

A COMPARISON OF ALTERNATIVE METHODS TO THE SHEWHART-TYPE
CONTROL CHART

by

Deborah A. Hall

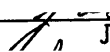
Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

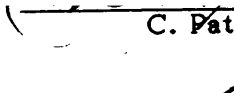
MASTER OF SCIENCE

in

Industrial Engineering and Operations Research

APPROVED:

 _____
Joel A. Nachlas, Chairman

 _____
C. Patrick Koelling

Marion R. Reynolds

January, 1989
Blacksburg, Virginia

6/7/80 W.A.

A COMPARISON OF ALTERNATIVE METHODS TO THE SHEWHART-TYPE
CONTROL CHART

by

Deborah A. Hall

Joel A. Nachlas, Chairman

Industrial Engineering and Operations Research

(ABSTRACT)

A control chart that simultaneously tracks the mean and variance of a normally distributed variable with no compensation effect is defined in this work. This joint control chart is compared to five other charts: an \bar{X} chart, an s^2 chart, a Reynolds and Ghosh chart, a Repko process capability plot, and a t-statistic chart. The criterion for comparison is the probability of a Type II sampling error. Several out-of-control cases are examined. In the case of Repko, an equation is defined to compute the Type II error probability. The results indicate that the Reynolds and Ghosh statistic is powerful for cases when the variance shifts out of control. The \bar{X} chart is powerful when the mean shifts with moderate changes in the variance. The joint chart is powerful for moderate changes in the mean and variance.

ACKNOWLEDGEMENTS

I would like to thank several people who have helped me in my work. Dr. Joel A. Nachlas' time and guidance were invaluable in completing this thesis. Without his influence, I would have never developed an interest in quality control and reliability. I also would like to thank Dr. Koelling and Dr. Reynolds for their time and advice. A final thanks to my family and friends for their support.

TABLE OF CONTENTS

CHAPTER ONE INTRODUCTION.....	1
CHAPTER TWO LITERATURE REVIEW.....	6
2.1 Introduction.....	6
2.2 Repko.....	7
2.3 Reynolds and Ghosh.....	10
2.4 Rao.....	12
CHAPTER THREE THE CONTROL CHARTS.....	14
3.1 Introduction.....	14
3.2 The Joint Control Chart.....	15
3.3 The T-Statistic Control Chart.....	25
3.4 Type II Error Probability Expressions.....	28
3.5 Repko's Plot.....	29
CHAPTER FOUR COMPARISON OF CONTROL CHART.....	33
4.1 Average Run Length.....	33
4.2 Type I Error Probability Vs. Type II Error Probability..	33
4.3 Computer Program.....	36
CHAPTER FIVE RESULTS AND CONCLUSION	39
5.1 Introduction	39
5.2 Results	40
5.3 Conclusion	66

BIBLIOGRAPHY.....	70
APPENDIX A Computer Program.....	71
APPENDIX B Equations Used in Computer Program.....	96
VITA.....	99

LIST OF FIGURES

Figure 2.1 Repko's Process Capability Plot.....	8
Figure 3.1 \bar{X} Chart.....	17
Figure 3.2 s^2 Chart.....	19
Figure 3.3 Joint Chart.....	20
Figure 3.4 Correlation Between Sample Mean and Variance.....	23
Figure 3.5 No Correlation Between Sample Mean and Variance....	24
Figure 3.6 T-Chart.....	27
Figure 4.1 Relationship Between α and β	35
Figure 5.1 Probability of a Type II Error $\alpha = .01 \lambda = 0.0$	42
Figure 5.2 Probability of a Type II Error $\alpha = .01 \lambda = 0.5$	44
Figure 5.3 Probability of a Type II Error $\alpha = .01 \lambda = 1.0$	46
Figure 5.4 Probability of a Type II Error $\alpha = .01 \lambda = 1.5$	48
Figure 5.5 Probability of a Type II Error $\alpha = .01 \lambda = 2.0$	50
Figure 5.6 Probability of a Type II Error $\alpha = .05 \lambda = 0.0$	52
Figure 5.7 Probability of a Type II Error $\alpha = .05 \lambda = 0.5$	54
Figure 5.8 Probability of a Type II Error $\alpha = .05 \lambda = 1.0$	56
Figure 5.9 Probability of a Type II Error $\alpha = .05 \lambda = 1.5$	58
Figure 5.10 Probability of a Type II Error $\alpha = .05 \lambda = 2.0$	60

LIST OF TABLES

Table 2.1	Power of the Sum Statistic Vs. Separate Charts.....	13
Table 5.1	Probability of a Type II Error $\alpha = .01$ $\lambda = 0.0$	41
Table 5.2	Probability of a Type II Error $\alpha = .01$ $\lambda = 0.5$	43
Table 5.3	Probability of a Type II Error $\alpha = .01$ $\lambda = 1.0$	45
Table 5.4	Probability of a Type II Error $\alpha = .01$ $\lambda = 1.5$	47
Table 5.5	Probability of a Type II Error $\alpha = .01$ $\lambda = 2.0$	49
Table 5.6	Probability of a Type II Error $\alpha = .05$ $\lambda = 0.0$	51
Table 5.7	Probability of a Type II Error $\alpha = .05$ $\lambda = 0.5$	53
Table 5.8	Probability of a Type II Error $\alpha = .05$ $\lambda = 1.0$	55
Table 5.9	Probability of a Type II Error $\alpha = .05$ $\lambda = 1.5$	57
Table 5.10	Probability of a Type II Error $\alpha = .05$ $\lambda = 2.0$	59
Table 5.11	The Most Powerful Chart $\alpha = .01$	64
Table 5.12	The Most Powerful Chart $\alpha = .05$	65

Chapter One

INTRODUCTION

Over the past thirty years, the need for total quality control has been discovered by industry. With the entire organization of a company exploring quality control, many tools have been developed. One successful tool has been the control chart. A control chart is a statistical device used for process control, specifically for a repetitive process. Use of a control chart may be focused upon monitoring the status of one of several parameters; an \bar{X} chart is used to monitor the mean of a process; an s^2 or R chart is used to control the variability of a process.

Variables control charts, assuming the variable is normally distributed, portray the behavior of a production process in terms of three quantities:

- 1) location of a variable
- 2) variability of a variable
- 3) time

Conventional charts, such as the \bar{X} and R charts, separate the location and variability statistics. When an \bar{X} chart is used, an s^2 or R chart has to be utilized first to verify that the variability of the process is in control

before the process location is examined. Often this is done by monitoring two charts, an \bar{X} chart and an R chart. First, a sample range statistic is plotted on the R chart. If the range is in control, then the sample mean is placed on the \bar{X} chart and a decision is made about the state of the process.

To avoid the process of plotting two statistics to examine only one quality characteristic, single statistics for use on one chart are being developed. Recent work by several authors has been focused upon combining the location and variability quantities. Recent work by Repko [Repko 1986] is based on the concept that the time variable is sometimes less important. He attempts to combine location and variability measures. This is especially appropriate when considering questions of process capability. However, the approximate nature of Repko's method limits its applicability.

In contrast, Reynolds and Ghosh [Reynolds and Ghosh 1981] have attempted to simultaneously represent all three quantities: location, variability, and time. This is an attractive idea but their method allows compensation effects between location and variability. Their method still has appeal for process monitoring applications but is not particularly well suited to the study of capability.

This thesis defines and compares two alternative control charts; one simultaneously tracks the mean and variance of a normally distributed variable while the other examines the mean without the need to examine the variance first. After these control charts are defined, they are

compared to the \bar{X} chart, the s^2 chart, the Reynolds and Ghosh sum statistic, and Repko's process capability plot. In order to make a comparison to Repko's plot, an appropriate Type II error probability expression is derived while recognizing the limitations of the assumptions of Repko's plot.

The first method examined is a joint control chart that contains separate limits for location and variability. One axis contains the control limits for the mean of a process variable and the other axis has limits for the variance of the same variable. The mean and variance are not numerically combined. Therefore, no compensation between the two statistics occurs.

The second approach is a control chart based on the t-statistic. The t-statistic is the distribution for the mean of a sample from a normally distributed population for which the variance is unknown. One problem encountered with some control chart statistics is the approximations made about the standard deviation when the sample size is less than thirty. The t-statistic may have some desirable statistical properties that can be used to control the mean without examining the variance first. Analyzing the utility of the t-statistic is the second question that is addressed.

In order to develop these two methods, the combined statistics of Repko and the statistics presented by Reynolds and Ghosh are examined. The two methods are defined exactly along with any assumptions or limitations. Also defined are the expressions for the Type I and Type II

sampling error probabilities for both the joint control chart and the t-statistic. A Type I sampling error is the observation of a statistic interpreted as signalling the process is out of control when the process is in control. A Type II sampling error is the observation of a statistic interpreted as indicating the process is in control when the process is out of control. The power of a test is the complement of the Type II error probability.

The method of comparison is to fix the Type I error probability at a specific value and to examine differences between the charts in the probability of a Type II error. The Type I error is preset since there must be some standard of equality. The Type II error probability is computed for different values of the mean and variance corresponding to several out of control cases. For each case, the Type II error probabilities for the \bar{X} chart, the S^2 chart, the Reynolds and Ghosh sum statistic, Repko's process capability plot, the joint chart and the t-chart are compared.

A computer program is constructed to perform the numerical computations to compare the different control charts. The output of the program is the probability of a Type II sampling error for each method. The input is the sample size, the Type I sampling error probability, the true value of the mean expressed in terms of the target mean, and the true value of the variance in terms of the target variance. If the appropriate cumulative density function is not in a closed form, either a series expansion or an approximating expression is used. The results are

presented in tabular form, and when feasible, in a graphical manner.

The comparison of Type II error probabilities shows that the Reynolds and Ghosh sum statistic, the \bar{X} chart, and the joint chart are the most powerful tools to use to detect shifts in the mean or variance. The Reynolds and Ghosh statistic is powerful in detecting changes in the variance. As is expected, the \bar{X} chart detects shifts in the mean quite rapidly. The joint chart is powerful for moderate changes in the mean and variance.

Chapter Two

LITERATURE REVIEW

2.1 INTRODUCTION

The control chart is a statistical tool that is widely used in quality control. It was first proposed by Walter Shewhart [Shewhart 1939]. As is now well known, a control chart is a graphical representation of the status of a variable such as a quality characteristic. The theoretical basis for a control chart is the differentiation between types of variability in a process. Variability exists in two forms: inherent and assignable. Inherent variation is a regular property of a statistical distribution and cannot be removed. Assignable variability is caused by a physical influence acting on the process. The purpose of a control chart is to distinguish between these variations and identify when assignable causes are affecting a production process.

A control chart is maintained by taking samples from a process and plotting on the chart in time order some statistic computed from the samples. Control limits on the chart represent the limits within which the plotted points would fall with high probability if the process is operating in control. In control means that the only source of variation is inherent

variability. A point outside the limits indicates that something (i.e. an assignable cause of variation) has happened to change the process.

All basic control charts examine one population statistic. Either location or variability are paired with a time dimension. The status, in control or out of control, is plotted over time. The authors who have studied a variation of the three quantities- location, variability, and time- include Repko, Reynolds and Ghosh.

2.2 REPKO

Repko [1986] develops a combined statistic for use in what he calls a process capability plot. A process capability plot graphically shows the capability of a process. It can be used as a control chart. Samples are taken from a normally distributed process. The sample mean and standard deviation of the quality characteristic are calculated, plotted, and compared to a prediction circle. Repko intends this circle to represent the boundary of in control behavior of a process.

The prediction circle is an $(1-\alpha)\%$ boundary for a χ^2 statistic. The definition of the χ^2 statistic begins with the relationship between normal standardized variables and χ^2 variables. The sum of the squares of normal standardized variables is distributed according to the χ^2 [Duncan 1986]: $Z_1^2 + Z_2^2 \sim \chi^2$. The sample means are normally distributed. If the sample size is assumed to be large, the sample standard deviation is assumed to be normally distributed as well. Therefore, $Z_{\bar{x}}^2 + Z_s^2 \sim \chi^2$ [Repko 1986],

$$R = \sqrt{\frac{\sigma_0^2 \chi_{1-\alpha,2}^2}{n}}$$

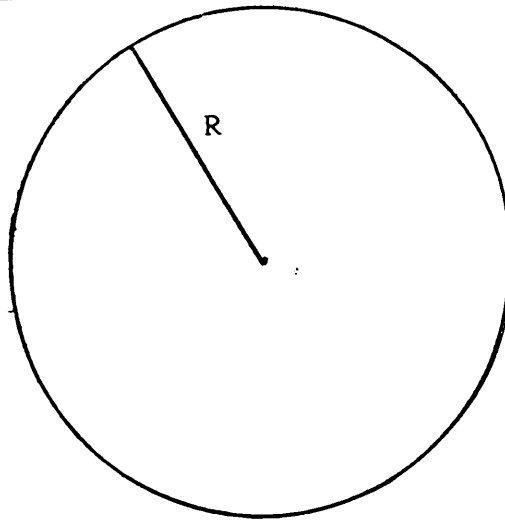


Figure 2.1 Repko's Process Capability Plot

where,

$$Z_{\bar{x}} = \frac{\bar{x} - M_p}{S_p/\sqrt{n}}$$

$$Z_s = \frac{s - S_p}{S_p/\sqrt{2n}}$$

M_p = grand mean, the mean of the sample means

S_p = pooled sample standard deviations

n = sample size, assumed large

Therefore,

$$Y = \left[\frac{\bar{x} - M_p}{S_p/\sqrt{n}} \right]^2 + \left[\frac{s - S_p}{S_p/\sqrt{2n}} \right]^2$$

$$Y \sim \chi_2^2$$

Y will take on values of the χ^2 distribution with 2 degrees of freedom. M_p and S_p can be replaced with target values of the mean and variance, μ_0 and σ_0 to obtain the equation:

$$Y = \left[\frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}} \right]^2 + \left[\frac{s - \sigma_0}{\sigma_0/\sqrt{2n}} \right]^2$$

By manipulating the equation for Y , the equation of an ellipse can be formed.

$$\left[\frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}} \right]^2 + \left[\frac{s - \sigma_0}{\sigma_0/\sqrt{2n}} \right]^2 = \chi^2$$

$$\frac{[\bar{x} - \mu_0]^2}{\sigma_0^2 \chi^2/n} + \frac{[s - \sigma_0]^2}{\sigma_0^2 \chi^2/2n} = 1$$

The intercept of one axis, a, is $\frac{\sigma_0}{\sqrt{n}} \sqrt{\chi^2}$, while the intercept of the other, b, is $\frac{\sigma_0}{\sqrt{2n}} \sqrt{\chi^2}$. If this equation were graphed, the resulting area bounded by the ellipse would be the area in which $(1-\alpha)\%$ of the plotted points from the process should be located if the process is in control. However, an ellipse is not as easy to draw with precision as a circle is. A circle is a special case of an ellipse where $a = b$. Therefore, the axis for the standard deviation can be scaled by a factor of $\sqrt{2}$. This will result in a circle that graphically shows the location of an in control process.

Repko's statistic is a compact manner for studying the behavior of the mean and variance. However, the assumption of a large sample size limits the robustness of this test. s is not normally distributed with mean σ and variance $\sigma^2/\sqrt{2n}$ when n is less than thirty. Repko's fundamental idea of a chart where the location and variability quantities are represented leads to the proposed joint control chart.

2.3 REYNOLDS AND GHOSH

Reynolds and Ghosh [1981] address several issues in designing control charts to detect changes in the mean and variance of a variable. The area discussed of use to this thesis is procedures to simultaneously control the mean and variance. This discussion yields a sum statistic to be used in quality control. The sum statistic is defined as follows:

$$U = \frac{n(\bar{X} - \mu_0)^2}{\sigma^2}$$

$$V = \frac{(n-1)s^2}{\sigma_0^2}$$

$$U \sim \chi^2_1$$

$$V \sim \chi^2_{n-1}$$

$$U + V \sim \chi^2_n, \text{ if } \mu = \mu_0 \text{ and } \sigma = \sigma_0$$

The sum $U + V$ is the sum statistic, which is distributed as a χ if $\mu = \mu_0$ and $\sigma = \sigma_0$. The sum is distributed as a non-central χ if $\mu \neq \mu_0$ or $\sigma \neq \sigma_0$.

Reynolds and Ghosh provide some numerical comparisons using a computer algorithm to compute the non-central χ probabilities:

$$\begin{aligned} \Pi(\mu_1, \sigma_1) &= P(U + V > \chi^2_{n,1-\alpha} \mid \mu_1, \sigma_1) \\ &= P\left(\left[\frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}}\right]^2 + \frac{(n-1)s^2}{\sigma_0^2} > \chi^2_{n,1-\alpha} \mid \mu_1, \sigma_1\right) \\ &= P\left(\left[\frac{\bar{X} - \mu_0}{\sigma_1/\sqrt{n}}\right]^2 + \frac{(n-1)s^2}{\sigma_1^2} > \chi^2_{n,1-\alpha} \frac{\sigma_0^2}{\sigma_1^2} \mid \mu_1, \sigma_1\right) \\ &= P(\chi^2_{(n,c)} > \chi^2_{n,1-\alpha} \frac{\sigma_0^2}{\sigma_1^2} \mid \mu_1, \sigma_1) \end{aligned}$$

$$c = n\left[\frac{\mu_1 - \mu_0}{\sigma_1}\right]^2$$

The power of the test is computed based on different values of the mean and variance of a variable. Results are obtained for the sum statistic and the separate \bar{X} and s^2 charts [Reynolds and Ghosh 1981]. It can be seen

from Table 2.1 that the Reynolds and Ghosh sum statistic performs better than the separate statistics when the variance shifts. However, if only the mean has deviated, the separate charts perform better.

2.4 RAO

Rao develops a control chart to be used under variable time sampling conditions. This chart simultaneously monitors the process mean and process variance. After defining this chart, which is an extension of Repko's chart, Rao compares it to other charts. Due to the approximate nature of the statistic, Rao uses simulation to examine the properties of the control chart. His results indicate that the Repko chart is comparable to the Reynolds and Ghosh chart.

Table 2.1 Power of the Sum Statistic Vs. Separate Charts

$\frac{\mu - \mu_0}{\sigma}$	$\sigma/\sigma_0 = 1$		$\sigma/\sigma_0 = 2$		$\sigma/\sigma_0 = 3$	
	\bar{X} & s^2	sum	\bar{X} & s^2	sum	\bar{X} & s^2	sum
0	.0010	.0010	.3457	.4004	.7700	.8093
1	.1070	.0473	.4723	.5398	.8071	.8431
2	.8394	.6520	.7719	.8119	.8882	.9133

Chapter Three

THE CONTROL CHARTS

3.1 INTRODUCTION

Control charts portray the behavior of a process in terms of the location of a variable, the variability of a variable, and time. Since the mean and variance, or standard deviation, of a normally distributed variable are the bases for the combined control charts examined here, the associated distributions should be reviewed.

The process variable being controlled is X . X is assumed to be normally distributed with mean μ , standard deviation σ , sample mean \bar{X} , and sample standard deviation s . \bar{X} is distributed normally with mean μ and variance σ^2/n , where μ is the mean of the universe being sampled, σ^2 is the variance of the universe, and n is the sample size [Duncan 1986].

$\frac{(n-1)s^2}{\sigma^2}$ is distributed as a χ^2 with $n-1$ degrees of freedom [Duncan 1986]. The expected value of s^2 is σ^2 and s^2 has a variance equal to $\frac{2}{n-1} \sigma^4$ and a standard deviation $\sqrt{\frac{2}{n-1}} \sigma^2$. The only assumption is that the universe of individual values is normal.

Although the standard deviation is the square root of the variance,

the mean of the distribution of sample standard deviations from a normal population is not, in samples less than thirty, the square root of the mean of the distribution of sample variances. Correspondingly, if n is small, the standard deviation of the distribution of sample standard deviations does not equal the square root of the standard deviation of sample variances [Duncan 1986]. The mean value of the distribution of sample standard deviations from a normal universe is [Duncan 1986]:

$$\sqrt{\frac{2}{n-1}} \frac{\left[\frac{n-2}{2}\right]!}{\left[\frac{n-3}{2}\right]!} \sigma$$

The coefficient of σ is called c_4 . The standard deviation of the distribution of sample standard deviations from a normal universe is approximately $\sigma/\sqrt{2(n-1)}$. The exact formula is [Duncan 1986]:

$$\sigma\sqrt{1 - (c_4)^2}$$

3.2 THE JOINT CONTROL CHART

A joint control chart combines the statistical limits and properties from an \bar{X} chart and an s^2 chart. Since the sample mean \bar{X} and the sample variance s^2 are statistically independent [Cowden 1957], no compensation effect occurs between the two estimates. Compensation could occur in a test statistic that adds two statistics. If one term is higher than its target value and the other is lower than its target value, the combined statistic would appear to be in control.

Control limits for an \bar{X} chart are $\mu_0 \pm k\sigma_{\bar{x}}$ [Duncan 1986]. μ_0 is the target value for the mean, $\sigma_{\bar{x}}$ is the standard deviation of the mean, and k is determined so that:

$$\text{probability}(\mu_0 - k\sigma_{\bar{x}} < \bar{x} < \mu_0 + k\sigma_{\bar{x}}) = 1 - \alpha_1.$$

Figure 3.1 shows a typical \bar{X} chart. The Type II error probability expression is:

$$\begin{aligned} \beta_{\bar{x}}(\mu_1, \sigma_1) &= \text{probability}(\text{no signal occurs} | \mu = \mu_1, \sigma = \sigma_1) \\ &= \text{Pr}(\mu_0 - k\sigma_{\bar{x}} < \bar{x} < \mu_0 + k\sigma_{\bar{x}} | \mu = \mu_1, \sigma = \sigma_1) \\ &= \text{Pr}(\mu_0 - \frac{k\sigma_0}{\sqrt{n}} - \mu_1 < \bar{x} - \mu_1 < \mu_0 + \frac{k\sigma_0}{\sqrt{n}} - \mu_1 | \mu_1, \sigma_1) \\ &= \text{Pr}\left(\frac{\mu_0 - \frac{k\sigma_0}{\sqrt{n}} - \mu_1}{\sigma_1/\sqrt{n}} < \frac{\bar{x} - \mu_1}{\sigma_1/\sqrt{n}} < \frac{\mu_0 + \frac{k\sigma_0}{\sqrt{n}} - \mu_1}{\sigma_1/\sqrt{n}} | \mu_1, \sigma_1\right) \\ &= \text{Pr}\left(\frac{\mu_0 - \mu_1}{\sigma_1/\sqrt{n}} - k\frac{\sigma_0}{\sigma_1} < Z < \frac{\mu_0 - \mu_1}{\sigma_1/\sqrt{n}} + k\frac{\sigma_0}{\sigma_1} | \mu_1, \sigma_1\right) [1] \end{aligned}$$

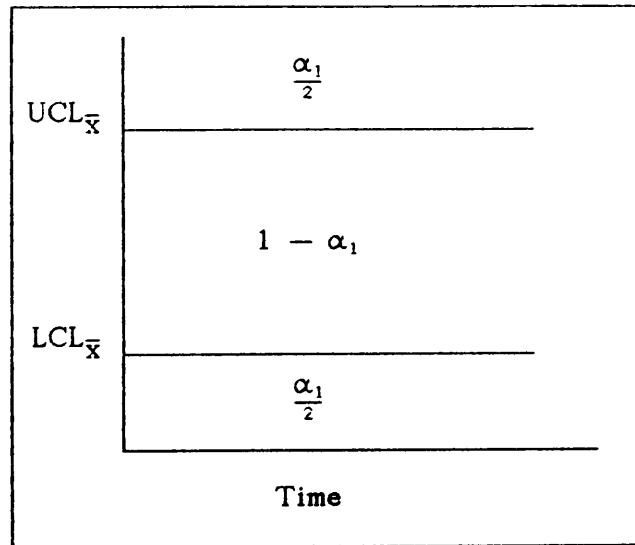


Figure 3.1 \bar{X} CHART

Control limits for an s^2 chart are set so that:

$$\text{probability} \left(\frac{s^2(n-1)}{\sigma^2} > \text{Upper Control Limit} \right) = \alpha_2$$

where the upper control limit is the χ^2 value for $n-1$ degrees of freedom and a tail area of α_2 . A lower limit is sometimes used. However most concerns are with the variance being too large and the lower limit is often taken to be 0. A s^2 chart is shown in Figure 3.2. The Type II error probability expression is:

$$\begin{aligned} \beta(\mu_1, \sigma_1) &= \text{probability}(\text{no signal occurs} | \mu = \mu_1, \sigma = \sigma_1) \\ &= \text{Pr} \left(\frac{s^2(n-1)}{\sigma_0^2} < \chi_{n-1, 1-\alpha}^2 | \mu = \mu_1, \sigma = \sigma_1 \right) \\ &= \text{Pr} \left(\frac{s^2(n-1)}{\sigma_1^2} < \chi_{n-1, 1-\alpha}^2 \frac{\sigma_0^2}{\sigma_1^2} | \mu = \mu_1, \sigma = \sigma_1 \right) \\ &= \text{Pr} \left(\chi^2 < \chi_{n-1, 1-\alpha}^2 \frac{\sigma_0^2}{\sigma_1^2} | \mu = \mu_1, \sigma = \sigma_1 \right) \quad [2] \end{aligned}$$

To construct the joint control chart, the \bar{X} -chart and the s^2 -chart are combined as illustrated in Figure 3.3.

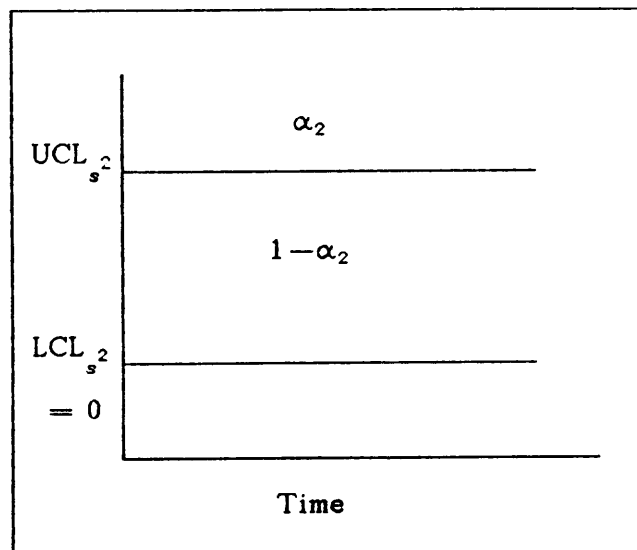


Figure 3.2 s^2 CHART

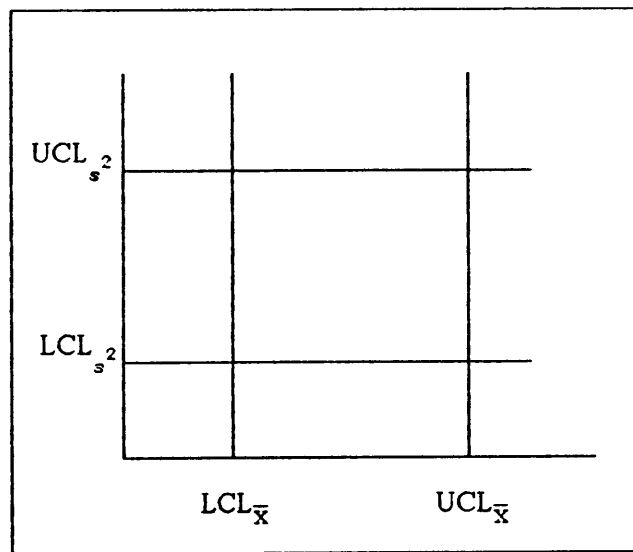


Figure 3.3 JOINT CHART

The alphas defined for the single charts are the Type I error probabilities. Knowing that \bar{X} and s^2 are statistically independent the expression for the Type I error probability for the joint control chart is:

$$\begin{aligned}\alpha &= 1 - \text{probability (observe a point within the limits)} \\ \alpha &= 1 - (1 - \alpha_1)(1 - \alpha_2) \\ \alpha &= (\alpha_1 + \alpha_2 - \alpha_1\alpha_2) \quad [3]\end{aligned}$$

α_1 is arbitrarily chosen to equal α_2 . No factors indicate to do otherwise.

The Type II error probability is defined as a function of the parameters being tested, in this case, the mean and variance. Let $\beta(\mu_1, \sigma_1)$ represent the Type II error. The power of a test, $\Pi(\mu_1, \sigma_1)$, is the complement of $\beta(\mu_1, \sigma_1)$.

$$\begin{aligned}\Pi_j(\mu_1, \sigma_1) &= \text{probability (a signal occurs, given } \mu = \mu_1, \sigma = \sigma_1) \\ &= 1 - (\text{probability no signal occurs}) \\ &= 1 - (\text{probability } \bar{x} \text{ does not signal})(\text{probability } s^2 \text{ does not signal})\end{aligned}$$

Using Equations [1] and [2], the expression is obtained for the power of the joint control chart.

$$\begin{aligned}\Pi_j(\mu_1, \sigma_1) &= [\text{Pr} \left(\frac{\mu_0 - \mu_1}{\sigma_1/\sqrt{n}} - k \frac{\sigma_0}{\sigma_1} < Z < \frac{\mu_0 - \mu_1}{\sigma_1/\sqrt{n}} + k \frac{\sigma_0}{\sigma_1} \right)] * \\ &\quad \{ \text{Pr} [\chi^2_{n-1} < (\chi^2_{\alpha_2, n-1}) \frac{\sigma_0^2}{\sigma_1^2}] \} \quad [4]\end{aligned}$$

The joint control chart combines the location and variability dimensions. The time dimension can be portrayed by either numbering the points in time order or drawing a line from point to point in time order. The advantages of combining the location and variability dimensions with no compensation effect are several.

The primary advantage of the joint chart is the visual relationship between the sample mean and sample variance. This information is lost when separate \bar{X} and s^2 charts are used. Since \bar{X} and s^2 are statistically independent, any correlation between the two is a result of the process. This knowledge can help the manufacturer in understanding and controlling his process. For example, an in-control process would be expected to yield random points within the limits. As an example, consider Figure 3.4. If sampling resulted in a chart like the one in Figure 3.4, a reasonable inference would be that the process is causing a correlation between the sample mean and variance.

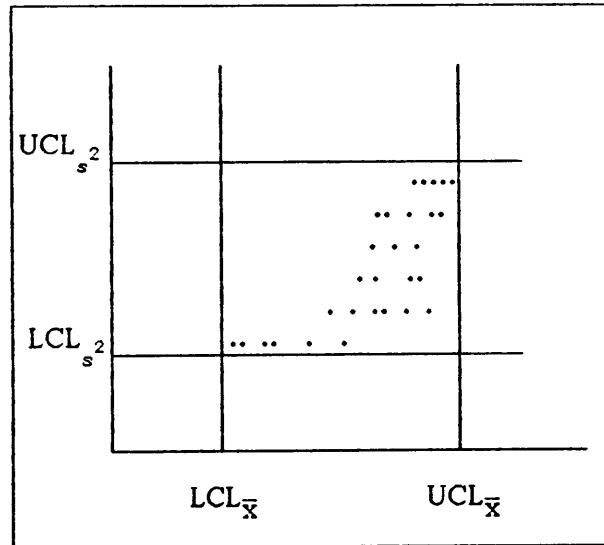


Figure 3.4 JOINT CHART

Correlation Between the Sample Mean and Variance

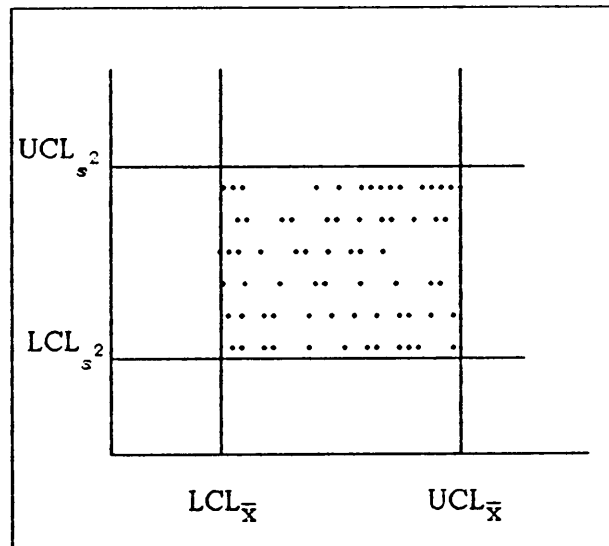


Figure 3.5 JOINT CHART

No Correlation Between the Sample Mean and Variance

The second advantage of the joint control chart is the definition of the Type I sampling error probability. When separate charts are used to control the mean and variance, an alpha is specified for the variability chart and an alpha is specified for the location chart. Hence the probability of a Type I error in controlling the mean is a function of the probability of a Type I error in the variability chart and the probability of a Type I error in the location chart. For example, if both alphas for the separate charts were set to .05, the combined alpha would be $1 - (1 - .05)^2 = .0975$. However, if a user specified an alpha of .05 for the joint control chart, the control limits may be set so that .05 rather than .0975 would be the final alpha.

3.3 THE T-STATISTIC CONTROL CHART

In order to understand the usefulness of the t-statistic in a control chart, the related hypothesis should be discussed. It is assumed that the sample comes from a normal universe with mean μ and standard deviation unknown. The t-statistic tests the null hypothesis:

$$H_0: \mu = \mu_0$$

$$H_1: \mu = \mu_1$$

The test statistic is:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Therefore, the t-statistic supports inferences concerning the mean without knowledge of the standard deviation. The mean is tested using the sample variance.

The t-chart should not arbitrarily be used in any situation. The t-chart can be used at the start of a production process while the variance is unknown. While the data for an estimation of the variance is being collected, the mean can be controlled using the t-statistic. This is better than performing no control procedures at all. Since the t-statistic has the sample variance in the denominator, it seems that the test will not signal when the variance increases. Therefore the t-chart would appear not to be powerful in any other situation. This statement is verified by comparisons to other charts.

As shown in Figure 3.6, the control chart is set up like any other, with,

$$UCL_t = t_{1-\alpha/2}$$

$$LCL_t = t_{\alpha/2}$$

The expression for the Type I error probability is:

$$\alpha = 1 - \text{probability (point is within the limits)}$$

$$\alpha = 1 - \text{probability } (t_{\alpha/2} \leq t \leq t_{1-\alpha/2}) \quad [5]$$

The expression for the power is:

$$P(\mu_1, \sigma_1) = 1 - \text{probability (no signal occurs} | \mu = \mu_1, \sigma = \sigma_1)$$

$$= 1 - \text{Pr } (t_{\alpha/2} \leq t \leq t_{1-\alpha/2} | \mu_1, \sigma_1)$$

$$= 1 - \text{Pr } (t_{\alpha/2} \leq \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \leq t_{1-\alpha/2} | \mu_1, \sigma_1)$$

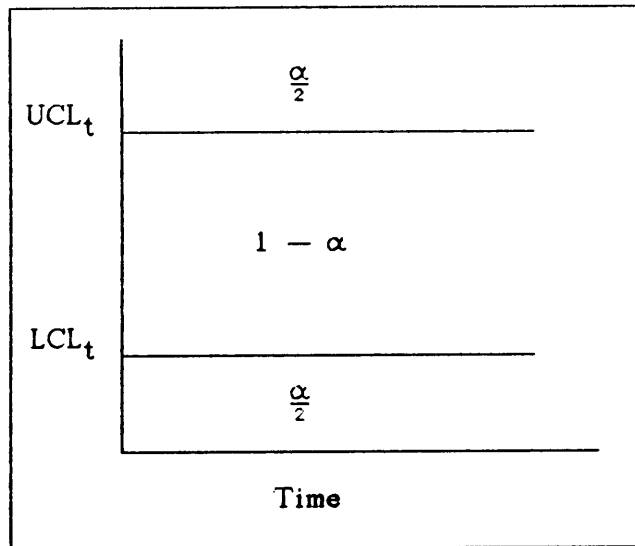


Figure 3.6 T - CHART

$$\begin{aligned}
&= 1 - \Pr \left(t_{\alpha/2} - \frac{\mu_1}{\sigma_1/\sqrt{n}} \leq \frac{\bar{x} - \mu_0}{\sigma_1/\sqrt{n}} - \frac{\mu_1}{\sigma_1/\sqrt{n}} \leq t_{1-\alpha/2} \right. \\
&\quad \left. - \frac{\mu_1}{\sigma_1/\sqrt{n}} \right) \\
&= 1 - \Pr \left(t_{\alpha/2} - \frac{\mu_1}{\sigma_1/\sqrt{n}} \leq \frac{\bar{x} - \mu_1}{\sigma_1/\sqrt{n}} - \frac{\mu_0}{\sigma_1/\sqrt{n}} \leq t_{1-\alpha/2} \right. \\
&\quad \left. - \frac{\mu_1}{\sigma_1/\sqrt{n}} \right) \\
&= 1 - \Pr \left(t_{\alpha/2} - \frac{\mu_1}{\sigma_1/\sqrt{n}} + \frac{\mu_0}{\sigma_1/\sqrt{n}} \leq \frac{\bar{x} - \mu_1}{\sigma_1/\sqrt{n}} \leq t_{1-\alpha/2} \right. \\
&\quad \left. - \frac{\mu_1}{\sigma_1/\sqrt{n}} + \frac{\mu_0}{\sigma_1/\sqrt{n}} \right) \\
&= 1 - \Pr \left(t_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma_1/\sqrt{n}} \leq t \leq t_{1-\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma_1/\sqrt{n}} \right) [6]
\end{aligned}$$

Due to the nature of the t-distribution, the power of the test will vary from sample to sample. This is because the power expression contains a random variable, s . In practical use, the power of the test is computed based on an estimate of the standard deviation, s . For the computations performed here, the estimate, s , is replaced with σ_1 .

3.4 TYPE II ERROR PROBABILITY EXPRESSIONS

Power expressions for the joint control chart and the t-statistic chart have been obtained [Equations 4 and 6]. From [4] the Type II error probability for the joint chart is:

$$\begin{aligned}
\beta_j(\mu_1, \sigma_1) &= [\Pr \left(\frac{\mu_0 - \mu_1 \sqrt{n}}{\sigma_1} - k \frac{\sigma_0}{\sigma_1} < Z < \frac{\mu_0 - \mu_1 \sqrt{n}}{\sigma_1} + k \frac{\sigma_0}{\sigma_1} \right) * \\
&\quad \Pr \left[\frac{s^2(n-1)}{\sigma_1^2} < (\chi^2_{\alpha_2, n-1}) \frac{\sigma_0^2}{\sigma_1^2} \mid \mu = \mu_1, \sigma = \sigma_1 \right)] [7]
\end{aligned}$$

From [6] the probability of a Type II error for the t-statistic chart is:

$$\beta_t(\mu_1, \sigma_1) = [\Pr \left(t_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma_1/\sqrt{n}} \leq t \leq t_{1-\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma_1/\sqrt{n}} \mid \mu_1, \sigma_1 \right)] [8]$$

From equations [1] and [2], the Type II error probability expressions for the \bar{X} chart and the s^2 chart are:

$$\beta_{\bar{x}}(\mu_1, \sigma_1) = [\text{Pr } \bar{x} \text{ does not signal } \mu = \mu_1] \\ = [\text{Pr}(\frac{\mu_0 - \mu_1}{\sigma_1/\sqrt{n}} - k\frac{\sigma_0}{\sigma_1} < Z < \frac{\mu_0 - \mu_1}{\sigma_1/\sqrt{n}} + k\frac{\sigma_0}{\sigma_1}) | \mu_1, \sigma_1] \quad [9]$$

$$\beta_{s^2}(\mu_1, \sigma_1) = [\text{probability } s^2 \text{ does not signal}] \\ = [\text{Pr} [\frac{s^2(n-1)}{\sigma_1^2} < (\chi^2_{2, 1-\alpha}) \frac{\sigma_0^2}{\sigma_1^2}] \quad [10]$$

3.5 REPKO'S PLOT

In order to compare Repko's chart to the other control charts using Type II error probabilities, the error introduced by the small sample size assumption has to be recognized. Repko's control limit is defined by an alpha and the sample size:

$$P\left(\left[\frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}}\right]^2 + \left[\frac{s - \sigma_0}{\sigma_0/\sqrt{2n}}\right]^2 > \chi^2_{2, 1-\alpha}\right) = \alpha \quad [11]$$

For a small sample size, this probability expression is not accurate. To rectify the error, [11] is redefined.

$$U^2 = \left[\frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}}\right]^2 \\ V_1^2 = \left[\frac{s - \sigma_0}{\sigma_0/\sqrt{2n}}\right]^2 \\ V_2^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

$$U^2 + V_2^2 \sim \chi^2_n$$

$$\begin{aligned}
& P\left(\left[\frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}}\right]^2 + \left[\frac{s - \sigma_0}{\sigma_0/\sqrt{2n}}\right]^2 > \chi^2_{2,1-\alpha}\right) \\
&= P\left(U^2 + \left[\frac{s - \sigma_0}{\sigma_0/\sqrt{2n}}\right]^2 > \chi^2_{2,1-\alpha}\right) \\
&= P\left\{\frac{s - \sigma_0}{\sigma_0/\sqrt{2n}} > (\chi^2_{2,1-\alpha} - U^2)^{1/2}\right\} \\
&= P\{s > (\chi^2_{2,1-\alpha} - U^2)^{1/2} \sigma_0/\sqrt{2n} + \sigma_0\} \\
&= P\{s^2 > [(\chi^2_{2,1-\alpha} - U^2)^{1/2} \sigma_0/\sqrt{2n} + \sigma_0]^2\} \\
&= P\left\{\frac{(n-1)s^2}{\sigma_0^2} > \frac{n-1}{\sigma_0^2} [(\chi^2_{2,1-\alpha} - U^2)^{1/2} \sigma_0/\sqrt{2n} + \sigma_0]^2\right\} \\
&= P\{V_2^2 > n-1 [(\chi^2_{2,1-\alpha} - U^2)^{1/2}/\sqrt{2n} + 1]^2\} \\
&= P\{U^2 + V_2^2 > n-1 \left[\left(\frac{\chi^2_{2,1-\alpha} - U^2}{2n}\right)^{1/2} + 1\right]^2 + U^2\} \tag{12}
\end{aligned}$$

The Type I error probability, α' , can be calculated for Repko's plot using [12]. This can be compared to the alpha intended by Repko in [11]. However, equation [12] can not be algebraically solved for one answer. Contours are developed depending on values of U^2 . If $\mu = \mu_0$, U^2 will be between 0 and 10.89 with probability .9995. Two values of U^2 are used: the expected value of U^2 , 1, and the value of U^2 that yields the minimum value of α' . To find the value of U^2 that minimizes alpha, the maximum value of the probability limit term, T, in [12] is found.

$$T = (n-1) \left[\left(\frac{\chi^2_{2,1-\alpha} - U^2}{2n} \right)^{1/2} + 1 \right]^2 + U^2$$

$$T = (n-1) \left[\frac{\chi^2_{2,1-\alpha} - U^2}{2n} + 2 \left(\frac{\chi^2_{2,1-\alpha} - U^2}{2n} \right)^{1/2} + 1 \right] + U^2$$

$$\frac{dT}{\delta U^2} = 0$$

$$\frac{dT}{\delta U^2} = (n-1) \left[-1/2n + \frac{1}{2}(2) \left[\frac{\chi^2_{2,1-\alpha} - U^2}{2n} \right]^{-1/2} (-1/2n) \right] + 1 = 0$$

$$(n-1)(-1/2n) \left[1 + \left(\frac{\chi^2_{2,1-\alpha} - U^2}{2n} \right)^{-1/2} \right] + 1 = 0$$

$$(n-1)(-1/2n) \left[1 + \left(\frac{\chi^2_{2,1-\alpha} - U^2}{2n} \right)^{-1/2} \right] = -1$$

$$\left[1 + \left[\frac{\chi^2_{2,1-\alpha} - U^2}{2n} \right]^{-1/2} \right] = \frac{2n}{n-1}$$

$$\left[\frac{\chi^2_{2,1-\alpha} - U^2}{2n} \right]^{-1/2} = \frac{2n}{n-1} - 1$$

$$\frac{\chi^2_{2,1-\alpha} - U^2}{2n} = \left[\frac{2n - n + 1}{n-1} \right]^{-2}$$

$$U_{\min}^2 = \chi^2_{2,1-\alpha} - \left[\frac{n-1}{n+1} \right]^2 (2n) \quad [13]$$

U_{\min}^2 is the value that gives the lower bound of α' . Type II error probabilities for the intended alpha, α , are computed for two contours of the actual alpha. Therefore two values of alpha are computed based on the expected value of U^2 and the value of U^2 that results in the minimum contour value of alpha. α^1 is the alpha corresponding to $U^2 = 1$. α^m is the lower bound on alpha. To compute the probability of not signalling, a non-central χ distribution is used.

$$\begin{aligned} \beta(\mu_1, \sigma_1, U^2) &= P(U^2 + V_2^2 < T / \mu = \mu_1, \sigma = \sigma_1) \\ &= P\left(\frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}} \right)^2 + \frac{(n-1)s^2}{\sigma_0^2} < T / \mu = \mu_1, \sigma = \sigma_1 \end{aligned}$$

$$= P\left(\left[\frac{\bar{x} - \mu_0}{\sigma_1/\sqrt{n}}\right]^2 + \frac{(n-1)s^2}{\sigma_1^2} < T \frac{\sigma_0^2}{\sigma_1^2} \mid \mu = \mu_1, \sigma = \sigma_1\right)$$

$$\beta(\mu_1, \sigma_1, U^2) = P(\chi^2_{(n,c)} < T \frac{\sigma_0^2}{\sigma_1^2} \mid \mu = \mu_1, \sigma = \sigma_1)$$

where,

$$c = n \left[\frac{\mu_1 - \mu_0}{\sigma_1} \right]^2$$

$$T = (n-1) \left[\left(\frac{\chi^2_{2,1-\alpha} - U^2}{2n} \right)^{1/2} + 1 \right]^2 + U^2$$

$$U_e^2 = 1, \quad U_{max}^2 = \chi^2_{2,1-\alpha} - \left[\frac{n-1}{n+1} \right]^2 (2n) \quad [14]$$

Expressions for the probability of a Type II error have been defined for all six control charts. The charts now are compared based on several values of out of control cases.

Chapter Four

COMPARISON OF CONTROL CHARTS

4.1 AVERAGE RUN LENGTH

The basis of comparison for control charts is the Average Run Length (ARL). The ARL is the average length of a run of in-control points [Duncan 1986].

$$ARL = \frac{1}{1 - \text{probability(no signal)}}$$

Note that the ARL can be defined for an in-control case or an out-of-control situation. The probability of no signal when the process is in control is $1 - \alpha$. The probability of no signal when the process is out of control is $\beta(\mu_1, \sigma_1)$. Therefore to compare charts, the Type II error probabilities $\beta(\mu_1, \sigma_1)$ are used. These equations already have been defined in Chapter Three.

4.2 TYPE I ERROR PROBABILITY VS. TYPE II ERROR PROBABILITY

When developing the probability expression for Repko's plot, equation [12] was defined for the minimum value of α' . This was done

because of the relationship between the probability of a Type I error and the probability of a Type II error. If a minimum α is defined, the corresponding β expression will be a maximum. This relationship can be understood graphically. The smaller the α , the larger the circle that is the limit for Repko's chart. The larger the circle, the higher the probability of being within the circle and not signalling when the process is out of control. Therefore, the minimum α yields the upper bound on the Type II error probability for Repko's plot.

For any sampling with a fixed sample size, the Type I sampling error probability and the Type II sampling error probability have the inverse relationship already described. However, with a fixed Type I error probability, a target Type II error probability can be obtained by changing the sample size. An increase in the sample size will decrease the Type II error probability and vice versa.

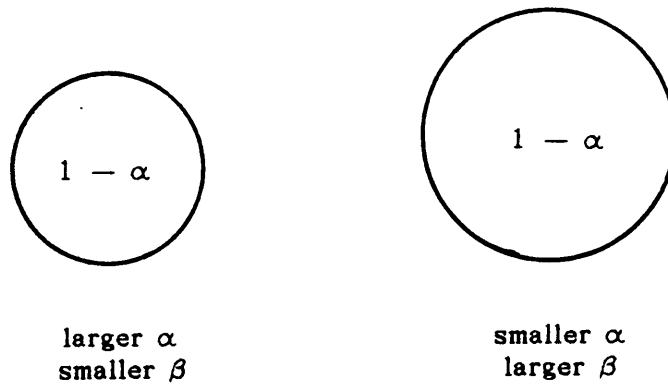


Figure 4.1 Relationship Between α and β

4.3 COMPUTER PROGRAM

A computer program was developed to perform the computations to make the comparison between the six control charts. Based on inputs of the sample size, alpha, and the population mean and variance, the program computes Type II errors probabilities for:

- 1) the joint control chart
- 2) the t-statistic chart
- 3) the \bar{X} chart
- 4) the s^2 chart
- 5) the sum statistic chart
- 6) Repko's plot

All the error probabilities have been expressed as a function of μ_1 and σ_1 . The probability expressions are rewritten as a function of absolute measures of μ_1 and σ_1 : λ and γ , where

$$\lambda = \frac{\mu_1 - \mu_0}{\sigma_1}$$

$$\gamma = \frac{\sigma_1^2}{\sigma_0^2}$$

$$\beta_{\bar{x}}(\mu_1, \sigma_1) = [\text{Pr} (-\lambda\sqrt{n} - k/\sqrt{\gamma} < Z < -\lambda\sqrt{n} + k/\sqrt{\gamma}) \mid \mu_1, \sigma_1] \quad [15]$$

$$\beta_{s^2}(\mu_1, \sigma_1) = [\text{Pr} [\chi^2 < (\chi^2_{1-\alpha, n-1})/\gamma] \mid \mu_1, \sigma_1] \quad [16]$$

$$\beta_j(\mu_1, \sigma_1) = [\text{Pr} (-\lambda\sqrt{n} - k/\sqrt{\gamma} < Z < -\lambda\sqrt{n} + k/\sqrt{\gamma}) *]$$

$$\Pr [\chi^2 < (\chi^2_{1-\alpha, 2, n-1})/\gamma \mid \mu_1, \sigma_1] \quad [17]$$

$$\beta_z(\mu_1, \sigma_1) = [\Pr (t_{\alpha/2} - \lambda\sqrt{n} \leq t \leq t_{1-\alpha/2} - \lambda\sqrt{n}) \mid \mu_1, \sigma_1] \quad [18]$$

$$\beta_{r,c}(\mu_1, \sigma_1) = P(\chi^2_{(n,c)} < (\chi^2_{n,1-\alpha})/\gamma \mid \mu_1, \sigma_1)$$

$$c = n\lambda^2 \quad [19]$$

$$\beta_{rep}(\mu_1, \sigma_1, U^2) = P(\chi^2_{(n,c)} < T/\gamma \mid \mu_1, \sigma_1)$$

$$c = n\lambda^2$$

$$T = (n-1) \left[\left(\frac{\chi^2_{2,1-\alpha} - U^2}{2n} \right)^{1/2} + 1 \right]^2 + U^2$$

$$U_0^2 = 1, \quad U_{max}^2 = \chi^2_{2,1-\alpha} - \left[\frac{n-1}{n+1} \right]^2 (2n) \quad [20]$$

Type II error probabilities are computed for λ ranging from 0 to 2 and γ ranging from 1 to 3. These values were chosen because they provide sufficient data about the behavior of the control charts. Initially, a higher range of out-of-control cases was examined; however, the increase in data did not yield additional information. To compute the Type II error probabilities, several probability distributions are needed within the program. They are :

- 1) the normal distribution
- 2) the χ^2 distribution
- 3) the non-central χ^2 distribution

4) the t-distribution

Series expansions are used for the t-distribution and the χ^2 distribution. An approximation for the normal is used. The error for the approximation is less than 7.5×10^{-8} . A truncated infinite series is the approximation for the non-central χ^2 [Abramowitz and Stegun 1972]. Refer to Appendices A and B for a listing of the program and a complete explanation of the approximations and expansions used.

Chapter Five

RESULTS AND CONCLUSION

5.1 INTRODUCTION

In order to compare control charts, a standard of equality must be established and a criterion for comparison defined. The standard of equality selected is the probability of a Type I sampling error and the sample size. The criterion for comparison is the probability of a Type II sampling error. A Type II sampling error is the observation of a statistic interpreted as indicating the process is in control when the process is out of control. Therefore, the probability of a Type II error is a function of the out-of-control parameters: the population mean μ_1 and the population standard deviation σ_1 .

Six control charts are compared:

- 1) the \bar{X} chart
- 2) the s^2 chart
- 3) the joint chart
- 4) the t chart
- 5) the Reynolds and Ghosh chart
- 6) the Repko plot

Several values of the Type I error probability and the sample size are used in the comparison of the charts.

5.2 RESULTS

The comparisons of the Type II error probabilities are as a function of λ and γ . λ is the number of standard deviations that the population mean exceeds the target mean. γ is the ratio of the population variance to the target variance. The control charts are compared for λ ranging from 0 to 2 in increments of 0.5 and γ ranging from 1 to 3 in increments of 0.25. Two initial conditions are used:

- 1) $\alpha = .01, n = 5$
- 2) $\alpha = .05, n = 5$

The results for the first case are numerically presented in Tables 5.1 through 5.5 and graphically shown in Figures 5.1 through 5.5. Tables 5.6 through 5.10 and Figures 5.5 through 5.10 contain the information for the second case. The results from the first case are discussed. Any changes caused by the increase of α to .05 are noted when appropriate.

In the case that the mean does not shift, the Reynolds and Ghosh statistic is the most powerful method. This result corresponds to the results obtained in 1981 by Reynolds and Ghosh [Reynolds 1981].

TABLE 5.1 PROBABILITY OF A TYPE II SAMPLING ERROR

$$\lambda = 0$$

		Gamma							
Chart	1.00*	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
\bar{X}	.9900	.9787	.9645	.9484	.9313	.9139	.8965	.8794	.8628
s^2	.9900	.9688	.9351	.8921	.8437	.7933	.7431	.6946	.6485
Joint	.9900	.9696	.9363	.8927	.8423	.7888	.7348	.6822	.6320
T	.9900	.9900	.9900	.9900	.9900	.9900	.9900	.9900	.9900
R&G	.9900	.9661	.9264	.8748	.8167	.7565	.6971	.6405	.5876
Repko	.9928	.9737	.9397	.8938	.8403	.7835	.7264	.6711	.6187

$$\alpha = .01, n = 5$$

*This is the in control case. The probability of no signal is $1 - \alpha$.

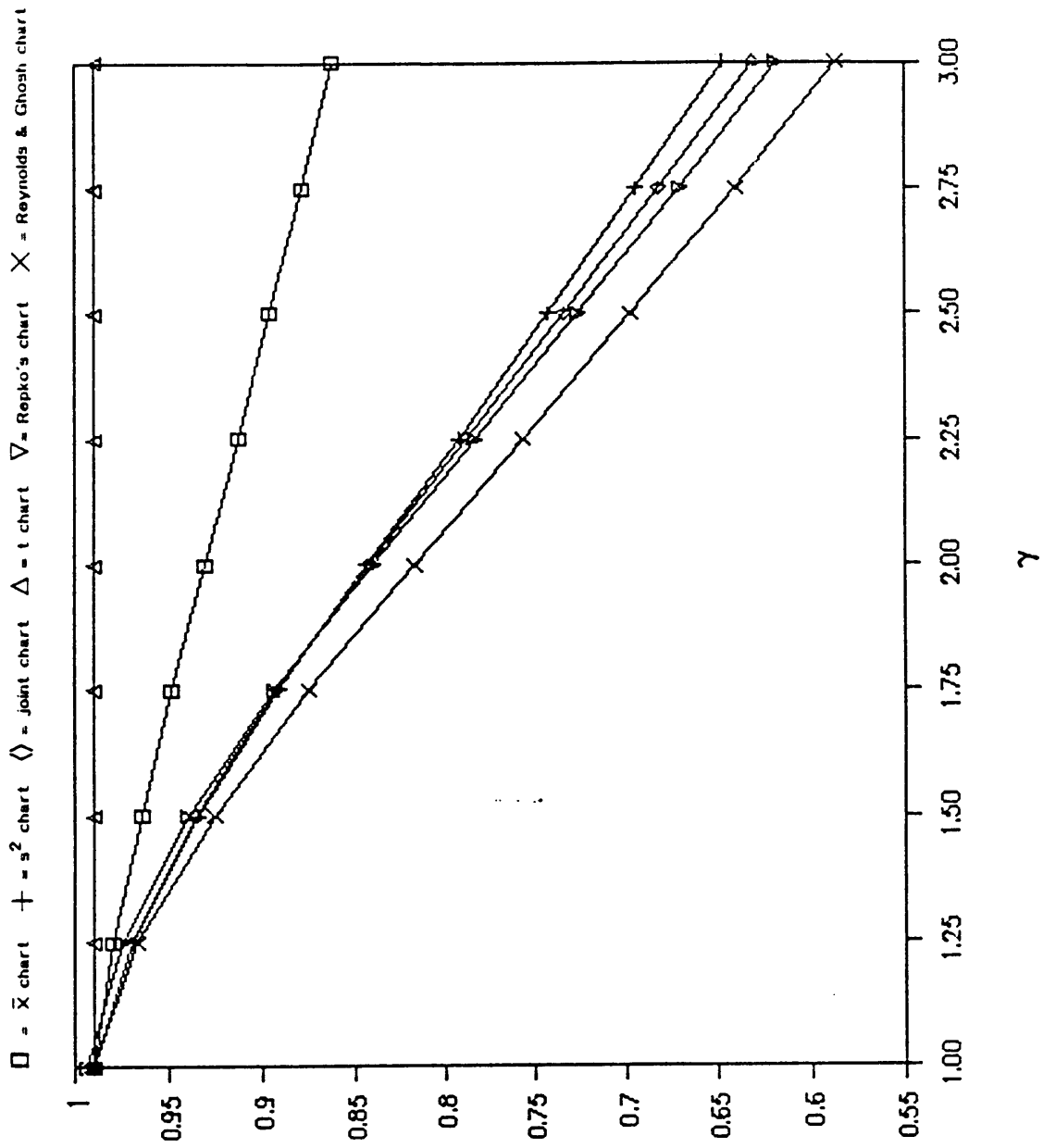


Figure 5.1
Probability of a Type II Error
 $\alpha = .01, n = 5, \lambda = 0.0$

TABLE 5.2 PROBABILITY OF A TYPE II SAMPLING ERROR

$$\lambda = 0.5$$

Gamma									
Chart	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
\bar{X}	.9273	.8816	.8368	.7951	.7571	.7229	.6920	.6642	.6390
s^2	.9900	.9688	.9351	.8921	.8437	.7933	.7431	.6946	.6485
Joint	.9488	.9002	.8411	.7768	.7117	.6487	.5897	.5354	.4862
T	.9851	.9851	.9851	.9851	.9851	.9851	.9851	.9851	.9851
R&G	.9686	.9172	.8482	.7721	.6960	.6244	.5590	.5005	.4486
Repko	.9759	.9320	.8700	.7989	.7260	.6558	.5907	.5317	.4788

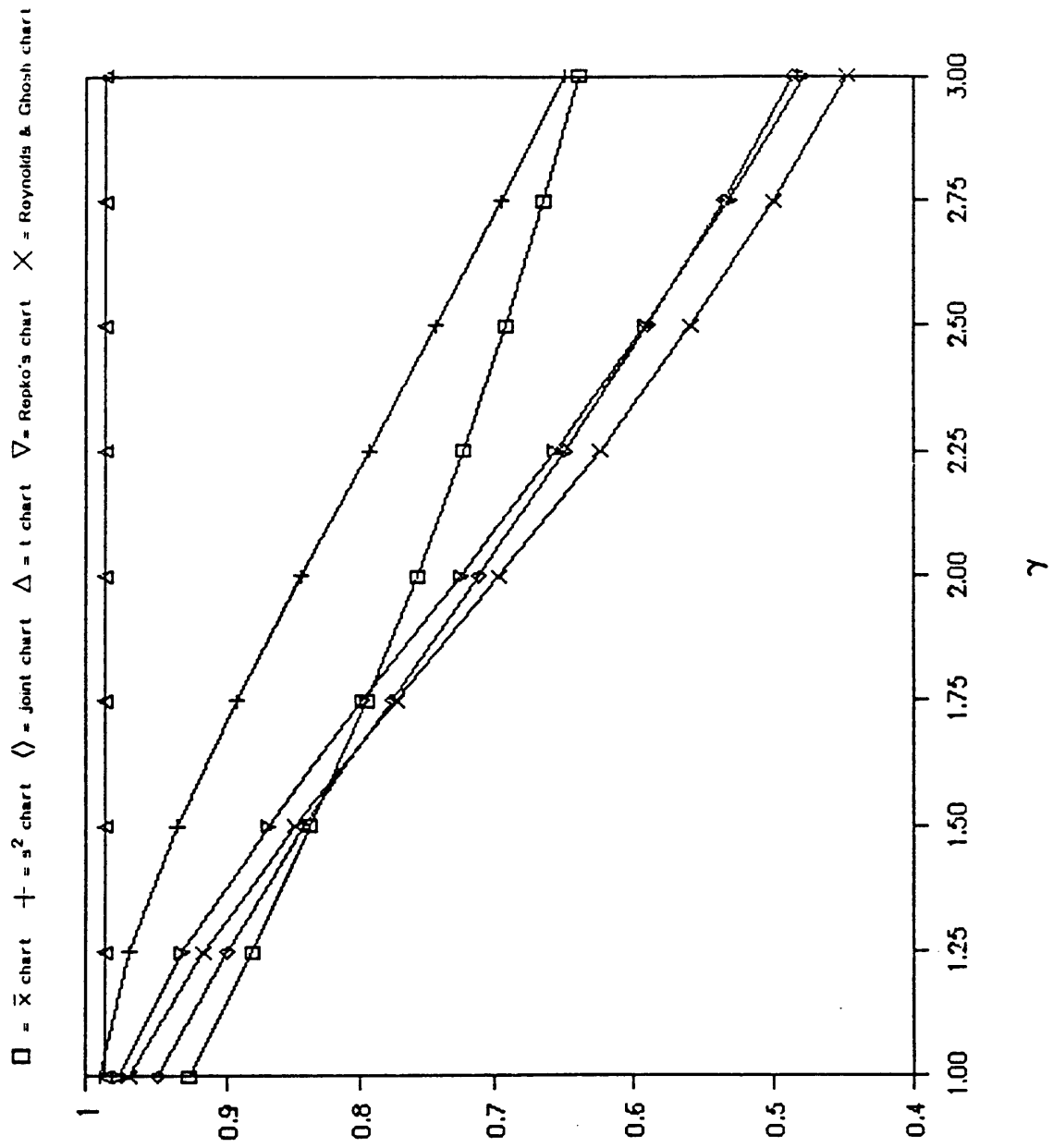


Figure 5.2
 Probability of a Type II Error
 $\alpha = .01, n = 5, \lambda = 0.5$

TABLE 5.3 PROBABILITY OF A TYPE II SAMPLING ERROR

$$\lambda = 1.0$$

		Gamma							
Chart	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
\bar{X}	.6325	.5266	.4470	.3862	.3391	.3018	.2718	.2473	.2268
s^2	.9900	.9688	.9351	.8921	.8437	.7933	.7431	.6946	.6485
Joint	.7009	.5945	.4978	.4184	.3530	.2995	.2555	.2193	.1894
T	.9603	.9603	.9603	.9603	.9603	.9603	.9603	.9603	.9603
R&G	.8339	.6970	.5696	.4630	.3774	.3099	.2566	.2145	.1809
Repko	.8599	.7310	.6057	.4976	.4091	.3382	.2816	.2364	.2001

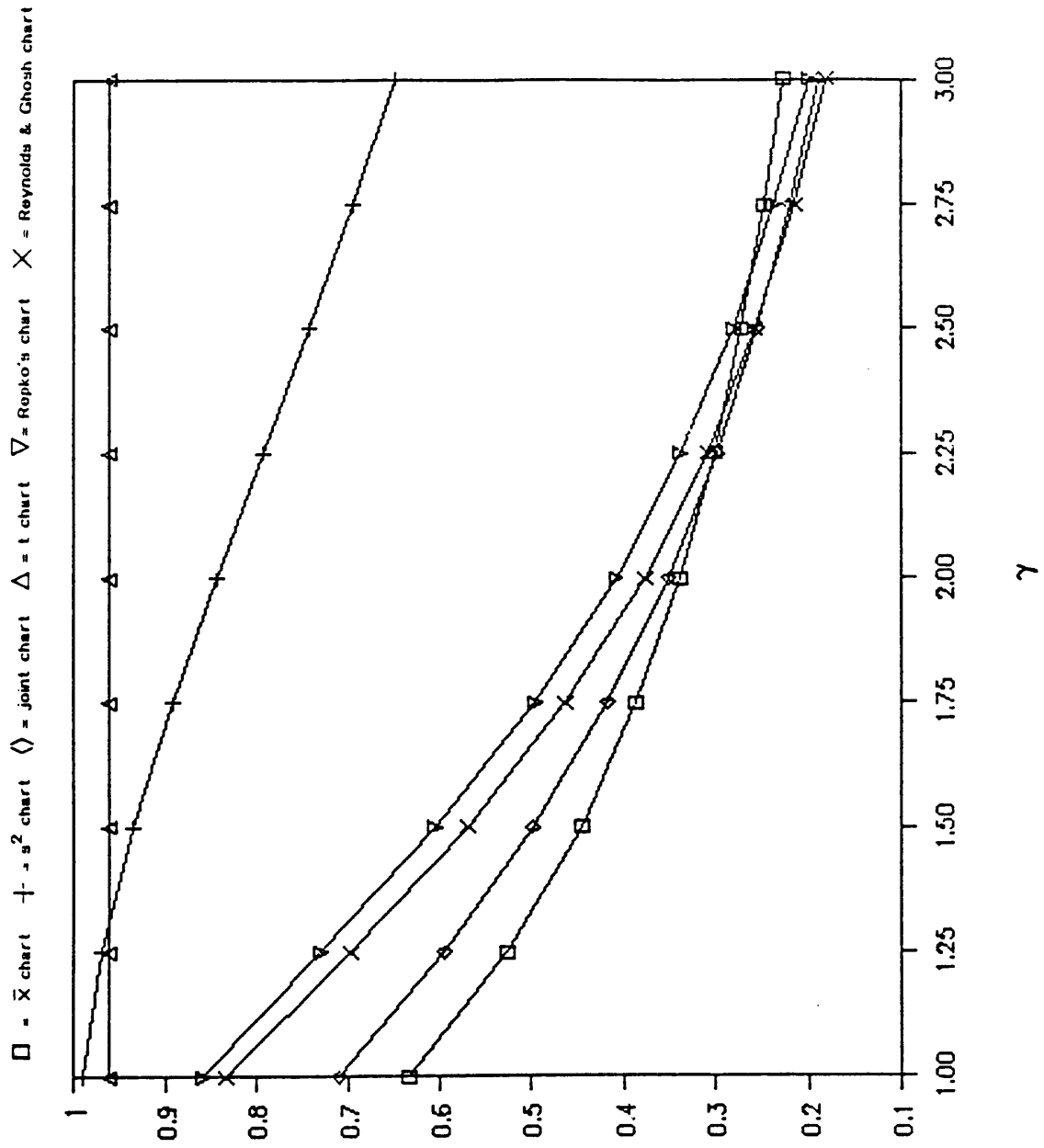


Figure 5.3
 Probability of a Type II Error
 $\alpha = .01, n = 5, \lambda = 1.0$

TABLE 5.4 PROBABILITY OF A TYPE II SAMPLING ERROR

$$\lambda = 1.5$$

		Gamma								
Chart	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	
\bar{X}	.2181	.1467	.1054	.0797	.0626	.0508	.0422	.0359	.0309	
s^2	.9900	.9688	.9351	.8921	.8437	.7933	.7431	.6946	.6485	
Joint	.2885	.1943	.1368	.0999	.0750	.0576	.0452	.0360	.0291	
T	.8596	.8596	.8596	.8596	.8596	.8596	.8596	.8596	.8596	
R&G	.4881	.3178	.2103	.1433	.1006	.0728	.0540	.0411	.0318	
Repko	.5311	.3534	.2373	.1632	.1154	.0839	.0625	.0476	.0370	

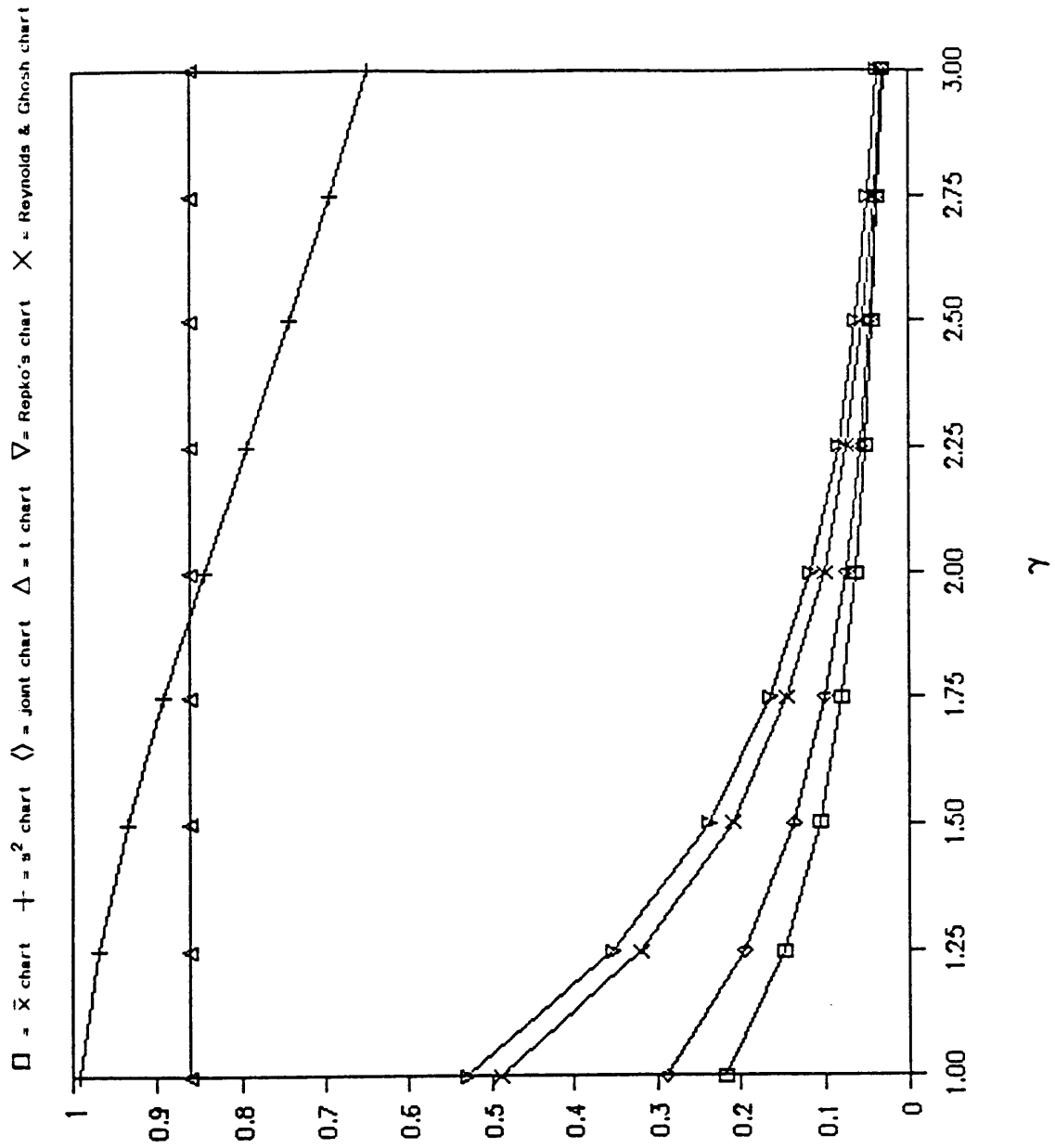


Figure 5.4
 Probability of a Type II Error
 $\alpha = .01, n = 5, \lambda = 1.5$

TABLE 5.5 PROBABILITY OF A TYPE II SAMPLING ERROR

$$\lambda = 2.0$$

		Gamma								
Chart	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	
\bar{X}	.0289	.0151	.0089	.0058	.0040	.0029	.0022	.0018	.0014	
s^2	.9900	.9688	.9351	.8921	.8437	.7933	.7431	.6946	.6485	
Joint	.0471	.0241	.0138	.0086	.0056	.0039	.0028	.0020	.0015	
T	.5489	.5489	.5489	.5489	.5489	.5489	.5489	.5489	.5489	
R&G	.1418	.0655	.0332	.0182	.0107	.0067	.0044	.0030	.0021	
Repko	.1671	.0787	.0403	.0223	.0131	.0082	.0054	.0037	.0026	

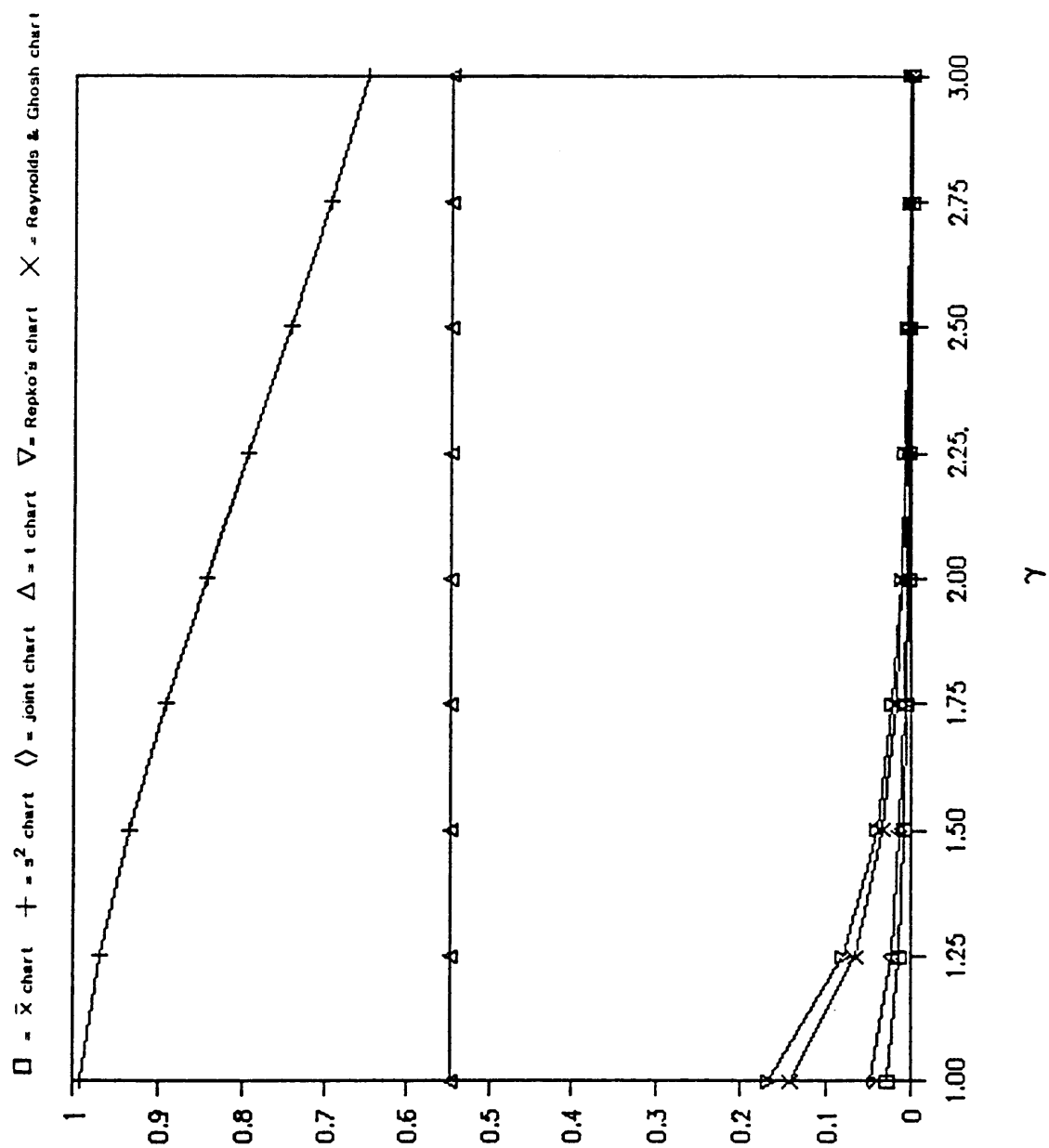


Figure 5.5
 Probability of a Type II Error
 $\alpha = .01, n = 5, \lambda = 2.0$

TABLE 5.6 PROBABILITY OF A TYPE II SAMPLING ERROR

$$\lambda = 0$$

Chart	Gamma								
	1.00*	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
\bar{X}	.9500	.9204	.8904	.8615	.8341	.8086	.7847	.7626	.7420
s^2	.9500	.8922	.8239	.7533	.6854	.6226	.5656	.5145	.4690
Joint	.9500	.8936	.8242	.7505	.6781	.6103	.5486	.4931	.4438
T	.9500	.9500	.9500	.9500	.9500	.9500	.9500	.9500	.9500
R&G	.9500	.8850	.8061	.7242	.6459	.5743	.5104	.4542	.4051
Repko	.9732	.9283	.8663	.7961	.7245	.6556	.5917	.5337	.4817

$$\alpha = .05, n = 5$$

*This is the in control case. The probability of no signal is $1 - \alpha$.

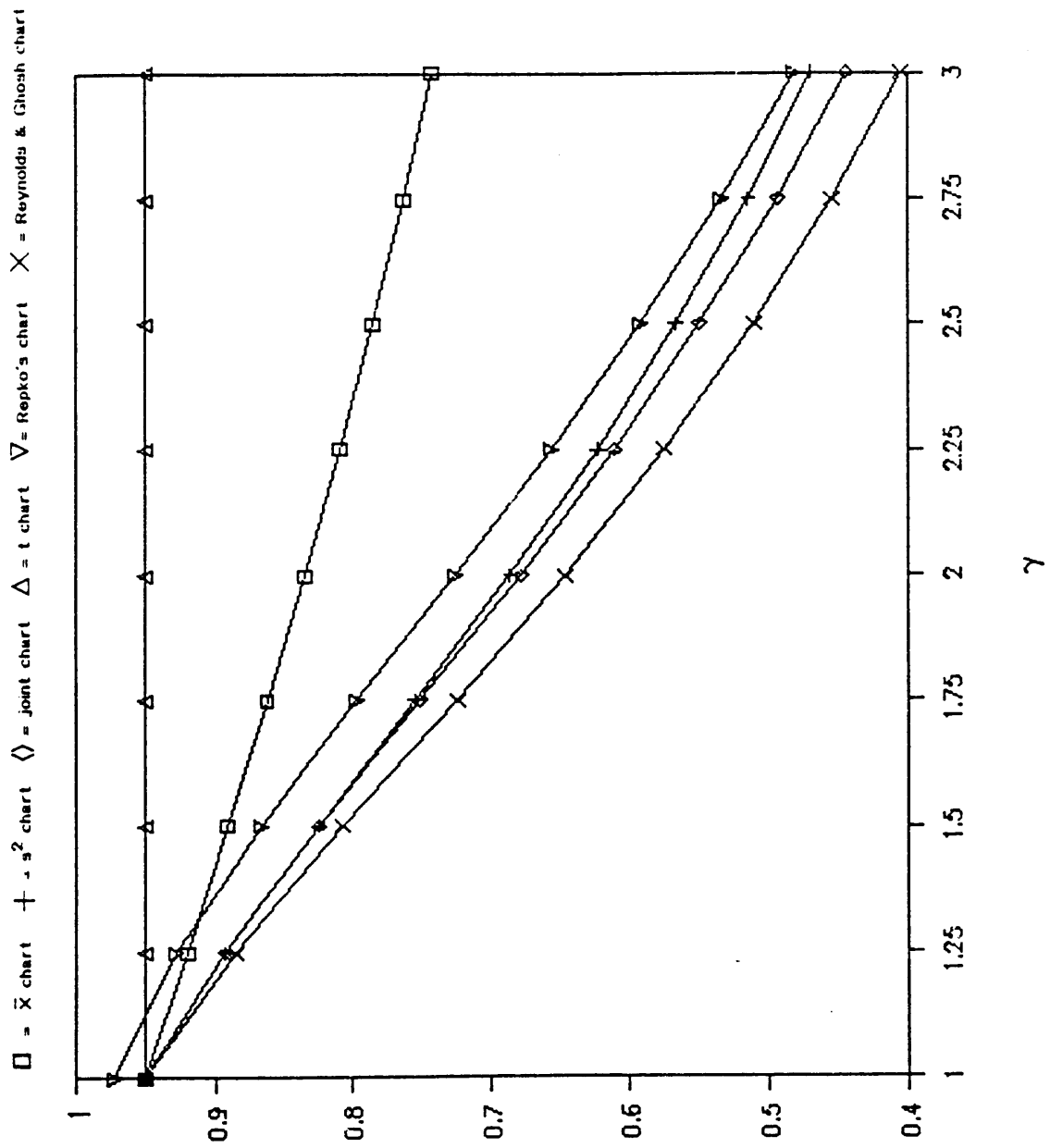


Figure 5.6
 Probability of a Type II Error
 $\alpha = .05, n = 5, \lambda = 0.0$

TABLE 5.7 PROBABILITY OF A TYPE II SAMPLING ERROR

$$\lambda = 0.5$$

		Gamma								
Chart	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	
\bar{X}	.7989	.7351	.6818	.6371	.5993	.5670	.5390	.5146	.4930	
s^2	.9500	.8922	.8239	.7533	.6854	.6226	.5656	.5145	.4690	
Joint	.8462	.7588	.6712	.5896	.5166	.4528	.3977	.3504	.3098	
T	.9049	.9049	.9049	.9049	.9049	.9049	.9049	.9049	.9049	
R&G	.8876	.7864	.6830	.5883	.5059	.4360	.3772	.3279	.2866	
Repko	.9311	.8513	.7604	.6708	.5886	.5158	.4526	.3983	.3517	

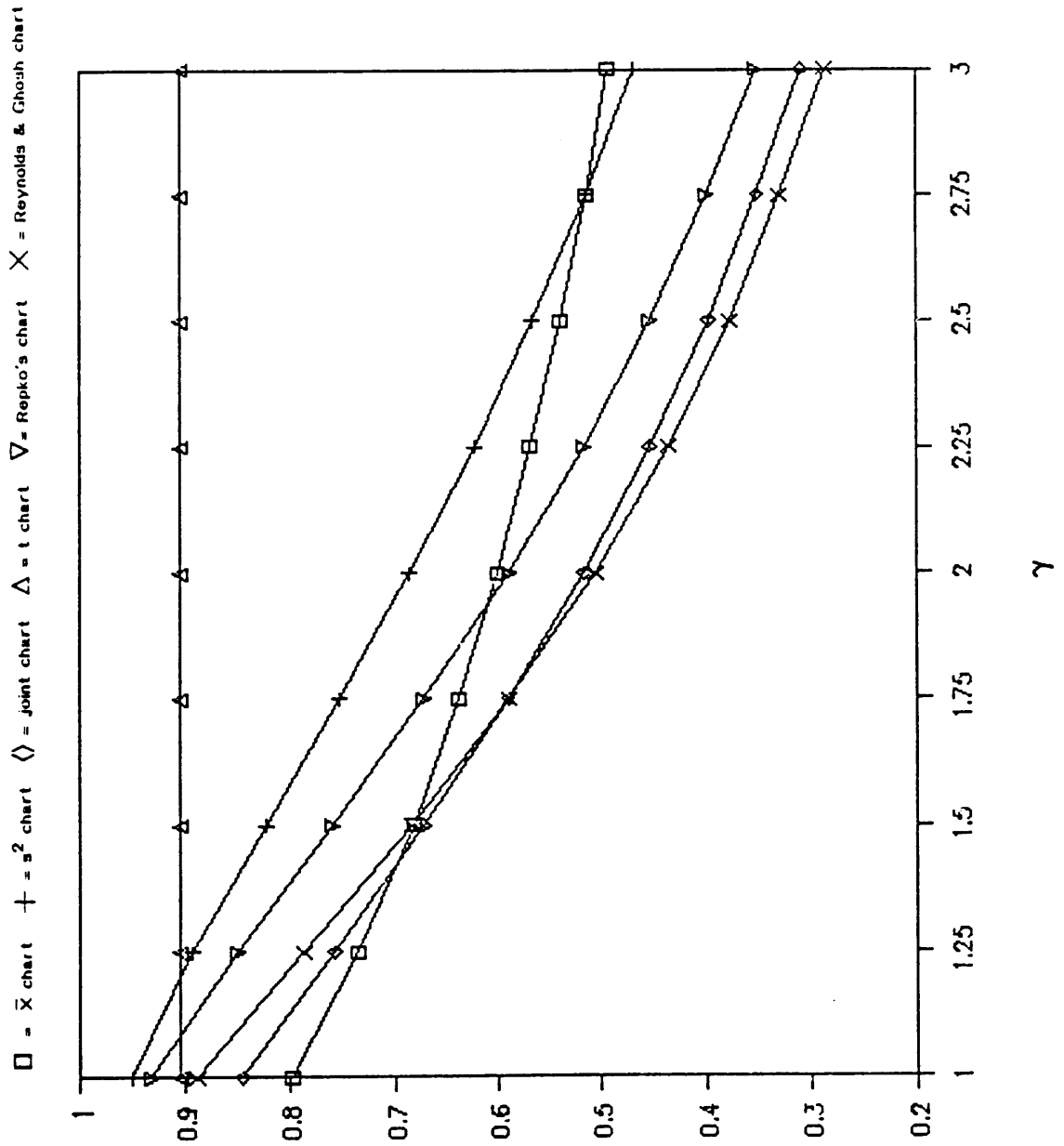


Figure 5.7
Probability of a Type II Error
 $\alpha = .05, n = 5, \lambda = 0.5$

TABLE 5.8 PROBABILITY OF A TYPE II SAMPLING ERROR

$$\lambda = 1.0$$

		Gamma							
Chart	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
\bar{X}	.3914	.3147	.2625	.2253	.1976	.1762	.1593	.1457	.1344
s^2	.9500	.8922	.8239	.7533	.6854	.6226	.5656	.5145	.4690
Joint	.4880	.3815	.3020	.2421	.1965	.1614	.1340	.1125	.0953
T	.6873	.6873	.6873	.6873	.6873	.6873	.6873	.6873	.6873
R&G	.6373	.4811	.3643	.2796	.2182	.1732	.1396	.1142	.0946
Repko	.7288	.5744	.4487	.3524	.2798	.2251	.1834	.1512	.1262

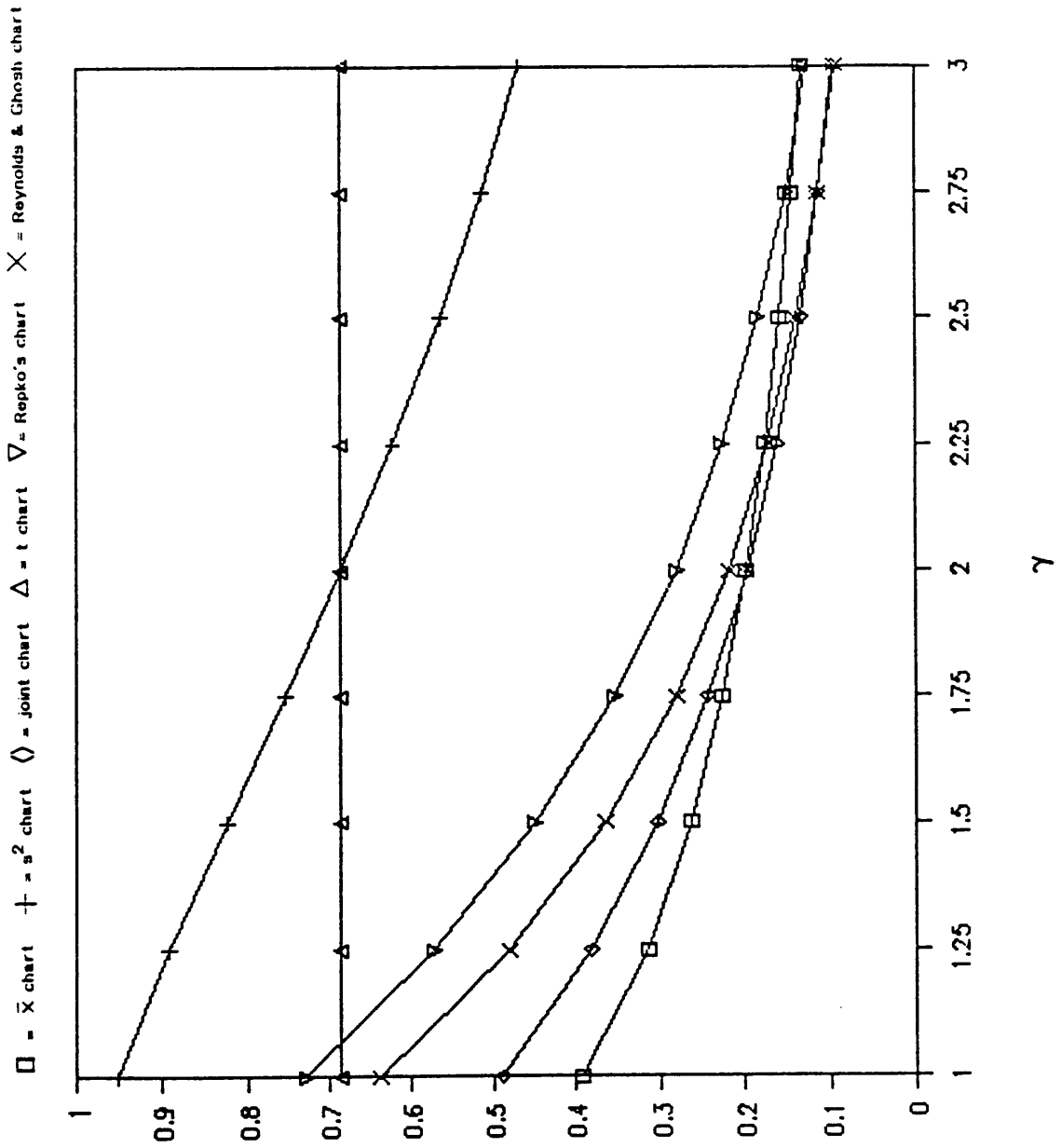


Figure 5.8
Probability of a Type II Error
 $\alpha = .05, n = 5, \lambda = 1.0$

TABLE 5.9 PROBABILITY OF A TYPE II SAMPLING ERROR

$$\lambda = 1.5$$

		Gamma								
Chart	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	
\bar{X}	.0817	.0547	.0398	.0306	.0245	.0203	.0172	.0149	.0131	
s^2	.9500	.8922	.8239	.7533	.6854	.6226	.5656	.5145	.4690	
Joint	.1290	.0826	.0561	.0398	.0293	.0221	.0171	.0135	.0108	
T	.2953	.2953	.2953	.2953	.2953	.2953	.2953	.2953	.2953	
R&G	.2630	.1535	.0948	.0618	.0421	.0299	.0219	.0164	.0127	
Repko	.3510	.2138	.1355	.0898	.0619	.0442	.0325	.0245	.0189	

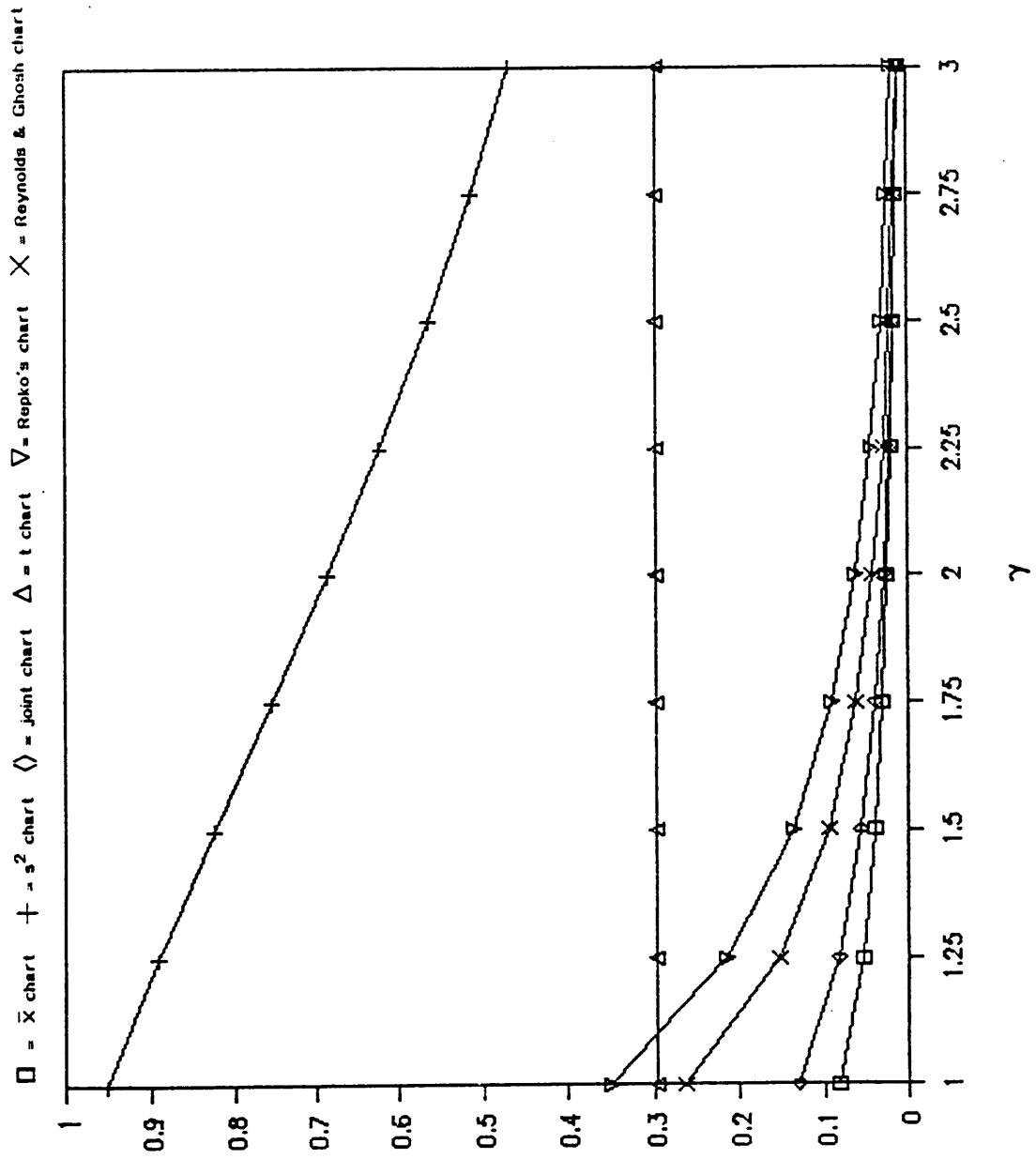


Figure 5.9
Probability of a Type II Error
 $\alpha = .05, n = 5, \lambda = 1.5$

TABLE 5.10 PROBABILITY OF A TYPE II SAMPLING ERROR

$$\lambda = 2.0$$

Gamma									
Chart	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
\bar{X}	.0060	.0033	.0020	.0014	.0010	.0008	.0006	.0005	.0004
s^2	.9500	.8922	.8239	.7533	.6854	.6226	.5656	.5145	.4690
Joint	.0124	.0063	.0036	.0022	.0015	.0010	.0007	.0005	.0004
T	.0816	.0816	.0816	.0816	.0816	.0816	.0816	.0816	.0816
R&G	.0477	.0202	.0098	.0053	.0031	.0019	.0012	.0008	.0006
Repko	.0777	.0340	.0167	.0091	.0053	.0033	.0021	.0015	.0010

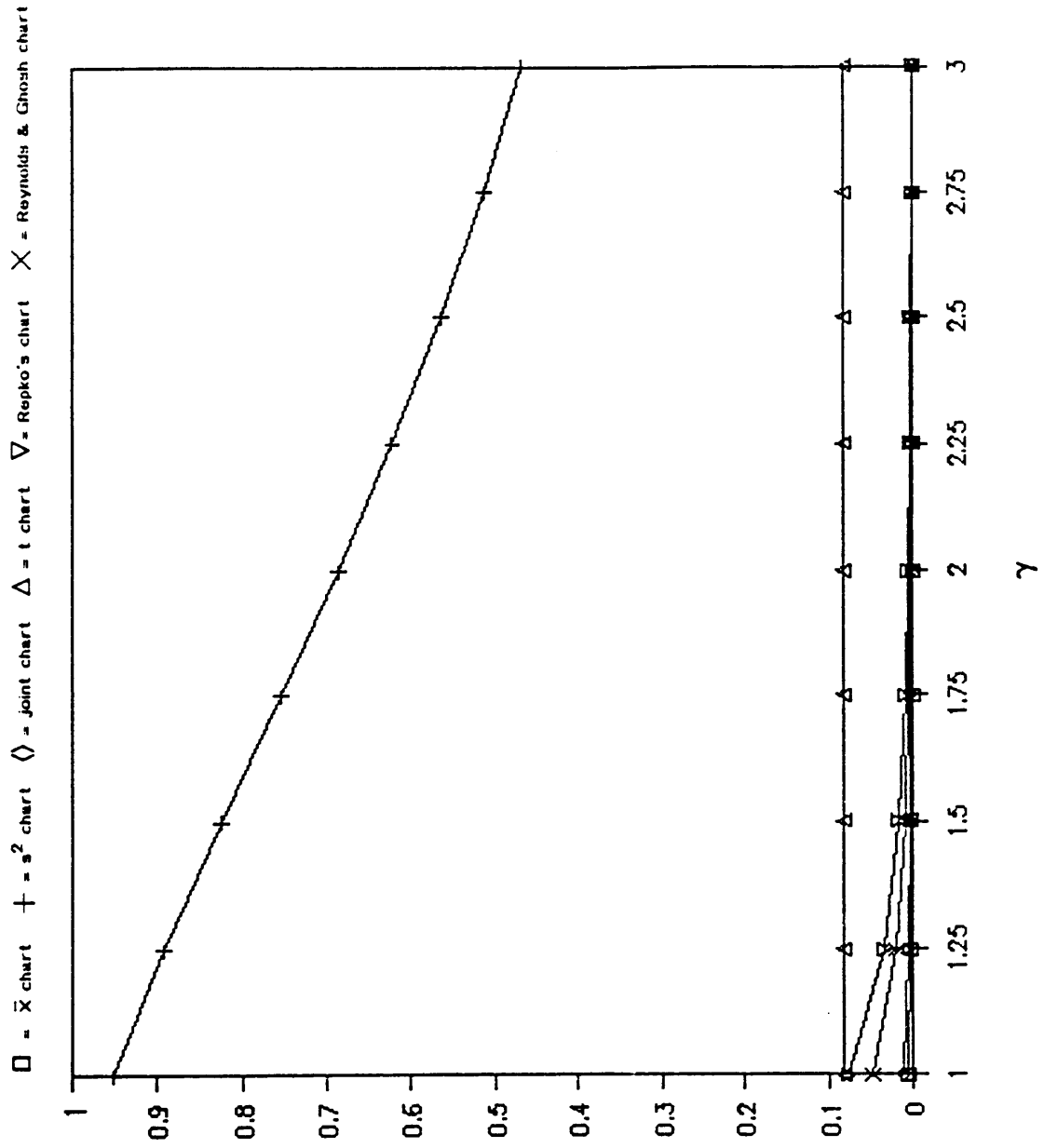


Figure 5.10
Probability of a Type II Error
 $\alpha = .05, n = 5, \lambda = 2.0$

In the case that the relative shift in the mean is 50% , the \bar{X} chart is most powerful if the shift in the variance is not more than 50%. Any higher shift in the variance causes the Reynolds and Ghosh statistic to become the most powerful control chart.

As the shift in the mean increases, the Reynolds and Ghosh statistic becomes less powerful compared to other charts. When $\lambda = 1.0$ and the variance increases upto 200%, the \bar{X} chart is most powerful. For the case where the γ is 2.25 and 2.50, the joint control chart is the best method to use. However, after γ exceeds 2.75, the Reynolds and Ghosh chart once again is the most powerful. The effect on increasing α to .05 is that the joint chart becomes most powerful for all high shifts in the variance except for when $\gamma = 3$. In this case, the Reynold and Ghosh is most powerful.

In the case $\lambda = 1.5$, the \bar{X} chart is most powerful except where γ is 3.00. For that situation, the joint control chart is more powerful. When $\alpha = .05$, the joint becomes the most powerful chart when $\gamma = 2.50$.

When $\lambda = 2.0$, regardless of the magnitude of the shift in the variance, the \bar{X} chart is the most powerful tool to use.

As can be seen from Table 5.1, Repko's chart does not have the dictated alpha of .01. The control limit for the Repko's plot is $\chi_{2,.99}^2$ but due to the approximations made related to the sample size, the actual alpha is not .01. The alpha presented in Table 5.1 is the lower bound on the actual alpha. Two values of the actual alpha are computed: the alpha resulting from the expected value of the term U^2 and the lower bound of alpha resulting from the maximum of U^2 . The first alpha is .0083 and the

lower bound alpha is .0072. The corresponding Type II error probabilities differ at the most by .0005. Therefore, only the results from the lower bound of alpha are presented.

If the control limit for Repko's chart could be calculated exactly so that the probability of a Type I error were .01, the Type II error probabilities would be identical to the results of Reynolds and Ghosh. This can be seen from [19] and [20]:

$$\beta_{r,g}(\mu_1, \sigma_1) = P(\chi^2_{(n,c)} < (\chi^2_{n,1-\alpha})/\gamma, / \mu = \mu_1, \sigma = \sigma_1)$$

$$c = n\lambda^2 \quad [19]$$

$$\beta_{rep}(\mu_1, \sigma_1, U^2) = P(\chi^2_{(n,c)} < T/\gamma, / \mu = \mu_1, \sigma = \sigma_1)$$

$$c = n\lambda^2$$

$$T = (n-1) \left[\left(\frac{\chi^2_{2,1-\alpha} - U^2}{2n} \right)^{1/2} + 1 \right]^2 + U^2$$

$$U_e^2 = 1, \quad U_{max}^2 = \chi^2_{2,1-\alpha} - \left[\frac{n-1}{n+1} \right]^{2(2n)} \quad [20]$$

To obtain the correct alpha for Repko's plot, T must equal $\chi^2_{n,1-\alpha}$. Therefore [20] becomes:

$$\beta_{rep}(\mu_1, \sigma_1, U^2) = P(\chi^2_{(n,c)} < (\chi^2_{n,1-\alpha})/\gamma, / \mu = \mu_1, \sigma = \sigma_1)$$

$$c = n\lambda^2$$

If the control limit for Repko's chart is set so that $T = \chi^2_{n,1-\alpha}$, the probability of a Type II error for the Reynolds and Ghosh statistic is the same as the probability of a Type II error for Repko's plot.

Tables 5.11 and 5.12 show the most powerful chart for each out-of-control case examined. The \bar{X} chart, the joint chart, and the Reynolds and

Ghosh chart perform the best. The Reynolds and Ghosh statistic is the most powerful when:

- * The mean does not shift and the variance does.
- * The mean shifts slightly and the variance shifts greatly.

The \bar{X} chart is the most powerful when:

- * The mean shifts and the variance does not.
- * The mean shifts and the variance shifts slightly.

The joint chart is most powerful for the cases between the \bar{X} chart and the Reynolds and Ghosh chart:

- * The mean shifts a moderate amount and so does the variance.
- * The mean shifts a moderately large amount and the variance shifts a large amount.

5.3 CONCLUSION

Control charts portray two or three of the following dimensions:

- 1) location
- 2) variability
- 3) time

A comparison of charts has been made based on several out-of-control cases of the location and variability dimensions.

The \bar{X} chart plots the location statistic versus time. The s^2 chart plots a variability statistic against time. Often the \bar{X} and the s^2 charts are used together to control the mean and variance of a variable. However, the results

show that the \bar{X} chart alone is the most powerful tool for over 50% of the out-of-control cases examined.

The Reynolds and Ghosh chart is a method of combining the location and variability dimensions of a variable with the time element. Repko's plot is another method of combining the location and variability dimensions of a variable, but ignores time. Due to approximations made about the sample size, Repko's plot is not accurate for small sample sizes. However, the control limit can be altered and the statistical properties of the Reynolds and Ghosh statistic can be obtained. The advantage of the Repko plot is the graphical ease with which it can be used. It also has the advantage of direct application to process capability analysis.

The Reynolds and Ghosh statistic is the most powerful tool when the variance shifts with a small shift in the mean. The Reynolds and Ghosh is not sensitive to changes in the mean due to the compensation effect. This can be understood intuitively. The Reynolds and Ghosh statistic is the sum of two χ^2 s. If one stays in control, the other must become quite large to create a signal. When only the mean changes, the Reynolds and Ghosh statistic is not sensitive. However, a shift in the variance affects both terms, which is why the statistic is sensitive to changes in the variance.

The joint control chart combines independent limits for the mean and variance into one chart. No compensation effect between the two variables occurs because the two statistics are not being numerically combined. The time variable is ignored. The joint chart with a probability of a Type I error α has the statistical properties of the \bar{X} chart and the s^2 chart each with $\alpha' = 1 - (1-\alpha)^2$.

The joint chart is not the most powerful tool for all of the cases. However, when moderate changes in the mean and large changes in the variance occur, the joint chart is the best tool to use. The joint chart performs close to the \bar{X} chart in cases where the \bar{X} chart is the most powerful. The reason the joint chart is not more powerful than the \bar{X} chart is because when the probability of a Type I error decreases the probability of a Type II error increases. Therefore, the probability of a Type II error for the \bar{X} limits within the joint chart is larger than the probability of a Type II error for the \bar{X} . Note that the joint chart can be

defined to be closer to the limits of the \bar{X} chart. In this case, the joint chart would become closer in power to the \bar{X} chart.

The joint chart has a graphical advantage over the \bar{X} chart. It shows any correlation between the two statistically independent sample statistics \bar{X} and s^2 . If any correlation is observed, it would be a result of either the process or an outside influence. The correlation could provide useful information for future process control. Since the joint chart has the statistical properties of an \bar{X} chart and has a graphical advantage over the mean chart, it should be considered for use in cases where both the mean and variance will change.

The t-statistic chart controls the sample mean of a normally distributed variable without knowing the standard deviation. The t-statistic proves to be a very weak tool for controlling the mean of a variable. However, if nothing is known about the variability of the process, the t test may be used to control the mean while data is being collected about the variance.

In summary, the Reynolds and Ghosh sum statistic has been shown to be the most effective procedure to control changes in the variance of a variable. If Repko's circular limit can be defined for small sample sizes, it will have the same properties as the sum statistic. Another approach would be to utilize the sum statistic so that it may be plotted like Repko's.

For changes in the mean, the \bar{X} chart is sensitive if the change in the variance is small. The joint chart is powerful for cases in between the sum statistic and the \bar{X} chart. The joint chart has similar probabilities to

the \bar{X} chart and has the additional graphical information about the relationship between the mean and variability.

Several areas exist to be examined further. In this work, the probabilities of a Type I error for the limits within the joint chart were set to be equal. Since the \bar{X} chart is very powerful, it seems that the joint chart limits could be "biased" towards the \bar{X} limits. What would the criteria be for setting the limits? Second, a cost analysis of the joint chart could be developed. Third, further work could be done examining data patterns in the joint chart. Interpreting correlation between the mean and variance would lead to valuable knowledge about the production process. The recapture of the time dimension by numbering the points through time or some other method could also be explored.

BIBLIOGRAPHY

- Abramowitz, Milton and Stegun, Irene A. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Washington, D.C. : U.S. Government Printing Office, 1972.
- Chengalur, I.N., Arnold, J.C, and Reynolds, M.R. Jr. Variable Sampling Intervals For Multiparameter Shewhart Charts. Blacksburg, VA: Department of Statistics and Statistical Consulting Center, Virginia Polytechnic Institute and State University, 1987.
- Cowden, Dudley J. Statistical Methods in Quality Control. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1957.
- Duncan, Acheson J. Quality Control and Industrial Statistics. Illinois: Richard D. Irwin, 1986.
- Lancaster, H.O. The Chi-Squared Distribution. New York: John Wiley & Sons, Inc, 1969.
- Rao, Naresh. A Variable Sampling Interval Chart for a Combined Statistic. Virginia Polytechnic Institute and State University, July, 1988.
- Repko, Joop. "Process Capability Plot: Value to Customer and Manufacturer". 30th EOQC Conference. Stockholm, Sweden, pp. 373-382, 1986.
- Reynolds, M.R. and Ghosh, B.K. "Designing Control Charts for Means and Variances" ASQC Quality Congress Transactions. San Francisco, pp. 400-407, 1981.
- Shewhart, W.A. Statistical Method from the Viewpoint of Quality Control. Washington D.C. :Graduate School, U.S. Department of Agriculture, 1939.

APPENDIX A

COMPUTER PROGRAM

.....

- COMPARISON OF ALTERNATIVE METHODS TO THE SHEWHART TYPE CHART

- BY DEBORAH A. HALL

- SUPPORTING COMPUTER PROGRAM

- This program computes the probability of a Type II error for the following

- control charts:

- 1. a mean chart

- 2. a variance control chart

- 3. a joint control chart

- 4. a t-statistic chart

- 5. a Reynolds and Ghosh sum statistic

- 6. a Reynolds process capability plot

- The following variables are defined for the charts:

- N = sample size

- ALPHA = probability of a type I error

- LAMBDA = (population mean - target mean) / standard deviation

- GAMMA = population variance / target variance

- EONE = probability of a type II error for the mean control chart
- ETWO = probability of a type II error for the variance control chart
- ETHREE = probability of a type II error for the joint control chart
- EFOUR = probability of a type II error for the t-chart
- EFIVE = probability of a type II error for the Reynolds and Ghosh chart
- ESIX = probability of a type II error for the Repko plot

```
REAL K,N,ALPHA,LAMBDA,GAMMA,ONE,ETWO,ETHREE,EFOUR,EFIVE,ESIX
```

```
DIMENSION USQR(50),ESIX(10)
```

```
REAL JONE,JTWO
```

```
INTEGER O
```

```
OPEN(9,FILE='DATA',STATUS = 'NEW')
```

```
WRITE(*,*) 'Input the sample size.'
```

```
READ(*,*) N
```

```
WRITE(*,*) 'Input alpha, the probability of a type I error.'
```

```
READ (*,*) ALPHA
```

```
WRITE(9,19) INT(N), ALPHA
```

```
19 FORMAT(2X,'The sample size is',i4,/, 'Alpha is',f7.4,/,
```

```
$'LAMBDA GAMMA EONE ETWO ETHREE EFOUR EFIVE ESIX')
```

- XK = k used in limits of a mean chart given a desired alpha
- Given an alpha, k is found by: $\text{prob}(z > k) = \alpha/2$

```
WRITE(*,*) 'Input k so that  $\text{prob}(z > k) = \alpha/2$ '
```

```
READ(*,*) XK
```

- UCL = the upper control limit for a variance chart given an alpha

```
DF = N-1.
```

WRITE(*,*) 'Input the upper control limit for the variance chart'

READ(*,*) UCL

- TVALUE = limit for t chart

WRITE(*,*) 'Input the value for the t chart'

READ(*,*) TVALUE

- SCL = limit used in sum statistic chart

WRITE(*,*) 'Input the chi square for n d.f.'

READ(*,*) SCL

- CHI = limit used in Repko's chart

WRITE(*,*) 'Input the value of chi square for 2 d.f.'

READ(*,*) CHI

- K = k used in limits of joint control chart

ALPHA1 = 1. - (1. - ALPHA)**.5

CALL VNORMAL(ALPHA1/2.,K)

- CL = used in joint control chart

FIVE = 1. - ALPHA1

CALL VCHISQR(FIVE,DF,CL)

.....

- A do loop is performed. Lambda is incremented from 0 to 2 by .5. Gamma

- is incremented from 1 to 3 by .25.

.....

LAMBDA = -.5

DO 69 I = 1,5

LAMBDA = LAMBDA + .5

GAMMA = .75

DO 79 J = 1,9

GAMMA = GAMMA + .25

.....

• THE X CONTROL CHART

- ZONE and ZTWO are the expressions subtracted to compute the probability
- of a type II error, EONE. The function RNORMAL computes the cumulative
- probability of the normal distribution.

ZONE = (-LAMBDA*(n*.5) - ABS(XK)/(GAMMA*.5)

ZTWO = (-LAMBDA*(N*.5) + ABS(XK)/(GAMMA*.5)

EONE = RNORMAL(ZTWO) - RNORMAL(ZONE)

.....

• THE VARIANCE CONTROL CHART

- Given an alpha, the upper control limit is the chi-square value corresponding
- to alpha and the degrees of freedom, N - 1. The subroutine VCHISQR is used to
- find the upper limit.
- DF = degrees of freedom
- VALUEC = value used to compute probability of type II error
- ETWO = probability of a type II error

VALUEC = UCL/GAMMA

CALL CHISQR(DF,VALUEC,ETWO)

.....

• THE JOINT CONTROL CHART

- Given an alpha, the alphas for the two sets of limits has to be computed.
- ALPHA1 = the alpha for the separate limits
- ALPHA = the overall alpha

- ONE = first term
- = probability that the mean doesn't signal
- TWO = second term multiplied in computing the probability of a type II error
- error
- = probability that the variance doesn't signal
- ETHREE = probability of a type II error

JONE = (-LAMBDA * (N** .5)) - (ABS(K)/GAMMA** .5)

JTWO = (-LAMBDA * (N** .5)) + (ABS(K)/GAMMA** .5)

ONE = RNORMAL(JTWO) - RNORMAL(JONE)

SIX = CL/GAMMA

CALL CHISQR(DF,SIX,TWO)

ETHREE = ONE*TWO

.....

- THE T CHART
- Given an alpha, the t-value is computed. The type II error probability then
- computed.
- TVALUE = absolute value of the upper and lower limits
- COMP = term used in computing type II error probability
- EFOUR = type II error probability

COMP = TVALUE - LAMBDA*(N** .5)

IF (COMP.GE.0.) THEN

CALL TSTAT(DF,COMP,AREA1)

PROB1 = 1. - ((1.-AREA1)/2.)

COMP2 = ABS(-TVALUE - LAMBDA*N** .5)

CALL TSTAT(DF,COMP2,AREA2)

```

PROB2 =(1.-AREA2)/2.

EFOUR = PROB1-PROB2

END IF

IF (COMP.LT.0) THEN

CALL TSTAT(DF,ABS(COMP),AREA1)

PROB1 = (1.-AREA1)/2.

COMP2 = ABS(-TVALUE-LAMBDA*N**.5)

CALL TSTAT(DF,COMP2,AREA2)

PROB2 = (1.-AREA2)/2.

EFOUR = PROB1-PROB2

END IF

```

.....

- THE SUM STATISTIC
- To compute the type II error probability for the sum statistic, the non-
- central chi-square distribution is used.
- SCL = chi-square value for alpha and n degrees of freedom

```

100 PARA = LAMBDA**2*N

TSUM = SCL/GAMMA

CALL CHINON(N,PARA,TSUM,EFIVE)

```

.....

- THE REPKO PROCESS CAPABILITY PLOT
- This subprogram computes the "true" alpha in Repko's process capability plot
- when the sample size is less than thirty. Contours are done with regard to
- values of U.

```

USQR(1) = 1.

```

```

USQR(2) = CHI - ((N-1.)/(N+1.))**.2.**2.**N

do 321 o = 1,2

• RTERM = chi-square value for alpha and n degrees of freedom

200 RTERM = (N-1.) * (((CHI-USQR(o))/(2.**N))**.5 + 1. )**2 + USQR(o)

CALL CHISQR(N,RTERM,REPKO)

B = 1. - REPKO

TREPKO = RTERM/GAMMA

PARA = LAMBDA**.2**N

CALL CHINON(N,PARA,TREPKO,ESIX(O))

321 CONTINUE

WRITE(6,49)LAMBDA,GAMMA,EFOUR,ESIX(1),ESIX(2)

WRITE(9,49)LAMBDA,GAMMA,EFOUR,ESIX(1),ESIX(2)

49 FORMAT(F4.2,7F6.4)

79 CONTINUE

69 CONTINUE

STOP

END

```

.....

• SUBROUTINES

.....

• THE NON-CENTRAL CHI-SQUARE DISTRIBUTION

• The non central chi-square distribution is calculated in the following

• subroutine. A infinite series expansion is used. The expansion is terminated

- after twenty-five terms. At that point, the sum is being increased by only
- .0000001 or less. If not, the sum is increased by five terms.
- The expansion is equation 26.4.25 on page 942 of Abramowitz and Stegun.
- V = degrees of freedom of noncentral
- P = non centrality parameter
- $VALUE$ = central chi-square value for alpha and ndf degrees of freedom

```

SUBROUTINE CHINON(V,P,VALUE,PROB)

DOUBLE PRECISION T1,TERM,SUM,FACT

SUM = 0.

111 N = 25

DO 100 J = 0, N

DF = V+2.*REAL(J)

CALL CHISQR(DF,VALUE,CHI)

T1 = EXP(-P/2.)

T2 = (P/2.)**J

TERM = (T1*T2) /FACT(J) * CHI

SUM = SUM + TERM

100 CONTINUE

PROB = SUM

if (term.gt..0000001) then

N = N + 5

GOTO 111

END IF

RETURN

END

```

DOUBLE PRECISION FUNCTION FACT(IT)

DOUBLE PRECISION EVEN

EVEN = 1

DO 30 I = 1,IT

EVEN = I * EVEN

30 CONTINUE

FACT = EVEN

RETURN

END

.....

- INVERSE INTERPOLATION
- These subroutines compute the value x from a known distribution associated
- with a known cumulative probability. (Reference: Random Variables in
- Digital Simulation by J.W. Schmidt)
- R = the probability known
- X_0 = the starting value for the search, let it be the expected value
- F = the cumulative probability
- S = the increment by which x will increase or decrease. S was calculated
- so that A is obtainable
- A = error in estimate, .0005
- D = indicator to increment or decrement the estimate of x

- U = upper limit used in interpolation
- L = lower limit used in interpolation
- X = the value sought

- The Normal Distribution

```
SUBROUTINE VNORMAL(R,X)

REAL L

INTEGER D

X0 = -2.5

A = .00005

F = RNORMAL(X0)

S = .0001

QUESTION = ABS(R-F)

IF(QUESTION.LT.A) THEN

GO TO 15

END IF

IF ((R-F).LT.0.0) THEN

D = -1

U = X0

FU = F

END IF

IF ((R-F).GT.0.0) THEN

D = 1

L = X0

FL = F
```

```
END IF

5  X1 = X0 + D*S

F1 = RNORMAL(X1)

QUESTION = (R-F1)

IF(ABS(QUESTION).LT.A) THEN

GO TO 14

END IF

IF(D.EQ.-1 .AND. QUESTION.LT.0) THEN

U = X1

FU = F1

X0 = X1

GO TO 5

END IF

IF(D.EQ.1 .AND.QUESTION.GT.0) THEN

L = X1

FL = F1

X0 = X1

GO TO 5

END IF

IF(D.EQ.-1 .AND.QUESTION.GT.0) THEN

L = X1

FL = F1

END IF

IF(D.EQ.1 .AND.QUESTION.LT.0) THEN

U = X1
```



```
FU = F1  
END IF  
  
11 X0 = (U*(R-FL) + L*(FU-R))/(FU-FL)  
  
F = RNORMAL(X0)  
  
QUEST = (R-F)  
  
IF(ABS(QUEST).LT.A) THEN  
  
GO TO 15  
  
END IF  
  
IF(QUEST.LT.0.) THEN  
  
U = X0  
  
FU = F  
  
END IF  
  
IF(QUEST.GT.0.) THEN  
  
L = X0  
  
FL = F  
  
END IF  
  
GO TO 11  
  
14 X0 = X1  
  
15 X = X0  
  
RETURN  
  
END
```

• The Chi-square Distribution

```
SUBROUTINE VCHISQR(R,DF,X)

REAL L

INTEGER D

X0 = DF

A = .00005

CALL CHISQR(DF,X0,F)

S = .0005

QUESTION = ABS(R-F)

IF(QUESTION.LT.A) THEN

GO TO 15

END IF

IF ((R-F).LT.0.0) THEN

D = -1

U = X0

FU = F

END IF

IF ((R-F).GT.0.0) THEN

D = 1

L = X0

FL = F

END IF

5 X1 = X0 + D*S

CALL CHISQR(DF,X1,F1)

QUESTION = (R-F1)
```

```
IF(ABS(QUESTION).LT.A) THEN

GO TO 14

END IF

IF(D.EQ.-1 .AND. QUESTION.LT.0) THEN

U = X1

FU = F1

X0 = X1

GO TO 5

END IF

IF(D.EQ.1.AND.QUESTION.GT.0) THEN

L = X1

FL = F1

X0 = X1

GO TO 5

END IF

IF(D.EQ.-1.AND.QUESTION.GT.0) THEN

L = X1

FL = F1

END IF

IF(D.EQ.1.AND.QUESTION.LT.0) THEN

U = X1

FU = F1

END IF

11 X0 = (U*(R-FL) + L*(FU-R))/(FU-FL)

CALL CHISQR(DF,X0,F)
```

```
QUEST = (R-F)

IF(ABS(QUEST).LT.A) THEN

    GO TO 15

END IF

IF(QUEST.LT.0.) THEN

U = X0

    FU = F

END IF

IF(QUEST.GT.0.) THEN

L = X0

    FL = F

END IF

GO TO 11

14 X0 = X1

15 X = X0

RETURN

END
```

- The T Distribution

```
SUBROUTINE VTSTAT(R,DF,X)

REAL L

INTEGER D

X0 = 1.282

A = .0005
```

```
CALL TSTAT(DF,X0,F)

S = .001

QUESTION = ABS(R-F)

IF(QUESTION.LT.A) THEN

GO TO 15

END IF

IF ((R-F).LT.0.0) THEN

D = -1

U = X0

FU = F

END IF

IF ((R-F).GT.0.0) THEN

D = 1

L = X0

FL = F

END IF

5 X1 = X0 + D*S

CALL TSTAT(DF,X1,F1)

QUESTION = (R-F1)

IF(ABS(QUESTION).LT.A) THEN

GO TO 14

END IF

IF(D.EQ.-1 .AND. QUESTION.LT.0) THEN

U = X1

FU = F1
```

X0 = X1

GO TO 5

END IF

IF(D.EQ.1.AND.QUESTION.GT.O) THEN

L = X1

FL = F1

X0 = X1

GO TO 5

END IF

IF(D.EQ.-1.AND.QUESTION.GT.O) THEN

L = X1

FL = F1

END IF

IF(D.EQ.1.AND.QUESTION.LT.O) THEN

U = X1

FU = F1

END IF

11 X0 = (U*(R-FL) + L*(FU-R))/(FU-FL)

CALL TSTAT(DF,X0,F)

QUEST = (R-F)

IF(ABS(QUEST).LT.A) THEN

GO TO 15

END IF

IF(QUEST.LT.O.) THEN

U = X0

```

      FU = F
END IF

      IF(QUEST.GT.0.) THEN

L = XO

      FL = F
END IF

      GO TO 11

      14 X0 = X1

      15 X = X0

      RETURN

      END

```

.....

- THE T-DISTRIBUTION
- This subroutine computes the probability, TPROB, that the absolute value
- of a t-statistic is less than the value T- the probability that t is between
- -T and + T. The degrees of freedom are V. The probability is computed from
- a series expansion. The equation is found on page 948 of Abramowitz and
- Stegun, eqs. 26.7.3 and 26.7.4. Three cases are examined: the degrees of
- freedom are even, odd and not equal to 1, or equal to 1.
-
- TSTAT = subroutine name
- V = degrees of freedom
- S = value for which probability is computed for
- TPROB = probability

.....

SUBROUTINE TSTAT(V,S,TPROB)

INTEGER DENOM

DIMENSION TERM(100)

SUM = 0.0

PI = 2.*ASIN(1.)

THETA = ATAN(S/(V*.5))

.....

• Case 1: degrees of freedom are odd and greater than 1

.....

IF((INT(V/2.).NE.V/2. .and. V.NE.1.) THEN

DO 110 I = 1,INT(V-2.),2

NUM = 1

DENOM = 1

DO 120 K = 2, I-1,2

NUM = NUM * K

120 CONTINUE

DO 130 J = 1,I,2

DENOM = DENOM * J

130 CONTINUE

TERM(I) = COS(THETA) ** (V-2.) * REAL(NUM)/REAL(DENOM)

SUM = SUM + TERM(I)

110 CONTINUE

TPROB = 2./PI*(THETA + SIN(THETA)*SUM)

END IF

.....

• Case II: 1 degree of freedom

.....

IF(INT(V).EQ.1) THEN

TPROB = 2./PI * THETA

END IF

.....

• Case III: degrees of freedom are even

.....

IF(INT(V/2.).EQ.V/2.) THEN

DO 140 I = 0, INT(V-2.),2

NUM = 1

DENOM = 1

DO 150 J = 1, I - 1, 2

NUM = NUM * J

150 CONTINUE

DO 160 K = 2, I, 2

DENOM = DENOM * K

160 CONTINUE

TERM(I) = COS(THETA)**I * REAL(NUM)/REAL(DENOM)

SUM = SUM + TERM(I)

140 CONTINUE

TPROB = SIN(THETA)*SUM

END IF

RETURN

END

.....

• THE CHI-SQUARE DISTRIBUTION

•

• The subroutine CHISQR computes the probability, CPROB, of being less than the

• value CHI2. The probability distribution is chi-square with V degrees of

• freedom. The two series expansions for the cumulative density function

• are found on page 941 of The Handbook of Mathematical Functions, edited by

• M. Abramowitz and I.E. Stegun. The first expansion, (Eq. 26.4.4, p.941),

• is used when the degrees of freedom are odd. The second expansion,

• (Eq. 26.4.5, p.941), is used when the degrees of freedom are even.

•

• CHISQR = subroutine name

• P = degrees of freedom

• CHI2 = value probability is computed for

• PROB = probability

.....

SUBROUTINE CHISQR(P,CHI2,PROB)

INTEGER V

DOUBLE PRECISION E1,E2,CZ,TERM,SUM,ETERM,ESUM,CPROB

CQ = 1.-RNORMAL(CHI2*.5)

PI = 2.*ASIN(1.)

CZ = 1./(2.*PI)*.5*EXP(-CHI2/2.)

V = NINT(P)

.....

• Case I: degrees of freedom are odd

.....

IF(INT(V/2).NE.V/2.) THEN

SUM = 0.

DO 100 I = 1, INT((V-1.)/2.)

tDENOM = 1.

DO 200 J = 1,2*I-1,2

tDENOM = tDENOM * real(J)

200 CONTINUE

E1 = (CHI2*.5)**(2*I-1)

TERM = E1/tDENOM

SUM = SUM + TERM

100 CONTINUE

CPROB = 2.*CQ + 2.*CZ*SUM

END IF

.....

• Case II: degrees of freedom are even

.....

```

IF(INT(V/2).EQ.V/2.) THEN

  ESUM = 0.

  DO 300 K = 1, INT((V-2.)/2.)

    DENOM = 1.

    DO 400 L = 2,2*K,2

      DENOM = DENOM * real(L)

400  CONTINUE

    E2 = (CHI2**5)**(2*K)

    ETERM = E2/DENOM

    ESUM = ESUM + ETERM

300  CONTINUE

  CPROB = (2.*PI)**.5*CZ*(1.+ESUM)

  END IF

  PROB = 1. - CPROB

  IF(PROB.GT.1.) THEN

    PROB = 1.

  END IF

  RETURN

  END

```

.....

• THE NORMAL DISTRIBUTION

•This function computes the cumulative density function value, RNORMAL,

- of the Z value X. An approximation is used, with error of less than 8
- decimal places. The equation used can be found on page 932 of The Handbook
- of Mathematical Functions, edited by Abramowitz and Stegun, equation 26.2.17.
- The approximation is to be used for only positive values of Z. However,
- due to the symmetry of the normal distribution, negative values of Z are easy
- to compute.
-
- RNORMAL = name of function, probability
- X = value probability is computed for

.....

FUNCTION RNORMAL(X)

DOUBLE PRECISION T,PROB,Z,PI

N = 0

.....

- If the z coordinate is negative, it is changed to positive. At the end of
- the subprogram, the probability computed is subtracted from 1 so that the
- answer is correct for the negative z .

.....

IF (X.LT.0.) THEN

X = - X

N = 1

END IF

PI = 2. * ASIN(1.)

Z = 1./(2.*PI)**.5 * EXP(-X*X/2.)

T = 1./(1. + (.2316419 * X))

B1 = .3198381530

B2 = -.356563782

B3 = 1.781477937

B4 = -1.821255978

B5 = 1.330274429

PROB = 1. - Z * (B1*T+B2*T**2.+B3*T**3.+B4*T**4.+B5*T**5.)

IF (N.EQ.1) THEN

PROB = 1. - PROB

X = - X

END IF

RNORMAL = PROB

RETURN

END

APPENDIX B

EQUATIONS USED IN COMPUTER PROGRAM

THE NORMAL DISTRIBUTION

The cumulative density function $[P(x)]$ for the standard normal is not in a closed form. However, several approximations exist [Abramowitz and Stegun 1972 p.932]. The approximation with the smallest error is:

$$P(x) = 1 - Z(x)(b_1t + b_2t^2 + b_3t^3 + b_4t^4 + b_5t^5) + \epsilon(x)$$

$$t = \frac{1}{1 + px}$$

$$\epsilon(x) < 7.5 \times 10^{-8}$$

$$p = .2316419$$

$$b_1 = .319381530$$

$$b_2 = -.356563782$$

$$b_3 = 1.781477937$$

$$b_4 = -1.821255978$$

$$b_5 = 1.330274429$$

$$Z = \text{density function of the normal distribution}$$

THE CHI-SQUARE DISTRIBUTION

The cumulative density function $[Q(\chi^2|\nu)]$ for the chi-square distribution has an exact series expansion [Abramowitz and Stegun 1972, p.941]. One expansion is used when the degrees of freedom are even, another when the degrees of freedom are odd.

$$Q(\chi^2|\nu) = 2Q(\chi) + 2Z(\chi) \sum_{r=1}^{\frac{\nu-1}{2}} \frac{\chi^{2r-1}}{1 \cdot 3 \cdot 5 \dots (2r-1)}$$

ν are odd, $\chi = \sqrt{\chi^2}$

$$Q(\chi^2|\nu) = \sqrt{2\pi} Z(\chi) \left[1 + \sum_{r=1}^{\frac{\nu-2}{2}} \frac{\chi^{2r}}{2 \cdot 4 \cdot 6 \dots (2r)} \right]$$

ν are even, $\chi = \sqrt{\chi^2}$

THE T DISTRIBUTION

The probability density function for the t-distribution has an exact series expansion [Abramowitz and Stegun 1972, p.948].

$$A(t/\nu) = \text{Prob}(|t_\nu| < t)$$

$$A(t/\nu) = \frac{2}{\pi} \left\{ \theta + \sin\theta \left[\cos\theta + \frac{2}{3}\cos^3\theta + \dots + \frac{2 \cdot 4 \dots (\nu-3)}{1 \cdot 3 \dots (\nu-2)} \cos^{\nu-2}\theta \right] \right\}$$

($\nu > 1$ and odd)

$$A(t/\nu) = \frac{2}{\pi} \theta$$

($\nu = 1$)

$$A(t/\nu) = \sin\theta \left[1 + \frac{1}{2}\cos^2\theta + \frac{1 \cdot 3}{2 \cdot 4}\cos^4\theta + \dots + \frac{1 \cdot 3 \cdot 5 \dots (\nu-3)}{2 \cdot 4 \cdot 6 \dots (\nu-2)} \cos^{\nu-2}\theta \right]$$

(ν even)

THE NON-CENTRAL CHI-SQUARE DISTRIBUTION

The cumulative density function $P(\chi'^2|\nu, \lambda)$ can be expressed as an infinite series expansion [Abramowitz and Stegun 1972, p.942]. The expansion is terminated when the term being added is less than 10^{-7} .

$$P(\chi'^2|\nu, \lambda) = \sum_{j=0}^{\infty} e^{-\lambda/2} \frac{(\lambda/2)^j}{j!} P(\chi'^2|\nu + 2j)$$

λ = the non-centrality parameter

**The vita has been removed from
the scanned document**