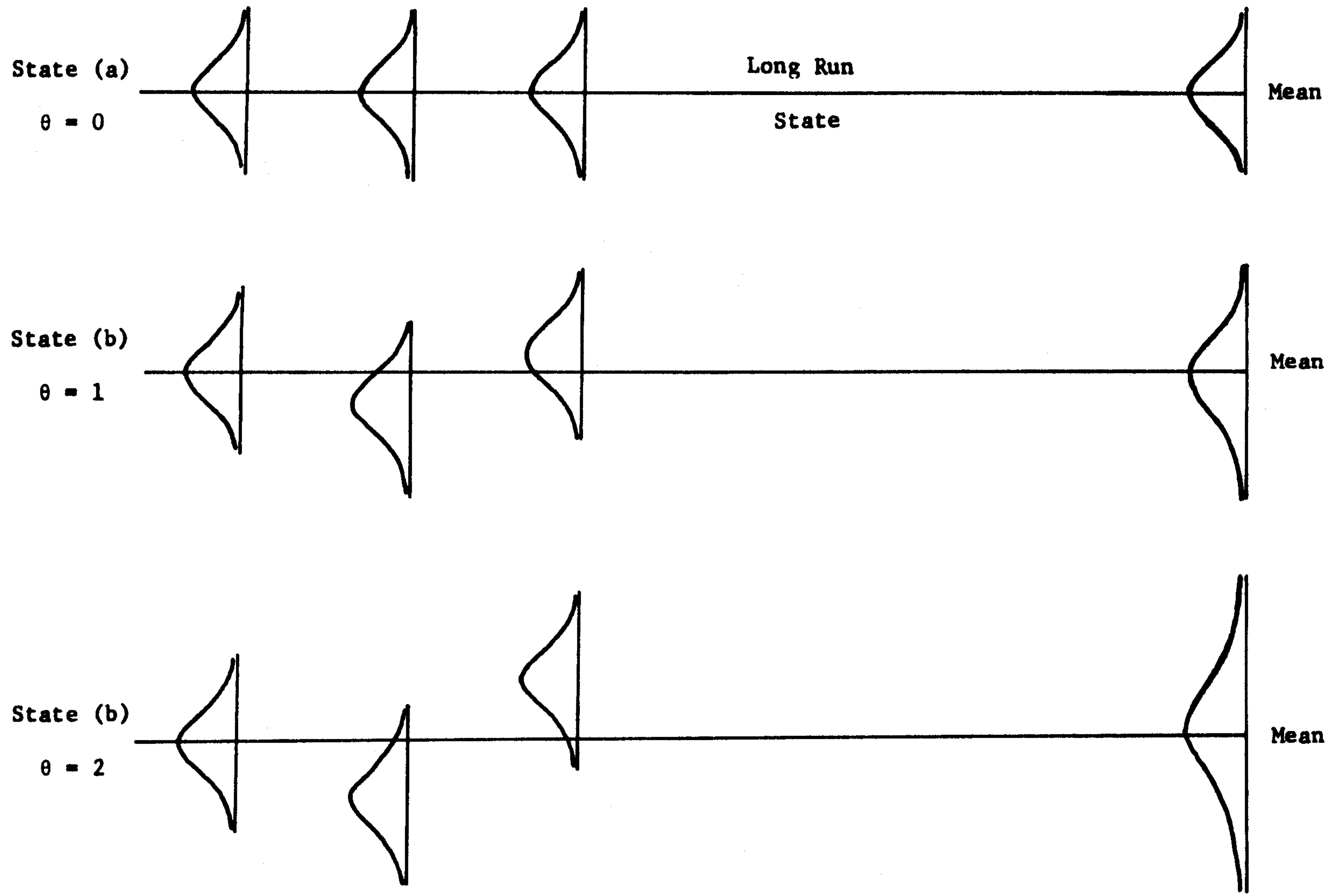

FIGURE 9. POSSIBLE STATES OF THE PROCESS

It was also shown that state (a) is a special case of state (b) when $\theta = 0$. Both of these states are illustrated in Figure 9, where several possibilities are shown corresponding to various values of θ (10,11).

Since either of these states may be present, we are confronted with two types of errors, as was stated previously. First, we may get out-of-control points on the chart when the process is actually in control. The chance of this occurring is the Type I error. On the other hand, we may get no out-of-control points when the process level is actually shifting. The chance of this happening is the Type II error.

In this chapter we will develop a probability function for both charts that will enable us to study these errors for various sample sizes, n, and for various control limit factors, K.

The probability function for the size of the Type II error, β, is called the operating characteristic (OC) function and associated with this function is the OC curve. The OC curve for a control chart used to study past output shows the probability of all m sample points falling inside the control limits. For the given set of sample data studied, this probability is expressed as a function of the actual process characteristics. Therefore, this curve gives a graphic picture of the ability of the control chart to detect trouble.

Let us then first turn to the development of the operating characteristic function for the $\overline{X}$-Chart.

Operating Characteristic Function for the $\overline{X}$-Chart

The operating characteristic function for the $\overline{X}$-Chart gives the probability that, for a selected value of the limit constant K, all of the m sample $\overline{X}$ values will fall within the control limits as a function of a given value of the process mean. This value can be represented by the process mean under control conditions plus some quantity δ, which indicates a shift in this mean. In symbolic notation this probability function may be stated as

$$\beta_{\overline{x}}(K)=P(\overline{\overline{X}}-K\sigma_{\overline{x}} < \overline{X}_1 < \overline{\overline{X}}+K\sigma_{\overline{x}}, \ldots , \overline{\overline{X}}-K\sigma_{\overline{x}} < \overline{X}_m < \overline{\overline{X}}+K\sigma_{\overline{x}}|\overline{\overline{X}}'_{\delta}), \qquad [15]$$

where $\overline{\overline{X}}'_{\delta} = \overline{\overline{X}}' + \delta$.

Thus we wish to evaluate [15] for various values of K, n, and δ.

First let us attempt to simplify [15]. By transposing the $\overline{\overline{X}}$ and dividing by $\sigma_{\overline{x}}$ in each inequality we obtain

$$\beta_{\overline{x}}(K)=P[-K < (\overline{X}_1 - \overline{\overline{X}})/\sigma_{\overline{x}} < K, \ldots , -K < (\overline{X}_m - \overline{\overline{X}})/\sigma_{\overline{x}} < K|\overline{\overline{X}}'_{\delta}].$$

The quantity $Z_j = (\overline{X}_j - \overline{\overline{X}})/\sigma_{\overline{x}}$ is distributed as a modified "t" distribution known as the Hotelling T^2 distribution (4). This distribution is often rather difficult to work with in solving theoretical problems. Therefore, we shall assume that Z_j is approximately distributed as a normal distribution with a mean of zero and a variance of one (see Appendix A, p. 87). We must realize that some error is introduced by this assumption. However, the error introduced due to the complexity of the solution of OC function when the Hotelling T^2 distribution is used in thought to be more than the error due to the normality assumption. Also, hypotheses testing when normality is assumed is more conservative than when a "t" distribution is assumed. Thus, we wish to evaluate, for

$Z_j \sim N(0,1)$,

$$\beta_{\bar{\bar{X}}}(K) = P(-K < Z_1 < K, \ldots , -K < Z_m < K|\bar{\bar{X}}_\delta').$$

Let us now examine the solution of the probability of one of the m inequalities for a given $\bar{\bar{X}}_\delta'$. This situation can be analyzed in terms of the two curves shown in Figure 10.

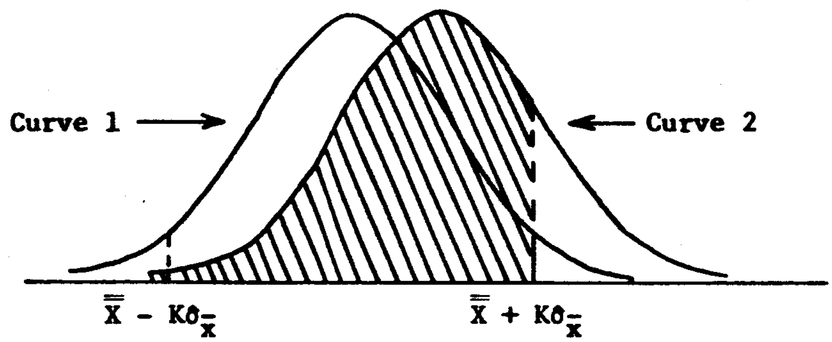

FIGURE 10. GRAPHIC REPRESENTATION OF A SHIFT IN THE PROCESS MEAN FROM $\bar{\bar{X}}'$ TO $\bar{\bar{X}}_\delta'$.

We wish to determine the shaded area for curve 2. That is, to evaluate

$$P(-K < Z_j < K|\bar{\bar{X}}_\delta')$$

we need to find the area under the normal curve with mean $\bar{\bar{X}}_\delta'$ (curve 2) between the -K and K limits that were placed on our original curve (curve 1). This can be done by expressing the -K and K limits in terms of their values on curve 2 and then integrating the standardized normal equation (see Appendix A, p.87) between these new limits. As was stated previously, the

Z equation for curve 1 is given by

$$Z_{(1)} = (\bar{X} - \bar{\bar{X}})/\hat{\sigma}_{\bar{\bar{X}}} \quad .$$

The Z equation for curve 2 is given by

$$Z_{(2)} = [\bar{X} - (\bar{\bar{X}} + \delta)]/\hat{\sigma}_{\bar{\bar{X}}} \quad .$$

By taking the difference, $Z_{(2)} - Z_{(1)}$, we can determine the amount that the standardized Z value has shifted.

$$\text{Shifted } Z = \frac{\bar{X} - (\bar{\bar{X}} + \delta)}{\hat{\sigma}_{\bar{\bar{X}}}} - \frac{(\bar{X} - \bar{\bar{X}})}{\hat{\sigma}_{\bar{\bar{X}}}}$$

$$= -\delta/\hat{\sigma}_{\bar{\bar{X}}} \quad .$$

We shall assume that any shift in the mean, δ, can be expressed as a parameter θ multiplied by $\sigma_{\bar{\bar{X}}}'$. Therefore,

$$\delta = \theta\sigma_{\bar{\bar{X}}}' \quad ,$$

and the shifted Z value becomes

$$\text{Shifted } Z = -\theta\sigma_{\bar{\bar{X}}}'/\hat{\sigma}_{\bar{\bar{X}}} \quad .$$

Since we do not know what $\sigma_{\bar{\bar{X}}}'$ is, it again must be estimated as $\hat{\sigma}_{\bar{\bar{X}}} = \sigma_{\bar{X}} \sqrt{m/(m-1)}$. Hence,

$$\delta = \theta\hat{\sigma}_{\bar{\bar{X}}} \quad , \qquad [17]$$

and $\quad \text{Shifted } Z = -\theta\hat{\sigma}_{\bar{\bar{X}}}/\hat{\sigma}_{\bar{\bar{X}}} = -\theta \quad .$

Since the Z value has shifted by an amount of $-\theta$, the $-K$ limit on curve 1 can be expressed as a point on curve 2 as $-K-\theta$. Likewise, the K limit will be, on curve 2, $K-\theta$.

Therefore,

$$P(-K < Z_j < K|\bar{\bar{X}}_{\delta}') = 1/\sqrt{2\pi} \cdot \int_{-K-\theta}^{K-\theta} e^{-Z^2/2}\, dZ$$

$$= 1 - 1/\sqrt{2\pi}\left(\int_{-\infty}^{-K-\theta} e^{-Z^2/2}\,dZ + \int_{K-\theta}^{\infty} e^{-Z^2/2}\,dZ\right).$$

Since we have assumed that the various $\overline{X}$ values are independent, then necessarily their corresponding Z values are independent, and equation [16] becomes (for a given value of δ)

$$\beta_{\overline{X}}(K) = \left[1 - 1/\sqrt{2\pi}\left(\int_{-\infty}^{-K-\theta} e^{-Z^2/2}\,dZ + \int_{K-\theta}^{\infty} e^{-Z^2/2}\,dZ\right)\right]^m . \quad [18]$$

This is due to the multiplication theorem of probability for independent events (see for instance 2, 4, 9).

Thus, for a given sample size n, θ can be determined for a selected value of δ from equation [17] and then equation [18] can be solved for various values of K.

The OC curve for the $\overline{X}$-Chart can be plotted by marking the horizontal axis with different values of δ and the vertical axis with $\beta_{\overline{X}}(K)$ and then computing $\beta_{\overline{X}}(K)$ for the different δ's and plotting the results.

We shall next consider the development of the operating characteristic function for the σ-Chart.

Operating Characteristic Function for the σ-Chart

The operating function for the σ-Chart gives the probability that, for a selected K value, all of the m sample σ's will fall within the control limits as a function of a given value of the process standard deviation. This value of the process standard deviation can be represented by the process standard deviation under control conditions plus the quantity δ, which now indicates some shift in this standard deviation. Symbolically we have

$$\beta_\sigma(K) = P(\bar{\sigma}-K\sigma_\sigma < \sigma_1 < \bar{\sigma}+K\sigma_\sigma, \ldots, \bar{\sigma}-K\sigma_\sigma < \sigma_m < \bar{\sigma}+K\sigma_\sigma | \sigma'_\delta), \qquad [19]$$

where $\sigma'_\delta = \sigma' + \delta$.

If we divide both sides of each inequality by σ'_δ we obtain

$$\beta_\sigma(K) = P\left(\frac{\bar{\sigma} - K\sigma_\sigma}{\sigma'_\delta} < \frac{\sigma_1}{\sigma'_\delta} < \frac{\bar{\sigma} + K\sigma_\sigma}{\sigma'_\delta}, \ldots, \frac{\bar{\sigma} - K\sigma_\sigma}{\sigma'_\delta} < \frac{\sigma_m}{\sigma'_\delta} < \frac{\bar{\sigma} + K\sigma_\sigma}{\sigma'_\delta} \middle| \sigma'_\delta\right) .$$

Now the quantity $n\sigma^2/\sigma'^2_\delta$ is known to be distributed as a chi square (χ^2) distribution (4). Therefore,

$$\sigma_j/\sigma'_\delta = \sqrt{\chi^2/n}$$

and equation [19] becomes (for a given value of δ)

$$\beta_\sigma(K) = [\, p\, (\, LCL_\sigma/\sigma'_\delta < \sqrt{\chi^2/n} < UCL_\sigma/\sigma'_\delta\,)\,]^m .$$

By squaring each side of the inequality and multiplying through by n, we finally obtain

$$\beta_\sigma(K) = \{\, P\, [\, (LCL_\sigma/\sigma'_\delta)^2 n < \chi^2 < (UCL_\sigma/\sigma'_\delta)^2 n \,]\, \}^m . \qquad [20]$$

This probability function can be solved for a selected value of K and σ'_δ by integrating the frequency function of the χ^2 distribution over the indicated limits. The frequency function of the χ^2 distribution is given by

$$f(\chi^2) = \frac{(\chi^2)^{v/2-1}\, e^{-\chi^2/2}}{2^{v/2}\, \Gamma(v/2)},$$

where v is the number of degrees of freedom given by $v = n-1$.

Thus,

$$\beta_\sigma(K) = \left[\frac{1}{2^{v/2}\,\Gamma(v/2)} \int_{x_1}^{x_2} (\chi^2)^{v/2 - 1} e^{-\chi^2/2} d\chi^2 \right]^m ,$$

where

$$x_1 = (LCL_\sigma/\sigma\hat{\delta})^2 \, n \text{ , and}$$

$$x_2 = (UCL_\sigma/\sigma\delta)^2 \, n \text{ .}$$

This integral can be evaluated by expanding $e^{-\chi^2/2}$ in a power series and then multiplying each term by $(\chi^2)^{v/2 - 1}$ and integrating term by term (see Appendix A, p. 88). However, most statistical textbooks contain tables of this integral and therefore, the OC function for the σ-Chart can be evaluated by using these tables.

The OC curve for the σ-Chart can be plotted by marking the horizontal axis with δ and the vertical axis with $\beta_\sigma(K)$ and then computing $\beta_\sigma(K)$ for the various δ values and plotting the results.

The probability function for the size of the Type I error, α, is relatively simple. This function gives the probability that sample points will fall outside the ±K control chart limits when the process is actually in control.

Probability Function for α for the $\overline{X}$-Chart

The probability that sample means will fall outside the ±K limits when the process mean is actually in control can be found by integrating the standardized normal equation over the area outside these limits.

Thus,

$$\alpha_{\bar{X}}(K) = 1/\sqrt{2\pi}\left[\int_{-\infty}^{-K} e^{-Z^2/2}\, dZ + \int_{K}^{\infty} e^{-Z^2/2}\, dZ\right],$$

or

$$\alpha_{\bar{X}}(K) = 2/\sqrt{2\pi}\int_{K}^{\infty} e^{-Z^2/2}\, dZ.$$

Probability Function for α for the σ-Chart

The probability that sample standard deviations will fall outside of the $\pm K$ limits when the process standard deviation is actually in control can be found by integrating the frequency function for σ over the area outside these limits.

Thus,

$$\alpha_{\sigma}(K) = \int_{-\infty}^{-K} f(\sigma)d\sigma + \int_{K}^{\infty} f(\sigma)d\sigma ,$$

where $f(\sigma)d\sigma$ is defined by equation [7].

Since the distribution of σ is not necessarily symmetrical, $\alpha_{\sigma}(K)$ cannot be further reduced.

CHAPTER V

THE DETERMINATION OF THE PROPER LIMIT CONSTANT

The determination of the proper limit constant, K, requires an economic evaluation of the risks involved in making an incorrect decision: i.e., the alpha and beta risks. We must select a K which will strike some economic balance between these risks. Therefore, to enable us to select the optimum K, we need to formulate a cost model which will represent the total cost attributed to these errors and then choose the value of K which will yield the lowest possible total cost.

The control chart cost model, using as a basis one telemetry package, may be stated as

$$TC = c_1\beta + c_2\alpha + c_3 n, \qquad [21]$$

where c_1 = unit cost of the β risk

c_2 = unit cost of the α risk

c_3 = unit cost of control chart sampling which is directly dependent on the sample size, n.

Note that in this model the cost of control chart sampling, $c_3 n$, will not be involved in the selection of the optimum K. However, it will be necessary to consider this term when we later select the best sample size, and thus this term is included in the cost model.

To determine the optimum K for the $\overline{X}$-Chart the β and α in equation [21] will be $\beta_{\bar{X}}$ and $\alpha_{\bar{X}}$. It was stated in CHAPTER I that we may use the same K factor for both the $\bar{X}$- and σ-Chart. Therefore, we will base our decision on K for both charts by using $\beta_{\bar{X}}$ and $\alpha_{\bar{X}}$ in equation [21] since these expressions are somewhat simpler than are β_σ and α_σ.

To find the optimum K it will be necessary to differentiate the total cost equation with respect to K, set this derivative equal to zero, and solve for K. This value for K will result in the minimum total cost if the second derivative of [21], evaluated at the optimum K, is positive. If the second derivative of [21] evaluated at the optimum K is negative, we will have found a value for K which maximizes the total cost. This, of course, is not our objective.

Let us proceed to determine K. Taking the first derivative of [21] with respect to K we obtain

$$\frac{d(TC)}{dK} = c_1\frac{d\beta}{dK} + c_2\frac{d\alpha}{dK} .$$

$$\text{Since } \beta_{\bar{X}} = \left[1 - 1/\sqrt{2\pi}\left(\int_{-\infty}^{-K-\theta} e^{-Z^2/2}\, dZ + \int_{K-\theta}^{\infty} e^{-Z^2/2}\, dZ\right)\right]^m$$

$$\text{and } \alpha_{\bar{X}} = \left[2/\sqrt{2\pi} \int_{K}^{\infty} e^{-Z^2/2}\, dZ\right] ,$$

the partial derivatives $\frac{d\beta}{dK}$ and $\frac{d\alpha}{dK}$ become

$$\frac{d\beta}{dK} = \frac{d}{dK}\left[1 - 1/\sqrt{2\pi}\left(\int_{\infty}^{-K-\theta} e^{-Z^2/2}\, dZ + \int_{K-\theta}^{\infty} e^{-Z^2/2}\, dZ\right)\right]^m$$

$$\frac{d\alpha}{dK} = \frac{d}{dK}\left[2/\sqrt{2\pi}\int_{K}^{\infty} e^{-Z^2/2}\, dZ\right].$$

To evaluate these two derivatives we must use the method of differentiation of integrals (see Appendix A, p. 90). Applying this method we obtain for $\frac{d\alpha}{dK}$

$$\frac{d\alpha}{dK} = \frac{2}{\sqrt{2\pi}}\,\frac{d}{dK}\int_{K}^{\infty} e^{-Z^2/2}\, dZ$$

$$= \frac{2}{\sqrt{2\pi}}\left[F(\infty)\,\frac{d\infty}{dK} - F(K)\,\frac{dK}{dK}\right]$$

$$= \frac{2}{\sqrt{2\pi}}\left[0 - e^{-K^2/2}\right]$$

$$= -\frac{2}{\sqrt{2\pi}}\, e^{-K^2/2}\,.$$

The evaluation of $\frac{d\beta}{dK}$ is somewhat more complicated. We first obtain

$$\frac{d\beta}{dK} = \frac{d}{dK}\left[1/\sqrt{2\pi}\int_{-K-\theta}^{K-\theta} e^{-Z^2/2}\, dZ\right]^m$$

$$= \frac{m}{\sqrt{2\pi}} \left\{ \left[\int_{-K-\theta}^{K-\theta} e^{-Z^2/2} \, dZ \right]^{m-1} \cdot \frac{d}{dK} \int_{-K-\theta}^{K-\theta} e^{-Z^2/2} \, dZ \right\}$$

$$= \frac{m}{\sqrt{2\pi}} \left\{ \left[\int_{-K-\theta}^{K-\theta} e^{-Z^2/2} \, dZ \right]^{m-1} \cdot \left(e^{-(K-\theta)^2/2} \frac{d(K-\theta)}{dK} \right.\right.$$

$$\left.\left. - e^{-(-K-\theta)^2/2} \frac{d(-K-\theta)}{dK} \right) \right\}$$

$$= \frac{m}{\sqrt{2\pi}} \left\{ \left[\int_{-K-\theta}^{K-\theta} e^{-Z^2/2} \, dZ \right]^{m-1} \cdot \left(e^{-(K-\theta)^2/2} + e^{-(-K-\theta)^2/2} \right) \right\}.$$

In order to eliminate the integral sign let us expand $e^{-Z^2/2}$ in a power series and integrate term by term.

$$\int_{-K-\theta}^{K-\theta} e^{-Z^2/2} \, dZ = \int_{-K-\theta}^{K-\theta} \left(1 - \frac{Z^2}{2} + \frac{Z^4}{4.2!} - \frac{Z^6}{8.3!} + \ldots \right) dZ$$

$$= Z - \frac{Z^3}{6} + \frac{Z^5}{40} - \frac{Z^7}{336} + \ldots \Bigg|_{-K-\theta}^{K-\theta}$$

$$= \left[(K-\theta) - \frac{(K-\theta)^3}{6} + \frac{(K-\theta)^5}{40} - \frac{(K-\theta)^7}{336} + \ldots \right]$$

$$- \left[(-K-\theta) - \frac{(-K-\theta)^3}{6} + \frac{(-K-\theta)^5}{40} - \frac{(-K-\theta)^7}{336} + \ldots \right].$$

Factoring (K-θ) from the first series and (-K-θ) from the second yields

$$\int_{-K-\theta}^{K-\theta} e^{-Z^2/2}\,dZ = (K-\theta)\left[1 - \frac{(K-\theta)^2}{2.3} + \frac{(K-\theta)^4}{4.5.2!} - \frac{(K-\theta)^6}{8.7.3!} + \ldots\right]$$

$$- (-K-\theta)\left[1 - \frac{(-K-\theta)^2}{2.3} + \frac{(-K-\theta)^4}{4.5.2!} - \frac{(-K-\theta)^6}{8.7.3!} + \ldots\right]$$

$$=(K-\theta)\left[\sum_{n=0}^{\infty} (-1)^n \frac{(K-\theta)^{2n}}{2^n n!(2n+1)}\right] - (-K-\theta)\left[\sum_{n=0}^{\infty} (-1)^n \frac{(-K-\theta)^{2n}}{2^n n!(2n+1)}\right].$$

if we remove $\sum_{n=0}^{\infty} 1/2^n$ from each bracketed expression, the remaining terms will define a function of e in series form. Thus,

$$\int_{-K-\theta}^{K-\theta} e^{-Z^2/2}\,dZ = \left[(K-\theta)e^{-(K-\theta)^2/2} \sum_{n=0}^{\infty} 1/2^n\right] - \left[(-K-\theta)e^{-(-K-\theta)^2/2} \sum_{n=0}^{\infty} 1/2^n\right].$$

The sum of the series $\sum_{n=0}^{\infty} 1/2^n$ is 2; therefore,

$$\int_{-K-\theta}^{K-\theta} e^{-Z^2/2}\,dZ = 2\left[(K-\theta)e^{-(K-\theta)^2/2} + (K+\theta)e^{-(-K-\theta)^2/2}\right].$$

Thus,

$$\frac{d(TC)}{dK} = \frac{2mc_1}{\sqrt{2\pi}}\left\{\left[(K-\theta)e^{-(K-\theta)^2/2} + (K+\theta)e^{-(-K-\theta)^2/2}\right]^{m-1}\right.$$

$$\cdot \left[e^{-(K-\theta)^2/2} + e^{-(-K-\theta)^2/2} \right] \Bigg\} - \frac{2c_2}{\sqrt{2\pi}}\, e^{-K^2/2} \quad . \qquad [22]$$

Setting this equation equal to zero and taking $\log_e$ of both sides to further simplify the equation gives

$$\log_e \frac{2mc_1}{\sqrt{2\pi}} + \log_e \Bigg\{ \left[(K-\theta)e^{-(K-\theta)^2/2} + (K+\theta)e^{-(-K-\theta)^2/2} \right]^{m-1}$$

$$\cdot \left[e^{-(K-\theta)^2/2} + e^{-(-K-\theta)^2/2} \right] \Bigg\} = \log_e \frac{2c_2}{\sqrt{2\pi}}\, e^{-K^2/2} \quad ,$$

or

$$\log_e \frac{2mc_1}{\sqrt{2\pi}} + (m-1)\log_e \left[(K-\theta)e^{-(K-\theta)^2/2} + (K+\theta)e^{-(-K-\theta)^2/2} \right] +$$

$$\log_e \left[e^{-(K-\theta)^2/2} + e^{-(-K-\theta)^2/2} \right] = \log_e \frac{2c_2}{\sqrt{2\pi}} + \log_e e^{-K^2/2} \quad .$$

Collecting constant terms on the right hand side,

$$(m-1)\log_e \left[(K-\theta)e^{-(K-\theta)^2/2} + (K+\theta)e^{-(-K-\theta)^2/2} \right] +$$

$$\log_e \left[e^{-(K-\theta)^2/2} + e^{-(-K-\theta)^2/2} \right] - \log_e e^{-K^2/2}$$

$$= \log_e \frac{2c_2}{\sqrt{2\pi}} - \log_e \frac{2mc_1}{\sqrt{2\pi}} \quad .$$

Factoring $e^{-(K-\theta)^2/2}$ from each bracketed expression, we obtain

$$(m-1)\left\{\log_e \left[e^{-(K-\theta)^2/2} + \log_e \; (K-\theta) + (K+\theta)\; e^{-4K\theta}\right]\right\} +$$

$$\log_e \left[e^{-(K-\theta)^2/2}\right] + \log_e \left[1 + e^{-4K\theta}\right] - \log_e e^{-K^2/2}$$

$$= \log_e \frac{2c_2}{\sqrt{2\pi}} - \log_e \frac{2mc_1}{\sqrt{2\pi}} \; .$$

Simplifying this expression yields

$$\frac{(1-m)(K-\theta)^2}{2} + (m-1) \log_e \left[(K-\theta) + (K+\theta)\; e^{-4K\theta}\right]$$

$$- \frac{(K-\theta)^2}{2} + \log_e (1 + e^{-4K\theta}) + \frac{K^2}{2} = \log_e \frac{2c_2}{\sqrt{2\pi}} - \log_e \frac{2mc_1}{\sqrt{2\pi}} \; ,$$

or

$$\frac{(1-m)(K-\theta)^2}{2} + (m-1) \log_e \left[(K-\theta) + (K+\theta)\; e^{-4K\theta}\right]$$

$$- \frac{K^2}{2} + \frac{2K\theta}{2} + \frac{K^2}{2} + \log_e (1 + e^{-4K\theta}) = \frac{\theta^2}{2} + \log_e \frac{2c_2}{\sqrt{2\pi}} - \log_e \frac{2mc_1}{\sqrt{2\pi}} \; .$$

This finally reduces to

$$\frac{(1-m)(K-\theta)^2}{2} + (m-1) \log_e \left[(K-\theta) + (K+\theta)\; e^{-4K\theta}\right] +$$

$$\frac{K\theta}{2} + \log_e (1 + e^{-4K\theta}) = \frac{\theta^2}{2} + \log_e \frac{2c_2}{\sqrt{2\pi}} - \log_e \frac{2mc_1}{\sqrt{2\pi}} \; . \qquad [23]$$

This equation may be solved for K by the method of trial and error for any selected values of c_1, c_2, and θ.

As stated previously, to prove that the K obtained by solving [23] is the optimum K, i.e., the one which minimizes [21], we must differentiate [22] with respect to K, substitute our optimum K and obtain a positive result. Differentiating [22] we obtain

$$
\begin{aligned}
\frac{d^2(TC)}{dK^2} = \frac{2mc_1}{\sqrt{2\pi}} \Bigg\{ & (m-1) \left[(K-\theta)\, e^{-(K-\theta)^2/2} + (K+\theta) e^{-(-K-\theta)^2/2} \right]^{m-2} \\
& \cdot \left[(-\theta)\, e^{-(K-\theta)^2/2} - (K-\theta)^2 e^{-(K-\theta)^2/2} + \theta e^{-(-K-\theta)^2/2} + \right. \\
& \left. (K+\theta)^2 e^{-(-K-\theta)^2/2} \right] \\
& \cdot \left[e^{-(K-\theta)^2/2} + e^{-(-K-\theta)^2/2} \right] + \left[(K-\theta) e^{-(K-\theta)^2/2} + (K+\theta) e^{-(-K-\theta)^2/2} \right]^{m-1} \\
& \cdot \left[(-K-\theta) e^{-(K-\theta)^2/2} + (K+\theta) e^{-(-K-\theta)^2/2} \right] \Bigg\} + \frac{2Kc_2}{\sqrt{2\pi}}\, e^{-K^2/2} \quad .
\end{aligned}
$$

CHAPTER VI

APPLICATION OF THE METHODOLOGY

The control chart methodology is now completely formulated; therefore, we may apply the methodology to a telemetry package experiment and hopefully obtain meaningful results.

Description of the Experimental Output

The experimental set-up was described in CHAPTER III and a block diagram of the experiment given by Figure 5. There were 14 subcarrier oscillators available in the experimental telemetry package. Two of these SCO's, channels 4 and 6, were purposely maladjusted SCO's. Also two of the SCO's, channels 12 and 13, were purposely caused to have higher variability than the others by injecting random noise into the system through these channels. This condition would correspond to a package having two "noisy" SCO's. These two types of malfunctioning components represent the conditions for which the control charts have been designed to detect. Therefore, if the methodology has been formulated correctly, the four SCO's (channels 4, 6, 12, and 13) will be judged out of control by the control charts. This will, of course, result in a decision to investigate these SCO's for assignable causes of variability.

Printed values of means and standard deviations for the 14 SCO's at the 0% level of input for four different sample sizes are given in Table 2. The selection of the optimum K factor will be based on these values rather than on all five input levels. It is believed that essentially the same decision as to the proper K would be reached regardless of the level of input voltage chosen. Also, the optimum sample size, n, will be determined from these values. Other sample sizes could have been chosen, but the four sizes listed seem to represent the most rational choices and thus, for computational simplicity, one of these four will be selected. After the choice of K and n is made the control charts will be applied at each of the five levels of input.

Table 1 contains values of $\bar{\bar{X}}$, $\sigma_{\bar{x}}$, $\bar{\sigma}$, and σ_σ for each of the four sample sizes for the data given in Table 2.

TABLE 1

VALUES OF $\bar{\bar{X}}$, $\sigma_{\bar{x}}$, $\bar{\sigma}$, AND σ_σ FOR 0% INPUT LEVEL

n	$\bar{\bar{X}}$	$\sigma_{\bar{x}}$	$\bar{\sigma}$	σ_σ
5	65.31	7.45	2.65	3.47
10	65.29	7.26	2.41	2.86
15	65.35	7.41	2.36	2.57
25	65.43	7.36	2.41	2.39

TABLE 2

MEANS AND STANDARD DEVIATIONS FOR 0% INPUT LEVEL

		SCO Channel													
		2	3	4	5	6	7	8	9	10	11	12	13	15	16
n= 5	$\overline{X}$	72.00	66.00	42.00	60.80	75.80	69.40	66.00	66.00	69.00	63.80	26.20	65.40	64.40	64.60
	σ	1.67	0.89	1.41	1.47	1.33	1.62	0.89	1.79	1.41	0.75	14.33	6.09	1.36	2.06
n=10	$\overline{X}$	71.30	65.80	42.20	61.70	76.20	69.70	66.30	65.90	68.50	63.40	68.10	65.70	64.00	65.20
	σ	0.90	0.98	0.98	1.00	1.54	1.55	1.73	1.37	1.28	1.43	11.76	5.87	1.26	2.14
n=15	$\overline{X}$	71.60	66.07	41.87	61.47	76.00	69.60	65.93	65.53	68.47	63.73	68.47	66.80	64.20	65.20
	σ	1.31	0.91	1.49	1.14	1.10	1.45	1.20	1.79	1.18	1.19	9.92	7.02	1.17	2.20
n=25	$\overline{X}$	71.40	65.72	42.04	61.80	76.08	69.52	65.68	65.60	68.28	63.48	69.64	67.12	64.36	65.28
	σ	1.30	1.40	1.28	1.13	1.41	1.70	1.69	1.26	1.40	1.24	9.01	7.03	1.23	2.42

Determination of Optimum K and n

Let us now consider the use of equation [23] in determining the optimum K. We shall consider three possible values of δ, the shift in the universe mean. These values of δ are 2, 5, and 10. Once again, it should be noted that many other values might have been chosen; however, these seem to represent values which we are most interested in detecting. Also, let us assume the following values for our cost parameters: $c_1 = 10$, $c_2 = 1$, $c_3 = 0.001$. The validity of this assumption cannot be readily assessed; however, the researcher intuitively believes that these cost parameters are fairly realistic in relation to each other. Using these assumptions, the trial and error method of solving [23] yields the optimum values of K listed in Table 3 as a function of n and δ.

TABLE 3

OPTIMUM VALUES OF K

δ \ n	5	10	15	25
2	1.54	1.57	1.57	1.59
5	1.61	1.68	1.68	1.70
10	2.09	2.15	2.15	2.18

A sample computation is given in Appendix A, pp. 90,91 to illustrate the method.

Table 4 gives values of total cost, determined from equation [21], for the assumed cost parameters.

TABLE 4

VALUES OF TOTAL COST ($TC = c_1\beta + c_2\alpha + c_3n$)

δ \ n	5	10	15	25
2	1.4576	1.5074	1.5124	1.5224
5	0.8614	1.1070	1.1120	1.1220
10	0.5896	0.7586	0.7636	0.7736

It can be seen from Table 4 that the optimum sample size (the value of n which results in the minimum total cost) is $n = 5$ regardless of the size shift in the universe mean that we wish to detect. Thus, Table 5 gives means and standard deviations based on a sample size of $n = 5$ for the other four levels of voltage input.

Before applying the control charts to the five levels of input we must decide on the amount of shift in the mean we wish to detect since different values of δ constitute different values of K. Let us assume that we are most concerned in detecting shifts of size $\delta = 5$. Under this assumption, the optimum K given in Table 3 is $K = 1.61$.

Analysis by Control Charts

We first analyze the data given given in Table 2 for $n = 5$ for

TABLE 5

MEANS AND STANDARD DEVIATIONS FOR 25%, 50%, 75%, AND 100% INPUT LEVELS

		SCO Channel 2	3	4	5	6	7	8
25%	$\overline{X}$	293.20	289.80	270.60	284.00	304.20	291.20	290.00
	σ	1.72	1.47	0.98	1.67	1.00	0.75	2.10
50%	$\overline{X}$	514.60	513.20	497.00	507.40	531.40	513.40	513.20
	σ	1.02	1.47	1.10	0.49	1.02	1.36	1.72
75%	$\overline{X}$	737.80	736.60	726.00	732.00	759.60	736.00	736.80
	σ	1.17	1.02	1.09	0.63	0.49	0.63	1.17
100%	$\overline{X}$	957.36	960.88	953.60	962.28	988.20	960.00	960.60
	σ	0.79	1.34	0.80	2.42	0.98	0.63	0.80

TABLE 5--Continued

		SCO Channel 9	10	11	12	13	15	16
25%	$\overline{X}$	289.00	289.40	286.80	281.40	291.80	288.00	289.00
	σ	1.41	1.36	0.98	7.76	3.31	1.67	2.61
50%	$\overline{X}$	512.00	512.60	511.20	512.60	510.40	512.20	512.40
	σ	0.89	1.36	1.17	14.97	6.02	1.17	1.62
75%	$\overline{X}$	736.20	736.80	736.60	734.60	739.40	736.40	736.60
	σ	1.60	0.40	1.62	5.64	7.12	0.80	1.74
100%	$\overline{X}$	960.20	960.00	960.40	963.00	963.60	961.00	958.60
	σ	0.75	0.63	1.36	8.37	5.53	2.10	1.74

variability by the use of the σ-Chart. From Table 2, $\bar{\sigma} = 2.65$ and $\sigma_\sigma = 3.47$. Thus, with $K = 1.61$, our control limits and center line become

$$\begin{aligned} UCL &= \bar{\sigma} + K\sigma_\sigma \\ &= 2.65 + (1.61)(3.47) \\ &= 5.59 \\ LCL &= \bar{\sigma} - K\sigma_\sigma \\ &= 2.65 - (1.61)(3.47) \\ &= 0 \quad \text{, since negative} \\ C_L &= \bar{\sigma} = 2.65 \end{aligned}$$

The σ-Chart is given in Figure 11, (a). SCO channels 12 and 13 are out of control. Therefore, these SCO's require investigation and both their $\bar{X}$ and σ values are eliminated from the data. Recomputing $\bar{\sigma}$ and σ_σ we have

$$\begin{aligned} \bar{\sigma} &= 1.39 \\ \sigma_\sigma &= 0.37 \end{aligned}$$

Thus, our revised limits and center line become

$$\begin{aligned} UCL &= 1.99 \\ C_L &= 1.39 \\ LCL &= 0.79 \end{aligned}$$

We now observe from Figure 11, (b) that channel 16 is out of control. Once again, we eliminate the data for this channel and revise our limits and center line.

$$\begin{aligned} \bar{\sigma} &= 1.33 \\ \sigma_\sigma &= 0.33 \end{aligned}$$

$$UCL = 1.86$$

$$C_L = 1.33$$

$$LCL = 0.80$$

Plotting the remaining values of σ in Figure 11, (c), we observe that the only point falling outside of the limits is channel 11. However, this point is only slightly below the lower limit and thus the SCO's are now judged to be in control with respect to variability at the 0% level. An estimate of σ' for this level is given by

$$\begin{aligned} \hat{\sigma}_1 &= \bar{\sigma}_1 / C_2 \\ &= (1.33)/(0.84) \\ &= 1.58 \end{aligned}$$

Now we apply the $\bar{X}$-Chart to the means listed in Table 2 for n = 5, remembering that the data for channels 12, 13, and 16 have been eliminated. Recomputing $\bar{\bar{X}}$ and $\hat{\sigma}_{\bar{X}}$ we obtain

$$\bar{\bar{X}} = 65.02$$

$$\hat{\sigma}_{\bar{X}} = 8.26$$

Thus our control limits become

$$\begin{aligned} UCL &= \bar{\bar{X}} + K\hat{\sigma}_{\bar{X}} \\ &= 65.02 + (1.61)(8.26) \\ &= 78.32 \end{aligned}$$

$$\begin{aligned} LCL &= \bar{\bar{X}} - K\hat{\sigma}_{\bar{X}} \\ &= 65.02 - (1.61)(8.26) \\ &= 78.32 \end{aligned}$$

$$C_L = \bar{\bar{X}} = 65.02$$

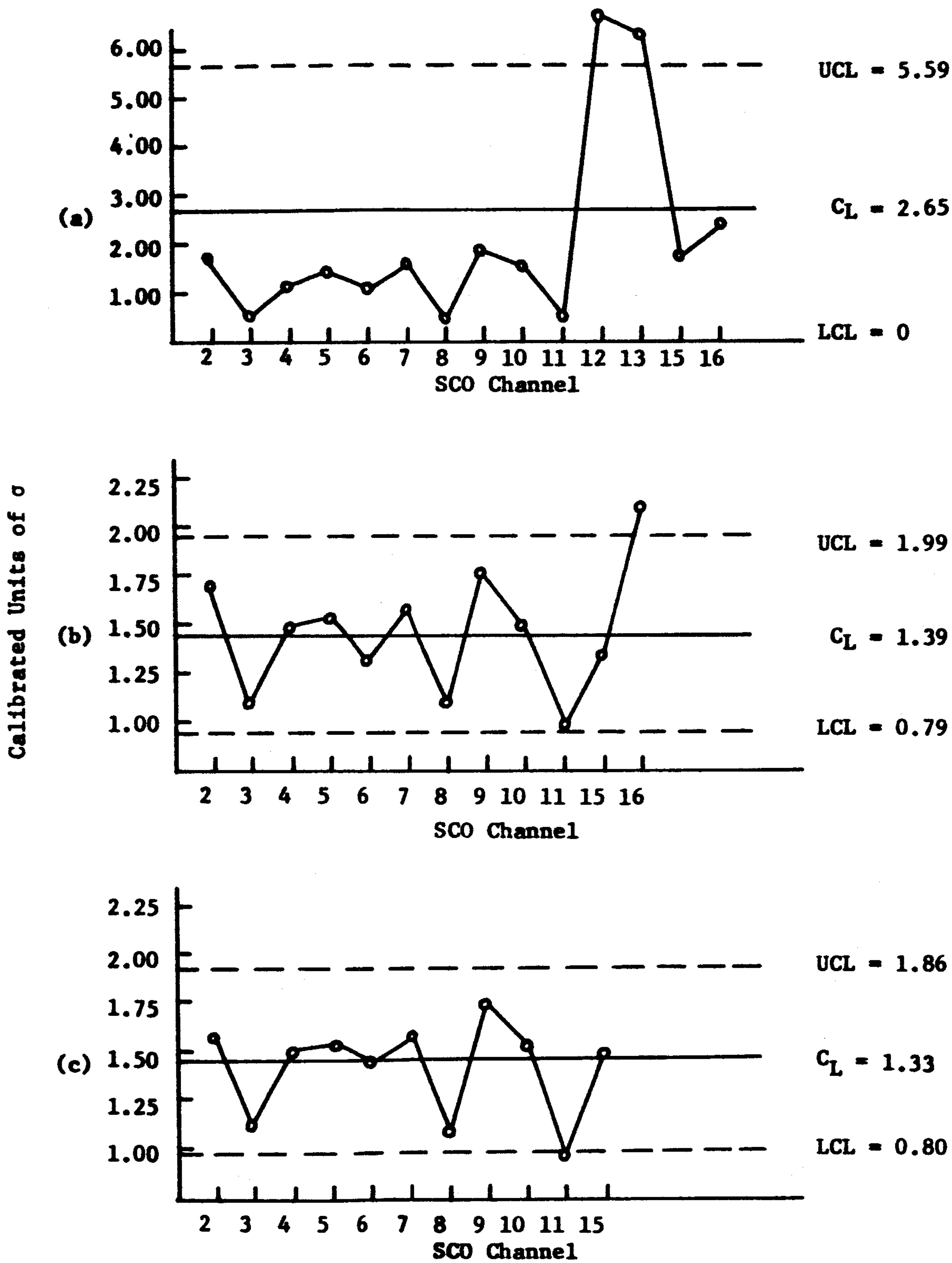

FIGURE 11. CONTROL CHARTS FOR σ, 0% INPUT
Data from Table 1, n = 5.
(a) gives original chart.
(b) and (c) are revised charts.

The $\bar{X}$-Chart is plotted in Figure 12, (a) with revisions given in Figure 12 (b), (c), and (d). The result of this analysis is an investigation of channels 2, 4, 5, and 6. Our final values for $\bar{\bar{X}}_1$ and $\sigma_{\bar{x}_1}$ are

$$\bar{\bar{X}}_1 = 66.37$$

$$\sigma_{\bar{x}_1} = 2.01$$

The control charts are applied to the data given in Table 5 in a similar manner and are shown in Figures 13, 14, 15, 16, 17, 18, 19, and 20. Values of σ_i, $\bar{\bar{X}}_i$, and $\sigma_{\bar{x}_i}$ as well as the SCO's requiring investigation are given in Table 6.

TABLE 6

VALUES OF σ'_i, $\bar{\bar{X}}_i$, AND $\sigma_{\bar{x}_i}$ AND CHANNELS REQUIRING INVESTIGATION FOR 5 INPUT LEVELS

i	σ'_i	$\bar{\bar{X}}_i$	$\sigma_{\bar{x}_i}$	SCO's Investigated
1	1.58	66.37	2.01	12, 13, 16, 2, 4, 5, 6
2	1.55	289.03	1.96	12, 13, 16, 8, 2, 4, 5, 6
3	1.43	512.20	4.88	12, 13, 4, 6
4	1.23	736.53	2.23	12, 13, 4, 5, 6
5	1.18	959.16	3.58	12, 13, 5, 6

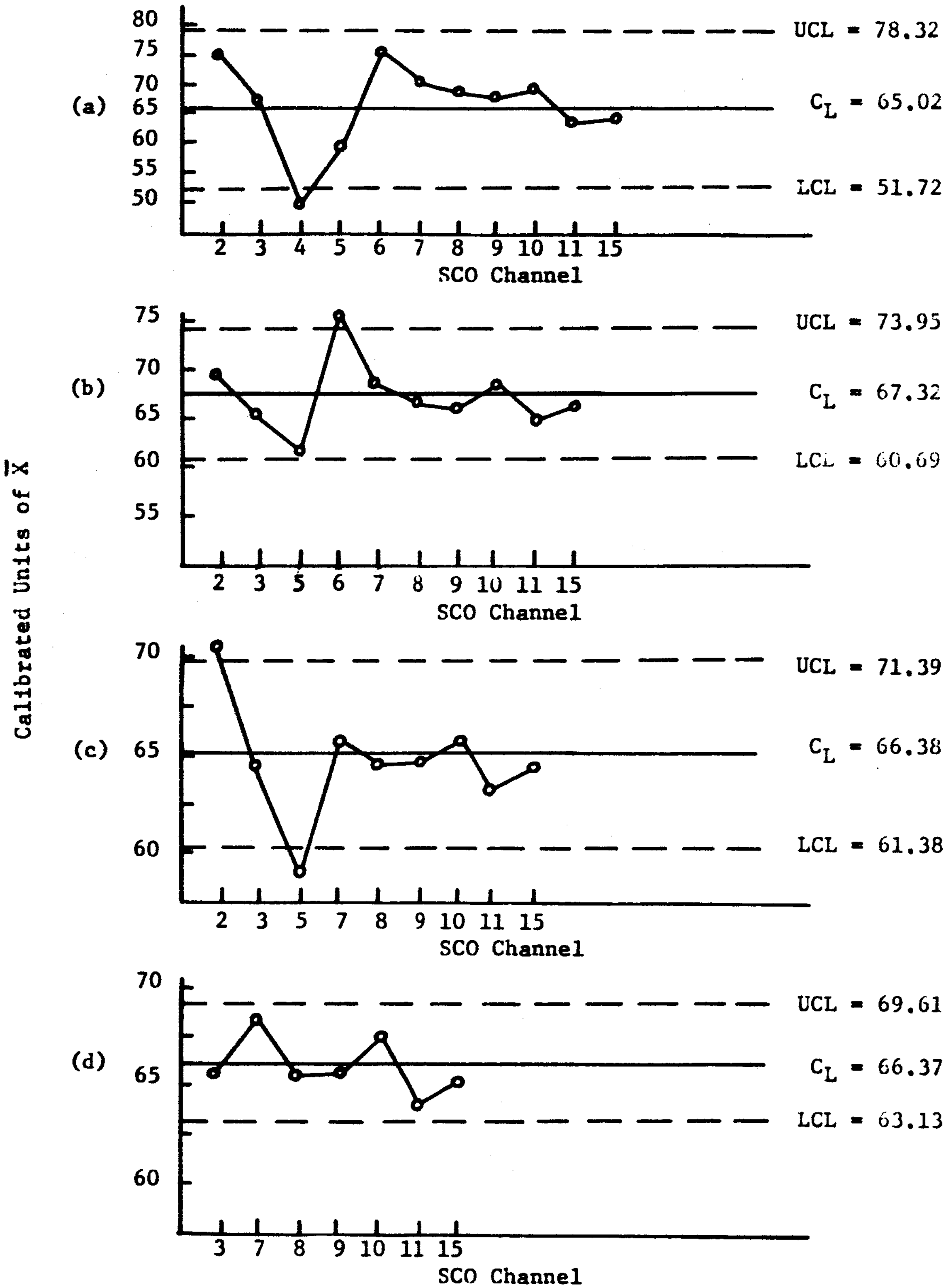

FIGURE 12. CONTROL CHARTS FOR $\overline{X}$, 0% INPUT. Data from Table 1, n = 5. (a) gives original chart. (b), (c), and (d) are revised charts.

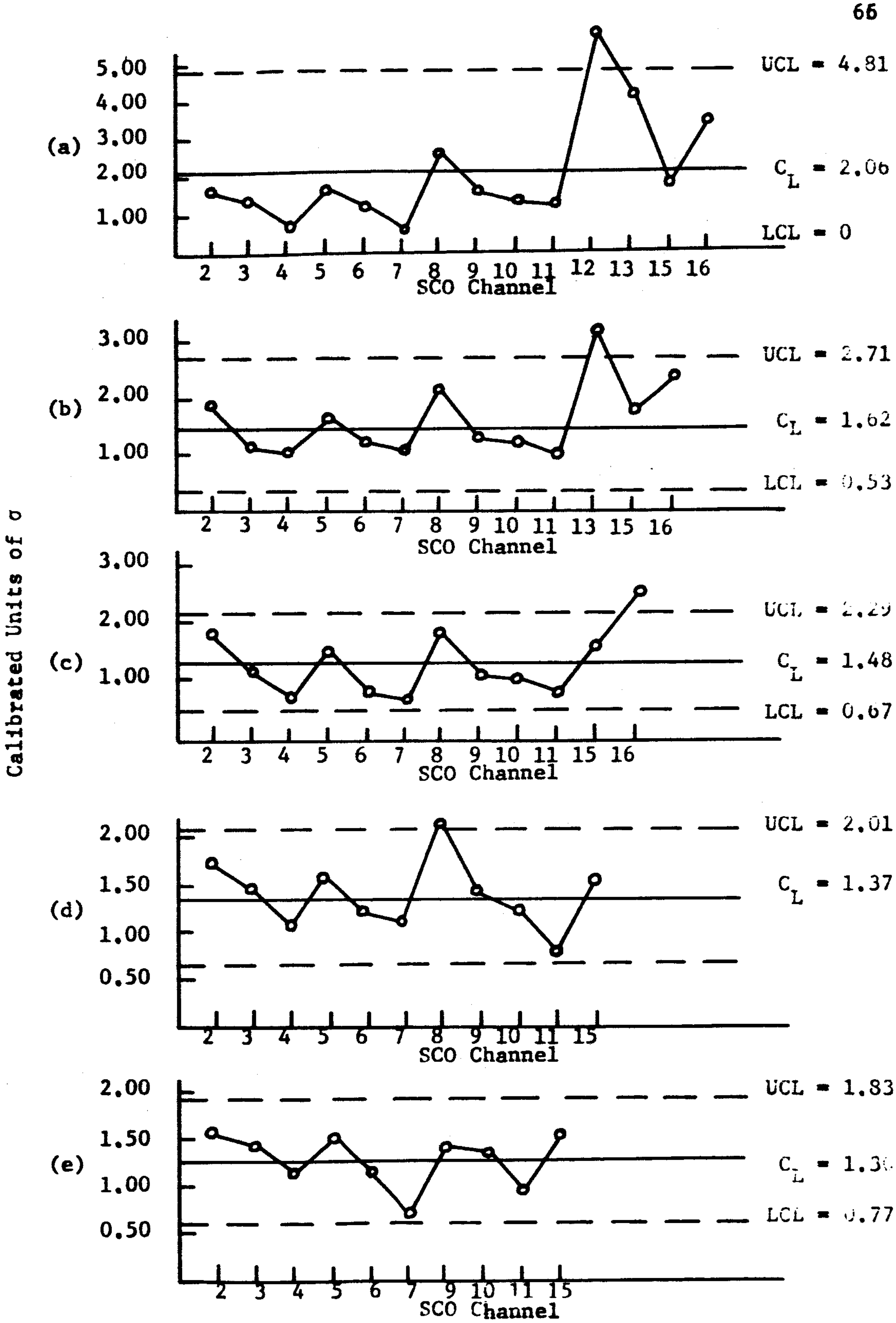

FIGURE 13. CONTROL CHARTS FOR σ, 25% INPUT. Data from Table 5. (a) gives original chart. (b), (c), (d), and (e) are revised charts.

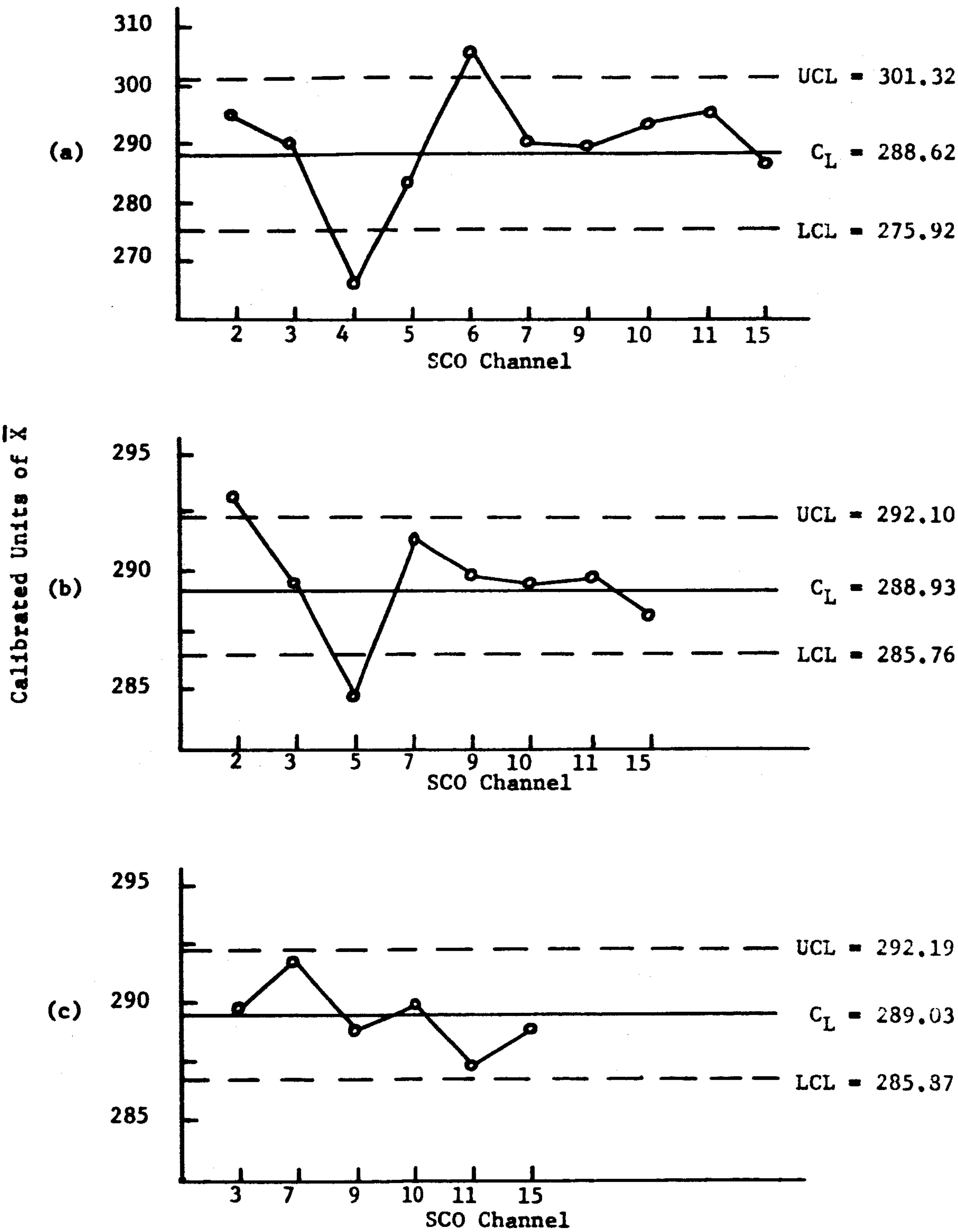

FIGURE 14. CONTROL CHARTS FOR $\overline{X}$, 25% INPUT. Data from Table 5. (a) gives original chart. (b) and (c) are revised charts.

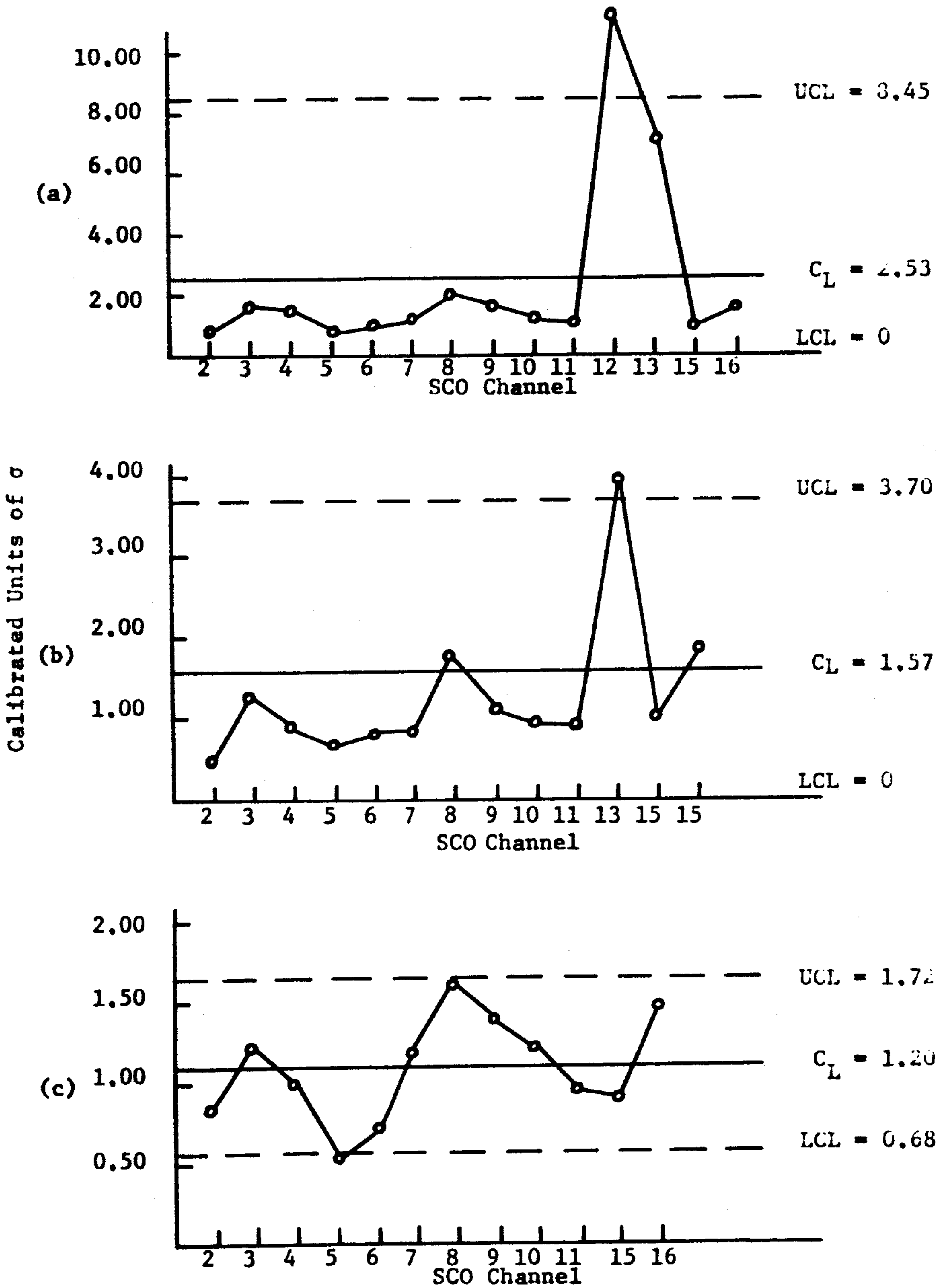

FIGURE 15. CONTROL CHARTS FOR σ, 50% INPUT. Data from Table 5. (a) gives original chart. (b) and (c) are revised charts.

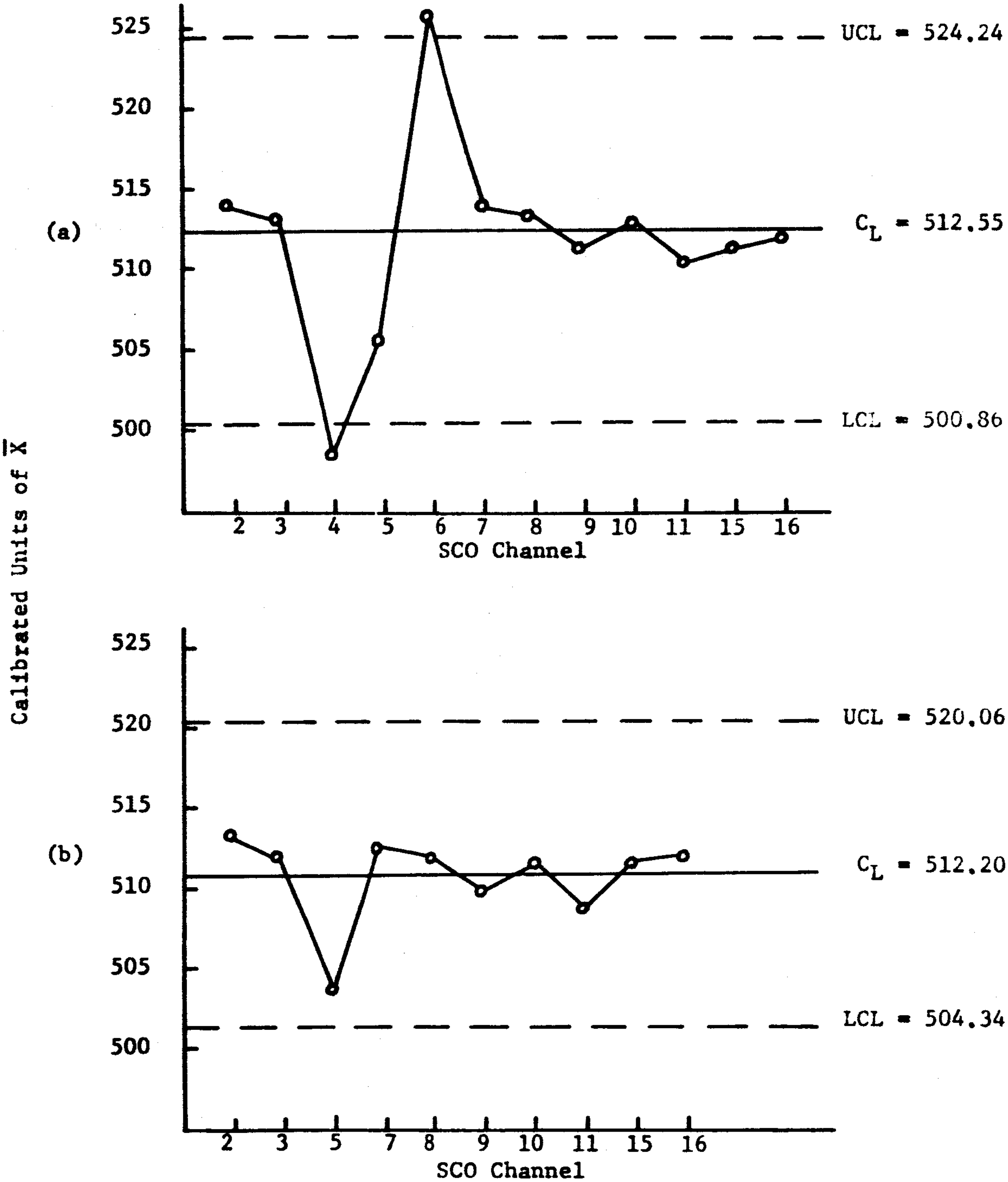

FIGURE 16. CONTROL CHARTS FOR $\overline{X}$, 50% INPUT. Data from Table 5. (a) gives original chart. (b) is the revised chart.

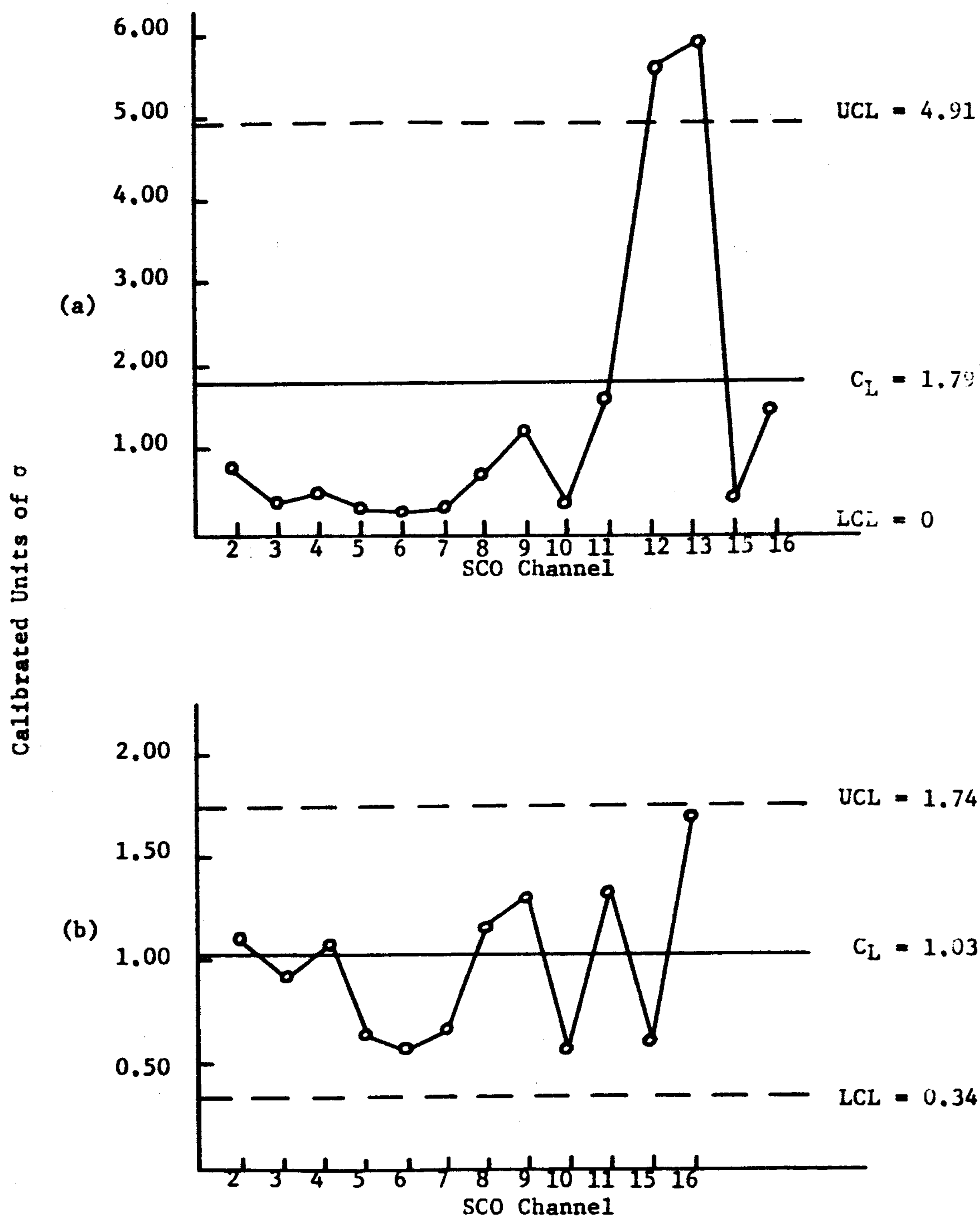

FIGURE 17. CONTROL CHARTS FOR σ, 75% INPUT. Data from Table 5. (a) gives original chart. (b) is the revised chart.

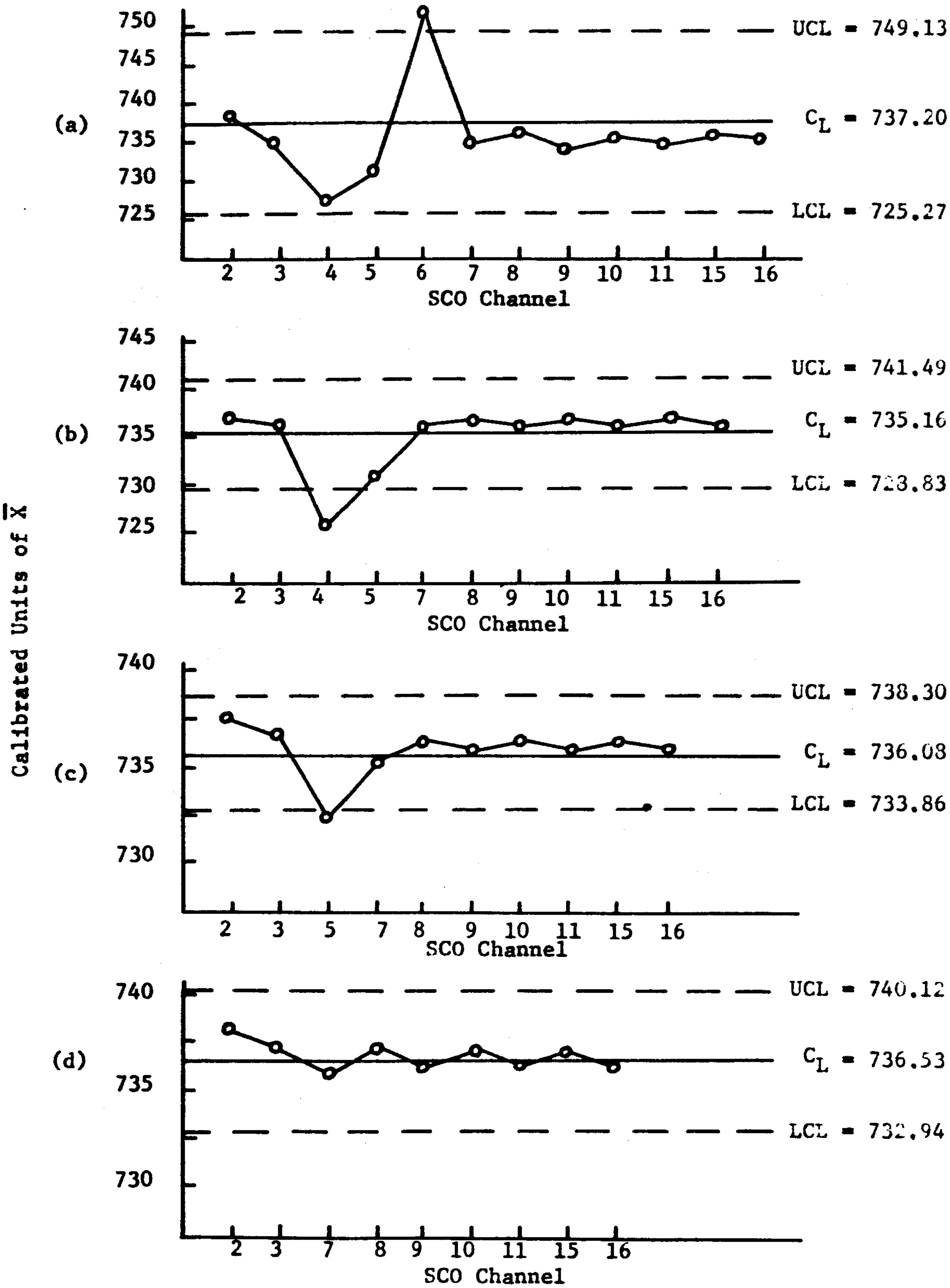

FIGURE 18. CONTROL CHARTS FOR $\overline{X}$, 75% INPUT. Data from Table 5. (a) gives original chart. (b), (c), and (d) are revised charts.

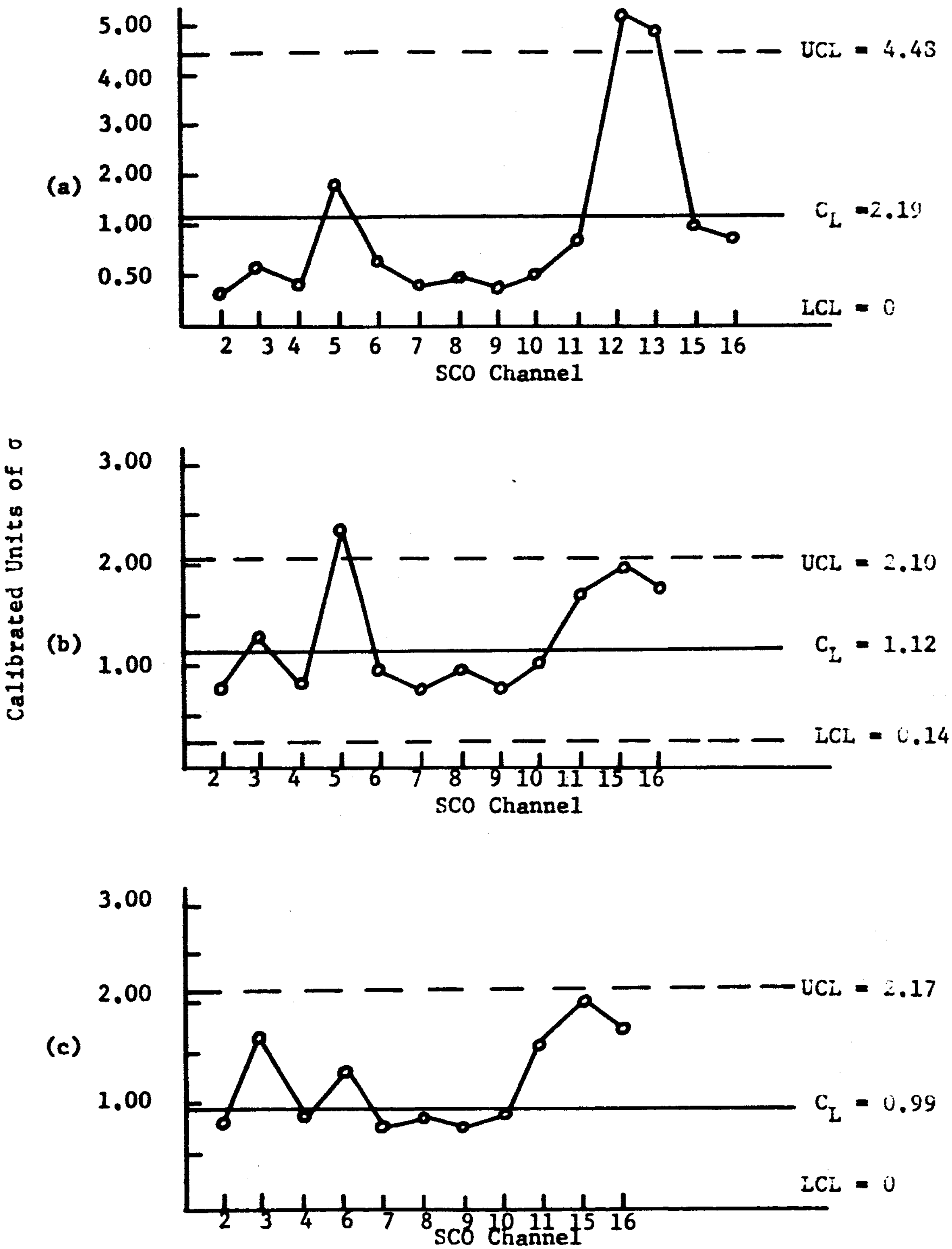

FIGURE 19. CONTROL CHARTS FOR σ, 100% INPUT. Data from Table 5. (a) gives original chart. (b) and (c) are revised charts.

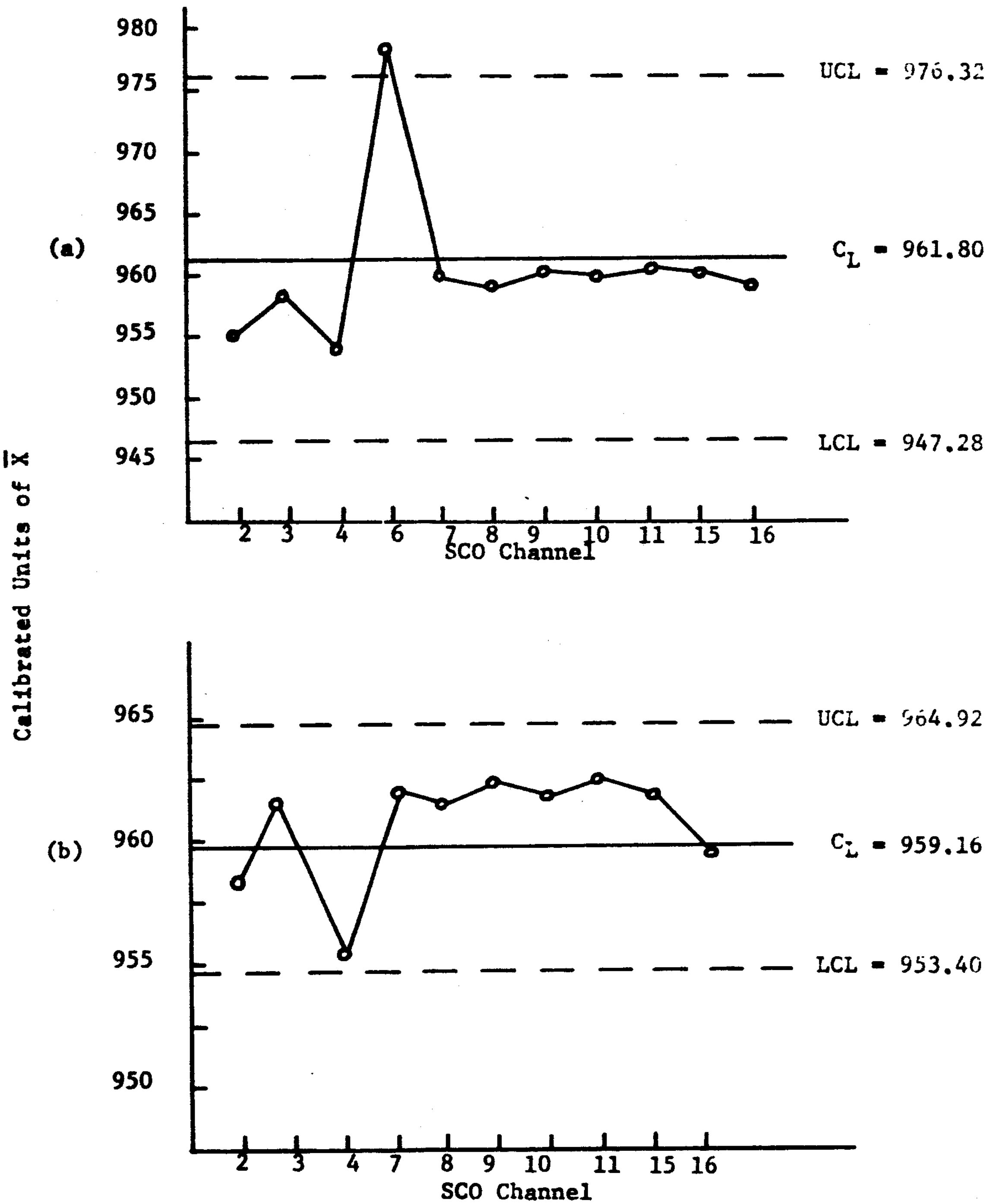

FIGURE 20. CONTROL CHARTS FOR $\overline{X}$, 100% INPUT. Data from Table 5. (a) gives original chart. (b) is the revised chart.

The best estimates for the population standard deviation at each voltage input level were given in Table 6. Therefore, we may estimate the universe standard deviation for the telemetry package as

$$\hat{\sigma}' = \frac{\sum_{i=1}^{5} \hat{\sigma}_i}{5}$$

$$= \frac{(1.58 + 1.55 + \ldots + 1.18)}{5}$$

$$= 1.39$$

Previous work has shown that a measure of telemetry system precision (random error) is given by the standard deviation expressed as a percentage of the range[9] of the process, $\sigma_p = 100\sigma/\overline{R}$ (5). Thus, we may estimate the precision of the telemetry package that we have analyzed as

$$\sigma_p = 100\hat{\sigma}'/\overline{R} \quad ,$$

where

$$\overline{R} = \overline{\overline{X}}_{max} - \overline{\overline{X}}_{min} \quad .$$

Obtaining $\overline{\overline{X}}_{max}$ and $\overline{\overline{X}}_{min}$ from Table 6,

$$\overline{R} = 959.16 - 66.37$$

$$= 892.79 \ .$$

Our estimate of the precision now becomes

$$\sigma_p = \frac{100(1.39)}{892.79}$$

$$= 0.16\%$$

[9]The range is merely the difference in the highest and lowest values.

Further work (6) has established that a measure of telemetry system accuracy (systematic errors) is given by

$$\sigma_\alpha = 100 \sqrt{(\sigma_M^2 - \sigma'^2)/mh} \Big/ \bar{\bar{X}} \quad ,$$

where σ_M^2 is a variance of the mean values about a theoretical curve fitted to the data, or

$$\sigma_M^2 = \frac{\sum_{i=1}^{h} (\bar{\bar{X}}_i - \hat{Y}_i)^2}{h - r - 1} \quad .$$

In the equation $\hat{Y}_i$ is the theoretical ordinate from the curve which best fits the $\bar{\bar{X}}_i$ points and r is the degree of this curve.

A regression analysis was performed with the aid of the University of Alabama Univac SS 80 computer and a linear equation was found to provide the best fit to the data. The correlation coefficient (degree of relationship) for the linear fit was 0.99799. Using the theoretical ordinates obtained from the linear equation the variance of the mean values about the curve was found to be

$$\sigma_M^2 = 1648.21 \quad .$$

Thus our estimate of the accuracy of the telemetry package becomes

$$\sigma_\alpha = \frac{100 \sqrt{(1648.21 - (1.39)\ /(14)(5)}}{892.79}$$

$$= \frac{100 \sqrt{1646.28/70}}{892.79}$$

$$= 0.54\% \quad .$$

Expressed as 99% confidence limits the values of accuracy and precision are

Average Precision = 0.48%

Average Accuracy = 1.62% .

This may be interpreted as meaning that we would be approximately 99% certain that the precision of the telemetry package that we have analyzed is no worse than 0.48% and that the accuracy is no worse than 1.62%.

We may now set standards for future control chart analysis of telemetry systems (based on $n = 5$, $\delta = 5$).

For the σ-Chart, at all 5 levels of input,

$$C_L = C_2\hat{\sigma}' = (0.84)(1.39) = 1.17$$

$$UCL = C_2\hat{\sigma}' + K\sigma_\sigma' = 1.17 + (1.61)(0.42) = 1.85$$

$$LCL = C_2\hat{\sigma}' - K\sigma_\sigma' = 1.17 - (1.61)(0.42) = 0.49$$

where, as implied in CHAPTER III, equation [14],

$$\sigma_\sigma' = \frac{\hat{\sigma}' \sqrt{2(n-1) - 2nC_2^2}}{\sqrt{2n}}$$

$$= \frac{1.39 \sqrt{2(4) - 2(5)(.84)^2}}{\sqrt{2(5)}}$$

$$= [(1.39)(0.95)] \,/\, 3.16$$

$$= 0.42 \;.$$

For the $\bar{X}$-Chart, at each input level separately,

0% Level: $C_L = \bar{\bar{X}}_1 = 66.37$

$$UCL = \bar{\bar{X}}_1 + K\hat{\sigma}_{\bar{x}_1} = 66.37 + (1.61)(2.01) = 69.61$$

$$LCL = \bar{\bar{X}}_1 - K\hat{\sigma}_{\bar{x}_1} = 66.37 - (1.61)(2.01) = 63.13$$

25% Level: $C_L = \bar{\bar{X}}_2 = 289.03$

$$UCL = \bar{\bar{X}}_2 + K\hat{\sigma}_{\bar{x}_2} = 289.03 + (1.61)(1.96) = 292.19$$

$$LCL = \bar{\bar{X}}_2 - K\hat{\sigma}_{\bar{x}_2} = 289.03 - (1.61)(1.96) = 285.87$$

50% Level: $C_L = \bar{\bar{X}}_3 = 512.20$

$$UCL = \bar{\bar{X}}_3 + K\hat{\sigma}_{\bar{x}_3} = 512.20 + (1.61)(4.88) = 520.06$$

$$LCL = \bar{\bar{X}}_3 - K\hat{\sigma}_{\bar{x}_3} = 512.20 - (1.61)(4.88) = 504.34$$

$$\begin{aligned}
\text{75\% Level:}\quad C_L &= \bar{\bar{X}}_4 = 736.53 \\
UCL &= \bar{\bar{X}}_4 + K\hat{\sigma}_{\bar{x}_4} = 736.53 + (1.61)(2.23) = 740.12 \\
LCL &= \bar{\bar{X}}_4 - K\hat{\sigma}_{\bar{x}_4} = 736.53 - (1.61)(2.23) = 732.94 \\
\text{100\% Level:}\quad C_L &= \bar{\bar{X}}_5 = 959.16 \\
UCL &= \bar{\bar{X}}_5 + K\hat{\sigma}_{\bar{x}_5} = 959.16 + (1.61)(3.58) = 964.92 \\
LCL &= \bar{\bar{X}}_5 - K\hat{\sigma}_{\bar{x}_5} = 959.16 - (1.61)(3.58) = 953.40
\end{aligned}$$

A final area of interest in applying the methodology will be to give the OC curves for the control charts based on the standard values that we have derived.

The OC function for the $\bar{X}$-Chart based on standard values is the same as that for past values except that the factor m in the exponent is eliminated since we are now interested in the probability of each new sample $\bar{X}$ value falling within the limits as they are computed. The function can thus be stated as

$$\beta_{\bar{x}}(K) = 1 - \left[\int_{-\infty}^{-K-\theta} f(Z)dZ + \int_{K-\theta}^{\infty} f(Z)dZ \right] .$$

Applying this function to the 0% level of input ($n = 5$, $K = 1.61$, $\hat{\sigma}_{\bar{x}} = 2.01$), the OC curve given in Figure 21 may be derived for various values of δ ($\theta = \delta/\hat{\sigma}_{\bar{x}}$).

The OC function for the σ-Chart based on standard values may be stated as

$$\beta_{\sigma}(K) = \int_{(LCL/\sigma'_{\delta})^2 n}^{(UCL/\sigma'_{\delta})^2 n} f(\chi^2)d\chi^2 ,$$

where $\sigma'_{\delta} = \hat{\sigma}' + \delta$.

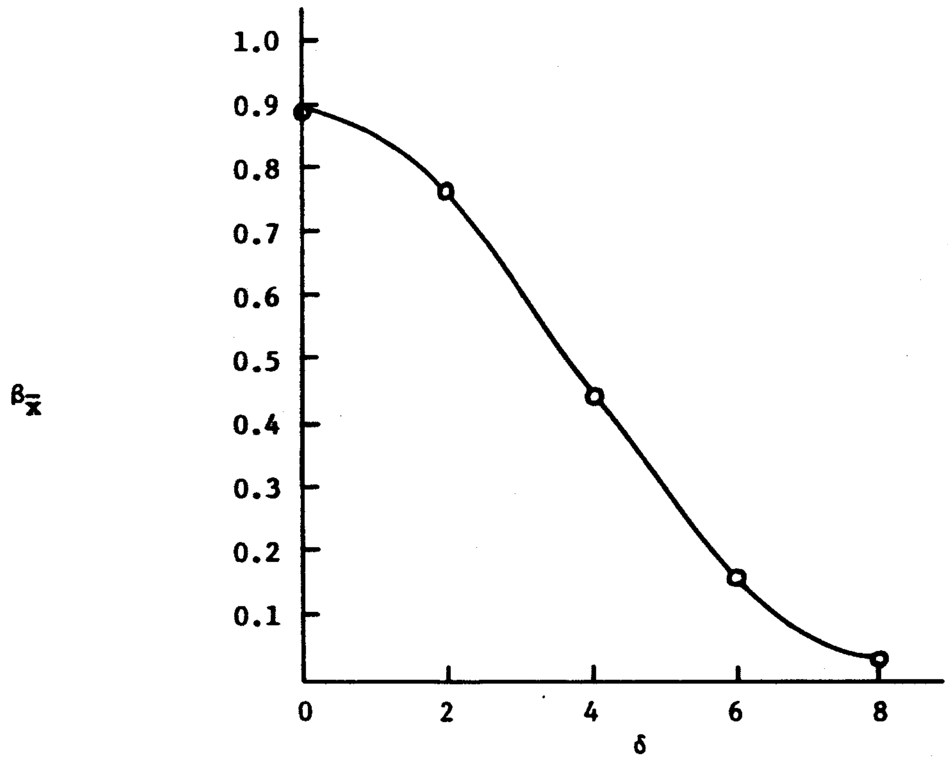

FIGURE 21. OC CURVE FOR $\bar{X}$-CHART BASED ON STANDARD VALUES, 0% INPUT.

Thus, for $\delta' = 1.39$, $\sigma'_\sigma = 0.42$, UCL = 1.85, LCL = 0.49, the OC curve given in Figure 22 may be derived for various values of δ.

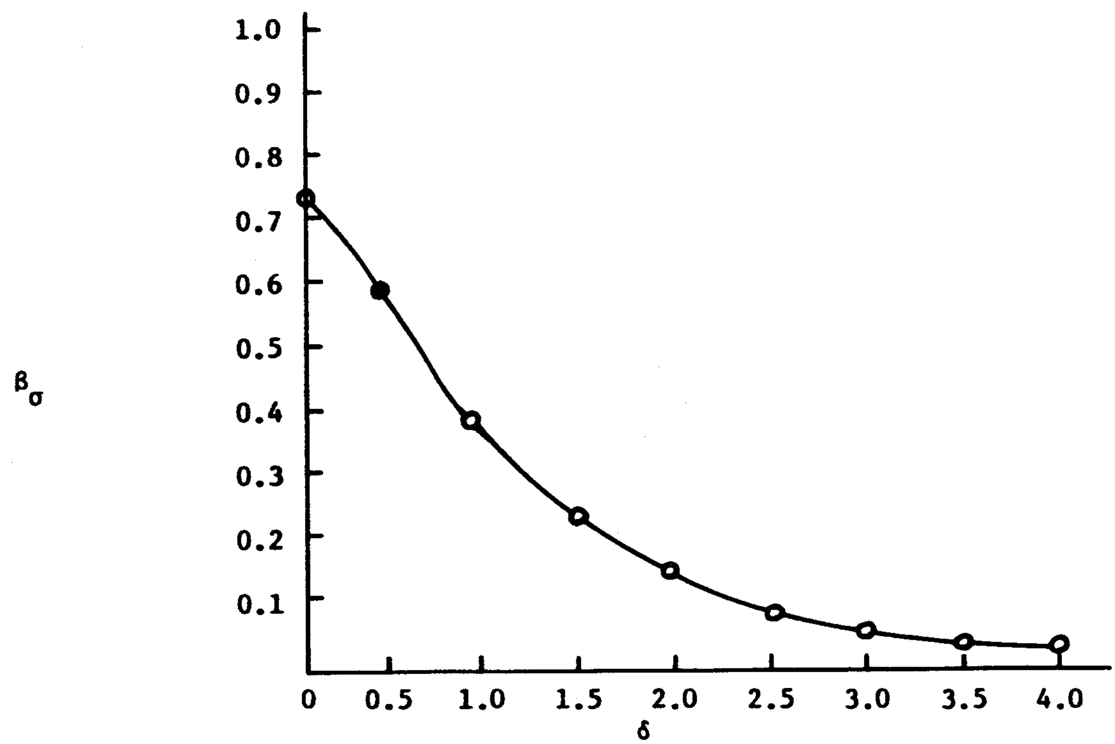

FIGURE 22. OC CURVE FOR σ-CHART BASED ON STANDARD VALUES.

CHAPTER VII

CONCLUSION

Summary

This thesis has presented a methodology for determining whether a telemetry package is in a state of statistical control. The development of the methodology has required research and experimentation in the following areas:

1. The theoretical foundation for the establishment of the $\overline{X}$ and σ charts;

2. The investigation of the operating characteristics of the $\overline{X}$ and σ charts based on past data;

3. The determination of the proper control limit constant, K, from an analysis of a control chart cost model;

4. The determination of the accuracy and precision of the analyzed telemetry system.

This research and experimentation has resulted in the establishment of a control chart methodology for analyzing past telemetry system performance and also the specification of control charts for future analysis of telemetry systems. On the basis of the control chart analysis, the accuracy and precision of telemetry systems can be estimated.

The results obtained in CHAPTER VI for the particular telemetry package that was analyzed are summarized for clarity as follows:

1. Combined use of the $\overline{X}$ and σ charts (considering all five levels of input voltage) resulted in the investigation of SCO channels 2, 4, 5, 6, 8, 12, 13, and 16. Of these eight SCO's, only channels 4, 6, 12, and 13 were deliberately caused to malfunction. The remaining channels were either detected incorrectly (due to the alpha error) or were in need of investigation due to assignable causes not immediately known.

2. Estimates of the system accuracy and precision, expressed as 99% confidence limits, were found to be

   Average Precision = 0.48%

   Average Accuracy = 1.62% .

3. Standards for future control chart analysis were established at all 5 levels of input. Selecting only the 0% level of input as a basis the standards are

σ-Chart:	UCL	= 1.85
	C_L	= 1.17
	LCL	= 0.49
$\overline{X}$-Chart:	UCL	= 69.61
	C_L	= 66.37
	LCL	= 63.13

4. The OC curves for both the $\overline{X}$ and σ charts based on standard values were given in Figures 21 and 22. The OC curves for

both charts proved that for shifts of 5 or more in the universe parameter ($\bar{\bar{X}}'$ or σ') the probability of not detecting this shift, β, is rather small. These curves and both control charts were based on a K factor of 1.61 which was discovered to be the optimum K.

Sources of Possible Error

Perhaps the most important assumption in the development of the methodology is that the individual values of output are normally distributed. This assumption has been made in other reports (5,6) although the data was often found to be only moderately normal. The more non-normal the telemetry data actually is, the more will be the error in utilizing the control chart methodology, or for that matter, any other type of parametric[10] statistical test.

Another source of possible error is in the assumed values of c_1, c_2, and c_3. We have assumed that the cost of the beta error is ten times as much as the cost of the alpha error, and the cost of the alpha error is 1000 times as much as the cost of control chart sampling. Since in the telemetry environment the cost of letting defective components pass inspection is much greater than is the cost of investigating satisfactory components, these cost estimates appear to be roughly accurate in their relationship to each other. However, if they are grossly inaccurate, the selection of K, the selection of

[10]Non-parametric tests are not affected by the assumption of normality. See, for example, (7).

the optimum sample size n, and consequently the development of the control chart limits must be made based on the new estimates. If these old cost estimates are used (resulting in $K = 1.61$, $n = 5$, etc.), and new estimates actually should be made, then some amount of error will be entered into the analysis. Furthermore, we have selected a shift of $\delta = 5$ as that which we wish to detect. If this assumption is not valid, our analysis must be made based on some other value of δ, and thus a possible source of error might be eliminated.

A final error source is in the assumption that the K factor selected for the $\bar{X}$-Chart will be satisfactory for the σ-Chart. Shifts in $\bar{\bar{X}}'$ of $\delta = 5$ seem to be fairly realistic as a governing factor for the determination of K for the $\bar{X}$-Chart. However, we may wish to make our analysis more precise by basing the determination of the control limits for the σ-Chart on some other value of K, or perhaps on a K determined for both $\bar{\bar{X}}'$ and σ' shifting simultaneously (see p. 83). Therefore, the selection of one value of K for both charts based on an analysis of the errors associated with the $\bar{X}$-Chart may have introduced some amount of error.

Recommendations for Future Research and Application

The researcher feels that the methodology presented in this thesis could easily be used to form the basis for a complete telemetry package checkout procedure. Once the necessary experimental equipment was set up, the actual implementation of the methodology would become a simple matter. The use of the SEL data reduction equipment in the telemetry ground station further enhances the possibility of this type of analysis. If the SEL equipment could be programmed to give means and standard deviations for samples of size n = 5, 10, 15, etc., these values could be plotted quickly on control charts whose limits were already set and components which were not in statistical control could be adjusted rapidly or could be immediately replaced.

This thesis has also suggested a major area for future statistical research. As was mentioned previously, it may be desirable to base the determination of the optimum value of K on the assumption that both $\bar{\bar{X}}'$ and σ' may shift simultaneously. This would require the investigation of the OC function for the control charts when both parameters are subject to simultaneous shifting. The development of the OC function for this situation has been accomplished when the charts are to be based on standard values for future production (2, 4). However, as far as this researcher knows, no investigation has been made of this function when the charts are based on past data. The function would take the form of a surface and would seem to be extremely complicated, yet this would no doubt be a major contribution to the field of mathematical and applied statistics.

Another area for statistical research would be in a sensitivity analysis of the control chart methodology to the assumption of normality. A significant contribution could be made if a non-parametric method could be found for analyzing product variability based on distributions whose form is unknown.

A final possible area for research and application in the telemetry environment is in the use of the control chart methodology during an actual flight calibration period. As was suggested in CHAPTER I, the analysis by statistical methods of in-flight calibration data would be a considerable aid in determining whether the telemetry package was performing satisfactorily in this phase. If a particular SCO was found to be out of control, adjustments could possibly be made for this malfunctioning component and rather than completely losing the data from this particular channel, only a short time period would be required for correcting the component.

APPENDICES

APPENDIX A

MISCELLANEOUS PROOFS, THEOREMS, AND SAMPLE CALCULATIONS

MISCELLANEOUS PROOFS, THEOREMS, AND SAMPLE CALCULATIONS

1. Elementary theorems of expectation.

$$E(c) = C$$

$$E(cx) = cE(x)$$

$$E(x) = \bar{X}$$

$$E(x \pm y) = E(x) \pm E(y)$$

$$E(\Sigma x) = \Sigma E(x)$$

$$\Sigma c = nc,$$

where c is a constant and x and y are variables.

2. Proof that $\displaystyle\int_{\bar{\bar{X}}-3\sigma_{\bar{X}}}^{\bar{\bar{X}}+3\sigma_{\bar{X}}} \frac{1}{\sigma_{\bar{X}}\sqrt{2\pi}}\, e^{-1/2[(\bar{X}-\bar{\bar{X}})/\sigma_{\bar{X}}]^2}\, d\bar{X} = 0.9973$

Let $Z = (\bar{X} - \bar{\bar{X}})/\sigma_{\bar{X}}$, then $dZ = d\bar{X}/\sigma_{\bar{X}}$ and

$$\int_{\bar{\bar{X}}-3\sigma_{\bar{X}}}^{\bar{\bar{X}}+3\sigma_{\bar{X}}} \frac{1}{\sigma_{\bar{X}}\sqrt{2\pi}}\, e^{-1/2[(\bar{X}-\bar{\bar{X}})/\sigma_{\bar{X}}]^2}\, d\bar{X} = 1/\sqrt{2\pi}\int_{-3}^{3} e^{-Z^2/2}\, dZ$$

or,

$$1/\sqrt{2\pi}\int_{-3}^{3} e^{-Z^2/2}\, dZ = 2/\sqrt{2\pi}\int_{0}^{3} e^{-Z^2/2}\, dZ.$$

The value of this last integral multiplied by the factor $1/\sqrt{2\pi}$ may be found in most statistical textbooks, for instance (9), to be

0.49865. Therefore, the value of the given integral is 2(0.49865) = 0.9973.

The function $f(Z) = (1/\sqrt{2\pi})e^{-Z^2/2}$ is known as the standardized normal equation and has a mean of 0 and a variance of 1.

3. Definition of the gamma function, $\Gamma[(n-1)/2]$.

By definition $\Gamma(K) = \int_0^\infty V^k e^{-v} dV$.

Let $U = v^k$, $dv = e^{-v} dv$, $dU = kv^{k-1} dv$, $V = -e^{-v}$ and

$$\Gamma(K + 1) = -v^k e^{-v} \Big|_0^\infty - \int_0^\infty kv^{k-1} (-e^{-v} dv)$$

$$= 0 - 0 + K\Gamma(K) ,$$

by the method of integration by parts.

Thus, $\Gamma(K + 1) = K\Gamma(K)$ $(K>0)$.

When we have a positive integer m for the argument of the function, repeated use of the above equation yields

$$\Gamma(m) = (m - 1)! \ \Gamma(1).$$

When m is not an integer, such as $\Gamma[(n-1)/2]$, we must use logarithms and tables of factorials.

Thus, $\Gamma(5.7) = (4.7)(3.7)(2.7)(1.7)\Gamma(1.7)$, and then tables are used for the last factor.

4. Development of method for evaluating

$$\beta_\sigma(K) = \frac{1}{2^{v/2}\,\Gamma(v/2)} \int_{X_1}^{X_2} (\chi^2)^{v/2 -1} e^{-\chi^2/2} d\chi^2 ,$$

where $X_1 = (LCL/\sigma_{\hat{\delta}})^2 n$, and

$X_2 = (UCL/\sigma_{\hat{\delta}})^2 n$.

Let $\chi^2 = y$ to simplify the expression.

$$\beta_\sigma(K) = \frac{1}{2^{v/2}\,\Gamma(v/2)}\left[\int_0^{X_2} y^{v/2-1} e^{y/2}\,dy - \int_0^{X_1} y^{v/2-1} e^{-y/2}\,dy\right].$$

$$\int_0^{X_1} Z^{v/2-1} e^{-y/2}\,dy = \int_0^{X_1} (y^{v/2-1}) \left[1 - \frac{y}{2} + \frac{y^2}{4} - \frac{y^3}{12} + \ldots + \frac{y^n}{2\cdot n!}\right] dy$$

$$= \int_0^{X_1}\left[y^{v/2-1} - \frac{y^{v-2}}{2} + \frac{y^{v-2}}{4} - \frac{y^{3v/2-3}}{12} + \ldots + \frac{y^{nv/2-n}}{2\cdot n!}\right] dy$$

$$= \sum_{k=1}^{\infty} \int_0^{X_1} (-1)^{k+1} \frac{y^{kv/2-k}}{2\cdot k!}$$

$$= \sum_{k=1}^{\infty} (-1)^{k+1} \left[\frac{(X_1)^{kv/2-k+1}}{2\cdot k!\,[k(v/2-1)+1]}\right].$$

A similar result is obtained for

$$\int_0^{X_2} Z^{v/2-1} e^{-y/2}\,dy$$

Thus,

$$\beta_\sigma(K) = \frac{1}{2^{v/2}\,\Gamma(v/2)} \sum_{k=1}^{\infty} (-1)^{k+1} \left[\frac{(X_2)^{k(v/2-1)+1} - (X_1)^{k(v/2-1)+1}}{[K(v/2-1)+1]\ 2\cdot K!} \right].$$

5. Method of Differentiation of integrals (12, 14).

The fundamental theorem of integral calculus states that whenever f(x) is a continuous function in the closed interval (a, b) and F(x) is a function such that $F'(x) = f(x)$, then

$$\int_{u_0}^{u_1} f(k)\, dX = F(u_1) - F(u_0)$$

for any two points u_0 and u_1 in the interval. If u_0 and u_1 are differentiable function of another variable, y, so that

$$u_0 = u_0(y), \quad u_1 = u_1(y),$$

the right hand member in the above integral is a function of y and the chain rule of differentiation gives

$$dF(u_1)/dy = F'(u_1)(du_1/dy) = f(u_1)(du_1/dy) \ .$$

Since a similar result holds for $F(u_0)$, differentiation of The original integral yields

$$d/d\alpha \int_{u_0(y)}^{u_1(y)} f(x)dX = f(u_1)(du_1/dy) - f(u_0)(du_0/dy) \ .$$

6. Sample calculations for finding optimum K from equation [23] by trial and error.

Assume $n = 5$, $m = 14$, $c_1 = 10$, $c_2 = 1$, $\sigma_{\bar{x}} = 7.45$, $\theta = 0.67$.

<u>Try K = 1.50:</u>

$$\frac{-13(0.83)^2}{2} + (13)\ \log_e(0.83 + 2.17e^{-4.02}) + 0.50 + \log_e(1 + e^{-4.02})$$

$$= 0.23 + \log_e(2/\sqrt{2\pi}) - \log_e(280/\sqrt{2\pi})\ .$$

Solving the left hand side of the equation we obtain -4.97. Solving the right hand side we obtain -4.76. We shall try for closer agreement.

<u>Try K = 1.64:</u>

$$\frac{-13(0.97)^2}{2} + (13)\log_e(0.97 + 2.31e^{-4.40}) + 0.55 + \log_e(1 + e^{-4.40})$$

$$= -4.76\ .$$

Solving the left hand side we obtain -4.56 .

<u>Try K = 1.61:</u>

$$\frac{-13(0.94)^2}{2} + (13)\log_e(0.94 + 2.28e^{-4.32}) + 0.54 + \log_e(1 + e^{-4.32})$$

$$= -4.76\ .$$

Solving the left hand side we obtain -4.78 . We, therefore, accept a value of K = 1.61.

7. Proof that value of K = 1.61 gives minimum cost by substitution into second derivative equation [24].

Assume K = 1.61, $\theta = 0.67$, m = 14, $c_1 = 10$, $c_2 = 1$.

$d^2(TC)/dK^2$ defined by equation [24].

$d^2(TC)/dK^2$ =

$$111.55\left\{13\ [0.94e^{-(0.94)^2/2} + 2.28e^{-(-2.28)^2/2}]^{12}\ \cdot\right.$$

$$[-0.67e^{-(0.94)^2/2} - 0.88e^{-(0.94)^2/2} + 0.67e^{-(-2.28)^2/2}$$

$$+ (2.28)\, e^{-(-2.28)^2/2}] \cdot [e^{-(0.94)^2/2} + e^{-(-2.28)^2/2}] +$$

$$[0.94e^{-(0.94)^2/2} + 2.28e^{-(-2.28)^2/2}]^{12} \cdot [-(0.94)e^{-(0.94)^2/2}$$

$$+ 2.28e^{-(-2.28)^2/2}] \Big\} + 1.27e^{-(0.94)^2/2} \quad .$$

This reduces to:

$$d^2(TC)/dK^2 = 111.55[13(0.798)^{12}(-0.498)(0.694)(0.798)^{13}(-0.520)]$$
$$+ 0.347 \quad .$$

As can easily be seen, this results in a positive number, and thus a minimum cost is obtained.

APPENDIX B

GLOSSARY OF SYMBOLS

GLOSSARY OF SYMBOLS

α	Probability of rejecting a true hypotheses
β	Probability of accepting a false hypothesis
c_1	Unit cost of the β error
c_2	Unit cost of the α error
c_3	Unit cost of control chart sampling which is directly dependent on the sample size.
C_2	Constant defined by $\sqrt{2/n}\ \dfrac{\Gamma(n/2)}{\Gamma(n-1)/2}$
Γ	Gamma function
C_L	Control chart center line
δ	Parameter reflecting amount of shift in some universe parameter
θ	Parameter reflecting amount of shift in some universe parameter in terms of the standard deviation of the universe parameter
h	Number of voltage input levels
i	Input level subscript ($i = 1, 2, \ldots, h$)
m	Number of subcarrier oscillators
j	Subcarrier oscillator subscript ($j = 1, 2, \ldots, m$)
n	Number of individual values at any particular SCO and input level
k	Individual value subscript ($k = 1, 2, \ldots, n$)
K	Control chart limit constant
OC	Operating Characteristic

SCO	Subcarrier oscillator
σ_{ij}	Sample standard deviation
σ'	Universe standard deviation
$\hat{\sigma}'$	Unbiased estimate of universe standard deviation
$\hat{\sigma}_i$	Unbiased estimate of standard deviation for each input level, i
$\bar{\sigma}$	Average standard deviation
σ_σ	Standard deviation of the distribution of standard deviations
$\hat{\sigma}_{\bar{x}}$	Unbiased estimate of the standard deviation among the mean values
$\bar{X}_i'$	A population mean for each input level, i
$\bar{\bar{X}}'$	A universe mean
$\bar{X}_{ij}$	A sample mean
$\bar{\bar{X}}$	An average of the sample means
X_{ijk}	Individual value
$\bar{\bar{X}}_\delta'$	Shifted universe mean defined by $\bar{\bar{X}}_\delta' = \bar{\bar{X}}' + \delta$
σ_δ'	Shifted universe standard deviation defined by $\sigma_\delta' = \sigma' + \delta$
σ_p	Precision of the telemetry package
σ_a	Accuracy of the telemetry package
σ_M^2	Variance of the mean values about a theoretical curve fitted to the data
$\bar{R}$	Average range defined by $\bar{\bar{X}}_{max} - \bar{\bar{X}}_{min}$
v	Degrees of freedom
UCL	Upper control chart limit
LCL	Lower control chart limit
χ^2	Chi square
Z	Standard normal deviate

LIST OF REFERENCES

LIST OF REFERENCES

(1) Irving W. Burr, Engineering Statistics and Quality Control (New York: McGraw-Hill Book Company, Inc., 1953).

(2) Dudley J. Cowden, Statistical Methods in Quality Control (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1957).

(3) Acheson J. Duncan, "The Economic Design of $\bar{X}$-Charts Used to Maintain Current Control of a Process," American Statistical Association Journal, June, 1956, 228-241.

(4) Acheson J. Duncan, Quality Control and Industrial Statistics 3rd edition (Homewood, Illinois: Richard D. Irwin, Inc., 1965).

(5) Marvin A. Griffin and Richard S. Simpson (Project Directors), Accuracy Analysis of FM/FM Telemetry System for the Saturn Vehicle, Bureau of Engineering Research, University of Alabama, University, Alabama, October, 1963.

(6) Marvin A. Griffin and Richard S. Simpson (Project Directors), Linearity Analysis of the FM/FM Telemetry System, Bureau of Engineering Research, University of Alabama, University, Alabama, August, 1965.

(7) Marvin A. Griffin and Richard S. Simpson (Project Directors), A Statistical Analysis of the Noise and Linearity Characteristics of DCS GFD-5, EMR 210, and EMR 67-D Discriminators, Bureau of Engineering Research, University of Alabama, University, Alabama, July, 1965.

(8) Eugene L. Grant, Statistical Quality Control 3rd edition (New York: McGraw-Hill Book Company, Inc., 1964).

(9) Paul G. Hoel, Introduction to Mathematical Statistics 3rd edition (New York: John Wiley and Sons, Inc., 1962).

(10) Edgar F. King, "The Operating Characteristic of the Average Chart," Industrial Quality Control, November, 1952, 30-32.

(11) Edgar P. King, "The Operating Characteristic of the Control Chart for Sample Means," Annals of Mathematical Statistics. Vol. 23, 1952, 384-395.

(12) Maurice Sasieni, Arthur Yaspan, and Lawrence Friedman, Operation Research---Methods and Problems (New York; John Wiley and Sons, Inc., 1959).

(13) W. A. Shewhart, Economic Control of Quality of Manufactured Product (Princeton, New Jersey: D. Van Nostrand Company, Inc., 1931).

(14) I. S. Sokolnikoff and R. M. Redheffer, Mathematics of Physics and Modern Engineering (New York: McGraw-Hill Book Company, Inc., 1958).

(15) Harry L. Stiltz (Editor), Aerospace Telemetry (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963).

BIBLIOGRAPHY

BIBLIOGRAPHY

Books and Reports

Bowker, Albert H. and Lieberman, Gerald J. Engineering Statistics. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1959.

Burr, Irving W. Engineering Statistics and Quality Control. New York: McGraw-Hill Book Company, Inc., 1953.

Cowden, Dudley J. Statistical Methods in Quality Control. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1957.

Duncan, Acheson J. Quality Control and Industrial Statistics. 3rd edition. New York: John Wiley and Sons, Inc., 1962.

Goodman, A. W. Analytic Geometry and the Calculus. New York: The Macmillan Company, 1963.

Grant, Eugene L. Statistical Quality Control. 3rd edition. New York: McGraw-Hill Book Company, Inc., 1964.

Griffin, Marvin A. and Simpson, Richard S. (Project Directors), Accuracy Analysis of FM/FM Telemetry System for the Saturn Vehicle, Bureau of Engineering Research, University of Alabama, University, Alabama, October, 1963.

_______. Linearity Analysis of the FM/FM Telemetry System, Bureau of Engineering Research, University of Alabama, University, Alabama, August, 1965.

_______. (Project Directors), A Statistical Analysis of the Noise and Linearity Characteristics of DCS GFD-5, EMR 210, and EMR 67-D Discriminators, Bureau of Engineering Research, University of Alabama, University, Alabama, July, 1965.

Hoel, Paul G. Introduction to Mathematical Statistics. 3rd edition. New York: John Wiley and Sons, Inc., 1962.

Juran, J. M. (Editor) Quality Control Handbook. 2nd edition. New York: McGraw-Hill Book Company, Inc., 1962.

Milne, William Edmund. Numerical Calculus, Princeton, New Jersey: Princeton University Press, 1949.

Owen, Donald B. Handbook of Statistical Tables. Reading, Massachusetts: Addison-Wesley Publishing Company, Inc., 1962.

Sasieni, Maurice, Yaspan, Arthur, and Friedman, Lawrence. Operation Research---Methods and Problems. New York: John Wiley and Sons, Inc., 1959.

Shewhart, W. A. Economic Control of Quality of Manufactured Product. Princeton, New Jersey: D. Van Nostrand Company, Inc., 1931.

Sokolnikoff, L. S. and Redheffer, R. M. Mathematics of Physics and Modern Engineering. New York: McGraw-Hill Book Company, Inc., 1958.

Stiltz, Harry L. (Editor) Aerospace Telemetry. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

Wine, Lowell R. Statistics for Scientists and Engineers. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964.

Articles

Craig, C. C. "Control Charts Versus Analysis of Variance in Process Control by Variables," Industrial Quality Control, January, 1947, 14-16.

Duncan, Acheson J. "The Economic Design of $\overline{X}$-Charts Used to Maintain Current Control of a Process," American Statistical Association Journal, June, 1956, 228-241.

King, Edgar P. "The Operating Characteristic of the Control Chart for Sample Means," Annals of Mathematical Statistics, Vol. 23, 1952, 384-395.

____. "The Operating Characteristic of the Average Chart," Industrial Quality Control, November, 1952, 30-32.

____. "Probability Limits for the Average Chart," Industrial Quality Control, May, 1954, 10-12.

Olds, Edwin G. "Power Characteristics of Control Charts," Industrial Quality Control, July, 1961, 4-9.

Scheffe, Henry. "Operating Characteristics of Average and Range Charts," Industrial Quality Control, May, 1949, 13-18.

Storm, Leo E. "Specification Study for Limited Production of Complex Equipment," Industrial Quality Control, February, 1960, 12-18.

Walsh, John E. "Operating Characteristics for Tests of the Stability of a Normal Population," American Statistical Association Journal, June, 1952, 191-202.

Weiler, H. "On the Most Economical Sample Size for Controlling the Mean of a Population," Annals of Mathematical Statistics, Vol. 23, 1952, 247-254.