1962

An analysis of methods used in estimating the CEP.

Hargrove, John Q.

Monterey, California: U.S. Naval Postgraduate School

http://hdl.handle.net/10945/11620

AN ANALYSIS OF METHODS USED IN ESTIMATING THE CEP

JOHN Q. HARGROVE

AN ANALYSIS OF METHODS

USED IN ESTIMATING THE CEP

by

John T. Hargrove

This work is accepted as fulfulling

the thesis requirements for the degree of

MASTER OF SCIENCE

from the

United States Naval Postgraduate School

ABSTRACT

This thesis presents a general discussion of the problems involved in estimating the Circular Probable Error, more commonly referred to as the CEP. A comparison is made between the estimates of the CEP under two distinct models. The models are identical except for the location of the mean vector in relation to the target. The assumption of dependence is made in both models and the resulting estimates are compared with the corresponding estimates obtained under the assumption of independence. Confidence interval estimates of the CEP are also presented. Two methods of removing outlier or "Maverick" observations are introduced and some of the possible effects on the estimated CEP are discussed. The different estimating procedures are illustrated with three numerical problems.

PREFACE

The term "CEP" is familiar to most Naval Officers, but the underlying assumptions upon which this measure of effectiveness rests are often misunderstood. Therefore, it is the objective of this thesis to explain as fully as possible what the CEP is and to illustrate some of the methods available to estimate the CEP.

The CEP was initially developed in order to give some criterion for measuring the expected effectiveness of a particular weapon system and to give some means for comparing similar weapon systems or weapons. In order to develop this criterion, it is essential that the assumptions used are well understood and established. The approach most often used is to assume that the errors in and across the line of sight are independent and that the variances are equal with the justification that these assumptions produce a negligible error. However, an error may be introduced and it is necessary to at least understand what is being assumed before making judgement on the legality of any assumption. This thesis therefore, attempts to explain such assumptions and to compare possible results of making certain assumptions in three example problems. The problems are all ficticious and utilized only for the purpose of explaining the estimating procedures and assumptions.

The thesis is primarily directed at the reader with a college background in calculus, some matrix theory, and some feel for basic probability and statistical procedures. The contents are arranged in six sections and three appendices. Section I is an introduction to the problem and the basic mathematical concepts which will be used. Sections II

and III introduce the most commonly used estimating procedures. Section IV explains the problem of deleting outlying observations from the determination of the estimate. Section V introduces the confidence interval. Section VI is a summary of the techniques used in the previous sections. Appendix A is concerned with the mathematical techniques which are used to explain and transform the true orientation of the dependent variables. Appendix B explains two methods of obtaining unbiased estimates of the CEP. Appendix C explains in detail the methods of integrations used. It is recommended that Appendices A and C be studied before starting Section II.

This thesis was written during the period January-June 1962 at the United States Naval Post Graduate School, Monterey, California. I wish to express my gratitude to Professor J. R. Borsting for his continued patience, encouragement, and most competent guidance while acting as faculty advisor, and to Professor Max Woods for his continuous aid and technical understanding of the problem while acting as second reader. I also wish to thank my wife for the moral, clerical, and artistic assistance given me during the writing of this thesis as well as the past two years.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

SECTION I

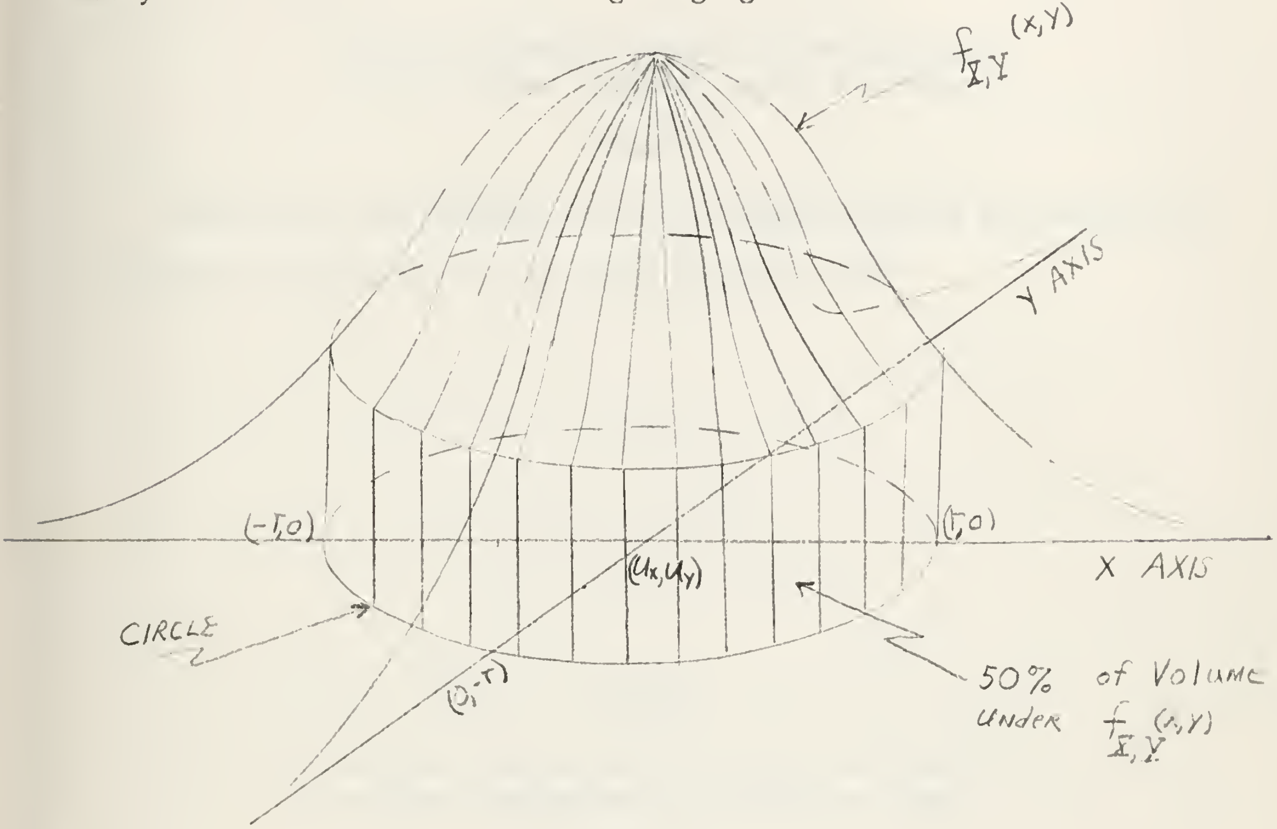
INTRODUCTION TO THE PROBLEM

1.1 General Discussion of the Circular Probable Error

The problem of determining useful estimates of the parameters which describe the distribution of the fall of shot about a target is directly related to the high cost of testing expensive weapon systems. Since relatively few tests are allowed because of this expense, it is not improbable that a good weapon system could be completely rejected because of inefficient utilization of the small amount of data available. Also, the size and yield of the warhead is directly related to the estimated parameters. If the estimated variance is large, the effective radius will also have to be large to cover the target complex, and in turn the missile will not be able to reach the range of the same missile with a smaller warhead. The most efficient use of the limited data will thus greatly reduce the risk involved in reducing the warhead size and increase the potential range. It also may aid in weapon deployment or assignment to larger targets because of the greater confidence that can be placed in the estimates. It seems logical that if a great deal of confidence can be placed in the weapon, fewer weapons will have to be assigned to a target, thus releasing some weapons for other targets. The important point is that the confidence placed on the estimators must be high enough to reduce the risk involved and provide a sound basis for decision.

One method, which is commonly used, to measure and compare the estimated parameters, is called the circular probable error or CEP method. The CEP is defined as the radius of the circle with center

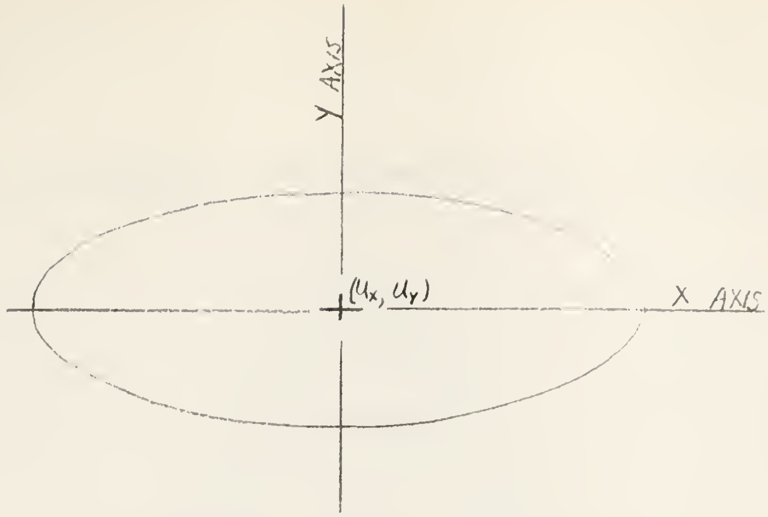
at (u_x, u_y) which includes 50% of a bivariate probability mass. The illustration in figure (1) shows the form of this function. It is to be noted that most of the volume under the curve is centered at the target and decreases as the distance increases from the target. This particular function is well founded historically on the basis of the analysis of observations from long range gun fire.



Bivariate Probability Mass

Figure 1

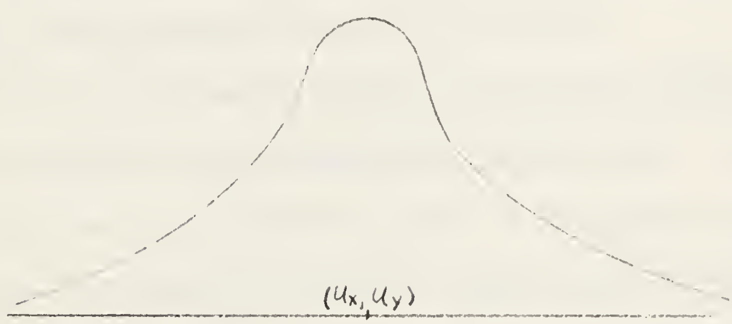
The bivariate normal distribution is a generalization of the normal distribution of a single variate and is bell shaped as shown in figure (1) above. Any plane parallel to the x,y plane that cuts the surface will intersect the bell in the elliptical curve shown in figure (2).



Bivariate Density Function which has been
Cut by a Plane Parallel to the x,y Plane.

Figure 2

Any plane perpendicular to the x,y plane will cut the surface in a curve of the normal form as shown in figure (3).

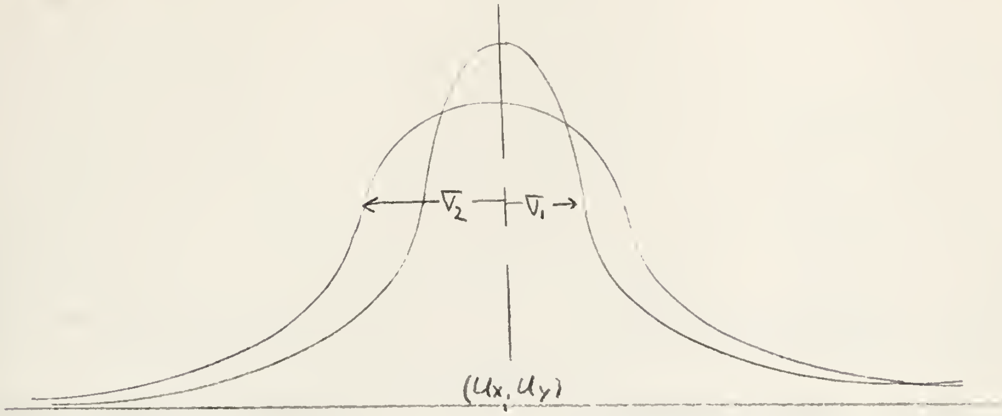


Bivariate Density Function which has been Cut
by a Plane Perpendicular to the x,y Plane.

Figure 3

The bivariate density function actually represents a five parameter family of distributions, the parameters being the means (u_x, u_y), the variances σ_x^2, σ_y^2, and the correlation coefficient ρ. This function is symmetric about the means and has its greatest value at the point (u_x, u_y). It should also be noted that if the errors in the x and y directions are independent and the variances equal, then the distribution will be in the shape of a bell with two of the opposites sides "pushed in" an

equal amount. The effect of the variance is shown in figure (4).

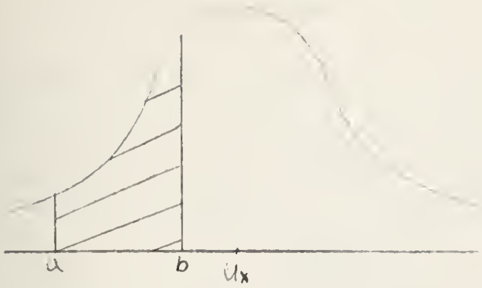


Two Bivariate Density Functions with
Different Variances about (u_x,u_y): Side View

Figure 4

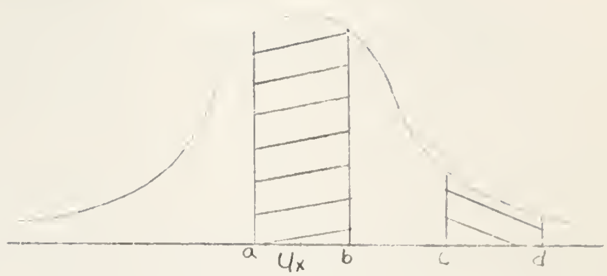
If the variances are equal, a plane cutting the surface, as in figure (2), will intersect the bell in a circle.

The height of the curve, forming the density function, at any point "a" is related to the probability of that point. Since this function is continuous, the probability must be expressed in the form of an interval since the probability of any single point is zero. However, the probability that a random variable X, in the distribution being considered, falls in an interval is equal to the area under the curve in the interval being considered. That is, the probability that $a \leq X \leq b$ is equal to the area shown under the curve in figure (5). Note that since the area under the curve about the point (u_x,u_y) is the greatest, the probability that the random variable X fall in this interval is greater than that of an interval of equal length away from the point (u_x,u_y). This is shown in figure (6).



Univariate Density Function Showing the Area Under Consideration When Determining P(a < X < b).

Figure 5



Univariate Density Function Showing the Areas Under Consideration In the Intervals (a,b) and (c,d) where b-a = d-c.

Figure 6

1.2 Mathematical Notation

X and Y are said to have a bivariate normal distribution if their joint density function, $f_{X,Y}(x,y)$, is given by

$$(1.1) \quad f_{X,Y}(x,y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-u_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-u_x}{\sigma_x}\right)\left(\frac{y-u_y}{\sigma_y}\right) + \left(\frac{y-u_y}{\sigma_y}\right)^2\right]\right\}$$

1.2.1

The quantity x is said to be an observed value of a numerical valued random phenomenon X if for every real number x there exists a probability that X is less than or equal to x. In this problem the observed values of the random variables X and Y are the coordinates of the data points with respect to the target. These coordinates can also be referred to as miss distances in and across the line of sight.

1.2.2

The parameters u_x and u_y are the mean values in the x and y directions respectively. The mean of a probability law is equivalent to the expected value of the random variable with respect to the probability law. This is written as:

$$(1.2) \quad u_x = E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \, f_{X,Y}(x,y)\,dxdy$$

$$(1.3) \quad u_y = E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y \, f_{X,Y}(x,y)\,dxdy$$

The mean value cannot be determined exactly in our problem even if all of the missiles have been fired but estimates of the mean values can be determined from the observations.

1.2.3

The expressions $(x - u_x)$ and $(y - u_y)$ are the deviations from the mean values in the x and y directions respectively.

1.2.4

σ_x and σ_y are the standard deviations in the x and y directions respectively. The standard deviation is defined as the square root of the variance of the probability law. The variance σ^2 is defined as the second central moment of the probability law and is defined by:

$$(1.4) \quad \sigma^2 = \left[(X - E(X))^2\right] = E\left[(X - u_x)^2\right] = E(X^2) - u_x{}^2$$

It should be noted that the mean values determine the location (u_x, u_y) of the center of the normal density function and the standard deviations (σ_x and σ_y) determine the shape of the function about the mean in the x and y directions respectively.

1.2.5

The correlation coefficient of two jointly distributed random variables X and Y is defined by $\rho = \frac{COV\ (X,Y)}{\sigma_x\ \sigma_y}$ where

(1.5) $$COV\ (X,Y) = E(X\ Y) - E(X)E(Y)$$

$$E(X\ Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)\,dxdy$$

The correlation coefficient provides a measure of how good a prediction can be formed on one of the random variables on the basis of the observed value of the other random variable. In other words, if the value of one of the random variables is given, the expected value of the other random variable can be determined. This may be written as $E(X|Y)$ where the value of Y is given. That is,

(1.6) $$E(X|Y) = \int_{-\infty}^{\infty} x\ f_{X|Y}(x|y)dx$$ where $f_{X|Y}(x|y)$

is the conditional density function of the random variable X given the value of the random variable Y. The conditional density function is derived from the conditional probability of a random event A, given a random variable X. This notion forms the basis of the mathematical treatment of jointly distributed random variables that are not independent.[1]

In the particular case where two random variables X and Y are jointly normally distributed, the conditional expected value of the random variable X given that the random variable Y is some particular value y, is a linear function. This linear function is related to the orientation of the shape of the density function as shown in Appendix A.

1 "Modern Probability Theory and Its Applications" by Emanuel Parzen /1/ of Stanford University.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

In order to simplify the notation, it will be convenient to represent the bivariate density function in matrix notation. The terms in formula 1.1 are first arranged in the form

$$(1.7)\quad f_{X,Y}(x,y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \exp -\left[\frac{1}{2(1-\rho^2)} (x-u_x,\ y-u_y) \begin{pmatrix} \frac{1}{\sigma_x^2} & -\frac{\rho}{\sigma_x \sigma_y} \\ -\frac{\rho}{\sigma_x \sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix} \begin{pmatrix} x-u_x \\ y-u_y \end{pmatrix}\right]$$

$$= \frac{1}{2\pi |A^{-1}|^{1/2}} \exp -\frac{1}{2} Z'AZ$$

where $Z = \begin{pmatrix} X - u_x \\ Y - u_y \end{pmatrix}$

$$A = \frac{1}{1-\rho^2} \begin{pmatrix} \frac{1}{\sigma_x^2} & -\frac{\rho}{\sigma_x \sigma_y} \\ -\frac{\rho}{\sigma_x \sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{pmatrix}.$$

Using this notation, we are now ready to look at several models investigating the CEP and confidence interval of the CEP.

1.3 The Basic Problem in Estimating the CEP.

The problem of estimating the CEP is essentially that of finding the radius of a circle with center at (u_x, u_y) such that the probability is .5 that a random point (X,Y) will lie inside this circle. This may be expressed as

$$(1.8)\quad P\left[(X-u_x)^2+(Y-u_y)^2 < r^2\right] = \iint_{\sqrt{(x-u_x)^2+(y-u_y)^2} < r} f_{X,Y}(x,y)\,dx\,dy$$

where $f_{X,Y}(x,y)$ is given by formula (1.1).

In order to introduce the problem, the assumptions will be made that the mean values are zero ($u_x = u_y = 0$), that the errors in the x and y directions are independent ($\rho = 0$), and that the standard deviations are equal ($\sigma_x = \sigma_y = \sigma$). The probability statement is thus simplified to

$$(1.9) \quad P\left[X^2 + Y^2 < r^2\right] = \frac{1}{2\pi\sigma^2} \iint_{\sqrt{X^2+Y^2}<r} \exp - \left[\frac{(x^2 + y^2)}{2\sigma^2}\right] dxdy = .5$$

In order to perform the integration let $R^2 = X^2 + Y^2$, $\tan\theta = \frac{Y}{X}$, $Y = R\sin\theta$, $X = R\cos\theta$.

Then $P(R \leq r) = \iint_{R<r} f_{X,Y}\left[(r\cos\theta),(r\sin\theta)\right] |J| \, dr d\theta$ where $J = \begin{vmatrix} dx/dr & dx/d\theta \\ dy/dr & dy/d\theta \end{vmatrix} = r$

$$(1.10) \quad P(R \leq r) = \frac{1}{2\pi\sigma^2} \int_0^{2\pi} \int_0^{r} r \exp\left(-\frac{r^2}{2}\right) dr \, d\theta = 1 - \exp \frac{-r^2}{2} = .5$$

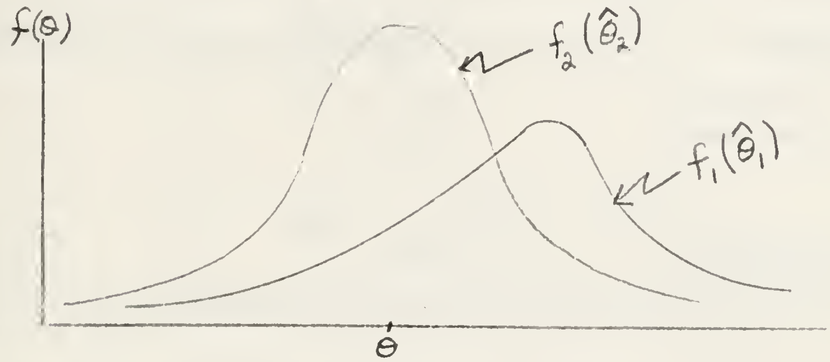
Therefore, the CEP $= r = 1.1774\sigma$.

The problem of estimating the CEP is thus one of obtaining a function of the n sample points $(x_1, y_1) \ldots\ldots\ldots (x_n, y_n)$ which will estimate the standard deviation σ. The estimators are functions of the observed values which are used to estimate the true values of the parameters. For example, if m points from a sample are given, the average or mean value is estimated by

$$(1.11) \quad \bar{x} = \frac{x_1 + x_2 + x_3 \ldots\ldots\ldots + x_m}{m}$$

The distribution of $\bar{x}$ becomes closely concentrated about the true value u_x as m becomes large.

There are many ways to estimate the parameters under investigation, and it is therefore necessary to specify certain properties which are desired in estimators. For example, the distribution of the estimator should be concentrated near the true parameter value. If $\hat{\theta}_1$ and $\hat{\theta}_2$ are different estimators of θ with density functions $f_1(\hat{\theta}_1)$ and $f_2(\hat{\theta}_2)$ as shown in figure (9), then $\hat{\theta}_2$ is a better estimator of θ than $\hat{\theta}_1$.



The Density Functions of Two Estimators

Figure 7

Other properties which are desired in estimators are defined as follows:

1.3.1 Relative Efficiency. The relative efficiency of two estimators is defined as a ratio of the mean square errors of the estimators. That is,

(1.12) $$\frac{E(\hat{\theta}_1 - \theta)^2}{E(\hat{\theta}_2 - \theta)^2} = R.F.$$ where R.F. is the ratio function.

If R.F. $<$ 1, then $\hat{\theta}_1^2$ is said to be a more efficient estimate of θ than $\hat{\theta}_2$.

1.3.2 Unbiased Estimator. An estimator, $\hat{\theta}$ is said to be an unbiased estimate of the parameter θ if $E(\hat{\theta}) = \theta$.

1.3.3 Consistent Estimator. An estimator $\hat{\theta}$ is said to be a consistent estimate of θ if $P(\hat{\theta} \to \theta) \to 1$, as $n \to \infty$.

1.3.4 Efficient Estimator. The estimators which have the smallest limiting variances are called efficient estimators of θ.

The estimators which will be used in the first part of this thesis are shown in Table a.

Table a

Properties of the Estimators Used in Models I and II

Parameter	Estimator	Properties
u_x	$\bar{x} = \sum_{i=1}^{N} \frac{x_i}{n}$	Unbiased, efficient, and consistent.
σ_x^2	$\hat{\sigma}_x^2 = \frac{1}{n-1} \sum_{i=1}^{N} (x_i - \bar{x})^2$	Unbiased, efficient and consistent.

A more detailed discussion of certain estimators under special assumptions is presented in Appendix B.

1.4 The Problem of Dependence

In the gunnery-problem, the errors introduced in the line of sight are due to variations in the range and projectile initial velocity. The error across the line of sight is due to bearing errors. Since bearing errors and range errors are independent of each other due to the fact that they are obtained from different sources, the mathematical assumption is generally made that these errors are also statistically independent. However, if we broaden the perspective to look at the major errors introduced in a missile trajectory, the major errors in the line of sight and across the line of sight are probably not independent of each other.

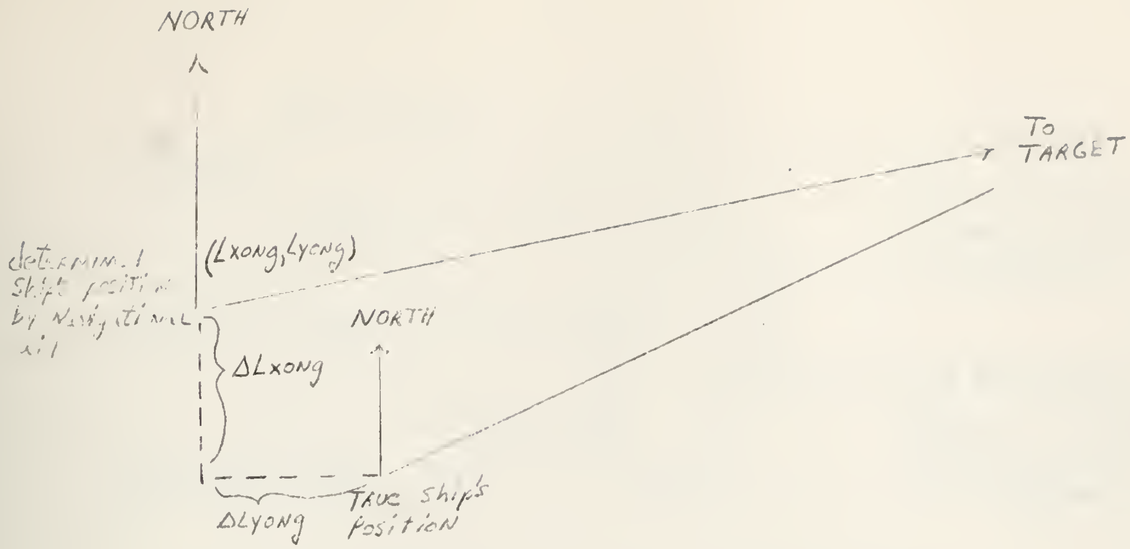
This is primarily due to factors which did not especially influence the gunnery fire control problem such as errors in ship's navigational

[illegible]

position, errors introduced by missile attitude during the time of powered flight, especially at cutoff, and weather conditions over the target.

In the gunnery problem there are two types of navigational problems. The first is the relative problem of firing from a moving object to another moving target where the fire control problem is one of obtaining relative bearings, ranges, courses, and speeds. But the firing ship's true navigational position relative to the target is not an influencing factor.

The second problem is one of shore bombardment where the ship's navigational position is determined by visual fix. This is closely related to missile launching except that the first shot in shore bombardment does not have to hit the target because the shore observer can tell the ship what spots to apply to the generating fire control solution. Therefore, this again becomes a relative fire control problem where errors introduced by the ship's and target's relative positions are corrected by spotting. This is not practical in long range missile launching because of the inability to obtain corrected visual navigational positions relative to the target due to lack of observers at the target area. What is done instead is that the probable errors must be predetermined and enough missiles launched to give a high probability of destruction of the target complex. If we assume that the launching ship is determined to be at the launch reference point then the errors introduced are as shown in figure (8).



Byrg = firing bearing from ship to position target will have at detonation.

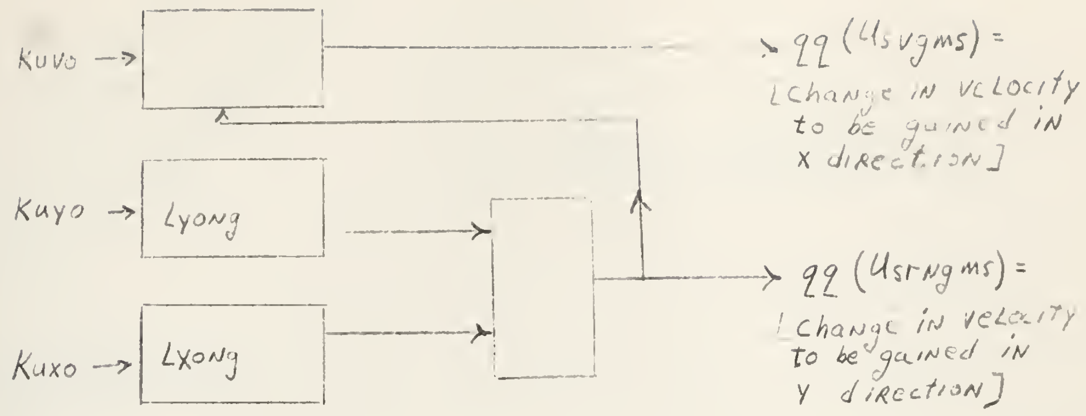
True Target Bearing Diagram

Figure 8

Byrg is proportional to $\begin{Bmatrix} \text{Lxong} \\ \text{Lyong} \end{Bmatrix}$

Byrg ' is proportional to $\begin{Bmatrix} \text{Lxong} + \text{Lxong} \\ \text{Lyong} + \text{Lyong} \end{Bmatrix}$

Since Byrg differs from Byrg' by the errors introduced in and across the line of sight, the errors are reflected in the q(Ugm)s interpolation computer as errors in velocity to be gained, which have not been entered. But the errors introduced are not independent because the inputs influence changes in velocity to be gained in both range and cross range directions as shown below:

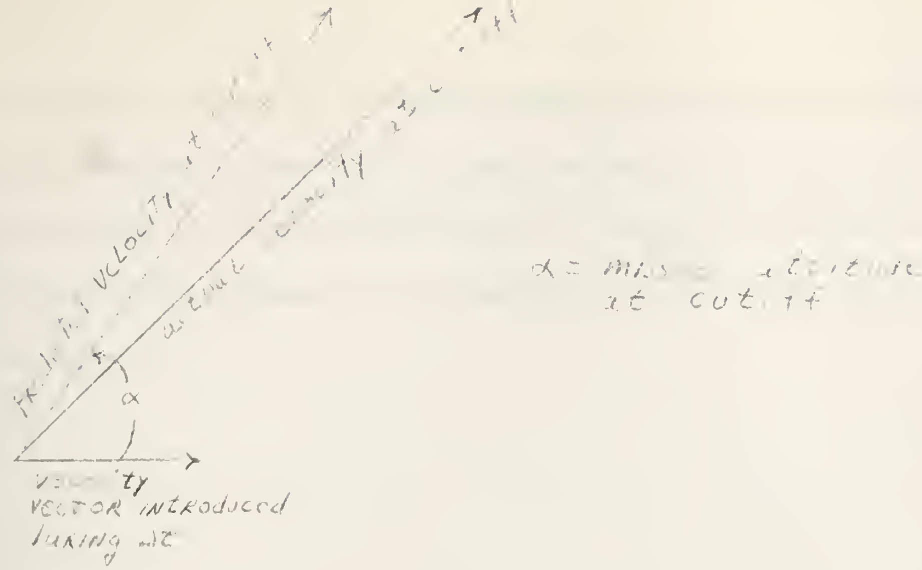


Flow Diagram of the Change in Velocity to be Gained

Figure 9

In the gunnery problem, the weather conditions over the firing ship's position are the same as the weather conditions over the target, therefore these values can be accurately estimated. The missile firing ship depends upon intelligence and weather forecasts to predict the inputs for target weather conditions. This information is therefore not as accurate as in the gunnery problem. Since the errors introduced by weather predictions influence the missile trajectory over the target, the re-entry body is most likely to be moved in any direction and the probability that the errors in and across the line of sight are independent of each other is low.

The errors introduced by missile attitude during cutoff can best be illustrated by a vector diagram.



Vector Diagram of Velocities at Cutoff

Figure 10

The missile attitude at cutoff can be regarded as a random variable because it can assume any attitude due to the fact that the requirements to initiate cutoff are due to past and present missile velocity and not to a predicted velocity at some Δ t after cutoff. Thus the errors introduced during the Δt of cutoff will influence the errors in and across the line of sight in a random manner. Therefore, the probability that the errors in and across the line of sight are independent is again lowered.

The conclusion is that due to the complexity of the fire control problem, the errors in and across the line of sight are probably not independent. If we approach the problem with this assumption and find that the increase in accuracy gained by this model is not sufficient to justify the increase in the mathematical difficulty, the independent model can be used.

1.5 Summary

Section one has been used to set up the environment of the problem that is to be analyzed. The basic assumption made is that the fall of

shot about the target is a random variable which obeys the bivariate normal probability laws. The assumption has been made that the errors in and across the line of sight are not independent and one of the objectives of this paper is to determine the effect of this assumption on the CEP.

SECTION II

DETERMINING THE CEP WITH THE DENSITY FUNCTION CENTERED AT THE TARGET: MODEL I

2.1 Introduction

The most important assumption made in this model is that u_x and u_y are zero. This means that the center of the bivariate density function is at the target. Although this is the desired condition, it may not be true initially due to the complexities of the fire control problem. One of the determinations that is made from the analysis of the firing data is whether a correction should be made to the fire control solution to bring the distribution of the fall of shot over the target. Therefore, by starting with the assumption that the center of the distribution is at the target and finding that this assumption is wrong, it becomes necessary to determine and apply the correction to the fire control solution. Also, it should be noted that although this assumption may not be true initially, it still may be true after correcting the initial fire control solution.

If the center of the distribution is close to the target, (0,0) in the coordinate system, or suspected of being so by analysis of the test data, the estimators determined from this model may be better estimators than the estimators used in Model II in Section III. A comparison can be made between model I and model II, using the criterion of relative efficiency to determine which model is theoretically the best. This criterion is explained in Section III.

In this model the errors in the x and y direction are assumed to be non-independent and distributed in accordance with the bivariate probability law. The probability that a random point (X,Y) will lie

[illegible]

[illegible]

[illegible]

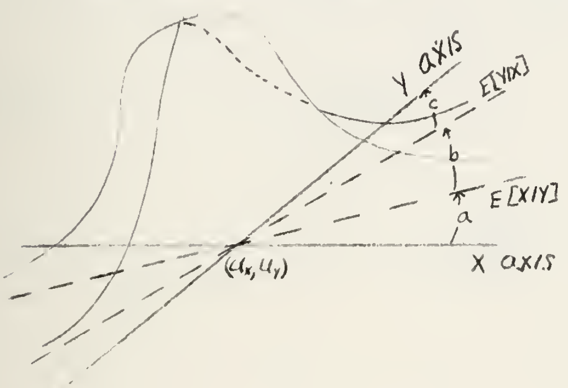
[illegible]

within a circle of radius $k\sigma_{max}$ is equal to

$$(2.1)\quad P(k,\sigma_x,\sigma_y) = P\left(\sqrt{X^2+Y^2} < k\sigma_{max}\right) = \iint\limits_{\sqrt{X^2+Y^2} < k\sigma\, max} f_{X,Y}(x,y)\,dxdy = \iint\limits_{\sqrt{X^2+Y^2} < k\sigma\, max} \frac{1}{2\pi\,|A^{-1}|^{\frac{1}{2}}} \exp -\tfrac{1}{2}Z'AZ\,dxdy$$

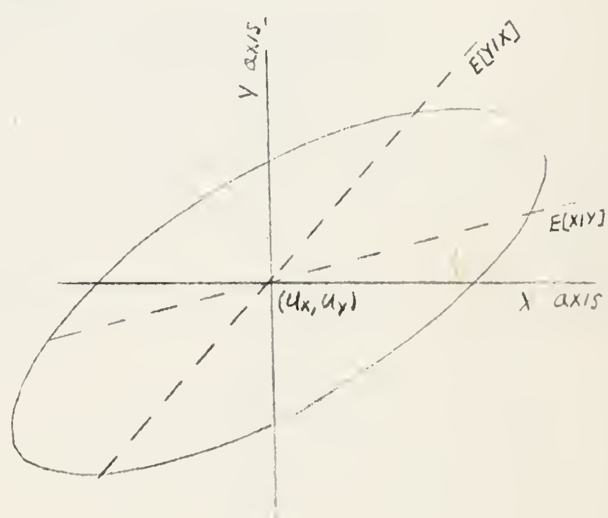
where Z and A are defined in (1.7).

In order to integrate over this form, it is necessary to first make a transformation to an orthogonal density function. The reason for this is that due to the assumption of non-independence ($\rho \neq 0$), this density function is oriented along non-orthogonal lines called the expected value of X given Y and the expected value of Y given X or in simpler notation $E(X/Y)$ and $E(Y/X)$ as defined in Section 1.2.5. This orientation is illustrated in figures 11a and 11b.



Three Dimensional Diagram of the Orientation of the Bivariate Density Function where $0 < \rho < 1$, $a+b+c = 90^o$.

Figure 11a



Two Dimensional Diagram of the Bivariate Density Function Formed by a Plane Parallel to the x,y Plane Cutting the Density Function.

Figure 11b

This reorientation is shown to be valid byproving that

(2.2) $\quad$ [illegible] $= W'A^*W$ $\quad$ where $W = \begin{pmatrix} u \\ v \end{pmatrix}$, $A^* = \begin{pmatrix} \frac{1}{\sigma_v^2} & 0 \\ 0 & \frac{1}{\sigma_u^2} \end{pmatrix}$, and [2]

$$(2.3) \quad \sigma_u^2 = \frac{\sigma_x^2 + \sigma_y^2 + \sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4\sigma_{xy}^2}}{2}$$

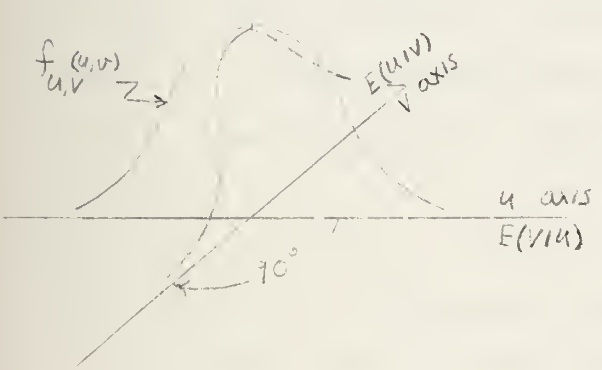
$$\sigma_v^2 = \frac{\sigma_x^2 + \sigma_y^2 - \sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4\sigma_{xy}^2}}{2}$$

The transformed density function thus becomes

$$g_{U,V}(u,v) = \frac{1}{2\pi\,|A^{*-1}|^{\frac{1}{2}}} \exp -\tfrac{1}{2}W'A^*W =$$

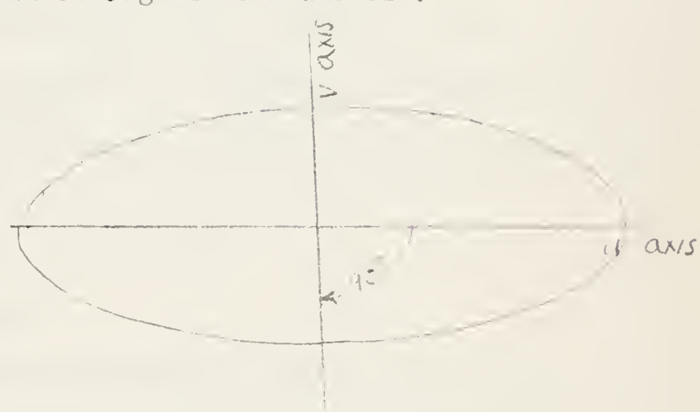
$$(2.4) \quad \frac{1}{2\pi\,\sigma_u\sigma_v} \exp\left[-\tfrac{1}{2}\left(\frac{u}{\sigma_u}\right)^2 + \left(\frac{v}{\sigma_v}\right)^2\right]$$

The reoriented axes are now as shown in figures 12a and 12b.



Three Dimensional Diagram of the Reoriented Axes of the Bivariate Density Function.

Figure 12a



Two Dimensional Diagram of the Reoriented Bivariate Density Function Formed by a Plane Parallel to the u,v Plane Cutting the Density Function.

Figure 12b

[2] The details of this transformation are contained in Appendix A.5.

This transformed density function can be handled more easily because the terms involving the correlation coefficient have been removed. The probability that a point (U,V) in the new coordinate system will lie within a circle with center at the origin and radius $k\sigma_u$ is

$$(2.5)\quad P(k,\sigma_u,\sigma_v) = \iint\limits_{\sqrt{U^2+V^2}<k\sigma_u} g_{U,V}(u,v)\,du\,dv = P(k,c) = \frac{1}{2\pi\sigma_u\sigma_v}\iint\limits_{\sqrt{U^2+V^2}<k\sigma_u} \exp -\tfrac{1}{2}\left[\left(\frac{u}{\sigma_u}\right)^2+\left(\frac{v}{\sigma_v}\right)^2\right]du\,dv$$

where $c = \frac{\sigma_v}{\sigma_u}$. This form is simplified in Appendix C.5 to

$$(2.6)\quad P(k,c) = \frac{2c}{\pi}\int_0^{\pi} \frac{1 - \exp\left\{\frac{-k^2}{4c^2}\left[(c^2+1)+(c^2-1)\cos\phi\right]\right\}}{(c^2+1)+(c^2-1)\cos\phi}\,d\phi$$

The values of (k) for various values of P(k,c) and (c) are tabulated in tables one and two. Table one is used by entering the table with $c = \frac{\sigma_v}{\sigma_u}$ in order to find k. This table can only be used for P(k,c)=.5. Table two is used by entering the table with $c = \frac{\sigma_v}{\sigma_u}$ and the probability P(k,c) in order to find k. This table can be used for various values of P(k,c).

2.2 Estimating the CEP using Model I

The first step is to find estimators for σx, σy and ρ from the n observed points (x_1,y_1)..........(x_n,y_n). This is done by computing the sample variances $\hat{\sigma}x^2$, $\hat{\sigma}y^2$, the sample covariance $\hat{\sigma}xy$, and the sample correlation coefficient $\hat{\rho}$ which are defined as follows:

$$\hat{\sigma}_x^2 = \frac{1}{N}\sum_{i=1}^{N} x_i^2 \qquad \hat{\sigma}_y^2 = \frac{1}{N}\sum_{i=1}^{N} y_i^2 \qquad \hat{\sigma}_{xy} = \frac{1}{N}\sum_{i=1}^{N} x_i y_i$$

$$\hat{\rho} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x\hat{\sigma}_y}$$

In these formulas, $\hat{\sigma}^2 x$, $\hat{\sigma}^2 y$, and $\hat{\sigma} xy$ are unbiased estimates of σx, σy, and σxy.

The transformed estimates of the variances are computed next.

$$(2.7)\quad \hat{\sigma}_u^2 = \frac{\hat{\sigma}_x^2 + \hat{\sigma}_y^2 + \sqrt{(\hat{\sigma}_x^2 - \hat{\sigma}_y^2)^2 + 4\hat{\sigma}_{xy}^2}}{2}$$

$$\hat{\sigma}_v^2 = \frac{\hat{\sigma}_x^2 + \hat{\sigma}_y^2 - \sqrt{(\hat{\sigma}_x^2 - \hat{\sigma}_y^2)^2 + 4\hat{\sigma}_{xy}^2}}{2},$$

Table two is entered with $P(k,c) = .5$ and $c = \frac{\hat{\sigma}_v}{\hat{\sigma}_u}$ to find k.

The estimate of the CEP $= \widehat{CEP}_1 = k\,\hat{\sigma}_u$

2.3 Estimating the CEP using the Assumption that the Errors in the x and y Directions are Independent.

If it has been assumed that the errors in the x and y directions are independent, an estimate of the CEP can be obtained by using the estimators in model I except that the estimated variances $\hat{\sigma}_x^2$ and $\hat{\sigma}_y^2$ are used instead of the estimated transformed variances $\hat{\sigma}_u^2$ and $\hat{\sigma}_v^2$

$$c^* = \frac{\hat{\sigma}_{min}}{\hat{\sigma}_{max}} \qquad \text{where} \qquad \begin{aligned}\hat{\sigma}_{min} &= \mathrm{Min}(\hat{\sigma}_x, \hat{\sigma}_y)\\ \hat{\sigma}_{max} &= \mathrm{Max}(\hat{\sigma}_x, \hat{\sigma}_y)\end{aligned}$$

$$P(k^*, c^*) = .5\ ,$$

Table two is entered with $P(k^*,c^*)$ and c^* in order to find k^*.

Then this estimate of the CEP $= \widehat{CEP}_1^* = k^*\,\hat{\sigma}_{max}$.

2.4 Information About the Problems.

In the problems which follow, both estimates of the CEP will be obtained in order to compare the results in the summary in Section VI.

2.5 Example Problems

The problems, which will be used to compare methods of estimating the CEP, have been set up in three cases. The first case will have ten sample points (x_1,y_1) ....(x_{10},y_{10}) and is representative of the point in time where some initial decision may be made as to whether the

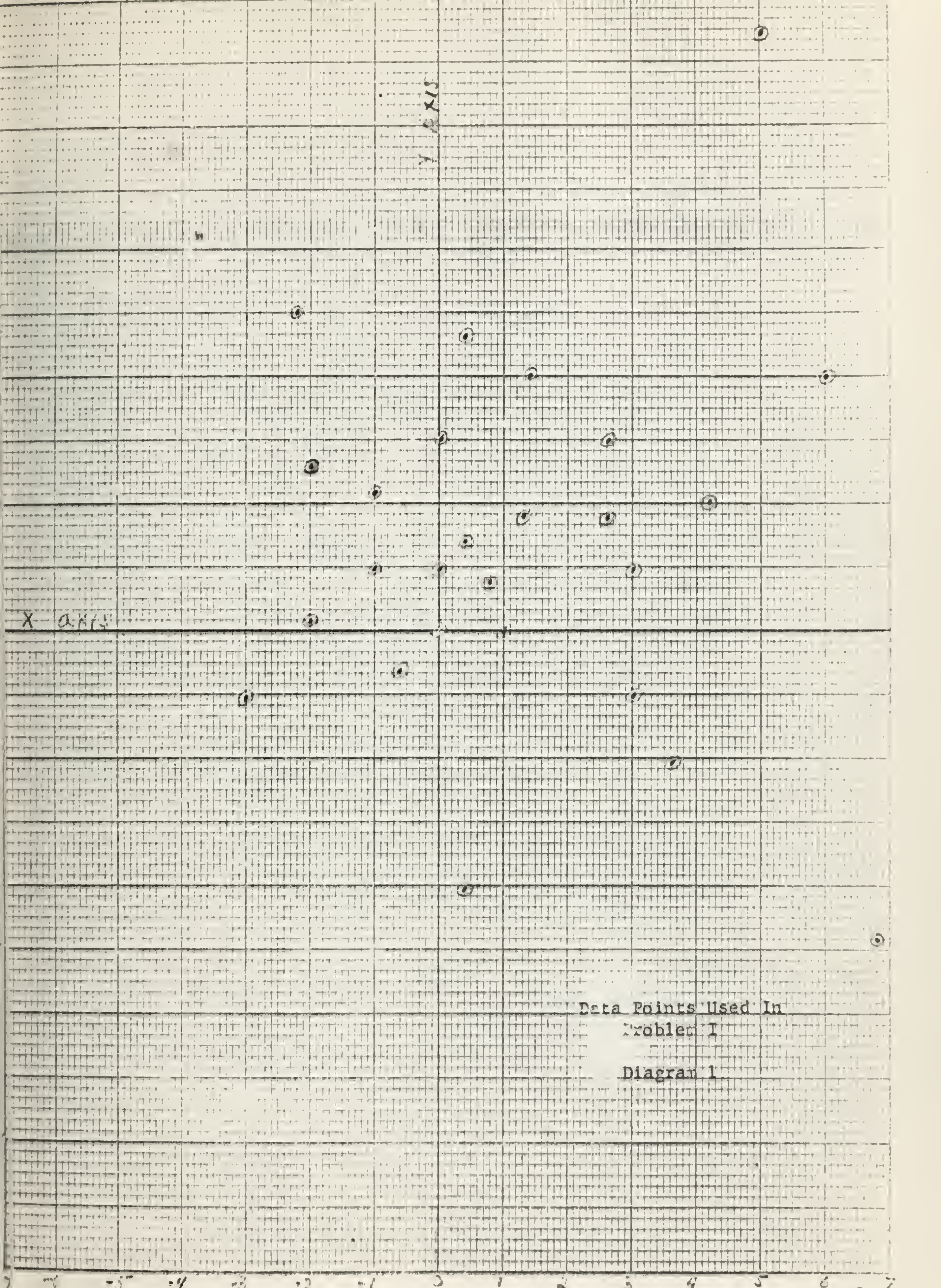
[illegible]

system should be accepted, rejected, or that more tests should be conducted. The second case will have fifteen sample points (x_1, y_1)..... .....(x_{15}, y_{15}) which will include the first ten sample points. This is intended to represent an intermediate point in time where some terminal decision may be made on the acceptance of the weapon system. The third case will consist of twenty five sample points (x_1, y_1)........(x_{25}, y_{25}). It should be noted that as the number of observations increase, the estimators are more likely to be closer to the true values. The actual distributions of the 25 points are shown in diagrams 1,2, and 3. The coordinates of the points are as follows:

	Problem I		Problem II		Problem III		Case
	x	y	x	y	x	y	
1.	-3.0	-1.0	-5.8	8.6	-8.6	-11.8	
2.	-2.2	5.0	-2.6	1.6	-3.6	3.2	
3.	-1.0	1.0	0	1.0	-1.6	- .2	
4.	- .6	- .6	1.3	1.0	-3.0	- .8	
5.	0	3.0	-1.6	-1.0	1.2	- 2.2	I
6.	1.0	0	.6	-1.0	0	- 1.2	
7.	3.6	-2.0	3.0	- .6	1.6	4.2	
8.	3.0	1.0	- .4	-2.4	.4	1.6	
9.	6.0	4.0	-1.0	-4.0	1.8	.4	
10.	5.0	[illegible].4	-4.0	-2.0	6.0	4.0	
11.	0	-1.0	-3.4	3.0	-2.6	- 3.6	
12.	1.4	4.0	0	2.8	1.6	- .6	
13.	.4	-4.0	-2.6	-1.8	.4	.2	II
14.	2.6	3.0	.2	-7.0	-1.8	1.4	
15.	6.0	-4.0	-2.6	0	- .4	2.6	
16.	-2.0	.2	-5.0	-5.8	-5.0	-4.0	
17.	-2.0	2.6	-6.0	- .8	-3.8	-2.0	
18.	-1.0	2.1	-5.8	2.2	-2.0	-2.0	
19.	.4	1.4	-1.4	5.0	-1.0	-1.0	
20.	.3	.6	- .6	0	- .6	-1.8	III
21	1.3	[illegible].3	-1.4	1.4	3.0	- .8	
22.	2.6	1.6	1.4	-1.0	- .4	.6	
23.	.2	4.6	2.8	2.0	.8	-5.0	
24.	4.2	2.6	2.2	5.8	1.4	1.6	
25.	3.0	-1.0	3.4	-3.6	3.4	2.2	

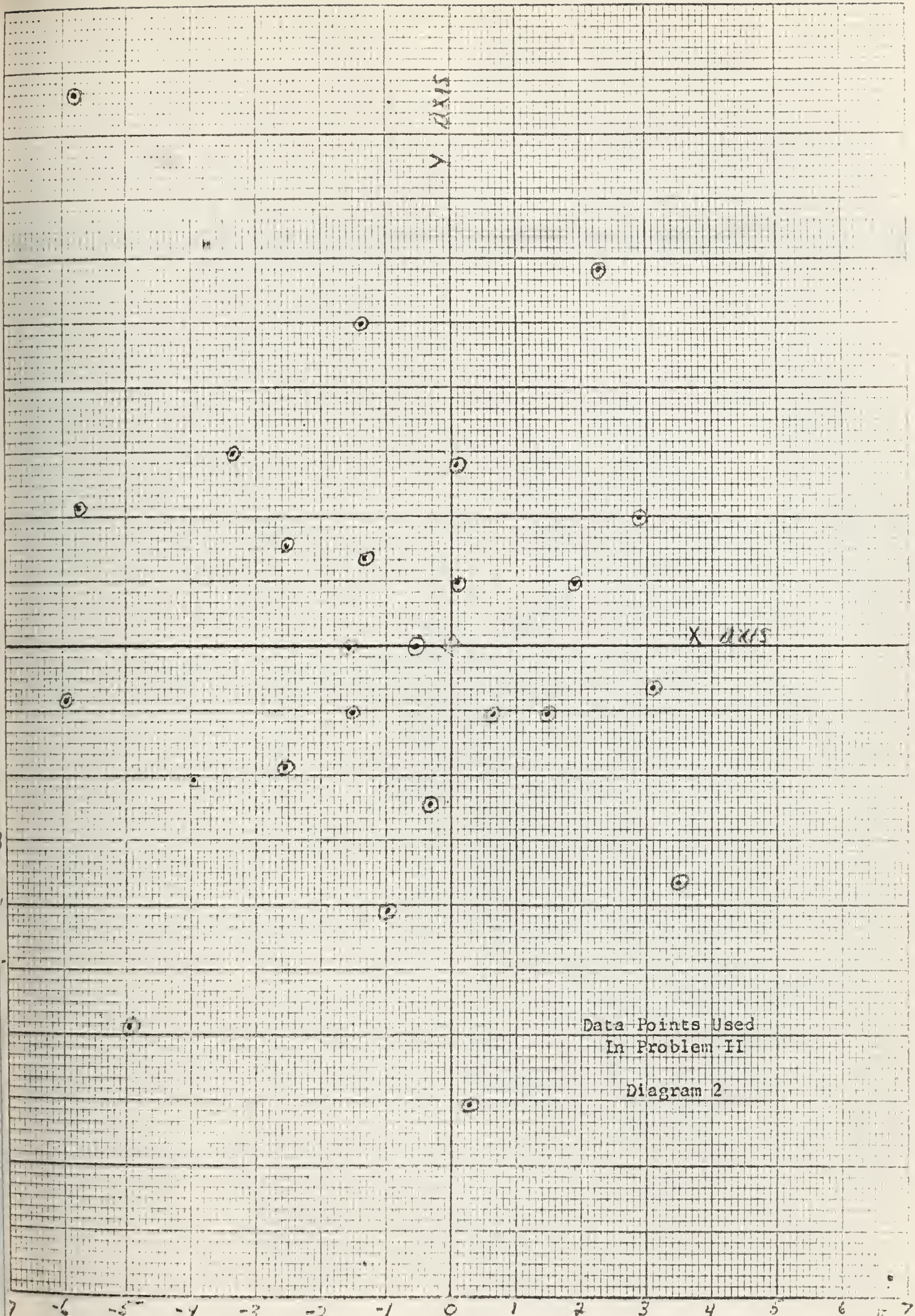
The value of the CEP obtained using the estimators from this section will be compared to the estimates of the CEP from Sections III, IV, and estimators which are explained in Appendix B. This comparison will extend to the problem of rejecting outliers and the comparison will be presented in Section VI.

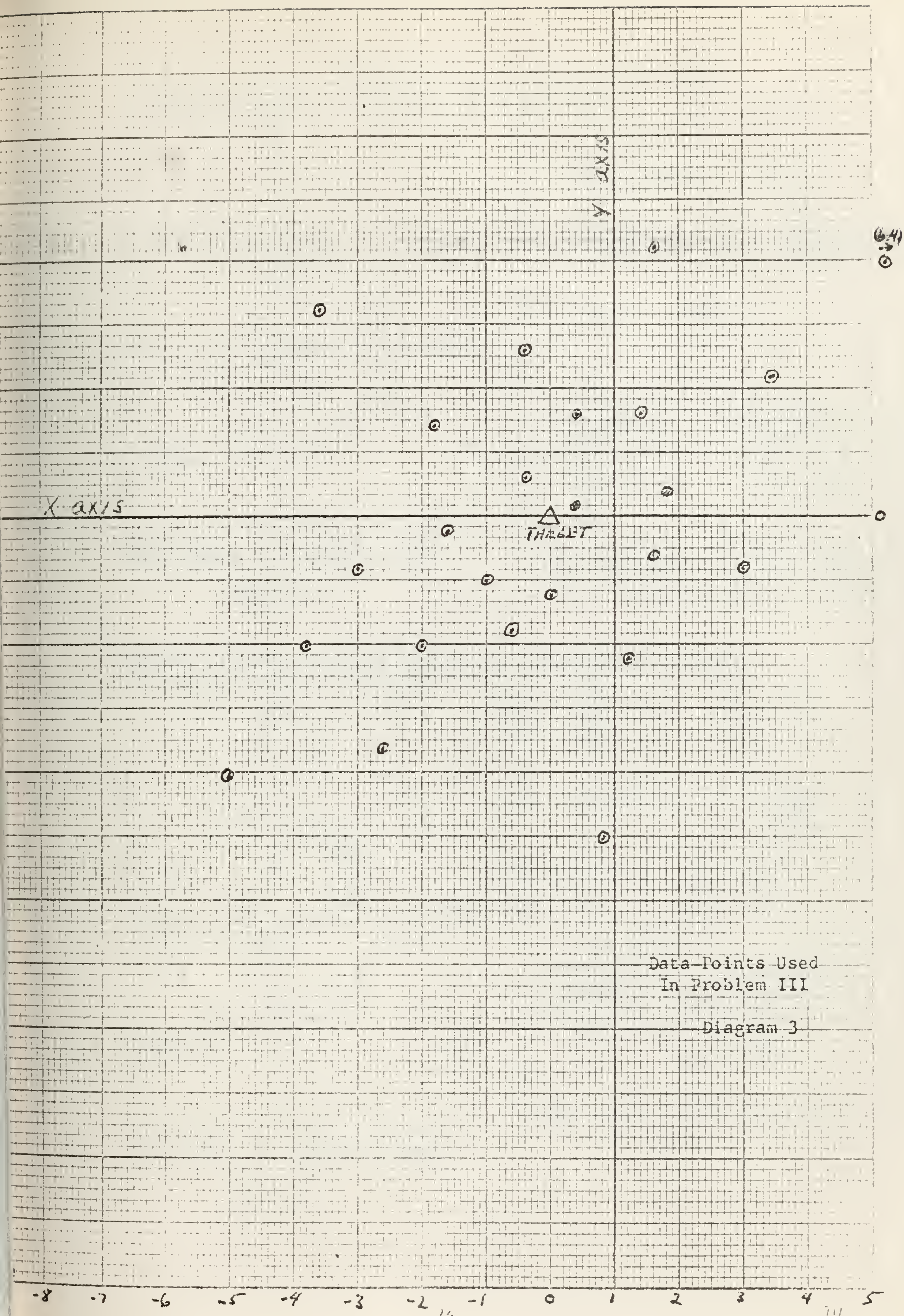
Although these problems are primarily oriented at tests involving the more expensive weapon systems, such as the IRBM, the environment can be extended to less expensive weapon systems which will naturally have more sample points. Although it was intended to make the problems as realistic as possible, no attempt was made to utilize data from actural missile firings.



Data Points Used In
Problem I

Diagram 1

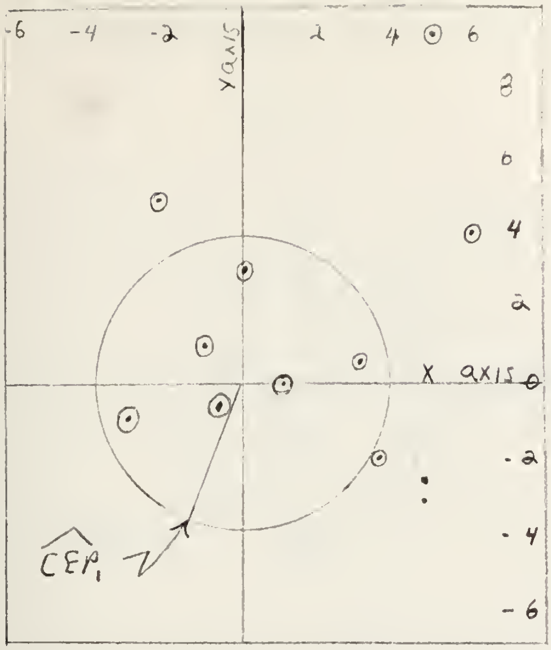

Y axis
X axis
Data Points Used
In Problem II
Diagram 2
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7



Data Points Used
In Problem III

Diagram 3

Problem I, Case I. Data points and computational results.



Data Points in Problem I, N=10

Diagram 4

$\hat{\sigma}_x^2 = \frac{\sum x_i^2}{N} = 10.3$ $\quad \hat{\sigma}_x = 3.2$

$\hat{\sigma}_y^2 = \frac{\sum y_i^2}{N} = 14.6$ $\quad \hat{\sigma}_y = 3.8$

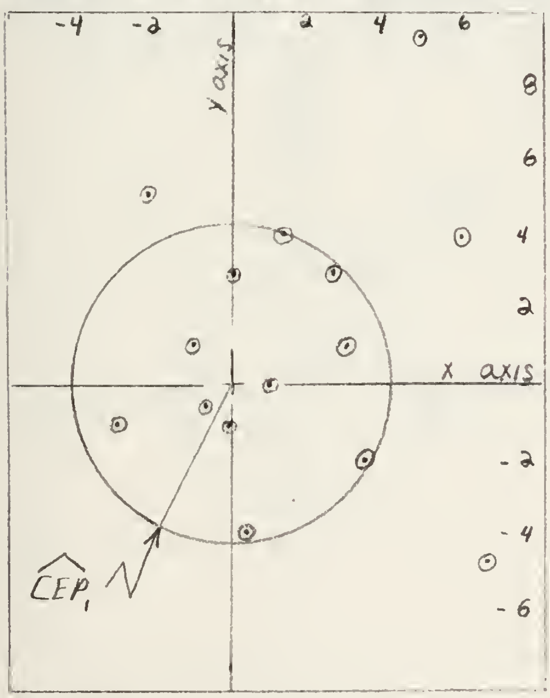
$\hat{\sigma}_{xy} = \frac{\sum x_i y_i}{N} = 5.8$ $\quad \hat{\rho} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x \hat{\sigma}_y} = .47$

$\hat{\sigma}_u^2 = 18.6$ $\quad \hat{\sigma}_u = 4.3$

$\hat{\sigma}_v^2 = 6.3$ $\quad \hat{\sigma}_v = 2.5$

Dependent Model	Independent Model
$C = \frac{\hat{\sigma}_v}{\hat{\sigma}_u} = .58$	$C^* = \frac{\hat{\sigma}_x}{\hat{\sigma}_y} = .84$
$K = .92$	$K^* = 1.08$
$\widehat{CEP}_1 = K\hat{\sigma}_u = 3.97$	$\widehat{CEP}_1^* = K^*\hat{\sigma}_y = 4.1$

Problem I, Case II. Data points and computational results



Data Points in Problem I, N=15

Diagram 5

$\hat{\sigma}_x^2 = \frac{\sum x_i^2}{N} = 10.3$, $\quad \hat{\sigma}_x = 3.2$

$\hat{\sigma}_y^2 = \frac{\sum y_i^2}{N} = 14.0$ $\quad \hat{\sigma}_y = 3.7$

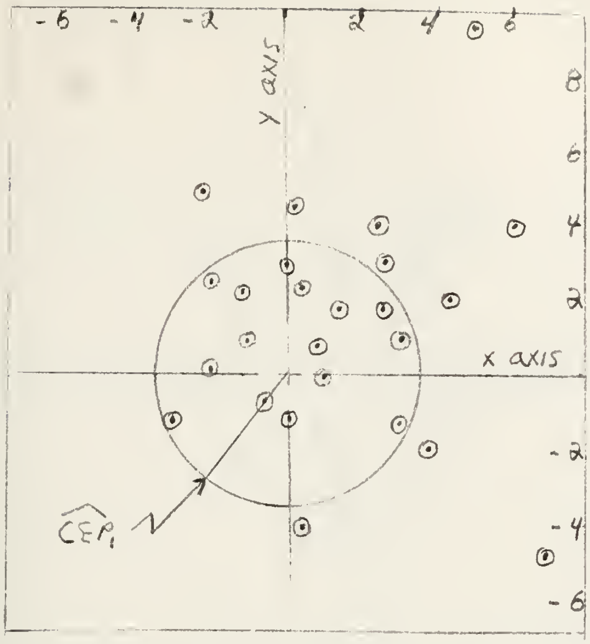
$\hat{\sigma}_{xy} = \frac{\sum x_i y_i}{N} = 2.5$ $\quad \hat{\rho} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x \hat{\sigma}_y} = .21$

$\hat{\sigma}_u^2 = 15.3$ $\quad \hat{\sigma}_u = 3.9$

$\hat{\sigma}_v^2 = 9.1$ $\quad \hat{\sigma}_v = 3.0$

Dependent Model	Independent Model
$C = \frac{\hat{\sigma}_v}{\hat{\sigma}_u} = .77$	$C^* = \frac{\hat{\sigma}_x}{\hat{\sigma}_y} = .85$
$K = 1.04$	$K^* = 1.09$
$\widehat{CEP}_1 = K\hat{\sigma}_u = 4.1$	$\widehat{CEP}^* = K\hat{\sigma}_y = 4.1$

Problem I, Case III. Data points and computational results.



Data Points in Problem I, N=25

Diagram 6

$\hat{\sigma}_x^2 = \frac{\sum x_i}{N} = 7.98 \qquad \hat{\sigma}_x = 2.7$

$\hat{\sigma}_y^2 = \frac{\sum y_i}{N} = 10.6 \qquad \hat{\sigma}_y = 3.3$

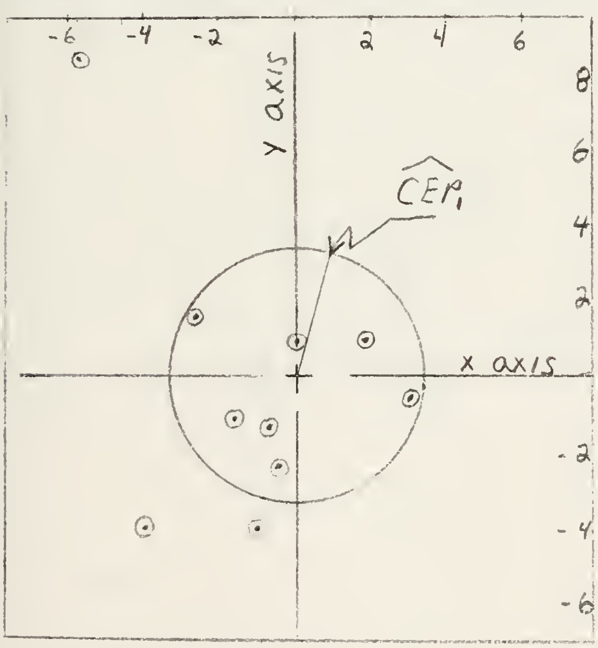
$\hat{\sigma}_{xy} = \frac{\sum x_i y_i}{N} = 1.8 \qquad \hat{\rho} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x \hat{\sigma}_y} = .20$

$\hat{\sigma}_u^2 = 11.5 \qquad \hat{\sigma}_u = 3.4$

$\hat{\sigma}_v^2 = 7.1 \qquad \hat{\sigma}_v = 2.7$

Dependent Model	Independent Model
$C = \frac{\hat{\sigma}_v}{\hat{\sigma}_u} = .8$	$C^* = \frac{\hat{\sigma}_x}{\hat{\sigma}_y} = .83$
$K = 1.05$	$K^* = 1.08$
$\widehat{CEP}_1 = K\hat{\sigma}_u = 3.55$	$\widehat{CEP}_1^* = K^*\hat{\sigma}_y = 3.5$ [illegible]

Problem II, Case I. Data points and computational results.



$\hat{\sigma}_x^2 = 7.27 \qquad \hat{\sigma}_x = 2.70$

$\hat{\sigma}_y^2 = 10.66 \qquad \hat{\sigma}_y = 3.27$

$\hat{\sigma}_{xy} = -4.0 \qquad \hat{\rho} = -.454$

$\hat{\sigma}_u^2 = 13.31 \qquad \hat{\sigma}_u = 3.64$

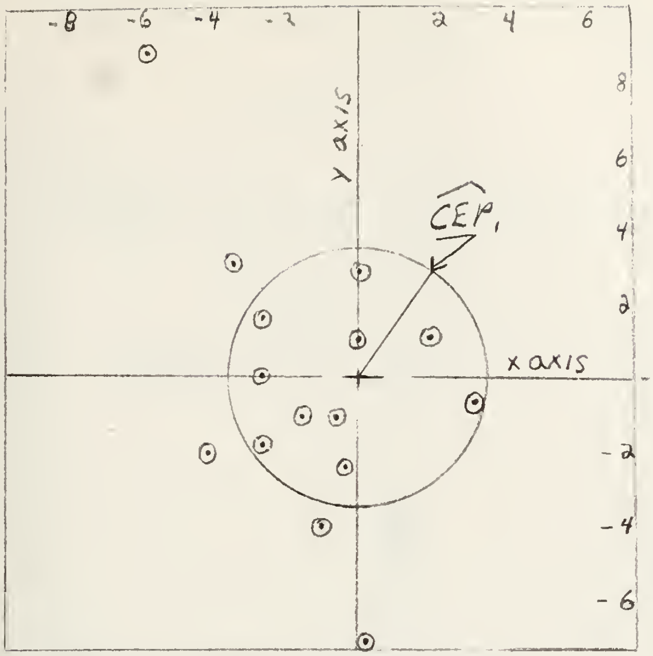
$\hat{\sigma}_v^2 = 4.62 \qquad \hat{\sigma}_v = 2.15$

Dependent Model	Independent Model
$C = .59$	$C^* = .825$
$K = .928$	$K^* = 1.07$
$\widehat{CEP}_1 = 3.37$	$\widehat{CEP}_1^* = 3.50$

Data Points in Problem II, N=10

Diagram 7

Problem II, Case II. Data points and computational results.



Data Points in Problem II, N=15

Diagram 8

$\hat{\sigma}_x^2 = 6.52$ $\hat{\sigma}_x = 2.56$

$\hat{\sigma}_y^2 = 11.7$ $\hat{\sigma}_y = 3.42$

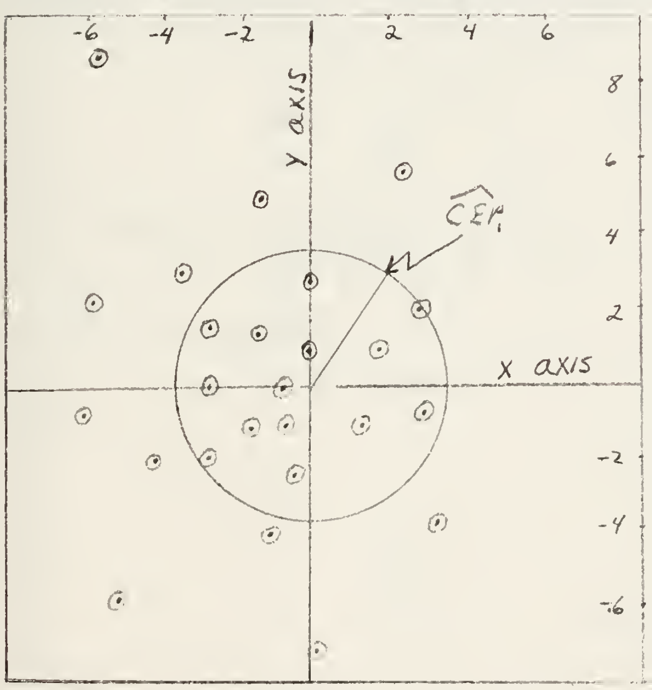
$\hat{\sigma}_{xy} = -2.21$ $\hat{\rho} = -.256$

$\hat{\sigma}_u^2 = 12.51$ $\hat{\sigma}_u = 3.54$

$\hat{\sigma}_v^2 = 5.71$ $\hat{\sigma}_v = 2.39$

Dependent Model	Independent Model
$C = .666$	$C^* = .712$
$K = .975$	$K^* = 1.026$
$\widehat{CEP}_1 = 3.45$	$\widehat{CEP}_1^* = 3.51$

Problem II, Case III. Data points and computational results.



Data Points in Problem II, N=25

Diagram 9

$\hat{\sigma}_x^2 = 8.92$ $\hat{\sigma}_x = 2.99$

$\hat{\sigma}_y^2 = 11.71$ $\hat{\sigma}_y = 3.42$

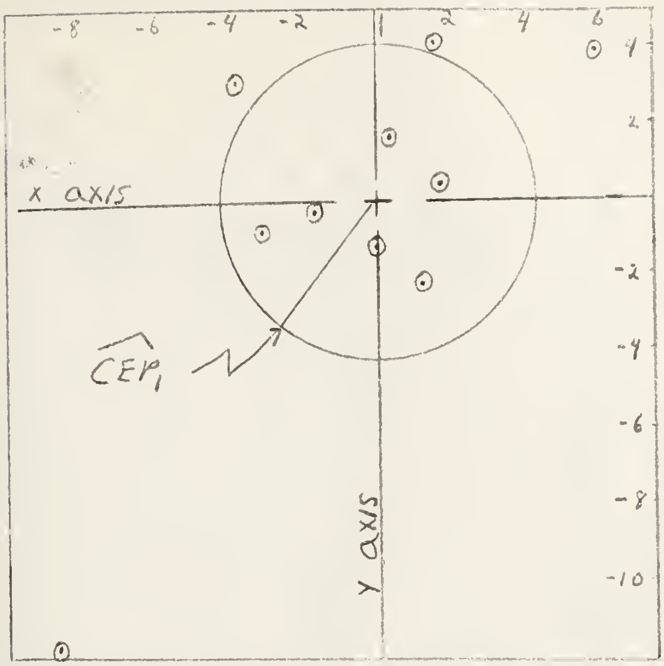
$\hat{\sigma}_{xy} = .325$ $\hat{\rho} = .031$

$\hat{\sigma}_u^2 = 11.75$ $\hat{\sigma}_u = 3.42$

$\hat{\sigma}_v^2 = 8.88$ $\hat{\sigma}_v = 2.98$

Dependent Model	Independent Model
$C = .872$	$C^* = .874$
$K = 1.10$	$K^* = 1.105$
$\widehat{CEP}_1 = 3.77$	$\widehat{CEP}_1^* = 3.78$

Problem III, Case 1. Data points and computational results.



Data Points in Problem III, N=10

Diagram 10

$\hat{\sigma}_x^2 = 14,2$ $\qquad \hat{\sigma}_x = 3,77$

$\hat{\sigma}_y^2 = 19,28$ $\qquad \hat{\sigma}_y = 4,40$

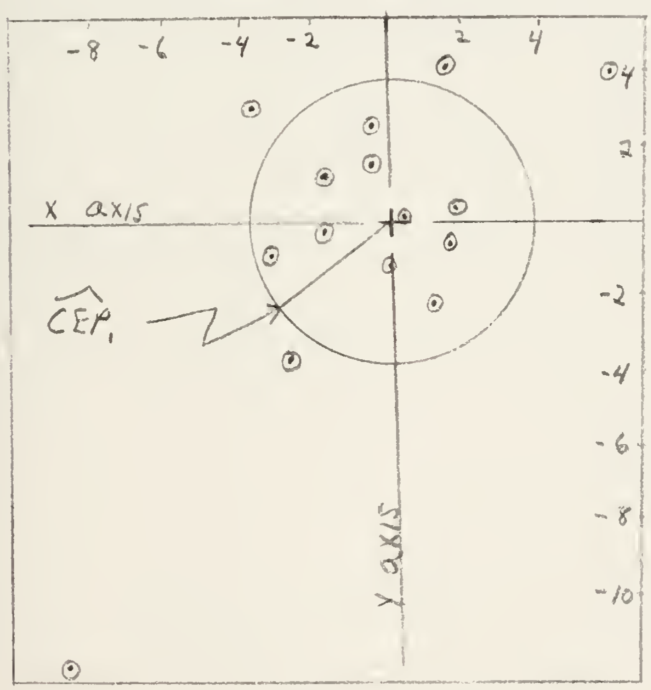
$\hat{\sigma}_{xy} = 12,2$ $\qquad \hat{\rho} = ,735$

$\hat{\sigma}_u^2 = 29.1$ $\qquad \hat{\sigma}_u = 5,4$

$\hat{\sigma}_v^2 = 4,35$ $\qquad \hat{\sigma}_v = 2,1$

Dependent Model	Independent Model
$C = ,387$	$C^* = ,855$
$K = ,80$	$K^* = 1,09$
$\widehat{CEP}_1 = 4,32$	$\widehat{CEP}_1^* = 4,80$

Problem III, Case II. Data points and computational results.



Data Points in Problem III, N=15

Diagram 11

$\hat{\sigma}_x^2 = 10.3$ $\qquad \hat{\sigma}_x = 3.21$

$\hat{\sigma}_y^2 = 14.33$ $\qquad \hat{\sigma}_y = 3.78$

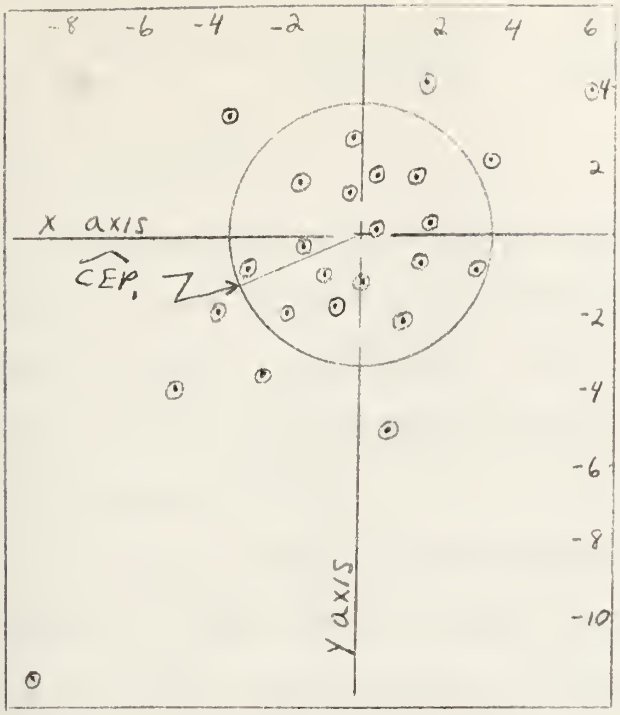
$\hat{\sigma}_{xy} = 8.43$ $\qquad \hat{\rho} = .695$

$\hat{\sigma}_u^2 = 20.9$ $\qquad \hat{\sigma}_u = 4.55$

$\hat{\sigma}_v^2 = 3.66$ $\qquad \hat{\sigma}_v = 1.91$

Dependent Model	Independent Model
$C = .42$	$C^* = ,85$
$K = ,82$	$K^* = 1,09$
$\widehat{CEP}_1 = 3,72$	$\widehat{CEP}_1^* = 4,10$

Problem III, Case III. Data points and computational results.



$\hat{\sigma}_x^2 = 8.82$ $\hat{\sigma}_x = 2.98$

$\hat{\sigma}_y^2 = 11.05$ $\hat{\sigma}_y = 3.33$

$\hat{\sigma}_{xy} = 6.54$ $\hat{\rho} = .66$

$\hat{\sigma}_u^2 = 16.6$ $\hat{\sigma}_u = 4.06$

$\hat{\sigma}_v^2 = 3.3$ $\hat{\sigma}_v = 1.82$

Dependent Model	Independent Model
$C = .446$	$C^* = .895$
$K = .835$	$K = 1.06$
$\widehat{CEP}_1 = 3.40$	$\widehat{CEP}^* = 3.53$

Data Points in Problem III, N=25

Diagram 12

SECTION III

DETERMINING THE CEP WHEN THE DENSITY FUNCTION IS CENTERED AT THE POINT (u_x, u_y): MODEL II

3.1 Introduction

The most important assumption made in model II is that if an infinite number of tests were conducted, the mean values of x and y would be u_x and u_y respectively. This means that the center of the bivariate normal density function is at some point (u_x, u_y) with respect to the target at (0,0).

If enough tests have been conducted to ascertain that this density function is offset from the target through the utilization of the estimators, then it may be possible to enter a spot $(-u_x, -u_y)$ to correct the fall of shot.

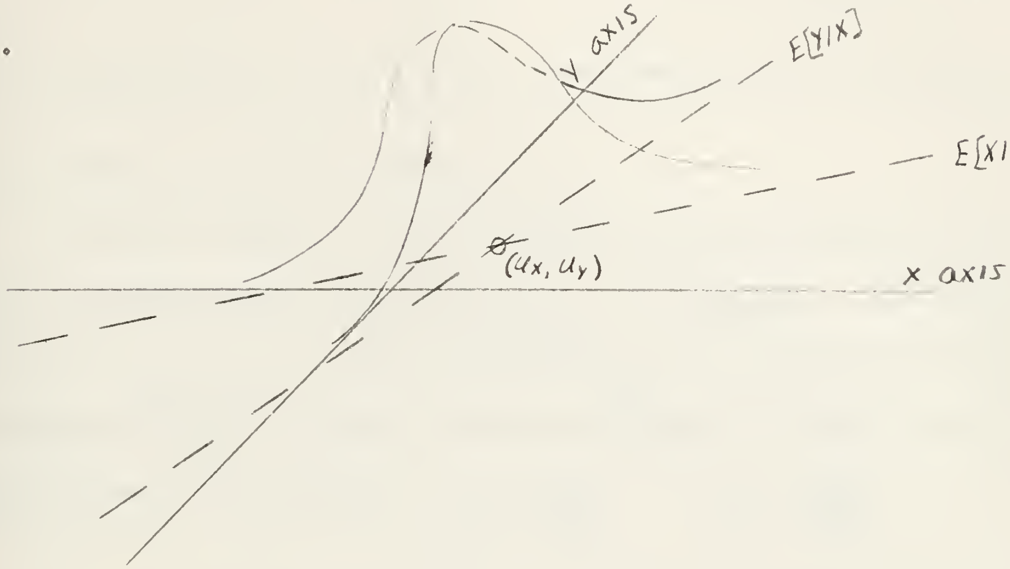
In this model the errors in the x and y directions are assumed to be non-independent but are distributed in accordance with the bivariate normal probability laws.

The probability that a point (x,y) whose coordinates are chosen at random will lie within a circle of radius $k\sigma_{max}$ with center at (u_x, u_y) is equal to

$$(2.1) \quad P\left[\sqrt{(X-u_x)^2+(Y-u_y)^2} \leq k\sigma_{max}\right] = \iint\limits_{\sqrt{(x-u_x)^2+(y-u_y)^2} \leq k\sigma_{max}} f_{X,Y}(x,y)\,dxdy =$$

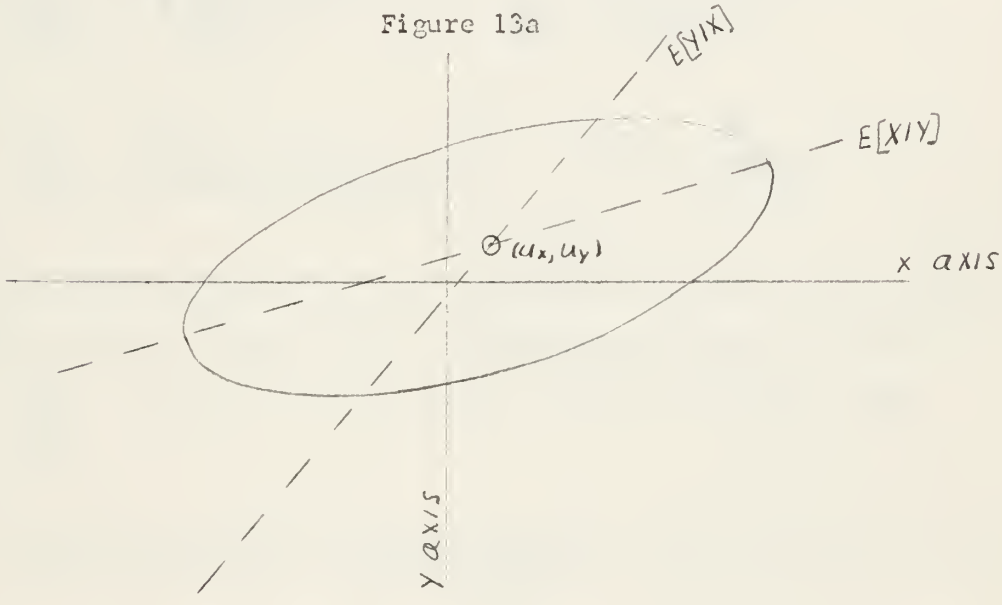
$$\frac{1}{2\pi|A^{-1}|^{1/2}} \iint\limits_{\sqrt{(x-u_x)^2+(y-u_y)^2} \leq k\sigma_{max}} \exp\left(-\tfrac{1}{2} Z'AZ\right)dxdy$$

In order to integrate over this form, it is necessary to first translate the axes before making the transformation because the density function is oriented along non-orthogonal lines away from the center of x,y coordinate system. This orientation is shown in figures 13a and 13b.



Three Dimensional Density Function with Center at (u_x, u_y) where $0 < \rho < 1$

Figure 13a



Ellipse Formed by a Plane Parallel to the x,y Plane Cutting the Density Function with Center at (u_x, u_y)

Figure 13b

The translation is made by subtracting the means (u_x, u_y) from their respective random variable X and Y. That is simply $(X - u_x)$ and $(Y - u_y)$ where in this case the matrix Z now becomes $Z = \begin{pmatrix} X - u_x \\ Y - u_y \end{pmatrix}$.

The transformation is then of the same form as the one in Section II.

3.2 Estimating the CEP Using Model II

The first step is to find estimators for $u_x, u_y, \sigma_x, \sigma_y$, and ρ from the n observed points $(x_1, y_1)(x_2, y_2) \ldots\ldots\ldots (x_n, y_n)$. This is done by first computing the sample means $\bar{x}, \bar{y}$ and then computing the sample variances $\hat{\sigma}_x^2, \hat{\sigma}_y^2$ the sample covariance $\hat{\sigma}_{xy}$ and the sample correlation coefficient $\hat{\rho}$ as follows:

$$\bar{x} = \frac{\sum x_i}{n} \qquad \bar{y} = \frac{\sum y_i}{n}$$

$$\hat{\sigma}_x^2 = \frac{\sum (x_i - x)^2}{n-1} \qquad \hat{\sigma}_y^2 = \frac{\sum (y_i - y)^2}{n-1}$$

$$\hat{\sigma}_{xy} = \frac{\sum (x_i - x)(y_i - y)}{n-1} \qquad \hat{\rho} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x \hat{\sigma}_y}$$

The transformed estimates of the variances are then computed using formulas (2.6). Table 1 or 2 is entered with $P(k,c) = .5$ and $c = \frac{\hat{\sigma}_v}{\hat{\sigma}_u}$ to find k. The estimate of the CEP $= \widehat{CEP}_2 = k\hat{\sigma}_u$.

3.3 Estimating the CEP Using the Assumption That the Errors in the x and y Directions Are Independent.

If it has been assumed that the errors in the x and y directions are independent, an estimate of the CEP can be obtained by using the estimators in model II except that the estimated variances $\hat{\sigma}_x^2$ and $\hat{\sigma}_y^2$

are used instead of the estimated transformed variances $\hat{V}_v^2$ and $\hat{V}_u^2$.

Then $c^* = \dfrac{\hat{V}_{min}}{\hat{V}_{max}}$ where $\begin{cases} \hat{V}_{MIN} = MIN(\hat{V}_x, \hat{V}_y) \\ \hat{V}_{max} = max(\hat{V}_x, \hat{V}_y) \end{cases}$

Table 1 or 2 is entered with $P(k^*, c^*)$ and c^* in order to find k^*.

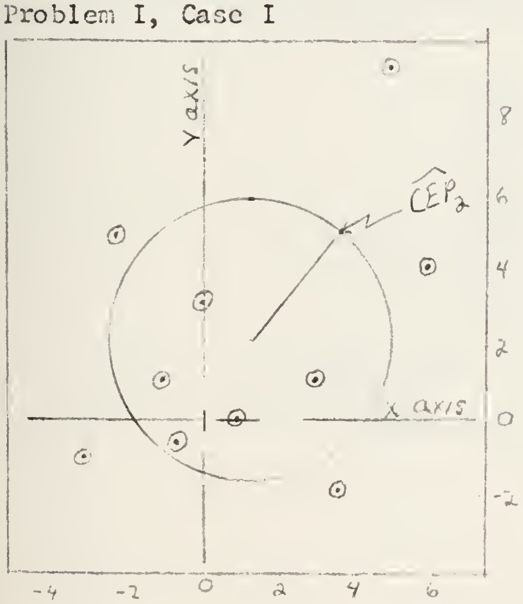
Then the estimate of the CEP is $\widehat{CEP}^*_2 = k^* \hat{V}_{max}$.

3.4 Comparison of Models I and II

If Model I is the true situation, then the estimator defined in Section II is the most efficient estimator. If the mean is not at (0,0), (Model II) then it still may be advantageous to use the estimate given for Model I if (u_x, u_y) is not too far away from the origin and if the sample size is small. This is because two degrees of freedom are lost in estimating (u_x, u_y). This problem is treated in Appendix B using the criterion of relative efficiency.

3.5 Problem Set

Problem I, Case I



Data Points in Problem I, N=10

Diagram 13

$\bar{X} = \dfrac{\sum X_i}{N} = 1.2$, $\bar{Y} = \dfrac{\sum Y_i}{N} = 2.0$

$\hat{V}_x^2 = \dfrac{\sum (X_i - \bar{X})^2}{N-1} = 9.48$, $\hat{V}_x = 3.08$

$\hat{V}_y^2 = \dfrac{\sum (Y_i - \bar{Y})^2}{N-1} = 10.7$, $\hat{V}_y = 3.27$

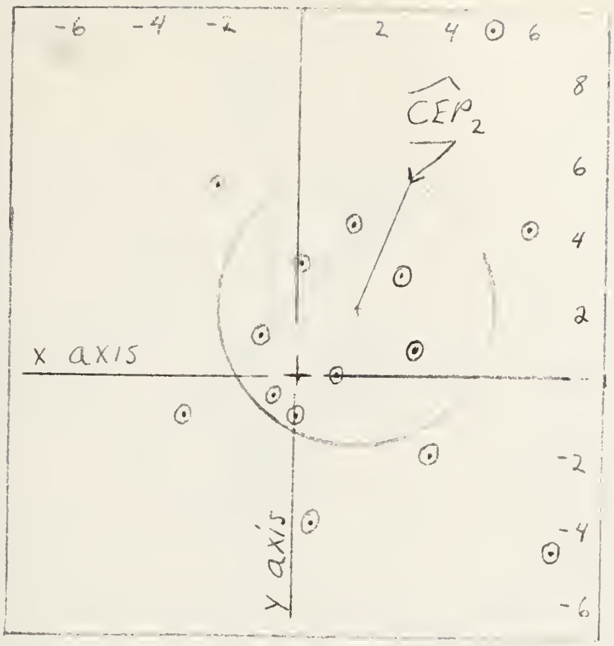
$\hat{V}_{xy} = \dfrac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{N-1} = 3.86$, $\hat{\rho} = \dfrac{\hat{V}_{xy}}{\hat{V}_x \hat{V}_y} = .38$

$\hat{V}_u^2 = 13.99$, $\hat{V}_u = 3.74$

$\hat{V}_v^2 = 6.16$, $\hat{V}_v = 2.48$

Dependent Model	Independent Model
$C = \dfrac{\hat{V}_v}{\hat{V}_u} = .664$	$C^* = \dfrac{\hat{V}_x}{\hat{V}_y} = .944$
$K = .973$	$K^* = 1.14$
$\widehat{CEP}_2 = 3.64$	$\widehat{CEP}^* = 3.72$

Problem I, Case II.



$\bar{X} = 1.53$ $\bar{Y} = 2.02$

$\hat{\sigma}_x^2 = 8.51$ $\hat{\sigma}_x = 2.92$

$\hat{\sigma}_y^2 = 13.6$ $\hat{\sigma}_y = 3.69$

$\hat{\sigma}_{xy} = .88$ $\hat{\rho} = .081$

$\hat{\sigma}_u^2 = 13.7$ $\hat{\sigma}_u = 3.70$

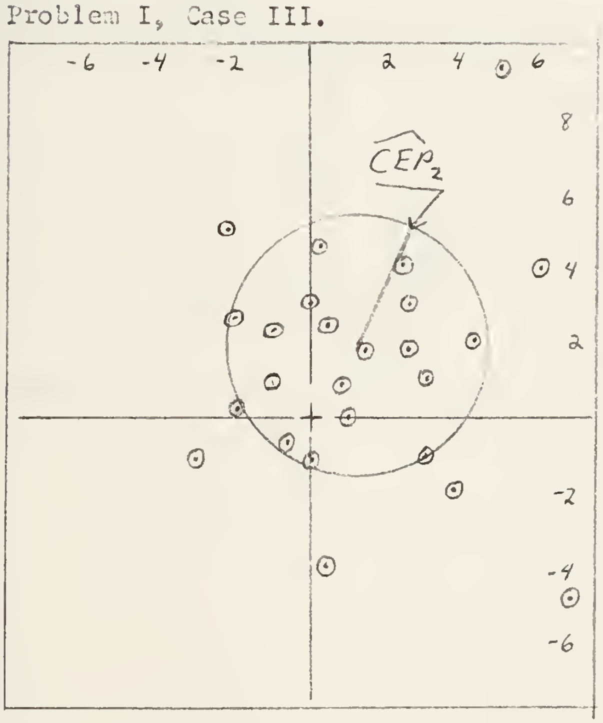
$\hat{\sigma}_v^2 = 8.4$ $\hat{\sigma}_v = 2.9$

Dependent Model	Independent Model
$C = .784$	$C^* = .791$
$K = 1.048$	$K^* = 1.052$
$\widehat{CEP}_2 = 3.87$	$\widehat{CEP}_2^* = 3.86$

Data Points in Problem I, N=15

Diagram 14

Problem I, Case III.



$\bar{X} = 1.6$ $\bar{Y} = 1.9$

$\hat{\sigma}_x^2 = 6.82$ $\hat{\sigma}_x = 2.62$

$\hat{\sigma}_y^2 = 8.63$ $\hat{\sigma}_y = 2.94$

$\hat{\sigma}_{xy} = -.374$ $\hat{\rho} = -.048$

$\hat{\sigma}_u^2 = 8.7$ $\hat{\sigma}_u = 2.96$

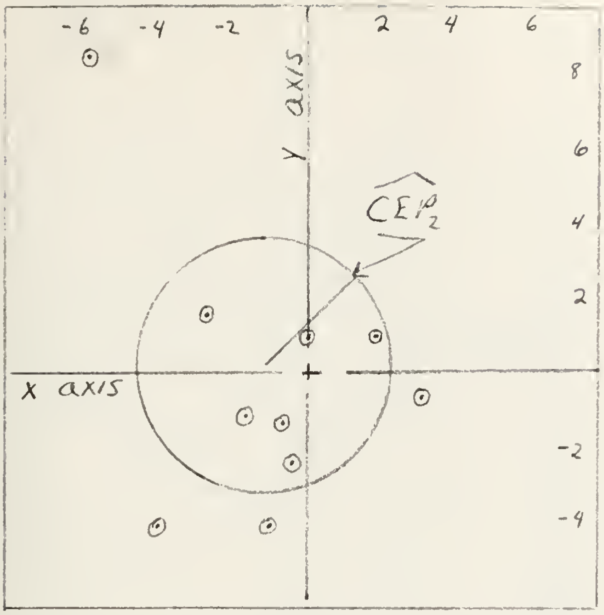
$\hat{\sigma}_v^2 = 6.8$ $\hat{\sigma}_v = 2.61$

Dependent Model	Independent Model
$C = .881$	$C^* = .891$
$K = 1.11$	$K^* = 1.11$
$\widehat{CEP}_2 = 3.28$	$\widehat{CEP}_2^* = 3.26$

Data Points in Problem II, N=25

Diagram 15

Problem II, Case I



$\bar{X} = -1.0$ $\bar{Y} = .1$

$\hat{\sigma}_x^2 = 6.97$ $\hat{\sigma}_x = 2.64$

$\hat{\sigma}_y^2 = 11.85$ $\hat{\sigma}_y = 3.45$

$\hat{\sigma}_{xy} = -5.7$ $\hat{\rho} = -.626$

$\hat{\sigma}_u^2 = 15.61$ $\hat{\sigma}_u = 3.95$

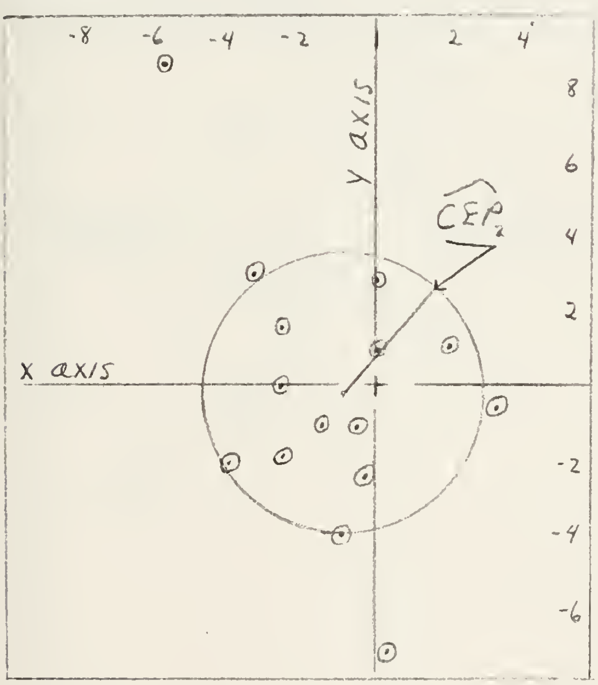
$\hat{\sigma}_v^2 = 3.21$ $\hat{\sigma}_v = 1.79$

Dependent Model	Independent Model
$C = .455$	$C^* = .765$
$K = .842$	$K^* = 1.03$
$\widehat{CEP}_2 = 3.33$	$\widehat{CEP}_2^* = 3.55$

Data Points in Problem II, N=10

Diagram 16

Problem II, Case II



$\bar{X} = -.9$ $\bar{Y} = -.2$

$\hat{\sigma}_x^2 = 6.16$ $\hat{\sigma}_x = 2.48$

$\hat{\sigma}_y^2 = 12.52$ $\hat{\sigma}_y = 3.54$

$\hat{\sigma}_{xy} = -3.47$ $\hat{\rho} = -.395$

$\hat{\sigma}_u^2 = 14.04$ $\hat{\sigma}_u = 3.74$

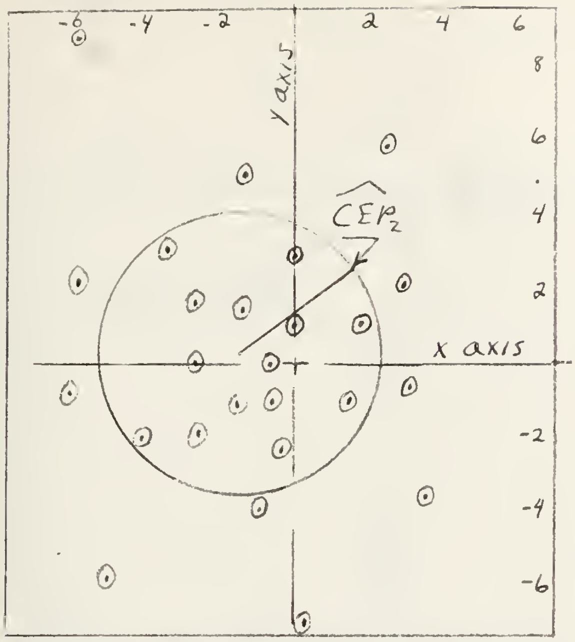
$\hat{\sigma}_v^2 = 4.64$ $\hat{\sigma}_v = 2.16$

Dependent Model	Independent Model
$C = .576$	$C^* = .7$
$K = .918$	$K^* = .996$
$\widehat{CEP}_2 = 3.39$	$\widehat{CEP}_2^* = 3.52$

Data Points in Problem II, N=15

Diagram 17

Problem II, Case III



$\bar{X} = -.9$ $\bar{Y} = .1$

$\hat{\sigma}_x^2 = 8.28$ $\hat{\sigma}_x = 2.88$

$\hat{\sigma}_y^2 = 12.2$ $\hat{\sigma}_y = 3.50$

$\hat{\sigma}_{xy} = -1.1$ $\hat{\rho} = -.107$

$\hat{\sigma}_u^2 = 12.25$ $\hat{\sigma}_u = 3.5$

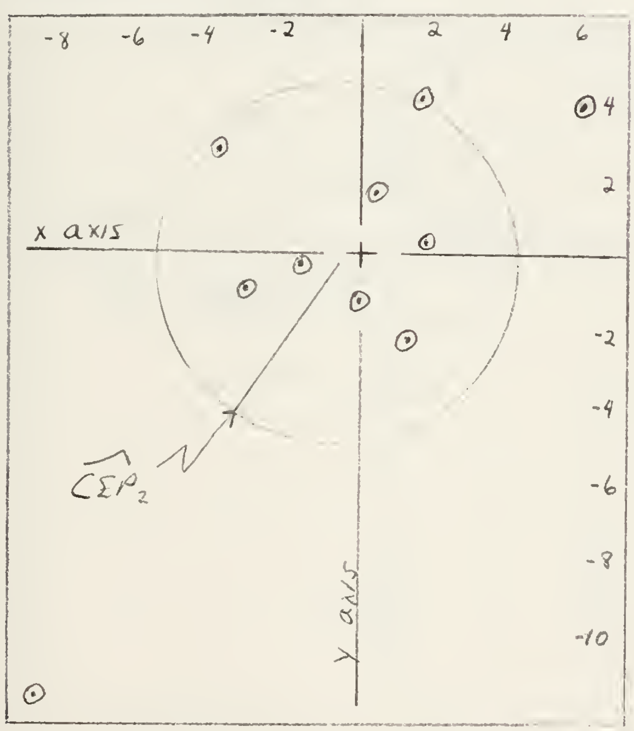
$\hat{\sigma}_v^2 = 7.98$ $\hat{\sigma}_v = 2.82$

Dependent Model	Independent Model
$C = .807$	$C^* = .822$
$K = 1.06$	$K^* = 1.07$
$\widehat{CEP}_2 = 3.71$	$\widehat{CEP}_2^* = 3.74$

Data Points in Problem II, N=25

Diagram 18

Problem III, Case I



$\bar{X} = -.6$ $\bar{Y} = -.3$

$\hat{\sigma}_x^2 = 15.4$ $\hat{\sigma}_x = 3.92$

$\hat{\sigma}_y^2 = 21.3$ $\hat{\sigma}_y = 4.61$

$\hat{\sigma}_{xy} = 11.3$ $\hat{\rho} = .625$

$\hat{\sigma}_u^2 = 30.1$ $\hat{\sigma}_u = 5.48$

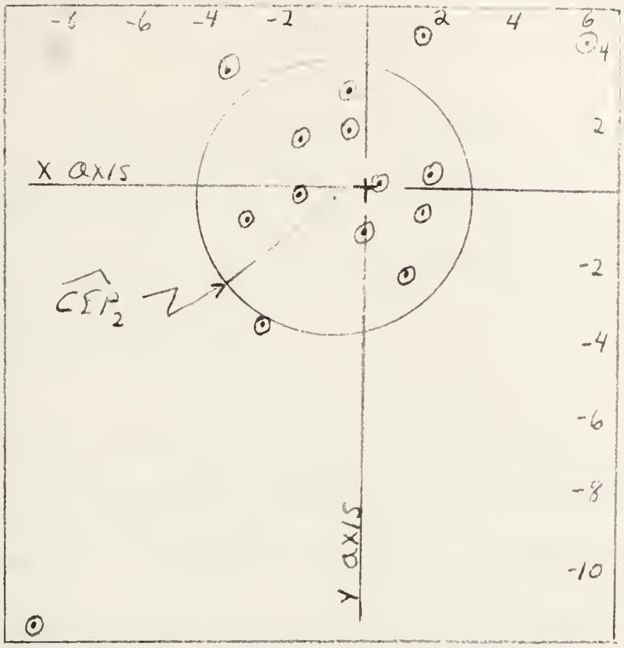
$\hat{\sigma}_v^2 = 6.65$ $\hat{\sigma}_v = 2.58$

Dependent Model	Independent Model
$C = .472$	$C^* = .848$
$K = .852$	$K^* = 1.087$
$\widehat{CEP}_2 = 4.66$	$\widehat{CEP}_2^* = 5.02$

Data Points in Problem III, N=10

Diagram 19

Problem III, Case II



Data Points in Problem III, N=15

Diagram 20

$\bar{X} = -.6$ $\bar{Y} = -.2$

$\hat{\sigma}_x^2 = 10.7$ $\hat{\sigma}_x = 3.27$

$\hat{\sigma}_y^2 = 15.35$ $\hat{\sigma}_y = 3.92$

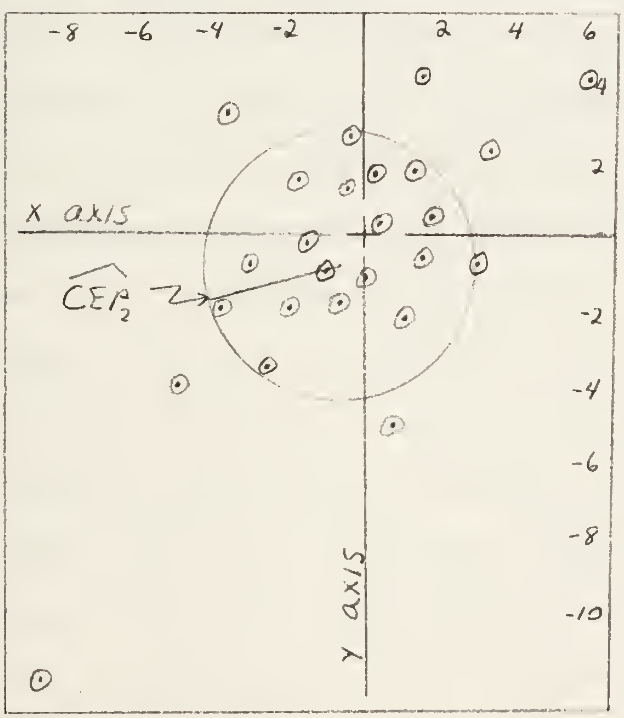
$\hat{\sigma}_{xy} = 11.6$ $\hat{\rho} = .903$

$\hat{\sigma}_u^2 = 24.82$ $\hat{\sigma}_u = 4.98$

$\hat{\sigma}_v^2 = 1.22$ $\hat{\sigma}_v = 1.11$

Dependent Model	Independent Model
$C = .222$	$C^* = .834$
$K = .715$	$K^* = 1.078$
$\widehat{CEP}_2 = 3.56$	$\widehat{CEP}_2^* = 4.21$

Problem III, Case III



Data Points in Problem III, N=25

Diagram 21

$\bar{X} = -.5$ $\bar{Y} = -.6$

$\hat{\sigma}_x^2 = 9.45$ $\hat{\sigma}_x = 3.08$

$\hat{\sigma}_y^2 = 10.7$ $\hat{\sigma}_y = 3.27$

$\hat{\sigma}_{xy} = 6.55$ $\hat{\rho} = .65$

$\hat{\sigma}_u^2 = 16.68$ $\hat{\sigma}_u = 4.21$

$\hat{\sigma}_v^2 = 3.48$ $\hat{\sigma}_v = 1.86$

Dependent Model	Independent Model
$C = .442$	$C^* = .942$
$K = .834$	$K^* - 1.14$
$\widehat{CEP}_2 = 3.52$	$\widehat{CEP}_2 = 3.72$

SECTION IV

REMOVING THE OUTLIER: MODEL III

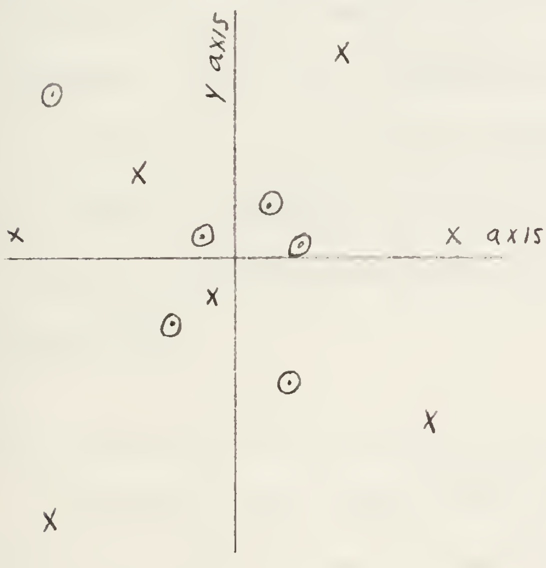
This model covers the problem of outliers and attempts to show some of the reasons for eliminating the outliers from consideration in the determination of the estimates as well as several methods for eliminating them.

4.1 Introduction to the Problem.

The general problem of removing outliers is related to the fact that it is desirable to obtain estimates of the parameters for the underlying bivariate density function which are not biased by observations from a distribution different from this underlying distribution. This in turn will yield more accurate estimates of the CEP. It is necessary to safeguard the estimate of the CEP from the ill effect of including information in the analysis that is not due to variations in the population of missiles, but is caused by some other factors such as weather or human errors. It is also possible that observations which have large deviations from the other observations may come from different distributions due to improvement in the missile design. This is especially true during the missile development stages where each succeeding missile has improved or different subsystem components than preceeding missiles. For example, an improved fuel may not be correctly compensated for in the missile guidance and fire control computers or a new type switch may not function quite as initially designed. The combination of changes may influence the range of the missile so that it lands farther from the target than predicted. If compensation is correctly made for the succeeding shot,

it seems reasonable that the observation for the first shot should not be included in the determination of estimates for the CEP.

Also, as improved subsystems are added to the missile, it is possible that the earlier missiles will not have the same density function as the later missiles and thus have a different CEP. In this case, it may become necessary to include only the later developed missiles in the determination of the CEP. Due to the fact that the missile development will be a continuing process with each missile slightly different than the preceeding one, it may not be easy to distinguish between these distributions. This is because both distributions will have some observations close to the target and others away from the target. The figures below may help to illustrate this point.



Observations from Two Distributions about the Target

x first population
0 second population

Figure 14a

Density Functions of Two Distributions about the Target

Figure 14b

It should be noted in figure 14b that distribution I has some probability of occuring in distribution II. If this probability is large, it may be extremely difficult to separate the two distributions. In fact, if it is desired to separate the two distributions, there is some probability that observations belonging to the underlying distribution under consideration will be removed along with the observations from the distribution that is not being considered. Thus one of the problems in removing outliers is to keep the probability, that the observations removed as outliers which do in fact belong to the underlying distribution, as low as possible. If this probability is small, it is possible that the observations belonging to the underlying distribution which are still removed will have such a low probability of occurence that their removal will still lead to a better estimate of the parameters. This may be especially true for small sample sizes where one sufficiently large or small observation can totally ruin an analysis of the data. Therefore, in order to eliminate an arbitrary result, it is necessary to establish some criteria for eliminating these outlying observations.

4.2 Criteria for Rejection of Outliers

Naturally shots which land at long distances from the target can be easily identified as wild shots or outliers with possible unknown errors. But as the observations move closer to the target, it becomes necessary to utilize some type of probabalistic consideration for the rejection of outlying observations. One way to approach a solution to this problem is to set it up as a hypothesis testing problem. On the basis of the observations (x_1,y_1)..........(x_n,y_n), a test is made of

the hypothesis that the observed point (x_i,y_i) belongs to the underlying distribution. The test is then conducted for each (x_i,y_i) for i = 1,2, ...n, one at a time. The alternate hypothesis is then that the observed point (x_i,y_i) does not belong to the underlying distribution but to some different distribution. This can be written as:

$$H_0: \ f_{X,Y}(x_i,y_i) = f_{X_o,Y_o}(x_i,y_i) \quad \text{for each } i = 1......n,$$

$$H_1: \ f_{X,Y}(x_i,y_i) \neq f_{X_o,Y_o}(x_i,y_i) \quad \text{where: } f_{X_o,Y_o}(x_i,y_i) \text{ is the}$$

true underlying distribution.

The probability of a Type I error will be called v where v is the probability of rejecting the hypothesis that the point (x_i,y_i) does belong to the underlying distribution when in fact it does belong to the underlying distribution. This can be expressed as

$$\text{Prob } [\text{Type I error}] = v$$

The probability of accepting the hypothesis that some point (x_i,y_i) does belong to the underlying distribution when the point does not belong to the underlying distribution and is called the Probability of a Type II error.

Thus the probability of the Type I error may be called the risk that the experimenter is willing to take in making a mistake by rejecting a point (x_i,y_i) as an outlier which does in fact belong to the underlying distribution even though the observed value does exceed some value specified by the criteria. Naturally, it is desirable to try to keep v small but if v is too small then the Type II error will increase and all outliers will be included in the determination of the parameters.

4.3 Method I For the Rejection of Outliers.

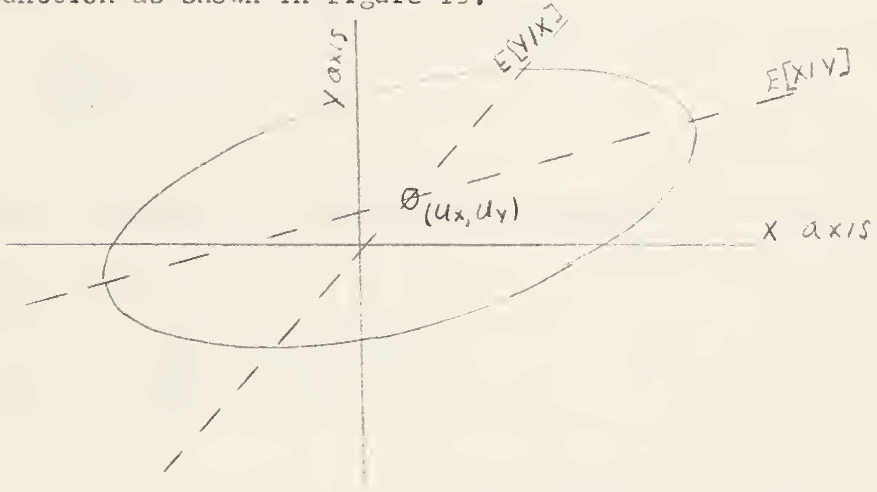
This method for the rejection of outliers is based on the probability that a random point (X,Y) will lie within the ellipse Z'AZ = k. Z'AZ is the matrix notation for the quadratic form of the dependent bivariate normally distributed random variables X and Y. That is,

$$(4.1) \quad Z'AZ = \frac{1}{1-\rho^2}\left[\left(\frac{X-u_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{X-u_x}{\sigma_x}\right)\left(\frac{Y-u_y}{\sigma_y}\right) + \left(\frac{Y-u_y}{\sigma_y}\right)^2\right] \quad \text{and}$$

k is defined by

$$(4.2) \qquad P\left(Z'AZ \leq k^2\right) = \iint_{Z'AZ \leq k^2} f_{X,Y}(x,y)\,dx\,dy = 1-v$$

Geometrically it is the probability that the point (X,Y) will lie inside the ellipse made by a plane parallel to the x,y axes cutting the density function as shown in figure 15.



Offset Ellipse Made by a Plane Parallel to x,y Axes
Cutting the Density Function

Figure 15

Due to the orientation of this density function, it is necessary to make the transformation to the orthogonal u,v coordinate system in order to integrate over this form. This transformation is made in the same

manner as in Sections II and III. The probability can now be expressed as

$$(4.3)\quad P(W'A^*W < k_1^2) = \iint_{W'A^*W<k_1^2} g_{U,V}(u,v)\,du\,dv = \frac{1}{2\pi\sigma_u\sigma_v}\iint_{W'A^*W<k_1^2} \exp(-\tfrac{1}{2}W'A^*W)\,du\,dv$$

where $W'A^*W = \frac{u^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}$

letting $T_2 = \frac{U^2}{\sigma_u^2} + \frac{V^2}{\sigma_v^2}$, (4.3) reduces to

$$(4.4)\quad P(T_2 < k_1^2) = \int_{T_2<k_1^2} \tfrac{1}{2}\exp(-\tfrac{1}{2}t)\ dt \qquad [3]$$

The random variable T has the Chi Squared distribution with two degrees of freedom. The above formula is a special case of the following result. If D_i are independent and normally distributed random variables with means u_i and variances , then

$$(4.5)\quad T_m = \sum_{i=1}^{m}\left(\frac{D_i - u_i}{\sigma_i}\right)^2$$

The degrees of freedom m is the number of independent terms in the sum. The density function of T_m is

$$(4.6)\quad f_{T_m}(t) = \frac{t^{(\frac{m}{2}-1)}\exp(-\frac{1}{2}t)}{(\frac{m}{2}-1)!\ 2^{(\frac{m}{2})}} \qquad t > 0$$

$$= \quad 0 \qquad t \leq 0$$

The areas under this density function are partially tabulated in Table 4. The desired percentage of the area under this curve is found by entering Table 4 with 1 - v and the degrees of freedom m.

The decision rule that is used for the elimination of outliers is to state that an observation is an outlier when

[3] "Introduction to the Theory of Statistics" by A. M. Mood /2/ of Rand Corporation.

$$(4.7) \quad k_1^2 < \frac{u_i^2}{\hat{\sigma}_u^2} + \frac{v_i^2}{\hat{\sigma}_v^2} = Z_i' A Z_i \quad \text{for } Z_i = \binom{x_i}{y_i}$$

4.4 Method II For the Rejection Of Outliers.

This method for the rejection of outliers is based on the probability that a random point (X,Y) will lie within a circle of radius $k\sigma_{max}$. Then, letting

$r = \left[(x - u_x)^2 + (y - u_y)^2\right]$, k is defined by

$$(4.8) \quad P\left\{\left[(X - u_x)^2 + (Y - u_y)^2\right] < k_\alpha \sigma_{max}\right\} = \int_{r < k_\alpha \sigma_{max}} f_{X,Y}(x,y)\,dxdy = 1-v$$

Geometrically, it is the probability that the random point (X,Y) will lie inside the circle imposed on the quadratic form made by a plane parallel to the x,y axes which cuts the density function as shown in Figure 16.

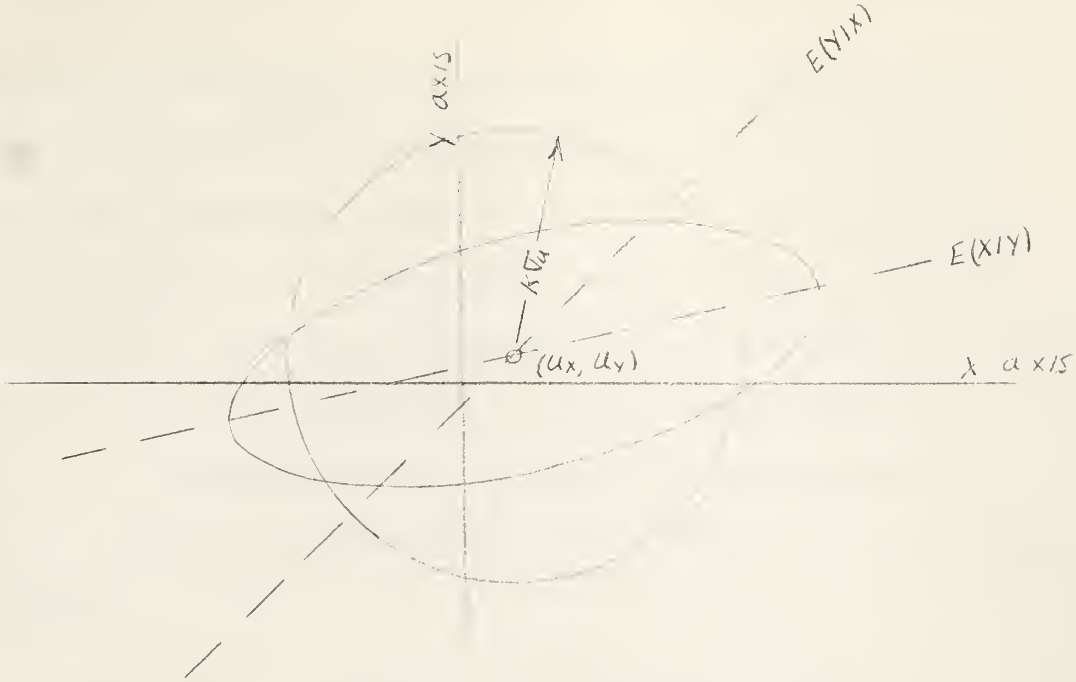


Illustration Showing the Circle of Interest
which is Imposed on the Ellipse Made by a Plane
Parallel to the x,y Plane Cutting the Density Function

Figure 16

Due to the orientation of this density function, it is also necessary to make the transformation to the orthogonal u,v coordinate system. The geometrical areas under consideration are shown below in figure 17 for this transformed density function.

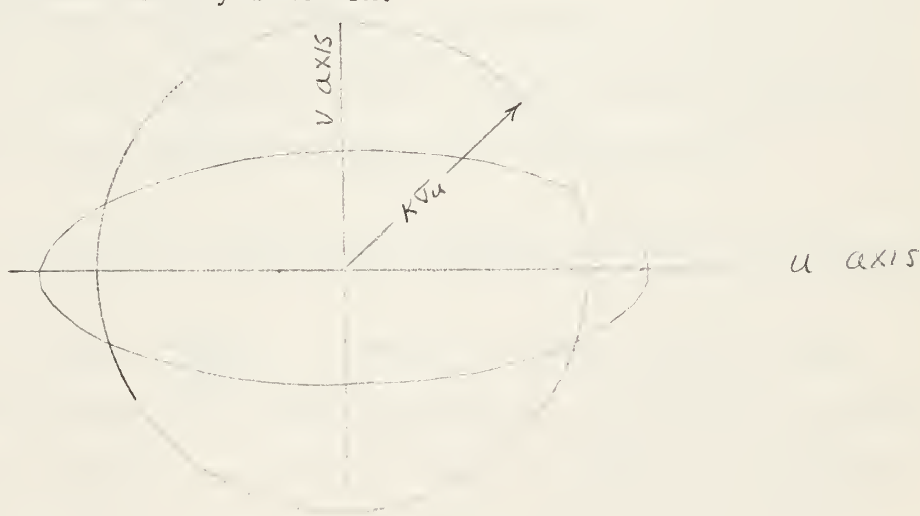


Illustration Showing the Circle of Interest
which is Imposed on the Ellipse made by a Plane
Parallel to the u,v Plane Cutting the Density Function

Figure 17

It should be noted that this method will reject points outside the circle but inside the ellipse which is estimated from the data points. Therefore, unless the variances are equal, this method will generally reject points farther from the target than method I, since some points on or near the major axis will be outside the circle as shown in figure 17. The circle is necessarily of smaller diameter than the major axis of the ellipse unless the variances are equal and then the circle and ellipse will be synonomous. This can be seen from the following inequality:

$$(4.9) \qquad \frac{x^2}{\nabla x^2} + \frac{y^2}{\nabla y^2} \geq \frac{x^2 + y^2}{\nabla^2_{max}} \qquad \text{where } \nabla^2_{max} = \max(\nabla x^2, \nabla y^2)$$

The probability that the point (U,V) in the transformed coordinate system will lie within the circle $\sqrt{U^2 + V^2} = k\nabla_u$ is expressed as

$$(4.10) \qquad P\left[\sqrt{U^2+V^2} \leq k\nabla_u\right] = 1-\gamma = \iint\limits_{\sqrt{u^2+v^2} \leq k\nabla u} g_{U,V}(u,v)\, du dv = P(k,c) \quad \text{where } c = \frac{\nabla_v}{\nabla_u}$$

This formula is the same general formula that was used for the determination of the CEP except that .5 has now been replaced by $(1-\gamma)$ in the range from $.5 \rightarrow 1$. The decision rule that is used for the elimination of outliers is to state that an observation is an outlier when

$(4.11) \quad W_i'A^*W_i = Z_i'AZ_i > k_2^2$ where k_2 is obtained from table 2 by entering with $1-\gamma$ and the value $c = \frac{\hat{\nabla} v}{\hat{\nabla} u}$. It should be noted that this value of k defines the radius of the circle centered at (u_x, u_y) which includes $(1-\gamma)100\%$ of the bivariate probability mass. The value of k obtained from method I defines the ellipse which includes $(1-\gamma)100\%$ of the bivariate probability mass.

4.5 Procedure for Removing Outliers Using Method I or Method II.

The actual procedure to remove the outliers differs from the discussion in Sections 4.3 and 4.4 in that the probability $1-\gamma$ is only exact if the true values of the parameters $u_x, u_y, \sigma_x^2, \sigma_y^2$ and $\sigma_{x,y}$ are used. It should be noted that both procedures substitute estimates of these parameters for the true values and therefore the probability of Type I error is not exactly equal to γ. The first step is to find estimators for $u_x, u_y, \sigma_x, \sigma_y$ and ρ from the n observed points $(x_1, y_1)\ldots$ $\ldots(x_n, y_n)$. This can be done using either model I or model II from Sections II and III respectively. The model used depends on which basic assumption is made about the true values of the means (u_x, u_y). If it is assumed that $u_x = u_y = 0$, then model I can be used. If it is assumed that $u_x \neq 0$, and/or $u_y \neq 0$, then model II can be used. Also, the criterion of relative efficiency can be used to determine whether model I or model II should be used. The estimates of the parameters $\bar{x}, \bar{y}, \hat{\sigma}_x^2, \hat{\sigma}_y^2, \hat{\sigma}_{xy}, \hat{\rho}, \hat{\sigma}_u^2, \hat{\sigma}_v^2$, are then computed by using the selected model. The estimated value of the matrix A is computed next using the above estimates.

$$\hat{A} = \frac{1}{1-\hat{\rho}^2}\begin{pmatrix} \frac{1}{\hat{\sigma}_x^2} & -\frac{\hat{\rho}}{\hat{\sigma}_x \hat{\sigma}_y} \\ -\frac{\hat{\rho}}{\hat{\sigma}_x \hat{\sigma}_y} & \frac{1}{\hat{\sigma}_y^2} \end{pmatrix} \tag{4.12}$$

Normally the value γ is predetermined by the experimenter and the outlier rejected on the basis of this value. It is advisable to delete the outliers one at a time until all of the data points are inside the region prescribed by the probability $1-\gamma$ and the method used. This is due to the fact that the estimated shape of the curve is dependent upon the data points and each deleted point will produce some change in the

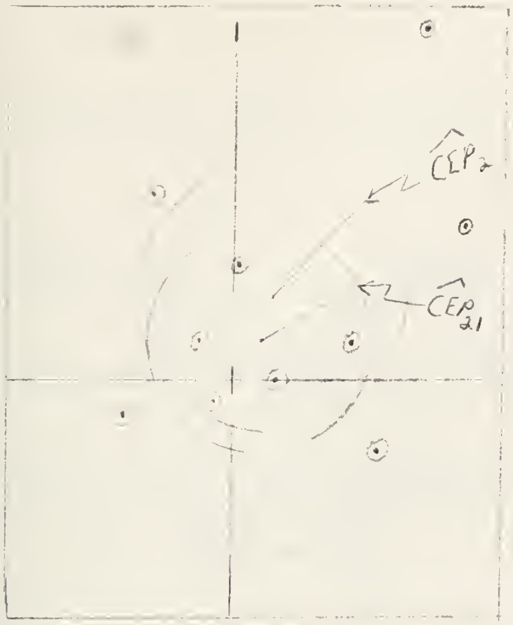
estimated shape of the density function. The first outlier is removed by investigating the points farthest from the estimated mean value and the point (x_i, y_i) is deleted whose estimated quadratic form $\underline{x}_i' \hat{A} \underline{x}_i$ is greater than k_1^2 (for method I) or k_2^2 (for method II). If there are two or more points which satisfy this requirement, the point is deleted first which has the greatest valued quadratic form.

It is then necessary to recompute the estimators and use the above procedures again, thus removing outliers one at a time, until there are no points left with estimated quadratic forms greater than k_i^2 (i=1or2). The final estimate of the CEP is then determined from the estimators derived using the data from the remaining observations. This estimate of the CEP will be referred to as $\widehat{CEP}_{2i}$ where the subscript i refers to the number of data points removed.

4.6 Information About the Problems.

In order to illustrate the above methods, the sample problems given in Section 2.5 were used. Model II was chosen arbitrarily for estimating the parameters for illustrative purposes. Both methods of rejecting outliers were set up for each problem case but instead of rejecting outliers with any specific probability, the tables were set up to show the probability that a specific data point could be rejected. This was done in order to compare the two methods.

Problem I, Case I. Data points and computational results.



Data Points in Problem I, N=10

Diagram 22

Step 1. In order to reduce computation, it is only necessary to find the maximum values of $Z_i'\hat{A}Z_i$ in each of the steps in removing the outliers.

$Z_{10}'\hat{A}Z_{10}=5.27$, $Z_2'\hat{A}Z_2=3.32$, $Z_1'\hat{A}Z_1=3.31$

1-V	Method 1 $\hat{k}_1$	Method 2 k_α^2
.[illegible]0	4.61	-
.95	5.[illegible]	4.5
.975	7.38	5.[illegible]

Conclusion: $Z'\hat{A}Z$ for point 10 is greatest and can be removed with 90% probability by Method 1 and 95% probability by Method.2.

The recomputed estimators, after deleting point 10 are then

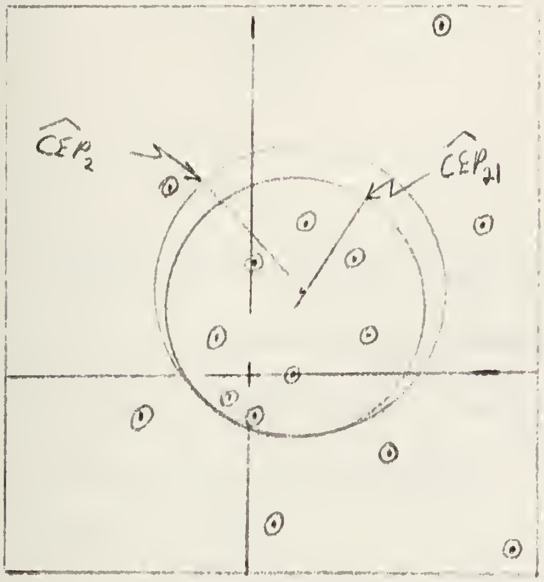
$\bar{x} = .[illegible]$, $\bar{y} = 1.2$, $\hat{V}_x = 2.[illegible]4$, $\hat{V}_y = 2.38$, $\hat{V}_{xy} = .41$, $\hat{\rho} = .06$, $\hat{V}_u = 2.9$, $\hat{V}_v = 2.4$

Dependent Model	Independent Model
c = .81	c* = .81
k = 1.06	k* = 1.06
$\widehat{CEP}_{21}=3.1$	$\widehat{CEP}^*_{21}=2.[illegible]$

Step 2. The procedure should now be continued with the 9 remaining data points to determine if any of the remaining data points can be removed

with a specified probability of 90% using method 1 or 95% probability using method 2. In this problem there are no more outliers.

Problem I, Case II. Data points and computational results.



Data Points in Problem 1, N=15

Diagram 23

Step 1.

$Z'_{10}\hat{A}Z_{10} = 5.04$, $Z'_{15}\hat{A}Z_{15} = 7.74$

1-V	Method 1 k_1^2	Method 2 k_2^2
.95	5.[illegible]	4.[illegible]3
.75	7.33	6.10
.90	[illegible].21	7.78

Conclusion: $Z'\hat{A}Z$ for data point 15 is greatest and can be removed with 97.5% probability by Method 1 and Method 2.

The recomputed estimators after deleting point 15 are then

$\bar{x} = 1.2$, $\bar{y} = 1.[illegible]$, $\hat{\sigma}_x = 2.62$, $\hat{\sigma}_y = 3.44$, $\hat{\sigma}_{xy} = 3.45$, $\hat{\sigma}_u = 3.60$, $\hat{\sigma}_v = 2.26$, $\hat{\rho} = .393$

Dependent Model	Independent Model
c = .[illegible]32	c* = .762
k = .954	k* = 1.054
$\widehat{CEP}_{21} = 3.41$	$\widehat{CEP}^*_{21} = 3.56$

Step 2. The procedure is now continued with the 14 remaining data points to determine if any of the remaining points can be removed with a specified probability of 95%.

$$Z'_{10}\hat{A}Z_{10} = 5.53$$

1-V	Method 1	Method 2
.95	5.[illegible]	4.41
.975	7.38	5.54

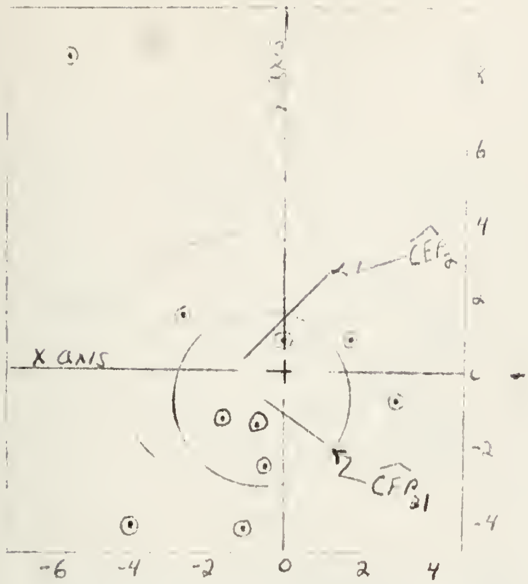
Conclusion: $Z'\hat{A}Z$ for data point 10 is the largest and can be removed with 95% probability by method 2 but would not be removed as an outlier by method 1. For purposes of illustration, this data point will be removed. The recomputed estimators after removing point 13 are

$\bar{x} = .3$, $\bar{y} = 1.7$, $\hat{V}_x = 2.47$, $\hat{V}_y = 2.71$, $\hat{V}_{xy} = .$[illegible]1, $\hat{V}_u = 2.79$, $\hat{V}_v =$ [illegible]$.37$, $\hat{\rho} = .136$

Dependent Model	Independent Model
$c = .85$	$c^* = .91$
$k = 1.09$	$k^* = 1.12$
$CEP_{22} = 3.04$	$CEP^*_{22} = 3.$[illegible]

Step 3. The procedure is again continued with the 13 remaining data points to determine if any of the remaining points can be removed with a specified probability of 95%. In this example there are no more outliers.

Problem II, Case I. Data points and computational results



Data Points in Problem II, N=10

Diagram 24

Step 1. $Z_1'\hat{\Lambda}Z_1 = 6.18$

1-V	Method 1	Method 2
.95	5.99	-
.975	7.38	5.22
.99	-	6.85

Conclusion: $Z'\hat{\Lambda}Z$ for point 1 is greatest and can be removed with 95% probability by method 1 and 97.5% probability by method 2.

The recomputed estimators after deleting point 1 are

$\bar{x} = -.5$, $\bar{y} = -.8$, $\hat{V}_x = 2.16$, $\hat{V}_y = 1.82$, $\hat{V}_{xy} = .79$, $\hat{\rho} = .201$, $\hat{V}_u = 2.23$, $\hat{V}_v = 1.72$

Dependent Model	Independent Model
c = .770	c* = .842
k = 1.039	k* = 1.083
$\widehat{CEP}_{21}$ = 2.31	$\widehat{CEP}^*_{21}$ = 2.34

Step 2. The procedure, using the 9 remaining points, does not reject any more data points in this problem.

Problem II, Case II. Data points and computational results.

Data Points in Problem II, N=15

Diagram 25

Step 1. $\hat{M'AZ}_1 = 7.33$

$1-\gamma$	Method 1 K_1^2	Method 2 K_2^2
.95	5.99	-
.975	7.38	-
.990	-	7.07
.995	-	8.30

Conclusion: $\hat{M'AZ}$ for point 1 is greatest and can be removed with 95% probability by method 1 and 99% probability by method 2.

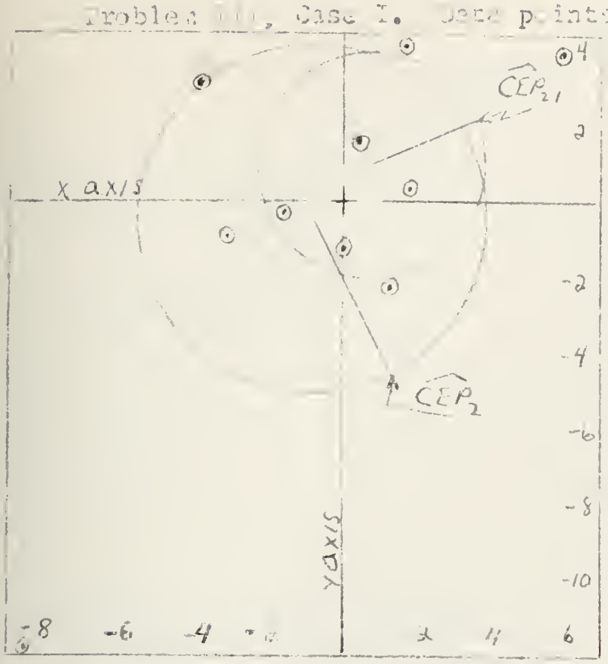
The recomputed estimators after deleting point 1 are

$\bar{x} = -.5$, $\bar{y} = -.7$, $\hat{V}_x = 2.16$, $\hat{V}_y = 2.68$, $\hat{V}_{xy} = -.2$, $\hat{\rho} = -.03$, $\hat{V}_u = 2.69$, $\hat{V}_v = 2.15$

Dependent Model	Independent Model
c = .803	c* = .805
k = 1.076	k* = 1.088
$\widehat{CEP}_{21} = 2.90$	$\widehat{CEP}^*_{21} = 2.92$

Step 2. The procedure using the 14 remaining points does not reject any more data points.

Problem III, Case I. Data points and computational results.



Data Points in Problem III, N=10
Diagram 26

Step 1. $\widehat{ZAZ}_1$ = [illegible]

1-V	Method 1	Method 2
.05	5.39	4.1
.975	7.33	5.33
.99	-	6.9

Conclusion: Z'AZ for point 1 is greatest and can be removed with 95% probability by method 1 and 97.5% probability by method 2.

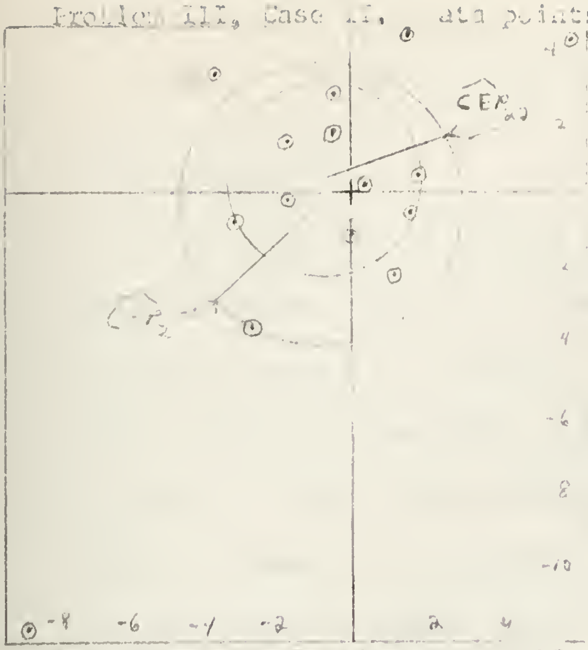
The recomputed estimators after deleting point 1 are

$\bar{x} = .3,\ \bar{y} = 1,\ \hat{V}_x = 2.8[illegible],\ \hat{V}_y = 2.36,\ \hat{V}_{xy} = 2.1[illegible],\ \hat{\rho} = .306,\ \hat{V}_u = 3.09,\ \hat{V}_v = 2.09$

Dependent Model	Independent Model
c = .667	c* = .816
k = [illegible]	k* = 1.1
$\widehat{CEP}_{21}$ = [illegible]	$\widehat{CEP}^*_{21}$ = 3.18

Step 2. The procedure using the 9 remaining points does not reject any more data points.

Problem III, Case II. Data points and computational results.



Data Points in Problem III, N=15

Diagram 27

Step 1. $Z_1'\hat{\Lambda}Z_1 = 8.9$, $Z_{10}'\hat{\Lambda}Z_{10} = 7.[illegible]$

1-γ	Method 1	Method 2
.95	5.99	-
.975	7.38	-
.99	-	6.12
.995	-	7.95

Conclusion: $Z'\hat{\Lambda}Z$ for point 1 is greatest and can be removed with greater than 97.5% probability for both methods. Also, Point 10 can be removed with 95% probability by method 1 and 99% probability by method 2. In this problem both points were removed in this step.

The recomputed estimators after deleting points 1 and 10 are

$\bar{x} = -.9$, $\bar{y} = .4$, $\hat{V}_x = 1.[illegible]1$, $\hat{V}_y = 2.20$, $\hat{V}_{xy} = .317$, $\hat{\rho} = .075$, $\hat{V}_u = 2.21$, $\hat{V}_v = 1.3[illegible]$

Dependent Model	Independent Model
$c = .957$	$c^* = .868$
$k = 1.099$	$k^* = 1.099$
$CEP_{22} = 2.42$	$CEP^*_{22} = 2.42$

The procedure, using the 13 remaining points, does not reject any more data points.

SECTION V

THE CONFIDENCE INTERVAL OF THE CEP

5.1 Introduction

The previously introduced estimates of the CEP are all called point estimates where the estimate of the CEP was defined by the locus of a point moving at a constant distance (the radius) from a fixed point (called the mean or (u_x, u_y)). This constant distance or radius is called the CEP. The confidence interval of the CEP attempts to give some measure of the possible error in the estimate of the CEP. The confidence is defined as the probability that the true value of the CEP lies in an interval between L_1 and L_2 where L_1 and L_2 are functions of the random observations (X_i, Y_i), i = 1,2,.....n. This expression in probability notation is

$$(5.1) \quad P\left[L_1(X_1 \ldots X_n, Y_1 \ldots Y_n) \leq CEP \leq L_2(X_1 \ldots X_n, Y_1 \ldots Y_n)\right] = 1 - \alpha$$

This interval estimate is a function of the confidence required, the number of observations, and the estimate of the standard deviation used.

5.2 Obtaining the Interval Estimate

In order to avoid lengthy computation in obtaining the interval estimate, it is assumed that the variances are equal. That is $\sigma_x^2 = \sigma_y^2 = \sigma^2$

The CEP was defined in Section 1.3 as being equal to $k\sigma$ where the value k is a function of the ratio of the variances and the probability that the mean centered circle contains 50% of the bivariate density mass.

4 See Section 2.1

Since the variances are assumed to be equal, the ratio of the variances is 1, and $P(k,1) = .5$, so that $k = 1.1774$ (from Table 1 with $c = 1$). Although the variances are assumed to be equal, the estimates of the variances are not necessarily equal.

The estimate of the standard deviation will be determined by the following two methods.

5.2.1 Determining the Confidence Interval, Method 1

In this method $\hat{\sigma}_z^2 = \max\left[\hat{\sigma}_x^2, \hat{\sigma}_y^2\right]$ will be selected to represent $\hat{\sigma}^2$. That is

$$(5.2) \qquad \hat{\sigma}_z^2 = \sum_{i=1}^{N} \frac{(Z_i - \bar{Z})^2}{n-1} = \max\left[\frac{\sum(X_i - \bar{X})^2}{n-1}, \frac{\sum(Y_i - \bar{Y})^2}{n-1}\right]$$

If $\hat{\sigma}_z^2$ is divided by the true value of the parameter and multiplied by $n-1$, this formula becomes

$$(5.3) \qquad \frac{(n-1)\,\hat{\sigma}_z^2}{\sigma^2} = \sum_{i=1}^{N} \frac{(Z_i - \bar{Z})^2}{\sigma^2}$$

Although the sum in (5.3) will not be an exact chi squared random variable because it is the maximum of two chi squared random variables, an approximate confidence interval can be obtained by treating (5.3) as though it were a chi squared random variable.[5]

The confidence interval defined by (5.1) thus becomes

$$(5.4) \qquad 1-\alpha = P\left(\chi^2_{N-1,\alpha/2} < \frac{(N-1)\,\hat{\sigma}_z^2}{\sigma^2} < \chi^2_{N-1,1-\alpha/2}\right)$$

$$= P\left(\frac{1}{\sqrt{\chi^2_{N-1,1-\alpha/2}}} < \frac{\sigma}{\sqrt{N-1}\;\hat{\sigma}_z} < \frac{1}{\sqrt{\chi^2_{N-1,\alpha/2}}}\right)$$

$$= P\left(\frac{1.1774\,\hat{\sigma}_z\sqrt{N-1}}{\sqrt{\chi^2_{N-1,1-\alpha/2}}} < CEP < \frac{1.1774\,\hat{\sigma}_z\sqrt{N-1}}{\sqrt{\chi^2_{N-1,\alpha/2}}}\right)$$

[5] See Section 4.3.

The values of $\chi^2_{N-1,1-\alpha/2}$ and $\chi^2_{N-1,\alpha/2}$ are obtained by entering table 4 with n-1 and either 1-α/2 or α/2 respectively.

5.2.2 Determining the Confidence Interval, Method 2.

The estimate of the variance in this method is the average of the two estimates. That is

$$(5.5) \quad \hat{\nabla}_Z^2 = \frac{\hat{\nabla}_X^2 + \hat{\nabla}_Y^2}{2} = \frac{1}{2}\left[\sum_{i=1}^{N}\frac{(X_i - \bar{X})^2}{N-1} + \sum_{j=1}^{N}\frac{(Y_j - \bar{Y})^2}{N-1}\right]$$

If (5.5) is divided by the true value of the parameter ∇^2 and multiplied by 2(n-1), the formula becomes

$$(5.6) \quad \frac{2(N-1)}{\nabla^2}\left(\frac{\hat{\nabla}_X^2 + \hat{\nabla}_Y^2}{2}\right) = \frac{1}{\nabla^2}\left[\sum(X_i - \bar{X})^2 + \sum(Y_j - \bar{Y})^2\right]$$

where x_i and y_i are normally and independently distributed and $\bar{x}$ and $\bar{y}$ are the sample means. This formula can be reduced by letting the values of i range from 1 to n and the values of j range from n + 1 to 2n. Then the formula becomes

$$(5.7) \quad \frac{2(N-1)}{\nabla^2}\hat{\nabla}_Z^2 = \sum_{K=1}^{2N}\frac{(Z_K - \bar{Z}_K)^2}{\nabla^2}$$

where $z_k = x_k$ for k = 1........n and $z_k = y_k$ for k = n + 1.......2n and there are 2(n-1) squares in the sum. Thus (5.7) has a χ^2 distribution with 2(n-1) degrees of freedom by the definition given in (4.5). The interval estimate is determined in the same way as in (5.4) and the formula becomes

$$(5.8) \quad P\left(\frac{1.1774\,\hat{\nabla}_{Z'}\sqrt{2(N-1)}}{\sqrt{\chi^2_{2(N-1),1-\alpha/2}}} < CEP < \frac{1.1774\,\hat{\nabla}_{Z'}\sqrt{2(N-1)}}{\sqrt{\chi^2_{2(N-1),\alpha/2}}}\right) = 1 - \alpha$$

The values of $\chi^2_{2(N-1),\alpha/2}$ and $\chi^2_{2(N-1),1-\alpha/2}$ are obtained by entering table 4 with 2(n-1) and either 1-α/2 or α/2 respectively. It should be noted that this method of interval estimation is not as conservative as method 1 because the average value is always less than the maximum of (σ_x^2, σ_y^2). Therefore, this interval estimate will be smaller.

5.3 Illustration

The estimates of the confidence interval of the CEP used in the following illustrations were obtained with the data from Section 3 and (1-α) = .95. A comparison is made between method 1 and method 2 as well as a variation of the two methods where the dependent estimate of the CEP ($\widehat{CEP}_2 = k\,\hat{\sigma}_u$) was substituted for $1.1774\,\hat{\sigma}_z$. It should be emphasized that none of the distribution theory used in method 2 holds when $\widehat{CEP}_2$ is used for $k\,\hat{\sigma}_u$. Therefore, it is hard to get a mathematically meaningful comparison between these methods.

Table 6 shows the various estimates of the CEP. The best estimate of the CEP is most likely to be $\widehat{CEP}_2$ due to the basic assumptions of dependence and unequal variances. The estimate of $1.1774\,\hat{\sigma}_{max}$ is the largest estimate of the CEP and therefore the most conservative estimate of the CEP.

Table b				
Problem	$1.1774\,\hat{\nabla}_{max}$	$1.1774\,\hat{\nabla}_{avg}$	$\widehat{CEP}_2$	Number of Observations
1				
Case 1	3.54	3.72	3.64	10
Case 2	4.34	3.89	3.97	15
Case 3	3.45	3.26	3.28	25
2				
Case 1	4.05	3.57	3.33	10
Case 2	4.15	3.53	3.39	15
Case 3	4.11	3.74	3.71	25
3				
Case 1	5.42	5.00	4.66	10
Case 2	4.60	4.21	3.36	15
Case 3	3.84	3.73	3.52	25

Tables c and d show the upper and lower bounds of the confidence interval estimates.

Table c				
$L_1(x_1 \ldots x_n,\ y_1 \ldots y_n)$ = Lower Bound of the Confidence Interval Estimate				
Problem	Method 1		Method 2	
	$\frac{1.1774\,\hat{\nabla}_{max}\sqrt{n-1}}{\sqrt{\chi^2_{N-1,1-\alpha/2}}}$	$\frac{\widehat{CEP}_2\sqrt{n-1}}{\sqrt{\chi^2_{N-1,1-\alpha/2}}}$	$\frac{1.1774\,\hat{\nabla}_{avg}\sqrt{2(n-1)}}{\sqrt{\chi^2_{2(N-1),1-\alpha/2}}}$	$\frac{\widehat{CEP}_2\sqrt{2(n-1)}}{\sqrt{\chi^2_{2(N-1),1-\alpha/2}}}$
1				
Case 1	2.65	2.51	2.82	2.76
Case 2	3.10	2.52	3.08	3.07
Case 3	2.70	2.56	2.72	2.74
2				
Case 1	2.80	2.30	2.70	2.52
Case 2	3.03	2.47	2.80	2.69
Case 3	3.22	2.90	3.12	3.10
3				
Case 1	3.74	3.22	3.79	3.53
Case 2	3.36	2.60	3.35	2.83
Case 3	3.00	2.75	3.11	2.94

Table d				
$L_2\ (x_1 \ldots x_n,\ y_1 \ldots y_n)$ = Upper Bounds of the Confidence Interval Estimate				
Problem	Method 1		Method 2	
	$\frac{1.1774 \hat{V}_{max} \sqrt{n-1}}{\sqrt{\chi^2_{N-1,\alpha/2}}}$	$\frac{\widehat{CEP}_2 \sqrt{n-1}}{\sqrt{\chi^2_{N-1,\alpha/2}}}$	$\frac{1.1774 \hat{V}_{avg} \sqrt{2(n-1)}}{\sqrt{\chi^2_{2(N-1),\alpha/2}}}$	$\frac{\widehat{CEP}_2 \sqrt{2(n-1)}}{\sqrt{\chi^2_{2(N-1),\alpha/2}}}$
1				
Case 1	7.00	6.65	5.48	5.36
Case 2	6.80	6.10	5.12	5.10
Case 3	4.80	4.56	4.16	4.20
2				
Case 1	7.40	6.10	5.26	4.[illegible]0
Case 2	6.53	5.34	4.73	4.60
Case 3	5.73	5.16	4.78	4.75
3				
Case 1	[illegible].00	[illegible].50	7.36	[illegible].86
Case 2	7.24	5.60	5.70	4.82
Case 3	5.35	4.50	4.76	4.50

It is noted that the lower bound estimates are for all practical purposes the same for both methods, with the average difference being only .31. However, the upper bound differences show that method 1 gives a greater estimate with the average difference being 1.59. The lengths of the confidence intervals are compared in Table e below.

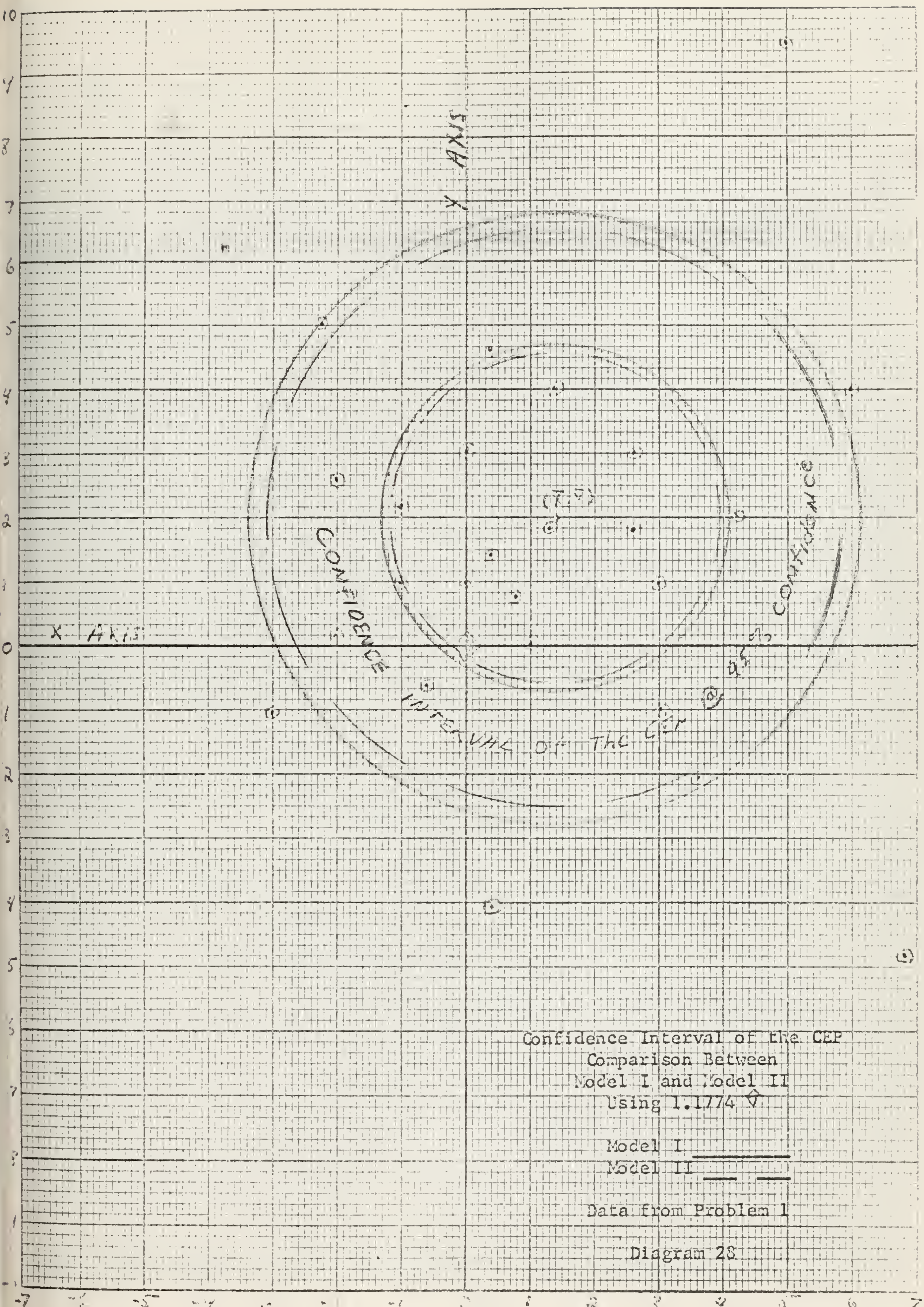
Table e						
Length of the Confidence Interval (Upper bound - Lower bound)						
	With $\hat{\nabla}_2$		1-2	With $\widehat{CEP}_2$		1-2
Problem	Method 1	Method 2	Difference	Method 1	Method 2	Difference
1						
Case 1	4.35	2.66	1.69	4.14	2.60	1.54
Case 2	3.65	2.04	1.61	3.28	2.03	1.25
Case 3	2.10	1.44	.66	2.00	1.46	.54
2						
Case 1	4.60	2.56	2.05	3.80	2.38	1.42
Case 2	3.50	2.00	1.50	2.87	1.91	.96
Case 3	2.51	1.66	.85	2.26	1.65	.61
3						
Case 1	6.16	3.58	2.58	5.28	3.33	1.95
Case 2	3.88	2.35	1.53	3.00	1.99	1.01
Case 3	2.35	1.65	.70	2.15	1.56	.59
Average difference			1.46			1.10

It should be noted that the confidence interval becomes smaller as the number of observations increase. This implies that the true value of the CEP is more likely to be within a smaller interval as the number of observations increase.

Diagrams 28,29 and 30 show the confidence interval using the different estimates. The confidence intervals were obtained by using the data from case III of each of the problems.

5.4 Conclusions

Method 1, using $1.1774\hat{\nabla}_{max}$ produces the largest estimates and therefore is the most conservative estimate of the confidence interval. However, $\widehat{CEP}_2$ and $1.1774\hat{\nabla}_{avg}$ are likely to be better estimates of the CEP and therefore method 2 or the approximate interval using the dependent estimate $\widehat{CEP}_2$ may be the best method for estimating the confidence interval. An analysis of actual missile data should give a more realistic insight into the best choice of methods to use in estimating the confidence interval. In order to come to any definite conclusions about the different methods, some comprehensive distribution theory problems must be solved.

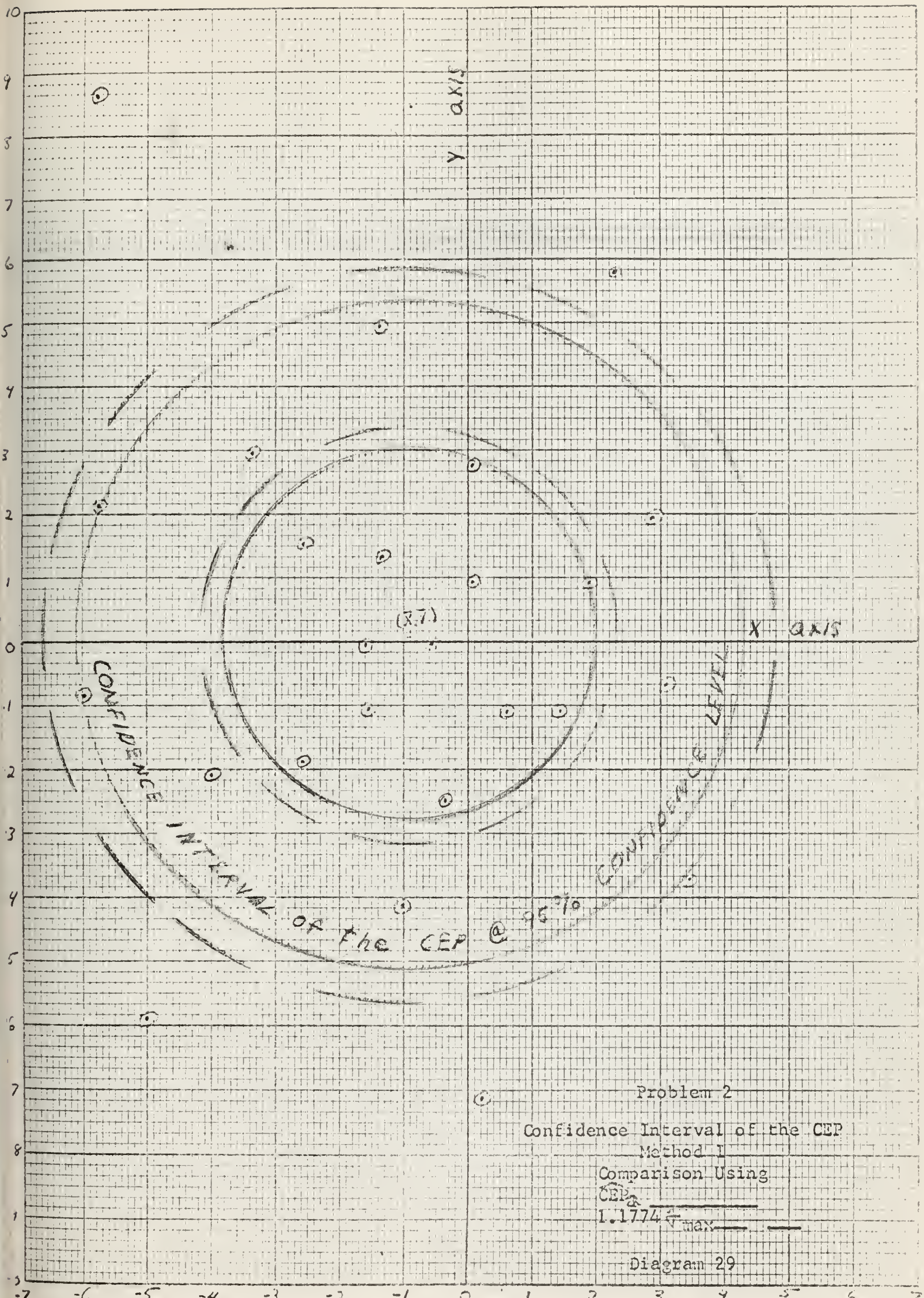


Confidence Interval of the CEP
Comparison Between
Model I and Model II
Using 1.1774 $\hat{\sigma}$

Model I ————
Model II — —

Data from Problem 1

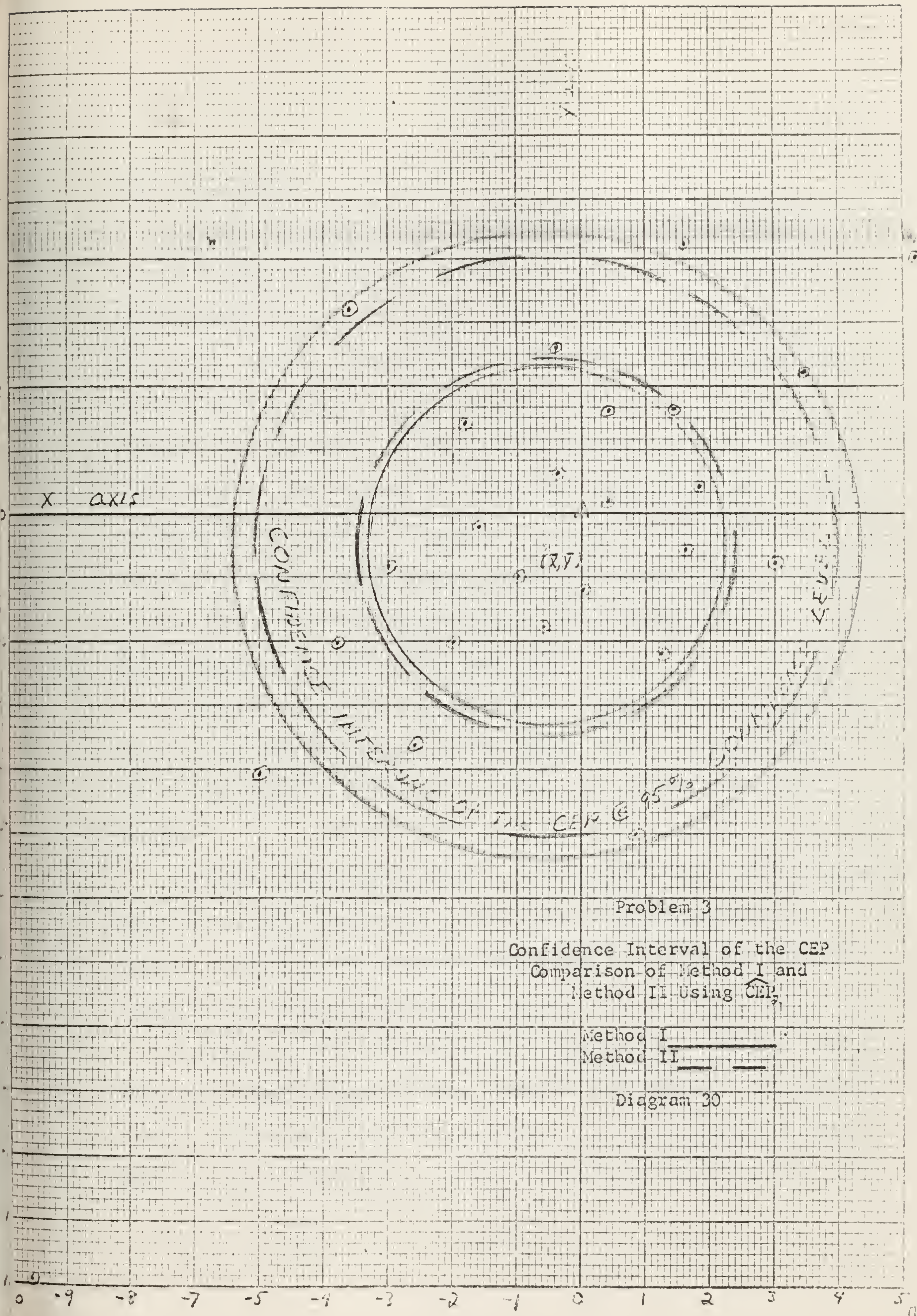
Diagram 28



Problem 2

Confidence Interval of the CEP
Method I
Comparison Using
$\widehat{CEP}_a$ ________
$1.1774\,\hat{\sigma}_{max}$ ____ ____

Diagram 29



Problem 3

Confidence Interval of the CEP
Comparison of Method I and
Method II Using $\widehat{CEP}_2$

Method I ———
Method II — —

Diagram 30

SECTION VI

SUMMARY

6.1 Introduction

The previous sections have been concerned with the development of different types of models and methods for estimating the radius of the mean centered circle which includes 50% of a bivariate probability mass. This section summarizes the different models and methods used in the previous sections, and includes an analysis of the results obtained from problems. Although the sample problems do not represent actual missile test results, an attempt has been made to make the data as realistic as possible. Therefore an analysis of the problems should show certain relationships between the models used to estimate the CEP that would also apply to actual missile test data.

6.2 Comparison Of Model I With Model II.

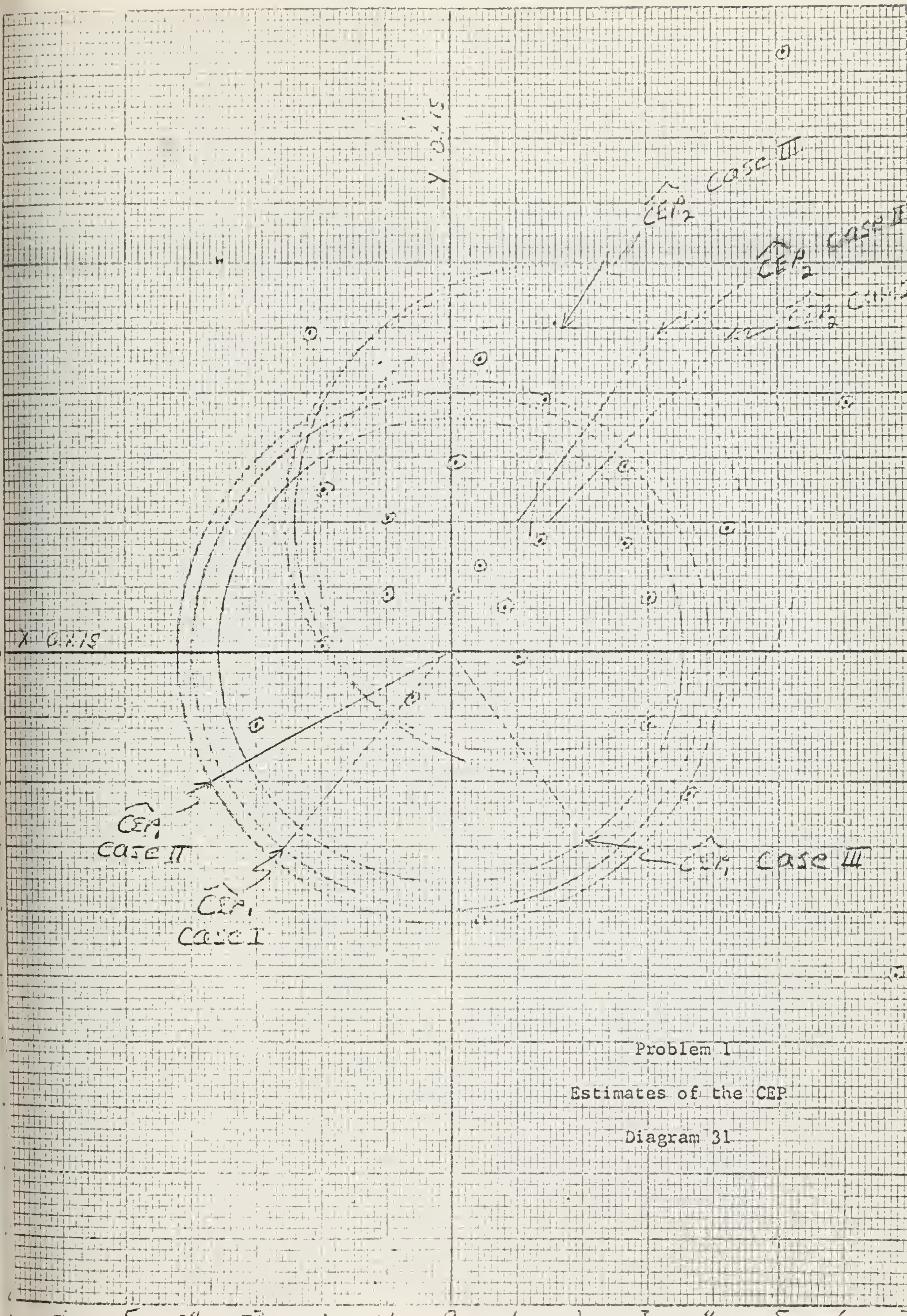
The basic underlying assumption made in Model I was that the true value of the mean was located at the target, (0,0). Therefore, the CEP in this model is defined as the radius of a circle around the target.

The basic underlying assumption made in Model II was that the true value of the mean was located at some point (u_x, u_y) away from the target. Therefore, the estimated CEP for this model is the radius of a circle with center at some point, $(\bar{x}, \bar{y})$.

A comparison of the estimate of the correlation coefficient shows that they change in much the same manner in both models. As suspected, a major difference between these models is in the location of $\bar{x}$ and $\bar{y}$. This is shown in Diagrams 31, 32, and 33 which illustrate the estimates

of the CEP. The estimate of the CEP for problems 2 and 3 is practically the same in all three cases. Therefore, when the center of the distribution is near the target, there is little practical difference between the two models. However, in problem 1, the distribution of data points is around some point $(\bar{x}, \bar{y})$ away from the center. If the procedure given in Appendix B is used to estimate the ratio function, then the values obtained indicate that $\widehat{CEP}_2$ gives the best estimate of the CEP for a sample size of 10 in problem 1. Also, as the sample size increases the ratio function increases, thus $\widehat{CEP}_2$ is also the best estimate for $n > 10$. The values of R.F. obtained for problems 2 and 3 show a preference for Model I for small sample sizes and are very close to 1 for large sample sizes and therefore either estimate may be used.

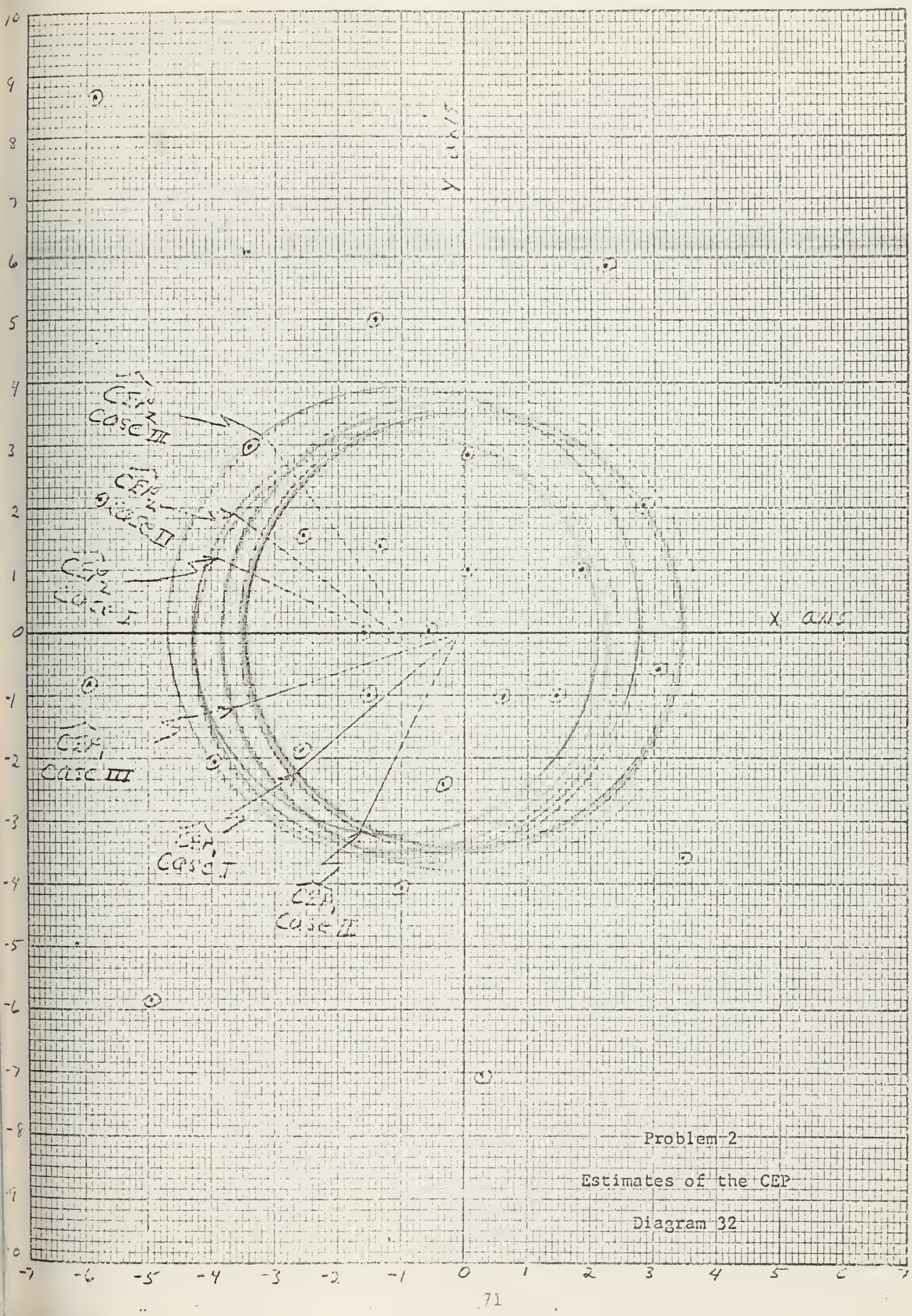
These problems tend to substantiate the fact that the procedure of Model II is superior to the procedure of Model I in large sample sizes. They also suggest that if Model I is used in analyzing a small number of observations, it might be advantageous to check the assumption of mean (0,0) by computing the sample means.



Problem I

Estimates of the CEP

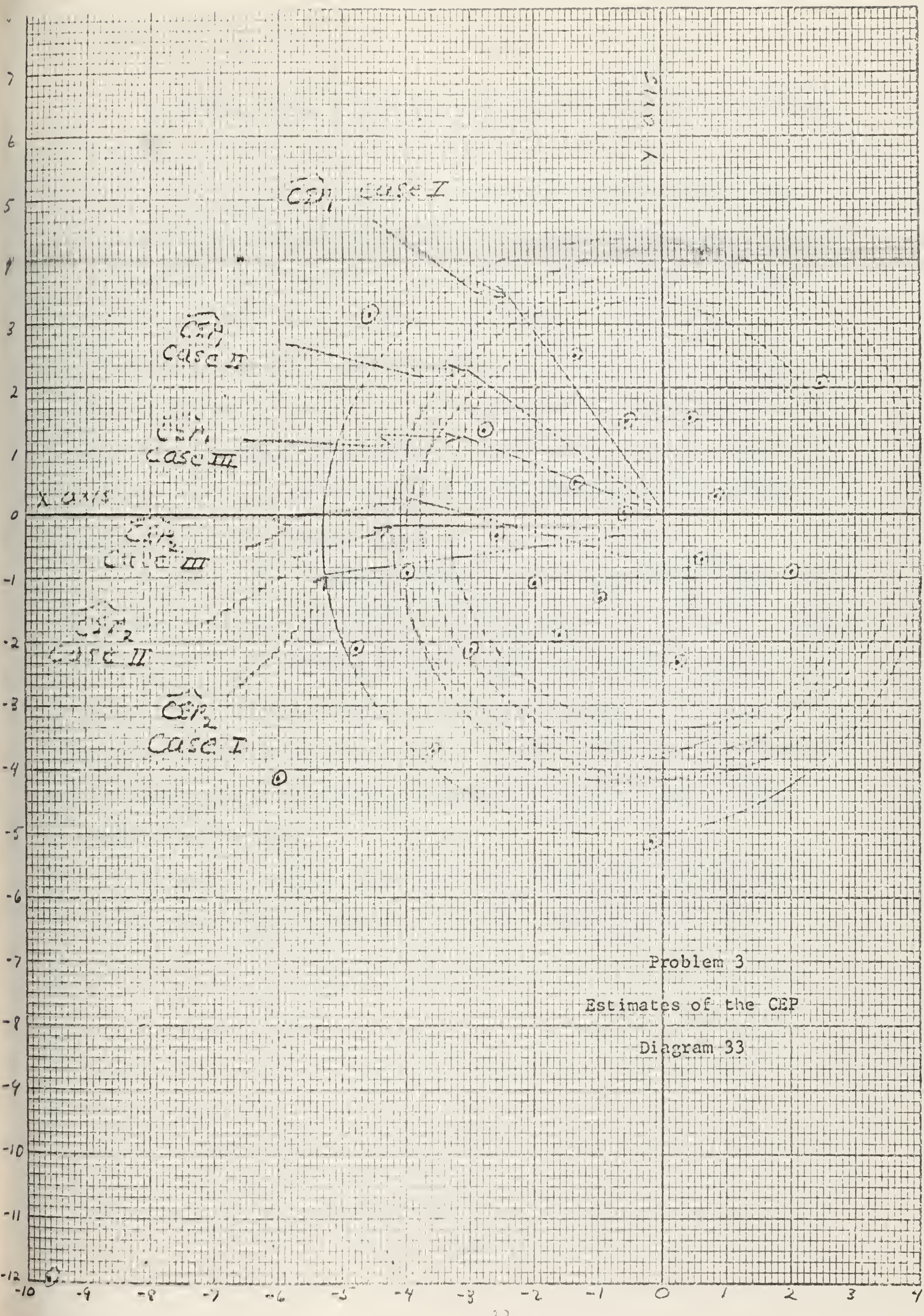
Diagram 31



Problem 2

Estimates of the CEP

Diagram 32



Problem 3

Estimates of the CEP

Diagram 33

6.3 Comparison Of The Independent And Dependent Methods Of Estimating The CEP.

In the introduction to the problem of estimating the CEP, the assumption was made that the errors in the x and y directions were not independent. This assumption is natural unless an apriori knowledge suggests that the errors in the x and y directions are independent. However, the assumption of independence in the fire control problem is quite difficult to justify due to its complexity. Therefore, it would seem wise to estimate the magnitude of the error involved in assuming independence in order to find out how much difference this assumption will mean in the determination of the CEP.

It was shown in Appendix A that the true orientation of the density function was related to the correlation coefficient. If the true shape of the density function is oriented at some angle with respect to the x and y axes and independence is assumed, the computed standard deviation is not the best estimate of the standard deviation. Consequently the independence assumption introduces an additional error in the estimate of the CEP.

Table 8 is used to illustrate some of the important differences in the results obtained from the problems using the two models.

Table f

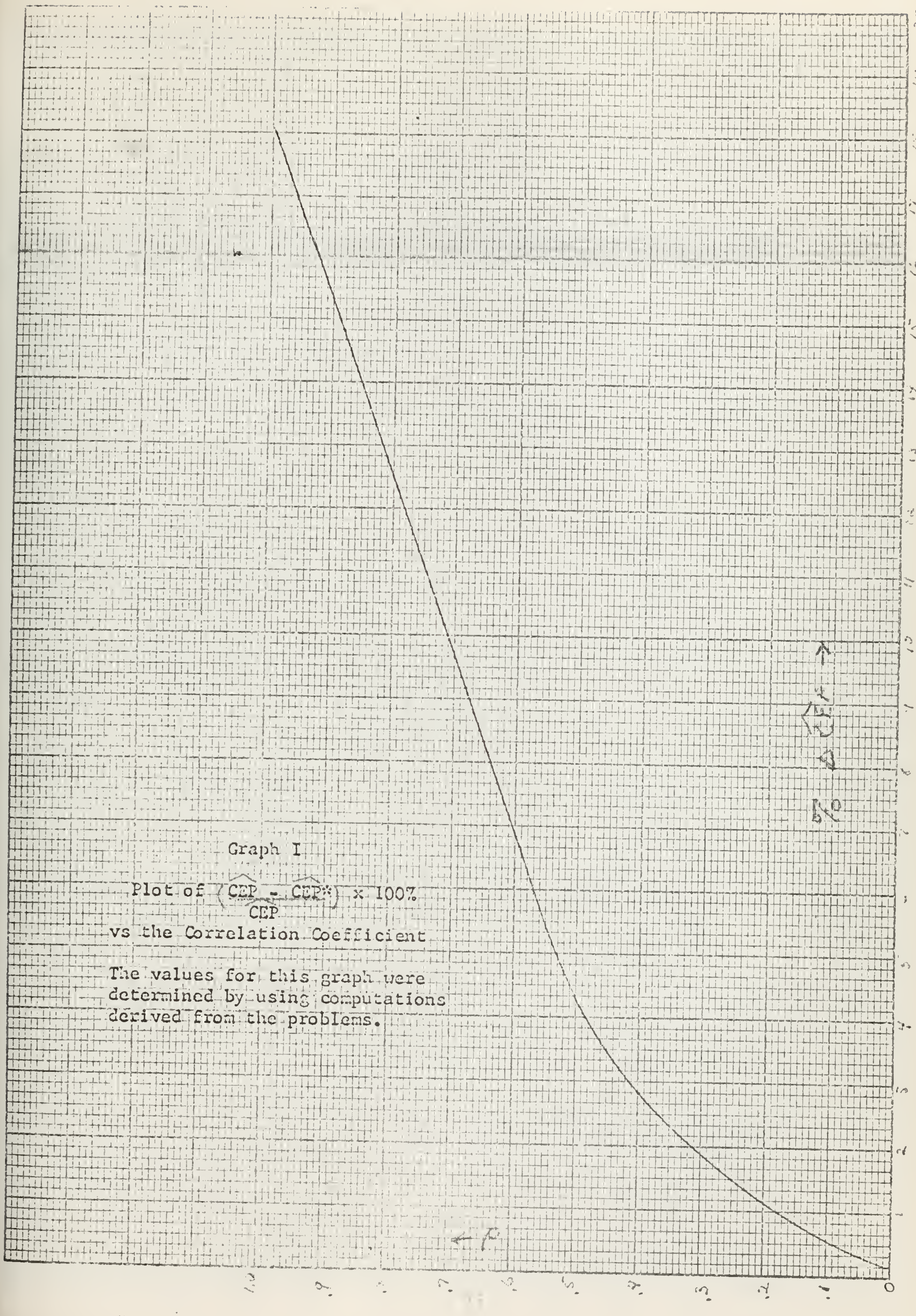
Computed Differences Between Models I and II

Problem	Model I			Model II			Differences (Model I - Model II)		
	Radius of CEP $\widehat{CEP}$	Diff in est. of stand. dev. $(\hat{\sigma}_x - \hat{\sigma}_y)$	Correl. coeff. $\hat{\rho}$	Radius of CEP $\widehat{CEP}$	Diff in est. of stand. dev. $(\hat{\sigma}_x - \hat{\sigma}_y)$	Correl. coeff. $\hat{\rho}$	Radius of CEP $\widehat{CEP}$	Est. of stand. dev.	Correl. coeff. $\hat{\rho}$
I									
Case 1	3.97	.60	.475	3.64	.15	.380	.33	.45	.095
2	4.15	.54	.206	3.37	.77	.081	.25	-.23	.125
3	3.55	.54	.204	3.28	.32	.320	.27	.22	-.116
II									
Case 1	3.37	.57	-.454	3.33	.81	-.626	.04	-.24	-.072
2	3.45	.86	-.256	3.39	1.06	-.395	.06	-.20	-.139
3	3.77	.43	.031	3.71	.62	-.107	.06	-.19	-.076
III									
Case 1	4.32	.63	.735	4.66	.69	.625	-.34	-.06	.110
2	3.72	.57	.695	3.56	.65	.903	.06	-.08	-.203
3	3.40	.35	.660	3.52	.65	.650	-.12	-.30	.010

The table shows some difference in the magnitude of the radius of the CEP as estimated by the two models with the maximum difference being .33/3.97 or 8.3%. Also, the trend in the size remains constant between the two models. That is, as the size of the estimated CEP changes in one model, it changes in the other model in the same direction.

Graph 1 shows a plot of the percent difference in the independent and dependent estimates versus the correlation coefficient. It should be emphasized that the points on the graph were obtained from data computed from the sample problem.

The differences in the estimates of the CEP from the problems are shown in Diagrams 34, 35, and 36. Some of the differences were so small that these estimates were left off. It is interesting to note that the distribution of data in problem 3 shows almost perfect correlation and the estimated differences were also a maximum.

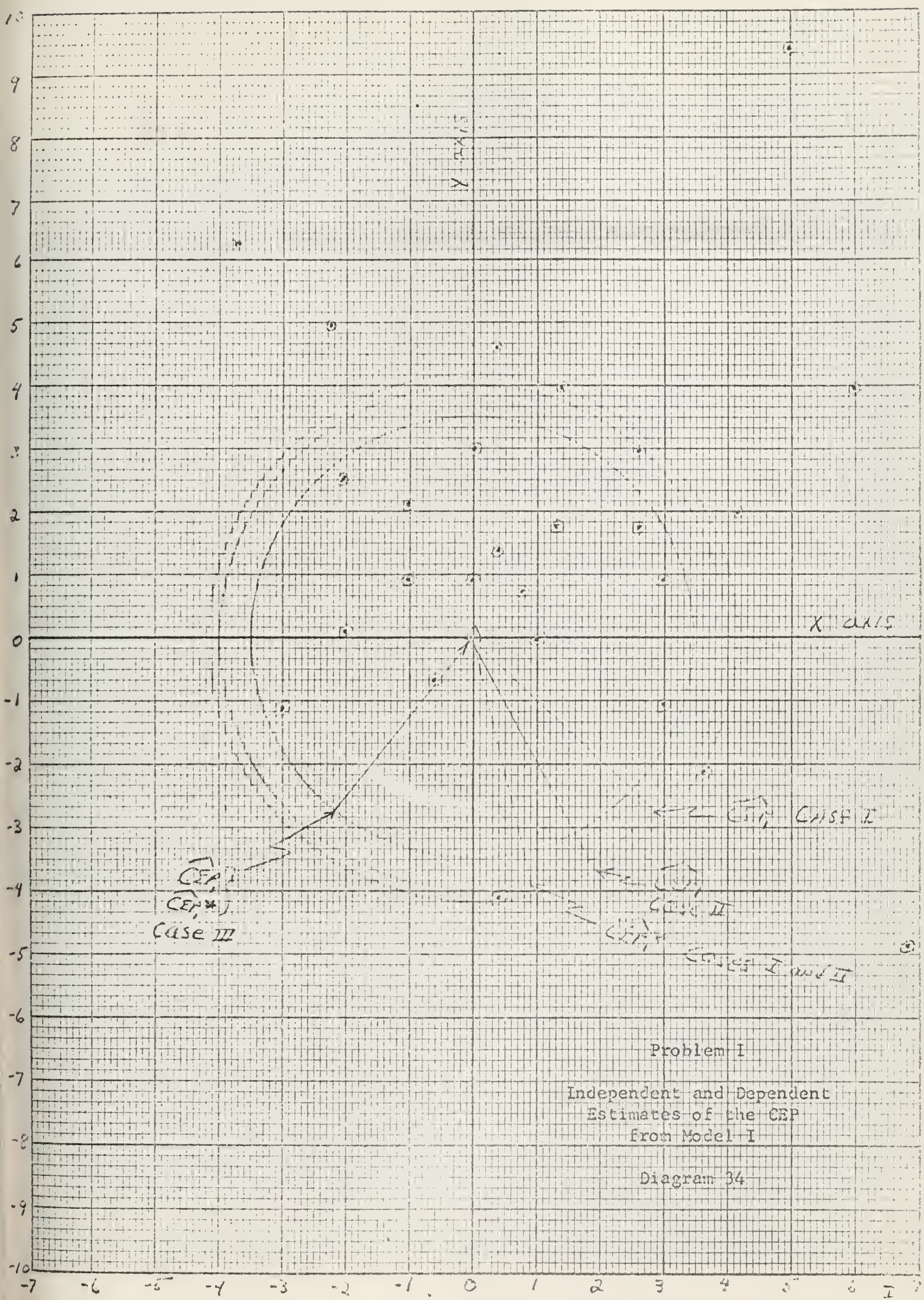


Graph I

Plot of $\left(\frac{\widehat{CEP} - \widehat{CEP^*}}{\widehat{CEP}}\right) \times 100\%$

vs the Correlation Coefficient

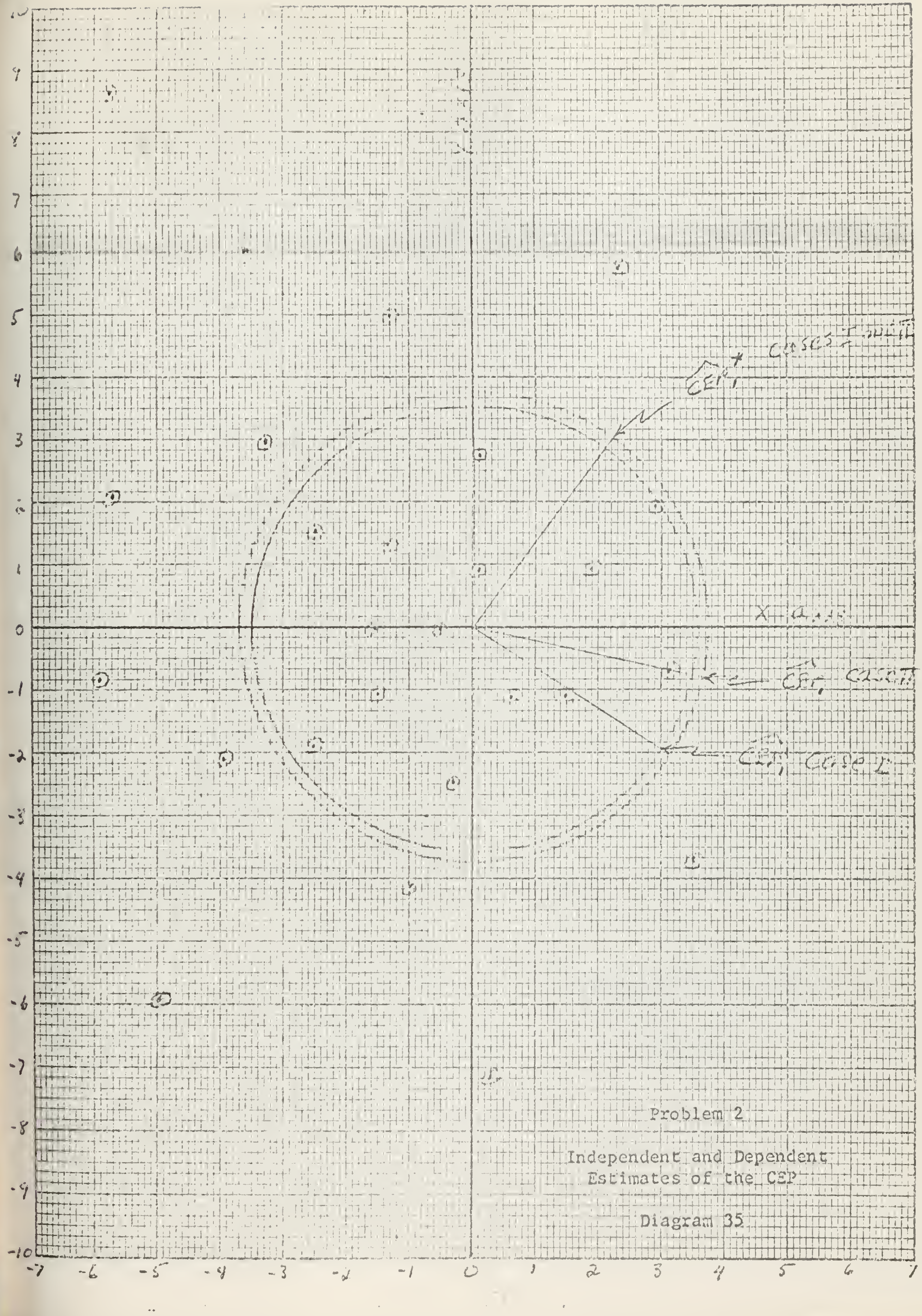
The values for this graph were determined by using computations derived from the problems.



Problem I

Independent and Dependent
Estimates of the CEP
from Model I

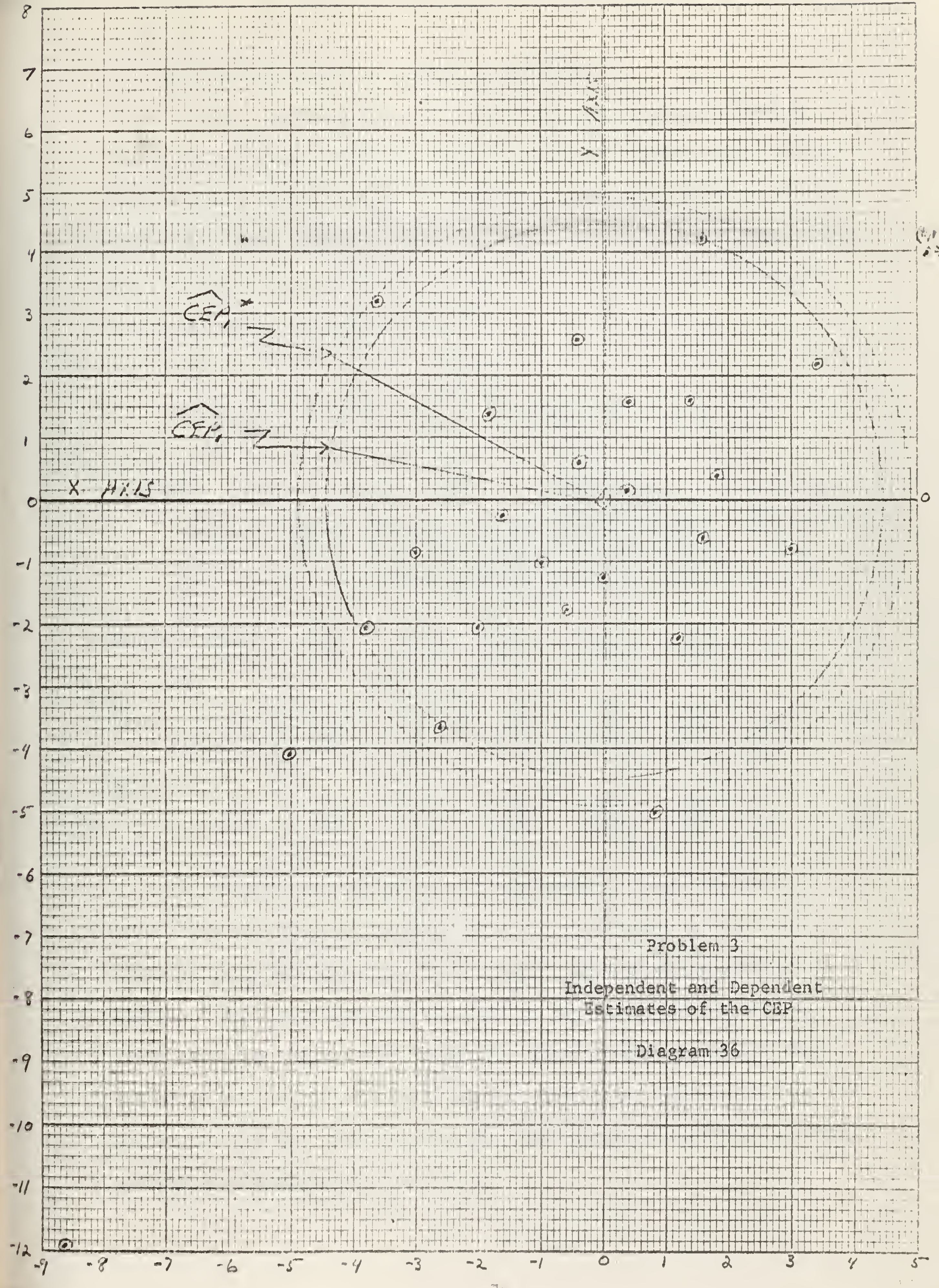
Diagram 34



Problem 2

Independent and Dependent Estimates of the CEP

Diagram 35



Problem 3

Independent and Dependent Estimates of the CEP

Diagram 36

6.4 Effects Caused By The Removal Of Outliers.

The most obvious effect on the CEP when outliers are removed is that the CEP becomes smaller. However, there are several other effects which are not obvious but may be important in determining which estimators can be used. Table g, using the sample problems, gives a comparison between Method I and Method II and the estimates of certain parameters before and after removal of outliers.

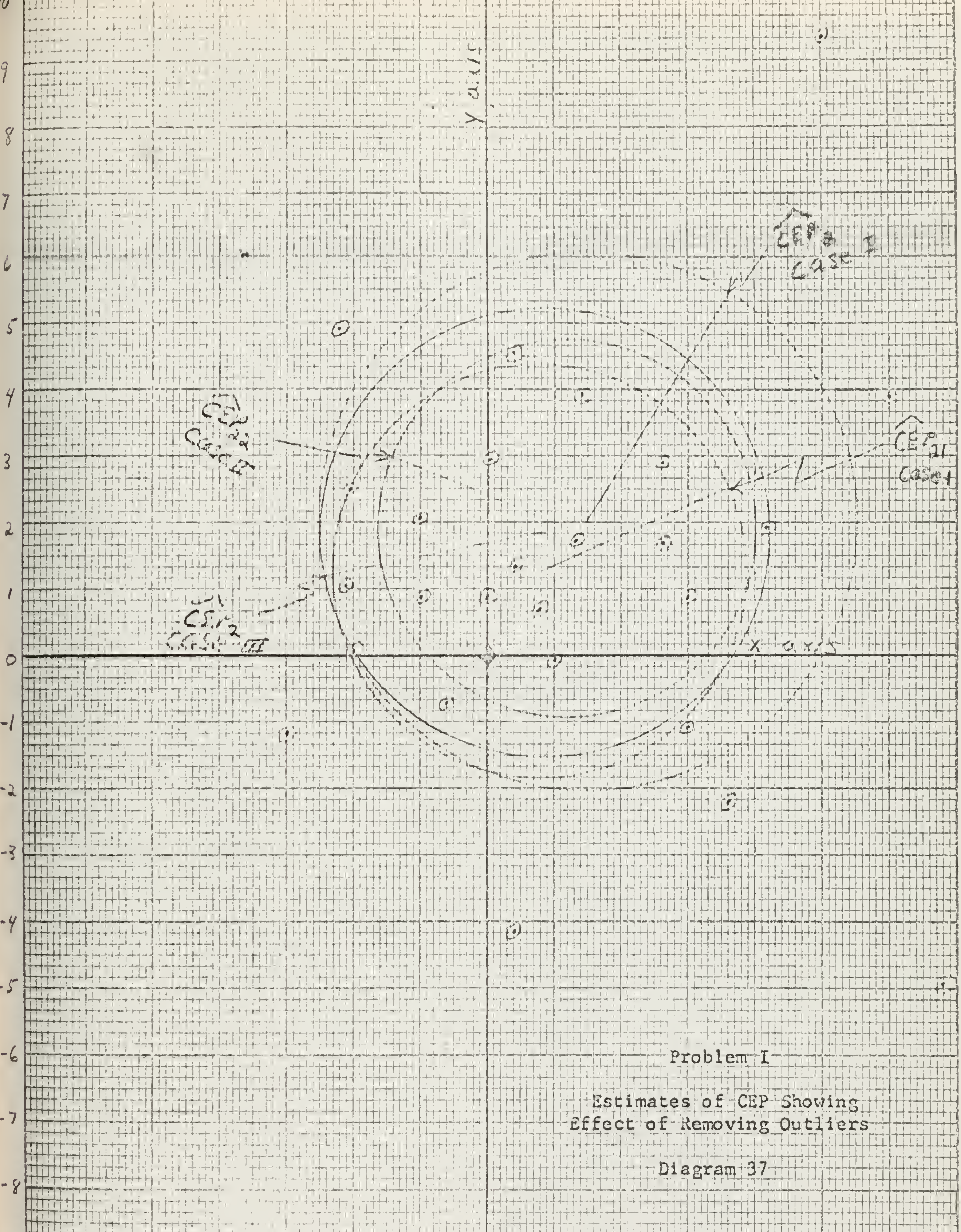
Table g

Effects Of Removing Outliers On $\widehat{CEP}_2$, $\hat{\rho}$, and $(\hat{\sigma}_x - \hat{\sigma}_y)$

Problem	P(Type I Error)		$\widehat{CEP}$		Correlation Coefficient $\hat{\rho}$		Difference in Stand. Dev. $(\hat{\sigma}_x - \hat{\sigma}_y)$	
	Method I	Method II	Before outlier removed	After outlier removed	Before outlier removed	After outlier removed	Before outlier removed	After outlier removed
I Case 1	.10	.05	3.64	3.12	.383	.059	.151	.56
2	.10	.05	3.87	3.04	.081	.136	.77	.24
II Case 1	.05	.025	3.33	2.31	-.626	.201	.81	.34
2	.05	.005	3.39	2.75	-.395	-.03	1.06	.52
III Case 1	.05	.025	4.66	3.03	.625	.306	.69	.43
2	.05	.01	3.56	2.42	.903	.075	.65	.29

The estimate of the CEP was reduced by from 14% to 36% in the problems by the removal of outliers. If a probability of the Type I error had been specified as .05, the point rejected as an outlier in problem 1 would not have been rejected by the elliptical method but would have been rejected by Method II. This is because Method I and Method II are not the same and will not necessarily reject the same points for the same confidence level. The effects of removing outliers are shown in Diagrams 37, 38, and 39.

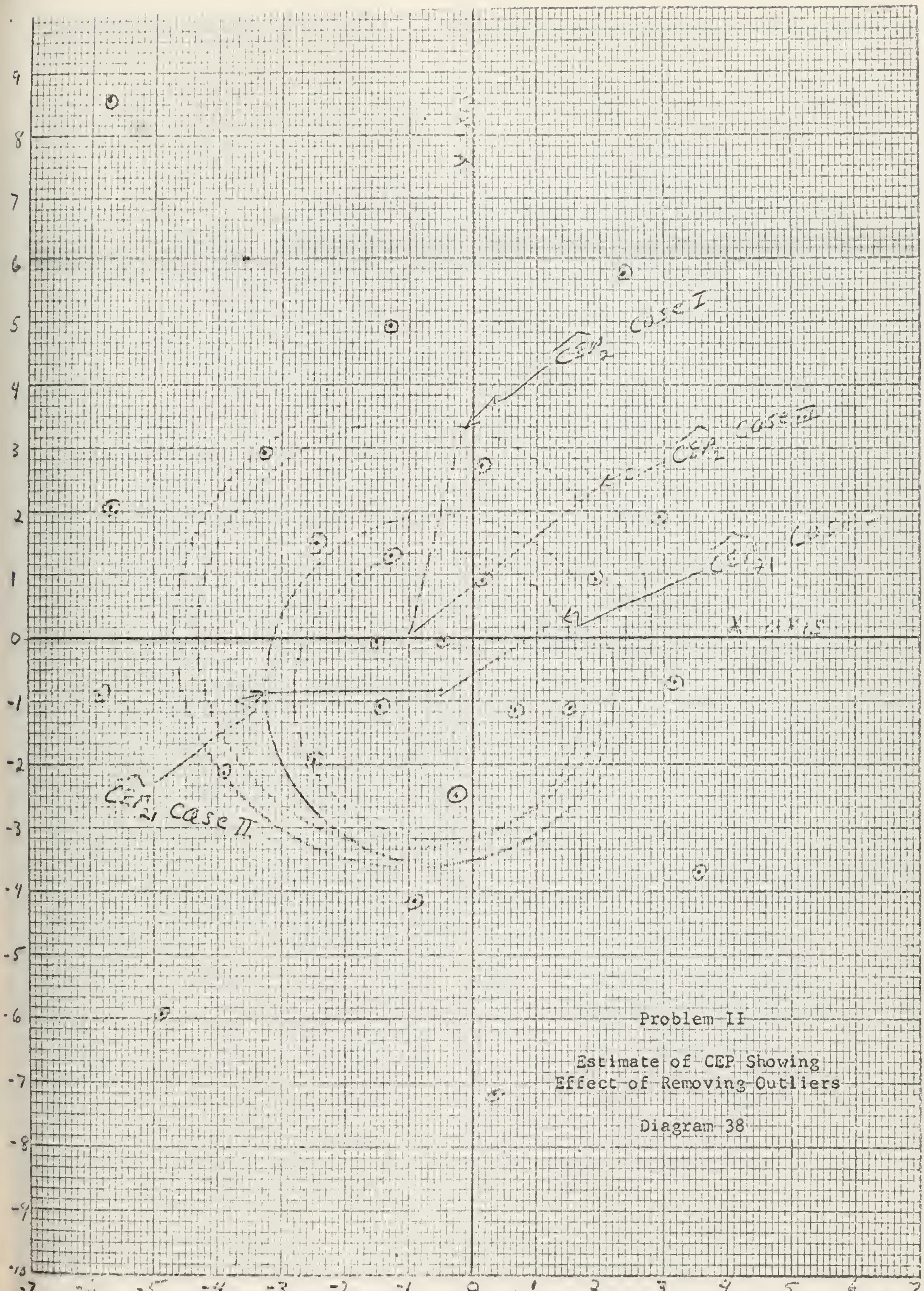
It should be noted that the removal of outliers may change both the correlation coefficients and the difference between the standard deviations in the x and y directions. This is due to the large effect that an outlier has upon the distribution parameters. Thus a large correlation coefficient may be due to the presence of an outlier and not due to correlation between the errors in the x and y directions. Therefore, before the independent method of estimation is rejected, an investigation should be made for outliers.



Problem I

Estimates of CEP Showing Effect of Removing Outliers

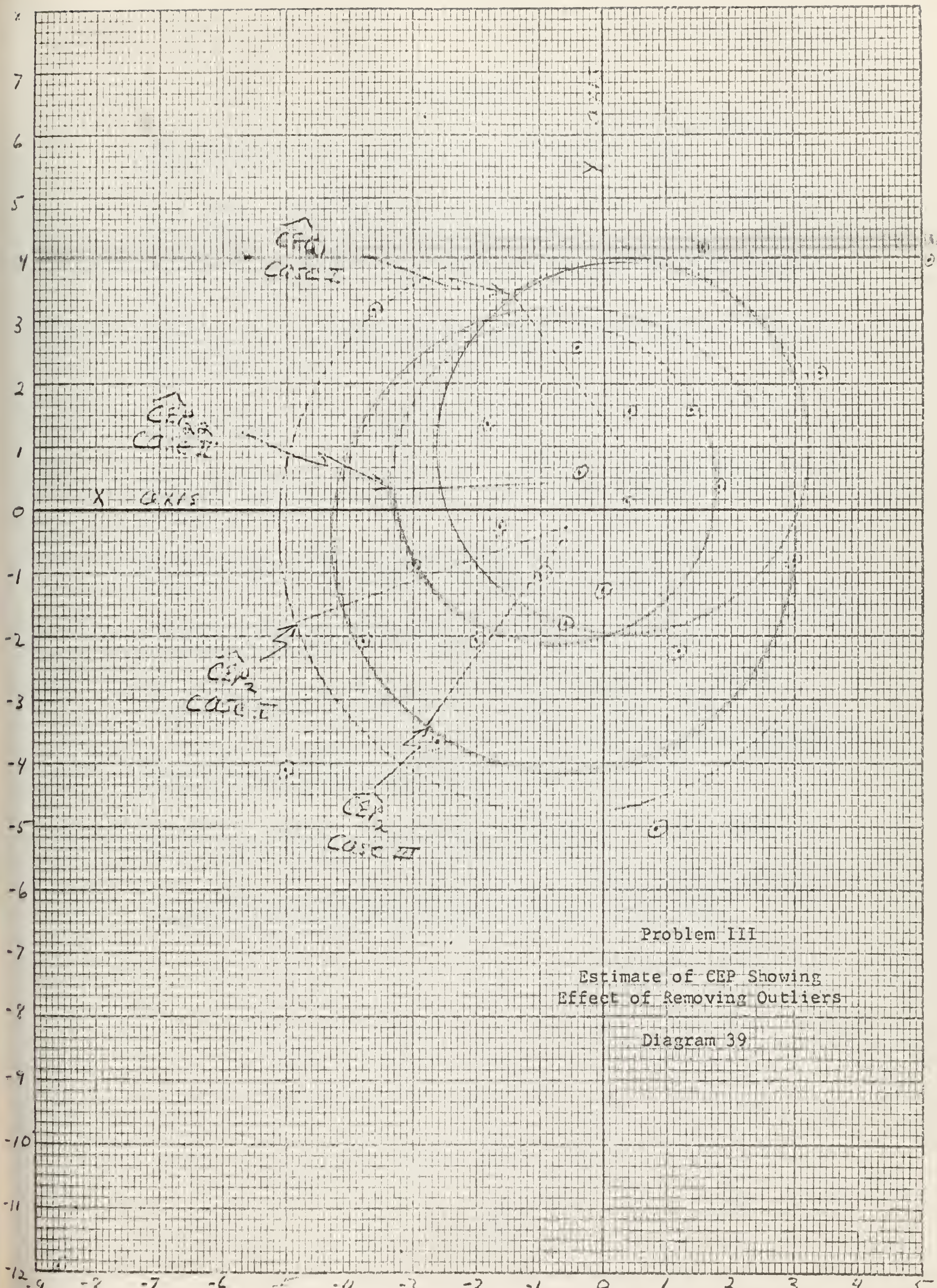
Diagram 37



Problem II

Estimate of CEP Showing
Effect of Removing Outliers

Diagram 38



Problem III

Estimate of CEP Showing Effect of Removing Outliers

Diagram 39

BIBLIOGRAPHY

1. Emanuel Parzen, Modern Probability Theory and Its Applications, John Wiley and Sons, Inc., Stanford University, 1960.

2. A. M. Mood, Introduction To the Theory of Statistics, McGraw-Hill Book Company, Inc., Rand Corporation, 1950.

3. P. B. Moranda, Comparison of Estimates of Circular Probable Error, American Statistical Association Journal 1959, Autonetics, 1959.

4. H. Leon Harter, Circular Error Probabilities, American Statistical Association Journal 1960, Aeronautical Research Laboratories, Wright Patterson Air Force Base, 1960.

5. A. R. Donato, Table of Generalized Circular Error, Mathematics of Computation 1961, Dalgren, Virginia, 1961.

6. C. P. Quesenberry, Some Tests For Outliers, Technical Report No. [illegible], Office of Ordnance Research, U S Army Virginia Polytechnical Institute, 1960.

7. R. A. Tucker, Notes On Estimating CEP and Reliability, Special Projects Report No. 9800.4, Washington D.C., 1960.

8. H. Solomon and L. Woods, Computation of the CEP for the Non-Circular Case in the Presence of Outliers, C-E-I-R, Inc., 1961.

9. F. A. Graybill, An Introduction to Linear Statistical Models, McGraw-Hill Book Company, Inc., Colorado State University, 1961.

10. P. B. Moranda, Effect of Bias on Estimates of the Circular Probable Error, American Statistical Association Journal 1960, Aeronutronic Division of Ford Motor Company, 1960.

11. E. Pairman, Tables of the Gamma, Digamma and Trigamma Functions, Cambridge University Press, London, 1919.

12. C. D. Hodgman, CRC Standard Mathematical Tables, Chemical Rubber Publishing Company, Case Institute of Technology, 1959.

TABLE I

Values of K for Given Values of $c = \frac{\sigma_v}{\sigma_u}$ CEP = K σ_u

C	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.00	.674	.675	.675	.675	.676	.676	.677	.678	.679	[illegible]
.10	.682	.683	.685	.687	.689	.691	.694	.696	.699	.703
.20	.706	.709	.713	.717	.721	.725	.730	.735	.740	.745
.30	.750	.755	.761	.766	.772	.778	.784	.790	.796	.802
.40	.808	.814	.820	.826	.833	.839	.845	.851	.858	.864
.50	.870	.877	.883	.889	.896	.902	.908	.915	.921	.927
.60	.934	.940	.946	.952	.959	.965	.971	.977	.984	[illegible]
.70	.996	1.002	1.009	1.015	1.021	1.027	1.033	1.039	1.045	[illegible]
.80	1.058	1.064	1.070	1.076	1.082	1.088	1.094	1.100	1.106	1.112
.90	1.118	1.124	1.130	1.136	1.142	1.148	1.154	1.160	1.166	1.171
1.00	1.177									

Table I is an excerpt from a table derived at the Naval Weapons Laboratory, Dahlgren, Virginia, under the direction of Harry Weingarten and A. R. DiDonato. Their table presents solutions for the value of K satisfying

$$\frac{1}{2\pi\,\sigma_u\sigma_v}\iint_C e\left[-1/2\left(\frac{u^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)\right]du\,dv = P$$

where C is the circle $u^2 + v^2 = K^2\sigma_u^2$, and $c = \frac{\sigma_v}{\sigma_u}$ for c=0(.01)

and P = 0(.01).99.

See Appendix C for volume of integration.

TABLE 2

VALUES OF K CORRESPONDING TO CUMULATIVE PROBABILITY P

P \ C	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
.5000	0.674	0.682	0.706	0.750	0.808	0.870	0.934	0.996	1.058	1.118	1.177
.7500	1.150	1.155	1.168	1.192	1.231	1.285	1.351	1.425	1.502	1.582	1.665
.9000	1.645	1.648	1.657	1.674	1.699	1.737	1.791	1.862	1.948	2.042	2.146
.9500	1.960	1.962	1.970	1.984	2.005	2.036	2.081	2.146	2.230	2.332	2.448
.9750	2.241	2.244	2.250	2.262	2.281	2.307	2.346	2.404	2.485	2.590	2.716
.9900	2.576	2.578	2.584	2.594	2.610	2.633	2.665	2.715	2.791	2.897	3.035
.9950	2.807	2.809	2.814	2.833	2.838	2.859	2.889	2.933	3.004	3.111	3.255
.9975	3.023	2.025	3.030	3.039	3.052	3.071	3.099	3.140	3.206	3.311	3.462
.9990	3.290	3.292	3.297	3.305	3.317	3.335	3.359	3.396	3.457	3.559	3.717

Table 3

Complete Γ-Functions

N	Log Γ(N)	N	Log Γ(N)	N	Log Γ(N)	N	Log Γ(N)	N	Log Γ(N)
2.0	0.000	5.0	1.380	11.0	6.559	17.0	13.320	23.0	21.050
2.1	0.019	5.2	1.512	11.2	6.764	17.2	13.564	23.2	21.321
2.2	0.042	5.4	1.649	11.4	6.971	17.4	13.809	23.4	21.593
2.3	0.066	5.6	1.789	11.6	7.179	17.6	14.055	23.6	21.865
2.4	0.094	5.8	1.932	11.8	7.389	17.8	14.302	23.8	22.138
2.5	0.123	6.0	2.079	12.0	7.601	18.0	14.551	24.0	22.412
2.6	0.155	6.2	2.228	12.2	7.814	18.2	14.800	24.2	22.687
2.7	0.188	6.4	2.381	12.4	8.028	18.4	15.050	24.4	22.962
2.8	0.224	6.6	2.537	12.6	8.244	18.6	15.301	24.6	23.238
2.9	0.261	6.8	2.696	12.8	8.461	18.8	15.553	24.8	23.515
3.0	0.301	7.0	2.857	13.0	8.680	19.0	15.806	25.0	23.792
3.1	0.341	7.2	3.021	13.2	8.900	19.2	16.060	25.2	24.070
3.2	0.384	7.4	3.187	13.4	9.121	19.4	16.315	25.4	24.349
3.3	0.428	7.6	3.356	13.6	9.344	19.6	16.570	25.6	24.629
3.4	0.474	7.8	3.528	13.8	9.568	19.8	16.827	25.8	24.909
3.5	0.521	8.0	3.702	14.0	9.794	20.0	17.085	26.0	25.190
3.6	0.570	8.2	3.878	14.2	10.021	20.2	17.343	26.2	25.472
3.7	0.620	8.4	4.057	14.4	10.249	20.4	17.602	26.4	25.754
3.8	0.671	8.6	4.237	14.6	10.478	20.6	17.863	26.6	26.037
3.9	0.724	8.8	4.420	14.8	10.708	20.8	18.124	26.8	26.321
4.0	0.778	9.0	4.605	15.0	10.940	21.0	18.386	27.0	26.605
4.1	0.833	9.2	4.792	15.2	11.173	21.2	18.648	27.2	26.890
4.2	0.889	9.4	4.981	15.4	11.407	21.4	18.912	27.4	27.176
4.3	0.947	9.6	5.172	15.6	11.642	21.6	19.176	27.6	27.462
4.4	1.005	9.8	5.365	15.8	11.878	21.8	19.442	27.8	27.749
4.5	1.065	10.0	5.559	16.0	12.116	22.0	19.709	28.0	28.036
4.6	1.126	10.2	5.756	16.2	12.355	22.2	19.975	28.2	28.325
4.7	1.188	10.4	5.954	16.4	12.594	22.4	20.242	28.4	28.613
4.8	1.251	10.6	6.154	16.6	12.835	22.6	20.511	28.6	28.903
4.9	1.315	10.8	6.356	16.8	13.077	22.8	20.780	28.8	29.193
5.0	1.380	11.0	6.559	17.0	13.320	23.0	21.050	29.0	29.484
								29.2	29.775
								29.4	30.067
								29.6	30.359
								29.8	30.652
								30.0	30.946

TABLE 4

Cumulative Chi-square Distribution

N \ F	.005	.010	.025	.050	.100	.250
1	.0	.0	.0	.0	.0158	.102
2	.0100	.0201	.0506	.103	.211	.575
3	.0717	.115	.216	.352	.584	1.21
4	.207	.297	.484	.711	1.06	1.92
5	.412	.554	.831	1.15	1.61	2.67
6	.676	.872	1.24	1.64	2.20	3.45
7	.989	1.24	1.69	2.17	2.83	4.25
8	1.34	1.65	2.18	2.73	3.49	5.07
9	1.73	2.09	2.70	3.33	4.17	5.90
10	2.16	2.56	3.25	3.94	4.87	6.74
11	2.60	3.05	3.82	4.57	5.58	7.58
12	3.07	3.57	4.40	5.23	6.30	8.44
13	3.57	4.11	5.01	5.89	7.04	9.30
14	4.07	4.66	5.63	6.57	7.79	10.2
15	4.60	5.23	6.26	7.26	8.55	11.0
16	5.14	5.81	6.91	7.96	9.31	11.9
17	5.70	6.41	7.56	8.67	10.1	12.8
18	6.26	7.01	8.23	9.39	10.9	13.7
19	6.84	7.63	8.91	10.1	11.7	14.6
20	7.43	8.26	9.59	10.9	12.4	15.5
21	8.03	8.90	10.3	11.6	13.2	16.3
22	8.64	9.54	11.0	12.3	14.0	17.2
23	9.26	10.2	11.7	13.1	14.8	18.1
24	9.89	10.9	12.4	13.8	15.7	19.0
25	10.5	11.5	13.1	14.6	16.5	19.9
26	11.2	12.2	13.8	15.4	17.3	20.8
27	11.8	12.9	14.6	16.2	18.1	21.7
28	12.5	13.6	15.3	16.9	18.9	22.7
29	13.1	14.3	16.0	17.7	19.8	23.6
30	13.8	15.0	16.8	18.5	20.6	24.5

TABLE 4 (cont)

N \ F	.500	.750	.900	.950	.975	.990	.995
1	.455	1.32	2.71	3.84	5.02	6.63	7.88
2	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	13.3	17.1	21.1	23.7	26.1	29.1	31.3
15	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	19.3	23.8	28.4	31.4	34.2	37.6	40.0
21	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	23.3	28.2	33.2	36.4	39.4	43.0	45.6
25	24.3	29.3	34.4	37.7	40.6	44.3	46.9
26	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	29.3	34.8	40.3	43.8	47.0	50.9	53.7

APPENDIX A

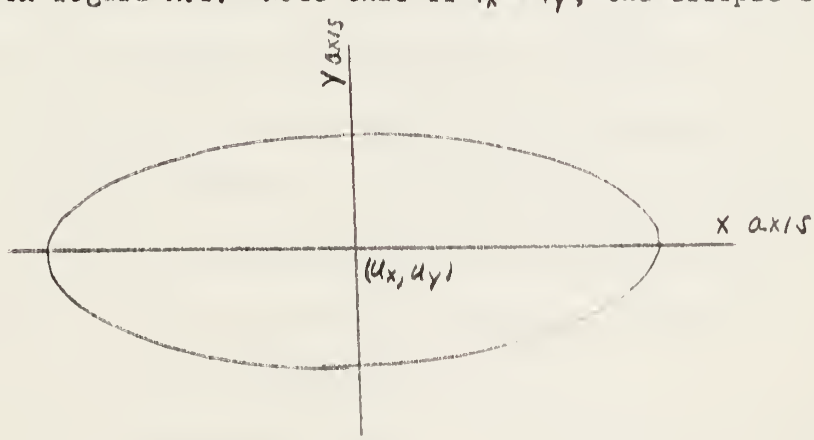
ORIENTATION AND TRANSFORMATION OF THE BIVARIATE DENSITY FUNCTION

A.1 Introduction

This appendix is concerned with the orientation of the bivariate normal density function over the x,y plane. Primarily this requires an investigation of the correlation between the random variables X and Y and once the correlation is determined, a transformation of axes so that the function can be integrated more easily.

A.2 Orientation of the Axes

If the correlation coefficient is zero, that is the random variables X and Y are independent, the orientation will be symmetrical with respect to the x and y axes. This means that a plane parallel to the x and y plane will cut the density function in the form of an ellipse whose minor and major axes are parallel to the x and y axes. This is shown below in figure A.1. Note that if $\sigma_x = \sigma_y$, the ellipse becomes a circle.

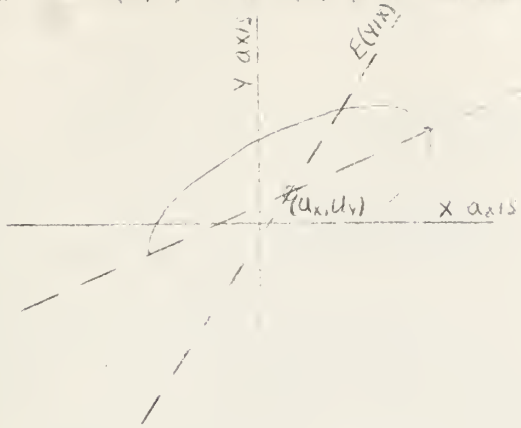


Orientation of the Ellipse When $\rho = 0$.

Figure A.1

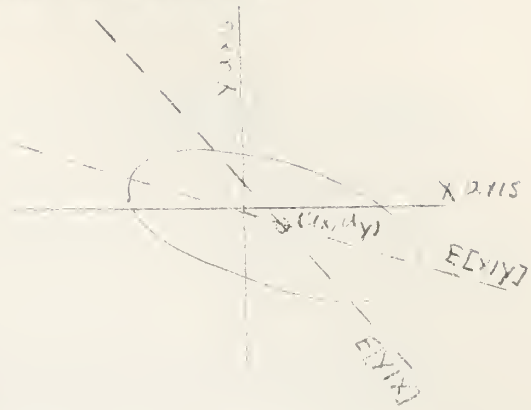
If the correlation coefficient is not zero and less than plus or minus 1, the orientation is offset from the x and y axes in the direction

of the E(X|Y) and E(Y|X) as shown in Figures A.2a and A.2b.



Orientation of the Ellipse
when $0 < \rho < 1$

Figure A.2a



Orientation of the Ellipse
when $-1 < \rho < 0$

Figure A.2b

The error introduced in assuming independence, when the random variables X and Y are not independent is a function of the correlation coefficient and is due to the true orientation of the density function with respect to the x and y axes. If it is assumed that the errors in the x and y directions are independent when in fact they are not, an additional error will be made in computing the estimates of the variances. This is due to the fact that the computation of $\hat{\sigma}_x, \hat{\sigma}_y$ will be in the direction of the assumed axes (x,y) instead of the direction of true orientation E(X|Y),E(Y|X). Therefore, it becomes important to obtain some knowledge of the true orientation in order to obtain the best estimates of the variances. This can be done by obtaining estimates of the angles (ϕ, θ) between the assumed axes and the true axes.

The following sections are devoted to the different possible orientations due to the different ranges of the correlation coefficient.

A.3 Determination of the Orientation of the Axes

A.3.1 Determining the True Orientation when the Correlation Coefficient is zero.

If $\rho = 0$, the random variables X and Y are independent. Therefore $E(X|Y) = E(X)$ and $E(Y|X) = E(Y)$. This can be proven by determination of either the value of the conditional expectation directly or indirectly using the linear predictor.

A.3.1.1 Direct Determination of the Conditional Expectation if $\rho = 0$.

The expected value of one random variable given the value of the other random variable was defined in section 1.2.5 as

$$\text{(A.1)} \quad E(X|Y) = \int_{-\infty}^{\infty} x\, f_{X|Y}(x|y)\, dx = \int_{-\infty}^{\infty} \frac{x\, f_{X,Y}(x,y)\, dx}{f_Y(y)} = \frac{\frac{1}{2\pi \sigma_x \sigma_y}}{\frac{1}{\sqrt{2\pi}\, \sigma_y}} \int_{-\infty}^{\infty} \frac{x\, e^{-\frac{1}{2}\left(\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right)}}{e^{-\frac{1}{2}\left(\frac{y}{\sigma_y}\right)^2}}\, dx$$

$$= \frac{1}{\sqrt{2\pi}\, \sigma_x} \int_{-\infty}^{\infty} x\, e^{-\frac{1}{2}\left(\frac{x}{\sigma_x}\right)^2} = E(X), \text{ as defined in Section 1.1.2}$$

The $E(Y|X)$ can be determined in the same way and is equal to $E(Y)$.

A.3.1.2 Indirect Determination of the Conditional Expectation Using the Best Linear Predictor.

The conditional expectation of one random variable given the value of the other random variable is a linear function of the known random variable when both random variables are jointly normally distributed. That is the $E(X|Y) = AY + B$, where A and B are constants which can be determined. The linear predictors for the conditional expectations under consideration will be defined as follows:

(A.2) $E(Y/X) = E(Y) + C_1[X-E(X)]$ where $C_1 = \frac{COV(X,Y)}{VAR(X)} = \frac{\sigma_{xy}}{\sigma_x^2}$

$E(X/Y) = E(X) + C_2[Y-E(Y)]$ where $C_2 = \frac{COV(X,Y)}{VAR(Y)} = \frac{\sigma_{xy}}{\sigma_y^2}$

$\rho = \frac{COV(X,Y)}{\sqrt{VAR(X)}\sqrt{VAR(Y)}}$ from 1.2.5

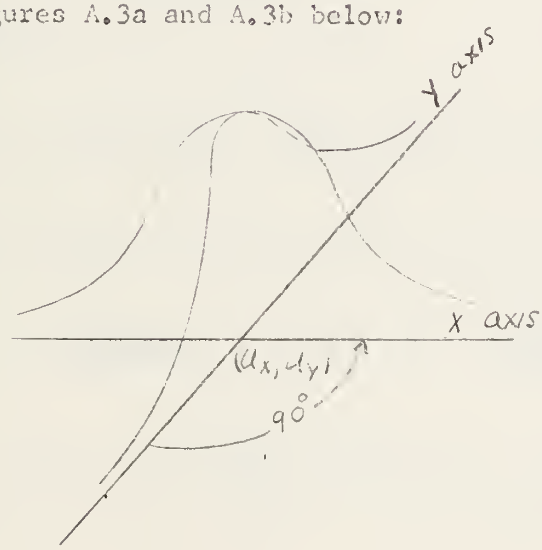
In the case where $\rho = 0 = \frac{COV(X,Y)}{VAR(X)}$, $COV(X,Y) = C = C_1 = C_2$ from (A.2)

Therefore, the results become the same as in Section A.2.1.1. That is

(A.3) $E(X/Y) = E(X)$ and

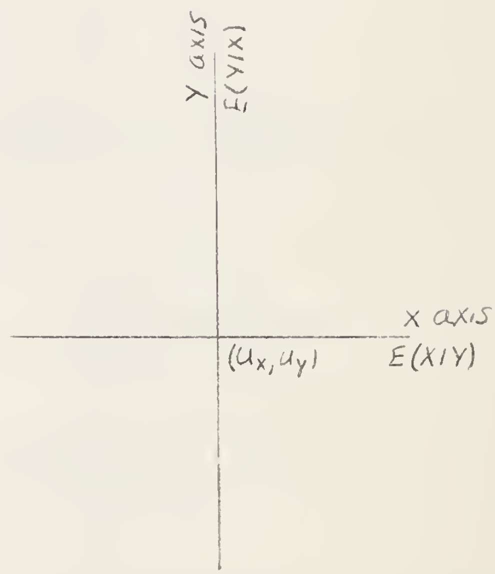
$E(Y/X) = E(Y)$

In the case that $\rho = 0$, the orientation of the axes will be as shown in figures A.3a and A.3b below:



Orientation of the Density Function when the Correlation Coefficient is Zero.

Figure A.3a



Orientation of the Axes when the Correlation Coefficient is Zero.

Figure A.3b

A.3.2 Determination of the Orientation if $\rho = 1$.

If $\rho = 1$, the $E(Y/X)$ and $E(X/Y)$ will lie along the same axis. This can be proven by using the best linear predictors in formula (A.2) and the definition of the correlation coefficient. That is

(A.4) $\rho = 1 = \dfrac{COV(X,Y)}{\sqrt{VAR(X)}\sqrt{VAR(Y)}}$ and therefore $COV(X,Y) = \sqrt{VAR(X)}\sqrt{VAR(Y)}$.

Then using formula (A.2) $E(Y/X) = E(Y) + C_1[X - E(X)] = E(Y) - C_1 E(X) + C_1 X$ where

(A.5) $C_1 = \dfrac{COV(X,Y)}{VAR(X)} = \sqrt{\dfrac{VAR(Y)}{VAR(X)}}$ using the result in formula A.4. Since

$E(X)$, $E(Y)$, and $\sqrt{\dfrac{VAR(Y)}{VAR(X)}}$ are constants, the random variable $E(Y/X)$ is of the form

(A.6) $E(Y/X) = AX + B$ where $A = C_1$ and $B = E(Y) - E(X)C_1$. The tangent of the angle between the y axis and the line $E(Y/X)$ is

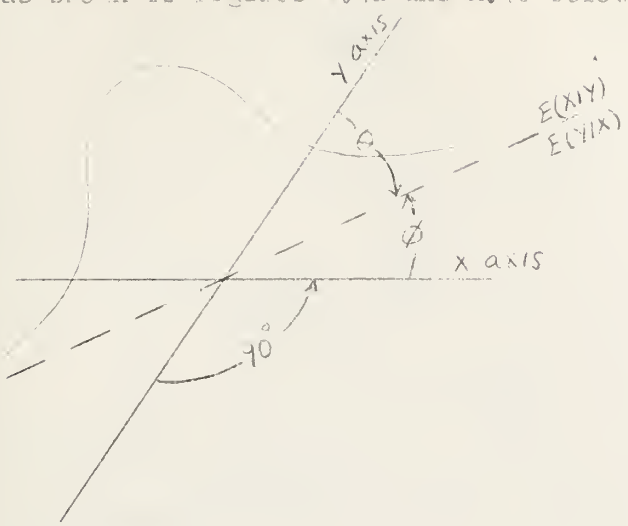
(A.7) $\dfrac{d[E(Y/X)]}{dx} = \text{Tan}\theta = C_1$ or $\sqrt{\dfrac{VAR(Y)}{VAR(X)}}$ in the case where $\rho = 1$.

The tangent of the angle between the x axis and the line $E(X/Y)$ is determined in the same way and in this case

(A.8) $\text{Tan}\phi = C_2 = \sqrt{\dfrac{VAR(X)}{VAR(Y)}}$ in the case where $\rho = 1$. Since $\text{Tan}\phi = \sqrt{\dfrac{VAR(X)}{VAR(Y)}}$

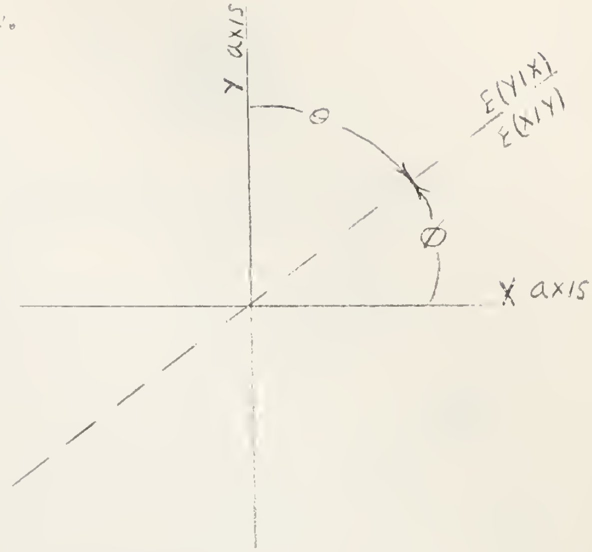
$\dfrac{1}{\sqrt{\dfrac{VAR(Y)}{VAR(X)}}} = \dfrac{1}{\text{Tan}\theta} = \text{COT}\theta$ the lines must be the same.

Therefore, in the case that $\rho = 1$ the orientation of the axes will be as shown is figures A.4a and A.4b below.



Orientation of the Density Function when the Correlation Coefficient is 1.

Figure A.4a



Orientation of the Axes when the Correlation Coefficient is 1.

Figure A.4b

A.3.3 Determination of the Orientation of the Axes if $0 < \rho < 1$.

If $0 < \rho < 1$, the two lines E(X|Y) and E(Y|X) will not be the same or perpendicular and will be oriented as shown in figures A.5a and A.5b. This can be proven by using the same method as in Section A.3.2, except that

(A.9) $$0 < \frac{COV(X,Y)}{VAR(X)\ VAR(Y)} < 1$$

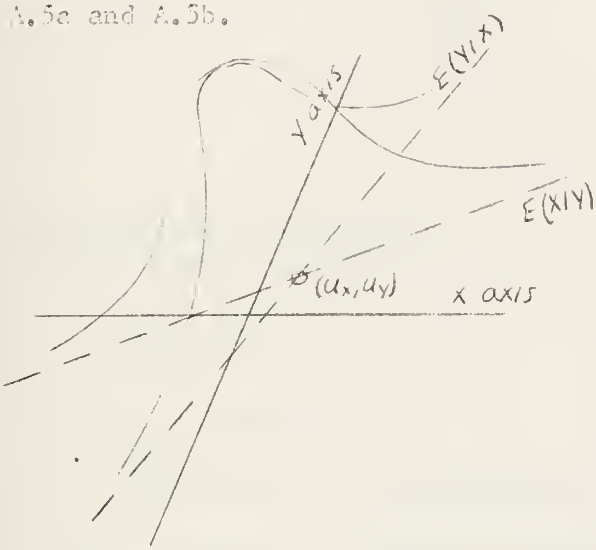
It follows from this that $0 < COV(X,Y) < \sqrt{VAR(X)}\sqrt{VAR(Y)}$. Then using formulas A.7 and A.8 with the definitions of the constants in formulas A.2, the desired angles are

(A.10) $$\tan\theta = C_1 = \frac{COV(X,Y)}{VAR(Y)} < \sqrt{\frac{VAR(Y)}{VAR(X)}}\ , \quad \tan\emptyset = C_2 = \frac{COV(X,Y)}{VAR(X)} < \sqrt{\frac{VAR(X)}{VAR(Y)}}$$

the complete range of possible values for the two angles using A.9 are

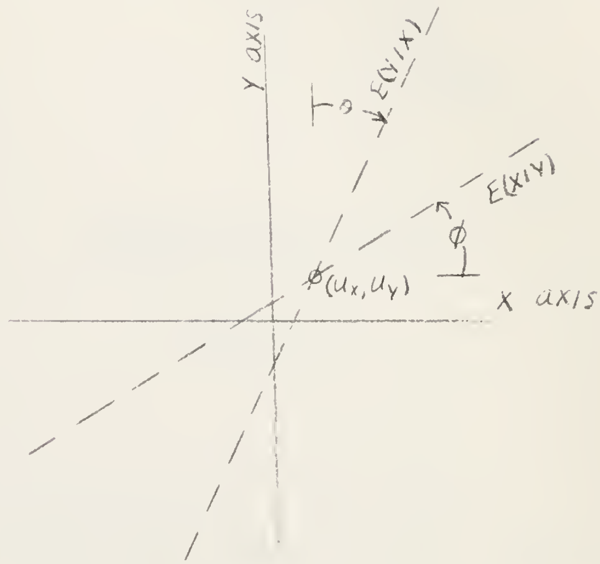
(A.11) $0 < \theta < \text{Tan}^{-1}\sqrt{\frac{VAR(Y)}{VAR(X)}}$ $\quad 0 < \phi < \text{Tan}^{-1}\sqrt{\frac{VAR(X)}{VAR(Y)}}$

In this case the orientation of the axes will be as shown in figures A.5a and A.5b.



Orientation of the Density Function when $0 < \rho < 1$.

Figure A.5a



Orientation of the Axes when the Correlation Coefficient is $0 < \rho < 1$.

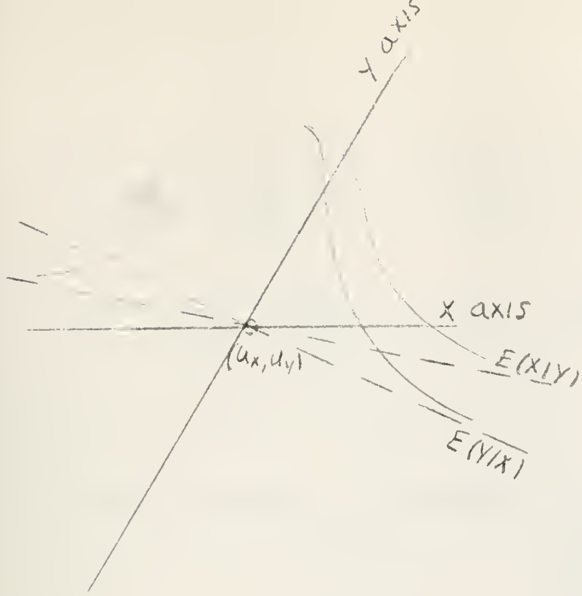
Figure A.5b

A.3.4 Determination of the Orientation of the axes if $-1 < \rho < 0$.

If $-1 < \rho < 0$, it follows from formulas A.9, A.10 and A.11 that

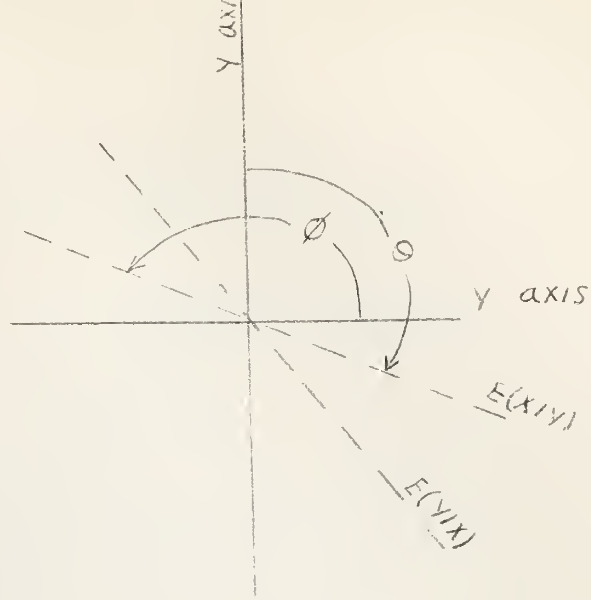
(A.12) $\text{Tan}^{-1}\left(-\sqrt{\frac{VAR(Y)}{VAR(X)}}\right) < \theta < 0, \quad \tan^{-1}\left(-\sqrt{\frac{VAR(X)}{VAR(Y)}}\right) < \phi < 0$

In this case the orientation of the axes will be as shown in figures A.6a and A.6b.



Orientation of the Density Function when $-1 < \rho < 0$

Figure A.6a



Orientation of the Axes when $-1 < \rho < 0$.

Figure A.6b

A.4 Illustrations

Although the true orientation of the axes will not be known, it can be estimated by using the various estimators shown in table h below:

Table h	
Estimators Used in Determining Estimated Axes Orientation	
Estimator used	Value Estimated
$\bar{x}$	E(X)
$\bar{y}$	E(Y)
$\hat{\sigma}_x^2$	VAR(X)
$\hat{\sigma}_y^2$	VAR(Y)
$\hat{\sigma}_x$	$\sqrt{VAR(X)}$
$\hat{\sigma}_y$	$\sqrt{VAR(Y)}$
$\hat{\sigma}_{xy}$	COV(X,Y)

The estimated parameters in the illustrations which follow are determined by using the data from the example problems in Section II.

A.4.1 Illustration (1)

The data is obtained from example problem no. 1 with a sample size of n = 25.

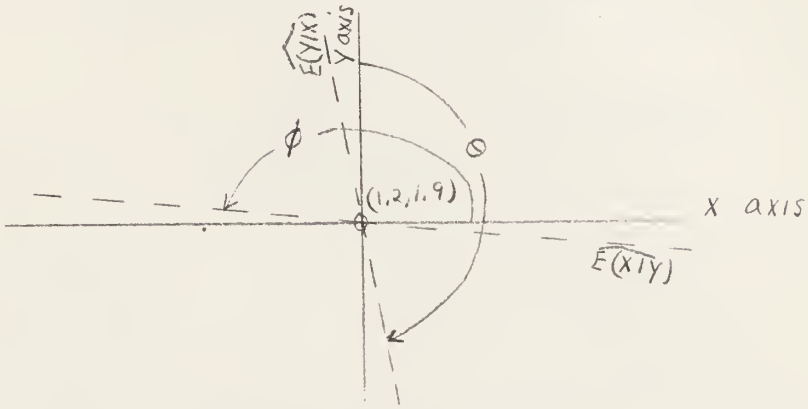
$\bar{x} = 1.2,\ \bar{y} = 1.9,\ \hat{\sigma}_x^2 = 6.8,\ \hat{\sigma}_y^2 = 8.6,\ \hat{\sigma}_{xy} = -.4,\ \hat{\rho} = -.05$

$$\widehat{E(Y|X)} = \widehat{E(Y)} + C_1[X - \widehat{E(X)}] \quad , \quad \widehat{E(X|Y)} = \widehat{E(X)} + C_2[Y - \widehat{E(Y)}]$$

$$\text{where} \quad C_1 = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x^2} = -.05 = \operatorname{Tan}\theta \qquad C_2 = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_y^2} = -.04 = \operatorname{Tan}\phi$$

$$\theta = 176^\circ 51' \qquad \phi = 177^\circ 31'$$

The orientation of the axes is shown in figure A.7



Estimated Orientation
of the Axes

Figure A.7

It should be noted that the orientation of the axes in figure A.7 implies that the random variables X and Y are nearly independent and that the independent model of computing the CEP can be used with only a small error due to the orientation. The computed values for the two different estimates of the CEP are $\widehat{CEP}_2 = 3.28$ and $\widehat{CEP}_2^* = 3.26$

A.4.2 Illustration (2)

The data obtained from problem 3 with a sample size of n = 15.

$\bar{x} = -.6,\ \bar{y} = -.2,\ \hat{\sigma}_x^2 = 10.7,\ \hat{\sigma}_y^2 = 15.4,\ \hat{\sigma}_{x,y} = 11.6,\ \hat{\rho} = .90$

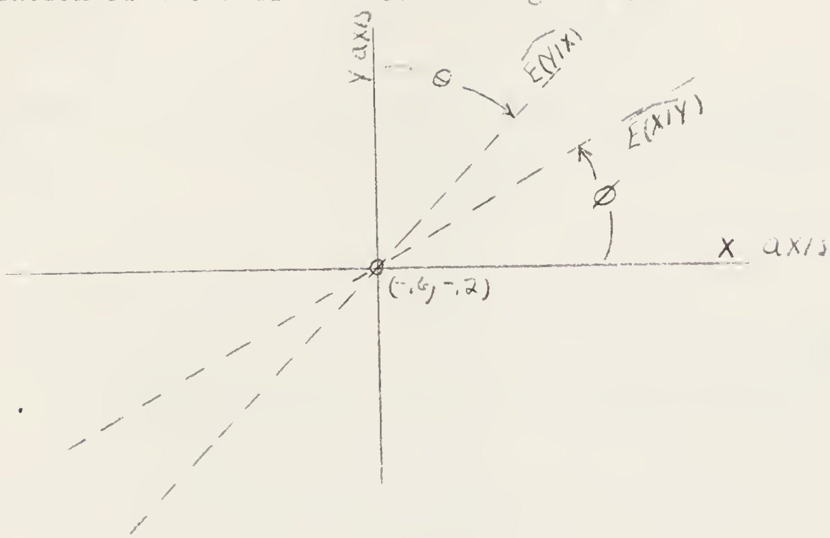
$$\widehat{E(X/X)} = \widehat{E(Y)} + C_1[X - \widehat{E(X)}] \quad , \qquad \widehat{E(X|Y)} = \widehat{E(X)} + C_2[Y - \widehat{E(Y)}]$$

where

$$C_1 = \frac{\widehat{COV(X,Y)}}{\widehat{VAR(X)}} = 1.082 = \tan\theta \qquad C_2 = \frac{\widehat{COV(X,Y)}}{\widehat{VAR(Y)}} = .755 = \tan\emptyset$$

$$\theta = 47^{\circ}15' \qquad \emptyset = 37^{\circ}3'$$

The orientation of the axes is shown in figure A.8



Estimated Orientation of the Axes
when Dependence is Implied

Figure A.8

It should be noted that if this were the true orientation, it implies almost perfect correlation between the random variables X and Y. This orientation will exhibit the greatest difference in the estimates of the CEP if independence were initially assumed. The computed values for the two different estimates of the CEP are $\widehat{CEP}_2 = 3.52$ and $\widehat{CEP}^*_2 = 3.72$.

A.4.3 Illustration (3)

The data for this illustration is also obtained from problem 3 with a sample size of n = 15. However, in this case, the two outliers have been removed and the sample size used for the computation is 13.

$\bar{x} = -.8,\ \bar{y} = .4,\ \hat{\sigma}_x^2 = 3.64,\ \hat{\sigma}_y^2 = 4.8,\ \hat{\sigma}_{xy} = .32,\ \hat{\rho} = .07$

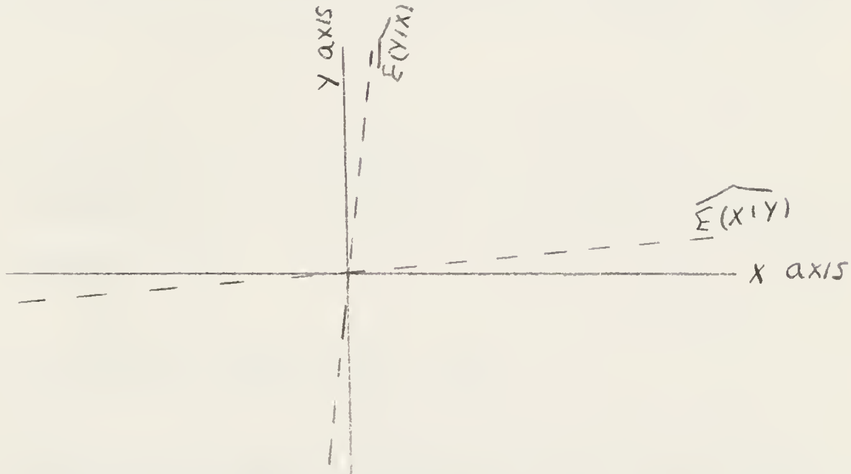
$\widehat{E(Y|X)} = \widehat{E(Y)} + C_1[X - \widehat{E(X)}] \qquad \widehat{E(X|Y)} = \widehat{E(X)} + C_2[Y - \widehat{E(Y)}]$

where

$$C_1 = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x^2} = .088 = \text{Tan}\theta \qquad C_2 = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_y^2} = .066 = \text{Tan}\phi$$

$$\theta = 5^\circ 3' \qquad \phi = 3^\circ 46'$$

The orientation of the axes after removal of the outliers is shown in figure A.9



Estimated Orientation of the Axes
After Removal of the Outliers

Figure A.9

It should be noted that the removal of the outliers rotated the axes enough so that independence could be assumed with only a small error in the estimate of the CEP. The computed values for the two different estimates of the CEP are $\widehat{CEP}_2 = 2.42$ and $\widehat{CEP}_2^* = 2.52$. Thus the removal of outliers will not only reduce the size of the CEP but may aid in the determination of whether the simpler model of independent estimates may be used or not.

A.5 Transformation of the Axes

In order to integrate formula (1.1) of Section 1, it is necessary to transform the axes. This transformation can be done in several ways

but the use of matrix notation can greatly simplify the procedure. It is necessary to first define some of the concepts which will be used.

A.5.1 Definitions:

A.5.1.1 The matrix $A=(a_{ij})$ where $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ will be used for simplification.

A.5.1.2 The transposed matrix is defined as $A'=\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$

A.5.1.2.1 Theorem 1. The transpose of $A'=(A')'=A$

A.5.1.3 The inverse of A is defined as the matrix A^{-1} such that $AA^{-1}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

A.5.1.4 The identity matrix $I=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

A.5.1.5 A symmetric matrix is defined as a matrix such that the transpose of the matrix A equals A. That is

$$A' = \begin{pmatrix} a_{11} & a_{22} \\ a_{12} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = A$$

A.5.1.6 If C is a 2x2 matrix such that $C'C = I$, then C is defined as an orthogonal matrix and $C'=C^{-1}$.

A.5.1.7 A characteristic root of a 2x2 matrix A is a scalar λ such that $AX=\lambda X$ and $AX-\lambda X=0$ for some vector $X \neq 0$. It follows that if λ is a characteristic root of A, then $(A-\lambda I)X=0$ and therefore $|A-\lambda I| = 0$.

A.5.1.8 The diagonal matrix D is defined as a square matrix whose off diagonal elements are all zero. That is, if $D=(d_{ij})$, the $d_{ij}=0$ if $i \neq j$. In this example $D=\begin{pmatrix} d_{11} & 0 \\ 0 & d_{22} \end{pmatrix}$

A.5.1.9 The quadratic form Q is defined as $Q=Z'AZ$

A.5.2 The bivariate normal density function in matrix notation is

$$f_{x,y}(x,y) = \frac{1}{2\pi\,|A^{-1}|^{\frac{1}{2}}} \exp -\tfrac{1}{2} Z'AZ \text{ where } A = \frac{1}{1-\rho^2}\begin{pmatrix} \frac{1}{\sigma_x^2} & -\frac{\rho}{\sigma_x \sigma_y} \\ -\frac{\rho}{\sigma_x \sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} \sigma_x^2 & \rho\,\sigma_x \sigma_y \\ \rho\,\sigma_x \sigma_y & \sigma_y^2 \end{pmatrix}$$

It should be noted that A and A^{-1} are both symmetric matrices. That is $A = A'$ and $A^{-1}=(A^{-1})'$. Thus the theorem applies that for every symmetric matrix A^{-1} there exists an orthogonal matrix C such that $C'A^{-1}C=D$ where D is a diagonal matrix whose diagonal elements are the characteristic roots of A^{-1}. The matrix D would thus be

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

where λ_1 and λ_2 are the characteristic roots of the matrix A.

In order to find the characteristic roots of A^{-1} we must first use the identity matrix $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, then $I\lambda = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$

The characteristic roots of a symmetric matrix are determined from the characteristic polynomial $f(\lambda) = |A^{-1} - \lambda I| = 0$

$$\left| \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right| = 0 = \begin{vmatrix} \sigma_x^2 - \lambda & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 - \lambda \end{vmatrix}$$

$$(\sigma_x^2 - \lambda)(\sigma_y^2 - \lambda) - \sigma_{xy}^2 = 0$$

$$\lambda^2 - (\sigma_x^2 + \sigma_y^2)\lambda + \sigma_x^2 \sigma_y^2 - \sigma_{xy}^2 = 0$$

This is solved using the general form for a quadratic equation $ax^2 + bx + c = 0$

(A.12)
$$\lambda = \frac{\sigma_x^2 + \sigma_y^2 \pm \sqrt{(\sigma_x^2 + \sigma_y^2)^2 - 4\sigma_x^2\sigma_y^2 + 4\sigma_{xy}^2}}{2}$$

$$\lambda_1 = \frac{\sigma_x^2 + \sigma_y^2 + \sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4\sigma_{xy}^2}}{2} = \sigma_u^2$$

$$\lambda_2 = \frac{\sigma_x^2 + \sigma_y^2 - \sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4\sigma_{xy}^2}}{2} = \sigma_v^2$$

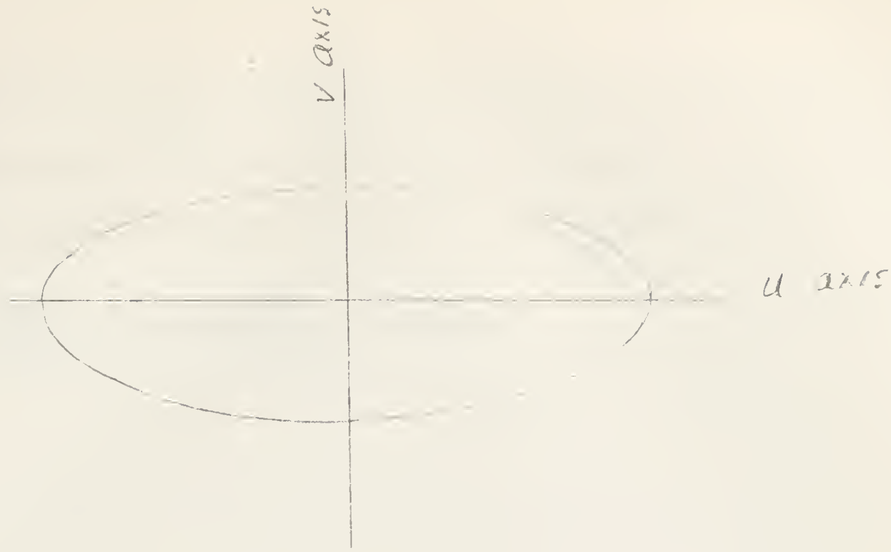
$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{pmatrix} = A^{*-1}$$

$$A^* = (A^{*-1})^{-1} = \begin{pmatrix} \frac{1}{\sigma_v^2} & 0 \\ 0 & \frac{1}{\sigma_u^2} \end{pmatrix}$$

The transformed density function is

(A.13)
$$g_{u,v}(u,v) = \frac{1}{2\pi |A^{*-1}|^{1/2}} \exp -\frac{1}{2} W'A^*W$$

We now have a normal bivariate density function with independent random variables U and V such that the matrix $W = \begin{pmatrix} V \\ U \end{pmatrix}$ is distributed $N(0, A^{*-1})$ and $W = C'Z$ where C is the orthogonal matrix defined which satisfies the relationship $C'A^{-1}C = A^{*-1}$. Therefore $C'Z$ is also distributed $N(0, A^{*-1})$ and the original terms involving the correlation constant do not appear in the quadratic form $Q = W'A^*W$ for this distribution. It should also be noted that $Q = (V,U)\begin{pmatrix} \frac{1}{\sigma_v^2} & 0 \\ 0 & \frac{1}{\sigma_u^2} \end{pmatrix}\begin{pmatrix} V \\ U \end{pmatrix}$. The orientation of the random variables U and V are thus orthogonal as shown in Figure A.11.



Ellipse Formed by Cutting the Bivariate
Normal Density Function $g_{uv}(u,v)$ by a Plane
Parallel to the u,v Axes.

Figure A.1

Fortunately it is not necessary to compute the orthogonal matrix C which satisfies the relationships above since the characteristics of the orthogonal matrix C requires that $C'C = I$ and $C' = C^{-1}$. Then

(A.14) $(W\ 'A^*(W\ = (C'Z\ 'A^*(C'Z\ =(Z\ 'CA^*C'(Z\ =(Z\ 'A(Z$

but $C'A^{-1}C=A^{*-1}$

therefore $(C'A^{-1}C)^{-1}=(A^{*-1})^{-1}=A^*$

$=C^{-1}AC=A^*$

$CC^{-1}ACC^{-1}=CA^*C^{-1}$

$A=CA^*C^{-1}$

Therefore, it can also be shown that the corresponding areas under the density functions are equal. That is

(A.15) $$\iint g_{U,V}(u,v)\,du\,dv=\iint f_{X,Y}(x,y)\,dx\,dy$$

where: $$g_{U,V}(u,v) = \frac{1}{2\pi\,|A^{*-1}|^{\frac{1}{2}}} \exp -\tfrac{1}{2}W'A^*W$$

$$f_{X,Y}(x,y) = \frac{1}{2\pi\,|A^{-1}|^{\frac{1}{2}}} \exp -\tfrac{1}{2}Z'AZ$$

This is because $W'A^*W = Z'AZ$ as shown above and $|A^{*-1}| = |A^{-1}|$. It is shown above that $C^{-1}A^{-1}C = A^{*-1}$ and the determinates of the two terms are $|C^{-1}AC| = |A^{*-1}| = |C^{-1}||A^{-1}||C| = |A^{*-1}| = |A^{-1}| \frac{|C|}{|C|} = |A^{-1}| = |A^{*-1}|$.

Therefore $\iint g_{U,V}(u,v)\,du\,dv = \iint f_{X,Y}(x,y)\,dx\,dy$

APPENDIX B

REMARKS ON ESTIMATION

B.1 Introduction

The theory of estimation is concerned with the problem of finding functions of the observations such that the distribution of these functions will be concentrated as closely as possible near the true values of the parameters estimated. The density function of the observations under consideration was described in Section I and the parameters which are to be estimated are $u_x, u_y, \sigma_x^2, \sigma_y^2$ and ρ. Some of the properties which are desired of the estimators were described in Section 1.3.

B.2 Maximum Likelihood Estimation

If $f(x_1, x_2 \ldots\ldots x_n, y_1, y_2 \ldots\ldots y_n, u_x, u_y, \sigma_x^2, \sigma_y^2, \rho)$ is the density function for a random sample of size n with unknown parameters $u_x, u_y, \sigma_x^2, \sigma_y^2$, and ρ, then the likelihood function is

$$\text{(B.1)} \quad L = \prod_{i=1}^{N} f(X_i, Y_i; u_x, u_y, \sigma_x^2, \sigma_y^2, \rho)$$

$$= \prod_{i=1}^{N} \frac{\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{X_i - u_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{X_i - u_x}{\sigma_x}\right)\left(\frac{Y_i - u_y}{\sigma_y}\right) + \left(\frac{Y_i - u_y}{\sigma_y}\right)^2\right]\right\}}{2\pi\,\sigma_x\,\sigma_y\sqrt{1-\rho^2}}$$

Since it is more convenient to deal with sums than products, it is easier to maximize the logarithm of the likelihood function rather than the likelihood function itself. It should be noted that the logarithm has its maximum at the same point as does the likelihood function. The log of (B.1) is

$$\text{(B.2)} \quad L' = -N \log 2\pi - \frac{N}{2}\log \sigma_x^2 - \frac{N}{2}\sigma_y^2 - \frac{N}{2}(1-\rho^2)$$

$$- \frac{1}{2(1-\rho^2)} \sum_{i=1}^{N} \left[\left(\frac{X_i - u_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{X_i - u_x}{\sigma_x}\right)\left(\frac{Y_i - u_y}{\sigma_y}\right) + \left(\frac{Y_i - u_y}{\sigma_y}\right)^2\right]$$

The maximum likelihood estimate of each of the unknown parameters is obtained by setting the derivative of the function with respect to each of the unknown parameters equal to zero and then solving the resulting equations simultaneously. To illustrate this procedure, the assumptions will be made that $\sigma_x^2 = \sigma_y^2 = \sigma^2$ and $\rho = 0$. For this special case, formula (B.2) becomes

$$\text{(B.3)} \quad L' = -N \log 2\pi - N \log \sigma^2 - \tfrac{1}{2} \sum_{i=1}^{N} \left[\frac{(X_i - u_x)^2 + (Y_i - u_y)^2}{\sigma^2} \right],$$

and the partial derivatives are

$$\frac{\partial (L')}{\partial \sigma^2} = -\frac{N}{\sigma^2} + \frac{1}{2} \sum_{i=1}^{N} \left[\frac{(X_i - u_x)^2 + (Y_i - u_y)^2}{\sigma^4} \right]$$

$$\frac{\partial (L')}{\partial u_x} = \sum_{i=1}^{N} (X_i - u_x) = \sum_{i=1}^{N} X_i - N u_x$$

$$\frac{\partial (L')}{\partial u_y} = \sum_{i=1}^{N} (Y_i - u_y) = \sum_{i=1}^{N} Y_i - N u_y .$$

Equating the partial derivatives to zero and solving simultaneously, it follows that,

$$\text{(B.4)} \quad \hat{u}_x = \frac{\sum_{i=1}^{N} X_i}{N} = \bar{X}$$

$$\text{(B.5)} \quad \hat{u}_y = \frac{\sum_{i=1}^{N} Y_i}{N} = \bar{Y}$$

$$\text{(B.6)} \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^{N} \left[(X_i - \bar{X})^2 + (Y_i - \bar{Y})^2 \right]}{2N} .$$

Since maximum likelihood estimators are in general biased estimators, it is necessary to examine them to see whether they are unbiased. For example, if the expected value of the estimator $\hat{\theta}$ is equal to $B_1 \theta$ where θ is the true parameter, then

(B.7) $E\left(\frac{\hat{\theta}}{B_1}\right) = \theta$ and $\frac{\hat{\theta}}{B_1}$ is an unbiased estimator.

Since $E(x_i) = u_x$ and $E(y_i) = u_y$ for $i = 1,\ldots\ldots,n$, it follows that $\bar{x}$ and $\bar{y}$ are unbiased estimators for u_x and u_y respectively.

The expected value of the estimator in formula (B.6) is obtained by recognizing the fact that there are 2(n-1) independent squares in the sum and therefore $2n\,\hat{\sigma}^2/\sigma^2$ is a chi squared random variable with 2(n-1) degrees of freedom as defined in formula (4.5). Since the expected value of a chi squared random variable is equal to its degrees of freedom, it follows that

(B.8) $$E\left(\frac{2N\hat{\sigma}^2}{\sigma^2}\right) = 2(n-1).$$

Therefore,

(B.9) $$E(\hat{\sigma}^2) = \frac{N-1}{N}\sigma^2 = B_1\sigma^2$$

and $$\frac{\hat{\sigma}_1^2}{B_1} = \sum_{i=1}^{N}\frac{[(X_i-\bar{X})^2+(Y_i-\bar{Y})^2]}{2(N-1)}$$ is an

unbiased estimator of σ^2 when the variances are equal. When the variances are not equal, the same procedure may be used and the unbiased estimators of σ_x^2 and σ_y^2 are

(B.10) $$\hat{\sigma}_x^2 = \sum_{i=1}^{N}\frac{(X_i-\bar{X})^2}{N-1}$$

$$\hat{\sigma}_y^2 = \sum_{i=1}^{N}\frac{(Y_i-\bar{Y})^2}{N-1}.$$

It should be noted that the estimators (B.10) are used in Model II. Also, if the assumption is made that the true values of the means are

zero, then the estimators for the variances in Model I are also unbiased. The estimators of the means and variance used in Model I and Model II are, apart from the biasing factors, maximum likelihood estimators.

However, it should be noted that the CEP is a function of the standard deviation and not the variance. The following section will determine unbiased estimators for the CEP using the procedure in this section.

B.3 Unbiased "Maximum Likelihood" Estimate Of The CEP: When

$$\sigma_x^2 = \sigma_y^2 = \sigma^2 \text{ and } \rho = 0.$$

The maximum likelihood function of σ when $u_x = u_y = 0$ is

$$\text{(B.11)} \qquad \hat{\sigma}_1 = \sqrt{\frac{1}{2N}\sum_{i=1}^{N}(X_i^2 + Y_i^2)}$$

Since the sum in (B.11) divided by σ^2 has a chi squared distribution, it follows that the square root of a chi squared random variable divided by its degrees of freedom has a chi distribution. The density function of a chi distributed random variable with 2n degrees of freedom is

$$\text{(B.12)} \qquad f_u(u) = \frac{2N^N u^{2N-1} \exp(-Nu^2)}{\Gamma(N)} \qquad u > 0$$

$$= \qquad 0 \qquad u \leq 0$$

where $\Gamma(n)$ is the gamma function with parameter n.

Then,

(B.13) $$E(\hat{\sigma}_1) = \frac{\Gamma(N+\frac{1}{2})\,\sigma}{\Gamma(N)\sqrt{N}}$$ and

$$\hat{\hat{\sigma}}_1 = \frac{\Gamma(N)\sqrt{N}}{\Gamma(N+\frac{1}{2})} \sqrt{\frac{1}{2N}\sum_{i=1}^{N}(X_i^2+Y_i^2)}$$ is an

unbiased estimate of σ and therefore

(B.14) $\widehat{CEP}^{**}_{1} = 1.1774\,\hat{\hat{\sigma}}_1$ is an unbiased estimate of the CEP.

The maximum likelihood estimator of σ when the means are not zero is

(B.15) $$\hat{\sigma}_2 = \sqrt{\frac{1}{2N}\sum_{i=1}^{N}\left[(X_i-\bar{X})^2+(Y_i-\bar{Y})^2\right]}$$

Therefore,

(B.16) $$E(\hat{\sigma}_2) = \frac{\Gamma(N-\frac{1}{2})}{\Gamma(N-1)\sqrt{N}}\,\sigma$$

and

(B.17) $$\hat{\hat{\sigma}}_2 = \frac{\Gamma(N-1)\sqrt{N}}{\Gamma(N-\frac{1}{2})}\sqrt{\frac{\sum_{i=1}^{N}\left[(X_i-\bar{X})^2+(Y_i-\bar{Y})^2\right]}{2N}}$$ is

an unbiased estimate of σ. Therefore

(B.18) $\widehat{CEP}^{**}_{2} = 1.1774\,\hat{\hat{\sigma}}_2$ is an unbiased estimate of the CEP.

The reader may be interested in the magnitudes of the biasing factors and a comparison of the biased and unbiased estimators of the CEP. The results obtained using the data from the sample problems are presented in Tables i and j.

Table i		
Comparison of the Biasing Factors of the Two Estimators		
Case	CEP_1^{**}	CEP_2^{**}
1	$B_{11} = \frac{\sqrt{10}\,\Gamma(10)}{\Gamma(10.5)} = 1.01$	$B_{21} = \frac{\sqrt{10}\,\Gamma(9)}{\Gamma(9.5)} = 1.09$
2	$B_{12} = \frac{\sqrt{15}\,\Gamma(15)}{\Gamma(15.5)} = 1.01$	$B_{22} = \frac{\sqrt{15}\,\Gamma(14)}{\Gamma(14.5)} = 1.04$
3	$B_{13} = \frac{\sqrt{25}\,\Gamma(25)}{\Gamma(25.5)} = 1.005$	$B_{23} = \frac{\sqrt{25}\,\Gamma(24)}{\Gamma(24.5)} = 1.03$

Table j						
Comparison of the Estimators with the Methods Used In Sections II and III						
Problem	Appendix B	Section II		Appendix B	Section III	
	$\widehat{CEP}_1^{**}$	$\widehat{CEP}_1$	$\widehat{CEP}_1^{*}$	$\widehat{CEP}_2^{**}$	$\widehat{CEP}_2$	$\widehat{CEP}_2^{*}$
1						
Case 1	4.20	3.97	4.13	3.86	3.64	3.72
Case 2	4.13	3.84	4.10	3.95	3.87	3.88
Case 3	3.58	3.48	3.51	3.29	3.28	3.26
2						
Case 1	3.55	3.37	3.50	3.47	3.33	3.55
Case 2	3.59	3.45	3.51	3.48	3.39	3.52
Case 3	3.83	3.77	3.78	3.78	3.71	3.74
3						
Case 1	4.76	4.17	4.69	4.76	4.66	5.02
Case 2	4.10	3.65	4.03	3.92	3.56	4.21
Case 3	3.69	3.36	3.66	3.78	3.52	3.72

B.4 Comparison Of The Two Estimates: Relative Efficiency

Throughout this section it is assumed that $\sigma_x^2 = \sigma_y^2 = \sigma^2$ and $\rho = 0$. It can be proven that $\widehat{CEP}_1^{**}$ has greater efficiency than any other unbiased linear sample statistic when the mean value is (0,0). In case

the mean is not zero but is known to be small, this estimate should be considered. $\widehat{CEP}_2^{**}$ is asymptotically efficient whatever the population mean may be, hence, if the mean is greatly different from (0,0), $\widehat{CEP}_2^{**}$ will be a better estimate than $\widehat{CEP}_1^{**}$. However, because 2 degrees of freedom are lost in estimating the coordinates of the mean, the estimate $\widehat{CEP}_2^{**}$ will not be as precise as $\widehat{CEP}_1^{**}$ for small values of (u_x, u_y).

In order to determine whether to use $\widehat{CEP}_1^{**}$ or $\widehat{CEP}_2^{**}$ when it is known that the true mean is close to (0,0), it is necessary to compare the two estimates by some criterion. The method which will be used is the ratio of the relative efficiencies. When $\widehat{CEP}_1^{**}$ and $\widehat{CEP}_2^{**}$ are used, the formula is

(B.19) $$R.F. = \frac{E[(\widehat{CEP}_1^{**} - CEP)^2]}{E[(\widehat{CEP}_2^{**} - CEP)^2]} = \frac{E[(\hat{\hat{\sigma}}_1 - \sigma)^2]}{E[(\hat{\hat{\sigma}}_2 - \sigma)^2]}$$

This comparison may be done by assuming that the true mean is either some point (u_x, u_y) or (0,0). In the case that the assumption is made that the true mean is (u_x, u_y) the joint density function is

(B.20) $$f(x,y;\, u_x, u_y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}\left[(x-u_x)^2 + (y-u_y)^2\right]\right\}$$

When it is assumed that the true mean is (0,0), the joint density function is

(B.21) $$g(x,y;\, 0,0,\sigma) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2}(x^2+y^2)\right]$$

The development of the ratio assuming that the true mean is (0,0) follows the procedure applied in formula (B.13). The result is

(B.22) $$E\left[(\hat{\hat{\sigma}}_1 - \sigma)^2\right] = \left[\frac{\Gamma(N)\,\Gamma(N+1)}{\Gamma^2(N+\frac{1}{2})} - 1\right]\sigma^2$$

$$(B.23)\quad E\left[\left(\hat{\hat{\sigma}}_2 - \sigma\right)^2\right] = \left[\frac{\Gamma(N)\Gamma(N-1)}{\Gamma^2(N-\frac{1}{2})} - 1\right]\sigma^2$$

Combining formulas (B.22) and B.23), the ratio function is

$$(B.24)\quad R.F. = \frac{\dfrac{\Gamma(N)\Gamma(N+1)}{\Gamma^2(N+\frac{1}{2})} - 1}{\dfrac{\Gamma(N)\Gamma(N-1)}{\Gamma^2(N-\frac{1}{2})} - 1}$$

When the mean is (0,0), the ratio function in (B.24) is less than 1 for all n. Table k presents values of the ratio function for n = 2(1)20, 25(5)50. P.B. Moranda tables this ratio for n = 2(1)8.

Table k

Values Of The Ratio Function When $u_x = u_y = 0$

n	R.F.	n	R.F.
2	.482	25	.959
3	.656	30	.966
4	.743	35	.971
5	.795	40	.974
6	.830	45	.977
7	.854	50	.979
8	.873		
9	.887		
10	.898		
11	.908		
12	.915		
13	.922		
14	.927		
15	.932		
16	.937		
17	.940		
18	.944		
19	.947		
20	.949		

If it is known that the true mean is at some point (u_x, u_y) then formula (B.20) is the joint density function of the component errors. The R.F. ratio for this case was developed by P.B. Moranda in reference

(3). In order to find the mean square deviation of $\widehat{CEP}_2^{**}$, the same procedure can be followed as in formula (B.13) and the result is the same as formula (B.22). The mean square error of $\widehat{CEP}_1^{**}$ is a function of u_x and u_y. Moranda assumed for ease of computation that $u_x = k_1 \overline{V}$ and $u_y = k_2 \overline{V}$.

Letting u be defined by

$$\text{(B.25)} \quad u = \frac{\sum_{i=1}^{n} (X_i^2 + Y_i^2)}{\overline{V}^2}$$

u has a non-central chi squared distribution. Values of R.F. shown in Table 1 (an erpt from Table (1) in reference (3)) were obtained by putting $k_1 = k_2$, and varying k from 0(.1)1.0. The results of this derivation show that as n increases, the ratio function decreases for a constant value of k. It can be ascertained from this table that for large n, $\widehat{CEP}_2^{**}$ will be the best estimate unless k equals zero and $\widehat{CEP}_1^{**}$ will be best for small n and small values of k. The practical use of the ratio under these assumptions require the use of estimates to obtain the values of k_1 and k_2 and although not exact, may still supply some useful information.

Table 1

$\left(E\,(CEP_1 - CEP)^2 \;/\; E\,(CEP_2 - CEP)^2 \right)$

n \ $k_1 (=k_2)$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
2	.482	.487	.503	.530	.575	.639	.723	.849	1.01	1.21	1.45
3	.656	.663	.685	.727	.796	.903	1.06	1.28			
4	.743	.751	.777	.829	.919	1.06	1.28				
5	.796	.804	.834	.893	1.00	1.18					
6	.830	.838	.860	.937	1.06	1.28					
7	.854	.864	.896	.972	1.12						
8	.873	.884	.917	1.00	1.15						

A possible procedure for using Table 1 is as follows:

The values of $\bar{x}$ and $\bar{y}$ are first computed. Then $\hat{V}_1$ and $\hat{V}_2$ are computed using formulas (B.11) and (B.15) respectively. The estimated value of k_1 and k_2 will then equal

$$(B.29) \quad k_1 = \frac{\bar{x}}{\hat{V}} \qquad \text{where} \quad \hat{V} = \frac{\hat{V}_1 + \hat{V}_2}{2}$$

$$k_2 = \frac{\bar{y}}{\hat{V}}$$

Using $k = \dfrac{k_1 + k_2}{2}$ and n, an analysis of table 1 may show when $\widehat{CEP}_1^{**}$ is not the best estimate. The reader should be cautioned that no attempt has been made to theoretically justify this procedure.

In order to better illustrate the above, the computed values from the example problems for case 1 are shown in table m.

Table m
Computed Values of k

Problem (N=10)	$\bar{x}$	$\bar{y}$	Model I	Model II	$\hat{V}$	k_1	k_2	$k=(k_1+k_2)/2$
			$\hat{V}_1$	$\hat{V}_2$				
1	1.2	2.0	3.5	3.2	3.3	.364	.605	.484
2	1.0	.1	3.0	3.0	3.0	.333	.033	.183
3	.6	.3	4.1	4.2	4.1	.146	.073	.109

Analysis of Table m Using the Above Values

1 for k = .484, R.F. $>$ 1 for all n $>$5, therefore $\widehat{CEP}_2^{**}$ is best estimate.

2 for k = .20, R.F. = .917 $<$ 1, for n = 8, therefore $\widehat{CEP}_1^{**}$ is slightly better.

3 for k = .109, R.F. = .884 $<$ 1 for n = 8. $\widehat{CEP}_1^{**}$ is better.

APPENDIX C

INTEGRATION OF THE BIVARIATE GAUSSIAN DISTRIBUTION

C.1 Introduction

This appendix discusses the details of the integration introduced in Sections II and IV. The integration of an ellipse or a circle over the bivariate normal density function may be simplified by making the transformation explained in Appendix A. That is from equation (A.15)

(C.1) $$\iint f_{x,y}(x,y)\,dx\,dy = \iint g_{u,v}(u,v)\,du\,dv \quad \text{where}$$

$$g_{u,v}(u,v) = \frac{1}{2\pi\,\sigma_u\,\sigma_v}\exp\left[-\frac{1}{2}\left(\frac{u^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)\right].$$

C.2 Integration Over Circle[6]

The probability that a random point (U,V) will lie within a circle with center at the origin and radius $k\,\sigma_u$ is written as

(C.2) $$P\left(\sqrt{U^2+V^2} \le k\,\sigma_u\right) = \iint_{\sqrt{u^2+v^2}\le k\sigma_u} g_{u,v}(u,v)\,du\,dv\ .$$

The two illustrations in Figure C.1 show the geometric area of integration.

[6] "Circular Error Probabilities" by H. Leon Harter /4/ of Aeronautical Research Laboratories.

Integration of the Bivariate Density Function Over a Circular Region

Figure C.1

In order to simplify equations (C.1) and (C.2) let

(C.3) $\frac{u}{\sigma_u} = m\cos\theta, \quad \frac{v}{\sigma_u} = m\sin\theta,$

then

$$P(K, \sigma_u \sigma_v) = \frac{1}{2\pi \sigma_u \sigma_v} \iint_{M<K} \exp\left\{-\frac{1}{2}\left[M^2\cos^2\theta + \frac{\sigma_u^2}{\sigma_v^2} M^2 \sin^2\theta\right]\right\} |J| \, dm\, d\theta$$

where

$$J = \begin{vmatrix} \frac{\partial u}{\partial m} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial m} & \frac{\partial v}{\partial \theta} \end{vmatrix} = m\sigma_u^2$$

now let $C = \frac{\sigma_v}{\sigma_u}$, $\sin^2\theta = 1 - \cos^2\theta$ and the probability is

$$P(K,C) = \frac{1}{2\pi C} \int_0^{2\pi}\int_0^{K} \exp\left\{-\frac{M^2}{C^2}\left[(C^2-1)\cos^2\theta + 1\right]\right\} m\, dm\, d\theta$$

let $\cos^2\theta = \frac{1}{2}(1+\cos 2\theta)$, $\phi = 2\theta$, $Z = \frac{m^2}{4C^2}$ and the probability is

$$P(K,C) = \frac{2C}{\pi} \int_0^{\pi}\int_0^{K^2/4C^2} \exp\left\{-Z\left[(C^2+1) + (C^2-1)\cos\phi\right]\right\} dZ\, d\phi$$

Then integrate with respect to ϕ and the probability becomes

$$(C.5) \quad P(K,C) = \frac{2C}{\pi} \int_0^{\pi} \frac{1 - \exp\left\{-\frac{K^2}{4C^2}\left[(C^2+1) + (C^2-1)\cos\phi\right]\right\}}{(C^2+1) + (C^2-1)\cos\phi} \, d\phi$$

This form is integrated using the trapozoidal rule and utilizing computers to do the integrating.

For example, the curve below represents some function that we wish to integrate over the designated interval. We can divide the interval into equal sub intervals ($d\phi$) and sum all of the sub intervals. As the sub intervals become smaller, the accuracy of this type of integration becomes better and this summation technique approaches the actual area under the curve.

Trapozoidal Technique of Integrating Under a Curve by Summation

Figure C.2

The formulation for integration with n intervals thus becomes

$$(C.6) \quad P(K,C) = \frac{2C}{\pi} \sum_{m=0}^{n} \frac{1-\exp\left\{-(K^2/4C^2)\left[(C^2+1) + (C^2-1)\cos m\theta\right]\right\}}{(C^2+1) + (C^2-1)\cos m\theta}.$$

where $\theta = \frac{\pi}{n}$

In computer techniques, the COSINE function is usually changed to the Chebyschev polynomial or some other scheme that converges more rapidly than the COSINE function. This can be shown by comparing the convergence of the two methods. If we let $\cos\theta = x = T_1(x)$, then

$T_m(x) = \cos m\theta = T_m(\cos\theta)$

The formula for integration now becomes

$$\text{(C.7)}\quad F(k,c) = \frac{2c}{\pi} \sum_{m=0}^{N} \frac{1 - \exp\left(-\frac{k^2}{4c^2}\left[(c^2+1)+(c^2-1)\,T_m(x)\right]\right)}{(c^2+1)+(c^2-1)\,T_m(x)} \;.$$

This summation is now made for different values of k,c, and F(k,c).

C.3 Integrating Over an Ellipse

The probability that a random point (U,V) will lie within an ellipse with center at the origin is written as

$$\text{(C.8)}\quad P\left[\sqrt{\frac{u^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}} < k\right] = \frac{1}{2\pi\,\sigma_u\,\sigma_v} \iint\limits_{\sqrt{\frac{u^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}}<k} \exp\left\{-\tfrac{1}{2}\left(\frac{u^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right\} du\,dv \;.$$

The two illustrations in figure C.3 show the geometric area of integration

Volume of Integration of Bivariate Density Function

Ellipse Formed by Plane Cutting $g_{U,V}(u,v)$ Parallel to u,v Plane

Figure C.3

Note that if the variances are equal, this form is also circular, the three dimensional form being a perfect bell and the two dimensional form being a circle.

In order to further simplify this form let

(C.9) $u = m\sigma_u \cos\theta$, $v = m\sigma_v \sin\theta$, then the probability becomes

$$P(k,m,\theta) = \frac{1}{2\pi\sigma_u\sigma_v}\iint_{m<k} \exp -\tfrac{1}{2}m^2 \, |J| \, dm\, d\theta, \quad \text{where } J = m\sigma_u\sigma_v$$

thus,

(C.10) $$P(k,m,\theta) = \frac{1}{2\pi}\int_0^k \exp\left\{-\tfrac{1}{2}m^2\right\} m\, dm\, d\theta,$$

Formula C.10 can be integrated directly by first integrating with respect to θ and then with respect to m. After integrating with respect to θ, the formula becomes

(C.11) $$P(k,m) = \int_0^k \exp(-\tfrac{1}{2}m^2)\, m\, dm.$$

If we let $t=m^2$, the probability statement becomes

(C.12) $P(k,t) = \frac{1}{2}\int_0^{k^2} \exp -\frac{1}{2}t \, dt$ where $f_c(t)$ is the chi squared density function with two degrees of freedom as defined in formula (4.5)

That is

(C.13) $$P[T_a < K^2] = P\left[\sqrt{\frac{u^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}} < K^2\right] = 1 - e^{-\frac{K^2}{2}}.$$

For any value of P(k,t) the value of k^2 can be obtained from table 4 by entering with P(k,t) and 2 degrees of freedom.