THE UNIVERSITY OF MANITOBA

STATISTICAL PROCESS CONTROL METHODOLOGIES FOR QUALITY IMPROVEMENT

by

Frederick Alfred Spiring

A Thesis

submitted to the Faculty of Graduate Studies

in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy

Department of Statistics
Winnipeg, Manitoba
May, 1988

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-47858-6

STATISTICAL PROCESS CONTROL METHODOLOGIES FOR QUALITY IMPROVEMENT

BY

FREDERICK ALFRED SPIRING

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

© 1988

Permission has been granted to the LIBRARY OF THE UNIVERSITY OF MANITOBA to lend or sell copies of this thesis, to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film, and UNIVERSITY MICROFILMS to publish an abstract of this thesis.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

ACKNOWLEDGMENTS

I would like to thank my advisors Drs. L. K. Chan and S. W. Cheng for their support and encouragement through my graduate studies. Their interest and confidence made some of the long hours bearable.

I would like to thank the University of Manitoba Fellowship program for their monetary awards.

I would especially like to thank my wife Britta for her support and for being such a good listener.

For the curious, this is the only mention of purple Smarties™ in my thesis.

Contents

ABSTRACT

A study conducted by the Ford Motor Company in late 1983 found only 50% of Ford suppliers' processes capable of meeting requirements. Common sense suggests that a process incapable of meeting requirements should not be used in the long run. Processes incapable of meeting requirements result in resources being allocated to the identification and repair/replacement of non-conforming output. Some aspects of process capability and its measure are examined.

Various properties of the process capability index are examined. The stochastic nature of the traditional estimator is stressed and analytical tools that promote stochastic interpretations and warnings are presented and discussed. The robustness of the traditional estimator with respect to departures from normality and a general procedure designed to detect departures from distributional assumptions are presented. As well, a Bayesian technique that alleviates some of the problems encountered in drawing stochastic inferences from the sampling results is suggested.

A measure of process capability is proposed that incorporates some of the new philosophies arising in quality control methodology. The new measure takes into consideration proximity to the target value. Some properties associated with a proposed estimator are presented and comparisons drawn among other competing measures. A multivariate analogue is presented and discussed as well.

Finally a graphical procedure for monitoring a process is presented. The procedure provides an alternative to the boxplot style of simultaneous control charting now being suggested in the literature and the traditional $\bar{x}$ and s control chart. It provides information regarding the process' proximity to the target value as well as the variability for both the univariate and multivariate cases.

Chapter 1
Introduction

1.1. OVERVIEW

Many early Quality Control (QC) procedures were developed for specific applications. These procedures often relied heavily on the intuition of the originator and the constraints associated with the problem being dealt with. Most of the early techniques have evolved, some through necessity others through general interest. Statistical theory and methods have had an impact on many techniques, however, there is still much that can be done.

In QC, the practicality of a method is as important as its mathematical elegance. Many methods are not the most statistically appealing, however a loss in efficiency is often sacrificed for administrative appeal. This is not to say advanced techniques should be disregarded. As practitioners become more sophisticated in their statistical and computing backgrounds, techniques once shunned as impractical will gain popularity. This is particularly true for those systems which are adaptable to a computer, where much of the more difficult material can be written into the operating system. Administrative ease and intuitive appeal must however be considered when proposing changes to existing procedures. QC procedures must be user friendly, for without co-operation and commitment, the most sophisticated technique will not be used. Most procedures will be administered by engineers and technicians not statisticians, hence discussions should be directed to them.

With this in mind two fundamental areas of statistical process control (SPC) are discussed. Current methods are examined and modifications suggested. The modifications are generally motivated from a statistical point of view but with attention paid to intuitive

and practical appeal. Emphasis has been placed on the study of process capability and simultaneous control charting.

1.2. Process Capability

Process capability has recently received a growing amount of attention. This new focus is due partly to the changing philosophy occurring in QC. Slogans such as "doing things right the first time" and "building a quality product" are good motivators but if a process is not capable of meeting requirements resources will be wasted. For example, if a mechanical process is not capable, the operators, regardless of their dedication and effort, will be unable to produce a quality product. Similarly if the operators are not capable of meeting the demands of the machinery a quality product will not result. Processes that are not capable regardless of their incapacity, waste resources.

Waste results from i) resources used to produce a non-conforming product, ii) the cost of identifying non-conforming product (either through inspection or customer dissatisfaction) and iii) repair/replacement of any non-conforming product. Some of these costs are tangible (such as repair/replacement) others (such as loss of business due to customer dissatisfaction) may be more difficult to quantify, but certainly exist.

Although it seems like common sense to use a process that is capable of meeting engineering requirements, it is not always the practice. In November 1983, Ford Motor Company reported that only 50% of those processes surveyed from suppliers with some sort of Quality Assurance program were capable! Regardless of how one defines process capability there is evidence that a problem exists.

In the past many companies have tried to inspect quality into the product. That is inspection teams were created whose role was to inspect the output for non-conforming product. This is both costly and in certain cases ineffective. By designing a process that is

capable and robust to input fluctuations, inspection becomes unnecessary. Once a process has been deemed capable, sampling is used only to monitor the procedure or to assess modifications made to the process. The ideal process would produce "identical" units under conditions which may include heterogeneous raw materials, different operators and a variety of operating conditions.

Process capability has become synonymous with process variability or process spread, while process capability indices relate process variability to the specification limits. There is however a growing demand to include proximity to the target value when considering process capability. Proximity to the target is part of the philosophy fostered by Dr. Genichi Taguchi.

Taguchi defines quality as the "the loss a product causes to society"[1] and promotes the use of a squared error loss function in assessing quality. Taguchi suggests such a loss function because it has zero loss only when the product is produced at the target value. As the product moves away from the target there is a loss in quality (Figure 1.2.1). Note that even small deviations from the target result in a loss of quality when using Taguchi's definition of quality. This approach to quality is substantially different from the classical approach where no loss in quality is assumed until the product is outside specifications (Figure 1.2.2).

[1] Introduction to Quality Engineering, by Genichi Taguchi (1986).

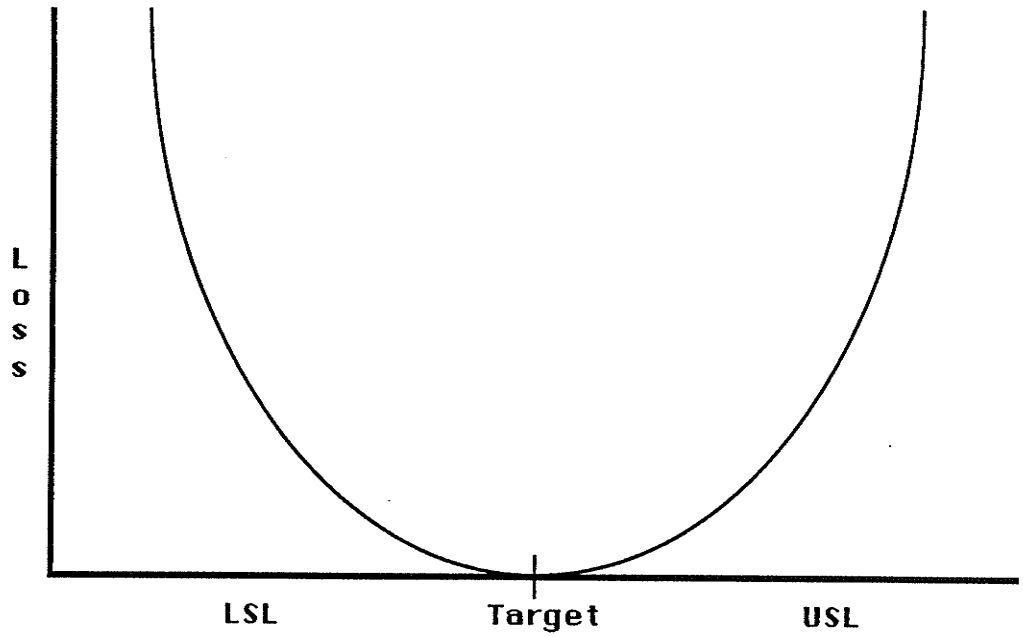


Figure 1.2.1. Squared Error Loss Function.

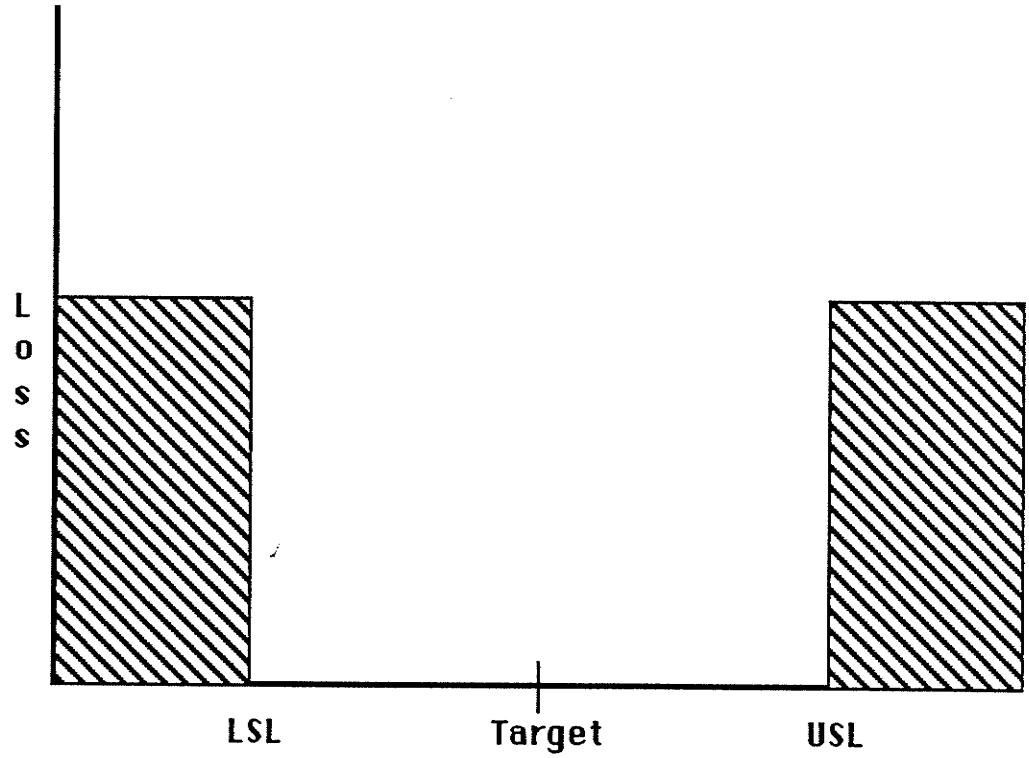


Figure 1.2.2. Traditional Loss Function.

Taguchi's loss function highlights the need to have small variability around the target. The best process in terms of Taguchi's definition of quality would be one that produced all its product at the target. When this is not the case the loss function suggests that both process variation and proximity to the target should be considered when assessing product quality. A process with all its product just inside the upper (or lower) specification may not be as desirable as a process with larger variation but centered on the target. Clearly the best process will be one that produces all of its product at the target, with the next best being the process with the smallest variability around the target. Figure 1.2.3 relates three populations with different levels of variation to the loss function.

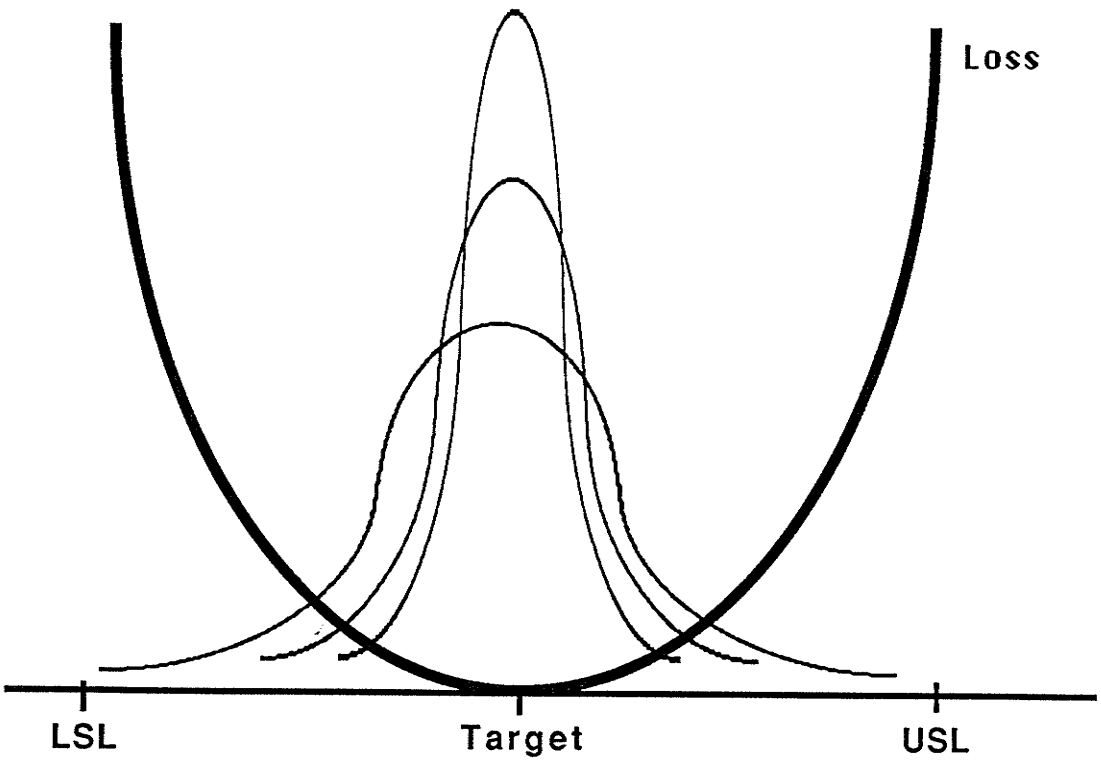


Figure 1.2.3 Loss Function with Three Normal Populations.

Changes in the definition of a quality product have forced changes in the procedures used to assess process capability. The ability of a process must now be measured in terms of process variability and proximity to the target. Small variability but not on target is just as undesirable as on target but with large variation.

Process capability has been defined in many ways and as a result several measures of process capability exist. The most common definition describes process capability as the range over which the output of a process varies. This quantity is also referred to as the actual process spread. Measures in this group depend upon the measuring units (i.e., meters, kilograms, ...) and hence do not encourage comparisons among processes with different quality characteristics.

The process capability index however relates the allowable process spread (usually an engineering requirement) to the actual process spread in the form of a ratio

$$\frac{\text{allowable process spread}}{\text{actual process spread}} .$$

The index will be unitless, thereby inviting comparisons among processes with different quality variables and promoting similar inferences regardless of the product or quality characteristic measured (i.e., widgets, televisions, ...). For example, an index value of one indicates that the allowable process spread is equivalent to the actual process spread. While a process capability index of two indicates that the allowable process spread is twice that of the actual process spread suggesting that the process is quite capable of producing within specifications (Figure 1.2.4).

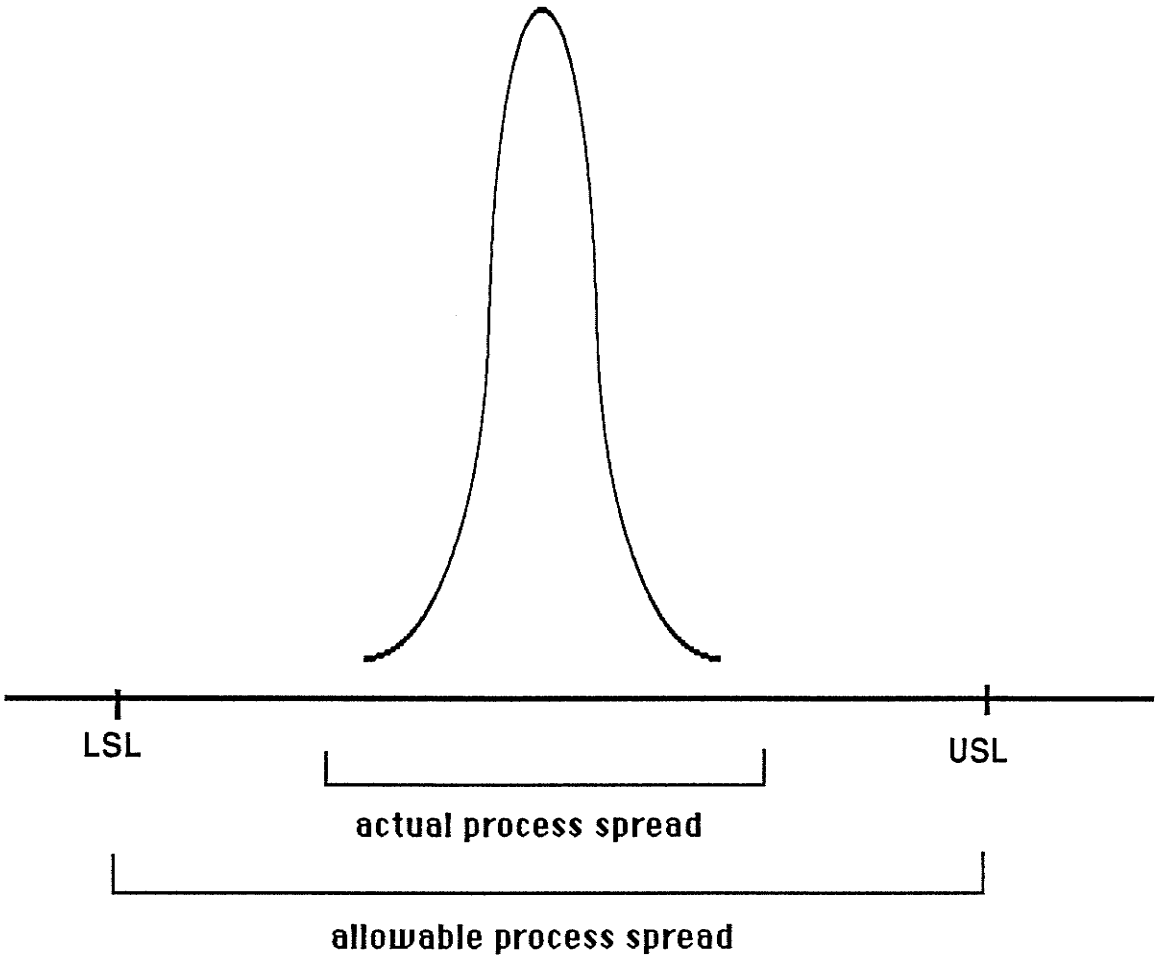


Figure 1.2.4. An Example of a Process with a Capability Index of 2.

Index values less than one indicate that the actual process spread is larger than the allowable process spread, suggesting that non-conforming product results (Figure 1.2.5).

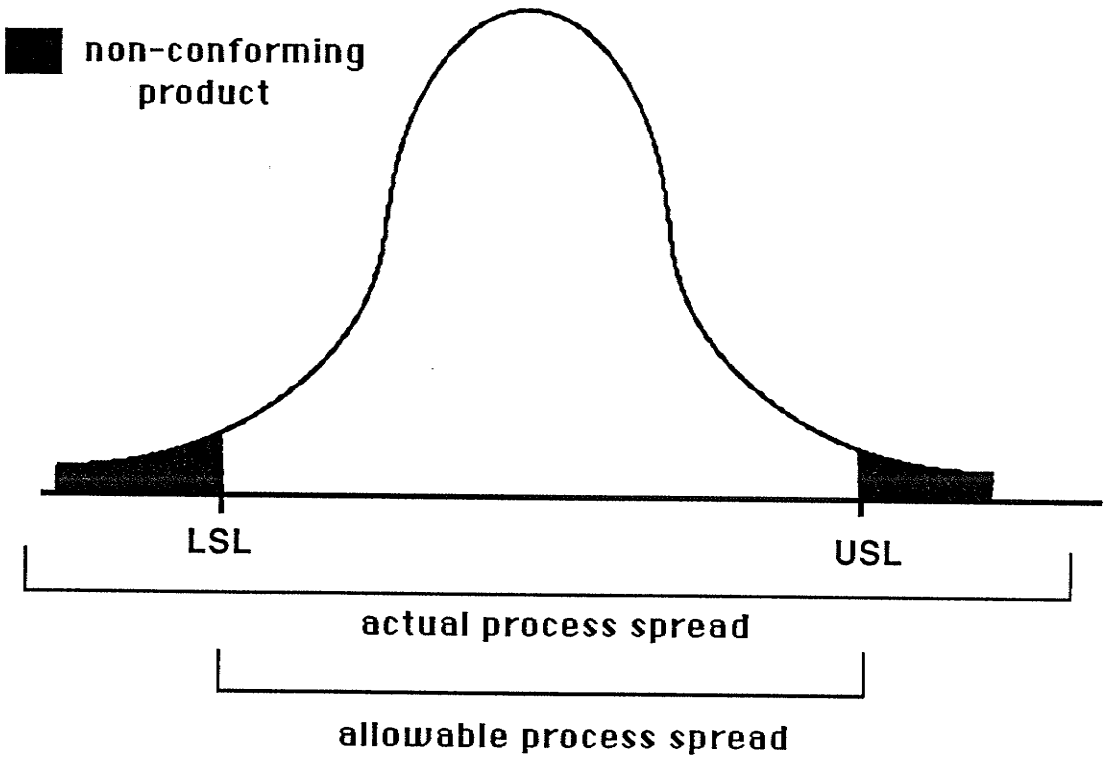


Figure 1.2.5 A Process with Actual Process Spread greater than Allowable Spread.

The actual process spread is generally taken to be 6σ which represents, in normal theory, the width of the interval that contains 99.73% of the population. The difference in the specification limits is used to indicate allowable process spread. The allowable process spread is considered fixed while the actual process spread in general must be estimated, hence the resultant measure of process capability will be stochastic.

It has become the practise among practitioners to ignore the stochastic nature of the estimated process capability index and to simply judge a process capable if the estimated process capability index is greater than one, and incapable if less than one. This is equivalent to drawing an inference from a point estimator without an accompanying confidence level.

There have been some recent attempts to rectify this unfortunate practise, most have met with limited success. The techniques suggested are statistically sound but lack administrative appeal. A Bayesian procedure is promoted here that has both administrative and intuitive appeal while providing sound statistical inferences.

A second practise whose abuse is not unique to QC practitioners is the failure to consider effects of non-normality on inferences drawn. The process capability index relies heavily on the assumption of normality. Clearly 6σ may not represent an interval that will contain 99.73% of the measurements outside of the normal family. However little or no warning is made regarding non-normal population characteristics.

For example Taguchi in the now famous example where a quality characteristic of televisions manufactured in Japan and the United States are compared, fosters the concept of comparing a normal distribution and an uniform distribution (Figure 1.2.6).

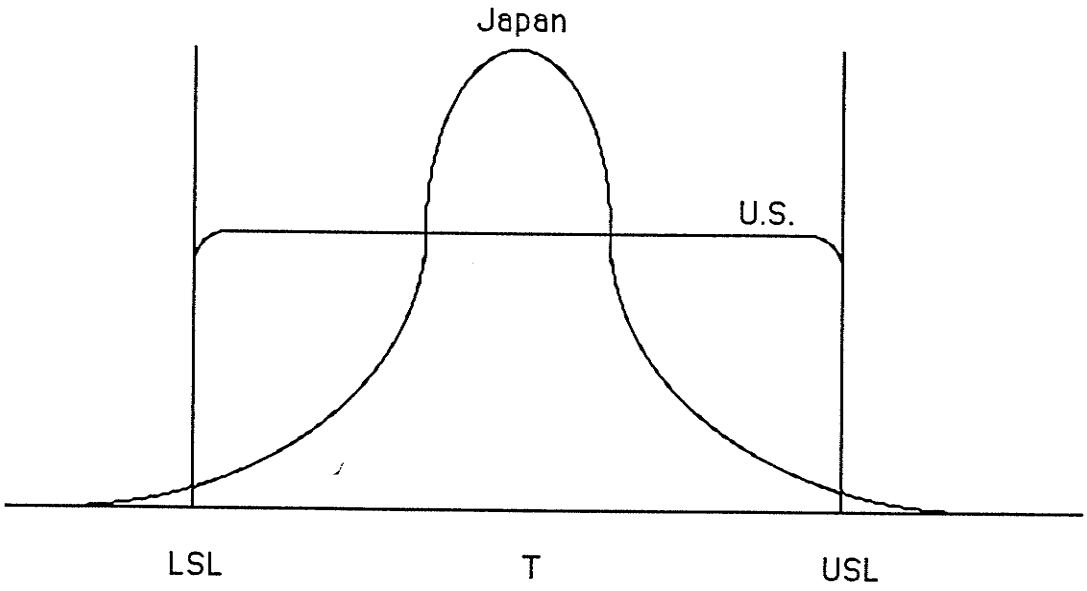


Figure 1.2.6 Taguchi's Television Example.

Although the example is useful in illustrating the problem associated with proximity to the target, it does not invite good statistical comparisons.

If 6σ is used to measure actual process spread it is not at all clear what 6σ represents outside the normal family of distributions. Hence the process capability index is not consistent in its interpretation when non-normal populations are encountered. Five populations (Figure 1.2.8), all such that Pr(LSL < x < USL) =0.9973, have process capabilities that range from 0.577 to 1.4030. Clearly a caution should be issued if distributional assumptions are not verified.

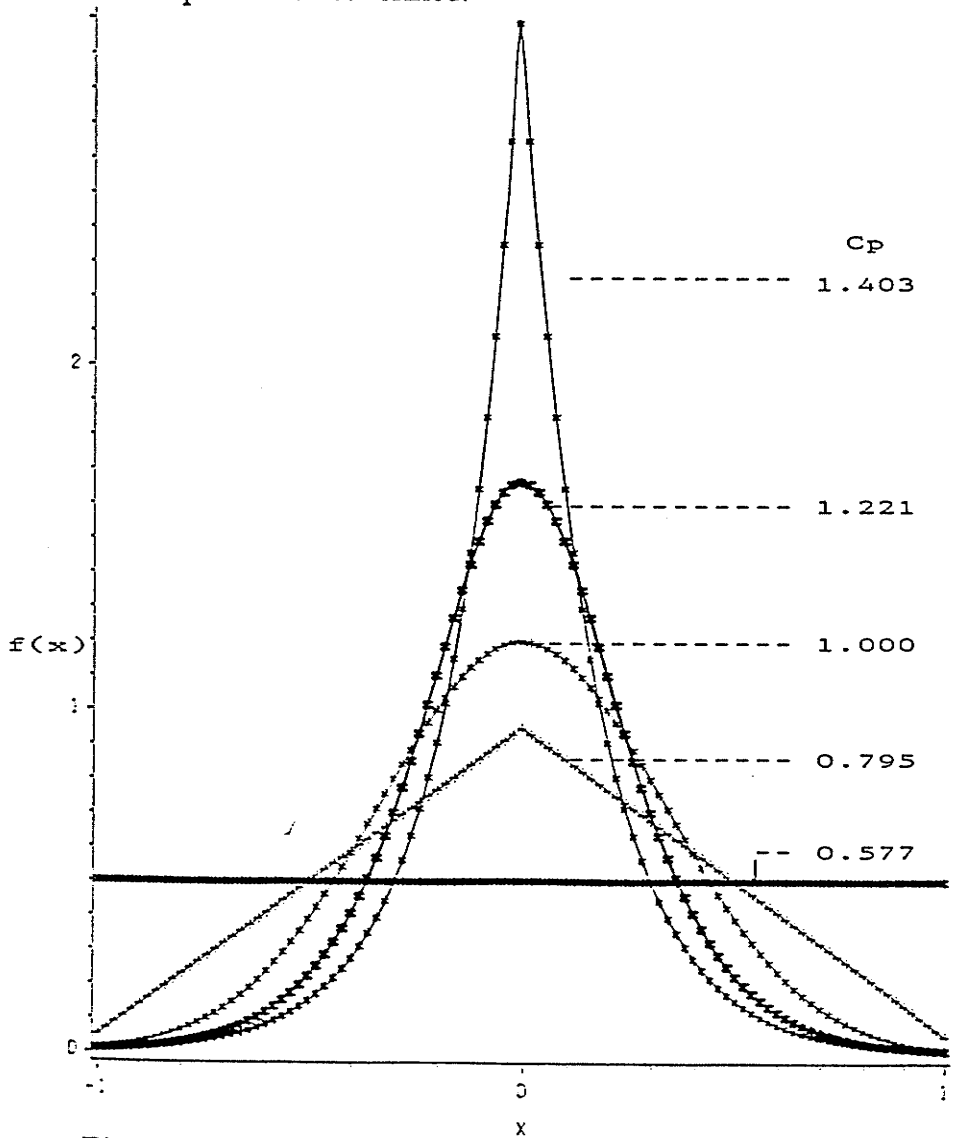


Figure 1.2.8. Five Populations with various Process Capabilities.

The non-robustness of the traditional process capability index estimator to departures from normality is also established and corrective procedures suggested for some identifiably non-normal populations. The non-robustness of the traditional estimator is not unexpected as it is, up to a constant, simply the inverse of the sample standard deviation, which is known to be non-robust with respect of departures from normality.

The process capability index can also be considered as a measure of non-conforming product. A value of one for the index represents 2700 parts per million (ppm) non-conforming, while 1.33 represents 63 ppm; 1.66 corresponds to .6 ppm; and 2 indicates <.1 ppm. These values are correct if the process measurement arise from a normal distribution centered on the midpoint of the specification limits. If this is not true the process capability index will underestimate the percent non-conforming. Processes 1 and 2 of Figure 1.2.9 have the same index values, but process 2 has roughly 30% non-conforming, where process 1 has near zero percent non-conforming.

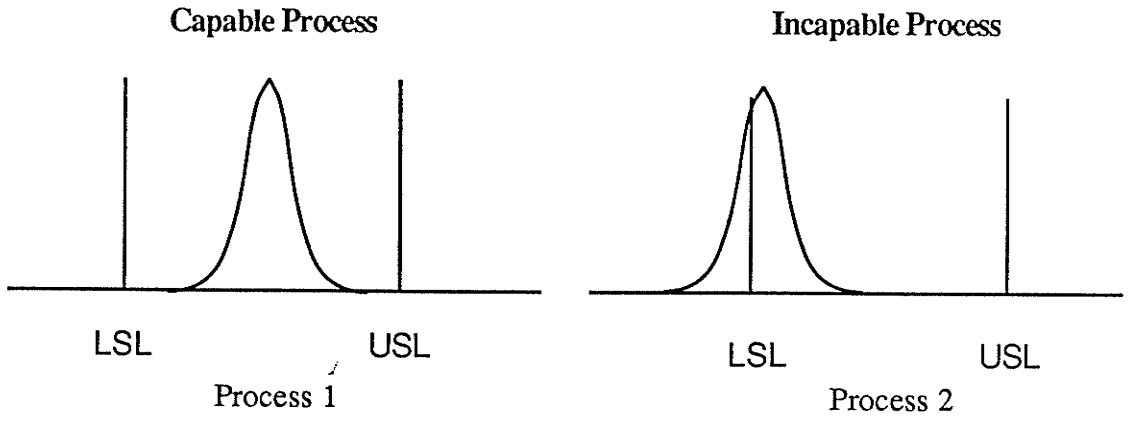


Figure 1.2.9. Processes with Equivalent Capability Index but different Non-conforming.

The process capability index's failure to consider proximity to the target value makes it incompatible with Taguchi's loss function concept of assessing quality. However a measure that is similar to the process capability index but with the ability to assess proximity to the target in addition to process variation is proposed. The changes are subtle but powerful, leaving the original measure largely unchanged but with superior statistical properties and more intuitive appeal than competing measures. This modified process capability index reflects the current feeling in QC and relates well with Taguchi's definition of quality. The modified process capability index behaves in much the same manner as the squared error loss function (inversely in terms of magnitude). At the target, for a given variability, the modified index is a maximum while the loss function is at is minimum. As the process drifts from the target the loss function increases, the modified index decreases. Similarly as the variability increases the modified index gets smaller while the loss function increases as the observations tend to move away from the target.

The modified index falls into a group of second generation measures of process capability that consider both process variation and proximity to the target when assessing the ability of a process. The shift from the original index to the modified index in terms of calculation is subtle but the inferences and interpretations are vastly different.

1.3. Control Charting

Control charts provide a graphical technique for monitoring the behaviour of some characteristic(s) of a process. There are many different categories of control charts whose use depends upon the nature of the quality characteristic(s) under investigation. However within a particular category of charts there are often several ways of displaying the sample results. For example when measuring a continuous quality variable practitioners often use $\bar{x}$ and R charts to monitor a process, while a statistician may use $\bar{x}$ and s charts. The first

pair of charts are more likely to be used on the manufacturing floor while the second group are more likely to be used by the data analyst in the Quality Assurance department.

Both the $\bar{x}$ and R procedure and the $\bar{x}$ and s procedure provide similar information regarding the behaviour of the quality characteristic under surveillance. The $\bar{x}$ chart monitors the central tendency of the process with respect to the grand mean $\bar{\bar{x}}$, while an R or s chart monitors the behaviour of the variability. Either set of charts monitor two characteristics of the process and involve plotting summary statistics for small groups of observations drawn periodically from the output of the process. Boundaries are created that aid in the identification of unusually large or small results. As well any trends or cycles that may occur can also be detected from these charts.

In either case (i.e., $\bar{x}$ and R or $\bar{x}$ and s) the plotting procedure requires two charts to illustrate the results. The latest trends in control charting procedures have focussed on fitting these charts on a single plot. Simultaneous control charts refer to the family of control charts that use a single plot to monitor a process. A plotting procedure is suggested that provides much of the information attainable from the traditional charts but with the added feature of requiring a single plot. The proposed plotting procedure brings control charts in step with the changing philosophy in QC. The univariate and multivariate procedures both provide good inferences, while being reasonably easy to use. Additional features include boundaries that are developed in a more statistically astute manner than the boundaries associated with the traditional charts. In terms of calculations the univariate procedure is easily performed using a hand calculator, while the multivariate case requires more computing power. For technicians working on the floor the procedure requires much the same level of sophistication as that of the $\bar{x}$ and s charts in the univariate case, but for the multivariate case a suitable software package is required.

1.4. Assessing Distributional Assumptions

The concept of distributional properties and assumptions is of general importance in the field of statistics and this carries over to the area of Quality Control. A test procedure is proposed that can be used in identifying the distribution from which a set of observations arise. The test procedure is motivated from probability plot results and is analogous to the Shapiro-Wilk test for normality. It suffers from some statistical drawbacks for small sample sizes but as sample sizes increase it behaves quite well. It has the advantage of being easy to calculate and its inference can be enhanced with an accompanying probability plot.

All too often an assumption regarding the underlying distribution of a population is made without verification. In most procedures distributional assumptions are made for theoretical reasons. However some procedures provide reasonable inferences when the distributional assumptions are not valid. The t and F statistics are good examples of robustness to moderate non-normal populations. On the other hand both Bartlett's and Hartley's tests for homogeneity of variance are extremely sensitive to departures from normality.

It has become the practise in many applications to ignore the distributional aspects of the underlying distribution. Seldom does one see the results of a test for normality included in a procedure that uses a t or F statistic. This can be quite dangerous even though both statistics have some propensity to operate in the face of non-normal distributions. In control charting and process capability procedures the assumption of normality plays a major role in ascertaining limits and distributional properties of estimators. Certain aspects of these procedures do not behave well when the assumption of normality is violated.

One major drawback in assessing distributional assumptions in QC procedures is the small sample sizes. In control chart procedures subgroup sizes are generally around

size five. There simply is not enough information in a sample of size five to provide an assessment of the underlying distribution. Hence the assumption is made but no warning is mentioned regarding non-normality and its consequences. Process capability studies suffer from the same practise but for different reasons. In many studies there is sufficient information to assess normality but this is seldom done. A graphical technique for assessing process capability, that is a modified normal probability plot, may help to alleviate this practise.

The inferences drawn from the graphical technique are subjective in that the linearity of the resultant plot provides information regarding the aptness of the distributional function in describing the population. In those cases where the linearity is borderline the graphical procedure becomes quite subjective. A procedure that can provide additional information in assessing the linearity of the plot is proposed.

The procedure is analogous to existing procedures for the Normal and exponential distributions, but is designed for use with the uniform distribution. For small sample sizes the procedure is not very powerful, however in conjunction with a uniform probability plot it can provide reasonably good inferences. The test procedure is quite easy to use and has the added features that i) a uniform probability plot does not require special probability paper and ii) the probability integral transformation results in U[0, 1] variates. A test statistic is proposed and several properties examined.

Chapter 2

Examining Goodness-of-fit

2.1. Introduction

Probability plots are an old and useful tool for examining the goodness-of-fit of a particular probability model to a data set. However probability plots provide no objective method for analyzing or testing goodness-of-fit. In fact, interpretation of probability plots is left to the judgement of the observer. In those cases where the resultant probability plot is quite obviously linear or definitely non-linear the probability plot procedure is very useful in assessing goodness-of-fit, however, in those cases where the linearity of the probability plot is borderline, subjectivity enters into the judgement regarding the aptness of a particular hypothesized distribution.

Shapiro and Wilk ([1], [2]) presented techniques for the normal and exponential distributions which examine the results of the probability plot procedure and quantify goodness-of-fit. The Shapiro-Wilk technique in conjunction with the probability integral transformation (PIT) and the uniform distribution allows investigation of distributional assumptions regarding almost every distribution imaginable.

With this thought an analogous test to the Shapiro-Wilk Analysis of Variance test for goodness-of-fit has been developed for the uniform distribution.

2.2. Probability Plots

Probability plots are used in several ways to aid in examining the nature of a set of observations arising from some population. Probability plots can be used to examine the data for outliers, goodness-of-fit, and systematic deviations from certain distributional assumptions. The probability plot procedure is quite straight forward in that it simply plots the order statistics of a data set versus the expected value of the order

statistics. If the resultant plot appears linear with no dramatic deviations, the data are said to have no outliers and arise from a population with probability density function (up to a constant) f(x).

For some probability distributions the expected value of the associated order statistics can be quite difficult to calculate, however, the uniform distribution lends itself quite nicely to the probability plot procedure by having a concise and uncomplicated algorithm for determining the expected value of its order statistics.

All distributions from the (continuous) uniform family can be expressed in the form

$$f(x) = \begin{cases} \frac{1}{a} & 0 < x < a \\ 0 & \text{elsewhere.} \end{cases} \qquad (2.2.1)$$

with cdf

$$F(x) = \begin{cases} 0 & x < 0, \\ \frac{x}{a} & 0 < x < a, \\ 1 & x > a\,. \end{cases} \qquad (2.2.2)$$

The expected value of the i^{th} order statistic $x_{[i]}$, from a sample of size n will be

$$E(x_{[i]}; n) = \frac{n!}{(i-1)!(n-i)!} \int_{-\infty}^{\infty} x_{[i]}[F(x_{[i]})]^{i-1}[1-F(x_{[i]})]^{n-i} dF(x_{[i]}) \qquad (2.2.3)$$

for i = 1, 2, 3, ... , n.

Substituting the results of (2.2.2) for $F(x_{[i]})$ in (2.2.3) and integrating, the expected value of the order statistics from any uniform distribution with pdf (2.2.1) is

$$E(x_{[i]};\, n) = \frac{n!}{(i-1)!(n-i)!} \int_0^a x_{[i]} \left[\frac{x_{[i]}}{a}\right]^{i-1} \left[1 - \frac{x_{[i]}}{a}\right]^{n-i} \frac{dx_{[i]}}{a} = \frac{a\ i}{n+1}$$

for i= 1, 2, 3, ... , n .

A probability plot (for any distribution) simply plots the order statistics ($x_{[i]}$) versus the expected value of the order statistics ($E(x_{[i]};n)$). Any inference regarding the aptness of the distribution function used to calculate the expected value of the order statistics is formulated from the shape of the plot.

If the probability plot does not readily avail itself to an inference (i.e., the plot is borderline linear) then an additional test can be performed which will allow quantification of the results. A natural inclination is to conduct a regression type analysis of the points ($x_{[i]}$,$E[x_{[i]};n]$) and to analyze them for linearity using traditional regression techniques.

Consider the following model

$$x_{[i]} = A + B[E(x_{[i]};n)] + e_i \qquad (2.2.4)$$

where i=1, 2, 3, ... n , and e_i is a suitable error term.

Equation (2.2.4) denotes a linear regression of the order statistics on their respective hypothesized expected values. Because the explanatory variable is the expected values of the order statistics, ordinary least squares (OLS) regression techniques will not generate the best linear unbiased estimates (BLUE) of A and B [3]. However using the general least squares (GLS) technique the resultant estimates for A and B will be BLUE (Searle [4]).

The GLS estimators of A and B, where the hypothesized distribution is uniform, are

$$\hat{A} = \frac{x_{[n]} + x_{[1]}}{2}$$

$$\hat{B} = \frac{n+1}{n-1}(x_{[n]} - x_{[1]}) .$$

Lloyd [5], as well as Kendall and Stuart [6], indicate that when regressing the order statistics on their respective standardized expected values the parameter A represents the population mean (μ) and B the population standard deviation (σ). Hence, $\hat{A}$ and $\hat{B}$ will provide estimates for μ and σ, respectively. Shapiro and Wilk use the result from [5] and [6] in conjunction with the sample variance, to examine distributional assumptions for data thought to have come from a normally distributed population and laterly for an exponentially distributed population.

2.3. THE W_U TEST PROCEDURE

The general form of the W-test as proposed by Shapiro and Wilk is

$$W = \frac{\hat{B}^2}{S^2}$$

where $\hat{B}$ is (up to a constant) the GLS estimate of B, when regressing the order statistics on their standardized expected values, and S^2 is the corrected sums of squares (i.e., $\sum_{i=1}^{n}(x_i - \bar{x})^2$) for the sample results.

In the case where the hypothesized distribution is from the uniform family, the proposed test statistic becomes

$$W_u = \frac{\left[\frac{n+1}{n-1}(x_{[n]}-x_{[1]})\right]^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} .$$

Computationally the W_u statistic is quite straight forward, however the exact distribution of the W_u statistic is difficult to determine.

2.4. TABULATION OF THE W_U STATISTIC

For each of the sample sizes n = 3(1)25, 30, 40, 50, 10,000 samples (of size n) were generated using a uniform(0,1) random number generator, and the W_u statistic calculated. Using these results the cumulative relative frequency distribution of the W_u statistic has been sketched for several sample sizes ranging (some have been left out for clarity) from 5 to 50 (Figure 2.4.1).

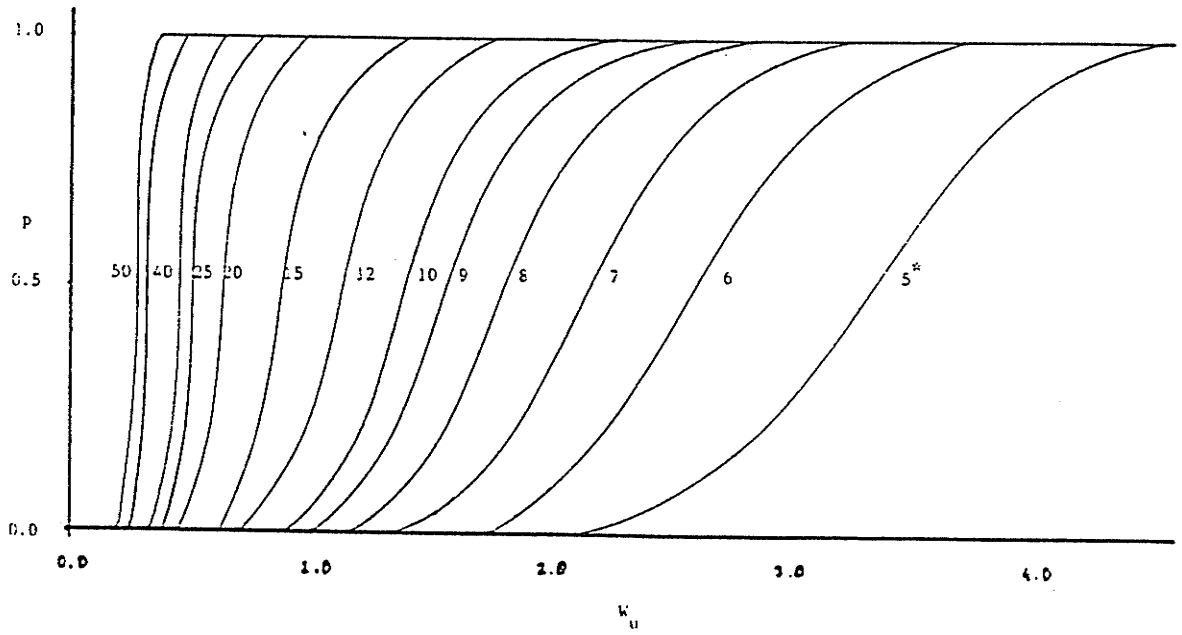


Figure 2.4.1. Cdfs of W_u for various sample sizes.*

In addition, certain critical quantiles (as calculated from the simulations) have been tabulated and included in Table 2.4.1. These quantiles provide approximate W_u acceptance-rejection regions for the various sample sizes included in the simulations.

Table 2.4.1

Quantiles and Bounds for the W_u distribution

		QUANTILES						
n	Lower Bound	$W_{.01}$	$W_{.05}$	$W_{.10}$	$W_{.90}$	$W_{.95}$	$W_{.99}$	Upper Bound
3	6.0000	6.0300	6.1500	6.3000	7.9700	7.9900	8.0000	8.0000
4	2.7778	3.0800	3.4400	3.7400	5.3100	5.4300	5.5300	5.5556
5	1.8750	2.3966	2.4256	2.5758	4.0199	4.1802	4.3884	4.5000
6	1.3067	1.7133	1.8801	2.0004	3.2287	3.4072	3.6698	3.9200
7	1.0370	1.3909	1.5456	1.6422	2.6763	2.8528	3.1341	3.5556
8	0.8265	1.1840	1.3168	1.4003	2.3176	2.4679	2.7716	3.3061
9	0.7031	1.0366	1.1525	1.2235	2.0403	2.1877	2.4592	3.1250
10	0.5975	0.9182	1.0161	1.0824	1.7929	1.9256	2.1799	2.9877
11	0.5280	0.8268	0.9177	0.9747	1.6046	1.7331	1.9909	2.8800
12	0.4656	0.7560	0.8386	0.8898	1.4530	1.5678	1.7882	2.7934
13	0.4213	0.6918	0.7656	0.8113	1.3186	1.4227	1.6358	2.7222
14	0.3804	0.6430	0.7101	0.7541	1.2164	1.3090	1.5181	2.6627
15	0.3499	0.6035	0.6639	0.7023	1.1157	1.2073	1.3901	2.6122
16	0.3211	0.5611	0.6184	0.6537	1.0345	1.1118	1.3062	2.5689
17	0.2988	0.5331	0.5803	0.6136	0.9663	1.0426	1.2047	2.5313
18	0.2776	0.5026	0.5498	0.5809	0.9054	0.9774	1.1244	2.4983
19	0.2605	0.4745	0.5223	0.5516	0.8534	0.9210	1.0461	2.4691
20	0.2443	0.4551	0.4978	0.5235	0.8036	0.8644	0.9849	2.4432
21	0.2310	0.4332	0.4733	0.4967	0.7603	0.8135	0.9250	2.4200
22	0.2181	0.4156	0.4510	0.4740	0.7206	0.7779	0.8917	2.3991
23	0.2074	0.3971	0.4313	0.4538	0.6809	0.7321	0.8362	2.3802
24	0.1969	0.3798	0.4169	0.4365	0.6524	0.6982	0.7998	2.3629
25	0.1881	0.3656	0.3979	0.4176	0.6238	0.6666	0.7632	2.3472
30	0.1524	0.3092	0.3348	0.3498	0.5091	0.5415	0.6121	2.2854
40	0.1105	0.2355	0.2535	0.2633	0.3689	0.3894	0.4359	2.2104
50	0.0867	0.1915	0.2048	0.2132	0.2878	0.3016	0.3310	2.1666
∞	0.0000							2.0000

2.5. THEORETICAL RESULTS OF THE W_U STATISTIC

Theorem 2.5.1: The central value of W_u with the assumption that the variable under investigation is from a uniform distribution will be

$$\frac{12(n+1)}{n(n-1)}.$$

Proof: Assume that x is a random variable with a standard uniform density function. Then the order statistics should be spread uniformly across the interval (0,1). If this is the case the squared deviations from the mean depend only on sample size. If n is odd the mean of the order statistics will be the $(n+1)/2^{th}$ ordered observation. If n is even the mean of the order statistic will be the mean of the $n/2^{th}$ and the $(n/2+1)^{th}$ ordered observations. In either case the squared deviation from the mean for the j^{th} order statistic will be

$$\left[j - \frac{n+1}{2} \right]^2 \qquad \text{for} \quad j=1,2,3, \ldots, n$$

Hence W_u will then be

$$W_u = \left[\frac{(n+1)^2}{(n-1)^2 \sum_{j=1}^{n} \left[\frac{j - \frac{n+1}{2}}{n-1} \right]^2} \right]$$

$$= \left[\frac{4(n+1)^2}{\sum_{j=1}^{n} (2j - n - 1)^2} \right]$$

$$= \left[\frac{4(n+1)^2}{\sum_{j=1}^{n} 4j^2 - \sum_{j=1}^{n} 4(n+1)j + \sum_{j=1}^{n} (n+1)^2} \right]$$

$$= \left[\frac{4(n+1)^2}{\frac{4n(n+1)(2n+1)}{6} - 2n(n+1)^2 + n(n+1)^2} \right]$$

$$= \left[\frac{12(n+1)}{n(n-1)} \right]$$

Theorem 2.5.2: The W_u statistic is location and scale invariant under the null hypothesis that the sampled population has an uniform distribution.

Proof: Let x be a random variable with a standard uniform distribution, and let y=ax+b where a, b are i) constants and ii) elements of the interval $(-\infty,\infty)$. Then the $W_u(y)$ statistic for the random variable y will be

$$= \left[\frac{(y_{[n]} - y_{[1]})^2 (n+1)^2}{(n-1)^2 \sum_{i=1}^{n} (y_i - \bar{y})^2} \right]$$

$$= \left[\frac{(n+1)^2 (ax_{[n]} + b - ax_{[1]} - b)^2}{(n-1)^2 \sum_{i=1}^{n} (ax_i + b - a\bar{x} - b)^2} \right]$$

$$= \left[\frac{(n+1)^2 (x_{[n]} - x_{[1]})^2}{(n-1)^2 \sum_{i=1}^{n} (x_i - \bar{x})^2} \right]$$

$$= W_u(x)$$

Theorem 2.5.3: For samples of size two, the pdf of the W_u statistic is degenerate.

Proof: For n=2, by definition $W_u = \dfrac{(n+1)^2 (x_{[n]} - x_{[1]})^2}{(n-1)^2 \sum_{i=1}^{n} (x_i - \bar{x})^2}$,

$$= \frac{9(x_{[2]} - x_{[1]})^2}{(x_{[2]} - \frac{x_{[2]} + x_{[1]}}{2})^2 + (x_{[1]} - \frac{x_{[1]} + x_{[2]}}{2})^2}$$

$$= \frac{36(x_{[1]} - x_{[2]})^2}{(x_{[1]} - x_{[2]})^2 + (x_{[1]} - x_{[2]})^2}$$

$$= 18 .$$

Theorem 2.5.4: For samples of size 3 the probability density function of the W_u statistic is

$$f(W_u) = \begin{cases} \dfrac{4\sqrt{3}}{W_u^{1.5}\sqrt{8 - W_u}} & 6 < W_u < 8, \\ 0 & \text{elsewhere.} \end{cases}$$

Proof: From Bol'shev [7]

$$g(\tfrac{w}{s}) = \begin{cases} \dfrac{4\sqrt{3}}{\left[\frac{w}{s}\right]^2 \sqrt{4 - \left[\frac{w}{s}\right]^2}} & \sqrt{3} < \frac{w}{s} < 2, \\ 0 & \text{elsewhere,} \end{cases}$$

where $w = x_{[3]} - x_{[1]}$ and $s^2 = \sum_{i=1}^{3} \dfrac{(x_i - \bar{x})^2}{2}$. Then since $W_u = \dfrac{(n+1)^2 w^2}{(n-1)^3 s^2} = 2\,\dfrac{w^2}{s^2}$, the theorem holds.

Theorem 2.5.5: For samples of size 4 the probability density function of the W_u statistic is

$$f(W_u) = \begin{cases} \frac{25\pi\sqrt{2}}{9W_u^2}\{1 - \frac{2}{\pi}\arcsin\{[1+2(\frac{25}{3W_u} - 2)^{.5}][3(\frac{50}{9W_u} - 1)^{.5}]^{-1}\}\} & \frac{25}{9} \le W_u < \frac{100}{27} \\ \frac{25\pi\sqrt{2}}{9W_u^2}\{1 - \frac{2}{\pi}\arcsin\{[2(\frac{50}{3W_u} - 4)^{.5}][3(\frac{50}{9W_u} - 1)]^{-1}\}\} & \frac{100}{27} \le W_u < \frac{225}{54} \\ \frac{25\pi\sqrt{2}}{9W_u^2} & \frac{225}{54} \le W_u < \frac{50}{9} \end{cases}$$

Proof: From Khakhubiya [8] the distribution of $x=(s^2/w^2)$ is

$$g(x) = \begin{cases} \sqrt{18\pi^2} & \frac{1}{6} \le x \le \frac{2}{9} \\ \sqrt{18\pi^2}\ \{1 - \frac{2}{\pi}\arcsin[\frac{\sqrt{8(9x-2)}}{3(6x-1)}]\} & \frac{2}{9} < x \le \frac{1}{4} \\ \sqrt{18\pi^2}\ \{1 - \frac{2}{\pi}\arcsin[\frac{1+2\sqrt{9x-2}}{3\sqrt{6x-1}}]\} & \frac{1}{4} < x \le \frac{1}{3} \end{cases}$$

Then since $W_u = \frac{(n+1)^2w^2}{(n-1)^3s^2}$ and n=4, $W_u = \frac{25}{27x}$, the theorem holds.

Theorem 2.5.6: For all values of n, W_u is bounded above by

$$\frac{2(n+1)^2}{(n-1)^2}$$

and below by

$$\frac{4(n+1)^2}{n(n-1)^2} \quad \text{if n is even, or}$$

$$\frac{4n(n+1)}{(n-1)^3} \quad \text{if n is odd.}$$

Proof: Under the assumption that W_u will be a minimum when half the data points reside at one end point (e.g. a) of the distribution and the other half of the points at the opposite end point (e.g. b) (if the sample size is odd then one end point will have one additional point) with d=|b-a|, it follows

$$\text{i) } W_u = \frac{(n+1)^2 d^2}{(n-1)^2 \sum \left[\frac{d}{2}\right]^2} = \frac{4(n+1)^2}{n(n-1)^2} \quad \text{where n is even, and}$$

$$\text{ii) } W_u = \frac{d^2(n+1)^2}{(n-1)^2\left[\frac{n+1}{2}\right]\left[\frac{(n-1)d}{2}\right]^2 + \left[\frac{n-1}{2}\right]\left[d - \frac{(n-1)d}{2n}\right]^2}$$

$$= \frac{4n(n+1)}{(n-1)^3} \quad \text{where n is odd.}$$

Assuming that W_u will be a maximum when all but two of the data points reside at the mean (e.g. (a+b)/n) of the data, while one point resides at one endpoint (a) of the data and the second point at the opposite end point (b) with d=|b-a|, W_u becomes

$$W_u = \frac{d^2(n+1)^2}{(n-1)^2\left[\left[\frac{d}{2}\right]^2 + \left[\frac{d}{2}\right]^2\right]}$$

$$= \frac{2(n+1)^2}{(n-1)^2}$$

Theorem 2.5.7: For a non-specific alternative hypothesis the W_u test will be two sided.

Proof: Assuming a fixed range, if the data follow a distribution with lighter tails than the uniform distribution, then in general

$$(x_i - \bar{x})^2 \leq \left[\frac{d}{2}\left[\frac{\frac{n+1}{2} - i}{\frac{n+1}{2} - 1}\right]\right]^2 \qquad \forall\, i$$

$$\Sigma (x_i - \bar{x})^2 \leq \Sigma \left[\frac{d}{2}\left[\frac{\frac{n+1}{2} - i}{\frac{n+1}{2} - 1}\right]\right]^2$$

$$S_n^2 \leq S^2$$

$$W_n \geq W$$

Thus if the data follow distributions with lighter tails than the uniform distribution, the rejection region will be in the upper tail of the W_u distribution.

Similarly for distributions with heavier tails

$$(x_i-\bar{x})^2 \geq \left[\frac{d}{2}\left[\frac{\frac{n+1}{2}-i}{\frac{n+1}{2}-1}\right]\right]^2 \qquad \forall\, i$$

$$\Sigma\,(x_i-\bar{x})^2 \geq \Sigma\left[\frac{d}{2}\left[\frac{\frac{n+1}{2}-i}{\frac{n+1}{2}-1}\right]\right]^2$$

$$S_n^2 \geq S^2$$

$$W_n \leq W$$

and hence the rejection region will be in the lower tail of the W_u distribution.

2.6. THE SENSITIVITY OF THE W_U TEST

The strength of any statistical test lies in its ability to provide correct inferences based on sample results. In order to assess the ability of the W_u statistic in discriminating against sample results from non-uniform distributions, random samples from various distributions were generated and analyzed using the W_u statistic. Samples from the normal, chi-square, exponential, and Weibull families of distributions were generated and the W_u statistic was used to test the hypothesis H_0: (Sample results arise

from a population with a uniform distribution). As a quality control measure, samples from the uniform family of distributions were generated and subjected to the W_u test.

Table 2.6.1

Percentage of Samples with W_u outside acceptance bounds for various distributions

	Sample size		
Distribution	10	20	40
Exponential(θ=1)	12.2%	47.6%	84.2%
Normal(0,1)	22.7%	57.5%	93.8%
Chi-Square(2)	12.2%	48.0%	84.3%
Chi-Square(5)	18.2%	51.7%	88.3%
Chi-Square(10)	21.4%	53.8%	90.9%
Chi-Square(20)	22.0%	56.5%	92.2%
Weibull(a=1,b=2)	18.6%	46.2%	84.8%
Weibull(a=1,b=5)	22.3%	56.5%	91.8%
Weibull(a=1,b=10)	22.6%	58.9%	93.1%
Cauchy (standard)	50.9%	97.2%	100.0%
Logistic (standard)	30.5%	72.7%	97.9%
Uniform	10.1%	9.8%	10.1%

Table 2.6.2

Comparison of Shapiro-Wilk W_e statistic and the W_u statistic for testing the hypothesis H_0: X~exponential(θ=1) for various simulated distributions

	W_e			W_u		
Sample size	10	20	40	10	20	40
Normal(4,1)	65.11%	96.14%	99.98%	12.17%	62.76%	94.29%
Weibull(a=1,b=2)	38.02%	76.80%	98.49%	15.04%	35.00%	74.17%
Weibull(a=1,b=5)	72.76%	98.13%	100.00%	21.38%	56.39%	92.00%
Chi-Square(2)	10.16%	10.02%	9.74%	12.18%	17.15%	20.25%

Table 2.6.3

Comparison of performance by Shapiro-Wilk W_e and W_u for the hypothesis

H_0: X~Weibull(a=1,b=2) *for various simulated distributions*

	W_e			W_u		
Sample size	10	20	40	10	20	40
Normal(4,1)	38.88%	76.99%	98.22%	02.24%	87.41%	99.87%
Exponential(θ=1)	52.71%	82.08%	97.68%	35.57%	73.22%	96.26%
Chi-Square(2)	53.39%	82.02%	97.33%	46.34%	77.48%	94.74%
Weibull(a=1,b=5)	50.30%	88.65%	99.79%	17.11%	43.62%	83.93%

Plotting the cumulative distribution curves of the simulated results (Figure 2.4.1) illustrates the relationship that exists between the test statistic and the sample size. For small sample sizes the 100(1-α)% intervals are much wider than those similar 100(1-α)% for large sample sizes, this result affects the discriminating power of the W_u test for small samples. For example, in samples of size 20, the percentage of identified samples from other distributions ranges from 46.2% (for the Weibull[1, 2]) to 97.2% (for the Cauchy distribution), while for samples of size 40, the percentages range from 84.2% (for the exponential) to 100% (for the Cauchy) (Table 2.6.1).

From the comparisons between the Shapiro-Wilk test statistic for the exponential distribution (W_e) and our proposed test statistic W_u for testing the exponential (Table 2.6.2) and Weibull distributions (Table 2.6.3), it is evident that the W_u statistic (at least for small sample sizes) does not do as well as other results currently available. For this reason W_u should not be used for small sample size cases. However it does have the advantage that it is easy to calculate. It (theoretically) can be used to test any continuous distribution, and for larger sample sizes it performs almost as well as the more widely known tests.

There are many situations where the distribution function is completely specified under H_0, however there are equally many situations where the parameter vector is not completely specified and/or the functional form of the distribution function is unknown. If the distribution function is not completely specified the PIT can not be used to produce uniform variates. A solution is presented for the case where the functional form of the distribution function is assumed but where the parameter vector has at least one unknown value.

2.7. THE GENERAL W_U PROCEDURE

The W_u statistic can be used to examine goodness-of-fit for any distribution function, provided the parameter values and the functional form of the distribution function are completely specified. The ability of the W_u statistic to test every continuous completely specified distribution function results directly from properties of the PIT. The general form of the PIT is

$$Y_i = F(X_i, \underset{\sim}{\theta}),$$

where under the hypothesis

$$H_0: X \sim f(x, \underset{\sim}{\theta}),$$

the Y_i's will be independent identically distributed (iid) U(0, 1) variates. Hence if the functional form (i.e., $f(\cdot,\cdot)$) and the values of the parameter vector (i.e., $\underset{\sim}{\theta}=(\theta_1,\theta_2,...,\theta_m)$) are known, a uniform probability plot of the transformed observations can be created and the W_u statistic calculated. In many cases the functional form of the distribution function is more important than the actual value of the parameters. For example most commercially produced probability paper assumes only the functional form of the observations, while

actually providing estimates for unknown parameters. Normal, Weibull and uniform probability paper are examples of such. In the normal and uniform cases, with pdfs

$$f(x, \underset{\sim}{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty<x<\infty, \sigma>0$$

and

$$f(x, \underset{\sim}{\theta}) = \frac{1}{\theta}$$

for $(\mu - \frac{\theta}{2}) < x < (\mu + \frac{\theta}{2})$ respectively, estimates for the location (μ) and scale (σ) parameters arise from the intercept and slope of the resultant probability plot. For the Weibull distribution with cdf

$$F(x; \sigma, \eta) = 1 - \exp\left\{-\left[\frac{x}{\sigma}\right]^{\eta}\right\} \qquad 0 \le x < \infty \text{ and } \sigma, \eta > 0$$

estimates for the scale (σ) and shape (η) parameters can be determined using the slope and intercept of the probability plot. In all cases only the functional form of the distribution function need be assumed.

The form of the W_u procedure is the same for all distribution functions. Under the null hypothesis the PIT results in a set of U(0, 1) observations regardless of the original distribution function. Hence the test procedure and critical values are those of the W_u procedure for all distribution functions. Thus the W_u procedure has the advantage of being unchanged regardless of the distribution function under investigation.

In many cases the null hypothesis does not completely specify the distribution function. When this occurs the PIT can not be used as it requires all parameter values as well as the functional form of the distribution function to be specified before it can be

performed. However in those cases where the functional form of the distribution function is known or assumed but where one or more of the parameter values are unspecified, a transformation analogous to the PIT exists that results in iid U(0, 1) variates. This transformation makes it possible to extend the W_u procedure to the case where only the functional form of the distribution function is specified.

The transformed variates obtained by substituting moment estimates for the missing parameter value(s) and performing the PIT are not iid U(0, 1)'s [9]. As a result the W_u test for

H_0: transformed observations are uniform

will not be a similar test for

H_0: non-transformed values are $F(x, \underset{\sim}{\theta})$

where $\underset{\sim}{\theta}$ is the parameter vector with at least one unknown value.

O'Reilly and Quesenberry [10] discuss this problem and provide a solution derived from a result first obtained by Rosenblatt [11]. Rosenblatt shows that if $(x_1, x_2, \ldots, x_m)$ is a vector of m random variables with absolutely continuous multivariate cdf F, then the U_i's, where

$$U_1 = F(x_1)$$

$$U_2 = F(x_2|x_1)$$

$$U_3 = F(x_3|x_1, x_2)$$

$$\vdots$$

$$U_m = F(x_m|x_1, x_2, \ldots, x_{m-1})$$

are a set of m iid U(0, 1)'s. O'Reilly and Quesenberry [10] use this result and a result from [9] to establish that if T_n, a p-component vector, is a complete and sufficient statistic for the parameter vector of missing values $\underset{\sim}{\theta}_p = (\theta_1, \theta_2, \ldots, \theta_p)$, the U_i's

$$U_1 = \tilde{F}_n(x_1)$$

$$U_2 = \tilde{F}_n(x_2|x_1)$$

$$\vdots$$

$$U_{n-p} = \tilde{F}_n(x_{n-p}|x_1, x_2, \ldots, x_{n-p-1}) \text{ will be iid U(0,1).}$$

$\tilde{F}_n(x_1)$, $\tilde{F}_n(x_2|x_1)$, ... , $\tilde{F}_n(x_{n-p}|x_1, x_2, \ldots, x_{n-p-1})$ are the marginal and conditional cdfs of $\tilde{F}_n(x_1, x_2, \ldots, x_n)$, where $\tilde{F}_n(x_1, x_2, \ldots, x_n)$ is the cdf of $(x_1, x_2, \ldots, x_n | T_n)$.

This general result provides the basis for extending the W_u procedure to the case where the distribution function is not completely specified. O'Reilly and Quesenberry refer to the resultant transformation as the conditional probability integral transformation (CPIT). By conditioning on the complete and sufficient statistic (T_n) for the missing parameter values, a subset of the original observations can be transformed into iid U(0, 1). Once the transformation is complete a probability plot can be created and the W_u statistic calculated for the transformed variates. Both may be of aid in assessing the hypotheses

$$H_0: U_i\text{'s} \sim U(0, 1)$$

versus

$$H_a: U_i\text{'s are not } U(0, 1)$$

which is a similar test for the hypotheses

$$H_0: X \sim F(X, \underset{\sim}{\theta})$$

versus

$$H_a: X \sim F(X, \underset{\sim}{\tau})$$

where $\underset{\sim}{\theta} \neq \underset{\sim}{\tau}$.

The CPIT can be mathematically intractable for many distribution functions. Quesenberry [12] has examined the CPIT for some of the more common distribution functions with the following results

1) **Exponential Distribution**

functional form $f(x, \underset{\sim}{\theta}) = \frac{1}{\theta} \exp\left[- \frac{x - \mu}{\theta} \right]$, $x > \mu$

$x_1, x_2, \ldots, x_n$ denote the observed sample values

$y_1, y_2, \ldots y_{n-1}$ denote the observed sample values with $x_{[1]}$ deleted

Case 1: $\underset{\sim}{\theta} = (\mu, \theta)$, $\mu = \mu_0$, $\theta = \theta_0$ known (PIT)

$$U_i = 1 - \exp\left[- \frac{x_i - \mu_o}{\theta_o} \right] \qquad \text{for } i = 1, 2, \ldots, n$$

Case 2: $\underset{\sim}{\theta} = (\mu, \theta)$, μ unknown, $\theta = \theta_0$ known (CPIT)

$$U_i = 1 - \exp\left[- \frac{y_i - x_{[1]}}{\theta_o} \right] \qquad \text{for } i = 1, 2, \ldots, n-1$$

Case 3: $\underset{\sim}{\theta} = (\mu, \theta)$, $\mu=\mu_0$ known, θ unknown (CPIT)

$$U_{i-1} = 1 - \left[\frac{\sum_{j=1}^{i-1}(x_j - \mu_o)}{\sum_{j=1}^{i}(x_j - \mu_o)} \right]^{i-1} \qquad \text{for } i = 2, 3, \ldots, n$$

Case 4: $\underset{\sim}{\theta} = (\mu, \theta)$, μ, θ unknown (CPIT)

$$U_{i-1} = 1 - \left[\frac{\sum_{j=1}^{i-1}(x_j - x_{[1]})}{\sum_{j=1}^{i}(x_j - x_{[1]})} \right]^{i-1} \qquad \text{for } i = 2, 3, \ldots, n-1$$

2) **Pareto Distribution**

functional form $f(x, \underset{\sim}{\theta}) = \dfrac{\alpha\delta^{\alpha}}{x^{1+\alpha}}$ for $x>\delta$, $\alpha,\delta>0$

$x_1, x_2, \ldots, x_n$ denote the observed sample values

$y_1, y_2, \ldots, y_{n-1}$ denote the observed sample values with $x_{[1]}$ deleted

Case 1: $\underset{\sim}{\theta} = (\alpha, \delta)$, $\alpha=\alpha_0$, $\delta=\delta_0$ known (PIT)

$$U_i = 1 - \left[\frac{\delta_o}{x_i} \right]^{\alpha_o} \qquad \text{for } i = 1, 2, \ldots, n$$

Case 2: $\underset{\sim}{\theta} = (\alpha, \delta)$, $\alpha=\alpha_0$ known, δ unknown (CPIT)

$$U_i = 1 - \left[\frac{x_{[1]}}{y_i} \right]^{\alpha_o} \qquad \text{for } i = 1, 2, \ldots, n-1$$

Case 3: $\underset{\sim}{\theta} = (\alpha, \delta)$, α unknown, $\delta=\delta_0$ known (CPIT)

$$U_i = 1 - \left[\frac{\sum_{j=1}^{i-1} (\ln(x_j) - \ln(\delta_o))}{\sum_{j=1}^{i} (\ln(x_j) - \ln(\delta_o))} \right]^{i-1} \qquad \text{for } i = 1, 2, \ldots, n$$

Case 4: $\underset{\sim}{\theta} = (\alpha, \delta)$, α, δ unknown (CPIT)

$$U_{i-1} = 1 - \left[\frac{\sum_{j=1}^{i-1} \ln(y_j - x_{[1]})}{\sum_{j=1}^{i} \ln(y_j - x_{[1]})} \right]^{i-1} \qquad \text{for } i = 1, 2, \ldots, n-1$$

3) **Normal Distribution**

$$\text{functional form } f(x, \underset{\sim}{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x - \mu)^2}{2\sigma^2}, \quad -\infty<x<\infty, \ \sigma>0$$

$x_1, x_2, \ldots, x_n$ denote the observed sample values

$$\bar{x}_i = \sum_{j=1}^{i} \frac{x_j}{i}, \quad s_i^2 = \sum_{j=1}^{i} \frac{(x_j - \mu_0)^2}{i}, \quad s_i^{2^*} = \sum_{j=1}^{i} \frac{(x_j - \bar{x}_i)^2}{i - 1}$$

$\Phi(z)$ denote the cdf of the N(0, 1) distribution

$t_\upsilon(z)$ denotes the cdf of the t distribution with υ degrees of freedom

Case 1: $\underset{\sim}{\theta} = (\mu, \sigma)$, $\mu=\mu_0$, $\sigma=\sigma_0$ known (PIT)

$$U_i = \Phi\left\{ \frac{x_i - \mu_o}{\sigma_o} \right\} \qquad \text{for } i = 1, 2, \ldots, n$$

Case 2: $\underset{\sim}{\theta} = (\mu, \sigma)$, μ unknown, $\sigma=\sigma_0$ known (CPIT)

$$U_{i-1} = \Phi\left\{\frac{\sqrt{i-1}\,(x_i - \bar{x}_{i-1})}{\sqrt{i\sigma_0^2}}\right\} \qquad \text{for } i = 2, \ldots, n$$

Case 3: $\underset{\sim}{\theta} = (\mu, \sigma)$, $\mu=\mu_o$ known, σ unknown (CPIT)

$$U_{i-1} = t_{i-1}\left\{\frac{x_i - \mu_o}{s_{i-1}}\right\} \qquad \text{for } i = 2, 3, \ldots, n$$

Case 4: $\underset{\sim}{\theta} = (\mu, \sigma)$, μ, σ unknown (CPIT)

$$U_{i-2} = t_{i-2}\left\{\frac{\sqrt{\frac{i-1}{i}}\,(x_i - \bar{x}_{i-1})}{\sqrt{s_{i-1}^{2*}}}\right\} \quad \text{for } i = 3, 4, \ldots, n$$

4) **Lognormal Distribution**

$$\text{functional form } f(x, \underset{\sim}{\theta}) = \frac{1}{\sqrt{2\pi x^2\sigma^2}}\exp\left\{-\frac{[\ln(x) - \mu]^2}{2\sigma^2}\right\}, \quad -\infty<x<\infty,\ \sigma>0$$

$x_1, x_2, \ldots, x_n$ denote the observed sample values

$y_i = \ln(x_i)$

$$\bar{y}_i = \sum_{j=1}^{i}\frac{y_j}{i}, \quad s_i^2 = \sum_{j=1}^{i}\frac{(y_j - \mu_0)^2}{i}, \quad s_i^{2^*} = \sum_{j=1}^{i}\frac{(y_j - \bar{y}_i)^2}{i-1}$$

$\Phi(z)$ denotes the cdf of the N(0, 1) distribution

$t_\upsilon(z)$ denotes the cdf of the t distribution with υ degrees of freedom

Case 1: $\underset{\sim}{\theta} = (\mu, \sigma)$, $\mu=\mu_0$, $\sigma=\sigma_0$ known (PIT)

$$U_i = \Phi\left\{ \frac{y_i - \mu_o}{\sigma_o} \right\} \qquad \text{for } i = 1, 2, \ldots, n$$

Case 2: $\underset{\sim}{\theta} = (\mu, \sigma)$, μ unknown, $\sigma=\sigma_0$ known (CPIT)

$$U_{i-1} = \Phi\left\{ \frac{\sqrt{i-1}\,(y_i - \bar{y}_{i-1})}{\sqrt{i\sigma_0^2}} \right\} \qquad \text{for } i = 2, \ldots, n$$

Case 3: $\underset{\sim}{\theta} = (\mu, \sigma)$, $\mu=\mu_O$ known, σ unknown (CPIT)

$$U_{i-1} = t_{i-1}\left\{ \frac{y_i - \mu_o}{s_{i-1}} \right\} \qquad \text{for } i = 2, 3, \ldots, n$$

Case 4: $\underset{\sim}{\theta} = (\mu, \sigma)$, μ, σ unknown (CPIT)

$$U_{i-2} = t_{i-2}\left\{ \frac{\sqrt{\frac{i-1}{i}}\,(y_i - \bar{y}_{i-1})}{\sqrt{s_{i-1}^{2*}}} \right\} \qquad \text{for } i = 3, 4, \ldots, n$$

5) Weibull Distribution

functional form $f(x, \underset{\sim}{\theta}) = \frac{\eta}{\sigma} x^{\eta-1} \exp\left[-\frac{x^{\eta}}{\sigma} \right]$

$x_1, x_2, \ldots, x_n$ denote the observed sample values

Case 1: $\underset{\sim}{\theta} = (\sigma, \eta)$, $\sigma=\sigma_0$, $\eta=\eta_0$ known (PIT)

$$U_i = 1 - \exp\left[-\frac{x_i^{\eta_o}}{\sigma_o} \right] \qquad \text{for } i = 1, 2, \ldots, n$$

Case 2: $\underset{\sim}{\theta} = (\sigma, \eta)$, σ unknown, $\eta=\eta_0$ known (CPIT)

$$U_{i-1} = 1 - \left[\frac{\sum_{j=1}^{i-1} x_j^{\eta_o}}{\sum_{j=1}^{i} x_j^{\eta_o}} \right]^{i-1} \qquad \text{for } i = 2, 3, \ldots, n$$

For each of the above distribution functions the proposed transformations result in iid U(0, 1) variates, allowing a uniform probability plot and the W_u statistic to be used to assess goodness-of-fit. The CPIT transformations allow the W_u statistic to perform exact tests for those cases where some or all of the parameter values associated with the distribution function are not specified.

Care must be exercised when performing the CPIT. For distribution functions other than the general exponential and uniform, different permutations of the data may lead to different values for the transformed variable. Quesenberry [12] discusses this topic and warns against ordering the data (in any fashion) prior to performing the CPIT.

2.8. Assessing Multivariate Normality

The univariate results for the PIT and CPIT can be extended to multivariate (MV) distribution functions. However, as is the case with many multivariate procedures that are adaptations of univariate results, both the PIT and CPIT can be difficult to administer.

For any completely specified multivariate distribution the PIT will continue to result in iid U(0, 1) variates. Hence the W_u procedure in conjunction with a uniform probability plot of the transformed results can be used to assess MV goodness-of-fit. Similar to the univariate case, if the functional form and/or the parameter vector of the multivariate

distribution function is unknown the PIT can not be made. Assuming the functional form to be specified and substituting estimates for the missing parameter values into the PIT, provides only asymptotically iid U(0, 1) variates. For large sample sizes and where (n-p) (p being the number of missing parameter values) is large this technique is a reasonable approximate test, however if an exact test is required this procedure can not be used.

O'Reilly and Quesenberry [10] and Rincon-Gallardo, Quesenberry and O'Reilly [13] have extended the CPIT to the MVN case. The resultant transformations in conjunction with the W_u procedure (or any test for uniformity) will result in an exact test procedure for assessing MVN goodness-of-fit. However like other MV test procedures, the CPIT can be difficult to administer. In addition, Rincon-Gallardo, Quesenberry and O'Reilly [13] suggest that sample sizes of between 100 and 200 are required in order to obtain good results, while again cautioning that different orderings of the observations will produce different transformed values. Although the PIT and CPIT are theoretically appealing, practical applications in the MV case are few.

There is a widely used graphical technique that is relatively easy to use while also being intuitively appealing. The procedure however depends on the subjective interpretation of a probability plot, while also belonging to the group of MVN procedures that rely on asymptotic theory. The general procedure proposed by Healy [14] uses a probability plot of the ordered Mahalanobis distances to assess MVN. However by using the PIT associated with the chi-square distribution function and the W_u procedure, a method for statistically assessing multivariate normality is presented that, in conjunction with the multivariate probability plot, is easy to use and eliminates the subjectivity involved in assessing borderline probability plots. The W_u procedure, in conjunction with the MVN plot , provides quantitative interpretation of the aptness of the MVN distribution function. Rather than subjective interpretation of the linearity of the MVN plot for those cases where

the linearity is suspect, a procedure that provides analytical results is possible, resulting in a general test for MVN.

Let $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ denote a sample of n independent υ dimensional vectors thought to come from a MVN distribution with dimension υ. The hypothesis of interest is

$$H_0: \underset{\sim}{Y} \sim MVN_\upsilon$$

versus

$$H_a: \underset{\sim}{Y} \text{ is not } MVN_\upsilon.$$

That is, we want to determine the aptness of the MVN_υ pdf in describing the underlying distribution of the sampled vectors.

The first step in the procedure will be to determine the ordered Mahalanobis distances [15] associated with each observed vector,

$$D_i^2 = (\underset{\sim}{Y}_i - \underset{\sim}{\bar{Y}})' S^{-1} (\underset{\sim}{Y}_i - \underset{\sim}{\bar{Y}}) \quad i = 1, 2, \ldots, n$$

where $\underset{\sim}{\bar{Y}} = n^{-1}(\underset{\sim}{Y}_1 + \underset{\sim}{Y}_2 + \ldots + \underset{\sim}{Y}_n)$ and $S = \sum_{i=1}^{n} (\underset{\sim}{Y}_i - \underset{\sim}{\bar{Y}})(\underset{\sim}{Y}_i - \underset{\sim}{\bar{Y}})'$. Under the null hypothesis the D_i^2's will follow a χ^2_υ distribution.

Healy's procedure plots the ordered D_i^2's versus the expected value of the ordered D_i^2's assuming a χ^2_υ distribution. The inference drawn regarding the MV normality of the data will depend upon the linearity of the plot. A reasonably linear plot indicates that the data do arise from a MVN distribution. A non-linear plot suggests that the MVN pdf is not appropriate in describing the population from which the observations were drawn.

Rather than plotting the ordered D_i^2's versus the quantiles of the expected values of the ordered statistics as determined from the χ^2_υ distribution, the PIT performed on the D_i^2's will result in iid U(0, 1) variates under the null hypothesis. A uniform probability plot of the transformed distances will then provide the same information as the χ^2_υ probability plot of the original distances. This result in turn permits an easy and general procedure for assessing MVN for any dimensionality under investigation.

The transformed procedure plots $H_\upsilon(D_i^2)$ where

$$H_\upsilon(t) = \int_0^t f_\upsilon(x)\, dx$$

$$f_\upsilon(x) = \frac{1}{\Gamma(\frac{\upsilon}{2})}\left[\frac{1}{2}\right]^{\frac{\upsilon}{2}} x^{\frac{\upsilon}{2}-1} e^{-\frac{x}{2}} \qquad 0 < x < \infty$$

versus the expected value of the order statistics now assumed to arise from a Uniform (0,1) distribution. If the resultant plot is linear, the assumption that the D_i^2's arise from a χ^2_υ distribution is not unjustified, suggesting that the MVN distribution is not inappropriate for representing the distributions of the sampled vectors. On the other hand if the plot is definitely non-linear the χ^2_υ distribution will not be representative of the Mahalanobis distances and hence the underlying distribution of the sampled vectors will not be MVN.

For those borderline linear plots, the W_u procedure can be used. Once the probability integral transformation has been made on the D_i^2's,

$$U_i = H_\upsilon(D_i^2) \qquad i = 1, 2, \ldots, n$$

the resulting U_i's will follow a Uniform (0, 1) distribution under the null hypothesis. Using the W_u procedure for testing the hypothesis

$$H_0: U \sim \text{Uniform } (0, 1)$$

versus

$$H_a: U \text{ is not Uniform } (0, 1)$$

results in the following test statistic

$$W_u = \frac{\left[\frac{(U_{[n]} - U_{[1]})(n+1)}{n-1}\right]^2}{\sum_{i=1}^{n}(U_i - \overline{U})^2}$$

where $\overline{U} = \frac{\sum_{i=1}^{n} U_i}{n}$.

Failing to reject H_0 suggests that the uniform distribution is adequate in describing the transformed distances, which in turn suggests that the original distances do arise from a χ^2_υ and hence that the original sample vectors are in fact from the MVN family.

2.9. EXAMPLES

Three examples of the applicability of the W_u statistic have been included. The first example illustrates the use of the W_u statistic when investigating results thought to arise from a Weibull distribution. Example 2.9.2 illustrates the use of the W_u statistic for investigating a set of data thought to arise from a population possessing an exponential distribution. Example 2.9.3 illustrates the W_u procedure when used in conjunction with Healy's multivariate normal plot.

Example 2.9.1: The data were taken from Example 2.2, page 46 of Sinha [16] and were said to have been generated from a two parameter Weibull distribution with pdf of the form

$$f(x; a, b) = \frac{b}{a} x^{b-1} e^{-\frac{x^b}{a}} \quad 0 < x < \infty \text{ with } a=4, b=2.$$

A probability plot (Figure 2.9.1) casts sufficient doubt as to whether the data arises from a population with an underlying Weibull (a=4, b=2) distribution to warrant further investigation.

The above Weibull distribution has a cdf of the form

$$F(x) = 1 - e^{-\frac{x^2}{4}} \quad 0 < x < \infty .$$

from which the following were found

Rank	x	F(x)
1	0.3761	.0347
2	0.5903	.0834
3	0.6288	.0950
4	0.6461	.0991
5	0.7500	.1312
6	0.7705	.1379
7	1.0509	.2413
8	1.3162	.3515
9	1.3545	.3679
10	1.3592	.3699
11	1.5319	.4438
12	1.5700	.4600
13	1.6173	.4800
14	1.6560	.4962
15	1.7172	.5215
16	1.7708	.5434
17	1.7961	.5536
18	1.8487	.5745
19	1.8802	.5868
20	1.8889	.5902
21	1.8889	.5902
22	1.9310	.6063
23	3.0349	.9018
24	3.3546	.9027
25	3.9558	.9865

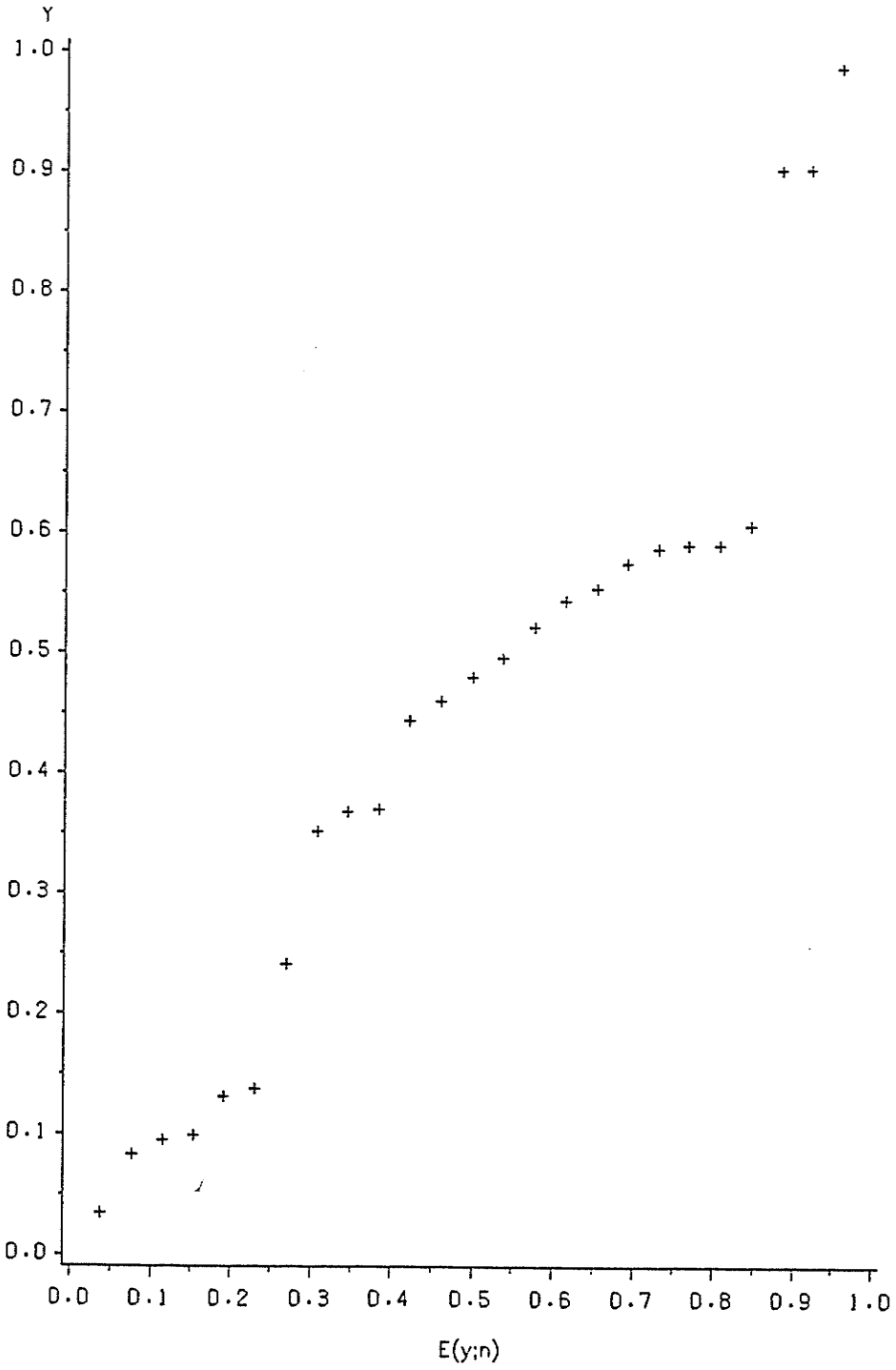


Figure 2.9.1 . Probability Plot for Example 2.9.1.

Treating $F(x_i)$, i=1,2, ... , 25 as the variable of interest ($F(x_i)=y_i$) results in

$$y_{[n]} - y_{[1]} = 0.9865 - 0.0347 = 0.9518$$

$$(n-1)s^2 = \sum_{i=1}^{25}(y_i - \bar{y})^2 = 1.64367$$

$$\text{and } W_u = \frac{(.9518)^2(26)^2}{1.64367(24)^2} = 0.64685 \ .$$

Then from Table 2.4.1, the approximate acceptance region for n=25, α=.1 is (.398, .667). Therefore, since W_u is an element of the acceptance region, one could conclude that there is not enough evidence to suggest that the data do not arise from a Weibull (a=4, b=2) distribution.

Example 2.9.2: Suppose that fifty-one observations (actually generated from a normal distribution with μ=4 and σ=1) were thought to arise from a population with an exponential distribution function (μ=0 and σ unknown), i.e.,

$$f(x, \underset{\sim}{\theta}) = \frac{1}{\sigma}\exp\left[-\frac{x}{\sigma}\right] \qquad x>0 \ .$$

Because σ is unknown in this case, the CPIT

$$U_{i-1} = 1 - \left[\frac{\sum_{j=1}^{i-1} x_j}{\sum_{j=1}^{i} x_j}\right]^{i-1} \qquad \text{for } i = 2, 3, \ldots , n$$

was used with the following results

i	x_i	U_i
1	3.46699	-
2	4.56077	0.568125
3	3.71026	0.532266
4	2.67077	0.459368
5	2.86051	0.515369
6	4.61174	0.693780
7	4.42081	0.668495
8	3.19812	0.552129
9	2.51814	0.480713
10	4.86938	0.720328
11	3.90443	0.634364
12	2.68995	0.504636
13	2.95665	0.545893
14	4.34931	0.687720
15	2.67675	0.512798
16	2.16189	0.448214
17	4.69168	0.726261
18	3.06715	0.569661
19	4.25339	0.689343
20	4.62955	0.715746
21	3.23884	0.583901
22	3.03853	0.563303
23	4.14997	0.677840
24	4.10856	0.672159
25	4.11375	0.670856
26	4.27908	0.683295
27	3.88213	0.646273
28	3.61455	0.619952
29	2.83205	0.533137
30	3.80442	0.641958
31	2.96548	0.551964
32	2.16314	0.446488
33	4.90936	0.739180
34	2.63017	0.512873
35	3.91565	0.658401
36	4.76996	0.727849
37	2.86131	0.541461
38	4.77197	0.727208
39	4.24827	0.683226
40	3.45063	0.606513
41	3.20351	0.580247
42	3.24542	0.586113
43	4.47897	0.703604
44	4.22079	0.680548
45	3.45160	0.606384
46	3.06617	0.564154
47	4.12335	0.672893
48	4.09604	0.669577
49	5.12119	0.747794
50	4.24984	0.679241
51	5.69706	0.780041

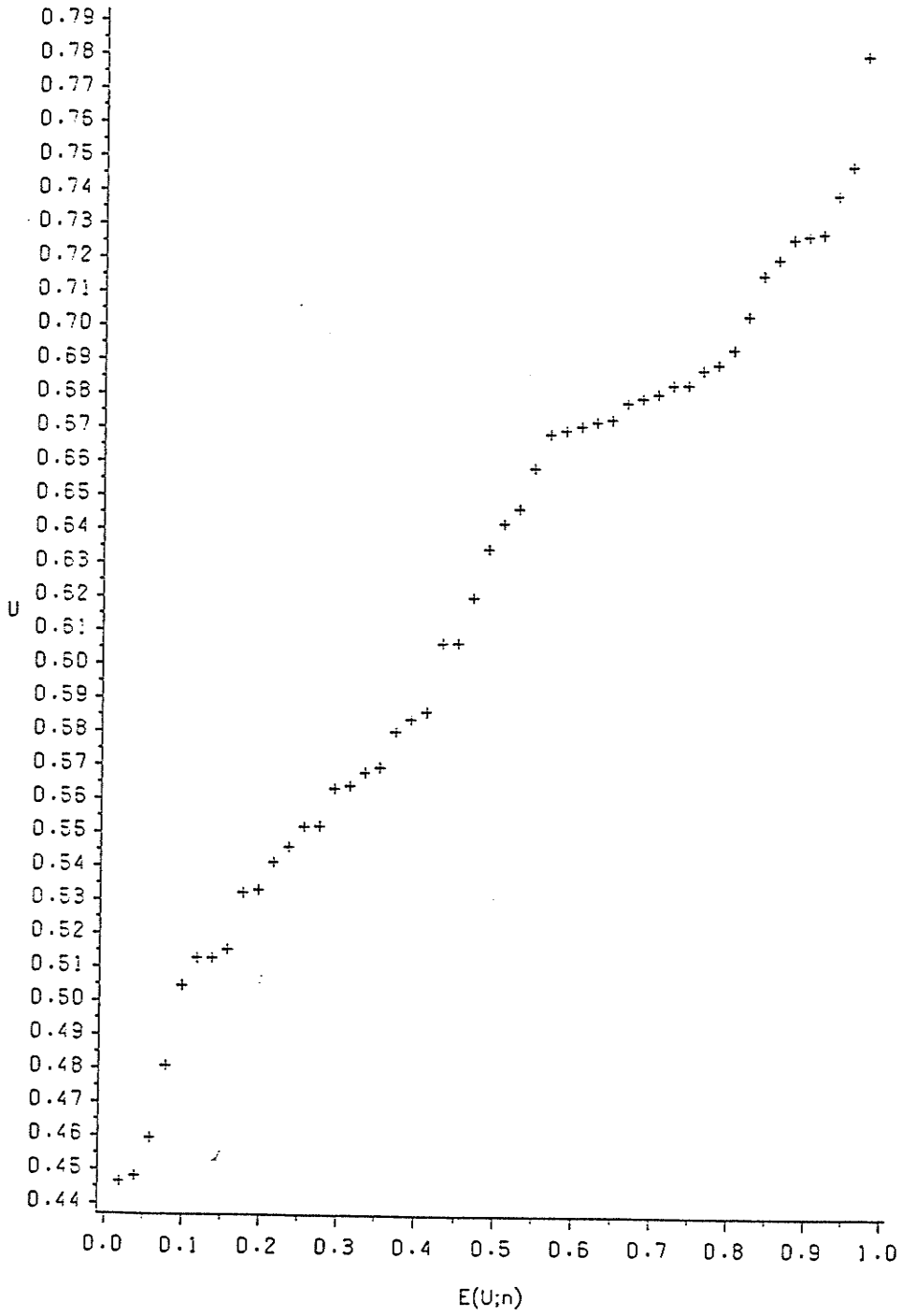


Figure 2.9.2 . Probability Plot for Example 2.9.2.

A probability plot (Figure 2.9.2) of the resultant U_i's casts doubt as to whether the data arise from a population with an exponential distribution function.

Using the W_u procedure,

$$W_u = \frac{\left\{\frac{n+1}{n-1}(U_{[n]} - U_{[1]})\right\}^2}{\sum_{i=1}^{50}(U_i - \bar{U})^2} = \frac{\left\{\frac{51}{49}(0.780041 - 0.446488)\right\}^2}{0.370053} = 0.3257 \ .$$

A linear interpolation from Table 2.4.1 suggests a p-value of 0.0344 for $W_u = 0.3257$ indicating that the data are in fact not exponentially distributed. The value of the Shapiro-Wilk test for the exponential distribution with known origin is W_e=0.2849 (p-value < 0.01), which is in agreement with the W_u procedure.

Example 2.9.3: Fifty pairs of observations ($\underset{\sim}{X} = (x_1, x_2)$) (generated independently from two U(0,1) populations) are thought to arise from a population possessing a Bivariate Normal distribution function, i.e.,

$$f(\underset{\sim}{X}, \underset{\sim}{\theta}) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{x_1-\mu_1}{\sigma_1}\frac{x_2-\mu_2}{\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right\}\right]$$

To examine this claim, Healy's procedure first calculates the sample Mahalanobis distances for each set of observations. The distances are then plotted on chi-square (υ=2) probability paper and the linearity of the plot used to assess goodness-of-fit. By employing the PIT associated with the χ_2^2 distribution function an uniform probability plot of the transformed values will provide the same information. The observed values ($\underset{\sim}{X} = (x_1, x_2)$), the Mahalanobis distances (D_i's) and the results of the PIT (U_i) for the fifty pairs of observations are

i	x_1	x_2	D_i	U_i
1	0.187892	0.602599	1.47748	0.522286
2	0.158867	0.884684	3.53558	0.829290
3	0.538657	0.390221	0.15089	0.072672
4	0.365064	0.245086	0.73670	0.308125
5	0.886303	0.093749	3.89823	0.857600
6	0.635174	0.362487	0.43070	0.193742
7	0.313890	0.543742	0.53013	0.232845
8	0.665011	0.833904	0.99304	0.391354
9	0.417821	0.313976	0.38766	0.176203
10	0.992919	0.991079	3.44292	0.821195
11	0.058372	0.050890	3.28287	0.806298
12	0.312426	0.946027	2.82103	0.755983
13	0.867790	0.945357	2.33278	0.688511
14	0.612469	0.769258	0.61615	0.265140
15	0.922171	0.935806	2.58568	0.725510
16	0.099271	0.442906	1.85734	0.604922
17	0.913334	0.411353	1.93769	0.620480
18	0.618213	0.313847	0.56650	0.246670
19	0.820695	0.414921	1.18847	0.448015
20	0.575715	0.046831	2.23935	0.673614
21	0.083141	0.346365	1.99596	0.631377
22	0.361104	0.077603	1.73644	0.580302
23	0.270909	0.165602	1.39253	0.501557
24	0.268493	0.566029	0.81760	0.335554
25	0.257458	0.095729	1.86134	0.605710
26	0.918656	0.849223	2.09958	0.649989
27	0.891344	0.826155	1.82044	0.597563
28	0.180900	0.394667	1.20319	0.452063
29	0.168098	0.215719	1.66996	0.566116
30	0.594946	0.253341	0.78990	0.326287
31	0.902094	0.495823	1.57761	0.545614
32	0.292795	0.012846	2.38026	0.695818
33	0.898434	0.973548	2.69182	0.739697
34	0.420613	0.250528	0.63051	0.270396
35	0.624813	0.234787	0.99133	0.390833
36	0.060599	0.486606	2.28513	0.681000
37	0.383403	0.847364	1.57159	0.544242
38	0.652861	0.637293	0.24977	0.117401
39	0.982719	0.564165	2.17665	0.663220
40	0.917294	0.967925	2.76407	0.748933
41	0.909504	0.026255	4.86434	0.912154
42	0.275453	0.546082	0.73345	0.307000
43	0.000593	0.966013	6.25673	0.956211
44	0.772428	0.194717	2.04535	0.640368
45	0.614798	0.910246	1.45351	0.516525
46	0.498718	0.945279	1.90590	0.614397
47	0.304777	0.383323	0.53425	0.234424
48	0.512108	0.990836	2.27802	0.679865
49	0.988910	0.610297	2.19323	0.666000
50	0.259543	0.139205	1.58265	0.546757

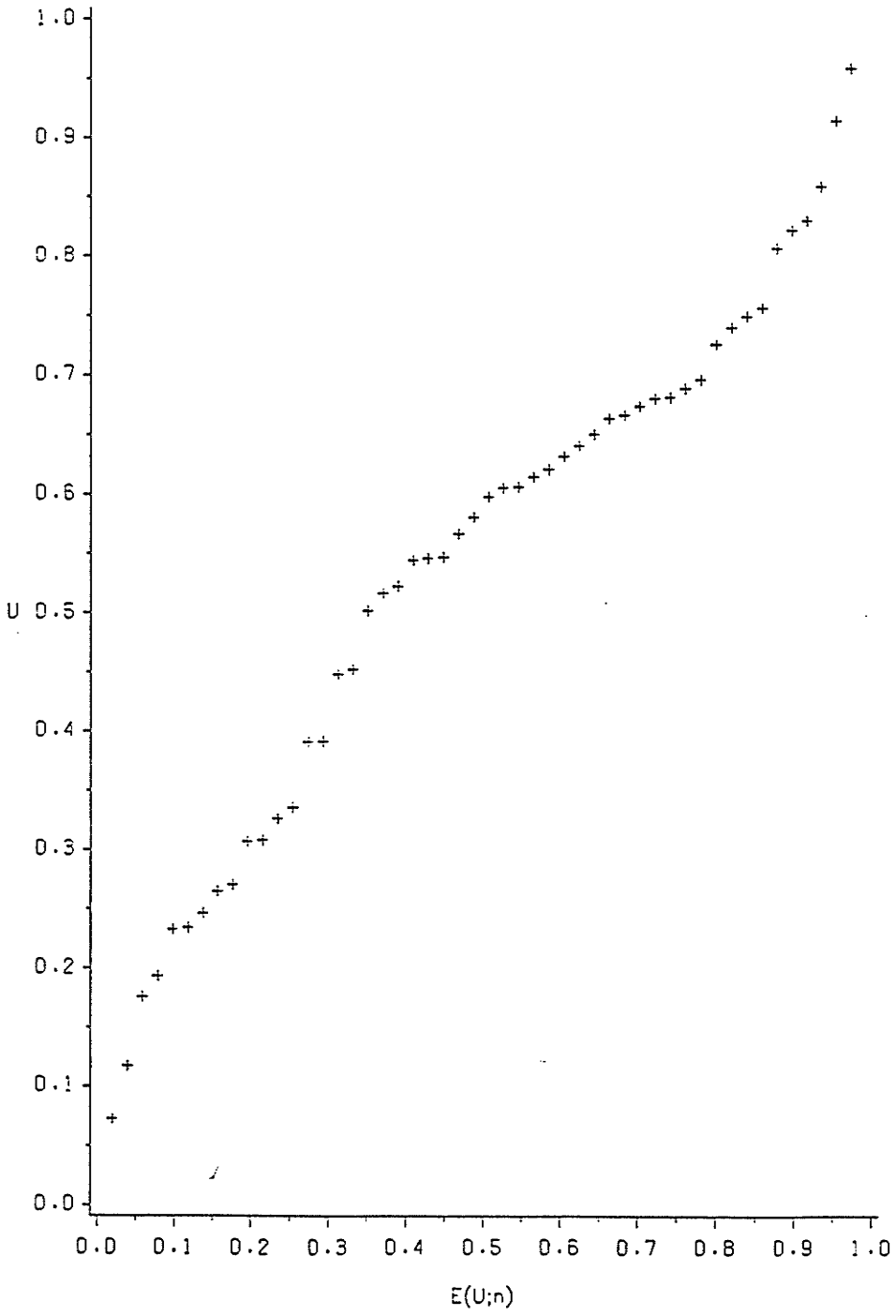


Figure 2.9.3. Probability Plot for Example 2.9.3.

An uniform probability plot (Figure 2.9.3) of the U_i's does not appear strongly linear.

The W_u statistic in this case is

$$W_u = \frac{\left[\frac{51}{49}(0.956211 - 0.0726715)\right]^2}{2.33194} = 0.36265$$

which for sample size 50 is not an element of the acceptance region ([0.1915, 0.3310]) resulting in a p-value < 0.02, suggesting that the observations do not arise from a population possessing a bivariate normal distribution function.

2.10. COMMENTS

A limiting requirement of the W_u test is that the cumulative distribution function (F(x)) must possess the ability to be evaluated at all points observed in the data set. For many distributions (e.g., exponential, logistic, extreme-value, Rayleigh, uniform, Weibull) computations are quite straight forward and $F(x_{[i]})$ is quite easily found, while for some distributions the calculations can be quite difficult. However, many of these distributions have been extensively studied and their cumulative distribution functions are well documented (e.g., normal, Student's t, chi-square, gamma) which again makes application of the W_u test quite easy.

In theory every continuous distribution can be represented, and a test, designed to determine whether a set of observations does in fact arise from a population possessing the distribution of interest, can also be conducted. The values associated with acceptance-rejection regions for these tests will depend on sample size as will the ability of the test to discriminate against "odd" distributions.

The performance of the W_u statistic can be enhanced by expressing a specific alternative hypothesis. However, when investigating distributional assumptions it is generally unrealistic to be able to assume a specific alternative.

D'Agostino and Stephens [17] point out that the Shapiro-Wilk test statistic (where they define $W=(n-1)(X_n-X_1)^2S^2$) (sic) is not consistent, which they go on to say (page 224 of [17]) is the case for most tests based on ratios of two variance estimators.

In conjunction with the probability-integral transformation the W_u test can be used to investigate the nature of any completely specified continuous distribution, which makes the distribution of the W_u statistic invariant with respect to hypothesized distribution. Hence, the W_u test may provide an alternative to the traditional chi-square test procedure, alleviating the need to subjectively create class boundaries necessary in the chi-square analysis.

The general W_u procedure is presented as a tool that is used in conjunction with an associated probability plot. Together the two procedures can provide good insights into the behavior of the data. In those cases where the sample sizes are small (i.e., $n<20$) the procedure is cautioned against. Both the W_u procedure and the probability plot procedure are not very powerful for small sample sizes.

The techniques used to extend the W_u procedure allow other inferences to be drawn from the probability plot as well. Outliers are more easily identified and alternative distributions may be suggested in those cases where the probability plot and the W_u procedure suggest that the assumed distribution function does not adequately describe the observed data.

The multivariate normal procedure is only an approximate procedure. The PIT is performed on the Mahalanobis distances which have been determined from sample results.

As a result the D_i's may not be iid. The PIT is performed on these transformed variables which Healy states are iid under the null hypothesis. Asymptotically the D_i's will be iid.

Healy's procedure is suitable for any MVN distribution function requiring only that χ^2_υ probability paper be available. Chi-square probability paper is unique for each value of υ hence commercially produced paper is not always available. However by performing the PIT associated with the χ^2_υ distribution function, uniform probability paper can be used to examine goodness-of-fit. Uniform probability paper is easily created from ordinary arithmetic graph paper and can be used to investigate all MVN distribution functions.

The PIT (for the completely specified distribution function) and the CPIT (for the case where the functional form is known) permit the W_u procedure to provide an exact test for goodness-of-fit. Approximate tests that involve substituting sample based estimates into the PIT for those missing parameter values are reasonably good for large samples. Results indicate that as sample size increases the transformation produces variates that are very nearly uniform. In addition as (n-p) (n being sample size, p being the number of missing parameter values) increases the dependence becomes negligible.

To examine the "ability" of the random number generator used in the simulations, the W_u statistic was calculated for 10,000 samples of size 10, 20 and 40 from the uniform distribution and the percentage of observations lying outside the W_u bounds (α=0.1) found in Table 2.4.1, calculated. The results were 10.1%, 9.8%, and 10.1% respectively.

As a second check on the generator 10,000 samples of size 3 were generated and the W_u values associated with certain quantiles of cumulative relative frequency calculated. The theoretical p-value (denoted p(distribution)) associated with the W_u result (determined from the simulations) has been determined using the theoretical results of

Theorem 2.5.4. The numerical difference that exists between the simulated and theoretical results is denoted by |p(s)-p(d)|. The results were as follows

W_u	p(simulations)	p(distribution)	|p(s)-p(d)|
1.73649	.01	.0103	.0003
1.75396	.05	.0509	.0009
1.77454	.10	.0996	.0004
1.99671	.90	.9005	.0005
1.99913	.95	.9489	.0011
1.99996	.99	.9890	.0010

Chapter 3

Bayesian Analysis of Process Capability

3.1. INTRODUCTION

The goal of a process capability study is to determine whether the production process is capable of reaching the required tolerance levels. If the actual process spread is greater than the allowable process spread the process is generally deemed incapable of reaching the required tolerances.

The process capability index, Cp, has been introduced as a tool to aid in the assessment of process performance. It is defined as

$$Cp = \frac{USL - LSL}{6\sigma}$$

where the allowable process spread is defined to be the difference between the maximum allowable upper limit (USL) and the minimum allowable lower limit (LSL) of the process, and 6σ, the actual process spread, is a function of the variance (σ^2) of the process.

A process will be judged capable if Cp is greater than some real valued constant c, and incapable if Cp is less than c. In many process studies, c is taken to be one (see Figure 3.1.1).

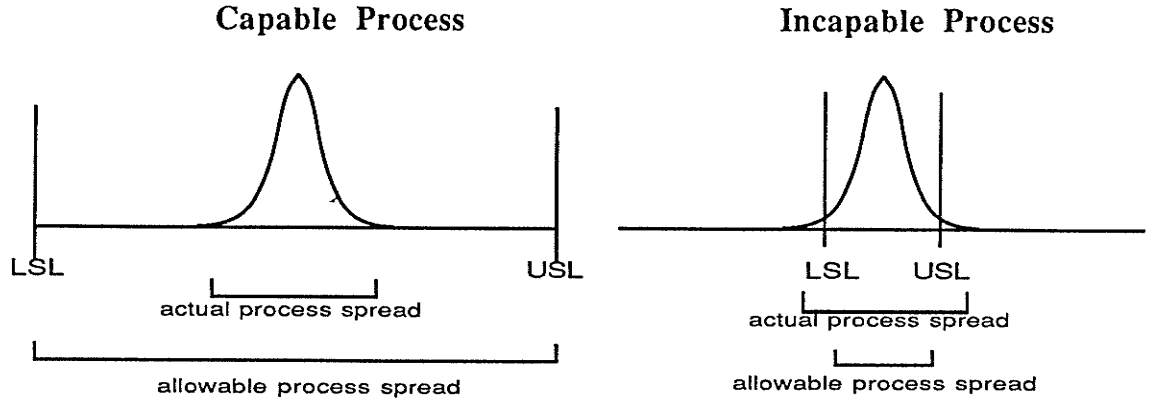


Figure 3.1.1 Examples of a capable and incapable Process.

The general form of the Cp estimator is

$$\hat{C}p = \frac{USL - LSL}{6\hat{\sigma}}$$

where $\hat{\sigma}$ is an estimate of σ. The most frequently used estimator of Cp is

$$\hat{C}p = \frac{USL - LSL}{6s} \qquad (3.1.1)$$

where $s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{(n-1)}$ is the sample variance of n process measurements $x_1, x_2, \ldots, x_n$.

There are many estimators for the population standard deviation, with the most appropriate estimator for any particular application depending upon such things as the distribution associated with the process measurements, the desired qualities of the estimator, and the number of sampling units examined. Due to its extensive use in process capability studies some of the statistical properties associated with the estimator in (3.1.1) are examined.

Practitioners often overlook the fact that $\hat{C}p$ will be stochastic. A bayesian technique for analyzing the outcome of a process capability study is presented as an alternative to the practise of judging a process capable based solely on the value of a point estimator.

In most process capability studies, the process measurements are assumed to come from a normal distribution with mean μ and variance σ^2. If this assumption is correct $\hat{C}p$ can be used to examine some of the characteristics of the production process. However when the assumption of normality can not be made, or is made incorrectly, the statistical properties and procedures formulated under the assumption of normality may be invalid.

The effect departures from normality have on the outcome of a process capability study are also investigated.

3.2. PROPERTIES OF $\hat{C}p$

Under the assumption that the process measurements come from a normal distribution, the following have been found

Theorem 3.2.1: The pdf of $\hat{C}p$ is

$$f(\hat{C}p) = \frac{1}{\Gamma(\frac{n-1}{2})} \frac{1}{2^{\frac{n-1}{2}-1}} \frac{(Cp^2[n-1])^{\frac{n-1}{2}}}{\hat{C}p^n} e^{\frac{-Cp^2(n-1)}{2\hat{C}p^2}} \qquad 0<\hat{C}p<\infty$$

Proof: $X \sim N(\mu,\sigma^2) \quad \Rightarrow \quad \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1} \quad \Rightarrow \quad \frac{(n-1)Cp^2}{\hat{C}p^2} \sim \chi^2_{n-1}$

Let $\hat{C}p = Cp\sqrt{\frac{n-1}{y}}$ where $y \sim \chi^2_{n-1}$.

Then $$g(\hat{C}p) = \frac{1}{\Gamma(\frac{n-1}{2})} \frac{1}{2^{\frac{n-1}{2}-1}} \frac{[Cp^2(n-1)]^{\frac{n-1}{2}}}{\hat{C}p^n} e^{\frac{-Cp^2(n-1)}{2\hat{C}p^2}} \qquad 0<\hat{C}p<\infty$$

Theorem 3.2.2: $\hat{C}p$ is a biased estimator of Cp.

Proof: By definition $E(\hat{C}p) = \int_0^\infty \hat{C}p\, g(\hat{C}p) d\hat{C}p$

Let $r = \frac{Cp^2(n-1)}{2\hat{C}p^2}$ then

$$E(\hat{Cp}) = Cp\sqrt{\frac{n-1}{2}}\frac{\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})}\int_0^{\infty}\frac{1}{\Gamma(\frac{n-2}{2})}r^{\frac{n-2}{2}-1}e^{-r}\,dr = Cp\sqrt{\frac{n-1}{2}}\frac{\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})}$$

$$\text{Hence} \quad E(\hat{Cp}) - Cp = Cp\left[\sqrt{\frac{n-1}{2}}\frac{\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})} - 1\right]$$

Theorem 3.2.3: $\hat{Cp}$ is asymptotically unbiased for Cp.

Proof: $\lim_{n\to\infty} E(\hat{Cp}) = \lim_{n\to\infty} Cp\sqrt{\frac{n-1}{2}}\frac{\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})} = Cp\lim_{n\to\infty}\sqrt{\frac{n-1}{2}}\frac{\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})} = Cp$

from [18] (page 257).

Theorem 3.2.4: $\hat{Cp}$ is mean square consistent.

Proof: Similar to the Proof of Theorem 3.2.2 one can show

$$MSE(\hat{Cp}) = \frac{(n-1)Cp^2}{2}\left[\frac{\Gamma(\frac{n-3}{2})}{\Gamma(\frac{n-1}{2})} - \frac{\Gamma^2(\frac{n-2}{2})}{\Gamma^2(\frac{n-1}{2})}\right]$$

Then the $\lim_{n\to\infty} MSE(\hat{Cp})$ becomes the difference between two limits we have shown to be one,

i.e., $\lim_{n\to\infty} MSE(\hat{Cp}) = 0$

The relative bias of $\hat{C}p$ (i.e., Bias/Cp) is quite small (less than 3%) for moderately sized samples (i.e., n>30) which in conjunction with its MSE consistency, makes $\hat{C}p$ a reasonable estimator of Cp. Note that the above properties all depend upon the assumption of normality and may be invalid when the normal assumption is not true.

Several authors ([19], [20]) point out that departures from normality may have a serious effect on the ability of the sample variance (s^2) to accurately depict the population variance (σ^2). These results cast doubt on the ability of $\hat{C}p$ to estimate Cp in the presence of non-normal process distributions.

3.3. A Test Procedure and its Robustness

Because $\hat{C}p$ is stochastic, the decision rule associated with judging a process capable using $\hat{C}p$ will depend upon the distribution of $\hat{C}p$ and the level of confidence that will be associated with the resulting decision. The hypotheses of interest is of the form

H_0: The process is capable versus H_a: The process is incapable

which can be rewritten in the following form

$$H_0: Cp \geq c \qquad \text{versus} \qquad H_a: Cp < c$$

where c is some real valued constant. Then by considering the conditional probability

$$Pr(\hat{C}p \geq b \mid Cp = c)$$

for some hypothesized value of c and for $p = 1 - \alpha$, where α is the probability of making a Type-I error, b will become the critical value associated with judging a process capable.

The decision rule associated with the hypothesis will be to reject $H_0: Cp \geq c$ and conclude that the process is not capable at the α level of significance if $\hat{C}p < b$ and fail to

reject H_0:Cp $\geq$ c, and conclude that the process is capable at the α level of significance if $\hat{Cp} \geq b$.

Under the assumption that the process measurements follow a normal distribution,

$$p = Pr(\hat{Cp} \geq b \mid Cp = c) = \int_b^{\infty} f(y \mid Cp = c)dy$$

$$= Pr(\hat{Cp} \geq b \mid Cp=c) = Pr((n-1)s^2 \leq (n-1)\frac{T^2}{6^2 b^2} \mid Cp=c)$$

$$= Pr((n-1)s^2 \leq \frac{(n-1)c^2\sigma^2}{b^2} \mid Cp=c)$$

where $v=(n-1)s^2$ will have the following pdf

$$h(v) = \left[\frac{1}{2\sigma^2}\right]^{\frac{n-1}{2}} \left[\frac{1}{\Gamma(\frac{n-1}{2})}\right] v^{\frac{n-1}{2}-1} e^{-\frac{v}{2\sigma^2}} \qquad 0<v<\infty \quad . \quad (3.3.1)$$

For fixed values of σ, c, n and p it is simply a matter of evaluating an incomplete gamma function to determine the critical value b .

When the distribution of the process measurements is not normal, the distribution of $\hat{Cp}$ will change, which may cause the critical value b to change as well. If the critical values change incorrect inferences may result. To avoid this, the robustness of the Cp procedure was examined. One general method that can be used involves finding the distribution of $\hat{Cp}$ for various distributions of the process measurements and then determining the appropriate critical value b' where

$$p = Pr(\hat{Cp} \geq b' \mid Cp = c) \quad ,$$

for these various distributions.

An alternate procedure derived from a technique first discussed in [21] for the t-distribution can be considered. In order to examine the robustness of the t-distribution with respect to departures from normality, Gayen found the joint distribution of $n\bar{x}$ and $(n-1)s^2$ for the population distribution specified by the truncated Edgeworth series. Using only those terms up to and including the fourth population cumulant, Gayen found an approximation of the t-distribution that allowed him to emulate changes in the population distribution by varying the numerical values of the third and fourth cumulants.

Consider a population with mean 0 and variance 1, integrating the joint distribution of $n\bar{x}$ and $(n-1)s^2$ for $n\bar{x}$ over the range $(-\infty, \infty)$, the marginal Edgeworth series for $v=(n-1)s^2$ is

$$g(.,v)=h(v)[1+\lambda_4\frac{(n-1)^2}{8n}\{\frac{v^2}{(n+1)(n-1)}-\frac{2v}{n-1}+1\}$$

$$+\lambda_3^2\frac{(n-1)(n-2)}{12n}\{\frac{v^3}{(n+3)(n+1)(n-1)}-\frac{3v^2}{(n+1)(n-1)}+\frac{3v}{n-1}-1\}]$$

where h(v) is equivalent to (3.3.1) with $\sigma^2=1$, λ_3 is the third cumulant and λ_4 the fourth cumulant.

Using the relationship

$$p = Pr(\hat{Cp} \geq b \mid Cp = c) = Pr((n-1)s^2 \leq (n-1)\frac{T^2}{6^2b^2} \mid Cp = c)$$

$$= \int_0^{(n-1)\frac{T^2}{6^2b^2}} g(\cdot, v)dv = \int_0^{(n-1)\frac{c^2}{b^2}} g(\cdot, v)dv$$

the effect non-normality has on the conditional probability statement for values of n=5(5)100, 200, 300 and c/b = 1, 1.5, 2 has been examined. Tables 3.3.1, 3.3.2 and 3.3.3 include the value of p associated with the normal theory result, the maximum and minimum values of p associated with varying λ_3 (0.0 (0.1) 0.4) and λ_4 (0.0 (0.5) 4.0) (chosen to ensure g(·,v) is a proper pdf (see [22])), the maximum change in p due to variations in λ_3 only, and the maximum change in p due to variations in λ_4 only.

Table 3.3.1

The values of p with c/b=1.0 for various sample sizes.

n	Normal p	min. p	max. p	maxΔp varying λ_3 only	maxΔp varying λ_4 only
5	.5940	.5940	.7384	.0000	.1444
10	.5627	.5627	.7018	.0031	.1359
15	.5503	.5503	.6762	.0042	.1217
20	.5432	.5432	.6580	.0045	.1103
25	.5384	.5384	.6444	.0046	.1013
30	.5349	.5349	.6337	.0046	.0920
35	.5323	.5323	.6251	.0045	.0883
40	.5301	.5301	.6180	.0044	.0834
45	.5284	.5284	.6119	.0043	.0792
50	.5269	.5269	.6067	.0042	.0756
100	.5189	.5189	.5772	.0033	.0549
200	.5133	.5133	.5552	.0025	.0394
300	.5109	.5109	.5453	.0021	.0323

Table 3.3.2

The values of p with c/b=1.5 for various sample sizes.

n	Normal p	min. p	max. p	maxΔp varying λ_3 only	maxΔp varying λ_4 only
5	.9389	.8489	.9417	.0028	.0900
10	.9836	.8953	.9836	.0011	.0872
15	.9953	.9502	.9953	.0014	.0436
20	.9986	.9793	.9986	.0008	.0185
25	.9996	.9919	.9996	.0004	.0073
30	.9999	.9969	.9999	.0002	.0028
35	.99996	.9989	1.0000	.0001	.0010
40	.99999	.9996	1.0000	.0000	.0004
45	1.0000	.9999	1.0000	.0000	.0001
50	1.0000	.9999	1.0000	.0000	.0000
100	1.0000	1.0000	1.0000	.0000	.0000
200	1.0000	1.0000	1.0000	.0000	.0000
300	1.0000	1.0000	1.0000	.0000	.0000

Table 3.3.3

The values of p with c/b=2.0 for various sample sizes.

n	Normal p	min. p	max. p	maxΔp varying λ_3 only	maxΔp varying λ_4 only
5	.9970	.9675	.9970	.0009	.0286
10	.9999	.9987	1.0000	.0001	.0012
15	1.0000	1.0000	1.0000	.0000	.0000
20	1.0000	1.0000	1.0000	.0000	.0000
25	1.0000	1.0000	1.0000	.0000	.0000
30	1.0000	1.0000	1.0000	.0000	.0000
35	1.0000	1.0000	1.0000	.0000	.0000
40	1.0000	1.0000	1.0000	.0000	.0000
45	1.0000	1.0000	1.0000	.0000	.0000
50	1.0000	1.0000	1.0000	.0000	.0000
100	1.0000	1.0000	1.0000	.0000	.0000
200	1.0000	1.0000	1.0000	.0000	.0000
300	1.0000	1.0000	1.0000	.0000	.0000

The results suggest that the value of p, particularly for smaller sample sizes, is sufficiently unstable to jeopardize inferences that may be drawn from a study. For example from Table 3.3.1 the value of p for a sample of size 10 ranges from 0.5627 to 0.7018 for various combinations of λ_3 and λ_4, with the normal theory value being 0.5627. The range of values that p can take on is quite substantial even for large sample sizes, as witnessed by the fact that for samples of size 100 the value of p ranges from 0.5189 to 0.5772 with the normal theory value being 0.5189. These results suggest that $\hat{C}p$ is non-robust with respect to departures from normality.

A further example of the non-robustness of $\hat{C}p$ to departures from normality is illustrated by the results summarized in Table 3.3.4. In this case the value of c/b that would ensure p=.95 for the normal distribution was fixed, and the probabilities associated with the various values of λ_3 (0.0 (0.1) 0.4) and λ_4 (0.0(1.0)4.0) calculated. This allows examination of the robustness of $\hat{C}p$ in the tail areas of the distribution. For example it was found for a sample of size 40 that the minimum value of p that arose for the various combinations of λ_3 and λ_4 was 0.7844, while the normal distribution value of p is 0.95. This represents a difference of 0.1656. Similar to the results observed above, the skewness of the distribution appears to have little effect on the value of p, while the kurtosis appears to affect the value of p significantly. Again using sample size 40 as an example, 0.1647 of the overall difference of 0.1656 in the normal distribution value and the minimum value of p is attributable to the kurtosis of the distribution, while only 0.0010 is attributable to the skewness.

Table 3.3.4.

Ranges of p by varying values of λ_3 and λ_4 where p=0.95 in the normal distribution.

n	c/b	min. p	maxΔp varying λ_3 only	maxΔp varying λ_4 only
5	1.5401	.8589	.0023	.0911
10	1.3711	.8186	.0011	.1314
15	1.3001	.8037	.0007	.1463
20	1.2596	.7960	.0002	.1538
25	1.2318	.7912	.0005	.1583
30	1.2114	.7881	.0007	.1743
35	1.1956	.7860	.0008	.1632
40	1.1829	.7844	.0009	.1647
45	1.1724	.7833	.0009	.1658
50	1.1636	.7823	.0010	.1667
55	1.1559	.7816	.0010	.1674
60	1.1493	.7811	.0010	.1679
65	1.1434	.7805	.0010	.1684
70	1.1382	.7802	.0010	.1688
75	1.1335	.7800	.0010	.1691
80	1.1293	.7796	.0010	.1694
85	1.1254	.7794	.0010	.1696
90	1.1219	.7792	.0010	.1698
95	1.1187	.7789	.0010	.1700
100	1.1157	.7788	.0010	.1702

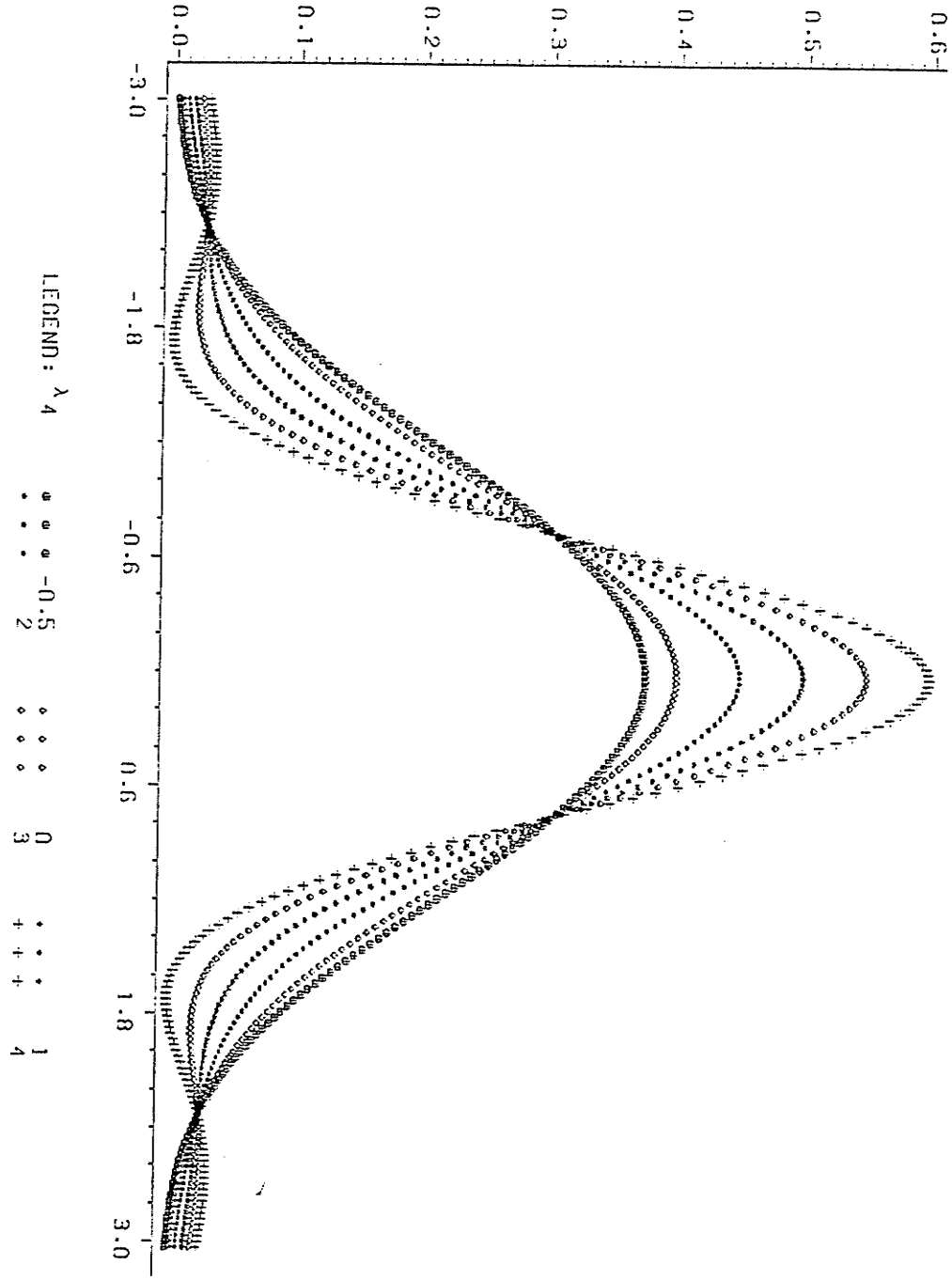


Figure 3.3.1 . Pdfs associated with the process measurements for various λ_4.

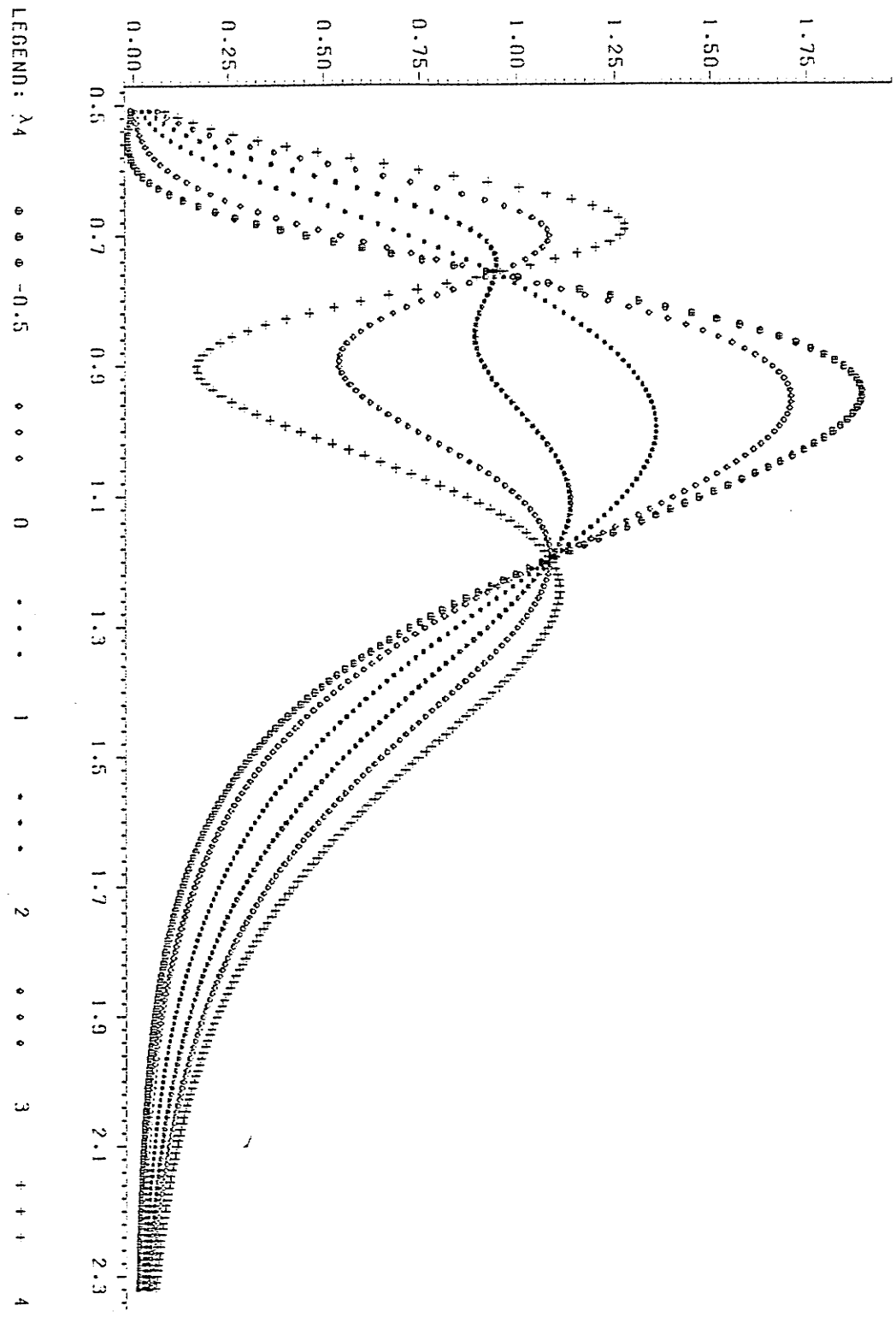


Figure 3.3.2 . Pdfs associated with $\hat{C}p$ for n=10 and various values of λ_4.

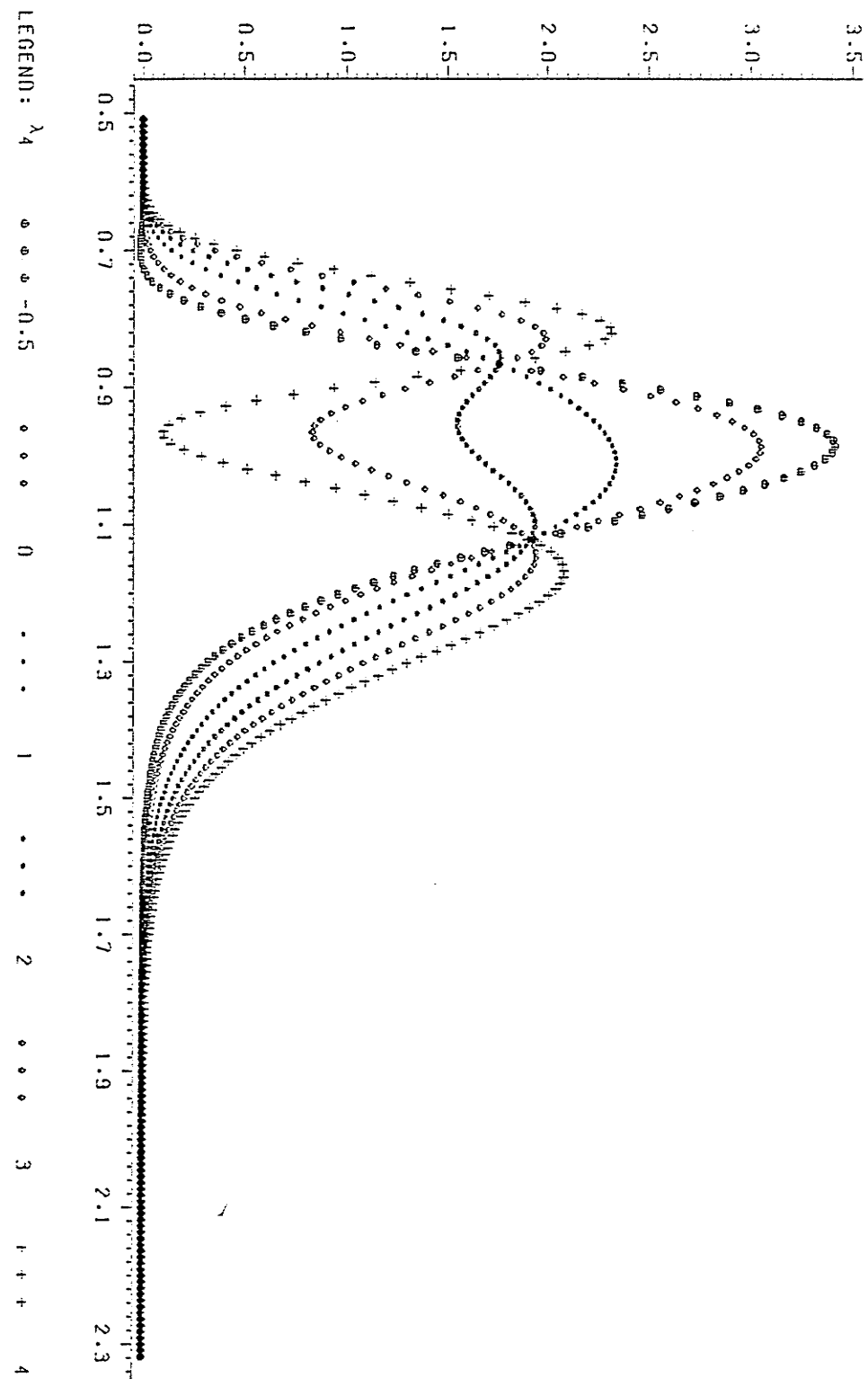


Figure 3.3.3 . Pdfs associated with $\hat{C}p$ for n=30 and various values of λ_4.

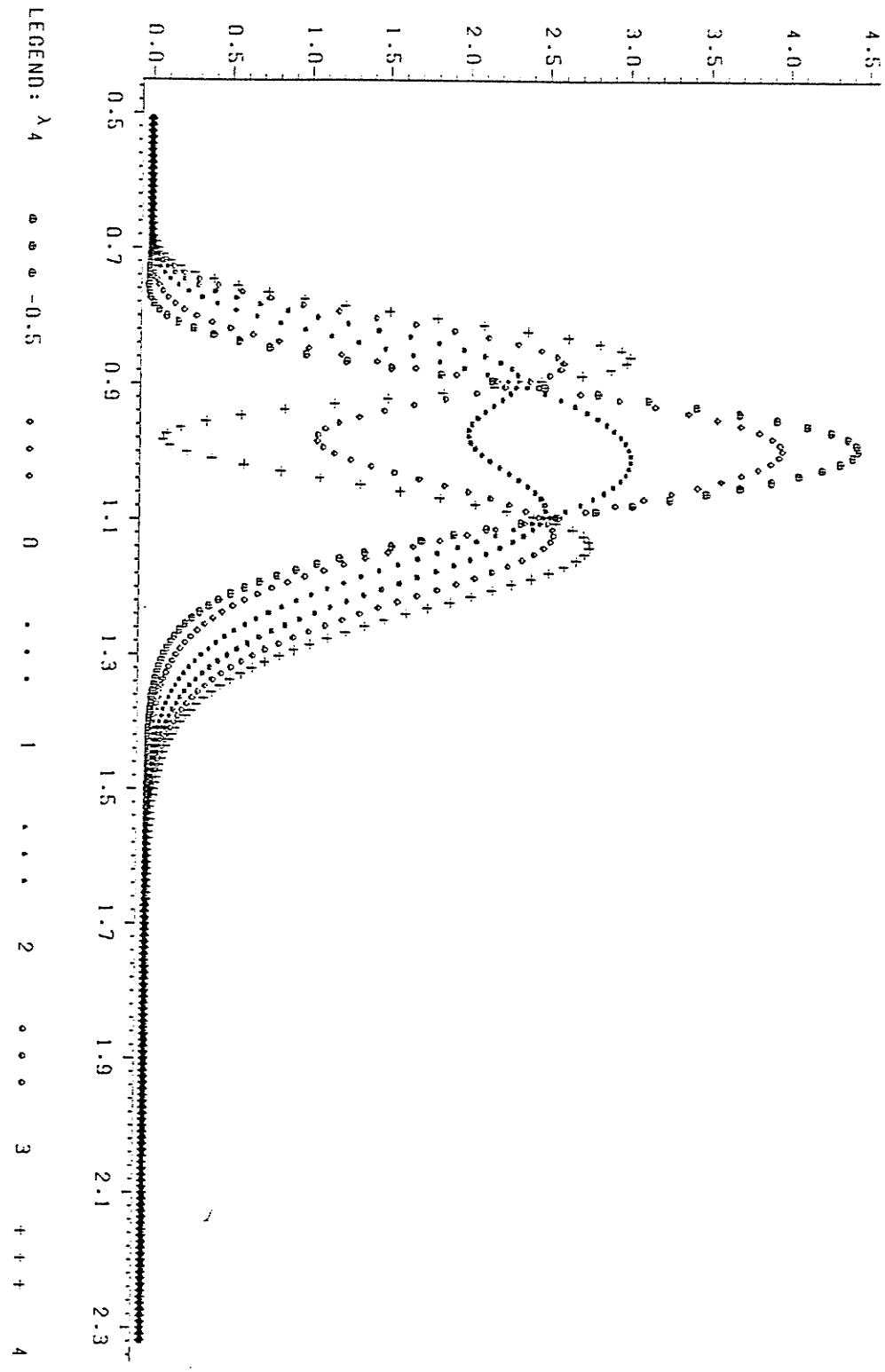


Figure 3.3.4 . Pdfs associated with $\hat{C}p$ for n=50 and various values of λ_4.

To further illustrate the effect that the standardized fourth moment has on the distribution of $\hat{C}p$, the pdfs associated with six values of λ_4 (-0.5, 0.0, 1.0, 2.0, 3.0 and 4.0) have been sketched, first for the distribution of the process measurements (Figure 3.3.1) and then for the distribution of $\hat{C}p$ for the samples sizes 10, 30 and 50 (Figures 3.3.2, 3.3.3 and 3.3.4). Several interesting results are apparent from these sketches. Relatively small changes in the shape of the pdf associated with the process measurements result in substantial changes in the shape of the pdf associated with $\hat{C}p$. Although the peakedness of the six distributions associated with the process measurements is markedly different, for moderately sized samples one would have difficulty in discerning between these distributions. The same is not true for the pdfs associated with $\hat{C}p$. In each of the Figures 3.3.2, 3.3.3 and 3.3.4 the shape of the pdfs range from a slightly skewed unimodal curve to a skewed bimodal curve.

The bimodal nature of the pdfs associated with $\hat{C}p$ is also of interest, as the central tendencies of those pdfs with λ_4=-0.5, 0.0 and 1.0 are reversed from those pdfs with λ_4=2.0, 3.0 and 4.0. The pdfs of $\hat{C}p$ associated with λ_4=2.0, 3.0 and 4.0 indicate that one should expect clustering of sample results away from $\hat{C}p$ =1, while for those pdfs associated with λ_4 = -0.5, 0.0 and 1.0 one would expect clustering of sample results in the vicinity of $\hat{C}p$ =1. Thus for those cases where λ_4 = 2.0, 3.0 and 4.0 estimated values of Cp in the neighborhood of one will be "relatively rare".

3.4. CORRECTION FACTORS FOR THE OC CURVE

Correction factors for the Cp procedure that take into account the degree with which the distribution of the process measurements departs from the assumed normal distribution can be determined. These correction factors allow one to use the $\hat{C}p$ procedure for analyzing a capability study when the process measurements are identifiably non-normal.

Formulating the correction factors involved finding the value of c/b that would ensure p=0.95 for various values of n, where the process measurements were assumed to come from a family of populations with non-zero values of the fourth cumulant. The effect of the third cumulant on the value of p was ignored as over the range considered the skewness of the distribution had little effect on the value of p. The corrected values for c/b of the conditional probabilities were found for λ_4 = (-0.5, 1, 2, 3 and 4) and included in Table 3.4.1.

Table 3.4.1

Corrected values of c/b for five values of λ_4 and the normal distribution (λ_4=0), p=0.95.

	λ_4					
n	-0.5	0	1.0	2.0	3.0	4.0
5	1.500	1.540	1.643	1.749	1.832	1.895
10	1.333	1.371	1.458	1.530	1.580	1.617
15	1.267	1.300	1.374	1.430	1.469	1.496
20	1.229	1.260	1.323	1.370	1.402	1.426
25	1.204	1.232	1.289	1.330	1.358	1.378
30	1.185	1.211	1.263	1.300	1.325	1.343
35	1.171	1.196	1.243	1.277	1.300	1.317
40	1.160	1.183	1.227	1.258	1.280	1.295
45	1.151	1.172	1.214	1.243	1.263	1.277
50	1.142	1.163	1.203	1.230	1.249	1.262
55	1.136	1.156	1.194	1.220	1.237	1.250
60	1.130	1.149	1.185	1.210	1.226	1.239
65	1.125	1.143	1.178	1.201	1.217	1.229
70	1.121	1.138	1.171	1.193	1.209	1.220
75	1.116	1.134	1.165	1.187	1.201	1.213
80	1.113	1.129	1.160	1.181	1.195	1.206
85	1.109	1.125	1.155	1.175	1.189	1.199
90	1.105	1.122	1.151	1.170	1.183	1.192
95	1.103	1.119	1.147	1.165	1.178	1.188
100	1.101	1.116	1.143	1.161	1.174	1.183

Corrections were also made to the analytical procedure Kane presented in [23], [24] that uses the Operating Characteristic (OC) curve approach to judge the capability of a process. In the OC curve approach the hypothesis of interest remains the same as the hypothesis discussed earlier. However rather than conditioning on a single value of Cp, the OC curve approach uses two values of Cp, Cp(A) and Cp(R) such that

$$p = Pr(\hat{Cp} \geq b \mid Cp = Cp(A)) = 1-\alpha$$

and

$$p = Pr(\hat{Cp} \geq b \mid Cp = Cp(R)) = \beta$$

where α is the probability of a Type-I error, β is the probability of a Type-II error and Cp(A) > Cp(R). The acceptable quality level (AQL) is often used for Cp(A) and the rejectable quality level (RQL) for Cp(R) but this is not absolutely necessary as any reasonable values may be used.

Fixing the values of α, β and one of either Cp(A) or Cp(R) allows the critical value b to be determined. However, rather than fixing an explicit value for either Cp(A) or Cp(R) the results can be summarized using the ratios

$$\frac{Cp(R)}{Cp(A)} = \sqrt{\frac{\chi^2_{n-1}(1-\beta)}{\chi^2_{n-1}(\alpha)}}$$

and

$$\frac{b}{Cp(R)} = \sqrt{\frac{n-1}{\chi^2_{n-1}(1-\beta)}}$$

where χ^2_{n-1} denotes the chi-squared distribution with n-1 degrees of freedom. The value of these ratios have been corrected (for $\lambda_4 = -0.5$, 1, 2, 3 and 4 with $\alpha=\beta=.05$) and included in Table 3.4.2.

Table 3.4.2

Correction Ratios for Kane's OC curve approach for $\hat{C}p$.

	λ_4									
n	-0.5		1.0		2.0		3.0		4.0	
5	3.39	.442	4.21	.390	4.78	.366	5.28	.347	5.71	.332
10	2.10	.634	2.57	.568	2.83	.541	3.04	.520	3.20	.505
15	1.78	.711	2.13	.646	2.31	.620	2.44	.601	2.55	.586
20	1.63	.754	1.91	.693	2.05	.668	2.15	.651	2.24	.637
25	1.54	.783	1.78	.725	1.90	.701	1.98	.685	2.05	.672
30	1.48	.803	1.69	.748	1.79	.726	1.87	.710	1.92	.699
35	1.43	.819	1.63	.766	1.71	.745	1.78	.731	1.83	.720
40	1.39	.832	1.57	.781	1.65	.761	1.71	.747	1.76	.737
45	1.37	.842	1.53	.794	1.61	.774	1.66	.761	1.70	.751
50	1.34	.850	1.50	.804	1.57	.785	1.62	.773	1.65	.763
55	1.32	.858	1.47	.813	1.53	.795	1.58	.783	1.62	.773
60	1.31	.864	1.44	.821	1.51	.803	1.55	.791	1.58	.783
65	1.29	.870	1.42	.828	1.48	.811	1.52	.799	1.55	.791
70	1.28	.875	1.40	.834	1.46	.818	1.50	.806	1.53	.798
75	1.27	.879	1.39	.839	1.44	.824	1.48	.813	1.51	.804
80	1.26	.883	1.37	.844	1.42	.829	1.46	.818	1.49	.810
85	1.25	.887	1.36	.849	1.41	.834	1.44	.824	1.47	.816
90	1.24	.890	1.35	.853	1.39	.839	1.43	.828	1.45	.821
95	1.24	.893	1.34	.857	1.38	.843	1.41	.833	1.44	.825
100	1.23	.896	1.33	.861	1.37	.847	1.40	.837	1.43	.830

3.5. EXAMPLES

Example 3.5.1: The data were taken from question 5.1 page 113 of [25], where it was stated that a sample of size 54 was taken from a process. Assuming that the upper and lower allowable limits of the process are 87 and 115 , a capability analysis was conducted. The following sample values were found:

minimum ---------- 90

maximum ---------- 115

mean ---------- 101

standard deviation ----- 5.40

kurtosis ---------- 0.9577

A normal probability plot was constructed (Figure 3.5.1) which along with an estimated kurtosis of 1, suggested that the critical value b should be adjusted to reflect the apparent departure of the process measurements from normality. For illustration purposes the moment estimator of λ_4 is used to adjust for non-normality.

The estimated value of Cp was found to be

$$\hat{Cp} = \frac{[115 - 87]}{6(5.40)} = 0.8641$$

Under the assumption that the data comes from a population with a normal distribution, for the hypotheses

H_0: $Cp \geq 1$ versus H_a: $Cp < 1$

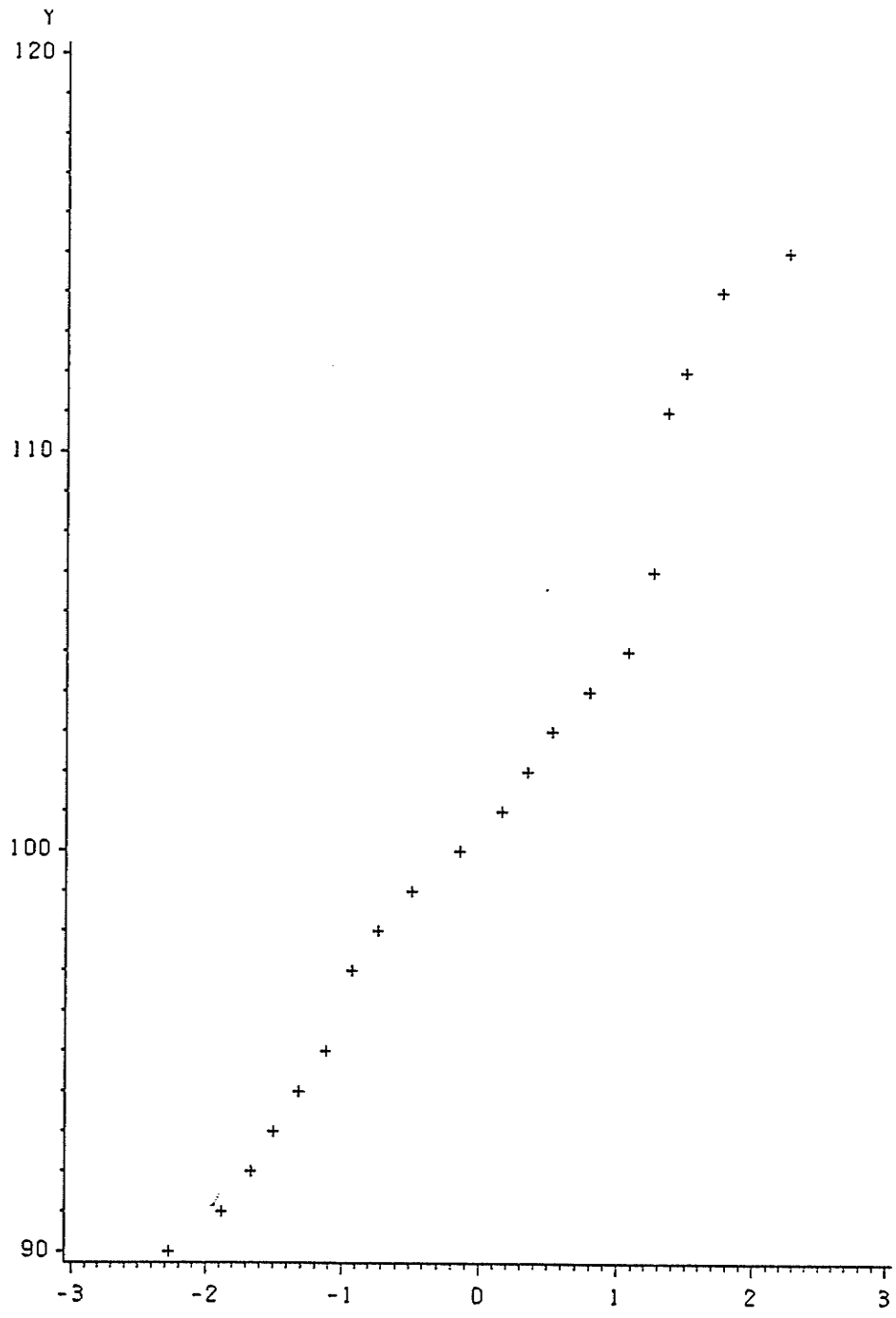


Figure 3.5.1. Probability Plot for Example 3.5.1.

with n=54 and $\alpha = .05$, the critical value b is 0.8643. Then since the estimated value of Cp is less than the critical value, the process would be judged incapable of meeting the imposed limits. However, adjusting the critical value to reflect the fact that the data appear to arise from a moderately non-normal distribution, the critical value for $\lambda_4 = 1$, $\alpha = 0.05$ and n=54 is 0.8368. On the basis of this result the process would be judged capable at the .05 level of significance.

Example 3.5.2: A second example arising from a generated dataset of 30 observations illustrates the same concept. Assuming that the following observations come from a process where the upper and lower specification limits are 2.80 and -2.80 respectively

-2.8227	-0.4670	0.5606	-0.3823	0.1166	0.5807
0.7059	0.9571	0.2129	-0.3924	-2.7676	-1.1087
0.1631	2.0094	-2.4940	0.0809	0.5543	0.9775
0.1323	-0.3611	0.2311	-0.2496	0.3560	-0.2607
-0.5868	0.3911	-0.0115	2.5092	-0.1194	-0.0349

the following estimates were found

minimum ---------- -2.8227

maximum ---------- 2.5092

mean ---------- -0.0507

standard deviation ------- 1.1484

kurtosis ---------- 1.8953

Then $\hat{Cp} = \frac{[2.80 - (-2.80)]}{6(1.1484)} = 0.813$, assuming that the measurements arise from a Normal distribution, the critical value (from Table 3.3.1) for the hypothesis

H_0: $Cp \geq 1$ versus H_a: $Cp < 1$

is 0.826, H_0 is rejected at the 0.05 level of significance. However adjusting for an estimated kurtosis of approximately 2 the critical value for the above hypotheses (from Table 3.4.1) will be 0.769 suggesting that the process is capable.

A probability plot (Figure 3.5.2) of the data suggests that the observations in the tail regions of the plot do not behave as though arising from a normal distribution . Hence the probability plot was again useful in indicating that the measurements were substantially non-normal, suggesting that the critical value for the hypothesis should be modified.

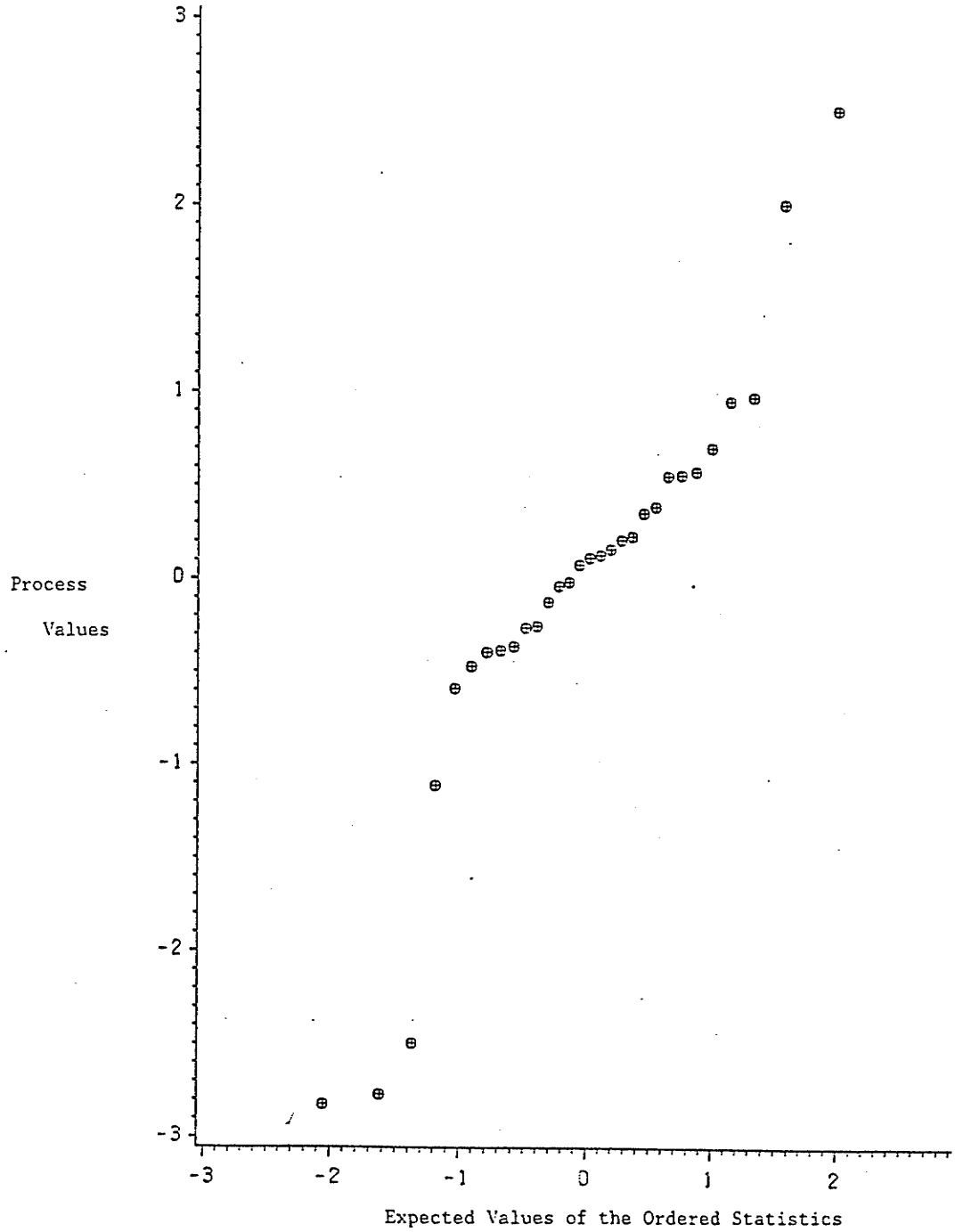


Figure 3.5.2. Probability Plot for Example 3.5.2.

3.6. ASSESSING PROCESS CAPABILITY: A BAYESIAN APPROACH

A second important concept, often overlooked by quality control practitioners, is that, regardless of form, $\hat{C}p$ will be stochastic. Consider a process that will be deemed capable if $Cp \geq 1$ and incapable if $Cp < 1$, but where the actual value of Cp is unknown. The current practice is to estimate Cp from the sampling results using an appropriate estimator and to judge the process capable if $\hat{C}p \geq 1$ or incapable if $\hat{C}p < 1$. The use of point estimators as the sole criteria for judging a process capable can be misleading.

Kane [23] presents a pair of techniques that consider the stochastic nature of $\hat{C}p$, while Wies and Burr [26] have developed a sequential technique based on the sample range that provides confidence levels for the estimated process capability. Both techniques circumvent the problem of having to know the explicit value of Cp, while neither completely eliminates the need to know, or at least condition on, the actual value of Cp.

An alternative to the approaches presented in [23] and [26] has been derived using Bayesian statistical techniques. The general Bayesian approach assumes that the parameter of interest (in this case a function of the process standard deviation, σ) is stochastic with an associated statistical distribution. The sampling results are used to adjust the assumed (prior) distribution to reflect the behaviour of the data. The end result is a posterior distribution for the parameter of interest, which is based on an assumed prior distribution and the likelihood function associated with the observed measurements.

In choosing a prior distribution that reflects a general "lack" of knowledge or information (i.e., a non-informative or diffuse prior), the posterior distribution is said to be data dominated [27]. This result reduces the subjectivity inherent in the general Bayesian approach (i.e., the choice of prior) from that encountered and criticized when an

informative prior distribution is used. Some informative priors require fairly large sample sizes before the posterior distribution becomes data dominated. For this reason conservatively choosing a non-informative prior distribution for the general Bayesian approach to analyzing Cp, will be the approach taken here. In those cases where good prior information exists an appropriate informative prior should enhance the inference drawn.

Under the assumption that the process measurements follow a normal distribution (i.e., $x \sim N(\mu,\sigma^2)$), the likelihood function for the sample $\underset{\sim}{X} = \{x_1, x_2, x_3, \ldots, x_n\}$ will be

$$L(\mu,\sigma^2)=\frac{1}{[2\pi\sigma^2]^{n/2}} \exp -\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2} ,$$

assuming a non-informative prior for μ and σ of the form

$$h(\mu,\sigma^2) = \frac{1}{\sigma} , \qquad -\infty<\mu<\infty,\ 0<\sigma<\infty$$

the marginal posterior distribution of σ is

$$f(\sigma|\underset{\sim}{X})= \frac{2}{\Gamma\left[\frac{n-1}{2}\right]}\left[\frac{(n-1)s^2}{2}\right]^{\frac{n-1}{2}} \frac{1}{\sigma^n} e^{-\frac{(n-1)s^2}{2\sigma^2}} . \qquad (3.6.1)$$

Now consider the following statement

$$\Pr(\text{process is capable} \mid \text{sample}) = \Pr(Cp > c \mid \underset{\sim}{X})$$

$$=\Pr\left(\frac{USL-LSL}{6\sigma} > c \mid \underset{\sim}{X}\right)$$

$$=\Pr(\sigma < \frac{\text{USL-LSL}}{6c} \mid \underset{\sim}{X}).$$

Therefore evaluating Pr(process is capable $\mid \underset{\sim}{X}$) is equivalent to determining a one-sided credible interval for σ. That is

$$\Pr(\text{process is capable} \mid \underset{\sim}{X}) = \int_0^{\frac{\text{USL-LSL}}{6c}} \frac{2}{\Gamma\left[\frac{n-1}{2}\right]} \left[\frac{(n-1)s^2}{2}\right]^{\frac{n-1}{2}} \frac{1}{\sigma^n} e^{-\frac{(n-1)s^2}{2\sigma^2}} d\sigma$$

$$= b \int_0^{\frac{\text{USL-LSL}}{6c}} \sigma^{-n} e^{-\frac{a}{\sigma^2}} d\sigma \qquad (3.6.2)$$

where a and b are constants such that $a = \frac{(n-1)s^2}{2}$ and $b = \frac{2}{\Gamma[\frac{n-1}{2}]} [\frac{(n-1)s^2}{2}]^{(n-1)/2}$. In order to evaluate the Pr(process is capable | sample) the sample size, n, the sample variance, s^2, the specification limits, USL and LSL, and the critical value, c, must be known. If this is the case then the Pr(process is capable | sample) can be determined simply by evaluating the incomplete gamma function. The minimum values of $\hat{C}p$ that will ensure probabilities of 0.90, 0.95, and 0.99 for c = 1, 4/3, 5/3 and n = 5(1)20, 25(5)50, 60(10)100 have been determined and included in Tables 3.6.1, 3.6.2 and 3.6.3.

Equation (3.6.2) in conjunction with the Wilson-Hilferty [28] transformation provides an approximation that is easy to evaluate and good for most values of n. The resulting algorithm requires evaluation of a standard normal function rather than an incomplete gamma function. Using the posterior distribution of σ (i.e., equation (3.6.1)) and letting $y=(n-1)s^2/\sigma^2$, results in

$$g(y|\underset{\sim}{X}) = \left[\frac{1}{2}\right]^{\frac{n-1}{2}} \frac{1}{\Gamma\left[\frac{n-1}{2}\right]} y^{\frac{n-1}{2}-1} e^{-\frac{y}{2}} \qquad 0<y<\infty$$

which is the pdf of a χ^2_{n-1} random variable. The Wilson-Hilferty [28] transformation states that in the case of a normal distribution $(s^2/\sigma^2)^{(1/3)}$ follows an approximate normal distribution (i.e., $N(1-\frac{2}{9(n-1)}, \frac{2}{9(n-1)})$). Hence

$$\Pr(Cp > c \mid \underset{\sim}{X}) = \Pr\left(\frac{s^2}{\sigma^2} > 36\frac{c^2 s^2}{T^2}\right)$$

$$= \Pr\left(\left(\frac{s^2}{\sigma^2}\right)^{(1/3)} > \left(\frac{c}{\hat{Cp}}\right)^{(2/3)}\right)$$

$$\cong 1-\Phi\left\{\frac{\left(\frac{c}{\hat{Cp}}\right)^{(2/3)} - \left(1-\frac{2}{9[n-1]}\right)}{\sqrt{\frac{2}{9(n-1)}}}\right\}$$

where Φ denotes the cdf of the standard normal distribution.

With the knowledge of the sample results, an approximate probability statement regarding the capability of a process can be determined. In addition, by fixing the probability of judging a process capable given the sample (e.g., p=0.95) it is easy to determine the approximate $\hat{Cp}$ (i.e., $\tilde{Cp}$) that will be required to attain this probability. For example suppose that one wishes to ensure that the $\Pr(Cp > 1 \mid \underset{\sim}{X}) = 0.95$ for a process under study, then using the Bayesian approach, $\Pr(Cp \geq 1 \mid \underset{\sim}{X}) = 0.95$

$$\Rightarrow 1 - \Phi\left\{\frac{\tilde{C}_p^{-2/3} - \frac{9n-11}{9n-9}}{\sqrt{\frac{2}{9n-9}}}\right\} = 0.95$$

$$\Rightarrow \Phi\left\{\frac{\tilde{C}_p^{-2/3} - \frac{9n-11}{9n-9}}{\sqrt{\frac{2}{9n-9}}}\right\} = 0.05$$

$$\Rightarrow \tilde{C}_p = \left(\sqrt{\frac{2}{9n-9}}\left[\Phi^{-1}(0.05)\right] + \frac{9n-11}{9n-9}\right)^{-3/2}$$

The value $\tilde{C}_p$ required to ensure that $\Pr(C_p \geq 1 | \underset{\sim}{X}) = 0.95$ can then be determined.

Values of $\tilde{C}_p$ associated with n = 5(1)20, 25(5)50, 60(10)100, p = 0.90, 0.95, 0.99 and for c = 1, 4/3, 5/3 have been included in Tables 3.6.1, 3.6.2 and 3.6.3.

Table 3.6.1

Minimum Values of $\hat{C}p$ and $\tilde{C}p$ required to ensure $Pr(Cp \geq 1 | \underset{\sim}{X}) = p$.

	p=0.99		p=0.95		p=0.90	
n	$\hat{C}p$	$\tilde{C}p$	$\hat{C}p$	$\tilde{C}p$	$\hat{C}p$	$\tilde{C}p$
5	3.6692	4.0242	2.3724	2.4074	1.9393	1.9406
6	3.0034	3.1603	2.0893	2.1054	1.7621	1.7611
7	2.6230	2.7093	1.9154	1.9243	1.6499	1.6484
8	2.3769	2.4310	1.7972	1.8026	1.5719	1.5703
9	2.2043	2.2413	1.7110	1.7146	1.5141	1.5126
10	2.0762	2.1032	1.6452	1.6477	1.4694	1.4681
11	1.9771	1.9977	1.5931	1.5949	1.4337	1.4324
12	1.8980	1.9143	1.5506	1.5520	1.4043	1.4032
13	1.8333	1.8465	1.5153	1.5164	1.3797	1.3786
14	1.7792	1.7902	1.4854	1.4863	1.3588	1.3578
15	1.7332	1.7425	1.4597	1.4604	1.3406	1.3397
16	1.6936	1.7017	1.4373	1.4379	1.3248	1.3239
17	1.6592	1.6662	1.4176	1.4181	1.3108	1.3100
18	1.6288	1.6350	1.4001	1.4006	1.2983	1.2976
19	1.6019	1.6074	1.3845	1.3849	1.2871	1.2864
20	1.5778	1.5827	1.3704	1.3707	1.2770	1.2763
25	1.4868	1.4900	1.3165	1.3166	1.2380	1.2375
30	1.4262	1.4286	1.2797	1.2798	1.2112	1.2108
35	1.3825	1.3843	1.2528	1.2528	1.1914	1.1910
40	1.3492	1.3506	1.2320	1.2320	1.1761	1.1758
45	1.3227	1.3240	1.2154	1.2154	1.1638	1.1635
50	1.3012	1.3023	1.2017	1.2018	1.1536	1.1534
60	1.2680	1.2688	1.1805	1.1805	1.1378	1.1375
70	1.2433	1.2440	1.1645	1.1646	1.1258	1.1256
80	1.2241	1.2247	1.1521	1.1521	1.1165	1.1163
90	1.2086	1.2092	1.1420	1.1420	1.1088	1.1087
100	1.1958	1.1963	1.1336	1.1336	1.1025	1.1023

Table 3.6.2

Minimum Values of $\hat{C}p$ and $\tilde{C}p$ required to ensure $Pr(Cp \geq 4/3 | \underset{\sim}{X}) = p$.

	p=0.99		p=0.95		p=0.90	
n	$\hat{C}p$	$\tilde{C}p$	$\hat{C}p$	$\tilde{C}p$	$\hat{C}p$	$\tilde{C}p$
5	4.8923	5.3656	3.1631	3.2100	2.5857	2.5875
6	4.0045	4.2138	2.7857	2.8072	2.3495	2.3482
7	3.4973	3.6124	2.5539	2.5657	2.2000	2.1979
8	3.1692	3.2413	2.3962	2.4034	2.0958	2.0937
9	2.9390	2.9884	2.2814	2.2861	2.0188	2.0168
10	2.7682	2.8042	2.1936	2.1970	1.9592	1.9574
11	2.6362	2.6636	2.1241	2.1266	1.9116	1.9099
12	2.5307	2.5524	2.0675	2.0694	1.8724	1.8709
13	2.4443	2.4620	2.0204	2.0219	1.8396	1.8382
14	2.3722	2.3869	1.9805	1.9817	1.8117	1.8103
15	2.3109	2.3234	1.9463	1.9472	1.7875	1.7863
16	2.2582	2.2689	1.9164	1.9172	1.7664	1.7652
17	2.2122	2.2216	1.8902	1.8908	1.7477	1.7467
18	2.1718	1.1800	1.8669	1.8674	1.7311	1.7301
19	2.1358	2.1432	1.8460	1.8465	1.7162	1.7152
20	2.1037	2.1103	1.8272	1.8276	1.7027	1.7018
25	1.9824	1.9867	1.7553	1.7555	1.6507	1.6500
30	1.9017	1.9048	1.7063	1.7064	1.6150	1.6144
35	1.8433	1.8457	1.6703	1.6705	1.5886	1.5881
40	1.7989	1.8008	1.6426	1.6427	1.5681	1.5677
45	1.7637	1.7653	1.6205	1.6206	1.5517	1.5513
50	1.7349	1.7364	1.6023	1.6023	1.5382	1.5378
60	1.6906	1.6917	1.5740	1.5740	1.5170	1.5167
70	1.6577	1.6586	1.5527	1.5528	1.5011	1.5008
80	1.6321	1.6329	1.5361	1.5361	1.4886	1.4884
90	1.6115	1.6122	1.5226	1.5226	1.4784	1.4782
100	1.5944	1.5951	1.5114	1.5114	1.4700	1.4698

Table 3.6.3

Minimum Values of $\hat{C}p$ and $\tilde{C}p$ required to ensure $Pr(Cp \geq 5/3 | \underset{\sim}{X}) = p$.

	p=0.99		p=0.95		p=0.90	
n	$\hat{C}p$	$\tilde{C}p$	$\hat{C}p$	$\tilde{C}p$	$\hat{C}p$	$\tilde{C}p$
5	6.1153	6.7070	3.9539	4.0124	3.2321	3.2344
6	5.0057	5.2672	3.4821	3.5090	2.9368	2.9352
7	4.3716	4.5155	3.1924	3.2071	2.7498	2.7473
8	3.9615	4.0517	2.9953	3.0043	2.6198	2.6172
9	3.6738	3.7355	2.8517	2.8577	2.5235	2.5211
10	3.4603	3.5053	2.7420	2.7462	2.4491	2.4468
11	3.2952	3.3295	2.6551	2.6582	2.3895	2.3873
12	3.1634	3.1905	2.5844	2.5867	2.3405	2.3386
13	3.0554	3.0775	2.5255	2.5274	2.2995	2.2977
14	2.9653	2.9836	2.4757	2.4771	2.2646	2.2629
15	2.8887	2.9042	2.4328	2.4340	2.2344	2.2328
16	2.8227	2.8361	2.3955	2.3965	2.2080	2.2065
17	2.7653	2.7770	2.3627	2.3635	2.1847	2.1833
18	2.7147	2.7250	2.3336	2.3343	2.1639	2.1626
19	2.6698	2.6790	2.3075	2.3081	2.1452	2.1440
20	2.6296	2.6378	2.2840	2.2845	2.1284	2.1272
25	2.4781	2.4834	2.1941	2.1944	2.0634	2.0625
30	2.3771	2.3809	2.1328	2.1330	2.0187	2.0180
35	2.3041	2.3072	2.0879	2.0881	1.9857	1.9851
40	2.2486	2.2510	2.0533	2.0534	1.9602	1.9596
45	2.2046	2.2066	2.0256	2.0257	1.9396	1.9391
50	2.1687	2.1705	2.0029	2.0029	1.9227	1.9223
60	2.1133	2.1147	1.9675	1.9675	1.8963	1.8959
70	2.0721	2.0733	1.9409	1.9409	1.8764	1.8760
80	2.0401	2.0411	1.9201	1.9201	1.8607	1.8604
90	2.0144	2.0153	1.9033	1.9033	1.8481	1.8478
100	1.9931	1.9939	1.8893	1.8893	1.8375	1.8372

3.7. COMPARISON BETWEEN THE BAYESIAN AND OC CURVE APPROACH

The Bayesian approach to analyzing process capabilities is philosophically different from that of the OC curve approach to analyzing process capabilities. However both techniques promote the use of a probabilistic statement or region in addition to a simple point estimator.

In the case where a non-informative prior is chosen for σ, the resultant posterior distribution is proportional to the likelihood function. Hence a credible interval for Cp is similar to a confidence interval for $\hat{C}p$ as derived from the OC curve approach. Boundary values associated (α=0.05) with sample sizes 10(10)100 have been determined for both procedures and are identical hence inferences regarding the capability of a process will be similar regardless of the approach taken. However the Bayesian approach with or without the Wilson-Hilferty transformation is more flexible and easier to administer than the OC curve approach. It should be pointed out that the Bayesian approach requires that a prior distribution be assumed for σ which is not the case for the OC curve approach. However the OC curve approach requires that maximum and minimum acceptable values of Cp be set, which may also introduce subjectivity into the analysis. As well when wishing to evaluate the probability that the estimated Cp index is greater than some value, the OC curve approach does not lend itself well to this type of analysis.

3.8. EXAMPLES

Example 3.8.1: Consider án example that deals with the parallelism and radial length of a machined hole in a transmission differential case [23]. The process is a new undertaking and capability study was performed. The sampling results were as follows

Stage	Parallelism				Radial Length			
	n	$\bar{x}$	s	$\hat{Cp}$	n	$\bar{x}$	s	$\hat{Cp}$
1	268	8.8	8.3	.80	201	4.7	8.7	.77
2	79	8.3	7.8	.85	96	10.4	21.1	.32
3	300	5.5	4.3	1.55	316	5.0	5.4	1.23

Now without some accompanying probabilistic statement there is little that can be said regarding the capability of the process at the different stages of the production. Using the Bayesian approach in conjunction with the Wilson-Hilferty transformation these results can be reported as follows assuming a non-informative (uniform) prior for unknown σ

Stage	Parallelism					Radial Length				
	n	$\bar{x}$	s	$\hat{Cp}$	$Pr(Cp>1\mid \hat{Cp})$	n	$\bar{x}$	s	$\hat{Cp}$	$Pr(Cp>1\mid \hat{Cp})$
1	268	8.8	8.3	.80	.0000	201	4.7	8.7	.77	.0000
2	79	8.3	7.8	.85	.0139	96	10.4	21.1	.32	.0000
3	300	5.5	4.3	1.55	1.0000	316	5.0	5.4	1.23	1.0000

Example 3.8.2: Consider example 5.3 from page 107 of [25] where a quality control department wanted to assess the capability of a process that involved the manufacturing of a shaft. The product manager set the upper and lower allowable limits as 2.15 and 2.01 respectively (with a target value of 2.08). A random sample (n=50) of shafts was taken and measurements recorded. The results were summarized as $\bar{x} = 2.08$ and $s = 0.022$, hence the estimated Cp index turned out to be

$$\hat{Cp} = \frac{(USL-LSL)}{6\hat{\sigma}} = \frac{2.15-2.01}{6(0.022)} = 1.06 \ .$$

Using the Bayesian approach with the Wilson-Hilferty transformation results in the following

$$\Pr(Cp > 1 \mid \hat{Cp} = 1.06) = 1-\Phi\left\{\frac{1.06^{-2/3} - \frac{439}{441}}{\sqrt{\frac{2}{441}}}\right\}$$

$$= 1-\Phi(-0.50)$$

$$= 0.6915 \quad .$$

Hence although the estimated process capability is greater than 1, the probability that the actual Cp value is greater than 1 is only about 0.7. From Table 3.6.1, the value of $\hat{Cp}$ required to judge the process capable at the 0.05 level of significance will be 1.20.

The OC curve approach requires that both the type-I and type-II error probabilities be fixed. Assuming $\alpha=\beta=0.05$, the ratios for n=50 are Cp(A)/Cp(R)=1.40 and b/Cp(R)=1.20. Letting the minimum requirement of Cp be 1 (i.e., Cp(R)=1), Cp(A)=1.40 and b=1.20. That is the minimum acceptable value of $\hat{Cp}$ that will ensure the actual Cp between 1.00 and 1.40 (with $\alpha=\beta=0.05$) will be 1.20. Hence in this case since $\hat{Cp}=1.06$, the process is not capable at the 0.05 level of significance. This is the same conclusion arrived at using the Bayesian procedure. Note however that this technique provides no method for attaching an actual probability statement to the outcome of the sampling results, that is, no interpretation of how "poor" $\hat{Cp}=1.06$ can be made. Secondly the technique forces one to consider a finite interval for Cp (in this case (1.00, 1.40)) which may not be appropriate in some applications.

3.9. COMMENTS

Some of the statistical properties of the estimated process capability index Cp have been examined. As was assumed, the traditional estimator of Cp is non-robust to

departures from normality. Some general correction factors have been included that may aid practitioners in assessing process capabilities when the process measurements do not arise from a normal distribution.

The Bayesian approach to analyzing process capability has also been presented as an alternative to techniques currently available. Under the assumption of a non-informative prior, critical regions have been created for the sample results, that allow a probabilistic statement to accompany estimates determined from the sample results. In addition to giving equivalent results to the OC curve, the Bayesian approach permits an easy method for determining how extreme the results of one's study are.

Chapter 4

A New Measure of Process Capability

4.1. Introduction

The process capability index Cp has been criticized for its inability to reflect departures from the target value while assessing process capability. Due to this inherent inability, several indices have been proposed that take the target value into account when assessing the capability of a process. The indices considered to date perform quite well in assessing the ability of a process to meet the required specification limits and be close to the target value. However the statistical distributions associated with the estimators of these indices can be quite complicated, which makes study of their statistical properties difficult.

A new index, Cpm, is proposed that takes into account departures from the target value as well as process variation when assessing process capability. The new index and its proposed estimator are presented and some of the associated properties investigated. Two analytical procedures that can be used to investigate process sampling results are examined and an example illustrating use of the new index is provided.

A graphical procedure that provides insights into process capability is presented. The procedure uses modified probability plots to examine the assumption of normality, proximity to target value and process variation. The general procedure is discussed and an example given which focuses on some of the highlights of the procedure.

4.2. Some Process Capability Indices and Their Estimators

As several authors have pointed out ([23], [29], [30]), the Cp index does not take into consideration proximity of the measurements to the target value. For example in Figure 4.2.1, samples arising from any of the five populations ($N(\mu_i, \sigma^2)$ for populations i=1, 2, 3, 4 and 5) would yield similar estimates for the Cp index. Because the

actual process spreads of the five populations are smaller than the allowable process spread, the corresponding estimates of Cp should all indicate that the processes are capable. However processes from populations 2, 3, 4 and 5 all deviate from the target value.

It can be argued that processes with characteristics similar to populations 2 and 3 are still within the specification limits and hence should be judged capable even though they are not centered at the target value. However additional work or costs may be incurred due to adjustments made necessary by these departures from the target value. Processes with population characteristics similar to 4 and 5 will be incapable of meeting the specifications required as they both have non-conforming output. Hence it is important to find an index that will reflect departures from the target value as well as changes in the process variation.

Processes with small variability, but poor proximity to the target value T, have sparked the derivation of several indices which are similar in nature to the Cp index. These indices attempt to take into account process variation as well as departures from the target value. Some of these indices include

$$CPU = \frac{USL - \mu}{3\sigma} \tag{4.2.1}$$

$$CPL = \frac{\mu - LSL}{3\sigma} \tag{4.2.2}$$

$$Cpk = \text{minimum}(CPL, CPU) \tag{4.2.3}$$

and

$$Cpk = (1-k)Cp \tag{4.2.4}$$

where $k = \frac{2\,|T - \mu|}{USL - LSL}$, with μ representing the process mean and $LSL < \mu < USL$.

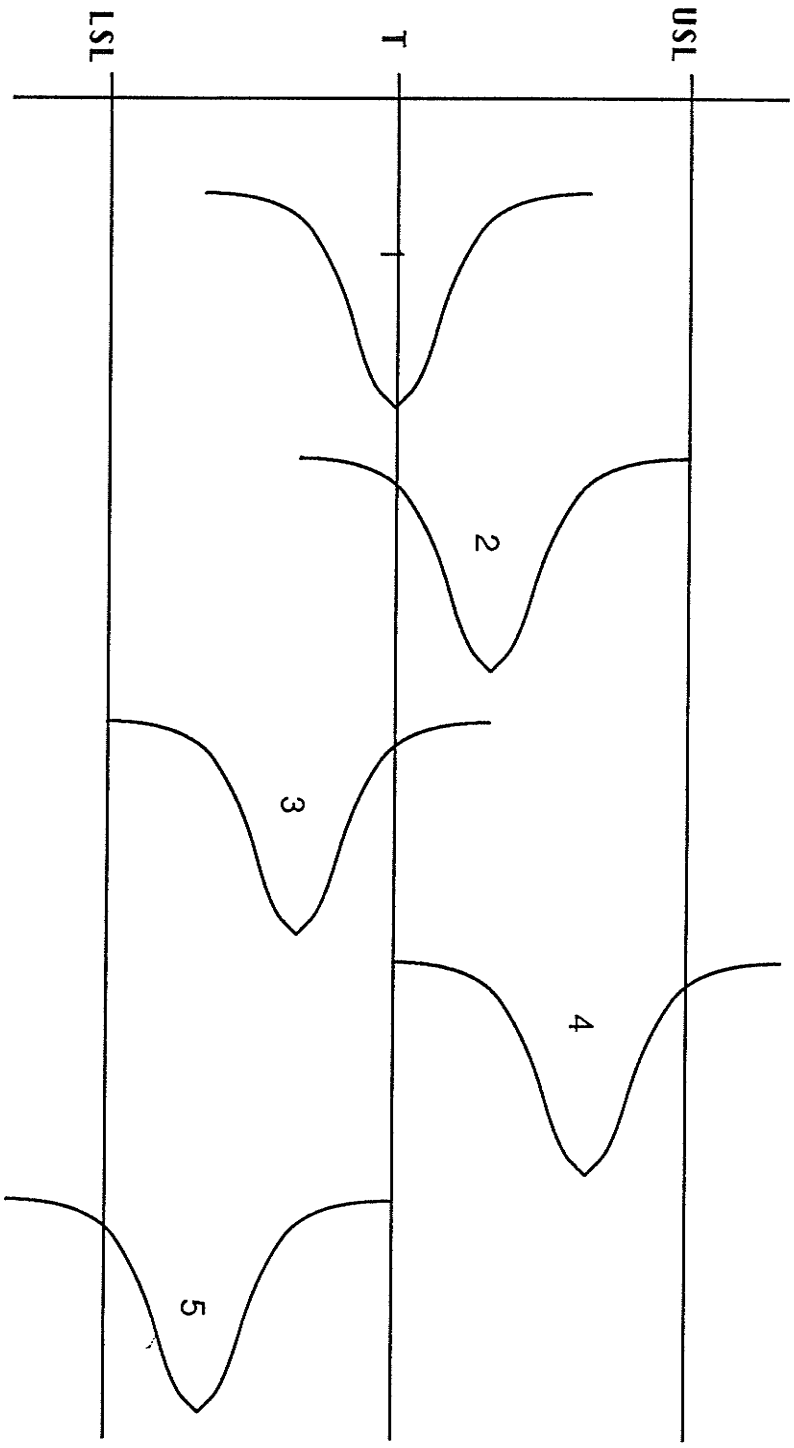


Figure 4.2.1. Five Normal Populations with identical values of Cp.

The two definitions of Cpk are presented interchangeably. They are algebraically equivalent for $0 \le k \le 1$ and where the target value is the midpoint of the specification limits.

Each of the indices involve the unknown parameters μ and σ^2 which generally must be estimated. The usual estimators are

$$\hat{C}PU = \frac{USL - \bar{x}}{3s}$$

$$\hat{C}PL = \frac{\bar{x} - LSL}{3s}$$

$$\hat{C}pk = \text{minimum}(\hat{C}PU, \hat{C}PL)$$

or

$$\hat{C}pk = (1 - \hat{k})\hat{C}p$$

where

$$\hat{k} = \frac{2\,|T - \bar{x}|}{USL - LSL} \quad .$$

These are the maximum likelihood estimators and provide reasonable point estimates for their respective indices. However the statistical distributions associated with these estimators are quite complicated, making inferences from the sampling results difficult. Assuming that the process measurements follow a normal distribution, $\hat{C}PU$ and $\hat{C}PL$ will have a probability density function proportional to that of the non-central t distribution. The pdf of $\hat{C}pk$ is a function of two dependent non-central t distributions and is difficult to derive.

4.3. THE Cpm INDEX, ITS PROPERTIES AND ESTIMATOR

The proposed Cpm index is defined as

$$Cpm = \frac{USL - LSL}{6\sigma'}$$

where

$$\sigma' = \sqrt{E(X - T)^2}$$

and T is the target value of the process. Similar to σ^2, σ'^2 is a population parameter that is usually unknown and will have to be estimated. The general form of the estimator will be

$$\hat{C}pm = \frac{USL - LSL}{6\hat{\sigma}'},$$

where

$$\hat{\sigma}' = \sqrt{\sum_{i=1}^{n} \frac{(x_i - T)^2}{n-1}}$$

Since $\sigma'^2 = E(X-\mu)^2 + (\mu-T)^2 = \sigma^2 + (\mu-T)^2$

$$Cpm = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}} \qquad (4.3.1)$$

It is easy to verify that Cpm will possess the necessary properties required for assessing process capability. If the process variance increases (decreases) the denominator of (4.3.1) increases (decreases) and Cpm will decrease (increase). If the process drifts from its target value (i.e., if μ moves away from T), the denominator of (4.3.1) will again increase causing Cpm to decline. In the case where the process variance changes and the process

mean drifts from T concurrently, the Cpm index possesses the ability to reflect these changes as well.

As an illustration (Figure 4.3.1) of how Cpm reacts to departures from the target value, consider an example from [23] with a target value of 14 rather than 16. As the process mean moves away from the target value, the value of Cpm decreases and although the comparison of absolute magnitudes will be subjective, it becomes obvious that the Cpm reacts to changes in the process in much the same manner as Cpk, while Cp remains constant regardless of the proximity to the target value. There is a one-to-one relationship between Cpk and Cpm for fixed values of Cp.

Theorem 4.3.1: For a fixed value of Cp and $T = \frac{USL+LSL}{2}$, Cpk and Cpm have a one-to-one relationship.

Proof: By definition Cpk=Cp(1-k) and $Cp=Cpm\sqrt{1+\frac{(\mu-T)^2}{\sigma^2}}$,

therefore $Cpk = (1-k)Cpm\sqrt{1+\frac{(\mu-T)^2}{\sigma^2}}$. Hence the relationship is one-to-one.

The indices Cpm and Cp will be identical when the process mean μ, and the target value T coincide (see Theorem 4.3.2). This implies that $\hat{C}pm$ and $\hat{C}p$ are estimates for the same measure when $T=\mu$. In this case, the magnitude of the bias associated with $\hat{C}pm$ is always less than that for $\hat{C}p$ given a fixed sample size (see Theorem 4.3.3) with $\hat{C}pm$ also having a smaller mean square error (MSE). Hence, in those cases where the target value is assumed to coincide with the process mean, $\hat{C}pm$ has better statistical properties than that of $\hat{C}p$.

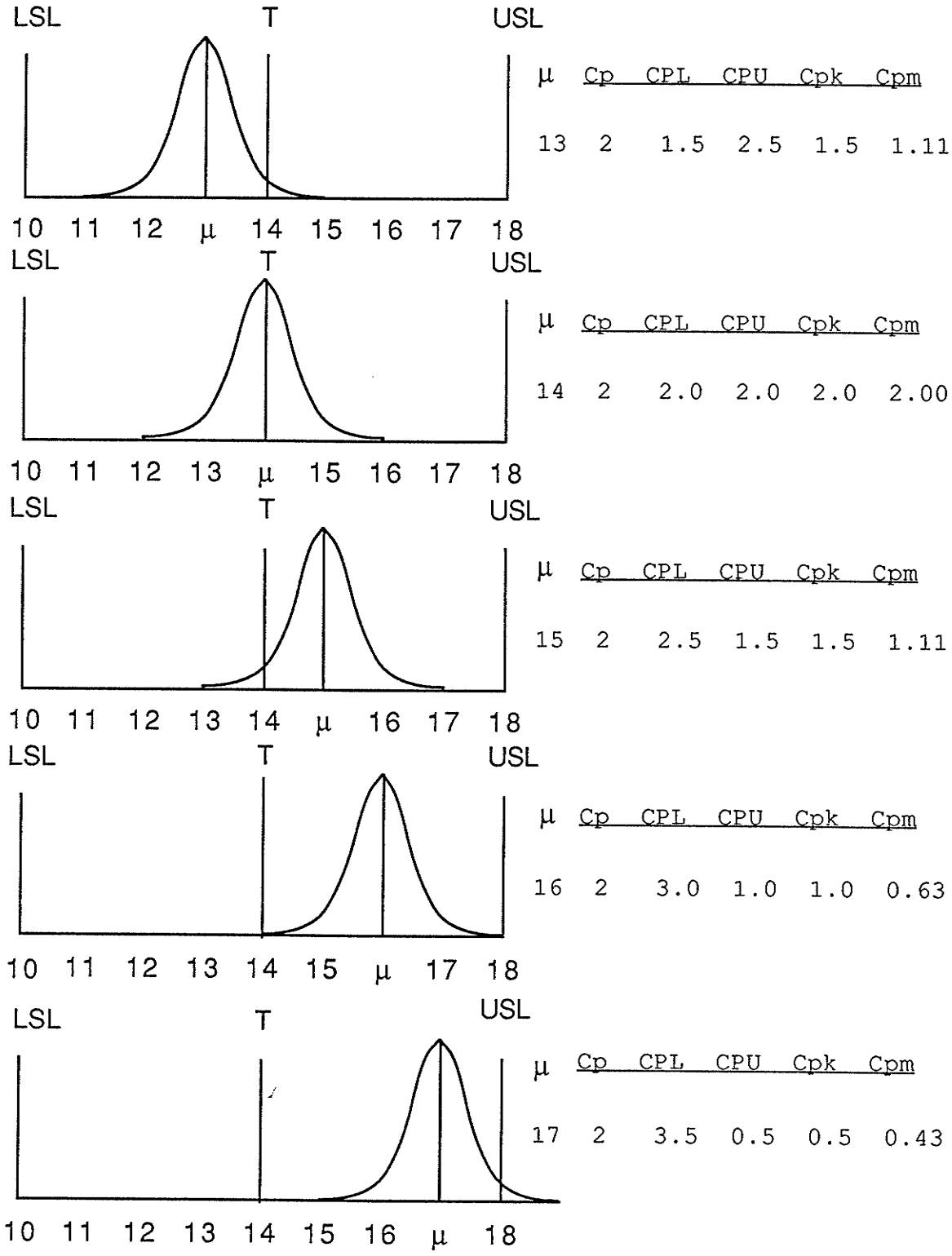


Figure 4.3.1 Comparisons of Cp, CPL, CPU, Cpk and Cpm.

Theorem 4.3.2: For $\mu=T$, Cpm = Cp.

Proof : When $\mu=T$, $\sigma'^2 = E(X-T)^2 = E(X-\mu)^2 = \sigma^2$.

Therefore Cp = Cpm.

Theorem 4.3.3: When $\mu=T$, $\hat{C}pm$ has smaller bias than $\hat{C}p$.

Proof: From Theorem 4.5.2, $E(\hat{C}pm)=\sqrt{\frac{n-1}{2}}\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})}Cp$,

while from Theorem 3.2.2 $E(\hat{C}p)=\sqrt{\frac{n-1}{2}}\frac{\Gamma(\frac{n}{2}-1)}{\Gamma(\frac{n-1}{2})}Cp$.

$$\therefore E(\hat{C}p)-E(\hat{C}pm)=\sqrt{\frac{n-1}{2}}Cp\left\{\frac{\Gamma(\frac{n}{2}-1)}{\Gamma(\frac{n-1}{2})}-\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})}\right\}$$

$$=\left\{\frac{\Gamma(\frac{n}{2})\Gamma(\frac{n}{2}-1)-\Gamma^2(\frac{n-1}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{n-1}{2})}\right\}\sqrt{\frac{n-1}{2}}Cp$$

$$\geq\left\{\frac{\Gamma(\frac{n}{2})\Gamma(\frac{n}{2}-1)-\Gamma^2(\frac{n-1}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{n}{2})}\right\}\sqrt{\frac{n-1}{2}}Cp\geq 0$$

because $MSE(\hat{C}pm)=\left\{\frac{\Gamma(\frac{n}{2})\Gamma(\frac{n}{2}-1)-\Gamma^2(\frac{n-1}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{n}{2})}\right\}\frac{n-1}{2}Cp^2\geq 0$

$$\Rightarrow\left\{\Gamma(\tfrac{n}{2})\Gamma(\tfrac{n}{2}-1)-\Gamma^2(\tfrac{n-1}{2})\right\}\geq 0$$

Hence $E(\hat{C}p)-Cp \geq E(\hat{C}pm)-Cp = E(\hat{C}pm)-Cpm$ since $T=\mu$ and Cp = Cpm.

Assuming that the process measurements arise from a normal distribution, the probability density function of $\hat{C}pm$ is

$$f(y)=\exp\left[-\frac{\frac{a}{y^2}+\lambda}{2}\right]\left[\frac{1}{2}\right]^{\frac{n}{2}-1}\left[\frac{a}{y^3}\right]\sum_{j=0}^{\infty}\left[\frac{\left[\frac{a}{y^2}\right]^{\frac{n}{2}+j-1}\lambda^j}{\Gamma(\frac{n}{2}+j)\,2^{2j}\,j!}\right], \qquad 0<y<\infty$$

where $a = Cpm^2(1+\frac{\lambda}{n})(n-1)$ and $\lambda = \frac{n(\mu - T)^2}{\sigma^2}$ (see Theorem 4.3.4). From the pdf of $\hat{C}pm$, the expected value (see Theorem 4.3.5) and the MSE (see Theorem 4.3.6) of $\hat{C}pm$ can be shown to be functions of the inverse moments of a non-central chi-square distribution. The probability $Pr(\hat{C}pm > c)$ (see Theorem 4.3.7) is

$$\exp-\left[\frac{\lambda}{2}\right]\sum_{j=0}^{\infty}\left[\frac{\left[\frac{\lambda}{2}\right]^j}{j!\;\Gamma(\frac{n}{2}+j)}\int_0^{\frac{a}{2c^2}} w^{\frac{n}{2}+j-1}\exp(-w)\,dw\right].$$

Theorem 4.3.4: If the process measurements are $N(\mu,\sigma^2)$, then the pdf of $\hat{C}pm$ is

$$f(y)=\exp\left[-\frac{\frac{a}{y^2}+\lambda}{2}\right]\left[\frac{1}{2}\right]^{\frac{n}{2}-1}\left[\frac{a}{y^3}\right]\sum_{j=0}^{\infty}\left[\frac{\left[\frac{a}{y^2}\right]^{\frac{n}{2}+j-1}\lambda^j}{\Gamma(\frac{n}{2}+j)\,2^{2j}\,j!}\right], \qquad 0<y<\infty.$$

Proof: When the process measurements are from a $N(\mu,\sigma^2)$, $x = (n-1)\frac{\hat{\sigma}'^2}{\sigma^2}$ follows a non-central chi-square distribution with n df and non-centrality parameter $\lambda = \frac{n(\mu-T)^2}{\sigma^2}$. The pdf of x is

$$f(x) = \exp\left[-\frac{x+\lambda}{2}\right]\left[\sum_{j=0}^{\infty} \frac{x^{\frac{n}{2}+j-1}\,\lambda^j}{\Gamma(\frac{n}{2}+j)\,2^{2j}\,j!}\right]\left[\frac{1}{2}\right]^{\frac{n}{2}}$$

The result follows from noting that $x = (n-1)\frac{\hat{\sigma}'^2}{\sigma^2} = \frac{a}{\hat{C}pm^2}$ and $\hat{C}pm>0$.

Theorem 4.3.5: $E(\hat{C}pm) = \sqrt{a}\,E(x^{-1/2})$.

Proof: Let $x = \frac{Cpm^2(n-1)(1+\frac{\lambda}{n})}{\hat{C}pm^2} = \frac{a}{\hat{C}pm^2}$.

$x \sim \chi^2_{n,\lambda}$ and $E(\hat{C}pm) = \sqrt{a}\int_0^{\infty} x^{-1/2} f(x)dx$

$= \sqrt{a}\,E(x^{-1/2})$.

Theorem 4.3.6: $MSE(\hat{C}pm) = a[E(x^{-1}) - E^2(x^{-1/2})]$

Proof: Let x be as defined in Proof 4.3.5.

$\hat{C}pm^2 = ax^{-1} \Rightarrow E(\hat{C}pm^2) = aE(x^{-1})$

Hence $MSE(\hat{C}pm) = E(\hat{C}pm^2)-[E(\hat{C}pm)]^2$

$=E(x^{-1})a-[E(x^{-1/2})\sqrt{a}\,]^2$

$=a[E(x^{-1})-[E(x^{-1/2})]^2]$.

Theorem 4.3.7: $$Pr(\hat{C}pm>w) = \exp-\left[\frac{\lambda}{2}\right]\sum_{j=0}^{\infty}\left[\frac{\left[\frac{\lambda}{2}\right]^j}{j!\ \Gamma(\frac{n}{2}+j)}\int_0^{\frac{a}{2w^2}} x^{\frac{n}{2}+j-1}\exp(-x)\,dx\right]$$

Proof: $$Pr(\hat{C}pm>w)=Pr(\hat{\sigma}'<\frac{USL-LSL}{6w})=Pr\left(\chi^2_{n,\lambda}<\frac{Cpm^2(n-1)(1+\frac{\lambda}{n})}{w^2}\right)$$

From [31] page 132,

$$Pr(\hat{C}pm>w)= \exp\left[-\frac{\lambda}{2}\right]\sum_{j=0}^{\infty}\left[\frac{\left[\frac{\lambda}{2}\right]^j}{j!\,2^{\frac{n}{2}+j}\,\Gamma(\frac{n}{2}+j)}\int_0^{\frac{a}{w^2}} y^{\frac{n}{2}+j-1}\exp-\left[\frac{y}{2}\right]dy\right]$$

Letting $x=\frac{y}{2}$ results in

$$Pr(\hat{C}pm>w) = \exp-\left[\frac{\lambda}{2}\right]\sum_{j=0}^{\infty}\left[\frac{\left[\frac{\lambda}{2}\right]^j}{j!\ \Gamma(\frac{n}{2}+j)}\int_0^{\frac{a}{2w^2}} x^{\frac{n}{2}+j-1}\exp(-x)\,dx\right].$$

4.4. THE OC CURVE APPROACH FOR ANALYZING Cpm

Suppose that the sample size and the critical value b are to be determined for the case where the probability of a Type I error is to be no more than α if the true value of Cpm is greater than the Acceptable Quality Level Cpm(A), and the probability of a Type II error is to be no more than β if the true value of Cpm is less than the Rejectable Quality Level Cpm(R). Under the assumption that the process measurements are from a $N(\mu,\sigma^2)$

$$\Pr(\hat{C}pm > b) = \Pr\left(\hat{\sigma}'^2 < \frac{(USL-LSL)^2}{6^2 b^2}\right)$$

$$= \Pr\left(\frac{(n-1)\hat{\sigma}'^2}{\sigma^2} < \frac{(n-1)(USL-LSL)^2}{6^2 b^2 \sigma^2}\right)$$

$$= \Pr\left(\chi^2_{n,\lambda} < \frac{(n-1)Cpm^2(1+\lambda/n)}{b^2}\right) .$$

Then the probabilities of making a Type I and Type II error are

$$\alpha = \Pr\left(\chi^2_{n,\lambda} > \frac{(n-1)\,(1+\lambda/n)\,Cpm(A)^2}{b^2} \mid Cpm(A)\right)$$

$$\beta = \Pr\left(\chi^2_{n,\lambda} < \frac{(n-1)(1+\lambda/n)Cpm(R)^2}{b^2} \mid Cpm(R)\right).$$

Therefore

$$\chi^2_{n,\lambda}(1-\alpha) = \frac{(n-1)Cpm(A)^2(1+\lambda/n)}{b^2}$$

$$\chi^2_{n,\lambda}(\beta) = \frac{(n-1)Cpm(R)^2(1+\lambda/n)}{b^2}$$

where $\chi^2_{n,\lambda}(\delta)$ is the $100\delta^{th}$ percentile of a non-central χ^2 distribution with n df and non-centrality parameter λ. The above results can be summarized for various values of α, β, λ, and n using the following ratios

$$\frac{Cpm(A)}{Cpm(R)}=\sqrt{\frac{\chi^2_{n,\lambda}(1-\alpha)}{\chi^2_{n,\lambda}(\beta)}}$$

and

$$\frac{b}{Cpm(R)}=\sqrt{\frac{(1+\lambda/n)(n-1)}{\chi^2_{n,\lambda}(\beta)}}.$$

As derived, the Cpm index possesses the property of reflecting changes in a process that affect the ability of the process to meet some preset specification limits. For the general case where the target value is not assumed to be the same as the process mean, an estimate for Cpm, along with some of its statistical properties have been found that allow stochastic analyses of the sampling results. The calculations associated with obtaining either a measure of variability or a probability statement for the estimate of Cpm can be quite difficult, however the ability to do so exists.

A special case of the Cpm where the target value and the process mean are identical (i.e., μ=T) has some additional properties.

4.5. SOME PROPERTIES OF $\hat{C}pm$ WHEN μ= T

Assuming μ=T, where T is known, the pdf of $\hat{C}pm$ for the case where the process measurements arise from a normal distribution is (see Theorem 4.5.1)

$$g(y)=\left[\frac{1}{2}\right]^{\frac{n}{2}-1}\left[\frac{1}{\Gamma(\frac{n}{2})}\right]\left[(n-1)C_{pm}^2\right]^{\frac{n}{2}}\left[\frac{1}{y}\right]^{n+1}\exp\left[-\frac{(n-1)C_{pm}^2}{2y^2}\right] \qquad 0<y<\infty$$

while the expected value of $\hat{C}pm$ is (see Theorem 4.5.2)

$$E(\hat{C}pm)=\sqrt{\frac{n-1}{2}}\left[\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})}\right]Cpm$$

and the MSE of $\hat{C}pm$ is (see Theorem 4.5.3)

$$MSE(\hat{C}pm)= \frac{n-1}{2}\left[\frac{\Gamma(\frac{n}{2}-1)\Gamma(\frac{n}{2}) - \Gamma^2(\frac{n-1}{2})}{\Gamma^2(\frac{n}{2})}\right] Cpm^2.$$

$\hat{C}pm$ is a biased estimator of Cpm (see Theorem 4.5.2) however $\hat{C}p$ is also a biased estimator of Cp (see Theorem 3.2.2). See Figure 4.5.1 for a comparison of the biases of both estimators. Asymptotically the bias associated with $\hat{C}pm$ is zero (see Theorem 4.5.3) implying that $\hat{C}pm$ is an asymptotically unbiased estimator of Cpm.

Theorem 4.5.1: If $\mu=T$ then the pdf of $\hat{C}pm$ is

$$g(\hat{C}pm) = \left[\frac{1}{2}\right]^{\frac{n}{2}-1}\left[\frac{1}{\Gamma(\frac{n}{2})}\right]\left[\frac{((n-1)C_{pm}^2)^{\frac{n}{2}}}{\hat{C}_{pm}^{n+1}}\right]\exp\left[\frac{(n-1)C_{pm}^2}{2\hat{C}_{pm}^2}\right], \quad 0<\hat{C}pm<\infty$$

Proof: When $\hat{\sigma}'^2=\Sigma\frac{(x_i-T)^2}{n-1}$,

$y=\frac{(n-1)\hat{\sigma}'^2}{\sigma^2}$ has a chi-square distribution with n df. Its pdf is

$$g(y) = (\frac{1}{2})^{n/2}\frac{1}{\Gamma(\frac{n}{2})}\, y^{n/2-1}\exp(-\frac{y}{2}) \qquad 0<y<\infty$$

$$\text{So } \hat{C}pm = \frac{USL-LSL}{6\,\hat{\sigma}'} = \frac{(USL-LSL)\sqrt{n-1}}{6\hat{\sigma}\sqrt{y}} = Cpm\sqrt{\frac{n-1}{y}}$$

$$\Rightarrow y = \frac{Cpm^2(n-1)}{\hat{C}pm^2}$$

$$\Rightarrow dy = \frac{-2Cpm^2(n-1)}{\hat{C}pm^3} d\hat{C}pm$$

$$\Rightarrow g(\hat{C}pm) = \left[\frac{1}{2}\right]^{\frac{n}{2}-1}\left[\frac{1}{\Gamma(\frac{n}{2})}\right]\left[\frac{((n-1)C_{pm}^2)^{\frac{n}{2}}}{\hat{C}_{pm}^{n+1}}\right]\exp\left[\frac{(n-1)C_{pm}^2}{2\hat{C}_{pm}^2}\right], \quad 0<\hat{C}pm<\infty$$

Theorem 4.5.2: If $\mu=T$, $E(\hat{C}pm)=Cpm\sqrt{\frac{n-1}{2}}\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})}$.

Proof: Let $Q=\frac{Cpm^2(n-1)}{2\hat{C}pm^2}$.

Then $E(\hat{C}pm) = \frac{Cpm\sqrt{n-1}\,\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})\sqrt{2}}\int_0^\infty \frac{1}{\Gamma(\frac{n-1}{2})} Q^{(n-1)/2-1}\exp(-Q)\,dQ$

$$=Cpm\sqrt{\frac{n-1}{2}}\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})}.$$

Theorem 4.5.3: If $\mu=T$, $MSE(\hat{C}pm)=Cpm^2\frac{(n-1)}{2}\left\{\frac{\Gamma(\frac{n}{2}-1)}{\Gamma(\frac{n}{2})} - \frac{\Gamma^2(\frac{n-1}{2})}{\Gamma^2(\frac{n}{2})}\right\}$.

Proof: Let Q be as in Proof 4.5.2.

Then $E(\hat{C}pm^2) = Cpm^2\,(n-1)\frac{\Gamma(\frac{n}{2}-1)}{2\,\Gamma(\frac{n}{2})}\int_0^\infty \frac{1}{\Gamma(\frac{n}{2}-1)} Q^{n/2-1}\exp(-Q)\,dQ$

So $MSE(\hat{C}pm)=E(\hat{C}pm^2)-[E(\hat{C}pm)]^2$

$$=Cpm^2\left[\frac{n-1}{2}\right]\left\{\frac{\Gamma(\frac{n}{2}-1)}{\Gamma(\frac{n}{2})}-\frac{\Gamma^2(\frac{n-1}{2})}{\Gamma^2(\frac{n}{2})}\right\}.$$

From Figure 4.5.1 it becomes evident that the magnitude of the relative bias decreases quite rapidly as the sample size increases. For example, the relative bias for $\hat{C}pm$ associated with samples of size 10 is about 3%. While for sample sizes greater than 25, the relative bias will be less than 1%. Note as well that the magnitude of the bias is sample size dependent while also being dependent upon the magnitude of the Cpm index.

$\hat{C}pm$ has been shown to be less biased than $\hat{C}p$ for all values of n. As a comparison, the relative bias for $\hat{C}p$ has been included in Figure 4.5.1. Again for small samples the bias associated with $\hat{C}p$ is quite significant while for larger samples the bias becomes irrelevant. In addition to having a smaller bias, $\hat{C}pm$ has a smaller MSE than $\hat{C}p$, making $\hat{C}pm$ a more efficient estimator than $\hat{C}p$ (see Figure 4.5.2).

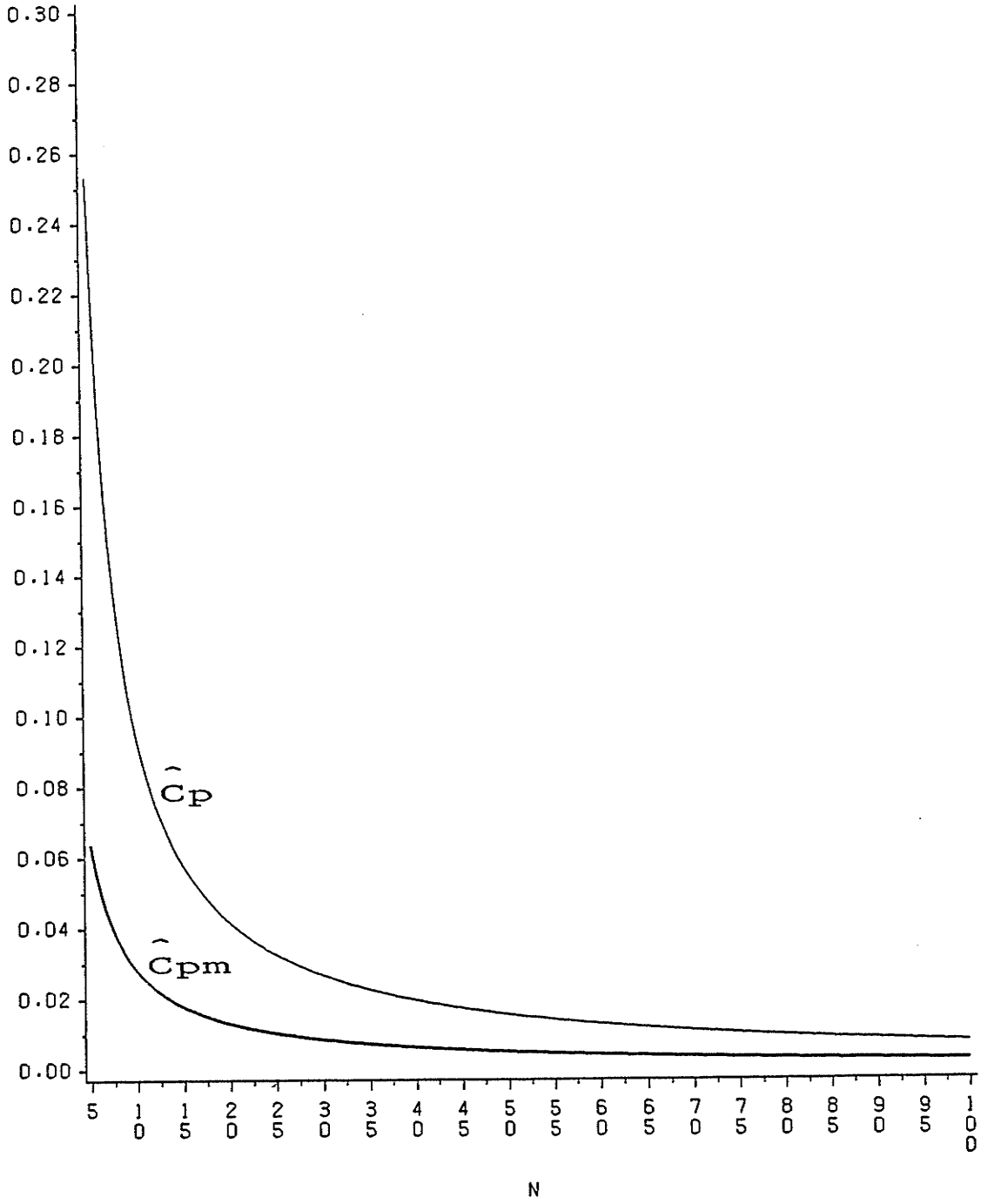


Figure 4.5.1. Relative Bias for $\hat{C}p$ and $\hat{C}pm$ for various values of n.

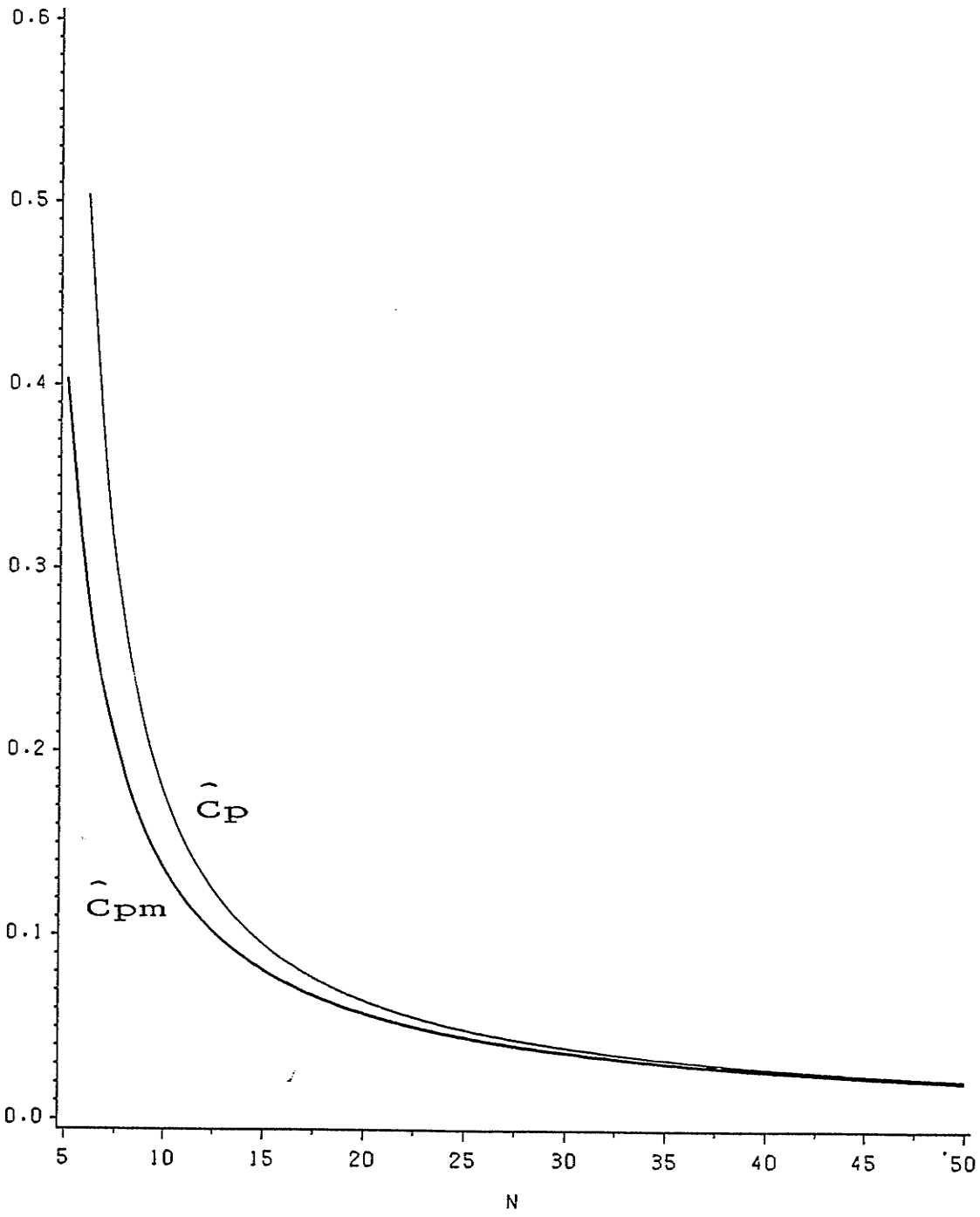


Figure 4.5.2. Relative MSE for $\hat{C}p$ and $\hat{C}pm$ for various values of n.

The mean square error associated with $\hat{C}pm$ is a function of the value of the Cpm index and the sample size (see Theorem 4.5.3) and is asymptotically 0, implying that $\hat{C}pm$ converges in probability to Cpm (see Theorem 4.5.4).

Theorem 4.5.4: $\hat{C}pm$ is MSE consistent.

Proof: From Proof 4.5.3 as n tends to infinity

$$\lim_{n\to\infty} MSE(\hat{C}pm) = \lim_{n\to\infty} \left\{ \frac{\Gamma(\frac{n}{2})\Gamma(\frac{n}{2}-1) - \Gamma^2(\frac{n-1}{2})}{\Gamma^2(\frac{n}{2})} \right\} \frac{(n-1)}{2} Cp^2$$

$$= \frac{Cp^2}{2} \lim_{n\to\infty} \left\{ \frac{\Gamma(\frac{n}{2}-1)\Gamma(\frac{n}{2}) - \Gamma^2(\frac{n-1}{2})}{\Gamma^2(\frac{n}{2})} \right\} (n-1)$$

$$= Cp^2 \left\{ \lim_{n\to\infty} (\frac{n}{2}-\frac{1}{2}) \frac{\Gamma(\frac{n}{2}-1)}{\Gamma(\frac{n}{2})} - \left[\lim_{n\to\infty} \sqrt{(\frac{n}{2}-\frac{1}{2})} \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})} \right]^2 \right\}$$

From [18], both limits are 1. Hence

$\lim_{n\to\infty} MSE(\hat{C}pm)=0$ and this implies that $\hat{C}pm$ converges in probability to Cpm.

4.6. THE OC CURVE APPROACH TO ANALYZING Cpm WHEN μ=T

When μ=T, a result analogous to the OC curve procedure for $\hat{C}p$ can be derived for $\hat{C}pm$. Suppose that the sample size and critical value b are to be determined for those cases where the probability of a Type I error is to be no more than α if the true value of Cpm is greater than the Acceptable Quality Level Cpm(A), and the probability of a Type II

error is to be no more than β if the true value of Cpm is less than the Rejectable Quality Level Cpm(R). Consider the following

$$Pr(\hat{C}pm \geq b) = Pr(\chi^2_n \leq \frac{Cpm^2(n-1)}{b^2})$$

where χ^2_n is a chi-square random variable with n degrees of freedom. The probabilities of making Type I and Type II errors become

$$\alpha = Pr(\chi^2_n > \frac{Cpm(A)^2(n-1)}{b^2} \mid Cpm(A))$$

$$\beta = Pr(\chi^2_n < \frac{Cpm(R)^2(n-1)}{b^2} \mid Cpm(R))$$

and as a result

$$\chi^2_n(\beta) = \frac{Cpm(R)^2(n-1)}{b^2}$$

$$\chi^2_n(1-\alpha) = \frac{Cpm(A)^2(n-1)}{b^2} .$$

Values for the ratios

$$\frac{Cpm(A)}{Cpm(R)} = \sqrt{\frac{\chi^2_n(1-\alpha)}{\chi^2_n(\beta)}}$$

$$\frac{b}{Cpm(R)} = \sqrt{\frac{(n-1)}{\chi^2_n(\beta)}}$$

have been found for various values of α, β, and n (Table 4.6.1).

This approach for evaluating the stochastic properties of the estimated Cpm index is rather cumbersome although it does allow a probabilistic statement to accompany the point

Table 4.6.1

Values of the ratios $\frac{Cpm(A)}{Cpm(R)}$ *and* $\frac{b}{Cpm(R)}$ *for* $\alpha=\beta=0.05$, $\alpha=\beta=0.10$.

	$\alpha=\beta=0.05$		$\alpha=\beta=0.10$	
n	$\frac{Cpm(A)}{Cpm(R)}$	$\frac{b}{Cpm(R)}$	$\frac{Cpm(A)}{Cpm(R)}$	$\frac{b}{Cpm(R)}$
10	2.1555	1.5113	1.8127	1.3601
20	1.7014	1.3233	1.5111	1.2357
30	1.5385	1.2523	1.3979	1.1865
40	1.4503	1.2129	1.3354	1.1587
50	1.3935	1.1872	1.2946	1.1402
60	1.3532	1.1688	1.2655	1.1259
70	1.3228	1.1548	1.2433	1.1167
80	1.2988	1.1437	1.2258	1.1086
90	1.2794	1.1347	1.2115	1.1020
100	1.2632	1.1271	1.1995	1.0964
110	1.2494	1.1207	1.1893	1.0916
120	1.2375	1.1151	1.1805	1.0875
130	1.2272	1.1102	1.1728	1.0839
140	1.2180	1.1059	1.1660	1.0806
150	1.2098	1.1020	1.1599	1.0778
160	1.2025	1.0985	1.1544	1.0752
170	1.1959	1.0954	1.1495	1.0728
180	1.1898	1.0925	1.1450	1.0707
190	1.1843	1.0899	1.1408	1.0687
200	1.1792	1.0875	1.1370	1.0669

estimate. The cumbersomeness arises from the requirement that the Type I and Type II error probabilities must be fixed and co-ordination with the sample size done prior to sampling in order to establish a critical value that can be used to judge the capability of a process. The number of tables required to document all possible results for the OC curve approach will be large.

4.7. A BAYESIAN APPROACH FOR ANALYZING Cpm WHEN μ=T

An approach similar to that proposed for the Cp index can be used to find exact and approximate credible intervals for the Cpm index. These intervals are in turn used in a procedure that is more general, easier to interpret, and less restrictive than those associated with the OC curve approach.

Assuming the measurements follow a $N(T, \sigma^2)$, the likelihood function for the sample $\underset{\sim}{X}=\{x_1, x_2, \ \dots \ , x_n\}$ is

$$L(\sigma^2) = (2\pi\sigma^2)^{\frac{-n}{2}} \exp\left[- \frac{(n-1)\hat{\sigma}'^2}{2\sigma^2} \right]$$

For the non-informative prior

$$h(\mu, \sigma) = \frac{1}{\sigma}, \qquad 0<\sigma<\infty, \ -\infty<\mu<\infty$$

the posterior distribution of σ is

$$f(\sigma|\underset{\sim}{X}) = \frac{2}{\Gamma(\frac{n}{2})} \left[\frac{(n-1)\hat{\sigma}'^2}{2} \right]^{\frac{n}{2}} \frac{1}{\sigma^{n+1}} \exp\left[- \frac{(n-1)\hat{\sigma}'^2}{2\sigma^2} \right] \qquad 0<\sigma<\infty \qquad (4.7.1)$$

where $\hat{\sigma}' = \sqrt{\sum_{i=1}^{n} \frac{(x_i - T)^2}{n-1}}$. The posterior probability becomes

$$p=Pr(Cpm>c \mid \underset{\sim}{X})=Pr(\frac{USL - LSL}{6\sigma}>c \mid \underset{\sim}{X})$$

$$=Pr(\sigma<\frac{USL - LSL}{6c} \mid \underset{\sim}{X})$$

$$=\int_0^a f(\sigma|\underset{\sim}{X})\, d\sigma$$

$$=\int_b^\infty \frac{1}{\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} e^{-y}\, dy$$

where $a = \frac{USL - LSL}{6w}$, $y = \frac{(n-1)\hat{\sigma}'^2}{2\sigma^2}$ and $b = \frac{(n-1)c^2}{2\hat{C}pm^2}$. Given an observed value of $\hat{C}pm$, the probability that Cpm is greater than c can be found. The minimum values of $\hat{C}pm$ required to ensure p=0.90, 0.95, 0.99 for c=1, 4/3, 5/3 and n=3(1)100 have been tabulated and included in Tables 4.7.1, 4.7.2 and 4.7.3. In order to judge the capability of a process using this procedure, one need only calculate $\hat{C}pm$ and compare it with the minimum value required as determined from the tables.

The Wilson-Hilferty transformation [28] provides a very good approximation to the posterior probability (equation (4.7.1)). It is easy to evaluate, good for large and small values of n, and uses the standard normal distribution instead of the incomplete gamma function. The approximation is

Table 4.7.1

Minimum value of $\hat{C}_{pm}$ for p=Pr(Cpm>1 | $\hat{C}_{pm}$).

	p				p		
n	0.90	0.95	0.99	n	0.90	0.95	0.99
3	1.8500	2.3842	4.1733	52	1.1372	1.1831	1.2776
4	1.6795	2.0545	3.1776	53	1.1358	1.1811	1.2744
5	1.5761	1.8687	2.6863	54	1.1344	1.1792	1.2713
6	1.5061	1.7485	2.3944	55	1.1331	1.1773	1.2683
7	1.4553	1.6638	2.2006	56	1.1318	1.1755	1.2654
8	1.4163	1.6005	2.0619	57	1.1305	1.1738	1.2626
9	1.3854	1.5511	1.9575	58	1.1293	1.1721	1.2598
10	1.3601	1.5113	1.8757	59	1.1281	1.1704	1.2572
11	1.3390	1.4785	1.8097	60	1.1269	1.1688	1.2546
12	1.3210	1.4508	1.7552	61	1.1258	1.1673	1.2521
13	1.3054	1.4271	1.7094	62	1.1247	1.1657	1.2496
14	1.2919	1.4066	1.6702	63	1.1236	1.1642	1.2473
15	1.2799	1.3886	1.6362	64	1.1225	1.1628	1.2449
16	1.2692	1.3726	1.6065	65	1.1215	1.1614	1.2427
17	1.2596	1.3583	1.5802	66	1.1205	1.1600	1.2405
18	1.2509	1.3455	1.5567	67	1.1195	1.1587	1.2384
19	1.2430	1.3339	1.5357	68	1.1186	1.1574	1.2363
20	1.2357	1.3233	1.5166	69	1.1176	1.1561	1.2342
21	1.2291	1.3136	1.4993	70	1.1167	1.1548	1.2323
22	1.2229	1.3046	1.4835	71	1.1158	1.1536	1.2303
23	1.2173	1.2964	1.4689	72	1.1150	1.1524	1.2284
24	1.2120	1.2887	1.4555	73	1.1141	1.1512	1.2266
25	1.2070	1.2806	1.4431	74	1.1133	1.1501	1.2248
26	1.2024	1.2750	1.4316	75	1.1125	1.1490	1.2230
27	1.1981	1.2688	1.4209	76	1.1117	1.1479	1.2213
28	1.1940	1.2629	1.4108	77	1.1109	1.1468	1.2196
29	1.1902	1.2575	1.4014	78	1.1101	1.1458	1.2179
30	1.1865	1.2523	1.3926	79	1.1094	1.1447	1.2163
31	1.1831	1.2474	1.3843	80	1.1086	1.1437	1.2147
32	1.1798	1.2428	1.3765	81	1.1079	1.1428	1.2132
33	1.1768	1.2384	1.3690	82	1.1072	1.1418	1.2116
34	1.1738	1.2342	1.3620	83	1.1065	1.1408	1.2102
35	1.1710	1.2302	1.3553	84	1.1058	1.1399	1.2087
36	1.1683	1.2265	1.3490	85	1.1051	1.1390	1.2073
37	1.1657	1.2228	1.3430	86	1.1045	1.1381	1.2059
38	1.1633	1.2194	1.3372	87	1.1038	1.1372	1.2045
39	1.1609	1.2161	1.3317	88	1.1032	1.1364	1.2031
40	1.1587	1.2129	1.3265	89	1.1026	1.1355	1.2018
41	1.1565	1.2099	1.3215	90	1.1020	1.1347	1.2005
42	1.1544	1.2070	1.3167	91	1.1014	1.1339	1.1992
43	1.1524	1.2042	1.3120	92	1.1008	1.1331	1.1980
44	1.1505	1.2015	1.3076	93	1.1002	1.1323	1.1967
45	1.1486	1.1989	1.3034	94	1.0996	1.1315	1.1955
46	1.1468	1.1964	1.2993	95	1.0991	1.1307	1.1943
47	1.1451	1.1940	1.2953	96	1.0985	1.1300	1.1932
48	1.1434	1.1917	1.2915	97	1.0980	1.1293	1.1920
49	1.1418	1.1894	1.2879	98	1.0974	1.1285	1.1909
50	1.1402	1.1872	1.2843	99	1.0969	1.1278	1.1898
51	1.1387	1.1851	1.2809	100	1.0964	1.1271	1.1887

Table 4.7.2
Minimum value of $\hat{C}pm$ for $p=Pr(Cpm>\frac{4}{3}|\hat{C}pm)$.

	p				p		
n	0.90	0.95	0.99	n	0.90	0.95	0.99
3	2.4667	3.1789	5.5645	52	1.5163	1.5774	1.7035
4	2.2393	2.7394	4.2368	53	1.5144	1.5748	1.6992
5	2.1014	2.4916	3.5818	54	1.5126	1.5723	1.6951
6	2.0082	2.3314	3.1926	55	1.5108	1.5698	1.6911
7	1.9404	2.2185	2.9341	56	1.5091	1.5674	1.6872
8	1.8884	2.1340	2.7492	57	1.5074	1.5650	1.6834
9	1.8472	2.0681	2.6099	58	1.5057	1.5628	1.6798
10	1.8135	2.0151	2.5009	59	1.5041	1.5606	1.6762
11	1.7853	1.9713	2.4129	60	1.5026	1.5584	1.6728
12	1.7613	1.9344	2.3403	61	1.5010	1.5563	1.6694
13	1.7406	1.9028	2.2791	62	1.4996	1.5543	1.6662
14	1.7225	1.8755	2.2269	63	1.4981	1.5523	1.6630
15	1.7065	1.8514	2.1816	64	1.4967	1.5504	1.6599
16	1.6922	1.8301	2.1420	65	1.4954	1.5485	1.6569
17	1.6794	1.8111	2.1069	66	1.4940	1.5467	1.6540
18	1.6678	1.7940	2.0756	67	1.4927	1.5449	1.6511
19	1.6573	1.7785	2.0476	68	1.4914	1.5431	1.6484
20	1.6476	1.7644	2.0222	69	1.4902	1.5414	1.6456
21	1.6388	1.7514	1.9991	70	1.4890	1.5398	1.6430
22	1.6306	1.7395	1.9780	71	1.4878	1.5381	1.6404
23	1.6230	1.7285	1.9586	72	1.4866	1.5365	1.6379
24	1.6159	1.7183	1.9407	73	1.4855	1.5350	1.6354
25	1.6094	1.7088	1.9242	74	1.4844	1.5335	1.6330
26	1.6032	1.7000	1.9088	75	1.4833	1.5320	1.6307
27	1.5974	1.6917	1.8945	76	1.4822	1.5305	1.6284
28	1.5920	1.6839	1.8811	77	1.4812	1.5291	1.6261
29	1.5869	1.6766	1.8686	78	1.4802	1.5277	1.6239
30	1.5820	1.6697	1.8568	79	1.4791	1.5263	1.6217
31	1.5774	1.6632	1.8457	80	1.4782	1.5250	1.6196
32	1.5731	1.6570	1.8353	81	1.4772	1.5237	1.6176
33	1.5690	1.6512	1.8254	82	1.4763	1.5224	1.6155
34	1.5650	1.6456	1.8160	83	1.4753	1.5211	1.6135
35	1.5613	1.6403	1.8071	84	1.4744	1.5199	1.6116
36	1.5577	1.6353	1.7987	85	1.4735	1.5187	1.6097
37	1.5543	1.6305	1.7906	86	1.4726	1.5175	1.6078
38	1.5510	1.6259	1.7830	87	1.4718	1.5163	1.6060
39	1.5479	1.6215	1.7757	88	1.4709	1.5152	1.6042
40	1.5449	1.6172	1.7687	89	1.4701	1.5140	1.6024
41	1.5420	1.6132	1.7620	90	1.4693	1.5129	1.6007
42	1.5392	1.6093	1.7556	91	1.4685	1.5118	1.5990
43	1.5365	1.6056	1.7494	92	1.4677	1.5108	1.5973
44	1.5340	1.6020	1.7435	93	1.4669	1.5097	1.5957
45	1.5315	1.5985	1.7378	94	1.4662	1.5087	1.5940
46	1.5291	1.5952	1.7324	95	1.4654	1.5077	1.5925
47	1.5278	1.5920	1.7271	96	1.4647	1.5067	1.5909
48	1.5246	1.5889	1.7220	97	1.4640	1.5057	1.5894
49	1.5224	1.5859	1.7171	98	1.4632	1.5047	1.5879
50	1.5203	1.5830	1.7124	99	1.4625	1.5038	1.5864
51	1.5183	1.5802	1.7079	100	1.4619	1.5028	1.5849

Table 4.7.3

Minimum value of $\hat{C}pm$ for $p=Pr(Cpm>\frac{5}{3}|\hat{C}pm)$.

	p				p		
n	0.90	0.95	0.99	n	0.90	0.95	0.99
3	3.0833	3.9736	6.9556	52	1.8954	1.9718	2.1293
4	2.7991	3.4242	5.2960	53	1.8930	1.9685	2.1240
5	2.6268	3.1145	4.4772	54	1.8907	1.9653	2.1188
6	2.5102	2.9142	3.9907	55	1.8885	1.9622	2.1138
7	2.4255	2.7731	3.6676	56	1.8863	1.9592	2.1090
8	2.3606	2.6675	3.4365	57	1.8842	1.9563	2.1043
9	2.3090	2.5852	3.2624	58	1.8821	1.9535	2.0997
10	2.2668	2.5189	3.1261	59	1.8801	1.9507	2.0953
11	2.2316	2.4641	3.0161	60	1.8782	1.9480	2.0910
12	2.2016	2.4180	2.9253	61	1.8763	1.9454	2.0868
13	2.1757	2.3786	2.8489	62	1.8745	1.9429	2.0827
14	2.1531	2.3443	2.7836	63	1.8727	1.9404	2.0788
15	2.1331	2.3143	2.7270	64	1.8709	1.9380	2.0749
16	2.1153	2.2877	2.6775	65	1.8692	1.9356	2.0711
17	2.0993	2.2639	2.6336	66	1.8675	1.9333	2.0675
18	2.0848	2.2425	2.5946	67	1.8659	1.9311	2.0639
19	2.0716	2.2231	2.5594	68	1.8643	1.9289	2.0604
20	2.0595	2.2054	2.5277	69	1.8627	1.9268	2.0571
21	2.0485	2.1893	2.4988	70	1.8612	1.9247	2.0537
22	2.0382	2.1744	2.4724	71	1.8597	1.9227	2.0505
23	2.0287	2.1606	2.4482	72	1.8583	1.9207	2.0474
24	2.0199	2.1479	2.4259	73	1.8569	1.9187	2.0443
25	2.0117	2.1360	2.4052	74	1.8555	1.9168	2.0413
26	2.0040	2.1250	2.3860	75	1.8541	1.9150	2.0383
27	1.9968	2.1146	2.3681	76	1.8528	1.9131	2.0354
28	1.9900	2.1049	2.3514	77	1.8515	1.9114	2.0326
29	1.9836	2.0958	2.3357	78	1.8502	1.9096	2.0299
30	1.9775	2.0871	2.3210	79	1.8489	1.9079	2.0272
31	1.9718	2.0790	2.3072	80	1.8477	1.9062	2.0245
32	1.9664	2.0713	2.2941	81	1.8465	1.9046	2.0219
33	1.9612	2.0640	2.2817	82	1.8453	1.9030	2.0194
34	1.9563	2.0570	2.2700	83	1.8442	1.9014	2.0169
35	1.9516	2.0504	2.2589	84	1.8430	1.8998	2.0145
36	1.9471	2.0441	2.2483	85	1.8419	1.8983	2.0121
37	1.9429	2.0381	2.2383	86	1.8408	1.8968	2.0098
38	1.9388	2.0323	2.2287	87	1.8397	1.8954	2.0075
39	1.9349	2.0268	2.2196	88	1.8387	1.8939	2.0052
40	1.9311	2.0215	2.2108	89	1.8376	1.8925	2.0030
41	1.9275	2.0165	2.2025	90	1.8366	1.8911	2.0008
42	1.9240	2.0116	2.1944	91	1.8356	1.8898	1.9987
43	1.9207	2.0070	2.1868	92	1.8346	1.8884	1.9966
44	1.9175	2.0025	2.1794	93	1.8337	1.8871	1.9946
45	1.9144	1.9982	2.1723	94	1.8327	1.8858	1.9925
46	1.9114	1.9940	2.1654	95	1.8318	1.8846	1.9906
47	1.9085	1.9900	2.1589	96	1.8309	1.8833	1.9886
48	1.9057	1.9861	2.1525	97	1.8300	1.8821	1.9867
49	1.9030	1.9823	2.1464	98	1.8291	1.8809	1.9848
50	1.9004	1.9787	2.1405	99	1.8282	1.8797	1.9830
51	1.8979	1.9752	2.1348	100	1.8273	1.8785	1.9811

$$p = Pr(Cpm > c \mid \underset{\sim}{X}) = 1 - \Phi\left\{ \frac{\left[\frac{(n-1)\,c^2}{n\hat{C}pm^2}\right]^{\frac{1}{3}} - \left[1 - \frac{2}{9n}\right]}{\sqrt{\frac{2}{9n}}} \right\}$$

where Φ denotes the cdf of the standard normal distribution. This approximation provides a reasonably easy method for attaining critical values for those sample sizes or values of c not included in the tables.

4.8. A GENERALIZATION OF Cpm

To this point T has been assumed to be the midpoint of the specification limits (i.e., (USL-T)=(T-LSL)). However the Cpm index can be generalized to the case where T is not the midpoint of the specification limits. Rather than considering the allowable process spread to be the difference between the USL and LSL, and the actual process spread to be $6\sigma'$, consider the following definition of Cpm

$$Cpm^* = \frac{\text{minimum}[USL-T,\ T-LSL]}{3\sigma'} \quad .$$

Clearly Cpm^* will continue to take the proximity to the target value into consideration while now also taking into account the non-symmetric specification limits. To illustrate how Cpm^* reacts to departures from the target value and non-central target values, an example (from [23]) has been appended to include Cpm^* (see Figure 4.8.1).

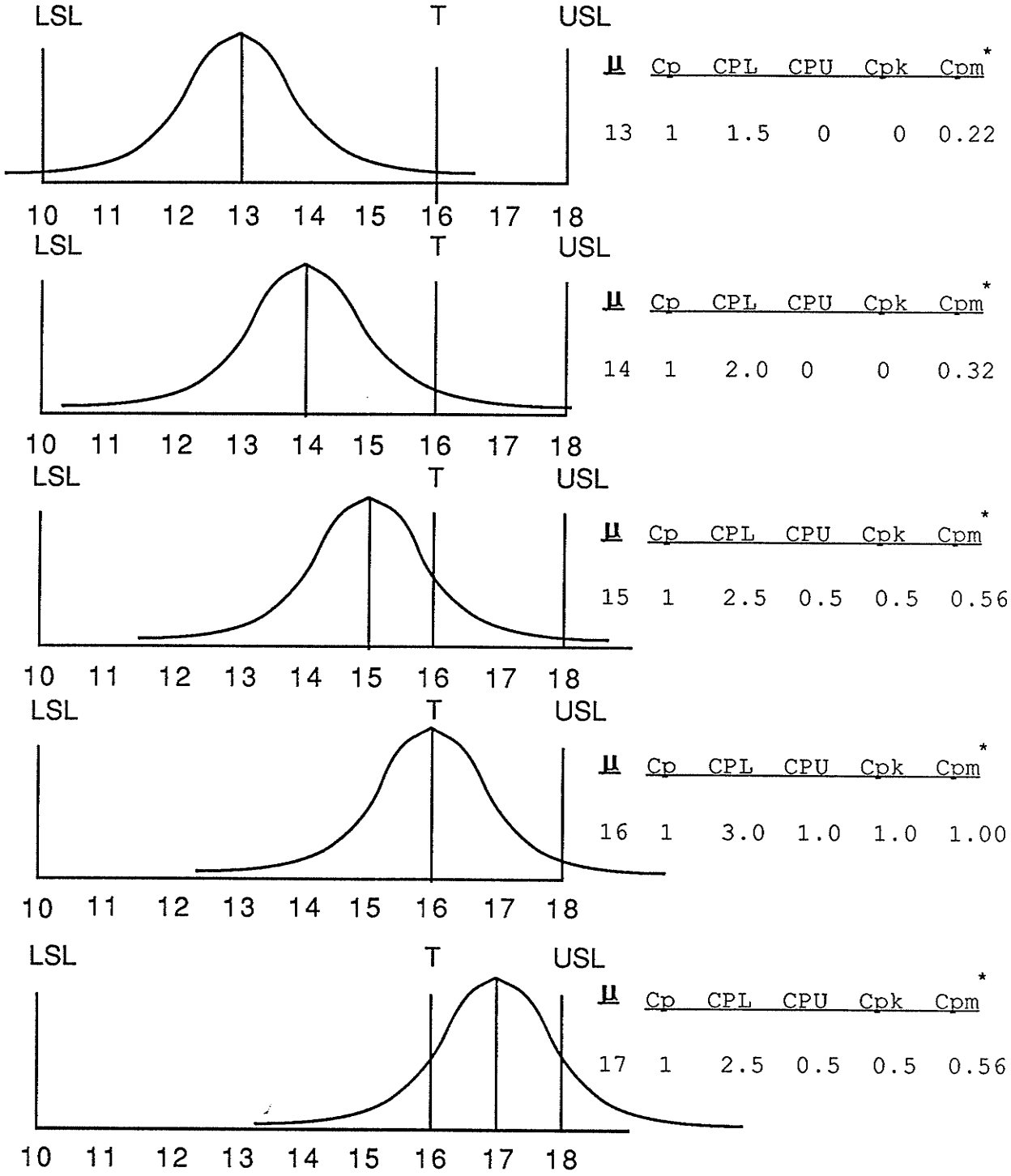


Figure 4.8.1. Comparisons of Cp, CPL, CPU, Cpk and Cpm.*

The proposed estimate of Cpm^* is

$$\hat{C}pm^* = \frac{\text{minimum[USL-T, T-LSL]}}{3\hat{\sigma}'}$$

where $\hat{\sigma}' = \sqrt{\sum_{i=1}^{n} \frac{(x_i - T)^2}{n-1}}$.

The pdf of $\hat{C}pm^*$ will depend upon the values T, USL, LSL and Cpm^*. Assuming the process measurements follow a normal distribution and the parameters T, USL and LSL are fixed for any particular process, the pdf of $\hat{C}pm^*$ is

$$f(x)=\exp\left[-\frac{\frac{(n-1)C\hat{pm}^{2*}(1+\lambda/n)}{x^2}+\lambda}{2}\right]\sum_{j=0}^{\infty}\left[\frac{\left[\frac{(n-1)C\hat{pm}^{2*}(1+\lambda/n)}{x^2}\right]^{\frac{n}{2}+j}\lambda^j}{x\,\Gamma(\frac{n}{2}+j)\,2^{\frac{n}{2}+2j-1}\,j!}\right] \qquad 0<x<\infty$$

where $\lambda=\frac{n(\mu-T)^2}{\sigma^2}$.

Knowing the pdf of $\hat{C}pm^*$ allows functional forms for $E(\hat{C}pm^*)$, $MSE(\hat{C}pm^*)$ and $Pr(\hat{C}pm^* > b \mid Cpm^*)$ to be determined, which in turn permits statistical analysis of the estimate. The OC curve procedure results in the following ratios

$$\frac{Cpm^*(A)}{Cpm^*(R)}=\sqrt{\frac{\chi^2_{n,\lambda}(1-\alpha)}{\chi^2_{n,\lambda}(\beta)}}$$

and

$$\frac{b}{Cpm^*(R)}=\sqrt{\frac{(1+\lambda/n)(n-1)}{\chi^2_{n,\lambda}(\beta)}}$$

where the probability of a Type I error is to be no more than α if the true value of Cpm^* is greater than $Cpm^*(A)$, the probability of a Type II error is to be no more than β if the true value of Cpm^* is less than $Cpm^*(R)$ and $\chi^2_{n,\lambda}$ denotes the non-central chi-square distribution with n degrees of freedom and non-centrality parameter $\lambda=\frac{n(\mu-T)^2}{\sigma^2}$.

Cpm^* permits analysis of one-sided specification limits as well. For example consider the situation where the LSL=10, T=12 and where there is no need to consider USL. For computational purposes let USL=∞, resulting in

$$Cpm^* = \frac{\min[USL-T, T-LSL]}{3\sigma'} = \frac{\min[12-10, \infty-12]}{3\sigma'} = \frac{2}{3\sigma'}.$$

It is then simply a matter of estimating σ' in order to find an estimate for process capability in the one-sided specification limit case. For the case where only the USL is of interest set LSL=$-\infty$.

Cpm* has the ability to assess process capability for a wide variety of situations. It can be used in the one-sided and non-symmetric target cases, while it is easy to show that Cpm* reduces to Cpm in the case where $T=\frac{USL+LSL}{2}$.

4.9. Process Capability Paper

The proposed graphical technique uses modified normal probability plots to first assess the assumption of normality and then to graphically attain estimates of process capability. The resultant probability plot also provides indications of proximity to the target value as well as an indication of the magnitude of the process capability (6σ). The graphical results allow visual comparisons of process capability and proximity to target

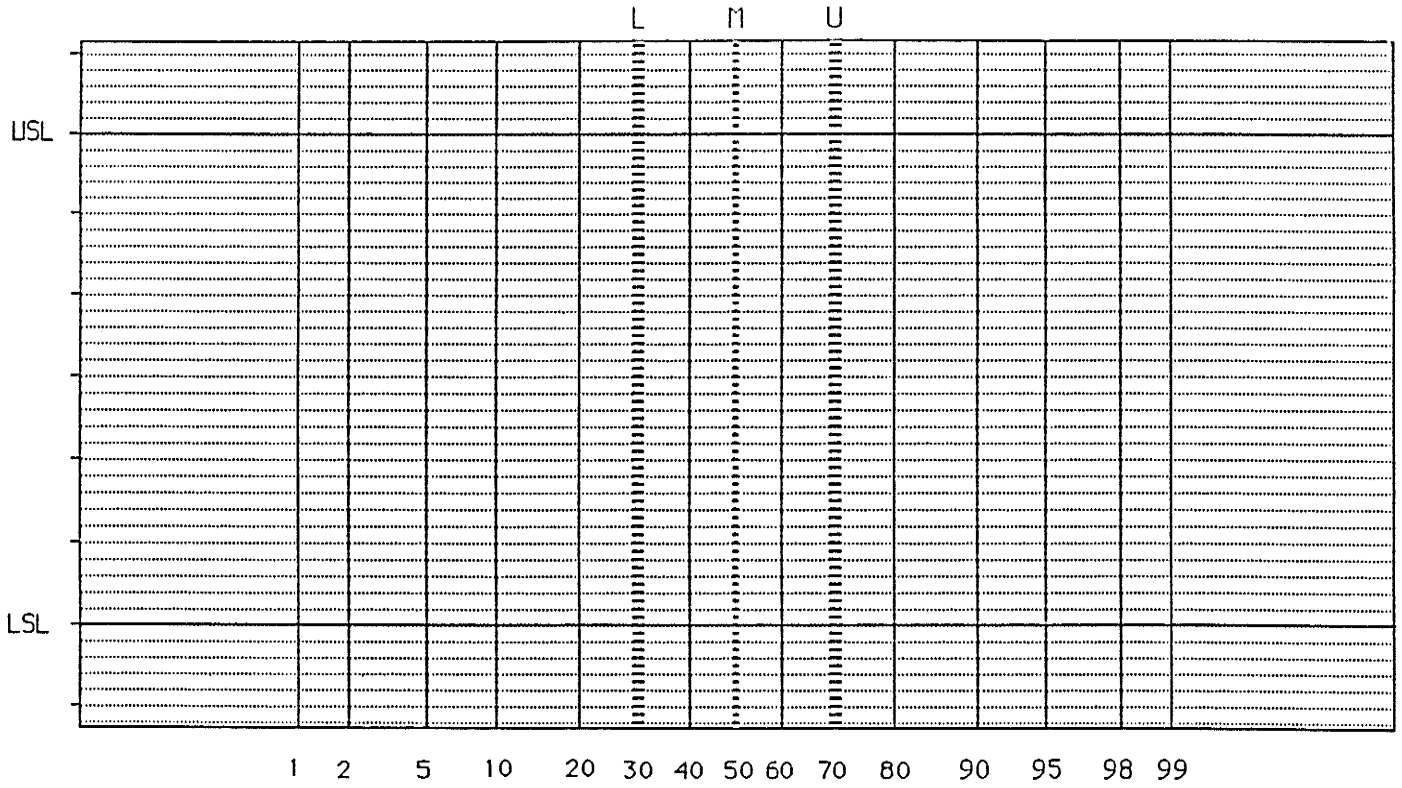
value for various stages or periods of time in order to get an indication of how the capability of the process is changing.

If for example we want to examine the capability of a process after some modification has been made we would simply examine the before and after modified probability plots to (a) assess the underlying distribution of the process measurements, (b) proximity to the target value, (c) process variability and (d) process capability.

The proposed procedure involves creating a series of modified normal probability plots from which inferences regarding the assumption of normality, the proximity to the target value, and the magnitude of the process variation can be made. Customized normal probability paper has been used to create process capability paper (PC paper) (Figure 4.9.1).

The horizontal axis of the proposed PC paper has been scaled identical to normal probability paper while the vertical axis has an arithmetic scale with the upper (USL) specification and lower (LSL) specification limits included. In addition three vertical lines labeled L, M and U have been included that are used to attain numerical estimates for the mean, standard deviation and the process capability indices Cp, Cpk or Cpm for any process under investigation.

The procedure requires a sample to be drawn from the process. The sample results are ordered and plotted versus their associated percentiles, similar to common probability plot procedures. Immediate insight into the nature of the distribution from which the process measurements arise becomes available. If the resultant PC plot (process capability plot) is linear the assumption of normality is not unjustified, which then allows one to draw further inferences regarding the capability of the process.



Sample size = _____ $\hat{C}pk$ = _____

$\hat{\mu}$ = _____ $\hat{C}pm$ = _____

$\hat{\sigma}$ = _____ $\hat{C}p$ = _____

Remarks _____

_____
Date

Figure 4.9.1. Process Capability Paper.

A measure of proximity to the target value is attained by examining the intersection of the PC plot with M (i.e., the 50^{th} percentile) and its location with respect to T. The point of intersection with M, in the case of the normal family, results in an estimate of the population mean (i.e., $\hat{\mu}$), hence as the intercept moves closer to T the estimated process mean moves closer to the target value. A numerical measure of proximity to the target value is obtained by taking the difference between the point of intersection $\hat{\mu}$ and T.

A visual, as well as numerical, measure of the process standard deviation can be determined from the plot. The slope of any probability plot which has been deemed linear represents a measure of the standard deviation associated with the population under investigation. Hence probability plots can be used to indicate those distributions which have larger or smaller standard deviations. Assuming equivalent scaling, those probability plots with "steeper" slopes will have larger standard deviations while those plots with "flatter" slopes will possess smaller standard deviations. Hence all other things being equal those probability plots with "steeper" slopes will have smaller values of Cp, while those with "flatter" slopes will have larger values of Cp. Therefore by keeping (i) the specification limits and (ii) the scaling of the vertical axis of the process capability paper constant, the slope of the PC plots at different stages in the process's history can be used to assess changes in the process standard deviation and hence changes in the value of Cp. A numerical estimate for the standard deviation associated with the process measurements can be determined by taking the difference between two values determined from the intersection of the PC plot with U and L. These lines are located such that the difference in their intercepts results directly in an estimate of the standard deviation. Using the results attained from the plot numerical estimates for Cp, Cpk or Cpm can be determined.

Repeating this procedure at various stages or periods in the history of the process then provides the practitioner with a series of process capability plots that can be used to

examine any changes that have occurred in the ability of the process to meet target values and fall within the specification limits.

4.10. EXAMPLES

Example 4.10.1: Consider those measurements taken on the radial length of machined holes in Example 3.8.1 and where the upper and lower specification limits were set at 20 and -20 units respectively with a target value of T=0. The estimate of Cpm is

$$\hat{C}pm = \frac{20 - (-20)}{6\hat{\sigma}'}$$

where

$$\hat{\sigma}' = \sqrt{\sum_{i=1}^{n} \frac{(x_i - T)^2}{n-1}}$$

$$= \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1} + \frac{n(\bar{x} - T)^2}{n-1}} .$$

The results can now be summarized as follows

Radial Length ($x10^3$ inches)

Stage	n	$\bar{x}$	s	$\hat{C}p$	$Pr(Cp>1\|\hat{C}p)$	$\hat{C}pk$	$\hat{C}pm$	$Pr(Cpm>1\|\hat{C}pm)$
1	201	4.7	8.7	0.77	0.0000	0.59	0.67	0.0000
2	96	10.4	21.1	0.32	0.0139	0.15	0.28	0.0000
3	316	5.0	5.4	1.23	1.0000	0.93	0.91	0.0067

The estimates associated with Cpm for the three stages are somewhat different from those estimates of Cpk and Cp. The $Pr(Cp>1|\hat{C}p)$ and $Pr(Cpm>1|\hat{C}pm)$ have been calculated using the Bayesian approach and the Wilson-Hilferty approximation.

To illustrate the Bayesian procedure, consider the results obtained in Stage 1, where USL=20, LSL=-20, T=0, $\bar{x}$ =4.7, s=8.7 and n=201. From these results the following were found

$$\hat{C}p = \frac{USL - LSL}{6s} = \frac{20 - (-20)}{6(8.7)} = \frac{40}{52.2} = 0.77$$

$$\hat{C}pm = \frac{USL - LSL}{6\sqrt{s^2 + \frac{n(\bar{x}-T)^2}{n-1}}} = \frac{20 - (-20)}{6\sqrt{8.7^2 + \frac{201(4.7-0)^2}{200}}} = 0.67$$

The p-value associated with the Bayesian approach (using the Wilson-Hilferty approximation) is

$$p = Pr(Cpm > c \mid \hat{C}pm) = 1 - \Phi\left\{ \frac{\left[\frac{(n-1)\,c^2}{n\hat{C}pm^2}\right]^{\frac{1}{3}} - \left[1 - \frac{2}{9n}\right]}{\sqrt{\frac{2}{9n}}} \right\}$$

where Φ denotes the cumulative distribution function of the standard normal distribution. Substituting the observed values from Stage 1 and assuming that the process will be judged capable if Cpm>1 (i.e., c=1),

$$p = 1 - \Phi\left\{ \frac{\left[\frac{200}{201(0.67)^2}\right]^{\frac{1}{3}} - \left[1 - \frac{2}{9(201)}\right]}{\sqrt{\frac{2}{9(201)}}} \right\}$$

$$= 1 - \Phi\left\{\frac{1.3039 - 0.9989}{0.0333}\right\}$$

$$= 1 - \Phi\{9.1592\} \ .$$

From any standard normal table $\Phi\{9.1592\}=1.00$, resulting in p=(1-1.00)=0.00, indicating that the probability of Cpm $\geq$ 1 given the sample results is 0.00, i.e., the process is not capable.

The OC curve approach for Cpm results in the same inference but in a slightly different manner. Again using the information USL=20, LSL=-20, T=0, $\bar{x}$=4.7, s=8.7, n=201 results in the estimates $\hat{C}p$=0.77 and $\hat{C}pm$=0.67. Letting α=Pr(Type I error)=β=Pr(Type II error)=0.05,

$$\frac{Cpm(A)}{Cpm(R)} = \sqrt{\frac{\chi^2_{201}(0.95)}{\chi^2_{201}(0.05)}}$$

and

$$\frac{b}{Cpm(R)} = \sqrt{\frac{200}{\chi^2_{201}(0.05)}} \ .$$

Solutions for the above equations involve finding the associated percentiles for the chi-square distribution with 201 df, which are $\chi^2_{201}(0.95)$=235.076 and $\chi^2_{201}(0.05)$=169.20. Substituting these values into the equations, finds

$$\frac{Cpm(A)}{Cpm(R)} = \sqrt{\frac{235.076}{169.20}} = 1.1787$$

and

$$\frac{b}{Cpm(R)} = \sqrt{\frac{200}{169.20}} \ .$$

Setting the rejectable quality level Cpm(R)=1 (i.e., the point where the probability of finding a process capable is no more than 0.05 if the actual value of Cpm<1), the acceptable quality level is Cpm(A)=1.1787 (i.e., the point where the probability of finding a process incapable is no more than 0.05 if the actual Cpm≥1.1787) and a critical value of b=1.0872. Thus since $\hat{C}pm=0.67<1.0872=b$, one can conclude that the process is not capable.

In order to determine the values of Cpm(A) and b in the above example, the 5^{th} and 95^{th} percentiles of the χ^2_{201} distribution were required. To avoid determining the percentiles of the χ^2_{201} distribution, an approximation using the results from Table 4.6.1 for n=200, could have been used

$$\frac{Cpm(A)}{Cpm(R)} = 1.1792 \text{ and } \frac{b}{Cpm(R)} = 1.0875 .$$

In this case, these ratios differ from the exact values only in the third decimal point, resulting in the same inference formulated above.

The nature of the estimates that consider both the process variance and the proximity to the target value (i.e., Cpk and Cpm) is highlighted in Stage 3, where a $\hat{C}p=1.23$ becomes misleading if proximity to the target value is important. Suppose that a process is considered to be capable if Cp>1, the value of $\hat{C}p=1.23$ indicates that with probability one the process is capable. On the other hand, if the proximity to the target value is important and the process is considered capable if Cpm>1, the value of $\hat{C}pm=0.91$ indicates that it is very unlikely (with probability 0.0067) the process is capable. This is essentially the finding of Kane [23] whose inference is based on the subjective

interpretations of the magnitudes of $\hat{C}p$ and $\hat{C}pk$. Note that in order to attach a probabilistic statement to $\hat{C}pk$ its distribution must be known (which is currently not the case).

Example 4.10.2: An example illustrating use of the proposed PC paper has been included. In this example the ability of a hypothetical process for three simulated stages in its history is examined. The data were generated independently for each stage using the normal random number generator resident in SAS, version 5. For stage one, thirty observations were generated from a N(-2, 2) distribution, twenty-five from a N(0, 2) distribution in stage two, and in stage three, 35 observations from a N(0, 0.5) distribution.

Assuming the target value of the process to be 0, and with USL=3.0 and LSL= -3.0 the vertical axis of the PC paper can now be scaled to reflect the magnitude of the USL and LSL with the target value suitably indicated. In order to compare the graphical results for the three stages the PC paper should be scaled identically at each stage.

The mechanics of the graphical procedure are identical for the three stages and involve plotting the order statistics at each stage versus their associated percentiles. In this example the plotting positions [(i - .5)/n]100% for the percentile where i denotes the rank of the i^{th} order statistic have been used. The choice of plotting positions is arbitrary as many possible plotting conventions exist. For a discussion of plotting positions see [32]. At each stage the order statistics (scaled arithmetically along the vertical axis) versus the percentiles (scaled to reflect standard normal probability paper along the horizontal axis) are plotted. This has been done for the three hypothetical stages denoted above (Figure 4.10.1).

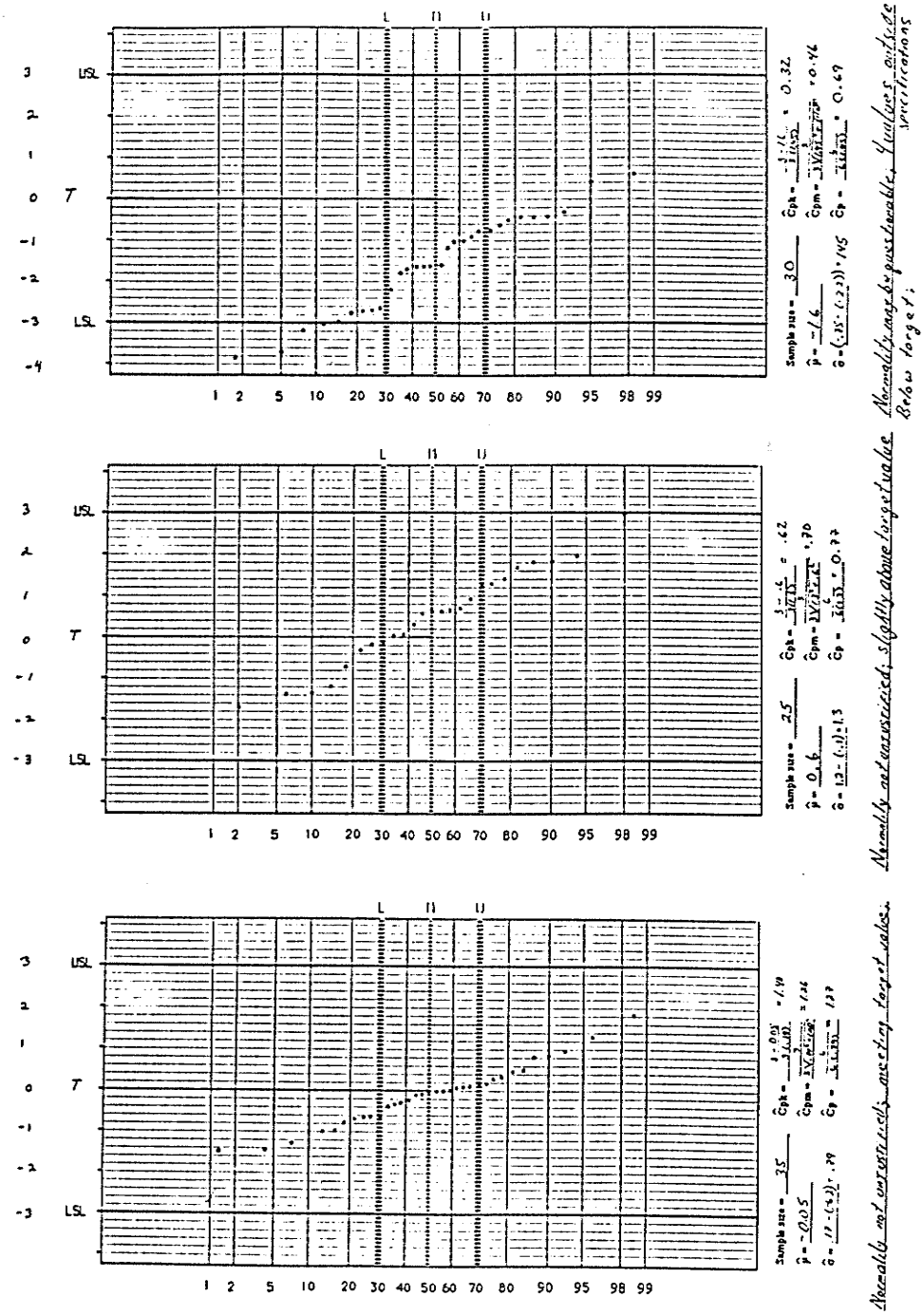


Figure 4.10.1. PC plots for Example 4.10.2.

Once the ordered statistics have been plotted a judgement as to the validity of the normal assumption can be ascertained from the linearity (non-linearity) of the plot. If the plot appears linear then further inferences can be drawn, while in those cases where the plot is non-linear any further inferences regarding the capability of the process are cautioned against. The plots for the three stages in the example all appear linear thereby supporting the assumption of normality and permitting additional inferences.

Prior to making any numerical estimates some general conclusions can be drawn from the plots. First, by noting the distance between the M intercept and T it appears that the mean of the process measurements moves closer to the target value. At stage one the mean of the process measurements appears substantially different from the target value where in stage two the process mean appears to move closer to the target value while in stage three the process mean and the target value appear to coincide.

A similar type of inference can be drawn from the slopes of the plots associated with the three stages. The slope of a PC plot provides insight into the relative magnitude of both σ and Cp. Disregarding location, the steeper the slope the greater the magnitude of σ and hence the smaller the value of Cp. It appears that the slopes at stage one and stage two are similar suggesting that the Cp does not appear to change over this period. However the slope of the line in stage three appears smaller than that of stage one and two indicating that the process variability has been reduced in stage three. Thus stage three appears to have smaller variability as well as attaining its target value suggesting that the process would be deemed most capable at stage three.

Numerical estimates of Cp, Cpk, and Cpm can be obtained from the plots by substituting the estimates of μ and σ, determined from the plot for each of the three stages. These values have been determined and are included in Figure 4.10.1. Note that these values reflect the visual results discussed above.

4.11. COMMENTS

The Cpm index and the Cpk index are two techniques which can be used to evaluate the ability of a process to attain a preset target value and to fall within required specification limits concurrently. The Cpm index can be estimated using $\hat{C}pm$ for those cases where $\mu=T$ as well for those cases where $\mu \neq T$. $\hat{C}pm$ has a distribution similar to that of $\hat{C}p$ when $\mu=T$. This result in conjunction with additional statistical techniques allows a probabilistic statement to be made regarding the likelihood of incorrectly judging the ability of a process. In addition, it has been shown that when the process attains its target value, the Cpm index is identical to the Cp index while $\hat{C}pm$ is less biased and more efficient than $\hat{C}p$.

The estimate for the Cpk index also possesses some of the properties associated with the Cpm index. The Cpk index is identical to the Cpm index when the process mean equals its target value and the target value is the midpoint of the upper and lower specification limits. However, the statistical distribution of $\hat{C}pk$ is difficult to determine. Under the assumption that the process measurements arise from a $N(T, \sigma^2)$, the distributions associated with $\hat{C}PU$ and $\hat{C}PL$ are proportional to non-central t distributions which can be computed. However $\hat{C}pk$ is defined as the minimum($\hat{C}PU$, $\hat{C}PL$) and as a result its distribution is further complicated. Simulations indicate that $\hat{C}pm$ is less biased and more efficient than $\hat{C}pk$.

A generalization of Cpm has been developed that permits assessment of the process capability for those cases where a) T is not the midpoint of the specification limits or b) one-sided specification limits are required. The distribution of the associated estimate (based on the assumption that the measurements of interest possess a normal distribution)

has been determined which in turn permits a stochastic statement to accompany the estimate determined from the sampled units.

A graphical result is examined that permits insights into the process capability. Cpk and Cpm are used as single measures of process capability that consider proximity to T and process variation when assessing process capability. The graphical result allows verification of normality while indicating whether the magnitude of Cpm is due strictly to process variation, proximity to T or some combination of the two.

Chapter 5

A Multivariate Measure of Process Capability

5.1. INTRODUCTION

For those cases where the process must be capable of attaining specification limits for more than one variable, the practice has been to examine each variable independently. A process is then considered capable if all components of interest are found capable. This procedure can be misleading in those situations where the variables being considered do not behave independently.

An index is proposed, $\underset{\sim}{C}pm$, that provides a general measure of process capability for those cases where any number of variables are used to assess capability. The new measure is analogous to the univariate measure Cpm in that $\underset{\sim}{C}pm$ considers both proximity to the target value and process variation while quantifying process capability.

In the multivariate case the specification limits and the actual process spread will be more difficult to define than their univariate counterparts. The specification limits will not simply be points on the number line and the actual process spread will no longer be a straightforward function of the process standard deviation.

Some of the troubles associated with assessing multivariate process capability are considered. These include problems associated with creating specification limits as well as some of the computational difficulties inherent in the procedure. Several properties associated with the proposed measure are examined and two examples illustrating use of the measure are given.

Discussions will focus on the bivariate case as explanations can be enhanced with the use of graphical aids. However all results, unless otherwise stated, hold for the general multivariate case.

5.2. MULTIVARIATE SPECIFICATION LIMITS

One of the major problems in assessing multivariate process capability is in establishing specification limits. In the univariate case the specification limits are boundaries representing the range of acceptable results arising from a process. These boundaries are generally a reflection of some engineering or manufacturing requirement often representing the maximum and minimum acceptable values (see Figure 5.2.1). In the multivariate case the specification boundaries will continue to represent the region of acceptable results, however these boundaries will be more complex than simply two points on a number line.

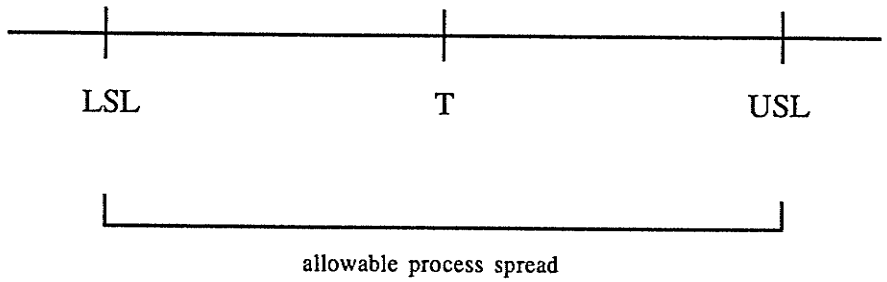


Figure 5.2.1. Typical Univariate Specification Limits.

In the case where two or more variables are used in assessing process capability, the most common practice is to examine the univariate capabilities on an individual basis, with the process being deemed capable if all variables are judged capable. This is equivalent to considering all variables independent, resulting in specification boundaries for the bivariate case similar to those illustrated in Figure 5.2.2.

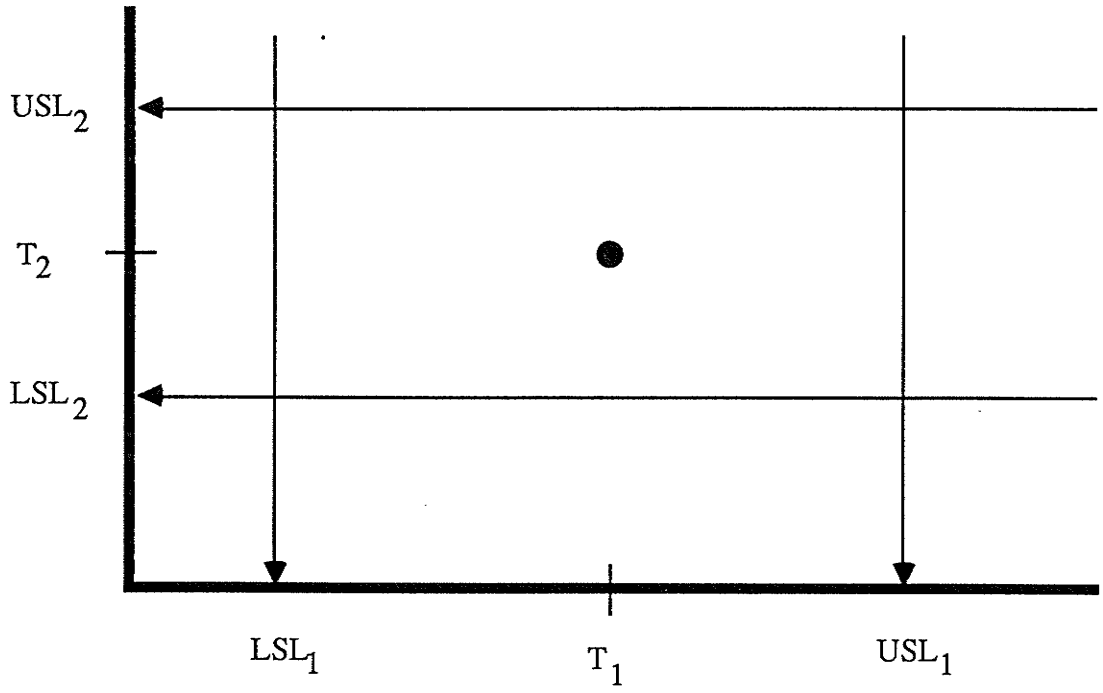


Figure 5.2.2. An Example of Bivariate Specifications for the Independent Variables Case.

In Figure 5.2.2 USL_1, LSL_1 and T_1 denote the univariate upper, lower and target values for variable 1 and USL_2, LSL_2 and T_2 the univariate upper, lower and target values for variable 2. An appropriate measure of multivariate process capability in this situation would be $Cpm_1 x\ Cpm_2$ where Cpm_1 is the value of Cpm for variable 1 and Cpm_2 the value of Cpm for variable 2. This measure of process capability, and the above boundaries, assume that the magnitude of variable 1 is not influenced by the magnitude of variable 2, and vice versa.

Where the specification limits of the variables do not behave independently, the product of individual process capability measurements will not provide an accurate indication of process capability. As an example consider the case where there is no physical relationship between the two variables, but where combinations of their extreme values are not acceptable in the manufacturing process. Specification boundaries reflecting this situation are shown in Figure 5.2.3.

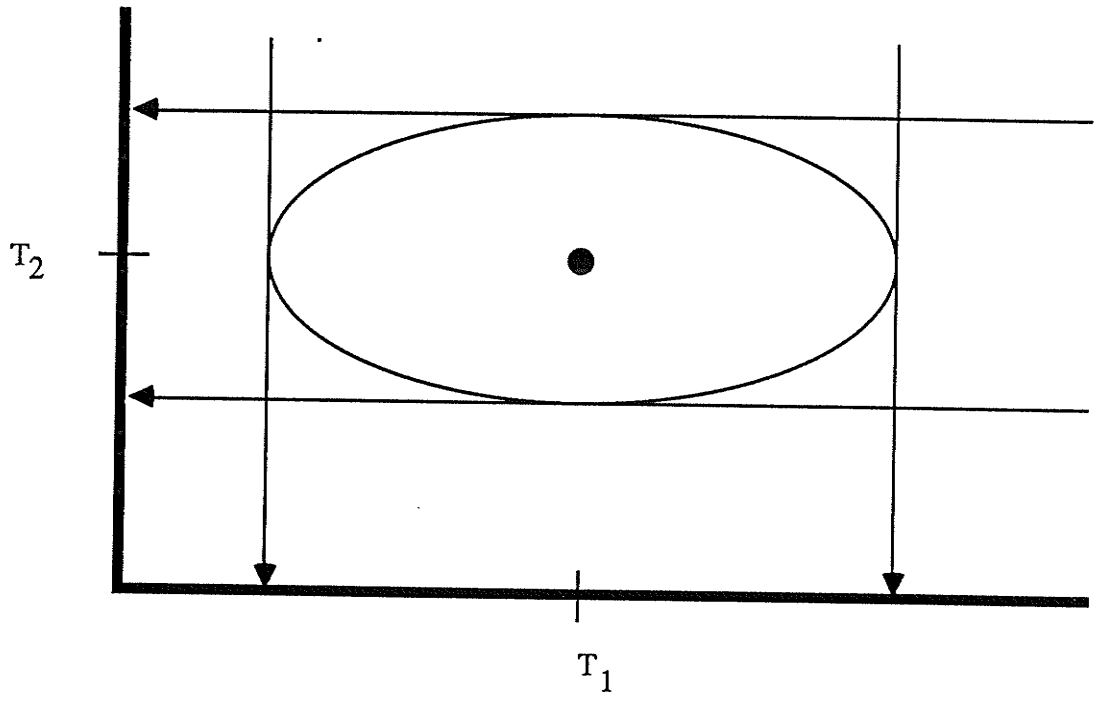


Figure 5.2.3. An Example of Ellipsoidal Specification Limits for Uncorrelated Case.

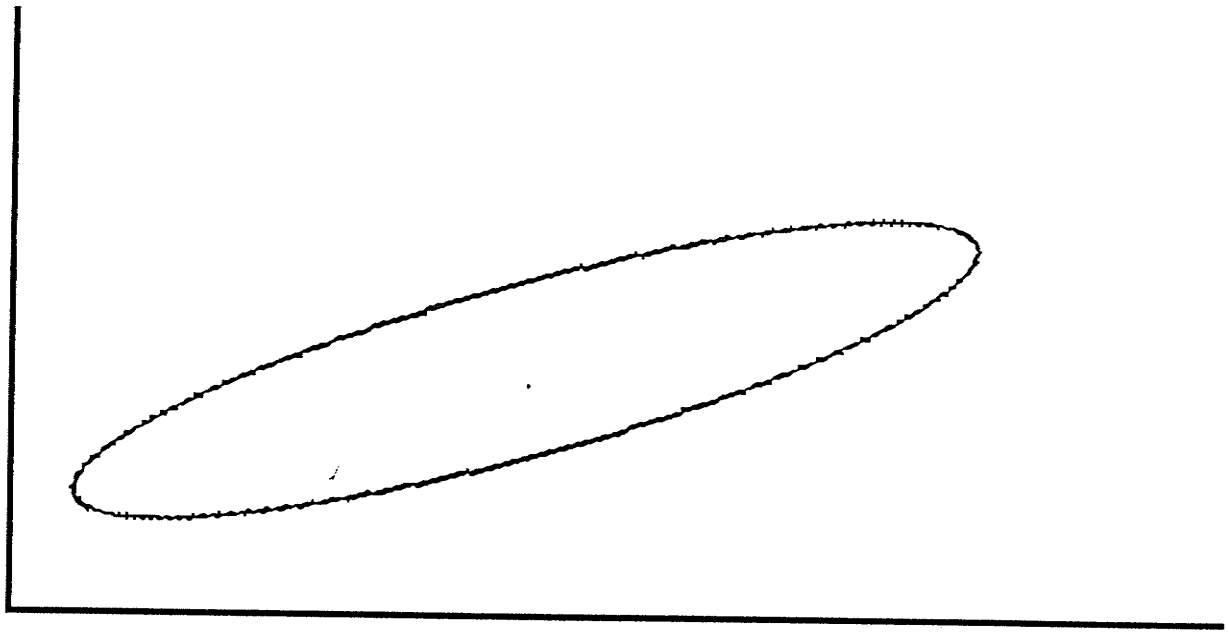


Figure 5.2.4. An Example of Ellipsoidal Specification Limits for Correlated Case.

The specification limits are denoted by the inscribed ellipse with the interior of the ellipse indicating the region of acceptable combinations of the two variables. Although the specification limits do not behave independently, the variables of interest are depicted as being uncorrelated. In such cases, the individual specification limits for either variable depend upon the level of the other variable used in assessing process capability.

The specification limits denoted by the ellipse in Figure 5.2.4 represent the acceptable combinations of variables 1 and 2 when the variables are assumed to have a positive linear relationship. Again the individual specification limits for any one variable will depend upon the value or level of the second variable. The product of the individual measures of process capability will not reflect the multivariate capabilities of the process in cases such as this. Hence the need for a multivariate measure of process capability when two or more related variables are to be used to assess the capability of a process.

5.3. CREATING MULTIVARIATE SPECIFICATION LIMITS

Other than the case where the capability variables are considered uncorrelated with independent boundaries, ellipsoidal specification limits are assumed. This assumption is made necessary by the assumption of multivariate normality (MVN) which will be made when examining some of the statistical properties associated with the proposed measure.

In the univariate case unilateral specification limits arise from time to time. Unilateral specification limits occur when only the upper or lower limit is of interest (e.g., specification regions of the form $(-\infty, USL)$ or (LSL, ∞)). Multivariate analogues of such regions cannot be represented using ellipsoids. In those cases where one or more of the capability variables have unilateral limits, fitting an ellipsoidal region within the required specification region will be equivalent to imposing more stringent engineering requirements than required (Figure 5.3.1).

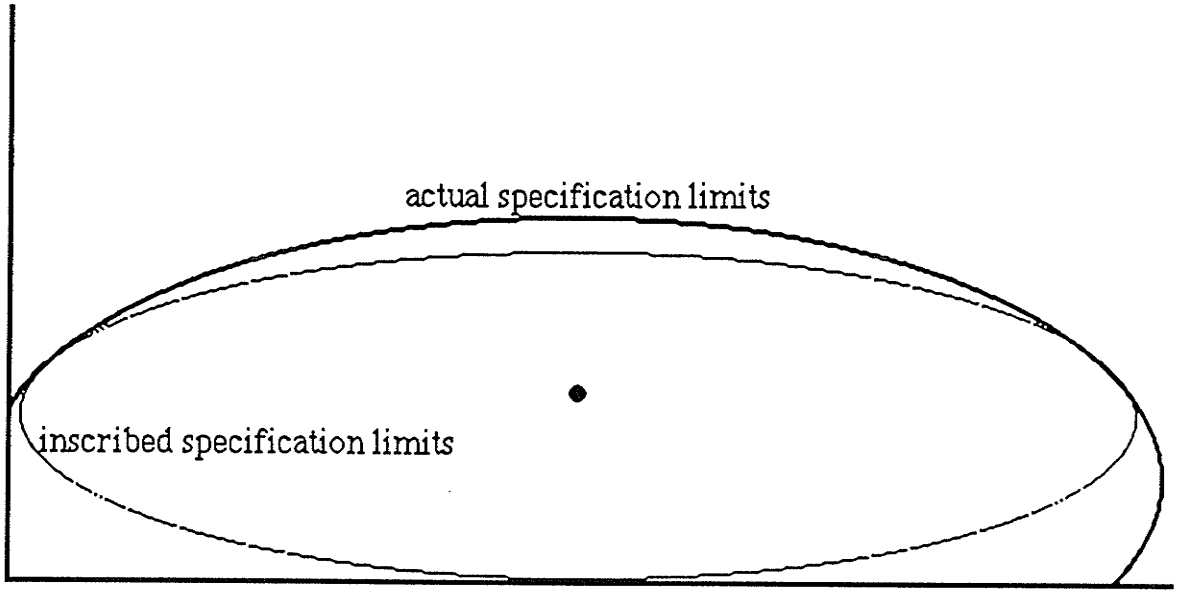


Figure 5.3.1. An Example of an Unilateral Specification.

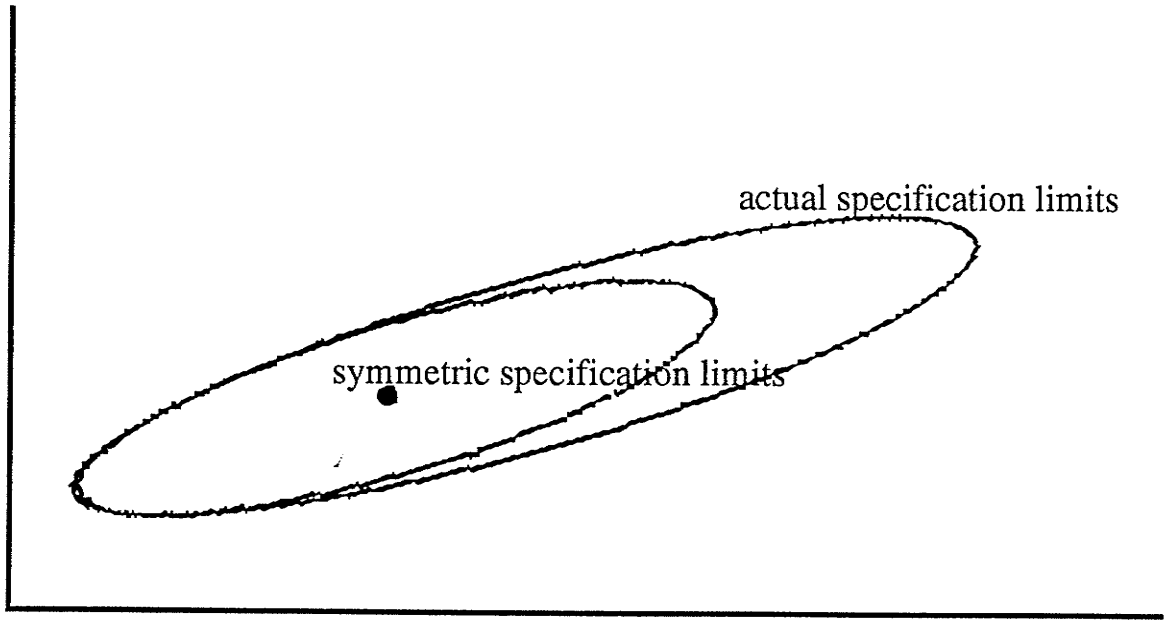


Figure 5.3.2. An Example of Non-symmetric Specifications.

A second univariate case for which no multivariate analogue currently exists, occurs when the target value is not the center of the specification ellipsoid. Because symmetric covariance structures will be assumed for the specification limits, when such cases occur, smaller symmetric ellipsoidal regions will again provide more stringent requirements than required (Figure 5.3.2).

The alternatives presented for these two cases are made in the presence of the assumption of MVN.

Given that the specification region is ellipsoidal all points lying on or within the specification boundary satisfy the inequality

$$(\underset{\sim}{X} - \underset{\sim}{T})' A^{-1} (\underset{\sim}{X} - \underset{\sim}{T}) \leq c^2$$

with those points lying precisely on the boundary represented by the strict equality

$$(\underset{\sim}{X} - \underset{\sim}{T})' A^{-1} (\underset{\sim}{X} - \underset{\sim}{T}) = c^2 \qquad (5.3.1)$$

where $\underset{\sim}{T}$ is the υ-dimensional target vector, A a $\upsilon x \upsilon$ matrix denoting the covariance structure of the capability variables and c a numerical constant. $\underset{\sim}{T}$ and A completely determine the shape and center of the ellipsoid, while c determines the coverage of the ellipsoid. The coverage is the percentage of observations expected to be within the boundaries. If the engineering requirements are specified in terms of $\underset{\sim}{T}$ and A, the specification limits are then fixed for some minimum level of coverage. Analogous to the univariate case, coverage will generally be taken to be 99.73%.

As an example, consider a hypothetical process with $\underset{\sim}{T} = \begin{bmatrix} 75 \\ 100 \end{bmatrix}$ being the vector representing the optimum values of the capability variables and $A = \begin{bmatrix} 144 & 90 \\ 90 & 225 \end{bmatrix}$ the covariance structure representing the univariate variance of both variables (i.e., $\sigma_1^2 = 144$

and σ_2^2=225) and the covariance between the two variables (i.e., σ_{12}=90). Substituting these values into equation (5.3.1) results in the equation

$$\left[\frac{1}{1-\left[\frac{90}{12(15)}\right]^2}\right]\left[\left[\frac{x-75}{12}\right]^2 - 2\frac{90}{12(15)}\left[\frac{x-75}{12}\right]\left[\frac{y-100}{15}\right]+\left[\frac{y-100}{15}\right]^2\right] = c^2.$$

In general $c^2 = \chi_v^2(a)$ will result in an ellipsoid with 100a% coverage. In the bivariate case letting c^2=-2ln(1-a) also results in an ellipse with 100a% coverage [33], which for a=0.9973 results in the following equation

$$\left[\frac{x-75}{12}\right]^2 - 2\frac{90}{12(15)}\left[\frac{x-75}{12}\right]\left[\frac{y-100}{15}\right]+\left[\frac{y-100}{15}\right]^2 = 8.87175$$

and specification boundaries as illustrated in Figure 5.3.3.

When the covariance structure is completely specified the specification boundaries are relatively easy to determine. Regardless of the number of capability variables considered, knowledge of $\underset{\sim}{T}$, A and c completely determine the specification boundaries.

In the bivariate case, the form and equation of the specification limits are of interest as graphical illustrations are often used. However in the general multivariate case no analogous graphical procedures currently exist, hence the actual form of the equation is of little consequence once $\underset{\sim}{T}$, A and c are such that the equation

$$(\underset{\sim}{X} - \underset{\sim}{T})' A^{-1} (\underset{\sim}{X} - \underset{\sim}{T}) = c^2$$

represents a 99.73% coverage region. The numerical values of $\underset{\sim}{T}$ and A are necessary for computation of the proposed measure of process capability.

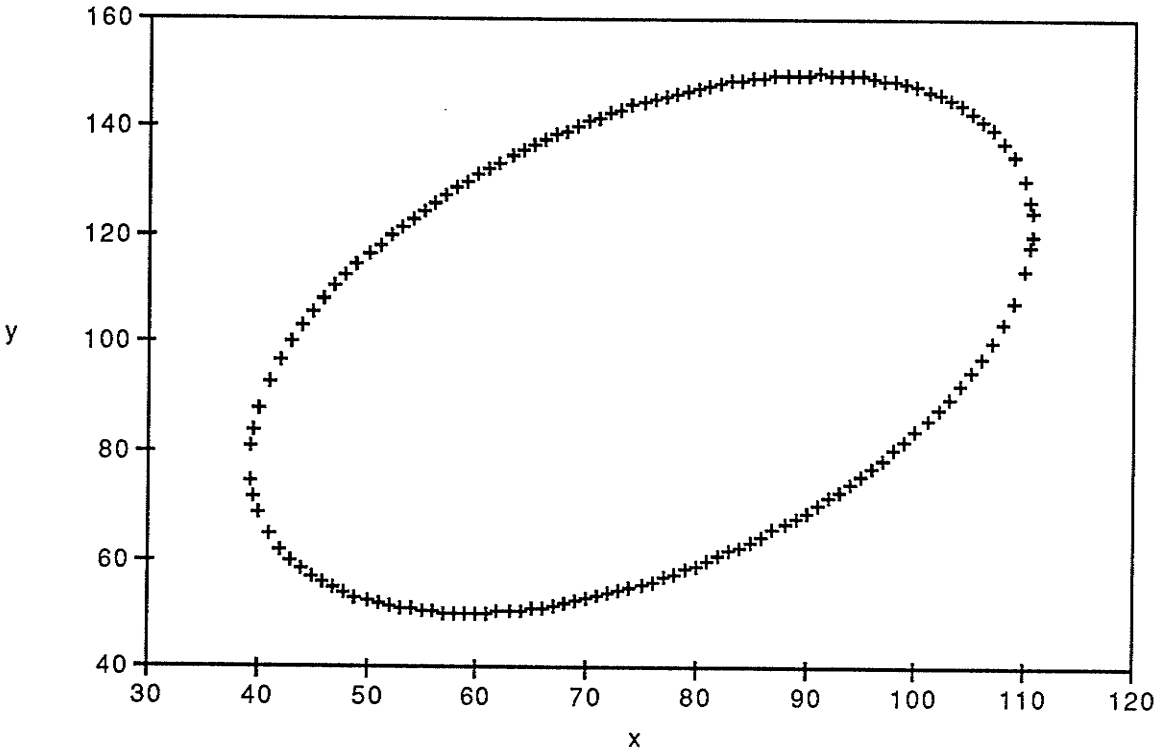


Figure 5.3.3. Example of Ellipsoidal Boundaries.

5.4. A MULTIVARIATE PROCESS CAPABILITY MEASURE

The proposed measure of multivariate process capability is

$$\underset{\sim}{C}pm = \sqrt{\frac{n\upsilon}{\sum_{i=1}^{n} (\underset{\sim}{X}_i - \underset{\sim}{T})' A^{-1} (\underset{\sim}{X}_i - \underset{\sim}{T})}} \qquad (5.4.1)$$

where $\underset{\sim}{X}_i$ denotes the i^{th} vector of dimension υ, υ the number of variables used in assessing capability, n the sample size, $\underset{\sim}{T}$ the target vector, and A a $\upsilon x \upsilon$ matrix representing the covariance structure as determined from the specification limits.

The numerator of (5.4.1) is the product of the sample size and the number of variables used in assessing capability. This value represents the degrees of freedom associated with the denominator. The denominator is the sum of the observed Mahalanobis distances (D^2) measured from the target. For any given process the numerator is a constant while the denominator reflects the "cluster" of the observations around the target.

The Mahalanobis distances are standardized measures of distance that possess the ability to account for any correlations that may exist among the variables. Letting A^{-1} denote the inverse of the covariance matrix of the capability variables, these distances are standardized to reflect the variability inherent in the various dimensions. These individual distances are analogous to the univariate distance measure

$$\left\{\frac{X-\mu}{\sigma}\right\}^2 .$$

For processes with observations clustered around the target (e.g., B of Figure 5.4.1), the denominator of equation (5.4.2) will be smaller in magnitude than that of the same process with observations more scattered and/or with a center of mass not at the target value (e.g., A and C of Figure 5.4.1). As the denominator grows, the value of $\underset{\sim}{C}pm$, for a fixed $n\upsilon$, will diminish. Smaller values of $\underset{\sim}{C}pm$ suggest that the process is unable to meet the specification limits, casting doubt on the capability of the process, while larger values of $\underset{\sim}{C}pm$ suggest that the process is indeed capable of meeting its specifications. Thus $\underset{\sim}{C}pm$ behaves in much the same fashion as the univariate measures of process capability.

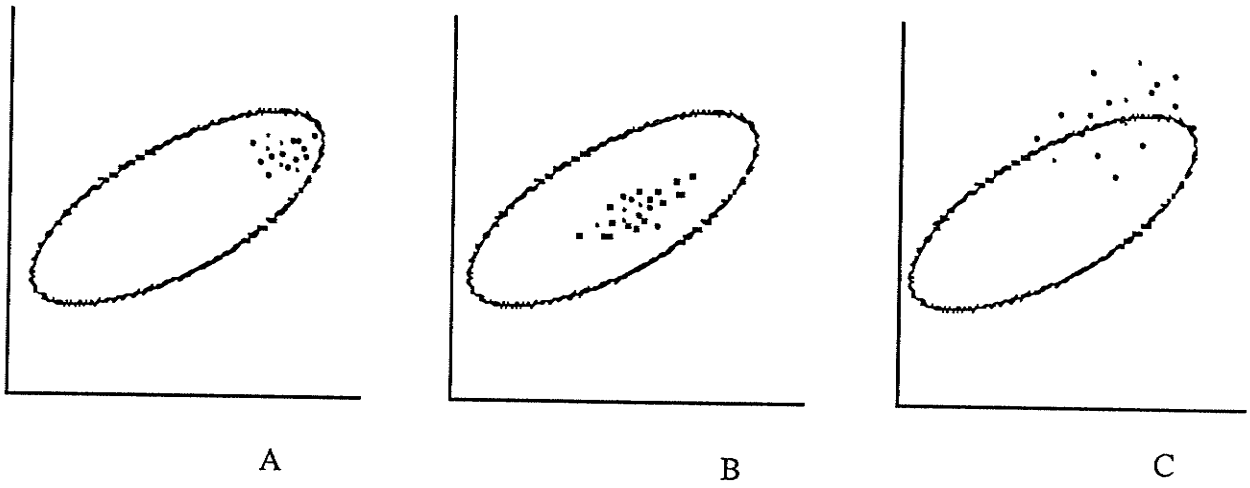


Figure 5.4.1. Three Examples of Bivariate Capability Studies.

The following statistical properties are derived under the assumption that the process measurements arise from a $MVN_{\upsilon}[\underset{\sim}{T}, A]$ population.

Theorem 5.4.1: If $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are iid $MVN_{\upsilon}[\underset{\sim}{T}, A]$ the pdf of $\underset{\sim}{C}pm$ is

$$f(x) = \frac{1}{\Gamma(\frac{n\upsilon}{2})}\left[\frac{1}{2}\right]^{\frac{n\upsilon}{2}-1} \frac{(n\upsilon)^{\frac{n\upsilon}{2}}}{x^{n\upsilon+1}} e^{-\frac{n\upsilon}{2x^2}} \qquad 0<x<\infty$$

Proof: Let $w = \sum_{i=1}^{n} (\underset{\sim}{Y}_i - \underset{\sim}{T})' A^{-1} (\underset{\sim}{Y}_i - \underset{\sim}{T})$. If $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are iid $MVN_\upsilon[\underset{\sim}{T}, A]$

then $w \sim \chi^2_{n\upsilon}$. Therefore since $\underset{\sim}{C}pm = \sqrt{\frac{n\upsilon}{w}}$,

it follows that $f(x) = \frac{1}{\Gamma(\frac{n\upsilon}{2})}\left[\frac{1}{2}\right]^{\frac{n\upsilon}{2}-1} \frac{(n\upsilon)^{\frac{n\upsilon}{2}}}{x^{n\upsilon+1}} e^{-\frac{n\upsilon}{2x^2}}$.

Theorem 5.4.2: If $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are iid $MVN_\upsilon[\underset{\sim}{T}, A]$ then

$$E(\underset{\sim}{C}pm) = \sqrt{\frac{n\upsilon}{2}} \frac{\Gamma(\frac{n\upsilon}{2} - \frac{1}{2})}{\Gamma(\frac{n\upsilon}{2})}.$$

Proof: If $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are iid $MVN_\upsilon[\underset{\sim}{T}, A]$ then

$E(x) = \int_0^\infty x\, f(x)dx$ where $f(x)$ is as defined in Theorem 5.4.1, hence

$$E(x) = \sqrt{\frac{n\upsilon}{2}} \frac{\Gamma(\frac{n\upsilon}{2} - \frac{1}{2})}{\Gamma(\frac{n\upsilon}{2})}.$$

Theorem 5.4.3: If $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are iid $MVN_\upsilon[\underset{\sim}{T}, A]$ then

$$\mathrm{Var}(\underset{\sim}{C}pm) = \frac{n\upsilon}{2}\left[\frac{\Gamma(\frac{n\upsilon}{2})\Gamma(\frac{n\upsilon}{2} - 1) - \Gamma^2(\frac{n\upsilon}{2} - \frac{1}{2})}{\Gamma^2(\frac{n\upsilon}{2})}\right].$$

Proof: If $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are iid $MVN_\upsilon[\underset{\sim}{T}, A]$, then

$$\mathrm{Var}(x) = \int_0^\infty x^2 f(x)dx \; - \left[\int_0^\infty x\, f(x)dx\right]^2, \text{ where } f(x) \text{ is as defined in Theorem 5.4.1.}$$

$$\text{Therefore } \mathrm{Var}(x) = \frac{n\upsilon}{2}\left[\frac{\Gamma(\frac{n\upsilon}{2})\Gamma(\frac{n\upsilon}{2} - 1) - \Gamma^2(\frac{n\upsilon}{2} - \frac{1}{2})}{\Gamma^2(\frac{n\upsilon}{2})}\right].$$

Theorem 5.4.4: If $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are iid $MVN_\upsilon[\underset{\sim}{T}, A]$, the asymptotic mean of $\underset{\sim}{C}pm$ is 1.

Proof: If $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are iid $MVN_\upsilon[\underset{\sim}{T}, A]$, from Theorem 5.4.2

$$E(x) = \sqrt{\frac{n\upsilon}{2}}\,\frac{\Gamma(\frac{n\upsilon}{2} - \frac{1}{2})}{\Gamma(\frac{n\upsilon}{2})} \quad . \text{ Then as } n\upsilon \text{ approaches } \infty$$

from [18] $\lim_{n\to\infty} \sqrt{\frac{n\upsilon}{2}}\,\frac{\Gamma(\frac{n\upsilon}{2} - \frac{1}{2})}{\Gamma(\frac{n\upsilon}{2})} = 1.$

Theorem 5.4.5: If $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are iid $MVN_\upsilon[\underset{\sim}{T}, A]$, the asymptotic variance of $\underset{\sim}{C}pm$ is 0.

Proof: If $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are iid $MVN_\upsilon[\underset{\sim}{T}, A]$, from Theorem 5.4.3

$$\text{Var}(x) = \frac{n\upsilon}{2}\left[\frac{\Gamma(\frac{n\upsilon}{2})\Gamma(\frac{n\upsilon}{2} - 1) - \Gamma^2(\frac{n\upsilon}{2} - \frac{1}{2})}{\Gamma^2(\frac{n\upsilon}{2})}\right], \text{ then the limit as } n\upsilon \to \infty \text{ is}$$

$$= \lim_{n\to\infty}\left[\frac{n\upsilon}{2}\frac{\Gamma(\frac{n\upsilon}{2} - 1)}{\Gamma(\frac{n\upsilon}{2})}\right] - \lim_{n\to\infty}\left[\frac{n\upsilon}{2}\frac{\Gamma^2(\frac{n\upsilon}{2} - \frac{1}{2})}{\Gamma^2(\frac{n\upsilon}{2})}\right]$$

From [18] both limits are 1, thus the asymptotic variance of $\underset{\sim}{C}pm$ is 0.

5.5. HYPOTHESIS TESTING

Knowledge of the pdf of $\underset{\sim}{C}pm$ permits statistically based inferences to be made regarding the capability of the process. The hypotheses

H_0: Process is capable

versus

H_a: Process is not capable

can be tested to determine whether or not a process is capable of meeting the specification limits for a particular process.

The specification limits completely determine the covariance matrix (i.e., A) used in the computations of $\underset{\sim}{C}pm$. If this matrix is an accurate reflection of the variations inherent in the process then the sample results should have a similar covariance structure. If this is the case, then the sum of the Mahalanobis distances will be close to $n\upsilon$ and hence $\underset{\sim}{C}pm$ should be in the vicinity of 1. If the sample covariance structure is not similar to that of A,

the value of $\underset{\sim}{C}pm$ should reflect these departures. If for example the variations in the sample results are larger than the variation specified in A, $\underset{\sim}{C}pm$ should be smaller than 1. Similarly if the covariance structure found in the sample results is different from that of A, $\underset{\sim}{C}pm$ should again be smaller than 1. On the other hand if the variation found in the sample is smaller than that specified in A the value of $\underset{\sim}{C}pm$ should be larger than 1.

Because only small values of $\underset{\sim}{C}pm$ are of interest, one-sided critical regions are appropriate for testing the hypotheses stated above. Using the value of $\underset{\sim}{C}pm$ determined from the sample results as a test statistic, the rejection region for the hypothesis will be [0, W], where $Pr(\underset{\sim}{C}pm<W)=\alpha$, given A to be an accurate measure of the inherent process variation. The value of W for values of $n\upsilon = 40(5)250$ and $\alpha=0.05$ are included in Table 5.5.1.

5.6. EXAMPLES

Example 5.6.1: Sultan [34] discusses an example in the context of control charts where the brinell hardness (H) and the tensile strength (S) of the output of a process are of interest. Assuming that the process is in control, consider the following hypothetical capability study of the process using the observed data.

Suppose that an engineering study suggests that the variance associated with H should be no more than 324, the variance of S no more than 25 with H and S having a covariance of 65, and with the target values of 177 for H and 53 for S. With this information specification limits for this capability study can be created (Figure 5.6.1).

The region created by the specification limits in Figure 5.6.1 being long and thin reflects the fact that there is greater variability inherent in H than in S. The positive covariance between H and S is depicted by the angle between the principal axis and the co-ordinate axis. The specification boundaries are completely specified by a=0.9973,

Table 5.5.1

Values of W for α=0.01, 0.025, 0.05 and 0.10 where Pr(C_{pm} < W_{α}) = α.

df	$W_{0.01}$	$W_{0.025}$	$W_{0.05}$	$W_{0.10}$
40	0.7925	0.8210	0.8470	0.8797
45	0.8020	0.8294	0.8543	0.8846
50	0.8103	0.8367	0.8606	0.8897
55	0.8175	0.8431	0.8662	0.8941
60	0.8239	0.8487	0.8710	0.8980
65	0.8297	0.8537	0.8754	0.9015
70	0.8349	0.8583	0.8793	0.9047
75	0.8396	0.8624	0.8829	0.9075
80	0.8439	0.8662	0.8861	0.9101
85	0.8479	0.8696	0.8891	0.9125
90	0.8515	0.8728	0.8919	0.9147
95	0.8549	0.8758	0.8944	0.9167
100	0.8581	0.8785	0.8968	0.9186
105	0.8611	0.8811	0.8990	0.9204
110	0.8638	0.8835	0.9011	0.9220
115	0.8664	0.8858	0.9030	0.9236
120	0.8689	0.8879	0.9048	0.9251
125	0.8712	0.8899	0.9066	0.9264
130	0.8734	0.8918	0.9082	0.9277
135	0.8755	0.8936	0.9097	0.9290
140	0.8774	0.8953	0.9112	0.9301
145	0.8793	0.8970	0.9126	0.9312
150	0.8811	0.8985	0.9139	0.9323
155	0.8828	0.9000	0.9152	0.9333
160	0.8845	0.9014	0.9164	0.9343
165	0.8860	0.9028	0.9176	0.9352
170	0.8875	0.9041	0.9187	0.9361
175	0.8890	0.9053	0.9198	0.9369
180	0.8904	0.9065	0.9208	0.9377
185	0.8917	0.9076	0.9218	0.9385
190	0.8930	0.9088	0.9227	0.9393
195	0.8942	0.9098	0.9236	0.9400
200	0.8954	0.9109	0.9245	0.9408
205	0.8966	0.9119	0.9254	0.9414
210	0.8977	0.9128	0.9262	0.9420
215	0.8988	0.9138	0.9270	0.9426
220	0.8998	0.9147	0.9277	0.9432
225	0.9008	0.9155	0.9285	0.9438
230	0.9018	0.9164	0.9292	0.9444
235	0.9028	0.9172	0.9299	0.9450
240	0.9037	0.9180	0.9306	0.9455
245	0.9046	0.9188	0.9312	0.9460
250	0.9054	0.9195	0.9319	0.9465

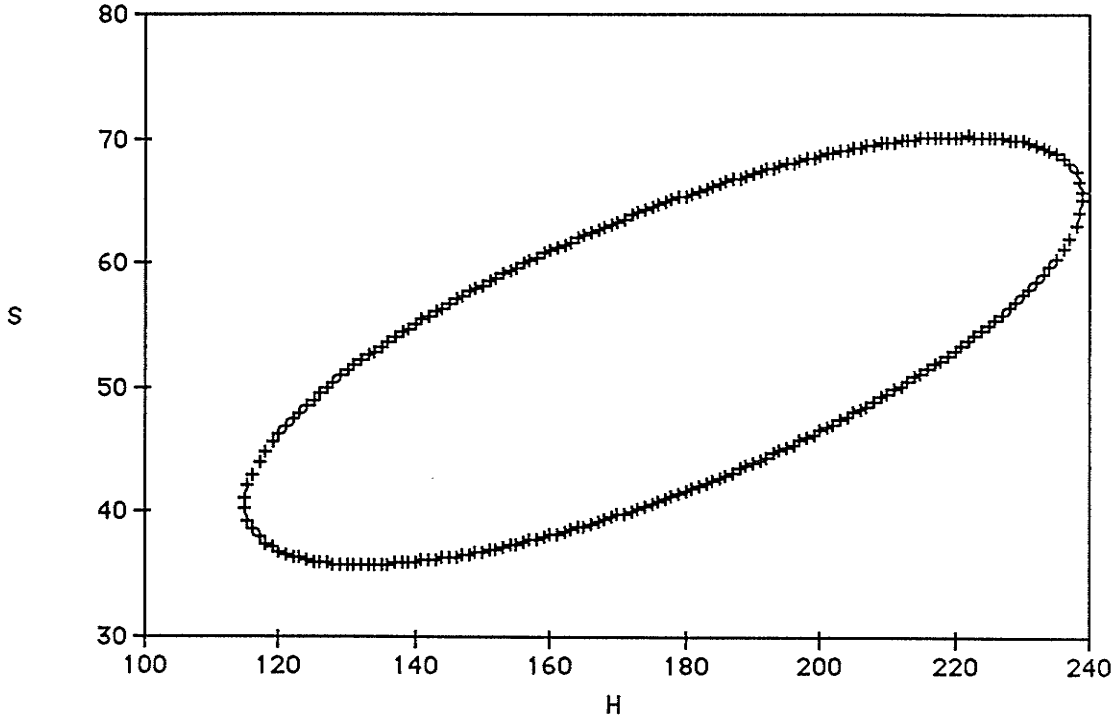


Figure 5.6.1 Boundaries for Example 5.6.1.

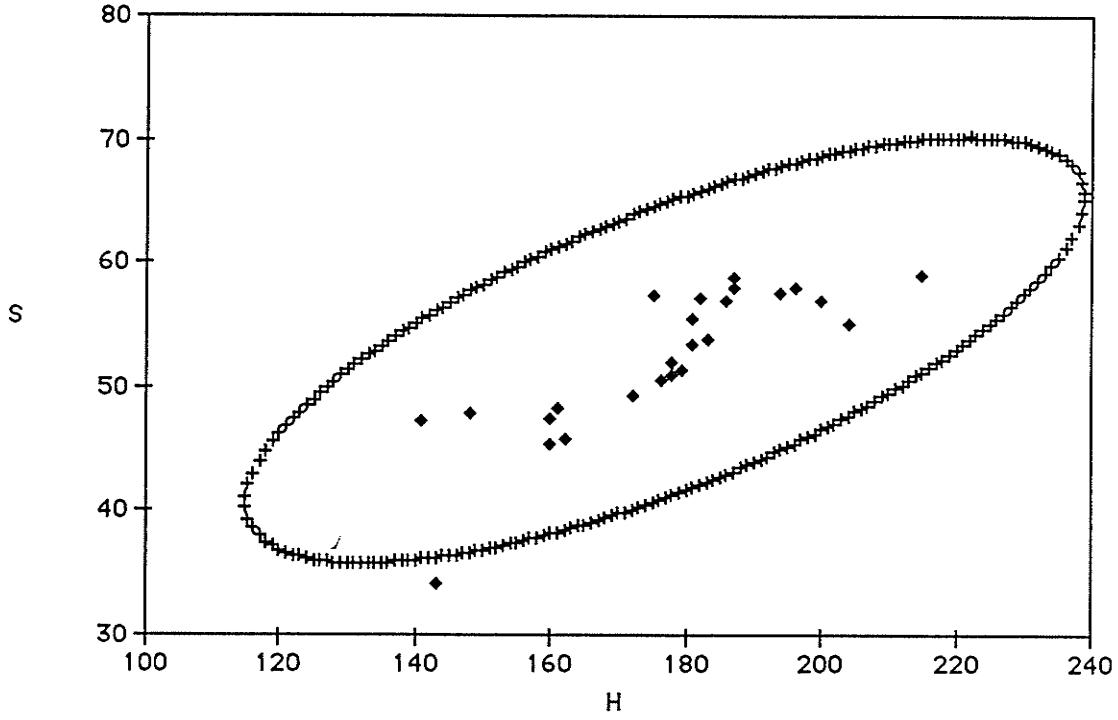


Figure 5.6.2 Boundaries values and observations for Example 5.6.1.

$T = \begin{bmatrix} 177 \\ 53 \end{bmatrix}$ and $A = \begin{bmatrix} 324 & 65 \\ 65 & 25 \end{bmatrix}$ and are of the form

$$\left[\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} - \begin{Bmatrix} 177 \\ 53 \end{Bmatrix}\right]' \begin{bmatrix} 324 & 65 \\ 65 & 25 \end{bmatrix}^{-1} \left[\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} - \begin{Bmatrix} 177 \\ 53 \end{Bmatrix}\right] = 11.829 \ .$$

Figure 5.6.2 contains both the specification limits and the plotted observations. From the plot it becomes clear that i) all but one observation falls within the specification boundaries and ii) the points tend to be clustered in the vicinity of the target. This graphical result appears to suggest that the process is capable of meeting the engineering requirements. In some situations the graphical result may be all that is required. However in order to draw comparisons with other processes, or to make judgements regarding changes in the capability of the process over time, a numerical measure of process capability is required.

Using the proposed numerical measure

$$\underset{\sim}{C}pm = \sqrt{\frac{n\upsilon}{\sum_{i=1}^{n} (\underset{\sim}{X}_i - \underset{\sim}{T})' A^{-1} (\underset{\sim}{X}_i - \underset{\sim}{T})}}$$

and $\underset{\sim}{T} = \begin{bmatrix} 177 \\ 53 \end{bmatrix}$, $A = \begin{bmatrix} 324 & 65 \\ 65 & 25 \end{bmatrix}$, $\upsilon=2$ and $n=25$. The observed values of brinell hardness and tensile strength along with their Mahalanobis distances (D^2) have been included

H	S	D^2
143	34.2	15.5661
200	57.0	1.6643
160	47.5	1.2570
181	53.4	0.0629
148	47.8	2.6276
178	51.5	0.2449
162	45.9	2.0936
215	59.1	4.6508
161	48.4	0.9517
141	47.3	4.1937
175	57.3	1.8603
187	58.5	1.3293
187	58.2	1.1615
186	57.0	0.6526
172	49.4	0.6410
182	57.2	0.9317
177	50.6	0.4816
204	55.1	3.1698
178	50.9	0.4456
196	57.9	1.2132
160	45.5	2.2903
183	53.9	0.1188
179	51.2	0.4175
194	57.5	0.9912
181	55.6	0.3195

resulting in a $\underset{\sim}{C}pm$ of 1.0067 . From Table 5.5.1 the critical value for $n\upsilon=50$ and $\alpha=0.05$ is 0.8606. Then because 1.0067>0.8606 we fail to reject H_0 and conclude that the process is capable.

Example 5.6.2: The same process is again considered, but with more stringent engineering specifications. Assume that in this case an engineering study suggested that the variance associated with H should be no more than 196, the variance of S no more than 9 with H and S having a covariance of 25, and with the target values of 175 for H and 55 for S. With a=0.9973, $\underset{\sim}{T}' = \begin{bmatrix} 175 \\ 55 \end{bmatrix}$ and $A = \begin{bmatrix} 196 & 25 \\ 25 & 9 \end{bmatrix}$ the equation specifying the specification boundaries is

$$\left[\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} - \begin{Bmatrix} 175 \\ 55 \end{Bmatrix}\right]' \begin{bmatrix} 196 & 25 \\ 25 & 9 \end{bmatrix}^{-1} \left[\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} - \begin{Bmatrix} 175 \\ 55 \end{Bmatrix}\right] = 11.829 .$$

From Figure 5.6.3 it is clear that i) one of the 25 observations falls outside the specification bounds and ii) the points are now much closer to the specifications and do not appear to be clustered around the target . These graphical results suggest that the process is incapable of meeting the more stringent engineering requirements.

The Mahalanobis distances (D^2) have been calculated using $\underset{\sim}{T} = \begin{bmatrix} 175 \\ 55 \end{bmatrix}$, $A = \begin{bmatrix} 196 & 25 \\ 25 & 9 \end{bmatrix}$, n=25 and υ=2 and along with H and S are

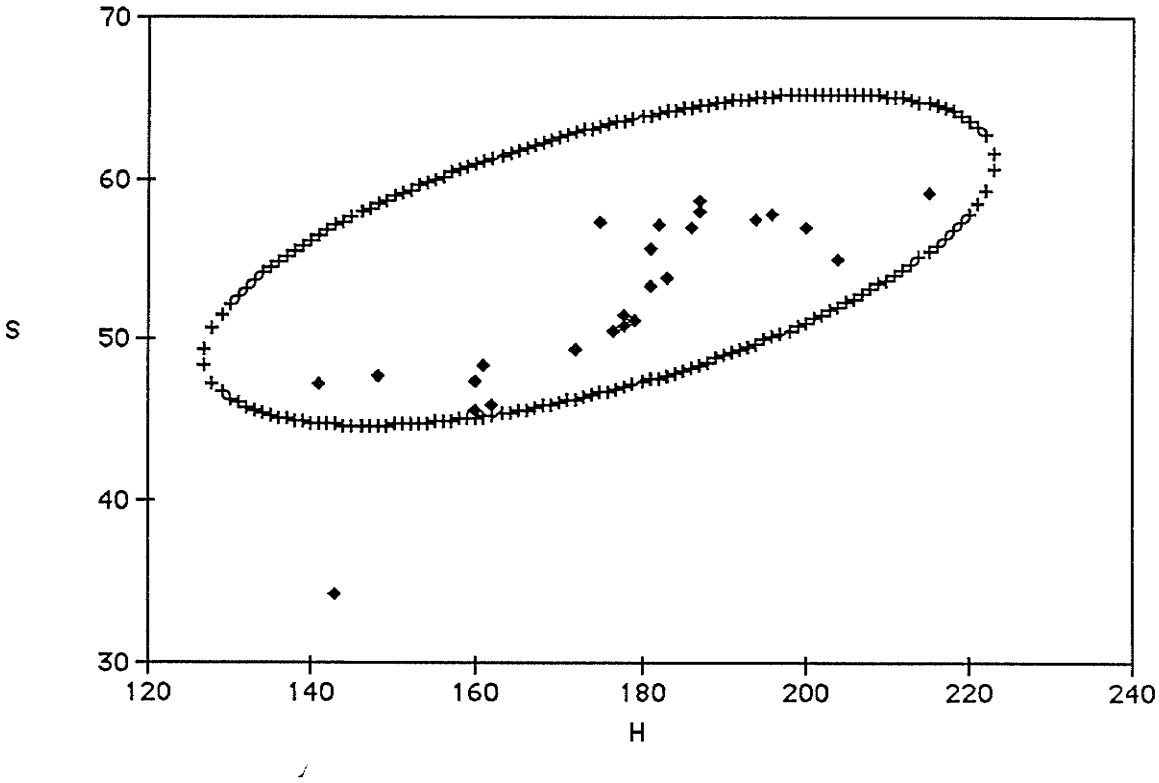


Figure 5.6.3 Boundaries and observations for Example 5.6.2.

H	S	D^2
143	34.2	53.3217
200	57.0	3.4320
160	47.5	6.5189
181	53.4	1.1464
148	47.8	6.1472
178	51.5	2.6400
162	45.9	10.3922
215	59.1	8.3360
161	48.4	4.9884
141	47.3	7.8445
175	57.3	0.9103
187	58.5	1.4021
187	58.2	1.2143
186	57.0	0.6787
172	49.4	4.7301
182	57.2	0.5440
177	50.6	3.7494
204	55.1	6.5197
178	50.9	3.5037
196	57.9	2.2584
160	45.5	11.0527
183	53.9	1.1002
179	51.2	3.2785
194	57.5	1.8428
181	55.6	0.1884

resulting in a $\underset{\sim}{C}pm$ of 0.5818. The critical value for the hypotheses

H_0: Process is capable

versus

H_a: Process is not capable

is 0.8606 (α=0.05). Then because 0.5818<0.8606, one would reject H_0 concluding that the process is not capable of meeting the more stringent engineering requirements (again agreeing with the graphical result).

5.7. COMMENTS

A multivariate measure of process capability is proposed that considers proximity to the target and dispersion around the target while having some reasonably good statistical

properties. The index provides a unitless measure that will permit comparisons within a process as well as among competing processes.

A benchmark of 1 for judging a process capable is suggested, as this represents the case where the sample covariance structure (centred on the target) is similar to the covariance structure (around the target) specified by the specification limits. As the engineering requirements change or the process changes, modifications to the specification boundaries can be made to maintain the benchmark of 1.

Because the degrees of freedom associated with the sum of the Mahalanobis distances are the product of the sample size and the number of capability variables used, this value will generally be reasonably large. As a result the asymptotic properties of $\underset{\sim}{C}pm$ will be of importance.

It has been assumed that the engineering specifications are given in terms of the required target and covariance structure. Situations may arise where these specifications are not given in terms of $\underset{\sim}{T}$, A and c. If only the physical region is given, the engineering information must be converted to an ellipsoid with known $\underset{\sim}{T}$ and A, as both must be known before $\underset{\sim}{C}pm$ can be determined. Shakun [35] and Jackson and Bradley [36] discuss two possible techniques.

Chapter 6

An Alternate Variables Control Chart

6.1. INTRODUCTION

Much attention has been focussed on the concept of simultaneous control charts. Boxplots ([37], [38]) are the latest in a series of improvements made to the traditional control chart originally developed by Shewhart [39]. In general boxplots can be very informative but tend to be quite overpowering. Unless users are familiar with the plots, the vast amount of information contained in a single plot can be confusing. Proposed additions to the charts regarding robust and resistant measures only serve to further clutter the inferences drawn. In the hands of an experienced data analyst boxplots and the proposed resistant and robust measures can be enlightening. However they may only serve to detract from the general inferences required on the manufacturing floor. Because of their complexity, boxplots are rarely performed by hand. In fact they are rather difficult and time consuming to construct, however most common statistical packages provide some form of the boxplot. In those cases where a computer is available or where the data are entered directly to a computer, boxplots may be quite convenient.

There is an abundance of newly developed control chart procedures designed to enhance the analysis of a process. However as pointed out by Woodall [40], none reflect the current change of philosophy in Quality Control. The proposed technique attempts to reflect the change of philosophy currently taking place in Quality Control on the North American continent. Both proximity to target and the variability are of concern in today's quality assurance programs. No longer is "conforming to specifications", without a nominal value, sufficient. The proposed procedure attempts to adopt this philosophy in a simultaneous control chart.

The technique proposed here is similar to the traditional control charts outlined by Shewhart. The resultant charts are clear and uncluttered (similar to the traditional charts) with the added advantage of appearing on a single plot. Control limits for the proposed measures are also included. The new control chart is highly visible with inferences easily drawn. The procedure is straightforward and can easily be carried out by employees whose major task is to monitor and adjust the process. The plots can be produced by hand with the univariate case requiring only the aid of a hand-held calculator while the multivariate case generally requires a computer capable of performing accurate matrix operations. In either case the procedure is adaptable to the computer and computer graphics.

6.2. The Proposed Procedure and Resultant Control Chart

Traditional North American Quality Control techniques are being challenged by innovative philosophies imported from Japan. One particular philosophy which has been discussed by many stresses the need of a nominal (target) value when assessing the quality levels of a process. In the past, attention has focussed on bringing a process within specification while placing little or no emphasis on attaining target values. Sullivan [29] among others gives examples that stress the importance of controlling target values as well as variability. Control charts that do not employ a target value are very much a part of the old guard. Good information can be drawn from the traditional control chart, however by integrating targets into the plotting procedure additional information is made available.

Traditional control charts plot the mean and either the sample range or standard deviation of each subgroup. If both of these measures reside within the control limits the process is deemed "in-control". The boundaries used to aid in drawing inferences regarding the process (i.e., the control limits) are also created from the sample results. The control limits for the subgroup means (i.e., the $\bar{x}$'s) are of the form

$$\bar{\bar{X}} \pm A_2 \bar{R} \quad \text{or} \quad \bar{\bar{X}} \pm A_3 \bar{S}$$

where $\bar{\bar{X}} = \sum_{j=1}^{k} \frac{\bar{x}_j}{k}$ is the mean of the k subgroup means, $\bar{R} = \sum_{j=1}^{k} \frac{R_j}{k}$ and $\bar{S} = \sum_{j=1}^{k} \frac{s_j}{k}$ are the mean of the k subgroup ranges and standard deviations respectively and A_2, A_3 are statistically determined constants. The control limits associated with the subgroup range, R, and the subgroup standard deviation, s, are of the form

$$D_3 \bar{R} \quad \text{and} \quad D_4 \bar{R},$$

$$B_3 \bar{S} \quad \text{and} \quad B_4 \bar{S}$$

respectively. D_3, D_4 and B_3, B_4 are again statistically determined constants.

The technique proposed here incorporates proximity to the target value into the plotting procedure while continuing to examine the inherent variability of the process. The procedure maintains the ability to draw similar inferences to those of the traditional control charts while being ideally suited to the new philosophies being incorporated into North American Quality circles.

The suggested measures of process performance are derived from the Mean Square Error (MSE). The MSE is a measure of the squared distances from some nominal value, in this case taken to be the target value. The MSE is then

$$MSE = \sum_{i=1}^{n} \frac{(x_i - T)^2}{n}$$

where x_i denotes the i^{th} measure of a particular subgroup, n the subgroup sample size and T the target or nominal value.

The MSE associated with the target value measures both the variability and proximity to the target value simultaneously, actually confounding the two measures.

However the MSE is easily partitioned into components which provide individual measures for the inherent variability and proximity to the target. Consider that the MSE

$$MSE=\frac{1}{n}\sum_{i=1}^{n}(x_i - T)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}+\bar{x} - T)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 + (\bar{x} - T)^2 = \frac{n-1}{n}s^2 + (\bar{x} - T)^2$$

breaks down into two terms, the first being a function of the sample variance and the other the squared difference between the subgroup mean and the target value. The proposed technique uses the numerical value of the MSE as a measure of overall variability and the squared difference between the subgroup mean and the target as a measure of proximity to the target value. Occasionally the subgroup variance will be of interest and can be added to the plot on an individual basis as required. Assuming the measures arise from a normal distribution, procedures for determining the control limits for the new measures will be illustrated and the associated constants determined.

The proposed monitoring measures are

$$(\bar{x} - T)^2$$

$$\text{and} \quad MSE = \sum_{i=1}^{n}\frac{(x_i - T)^2}{n}$$

($s^2 = \sum_{i=1}^{n}\frac{(x_i - \bar{x})^2}{n-1}$ will also receive occasional attention) where T denotes the general target value of the entire process, n the sample size associated with each subgroup sample taken, x_i the i^{th} measure of a subgroup, $\bar{x}$ the subgroup mean and s^2 the sample variance of the subgroup.

The statistical distributions associated with these measures can be found for the case where the process measurements are assumed to be iid $N(\mu,\sigma^2)$. Knowledge of the

distributions permits calculation of the control limits for various sample sizes. The control limits associated with the measures will be shown to be of the form

$$\text{i)} \quad A\sigma^2 \quad \text{for} \quad (\bar{x} - T)^2,$$

$$\text{ii)} \quad C\sigma^2 \quad \text{for} \quad MSE$$

$$\text{and} \quad \text{iii)} \quad B\sigma^2 \quad \text{for} \quad s^2$$

where A, B and C are constants derived from the statistical distributions associated with each of the measures.

6.3. Derivation of the Control Limits and Constants

In each of the following cases the process measurements are assumed to be independently distributed normal variates with mean μ and variance σ^2 (i.e., iid $N(\mu, \sigma^2)$). For reasons discussed later only upper control limits (UCL) will be developed.

UCL for $(\bar{x}-T)^2$

Assuming $X \sim N(\mu, \sigma^2)$ it follows that $(\bar{x} - T)^2 \sim \frac{\sigma^2}{n}\chi^2_{1,\lambda}$ (see Theorem 6.7.1), where $\chi^2_{1,\lambda}$ denotes the non-central chi-square distribution with one degree of freedom (df) and non-centrality parameter $\lambda = n\left[\frac{\mu - T}{\sigma}\right]^2$. Defining the UCL to be the $(1-\alpha)100\%$ percentile of the distribution function associated with $(\bar{x} - T)^2$ results in

$$UCL_{(\bar{x} - T)^2} = \frac{\sigma^2}{n}\chi^2_{1,\lambda,(1-\alpha)}$$

$$= \sigma^2 A \quad .$$

Table 6.3.1 contains the values of A, denoted A_3 for $(1-\alpha)100\%=99.73\%$, n=2(1)10 and values of $\left[\frac{\mu - T}{\sigma}\right]^2 = 0.00(0.05)0.95$.

In order to determine A for a particular process, $\left[\frac{\mu - T}{\sigma}\right]^2$ must be known. This is seldom the case as μ and σ are process population parameters. In practice $\bar{\bar{X}} = \sum_{j=1}^{k} \frac{\bar{x}_j}{k}$ and $\overline{S^2} = \sum_{j=1}^{k} \frac{s_j^2}{k}$, where k denotes the number of subgroups and $\bar{x}_j$, s_j^2 the sample mean and variance of the j^{th} subgroup, are substituted for μ and σ^2 respectively, resulting in a reasonable estimate (see Theorem 6.7.4 for discussion) for the value of A and hence for the UCL. As a result the control limit for $(\bar{x} - T)^2$ will generally be of the form

$$UCL_{(\bar{x} - T)^2} = A\,\overline{S^2} .$$

Other values of $(1-\alpha)100\%$ may be used when creating the UCL. If, for example, the practitioner wishes to use boundaries equivalent to $\pm 2\sigma$ then $(1-\alpha)100\%$ is set to 95.44% and the values of A redetermined using $\chi^2_{1,\,\lambda,\,0.9544}$.

Note that a lower control limit will not be required for this measure as the minimum value $(\bar{x} - T)^2$ can assume will be zero which arises when the subgroup mean is the same as the target (i.e., the optimal situation).

UCL for MSE

The $MSE = \sum_{i=1}^{n} \frac{(x_i - T)^2}{n}$ will follow a $\frac{\sigma^2}{n}\chi^2_{n,\lambda}$ distribution (see Theorem 6.7.2), where $\chi^2_{n,\lambda}$ denotes the non-central chi-square distribution with n df and $\lambda = n\left[\frac{\mu - T}{\sigma}\right]^2$. Similar to the UCL for $(\bar{x} - T)^2$ the (1-α)100% percentile associated with $\frac{\sigma^2}{n}\chi^2_{n,\lambda}$ will be used as the UCL for MSE, resulting in

$$UCL_{MSE} = \frac{\sigma^2}{n}\chi^2_{n,\lambda,(1-\alpha)}$$

$$= \sigma^2 \, C \quad .$$

Table 6.3.2 contains values associated with C, denoted C_3 when (1-α)100%=99.73%, for n=2(1)10 and $\left[\frac{\mu - T}{\sigma}\right]^2 = 0.00(0.05)0.95$. Analogous to the UCL for $(\bar{x} - T)^2$, the statistical distribution associated with MSE has a non-centrality parameter that is a function of the population parameters μ and σ^2. Since μ and σ^2 are rarely known the practice is to again substitute $\bar{\bar{X}}$ and $\overline{S^2}$ respectively and to use the result as a measure of $\left[\frac{\mu - T}{\sigma}\right]^2$. In practice the form of the UCL for MSE will be

$$UCL_{MSE} = C\,\overline{S^2}.$$

UCL for s^2

Again assuming the process measurements to be iid $N(\mu, \sigma^2)$, $s^2 \sim \chi^2_{n-1,0}$ (see Theorem 6.7.3) where $\chi^2_{n-1,0}$ denotes the central chi-square (i.e., λ=0) distribution.

Defining the UCL to represent the (1-α)100% percentile of the distribution associated with UCL_{s2} we get

$$UCL_{s2} = \frac{\sigma^2}{(n-1)} \chi^2_{(n-1),0,\,(1-\alpha)}$$

$$= \sigma^2 B \quad .$$

The value of B for n=2(1)20, (1-α)=0.9973 and 0.9544 (denoted B_3 and B_2 respectively) have been determined and included in Table 6.3.3. The coefficients in Table 6.3.3 result in more stringent UCL than those associated with the traditional σ charts. The reason is that it is our philosophy not to consider (at least physically) a lower control limit for s^2. In this way the control chart becomes a single-tailed problem. Identification and investigation of "sharp" declines in subgroup variation is encouraged, but not seen as an "error". For this reason only the UCL for both s^2 and MSE is considered.

The UCL_{s2} depends upon the value of the population parameter σ^2 and B. B is easily determined for a fixed n and (1-α), however σ^2 is rarely known. The practice is again to replace σ^2 with $\overline{S^2}$, resulting in

$$UCL_{s2} = B\,\overline{S^2}.$$

6.4. Computations, Plotting Strategies and Inferences

The procedure for creating the proposed plots is very similar to the procedure followed for traditional $\bar{x}$ and s charts. The subgroup means and variances, $\bar{x}$ and s^2, as well as $\bar{\bar{X}} = \sum_{j=1}^{k} \frac{\bar{x}_j}{k}$ and $\overline{S^2} = \sum_{j=1}^{k} \frac{s_j^2}{k}$ must be determined. In addition the practitioner must find $(\bar{x} - T)^2$ for each subgroup, noting the cases where $\bar{x} < T$. The MSE for each subgroup can then be determined using

Table 6.3.1. Coefficient A_3 used to determine the boundary value associated with $(\bar{x} - T)^2$ for various values of n and λ.

$$\lambda = \left[\frac{\mu - T}{\sigma}\right]^2$$

n	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
2	4.50	4.91	5.26	5.56	5.84	6.09	6.33	6.55	6.76	6.96	7.15	7.34	7.52	7.69	7.86	8.03	8.19	8.35	8.50	8.66
3	3.00	3.39	3.71	3.98	4.22	4.44	4.64	4.83	5.01	5.19	5.35	5.51	5.67	5.82	5.97	6.11	6.25	6.39	6.53	6.66
4	2.25	2.63	2.92	3.16	3.38	3.58	3.76	3.93	4.09	4.25	4.40	4.55	4.69	4.83	4.96	5.09	5.22	5.35	5.47	5.60
5	1.80	2.17	2.44	2.66	2.86	3.04	3.21	3.37	3.52	3.67	3.81	3.94	4.08	4.20	4.33	4.45	4.57	4.69	4.81	4.92
6	1.50	1.85	2.11	2.32	2.51	2.68	2.83	2.98	3.13	3.26	3.40	3.52	3.65	3.77	3.89	4.01	4.12	4.23	4.35	4.45
7	1.29	1.63	1.87	2.07	2.25	2.41	2.56	2.70	2.84	2.97	3.09	3.22	3.33	3.45	3.57	3.68	3.79	3.89	4.00	4.11
8	1.12	1.46	1.69	1.88	2.05	2.20	2.35	2.48	2.61	2.74	2.86	2.98	3.09	3.20	3.31	3.42	3.53	3.63	3.73	3.84
9	1.00	1.33	1.55	1.73	1.89	2.04	2.18	2.31	2.43	2.55	2.67	2.79	2.90	3.01	3.11	3.22	3.32	3.42	3.52	3.62
10	0.90	1.22	1.43	1.61	1.76	1.90	2.04	2.17	2.29	2.40	2.52	2.63	2.74	2.84	2.95	3.05	3.15	3.25	3.34	3.44

Table 6.3.2. Coefficient C_3 used to determine the boundary value associated with $\frac{1}{n}\sum_{i=1}^{n}(x_i - T)^2$ for various values of n and λ.

$$\lambda = \left[\frac{\mu - T}{\sigma}\right]^2$$

n	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
2	11.82	12.39	12.91	13.40	13.85	14.28	14.69	15.08	15.45	15.82	16.17	16.51	16.85	17.17	17.49	17.80	18.11	18.41	18.71	19.00
3	7.08	7.42	7.73	8.02	8.30	8.57	8.82	9.07	9.30	9.53	9.76	9.97	10.19	10.39	10.60	10.80	11.00	11.19	11.38	11.57
4	5.42	5.68	5.92	6.15	6.37	6.57	6.77	6.97	7.15	7.34	7.51	7.69	7.86	8.03	8.19	8.35	8.51	8.67	8.82	8.98
5	4.55	4.77	4.98	5.17	5.36	5.53	5.70	5.87	6.03	6.19	6.35	6.50	6.65	6.79	6.93	7.08	7.21	7.35	7.49	7.62
6	4.01	4.21	4.39	4.56	4.73	4.89	5.04	5.19	5.34	5.48	5.62	5.75	5.89	6.02	6.15	6.28	6.41	6.53	6.66	6.78
7	3.64	3.82	3.98	4.14	4.29	4.44	4.58	4.72	4.85	4.99	5.12	5.24	5.37	5.49	5.61	5.73	5.85	5.97	6.08	6.19
8	3.37	3.53	3.69	3.83	3.97	4.11	4.24	4.37	4.50	4.62	4.75	4.87	4.98	5.10	5.21	5.33	5.44	5.55	5.66	5.76
9	3.16	3.31	3.46	3.59	3.73	3.86	3.98	4.11	4.23	4.35	4.46	4.57	4.69	4.80	4.91	5.01	5.12	5.23	5.33	5.43
10	2.99	3.13	3.27	3.25	3.53	3.66	3.78	3.89	4.01	4.12	4.23	4.34	4.45	4.56	4.66	4.76	4.87	4.97	5.07	5.17

Table 6.3.3. Coefficients B_2 and B_3 used to determine the boundary value associated with s^2 for various values of n .

n	B_3	B_2
2	9.00	4.00
3	5.91	3.09
4	4.72	2.67
5	4.06	2.43
6	3.64	2.26
7	3.34	2.14
8	3.12	2.05
9	2.95	1.97
10	2.81	1.91
11	2.69	1.86
12	2.59	1.82
13	2.51	1.79
14	2.44	1.75
15	2.37	1.72
16	2.31	1.69
17	2.26	1.67
18	2.22	1.64
19	2.18	1.62
20	2.14	1.61

$$MSE = \frac{n-1}{n} s^2 + (\bar{x} - T)^2.$$

Finally the UCL for each of the measures must be found using the appropriate algorithms presented earlier.

The suggested plotting strategy is to first plot the values of $(\bar{x} - T)^2$ for each of the subgroups, along with the $UCL_{(\bar{x} - T)^2}$. Use of the plotting characters "-" if $\bar{x} < T$ and "+" if $\bar{x} \geq T$ when plotting $(\bar{x} - T)^2$ is suggested. In this way any trends which may have been obscured through examining squared differences from the target can be identified.

The practitioner may now proceed to plot MSE for each of the subgroups along with the UCL for both s^2 and MSE. The plotting characters used here are of little consequence and any useful character can be used. Any trends that may occur here should be obvious regardless of the units used. Plotting the MSE rather than s^2 is suggested, as the MSE must always be greater than or equal to $(\bar{x} - T)^2$ for each subgroup. This results in no overlapping of plots, which is not necessarily the case with s^2. One reason for $\bar{x}$ and s requiring two plots is because a single plot can become extremely difficult to draw inferences from when crossovers are common (see Figure 6.4.1). The boxplot technology eliminates the s chart and the associated UCL by using boxes proportional in size to the subgroup standard deviation. If the standard deviation of a subgroup exceeds the UCL the proportional box representing the subgroup standard deviation is then drawn in a slightly different manner. The boxplot technique does use a visible UCL for s.

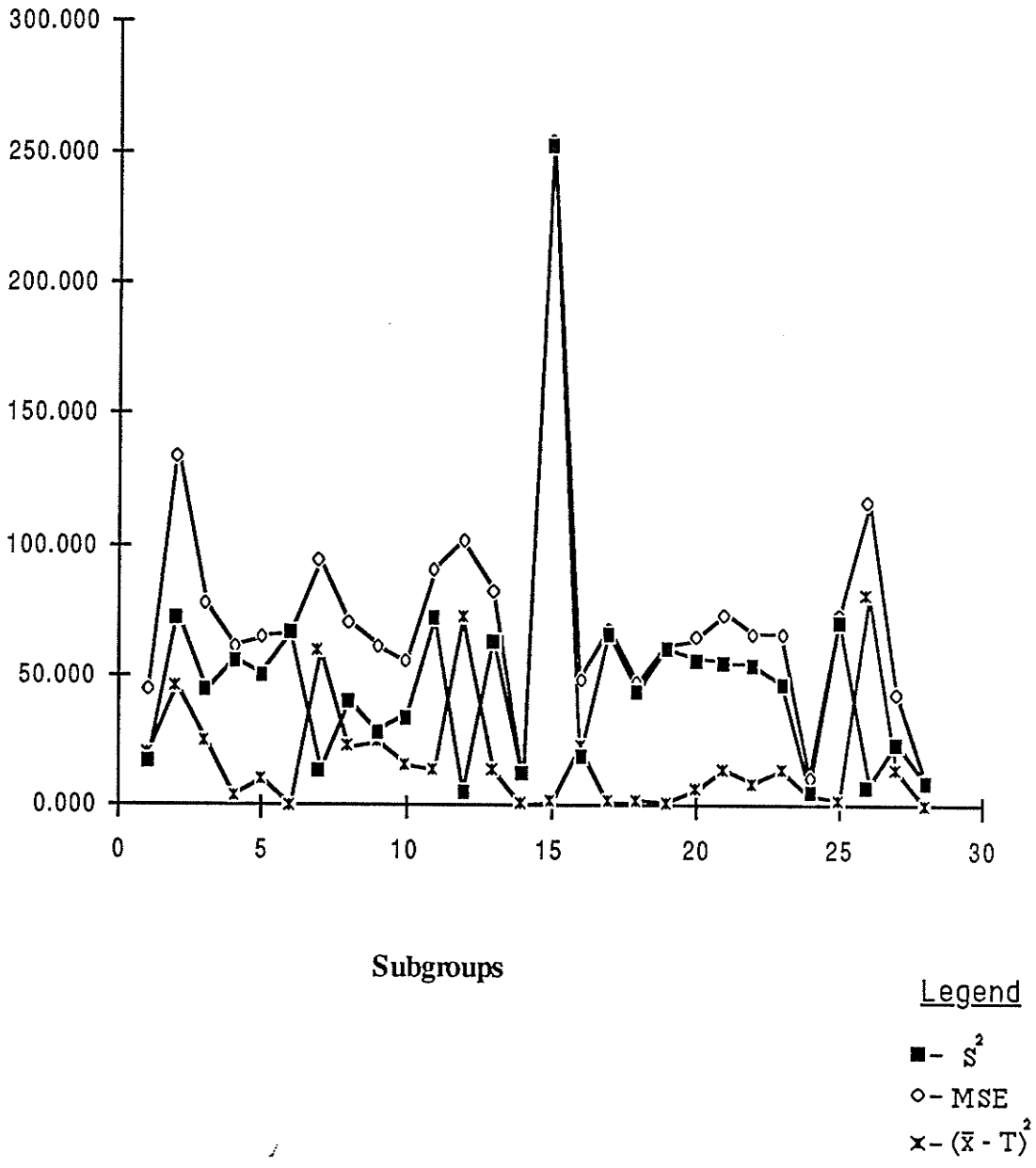


Figure 6.4.1. An Example of crossovers affecting the clarity of the plot.

Once the plot is complete inferences regarding the process are possible. Similar to the standard control charts those subgroups with either (or both) of the proposed measures above the appropriate UCL should be highlighted and investigated. As well if any sharp declines, unusual trends or actions appear they too should be investigated. Any decisions and actions made at this stage will be similar to those taken for the traditional control charts. If for example the subgroup means appear to be drifting away from the target or if the overall variability is changing in a systematic fashion, adjustments should be made.

As part of the plotting procedure it was suggested that the UCL for both s^2 and MSE be included on the chart however s^2 was not included in the procedure. s^2 can disturb the clarity of the plot to the extent that visual inferences become quite difficult. However MSE may not in appearance provide all necessary information. In some subgroups it may occur that both MSE and $(\bar{x} - T)^2$ are within the acceptable boundaries but s^2 exceeds its UCL. This may occur when the $(\bar{x} - T)^2$ is small. In cases such as this the difference between plotted points $(\bar{x} - T)^2$ and MSE will be large indicating that s^2 is large. Although the plot will in fact indicate this, the result is similar to boxplot representations of the standard deviation, in that the result does not immediately "jump off" the plot. Users familiar with the procedure will look for this phenomenon but others may miss it. To ensure detection, the value of s^2 for all cases where MSE exceeds the UCL_{S^2} should be included in the plot. In this way good visual inferences will result without the excess clutter associated with the entire set of values for s^2. An example has been selected that illustrates this particular situation (see Example 6.8.1).

6.5. MULTIVARIATE CASE

Similarly styled control charts can be created for those situations where more than one variable (i.e., multivariate case) is used to monitor a process. Target values continue to be important in the multivariate case, while Alt [41] states that "development of one control

chart for simultaneous monitoring of both location and dispersion is needed". Hence the motivation for a simultaneous control chart that reflects the changing nature of Quality Control for the multivariate case is identical to that for the univariate case.

Notation for the multivariate case will be slightly different from that of the univariate case. Vectors of measurements rather than a single value will result from the sampling schemes used to monitor the process. As the number of variables used to monitor a process increases, calculations become more complex. While it is possible to perform the required calculations by hand (using a hand-held calculator) for the case where two variables are used (i.e., bivariate case), the calculations can be tedious and time consuming. A computer capable of performing accurate matrix operations will be required for cases where more than two variables are used.

The methods and sampling schemes used to gather the measurements are assumed to be identical to the univariate case (i.e., k subgroups consisting of n samples are drawn from a process). However in the multivariate case a set of p measurements are made on each sampling unit resulting in a p-dimensional vector of the form

$$\underset{\sim}{X} = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_p \end{bmatrix}$$

for each sampling unit. A subgroup mean will now be a vector of means

$$\underset{\sim}{\bar{\bar{X}}} = \frac{1}{n}\sum_{i=1}^{n} \underset{\sim}{X}_i = \begin{bmatrix} \sum_{i=1}^{n} \frac{x_{1i}}{n} \\ \vdots \\ \sum_{i=1}^{n} \frac{x_{pi}}{n} \end{bmatrix},$$

with the grand mean , denoted $\underset{\sim}{\bar{\bar{X}}}$, also a vector of the form

$$\underset{\sim}{\bar{\bar{X}}} = \sum_{j=1}^{k} \frac{\underset{\sim}{\bar{X}}_j}{k} = \begin{bmatrix} \sum_{j=1}^{k}\sum_{i=1}^{n} \frac{x_{1ij}}{nk} \\ \vdots \\ \sum_{j=1}^{k}\sum_{i=1}^{n} \frac{x_{pij}}{nk} \end{bmatrix}.$$

The subgroup mean vector $\underset{\sim}{\bar{X}}$ and the grand mean vector $\underset{\sim}{\bar{\bar{X}}}$ are completely analogous to the subgroup mean ($\bar{x}$) and the grand mean ($\bar{\bar{X}}$) in the univariate model. For p=1 the multivariate results are exactly those of the univariate case.

The target must also be stated in the form of a vector (denoted $\underset{\sim}{T}$)

$$\underset{\sim}{T} = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_p \end{bmatrix}$$

where $T_1, T_2, \ldots, T_p$ represent the univariate target values associated with each of the p variables used in monitoring the process.

The measures used to monitor a multivariate process will be similar to those used in the univariate models. They are derived in much the same fashion as their univariate counterparts and have associated distribution functions that will be used to create control limits. The proposed measures, although determined from multivariate data, will be scalars.

The measure

$$\tilde{T}_p = (\underset{\sim}{\bar{X}} - \underset{\sim}{T})'\Sigma^{-1}(\underset{\sim}{\bar{X}} - \underset{\sim}{T})$$

where Σ is the variance-covariance matrix, will be used to assess proximity to the target value. For each of the k subgroups $(\underset{\sim}{\bar{X}} - \underset{\sim}{T})'\Sigma^{-1}(\underset{\sim}{\bar{X}} - \underset{\sim}{T})$ provides a scalar measure of proximity to the target. The multivariate analogue of $(\bar{x} - T)^2$ uses standardized squared distances, which on the surface appear different from the univariate case. However the univariate standardization is reflected in the control limits where σ^2 is used to modify the UCL. Hence the multivariate measure $(\underset{\sim}{\bar{X}} - \underset{\sim}{T})'\Sigma^{-1}(\underset{\sim}{\bar{X}} - \underset{\sim}{T})$ provides much the same information as its counterpart $(\bar{x} - T)^2$ in the univariate case.

The measure proposed for assessing the variability within a subgroup for the multivariate case is analogous to the univariate measure MSE. In appearance the multivariate measure differs from its univariate counterpart as it again uses standardized distance measures. However the inferences drawn will be identical to those of the univariate case. The proposed measure is

$$MSE_p = \frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X_i} - \underset{\sim}{T})'\Sigma^{-1}(\underset{\sim}{X_i} - \underset{\sim}{T})$$

and is completely consistent with the univariate measure MSE. The algorithm results in a scalar measure that incorporates both proximity to the target and the inherent variability associated with a particular subgroup into a single measure. Similar to the univariate case, MSE_p can be partitioned into the following components

$$\frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \underset{\sim}{T})'\Sigma^{-1}(\underset{\sim}{X}_i - \underset{\sim}{T}) = \frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \bar{\underset{\sim}{X}})'\Sigma^{-1}(\underset{\sim}{X}_i - \bar{\underset{\sim}{X}}) + (\bar{\underset{\sim}{X}} - \underset{\sim}{T})'\Sigma^{-1}(\bar{\underset{\sim}{X}} - \underset{\sim}{T})$$

with the first component being a measure of variation within a subgroup and the second measuring proximity to the target. Similar to the univariate case each of the terms in the above equation will be used to monitor the process. The proposed measures in the multivariate case are then

$$\text{i) } \tilde{T}_p = (\bar{\underset{\sim}{X}} - \underset{\sim}{T})'\Sigma^{-1}(\bar{\underset{\sim}{X}} - \underset{\sim}{T}),$$

$$\text{ii) } MSE_p = \frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \underset{\sim}{T})'\Sigma^{-1}(\underset{\sim}{X}_i - \underset{\sim}{T})$$

and occasionally

$$\text{iii) } S_p^2 = \frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \bar{\underset{\sim}{X}})'\Sigma^{-1}(\underset{\sim}{X}_i - \bar{\underset{\sim}{X}}).$$

The above algorithms result in scalar measures of proximity to the target, variability around the target and variability around the subgroup mean. Once these have been found they can be plotted using similar strategies to those in the univariate case. It is again suggested that only $\tilde{T}_p$ and MSE_p be plotted for each subgroup as S_p^2 can cause the plot to be cluttered.

Inferences drawn here will be slightly different from the univariate case but identical to those inferences drawn from the multivariate control charts based on Hotelling's T^2 results [42]. See Alt [41] for more discussion.

6.6. DERIVATION OF THE CONTROL LIMITS AND CONSTANTS

The UCL for the multivariate measurements will be based on the assumption that the process measurements (i.e., p-dimensional vectors) are independent with a $MVN_p(\underset{\sim}{\mu}, \Sigma)$ distribution.

UCL for $\tilde{T}_p$

Assuming $\underset{\sim}{X} \sim MVN_p(\underset{\sim}{\mu}, \Sigma)$ it follows that $(\bar{\underset{\sim}{X}} - \underset{\sim}{T})' \Sigma^{-1} (\bar{\underset{\sim}{X}} - \underset{\sim}{T}) \sim \frac{1}{n}\chi^2_{p,\lambda}$ (see Theorem 6.7.5) with $\lambda = n(\underset{\sim}{\mu} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\mu} - \underset{\sim}{T})$ and p the number of variables used to monitor the process. Defining the UCL to be the $(1-\alpha)100\%$ percentile of the distribution function associated with $\tilde{T}_p$, the UCL will be of the form

$$UCL_{\tilde{T}_p} = \frac{1}{n}\chi^2_{p,\lambda,(1-\alpha)}$$

$$= D \quad .$$

Table 6.6.1 contains the values of D, denoted D_3 when $(1-\alpha)100\% = 99.73\%$, for n=2(1)6, p=2, 3, 4 and values of $(\underset{\sim}{\mu} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\mu} - \underset{\sim}{T}) = 0.00(0.05)0.90$.

UCL for MSE_p

Assuming the process measurements follow a $MVN_p(\underset{\sim}{\mu}, \Sigma)$, the $MSE_p \sim \frac{1}{n}\chi^2_{np,\lambda}$ (see Theorem 6.7.6) where $\lambda = n(\underset{\sim}{\mu} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\mu} - \underset{\sim}{T})$. The UCL, being the $(1-\alpha)100\%$ percentile of the distribution function, will be

$$UCL_{MSE_p} = \frac{1}{n}\chi^2_{np,\lambda,(1-\alpha)}$$

$$= E \quad .$$

Table 6.6.2 contains the value of E, denoted E_3 when $(1-\alpha)100\%=99.73\%$, for n=2(1)6, p=2, 3, 4 and $(\underset{\sim}{\mu} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\mu} - \underset{\sim}{T}) = 0.00(0.05)0.90$.

UCL for S_p^2

Assuming $MVN_p(\underset{\sim}{\mu}, \Sigma)$ for the process measurements, $\frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \bar{\underset{\sim}{X}})' \Sigma^{-1} (\underset{\sim}{X}_i - \bar{\underset{\sim}{X}})$ will follow a $\frac{1}{n}\chi^2_{(n-1)p,0}$ (see Theorem 6.7.7). As a result the UCL associated with the $(1-\alpha)100\%$ percentile of the distribution function will be

$$UCL_{S_p^2} = \frac{1}{n}\chi^2_{(n-1)p,0,\,(1-\alpha)}$$

$$= F \quad .$$

Table 6.6.3 contains the value of F, denoted F_3 and F_2 for the case where $(1-\alpha)100\%=99.73\%$ and 95.44%, n=2(1)6 and p=2, 3, 4.

The population parameters $\underset{\sim}{\mu}$ and Σ will seldomly be known, analogous to the univariate case we suggest substituting $\bar{\bar{\underset{\sim}{X}}}$ and $\bar{\underset{\sim}{S}}^2$ respectively, where $\bar{\underset{\sim}{S}}^2 = \frac{1}{k}\sum_{j=1}^{k}\underset{\sim}{S}_j^2$ with $\underset{\sim}{S}_j^2 = \frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \underset{\sim}{T})(\underset{\sim}{X}_i - \underset{\sim}{T})'$ for j= 1, 2, ... , k.

Table 6.6.1. Coefficient D_3 used to determine the boundary value associated with $\tilde{T}_p$ for various values of n, p and λ.

$$\lambda = (\underset{\sim}{\mu} - \underset{\sim}{T})' \, \Sigma^{-1} \, (\underset{\sim}{\mu} - \underset{\sim}{T})$$

p	n	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90
2	2	5.91	6.20	6.46	6.70	6.92	7.14	7.34	7.54	7.73	7.91	8.09	8.26	8.42	8.59	8.75	8.90	9.06	9.21	9.35
	3	3.94	4.22	4.47	4.69	4.90	5.09	5.27	5.45	5.62	5.78	5.93	6.09	6.24	6.38	6.52	6.66	6.80	6.93	7.06
	4	2.96	3.23	3.46	3.67	3.86	4.04	4.21	4.37	4.53	4.68	4.82	4.96	5.10	5.23	5.36	5.49	5.61	5.74	5.86
	5	2.37	2.63	2.86	3.05	3.23	3.40	3.56	3.71	3.86	4.00	4.13	4.26	4.39	4.52	4.64	4.76	4.88	4.99	5.11
	6	1.97	2.23	2.45	2.64	2.81	2.97	3.12	3.26	3.40	3.53	3.66	3.78	3.91	4.03	4.14	4.26	4.37	4.48	4.59
3	2	7.08	7.31	7.52	7.73	7.93	8.12	8.30	8.48	8.65	8.82	8.99	9.15	9.30	9.46	9.61	9.76	9.90	10.04	10.19
	3	4.72	4.94	5.15	5.35	5.54	5.71	5.88	6.04	6.20	6.35	6.50	6.65	6.79	6.93	7.07	7.20	7.33	7.46	7.59
	4	3.54	3.76	3.96	4.15	4.33	4.49	4.65	4.80	4.95	5.09	5.23	5.37	5.50	5.63	5.75	5.88	6.00	6.12	6.24
	5	2.83	3.05	3.25	3.43	3.59	3.75	3.90	4.05	4.19	4.32	4.45	4.58	4.70	4.83	4.94	5.06	5.18	5.29	5.40
	6	2.36	2.58	2.77	2.94	3.10	3.25	3.40	3.53	3.67	3.79	3.92	4.04	4.16	4.28	4.39	4.50	4.61	4.72	4.83
4	2	8.13	8.32	8.52	8.70	8.88	9.05	9.22	9.39	9.55	9.71	9.86	10.01	10.16	10.31	10.45	10.59	10.73	10.87	11.00
	3	5.42	5.61	5.80	5.98	6.15	6.31	6.47	6.62	6.77	6.92	7.06	7.20	7.34	7.47	7.60	7.73	7.86	7.98	8.11
	4	4.06	4.26	4.44	4.61	4.77	4.93	5.08	5.22	5.37	5.50	5.64	5.77	5.89	6.02	6.14	6.26	6.38	6.50	6.62
	5	3.25	3.44	3.62	3.79	3.94	4.09	4.24	4.37	4.51	4.64	4.77	4.89	5.01	5.13	5.25	5.36	5.48	5.59	5.70
	6	2.71	2.90	3.07	3.24	3.39	3.53	3.67	3.80	3.93	4.05	4.18	4.30	4.41	4.53	4.64	4.75	4.86	4.96	5.07

Table 6.6.2. Coefficient E_3 used to determine the boundary value associated with MSE_p for various values of n, p and λ.

$$\lambda=(\underset{\sim}{\mu} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\mu} - \underset{\sim}{T})$$

p	n	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90
2	2	8.13	8.32	8.52	8.70	8.88	9.05	9.22	9.39	9.55	9.71	9.86	10.01	10.16	10.31	10.45	10.59	10.73	10.87	11.00
	3	6.69	6.85	7.01	7.16	7.31	7.46	7.60	7.74	7.88	8.01	8.14	8.27	8.40	8.53	8.65	8.77	8.89	9.01	9.13
	4	5.89	6.04	6.18	6.32	6.45	6.58	6.71	6.83	6.96	7.08	7.20	7.31	7.43	7.54	7.65	7.77	7.88	7.99	8.09
	5	5.38	5.51	5.64	5.77	5.89	6.01	6.13	6.24	6.36	6.47	6.58	6.69	6.80	6.90	7.01	7.11	7.22	7.32	7.42
	6	5.02	5.14	5.26	5.38	5.49	5.61	5.72	5.83	5.93	6.04	6.14	6.25	6.35	6.45	6.55	6.65	6.75	6.85	6.94
3	2	10.03	10.20	10.36	10.52	10.67	10.82	10.97	11.12	11.26	11.40	11.54	11.68	11.82	11.95	12.08	12.21	12.34	12.47	12.60
	3	8.42	8.56	8.69	8.83	8.96	9.09	9.21	9.34	9.46	9.59	9.71	9.82	9.94	10.06	10.17	10.29	10.40	10.51	10.62
	4	7.52	7.65	7.77	7.89	8.01	8.13	8.24	8.35	8.47	8.58	8.69	8.79	8.90	9.01	9.11	9.22	9.32	9.42	9.53
	5	6.94	7.06	7.17	7.28	7.39	7.50	7.61	7.71	7.82	7.92	8.02	8.12	8.22	8.32	8.42	8.52	8.62	8.71	8.81
	6	6.53	6.64	6.74	6.85	6.95	7.05	7.16	7.26	7.35	7.45	7.55	7.65	7.74	7.84	7.93	8.02	8.12	8.21	8.30
4	2	11.79	11.93	12.08	12.22	12.36	12.50	12.63	12.77	12.90	13.03	13.16	13.29	13.41	13.54	13.66	13.79	13.91	14.03	14.15
	3	10.03	10.16	10.28	10.40	10.52	10.64	10.76	10.87	10.99	11.10	11.21	11.32	11.43	11.54	11.65	11.76	11.87	11.97	12.08
	4	9.05	9.17	9.28	9.39	9.50	9.60	9.71	9.82	9.92	10.02	10.13	10.23	10.33	10.43	10.53	10.63	10.73	10.83	10.92
	5	8.42	8.52	8.62	8.73	8.83	8.93	9.03	9.13	9.22	9.32	9.42	9.51	9.61	9.70	9.80	9.89	9.98	10.08	10.17
	6	7.96	8.06	8.16	8.25	8.35	8.45	8.54	8.63	8.73	8.82	8.91	9.00	9.09	9.18	9.27	9.36	9.45	9.54	9.63

Table 6.6.3. Coefficients F_2 and F_3 for the boundary value associated with S_p^2 for various values of n and p .

p	n	F_3	F_2
2	2	5.91	3.09
	3	5.42	3.24
	4	5.02	3.21
	5	4.71	3.16
	6	4.48	3.10
3	2	7.08	4.01
	3	6.69	4.28
	4	6.31	4.30
	5	6.02	4.27
	6	5.79	4.22
4	2	8.13	4.86
	3	7.86	5.26
	4	7.52	5.34
	5	7.24	5.33
	6	7.01	5.30

6.7. PROPERTIES

Theorem 6.7.1: $(\bar{x} - T)^2 \sim \frac{\sigma^2}{n} \chi^2_{1,\lambda}$.

Proof: If $X \sim N(\mu, \sigma^2)$ then $\frac{(\bar{x} - T)}{\frac{\sigma}{\sqrt{n}}} - \frac{(\mu - T)}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$, from page 130 of [31]

$$\left[\frac{(\bar{x} - T)}{\frac{\sigma}{\sqrt{n}}} - \frac{(\mu - T)}{\frac{\sigma}{\sqrt{n}}} + \frac{(\mu - T)}{\frac{\sigma}{\sqrt{n}}} \right]^2 \sim \chi^2_{1,\lambda} \text{ with } \lambda = n\left[\frac{\mu - T}{\sigma} \right]^2 \text{ hence}$$

$$(\bar{x} - T)^2 \sim \frac{\sigma^2}{n} \chi^2_{1,\lambda} \quad \text{with } \lambda = n\left[\frac{\mu - T}{\sigma} \right]^2.$$

Theorem 6.7.2: $MSE \sim \frac{\sigma^2}{n} \chi^2_{n,\lambda}$ with $\lambda = n\left[\frac{\mu - T}{\sigma} \right]^2$

Proof: If $X \sim N(\mu, \sigma^2)$ then $\frac{(x_i - T)}{\sigma} - \frac{(\mu - T)}{\sigma} \sim N(0, 1)$ for all $i = 1, 2, \ldots, n$

If the x_i's are independent then, from page 130 of [31] ,

$$\sum_{i=1}^{n} \left[\frac{(x_i - T)}{\sigma} - \frac{(\mu - T)}{\sigma} + \frac{(\mu - T)}{\sigma} \right]^2 \sim \chi^2_{n,\lambda} \text{ , as a result}$$

$$\sum_{i=1}^{n} \frac{(x_i - T)^2}{n} \sim \frac{\sigma^2}{n} \chi^2_{n,\lambda} \text{ with } \lambda = n\left[\frac{\mu - T}{\sigma} \right]^2.$$

Theorem 6.7.3: $s^2 \sim \frac{\sigma^2}{n-1}\chi^2_{n-1}$.

Proof: If $X \sim N(\mu, \sigma^2)$ then from page 135 of [31] $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

As a result $s^2 \sim \frac{\sigma^2}{n-1}\chi^2_{n-1}$.

Theorem 6.7.4: $\frac{(\bar{\bar{X}} - T)^2}{\overline{S^2}} \sim \frac{1}{nk} F_{1,\, k(n-1),\, \lambda}$ where n is the subgroup sample size, k the number of subgroups and $F_{1,\, k(n-1),\, \lambda}$ the non-central F distribution with 1 and k(n-1) df and non-centrality parameter $\lambda = nk\left[\frac{\mu - T}{\sigma}\right]^2$.

Proof: If $X \sim N(\mu, \sigma^2)$ analogous to the proof for Theorem 6.7.1

$$(\bar{\bar{X}} - T)^2 \sim \frac{\sigma^2}{nk}\chi^2_{1,\,\lambda} \text{ with } \lambda = nk\left[\frac{\mu - T}{\sigma}\right]^2.$$

If the k subgroups are assumed to be independent and since $s_j^2 \sim \frac{\sigma^2}{n-1}\chi^2_{n-1}$ for all j

$$\overline{S^2} = \frac{1}{k}\sum_{j=1}^{k} s_j^2 \sim \frac{\sigma^2}{k(n-1)}\chi^2_{k(n-1)}.$$

Then from page 189 of [31] $\frac{(\bar{\bar{X}} - T)^2}{\overline{S^2}} \sim \frac{1}{nk} F_{1,\, k(n-1),\, \lambda}$.

From page 190 of [31] the expected value of $\frac{(\bar{\bar{X}} - T)^2}{\overline{S^2}}$ is $\frac{1}{1 - \frac{2}{k(n-1)}}\left\{\frac{1}{nk} + \left[\frac{\mu - T}{s}\right]^2\right\}$ which for a moderate number of subgroups is approximately equal to $\left[\frac{\mu - T}{\sigma}\right]^2$. This

result suggests that substituting $\bar{\bar{X}}$ and $\bar{S^2}$ for μ and σ^2 respectively should provide a reasonable estimate of $\left[\frac{\mu - T}{\sigma}\right]^2$.

Theorem 6.7.5: $(\underset{\sim}{\bar{X}} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\bar{X}} - \underset{\sim}{T}) \sim \frac{1}{n}\chi^2_{p,\lambda}$ with $\lambda = n (\underset{\sim}{\mu} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\mu} - \underset{\sim}{T})$.

Proof: If $\underset{\sim}{X} \sim MVN_p(\underset{\sim}{\mu}, \Sigma)$ then $\underset{\sim}{\bar{X}} \sim MVN_p(\underset{\sim}{\mu}, \frac{1}{n}\Sigma)$

From page 113 of [43] $n(\underset{\sim}{\bar{X}} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\bar{X}} - \underset{\sim}{T}) \sim \chi^2_{p,\lambda}$ with $\lambda = n (\underset{\sim}{\mu} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\mu} - \underset{\sim}{T})$.

Theorem 6.7.6: $MSE_p = \frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \underset{\sim}{T})'\Sigma^{-1}(\underset{\sim}{X}_i - \underset{\sim}{T}) \sim \frac{1}{n}\chi^2_{np,\lambda}$.

Proof: If $\underset{\sim}{X} \sim MVN_p(\underset{\sim}{\mu}, \Sigma)$ then from Theorem 6.7.5, $(\underset{\sim}{X} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{X} - \underset{\sim}{T}) \sim \chi^2_{p,\lambda}$ with $\lambda = (\underset{\sim}{\mu} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\mu} - \underset{\sim}{T})$. Assuming the $\underset{\sim}{X}$'s to be independent it follows that

$$\sum_{i=1}^{n}(\underset{\sim}{X}_i - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{X}_i - \underset{\sim}{T}) \sim \chi^2_{np,\lambda} \text{ where } \lambda = n (\underset{\sim}{\mu} - \underset{\sim}{T})' \Sigma^{-1} (\underset{\sim}{\mu} - \underset{\sim}{T}). \text{ Therefore}$$

$$\frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \underset{\sim}{T})'\Sigma^{-1}(\underset{\sim}{X}_i - \underset{\sim}{T}) \sim \frac{1}{n}\chi^2_{np,\lambda}.$$

Theorem 6.7.7: $\frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \underset{\sim}{\bar{X}})' \Sigma^{-1} (\underset{\sim}{X}_i - \underset{\sim}{\bar{X}}) \sim \frac{1}{n}\chi^2_{(n-1)p}$

Proof: If $\underset{\sim}{X} \sim MVN_p(\underset{\sim}{\mu}, \Sigma)$ then $(\underset{\sim}{X}_i - \underset{\sim}{\bar{\bar{X}}}) \sim MVN_p(\underset{\sim}{0}, \frac{n-1}{n}\Sigma)$

Then it follows that $\frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \underset{\sim}{\bar{X}})' \Sigma^{-1} (\underset{\sim}{X}_i - \underset{\sim}{\bar{X}}) \sim \frac{1}{n}\chi^2_{(n-1)p}$.

6.8. EXAMPLES

Example 6.8.1: The example presented is taken from question 6.10, page 144 of [25]. An arbitrary target value of 12 is assumed resulting in the following

subgroup #	X_1	X_2	X_3	X_4	$\bar{x}$	$(\bar{x} - T)^2$	MSE	s^2
1	5	12	10	3	7.5	20.250	44.67	17.667
2	1	0	2	18	5.25	45.563	133.67	72.917
3	1	8	3	16	7	25.000	78.00	44.667
4	7	14	18	1	10	4.000	62.00	56.667
5	3	19	5	8	8.75	10.563	65.00	50.917
6	18	16	14	0	12	0.000	66.67	66.667
7	9	0	3	5	4.25	60.063	94.33	14.250
8	14	11	0	4	7.25	22.563	71.00	40.917
9	20	12	23	13	17	25.000	62.00	28.667
10	16	8	18	22	16	16.000	56.00	34.667
11	2	0	14	17	8.25	14.063	91.00	72.250
12	5	1	2	6	3.5	72.250	102.00	5.667
13	5	20	6	2	8.25	14.063	83.00	64.250
14	9	16	12	8	11.25	0.563	13.67	12.917
15	3	35	15	0	13.25	1.563	54.33	252.250
16	11	11	3	4	7.25	22.563	49.00	18.917
17	19	4	9	21	13.25	1.563	67.67	65.583
18	17	14	4	19	13.5	2.250	47.33	44.333
19	5	6	22	11	11	1.000	62.00	60.667
20	4	3	19	12	9.5	6.250	64.67	56.333
21	8	18	0	7	8.25	14.063	73.67	54.917
22	20	21	5	14	15	9.000	66.00	54.000
23	16	0	6	11	8.25	14.063	65.67	46.917
24	13	8	11	8	10	4.000	11.33	6.000
25	11	5	25	12	13.25	1.563	73.00	70.917
26	1	3	1	7	3	81.000	116.00	8.000
27	4	12	4	13	8.25	14.063	43.00	24.250
28	11	17	12	10	12.5	0.250	10.00	9.667

For T=12 and with $\bar{\bar{X}} = \sum_{j=1}^{k} \frac{\bar{x}_j}{k} = 9.73$ and $\overline{S^2} = \sum_{j=1}^{k} \frac{s_j^2}{k} = 48.42$, the value of $\left[\frac{\mu - T}{\sigma}\right]^2 \cong \frac{[9.73 - 12]^2}{48.42} = 0.1064$. Rounding off to the nearest 0.05 for computational purposes, finds $\left[\frac{\mu - T}{\sigma}\right]^2 = 0.10$. The UCL for the measures of interest are then

$$\text{i) } A_3 \sigma^2 = 2.92 \sigma^2 \cong 2.92 \overline{S^2} = 141.4,$$

$$\text{ii) } C_3 \sigma^2 = 5.92 \sigma^2 \cong 5.92 \overline{S^2} = 286.6,$$

$$\text{iii) } B_3 \sigma^2 = 4.72 \sigma^2 \cong 4.72 \overline{S^2} = 228.5.$$

Plotting $(\bar{x} - T)^2$, and the UCL's for $(\bar{x} - T)^2$, s^2 and MSE results in the plot depicted in Figure 6.8.2.

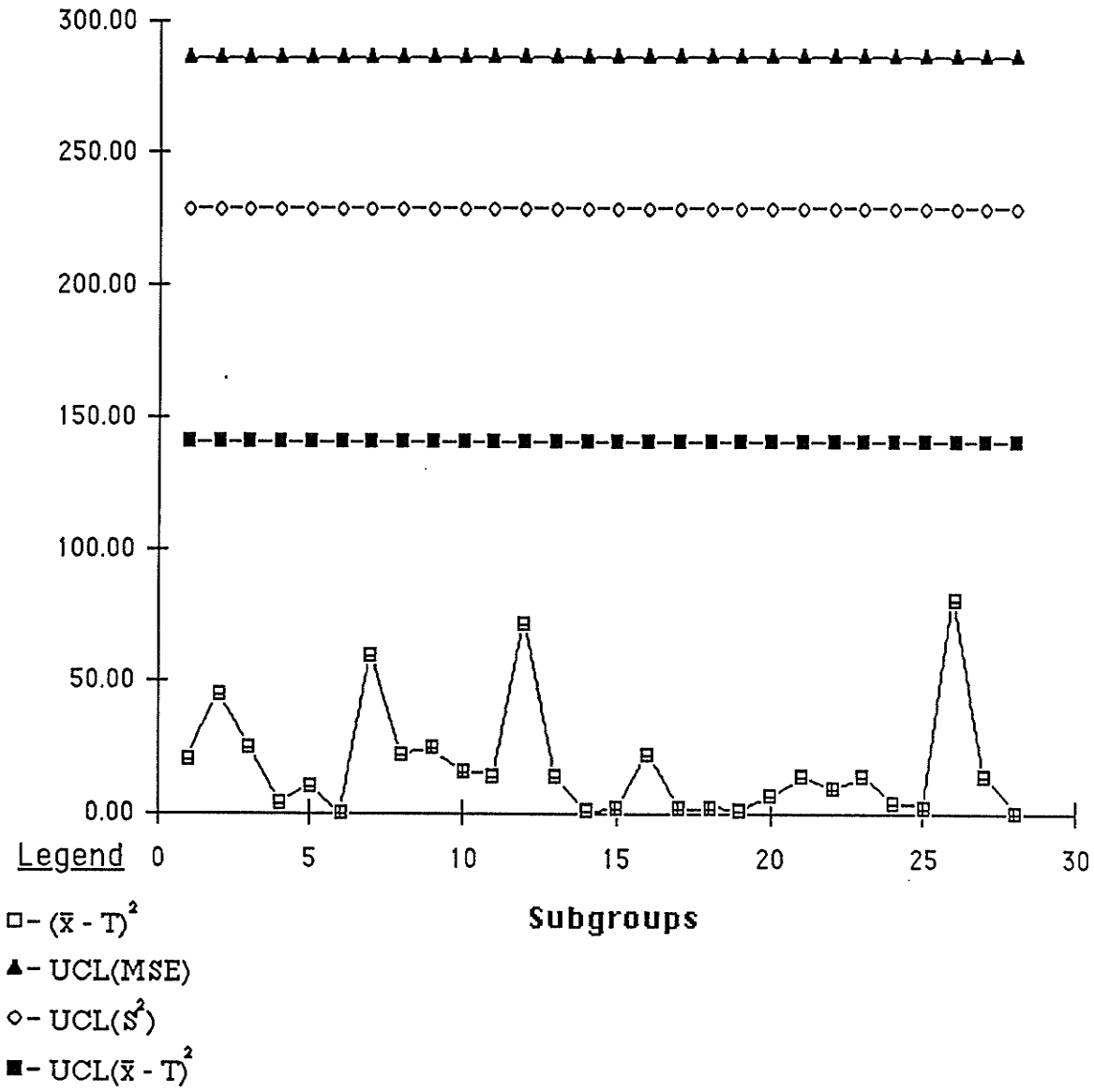


Figure 6.8.2. The Plot of $(\bar{x} - T)^2$ and the UCL's for $(\bar{x} - T)^2$, s^2 and MSE.

From Figure 6.8.2 it is apparent that none of the subgroups exceed the UCL for $(\bar{x} - T)^2$ and generally most of them appear to be relatively close to 0 indicating reasonable proximity to the target. Subgroups 7, 12 and 26 appear to have the largest departures from the target but are well within the control limits. Nineteen of the subgroups have sampling means below the target as denoted by the negative signs. The longest

sequence of similar signs is five (subgroups 1 to 5), indicating that the longest run of subgroup means either above or below (in this case below) the target is five. There does not appear to be any significant drifting from the target nor does there appear to be any cyclical relationships occurring. However before any formal inferences can be drawn the MSE for each subgroup must be plotted (see Figure 6.8.3).

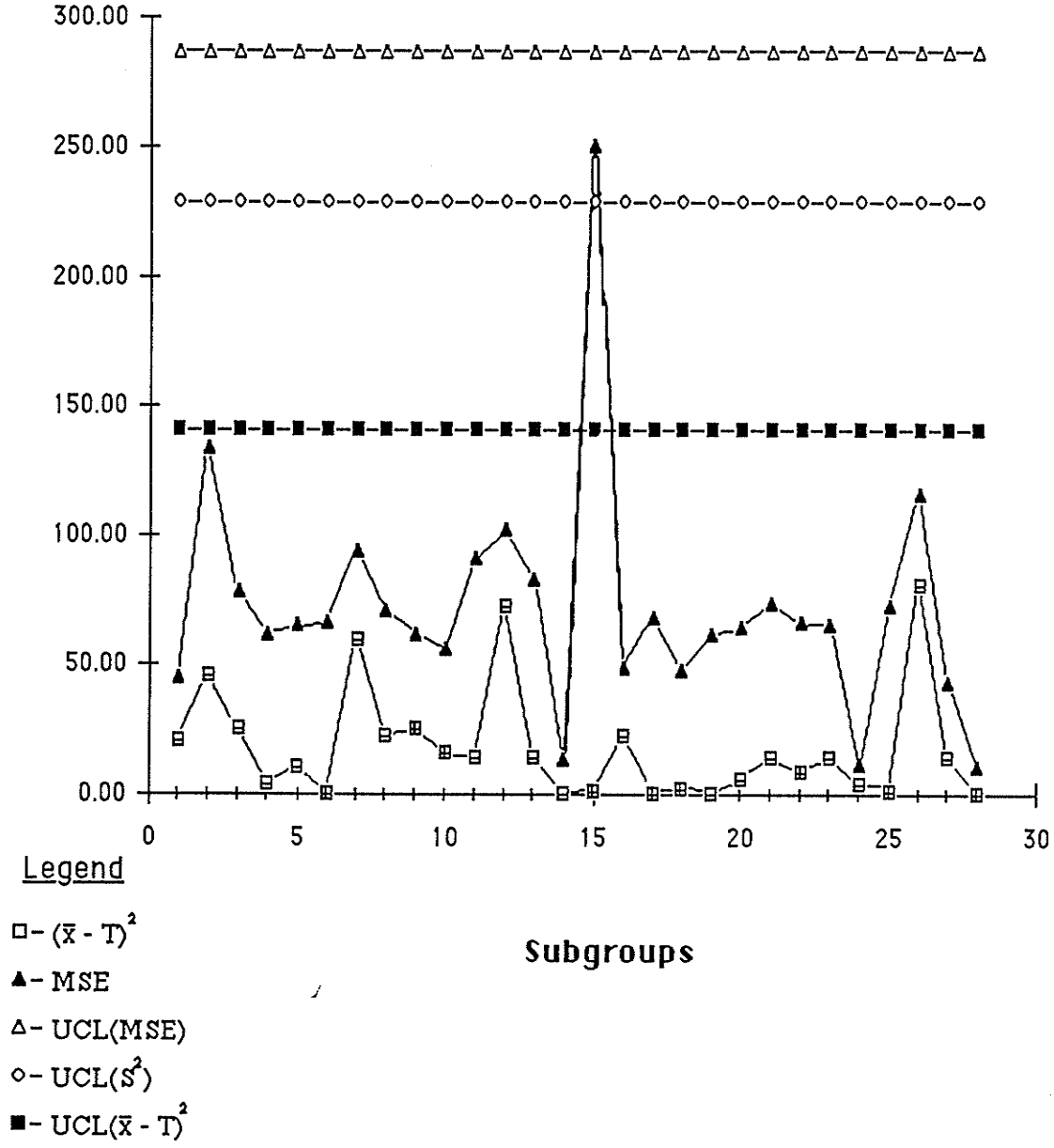


Figure 6.8.3. The Plot of $(\bar{x} - T)^2$, MSE and the UCLs.

From Figure 6.8.3, none of the subgroups exceed the UCL for MSE, but subgroup 15 has a MSE that exceeds the UCL for s^2. Closer investigation finds $(\bar{x} - T)^2$ to be quite small for subgroup 15 resulting in a large difference between $(\bar{x} - T)^2$ and MSE. Clearly this difference will exceed the UCL for s^2 however s^2_{15} should be included in the plot (see Figure 6.8.4). Analysis of the rest of the plot follows in a predictable fashion. Subgroups 14, 24 and 28 may require investigation as they appear to have small variability while being quite close to the target.

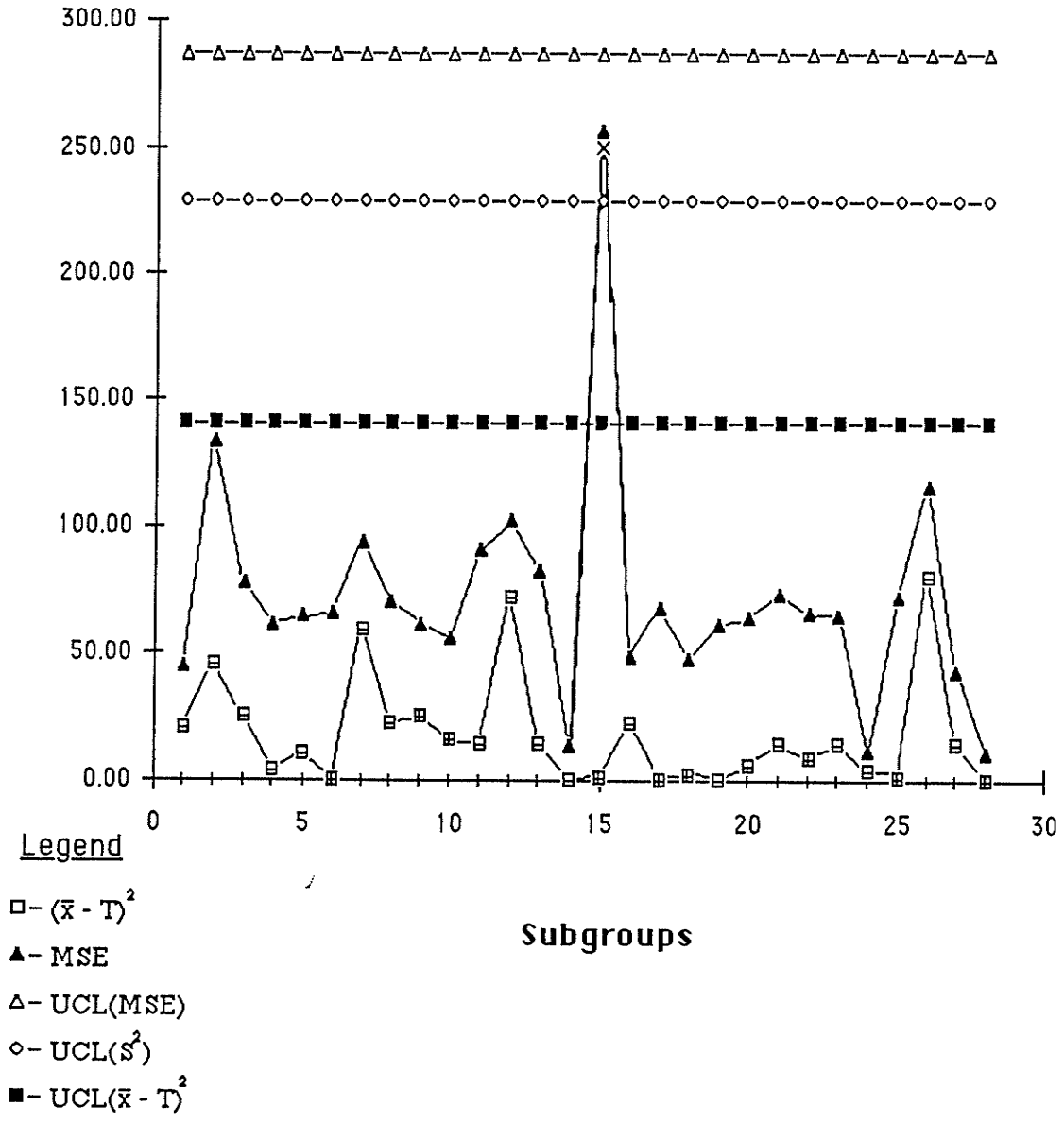


Figure 6.8.4. The Plot of $(\bar{x} - T)^2$, MSE and one extreme value of s^2.

Example 6.8.2: The data as taken from [34] consists of 30 bivariate observations taken from a steel manufacturing process. In order to illustrate the multivariate control chart procedure the 30 observations were taken in groups of size five with the groupings formed using the sequential sample numbers. An arbitrary target of (175, 55) was assumed with the following results

Subgroup	Sample Number	X=(x, y)	$\tilde{T}_p$	MSE
1	1	143, 34.2		
	2	200, 57.0		
	3	160, 47.5	1.12	2.92
	4	181, 53.4		
	5	148, 47.8		
2	6	178, 51.5		
	7	162, 45.9		
	8	215, 59.1	0.54	2.18
	9	161, 48.4		
	10	141, 47.3		
3	11	175, 57.3		
	12	187, 58.5		
	13	187, 58.2	0.10	0.43
	14	186, 57.0		
	15	172, 49.4		
4	16	182, 57.2		
	17	177, 50.6		
	18	204, 55.1	0.79	1.30
	19	178, 50.9		
	20	196, 57.9		
5	21	160, 45.5		
	22	183, 53.9		
	23	179, 51.2	0.42	0.86
	24	194, 57.5		
	25	181, 55.6		
6	26	195, 58.0		
	27	134, 47.5		
	28	187, 42.0	0.72	4.32
	29	135, 40.5		
	30	159, 58.0		

From Tables 6.6.1,6.6.2 and 6.6.3, $\underset{\sim}{T}_{2,\ 5,\ 0.35} = 3.71$, $MSE_{2,\ 5,\ 0.35} = 6.24$ and $S^2_{2,\ 5,\ 0.35} = 4.71$. The control chart associated with this set of data (Figure 6.8.5) indicates that none of the subgroups have centers of mass dramatically different from the

target ($\underset{\sim}{T}$) and although there appears to be a large jump in MSE_p in subgroup 6 all have MSE_p and S_p^2 values well within the upper control limits. These results tend to suggest that the process is in-control. If additional information regarding individual observations is required this could be included in Figure 6.8.5. For example a closer examination of the observations in Subgroup 6 may be warranted here.

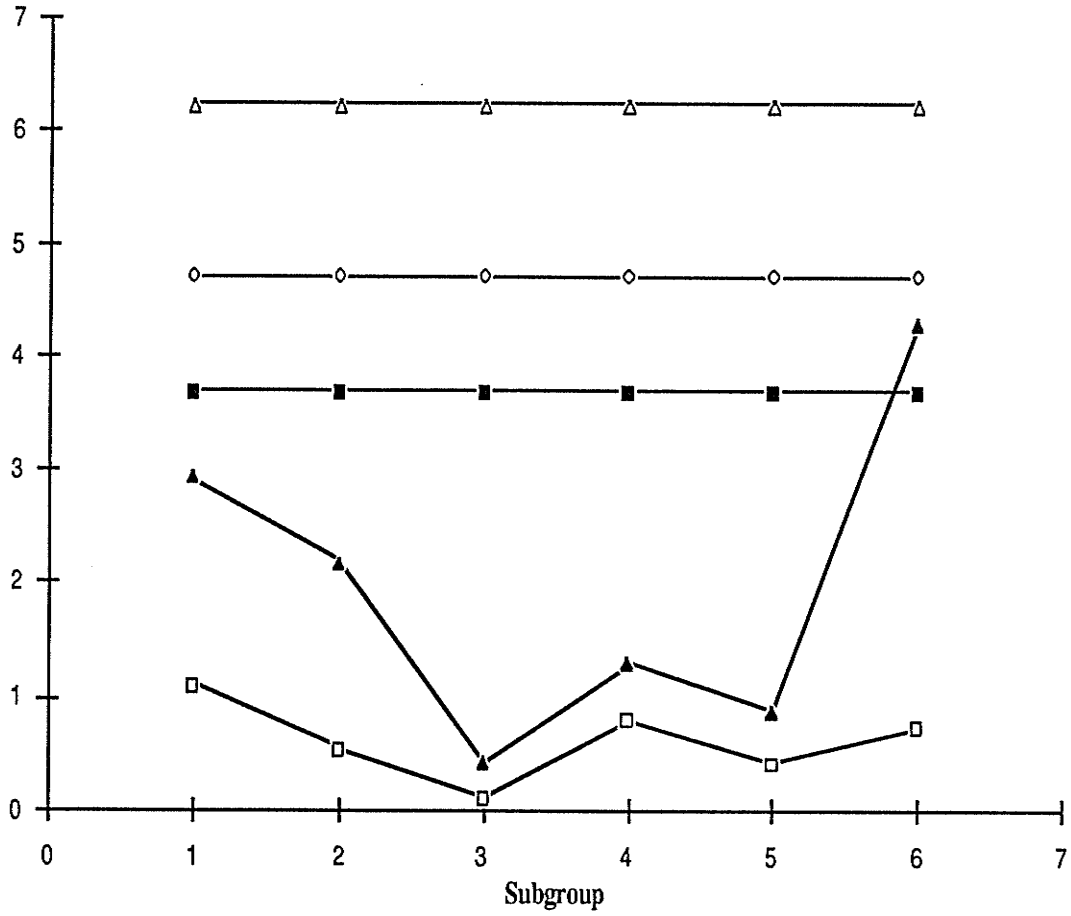


Figure 6.8.5. Control Chart for Example 6.8.2.

6.9. COMMENTS

A simultaneous control chart has been presented that mirrors the recent changes occurring in Quality Control circles. The proposed chart is analogous to the traditional $\bar{x}$ and s chart providing much of the information available from the traditional control charts while incorporating additional pertinent information. The additional information has been incorporated in a single plot without loss of clarity.

The calculations required for the univariate case are straightforward requiring only a hand-held calculator. The procedure is easily adapted to computer analysis and graphics.

The theory used to derive the results, in our opinion is clearer than that associated with the traditional charts as all UCLs represent a specific percentile of the distribution associated with the measurement.

Many of the properties derived and examined for the traditional control charts apply to the new procedure as the measures used are very similar in nature. Replacing the target value in the new procedure with $\bar{\bar{X}}$ results in a plot that provides all the information included in the traditional control chart. However in current philosophy much emphasis is placed on proximity to the target value when monitoring and assessing a process.

The simultaneous nature of the proposed procedure has been achieved without sacrificing clarity. The resultant control chart appears to purvey more information than is available in traditional charts while not being dramatically different from the traditional charts in their motivation or inference. Although the new procedure does not provide all the information that may be contained in a boxplot style chart it does have some features such as explicit boundary values for s^2 and MSE that are not available in a boxplot.

The proposed multivariate procedure is similar to the traditional Hotelling T^2 style of chart. The proposed plot allows investigation of both proximity to the target and overall variability where Hotelling's procedure confounds these measures.

REFERENCES

[1] Shapiro, S., and M. Wilk, "An Analysis of Variance Test for Normality (Complete Samples)", Biometrika, vol 52, 1965, pp 591-611.

[2] Shapiro, S., and M. Wilk, "An Analysis of Variance Test for the Exponential Distribution (Complete Samples)", Technometrics, vol 14, 1972, pp 355-370.

[3] Mood, A.M., F.A. Graybill and D.C. Boes, Introduction to the Theory of Statistics, 3rd Edition, McGraw-Hill Inc, 1974.

[4] Searle, S.R., Linear Models, John Wiley and Co., 1971.

[5] Lloyd, E.H., "Least Squares Estimation of Location and Scale Parameters Using Order Statistics", Biometrika, vol 39, 1952, pp 88-95.

[6] Kendall, M.G., and A. Stuart, The Advanced Theory of Statistics, Volume 2, Inference and Relationship, Hafner Pub. Co., 1960.

[7] Bol'shev, L.N., "On the question of distinguishing normal and uniform distributions from small samples", SIAM Theory of Probability and Applications, vol 10, 1965, pp 694-696.

[8] Khakhubiya, TS.G., "On the effectiveness of distinguishing between normal and uniform distributions on the basis of small samples", SIAM Theory of Probability and Applications, vol 11, 1966, pp 166-172.

[9] David, F.N., and N.L. Johnson, "The Probability Integral Transformation When Parameters are Estimated From the Sample", Biometrika, vol 35, 1948, pp 182-192.

[10] O'Reilly, F., C.P. Quesenberry, "The Conditional Probability Integral Transformation and Applications to obtain Composite Chi-square Goodness-of-Fit Tests", The Annals of Statistics, vol 1, 1973, pp 74 - 83.

[11] Rosenblatt, M., "Remarks on a Multivariate Transformation", Annals of Mathematical Statistics, vol 23, 1952, pp 470-472.

[12] Quesenberry, C.P., "Some Transformation Methods in Goodness-of-Fit", in Goodness-of-Fit Techniques, edited by R.B. D'Agostino, M.A. Stephens, 1986, pp 235-277.

[13] Rincon-Gallardo, S., Quesenberry, C.P. and O'Reilly, F.J., "Conditional Probability Integral Transformations and Goodness-Of-Fit Tests for Multivariate Normal Distributions", The Annals of Statistics, vol 7(5), 1979, pp 1052-1057.

[14] Healy, M.J.R., "Multivariate Normal Plotting", J. Royal Statistical Society (C), vol 17, 1968, pp 157-161.

[15] Mahalanobis, P.C., "On the Generalized Distance in Statistics", Proceedings of the National Institute of Sciences of India, vol 12, 1936, pp 49-55.

[16] Sinha, S.K., and B.K. Kale, Life Testing and Reliability Estimation, John Wiley and Sons, 1980.

[17] D'Agostino, R.B., and M.A. Stephens, Goodness-of-Fit Techniques, Marcel Dekker, 1986.

[18] Abramowitz, M., and Stegun, I.A., Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover Publication Incorporated, New York, 1970, pp 257.

[19] Tukey, J.. A Survey of Sampling From Contaminated Distributions, in Contributions to Probability and Statistics, eds. I. Olkin, S. Ghurye, W. Hoeffding,W. Madow, and H. Mann. Stanford University Press, Stanford, California, 1960.

[20] Lax, D.A., Robust Estimators of Scale: Finite-Sample Performance in the Long-Tailed Symmetric Distributions, Journal of the American Statistical Association, 80, 1985, pp 353-369.

[21] Gayen, A.K., The distribution of 'Student's' t in random samples of any size drawn from non-normal universes, Biometrika, 36, 1949, pp 353-369.

[22] Barton, D.E., and Dennis, K.E., The conditions under which Gram-Charlier and Edgeworth Curves are positive definite and unimodal, Biometrika, 39, 1951, pp 425-427.

[23] Kane, V.E., Process Capability Indices, Journal of Quality Technology, 18, 1986, pp 41-52.

[24] Kane, V. E., "Corrigenda", Journal of Quality Technology, 18, 4, 1986, pp 265.

[25] Braverman, J.D., Fundamentals of Statistical Quality Control. Prentice-Hall Co., Virginia, 1981.

[26] Wies, Jr., H. M., and Burr, I. W., "Simple Capability Analysis Acceptance Test", Industrial Quality Control, 21, 1964, pp 266-268.

[27] Winkler, R. L., Introduction to Bayesian Inference and Decision, Holt, Rinehart and Winston, Inc., Toronto, 1972 .

[28] Wilson, E. B., and Hilferty, M. M., "The Distribution of Chi-Square", Proceedings of the National Academy of Sciences, 17, 1931, pp 684-688.

[29] Sullivan, L. P., "Targeting Variability - A New Approach to Quality", Quality Progress, July, 1984, pp 15-21.

[30] Sullivan, L. P., "Letters", Quality Progress, April 1985, pp 7-8.

[31] Johnson, N. L., and Kotz, S., Continuous Univariate Distributions-2, John Wiley and Sons, New York, NY, 1970.

[32] Kimball, B. F., "On the choice of plotting positions in probability paper", Journal of the American Statistical Association, 55, 1960, pp 546-560.

[33] Johnson, N.L., and Kotz, S., Distributions in Statistics: Continuous Multivariate Distributions, John Wiley and Sons, Inc., New York, New York, 1972.

[34] Sultan, T.I., "An Acceptance Chart for Raw Materials of Two Correlated Properties", Quality Assurance, 12, 3, 1986, pp 70-72.

[35] Shakun, M.F., "Multivariate Acceptance Sampling Procedures for General Specification Ellipsoids", Journal of the American Statistical Association, 60, ,1986, pp 905-13.

[36] Jackson, J.E., and Bradley, R.A., "Sequential χ^2 and T^2 Tests and Their Application to an Acceptance Sampling Problem", Technometrics, 3, 1961, pp 519-34.

[37] White, E.M., and Şchroeder, R., "A Simultaneous Control Chart", Journal of Quality Technology, 19(1), 1987, pp 1-10.

[38] Iglewicz, B., and Hoaglin, D.C., "Use of Boxplots for Process Evaluation", Journal Of Quality Technology, 19(4), 1987, pp 180-190.

[39] Shewhart, W.A., Economic Control of Quality of Manufactured Product, Van Nostrand Company, Inc., 1931.

[40] Woodall, W.H., "The Statistical Design of Quality Control Charts", The Statistician, 34, 1985, pp 155-160.

[41] Alt, F.B., "Multivariate Quality Control", in Encyclopedia of Statistics, Vol. 6, edited by N.L. Johnson and S. Kotz, 1986, pp 110-122.

[42] Hotelling, H., "Multivariate Quality Control, Illustrated by the Air Testing of Sample Bombsights", in Techniques of Statistical Analysis, edited by C. Eisenhart, H. Hastay and W.A. Wallis, McGraw-Hill, New York, 1947, pp 111-184.

[43] Anderson, T.W., An Introduction to Multivariate Analysis, 2nd edition, John Wiley & Sons, New York, 1984.

BIOGRAPHICAL DATA

Frederick Alfred Spiring was born in Winnipeg, Manitoba on November 3, 1954. He received his B. Sc. (1981) and M. Sc.(1983) at the University of Manitoba, Winnipeg, Manitoba. He returned after two years as a Research Consultant with Health Protection Branch, Health and Welfare Canada to pursue his Ph.D. He is now an Assistant Professor in the Department of Statistical and Actuarial Sciences at The University of Western Ontario, London, Ontario where he is also Manager of the Ontario Quality Centre.

CONTRIBUTIONS TO RESEARCH

E. C. D. Todd, R. Szabo and F. Spiring. "Donairs(gyros) - potential hazards and control", Journal of Food Protection, 49(5), 1986.

U. Purvis, A. N. Sharpe, D. M. Bergener, G. Lachapelle, M. Milling and F. Spiring. "Recovery of Bacteria from Naturally Contaminated Foods by Stomacher", Canadian Journal Of Microbiology, 33(1), 1987.

S. W. Cheng and F. A. Spiring. "A Test to Identify the Uniform Distribution with Applications to Probability Plotting and Other Distributions", IEEE Transactions on Reliability, 36(1), 1987.

S. W. Cheng and F. A. Spiring. "Assessing Process Capability: A Bayesian Approach", Institute of Industrial Engineers Transactions, To appear 1988.

L. K. Chan, S. W. Cheng and F. A. Spiring. "A New Measure of Process Capability: Cpm", Journal of Quality Technology, July 1988.

L. K. Chan, S. W. Cheng and F. A. Spiring. "The Robustness of the Process Capability Index to Departures from Normality", Proceedings of the 2nd Pacific Area Statistical Conference (refereed), December 1986.

L. K. Chan, S. W. Cheng and F. A. Spiring. "Some Measures of Process Capability", Proceedings of the 46th Session of the International Statistical Institute, September 1987.

L. K. Chan, S. W. Cheng and F. A. Spiring. "A Graphical Technique for Process Capability", Proceedings of the ASQC 42nd Annual Quality Congress (refereed), May 1988.