AN ABSTRACT OF THE THESIS OF

Marvin Edwin Harrell for the Master of Science

in Mathematics presented on July 17, 1987

Title: Methods for Estimating The Radius of a Circle

Abstract approved: L. Scott

Statisticians are often confronted with the estimation of unknown parameters. This process of estimation, by far, is not unique. The purpose of this thesis is to examine various methods to estimate the radius of a circle when observing n data points that are assumed to lie on a semi-circle but are measured with error. The estimators mentioned in this paper involve the use of the circumscribed circle, the radius of curvature, and a residual approach. Computer simulations are used in the evaluation of the estimators. A representative listing of the simulations are provided.

METHODS FOR ESTIMATING THE RADIUS OF A CIRCLE

A Thesis

Presented to

the Department of Mathematics

EMPORIA STATE UNIVERSITY

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

Marvin E. Harrell

July, 1987

Approved for the Major Department

Approved for the Graduate Council

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

PAGE

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to Dr. Larry Scott, who suggested the problem, and for his helpful criticism and his direction in the writing of this thesis. Furthermore, I'd like to thank the instuctors in the Mathematics Department for their help and guidance during my time at E.S.U. Most of all, I'd like to thank my wife for tolerating the many sleepless nights while I was working on this thesis.

CHAPTER 1

INTRODUCTION

Statisticians are often confronted with the estimation of unknown parameters. The process of estimating such a parameter is not, by far, unique. Many times the job of the statistician is to find a method that gives a reasonable estimate of the parameter.

The purpose of this thesis is to investigate methods that estimate the radius R of a circle when observing n points that lie on the circle. Chapter 1 will discuss the underlying assumptions of the model and provide meaningful applications. Methods for estimating the radius of the circle by averaging the radii of known circles are given in Chapter 2. On the other hand, Chapter 3 deals with the radius of curvature to determine an estimate for R. In Chapter 4 the writer will look at the minimization of the sum of squared residuals to aid in estimating the radius. The feasibility and comparisons of methods through computer simulations will be discussed in Chapter 5. Chapter 6 will include a summary of the thesis and final comments about estimating parameters.

Statement of the Problem and Assumptions

As with any statistical model the underlying assumptions must be known. The problem is to estimate the radius R of a circle when given n ($n \geq 3$) data points that lie on a semi-circle but are measured with error. Let the points have coordinates (X_i, Y_i). The Y_i's are assumed to be measured with negligible error. However, the X_i 's are from a random variable and measured with error ε_i. The error term is a random variable that is normally distributed with zero mean and unknown variance σ^2. As can be seen from ILLUSTRATION 1, $R = [(X_i - h - \varepsilon_i)^2 + (Y_i - k)^2]^{\frac{1}{2}}$ where (h, k) is the unknown center of the circle.

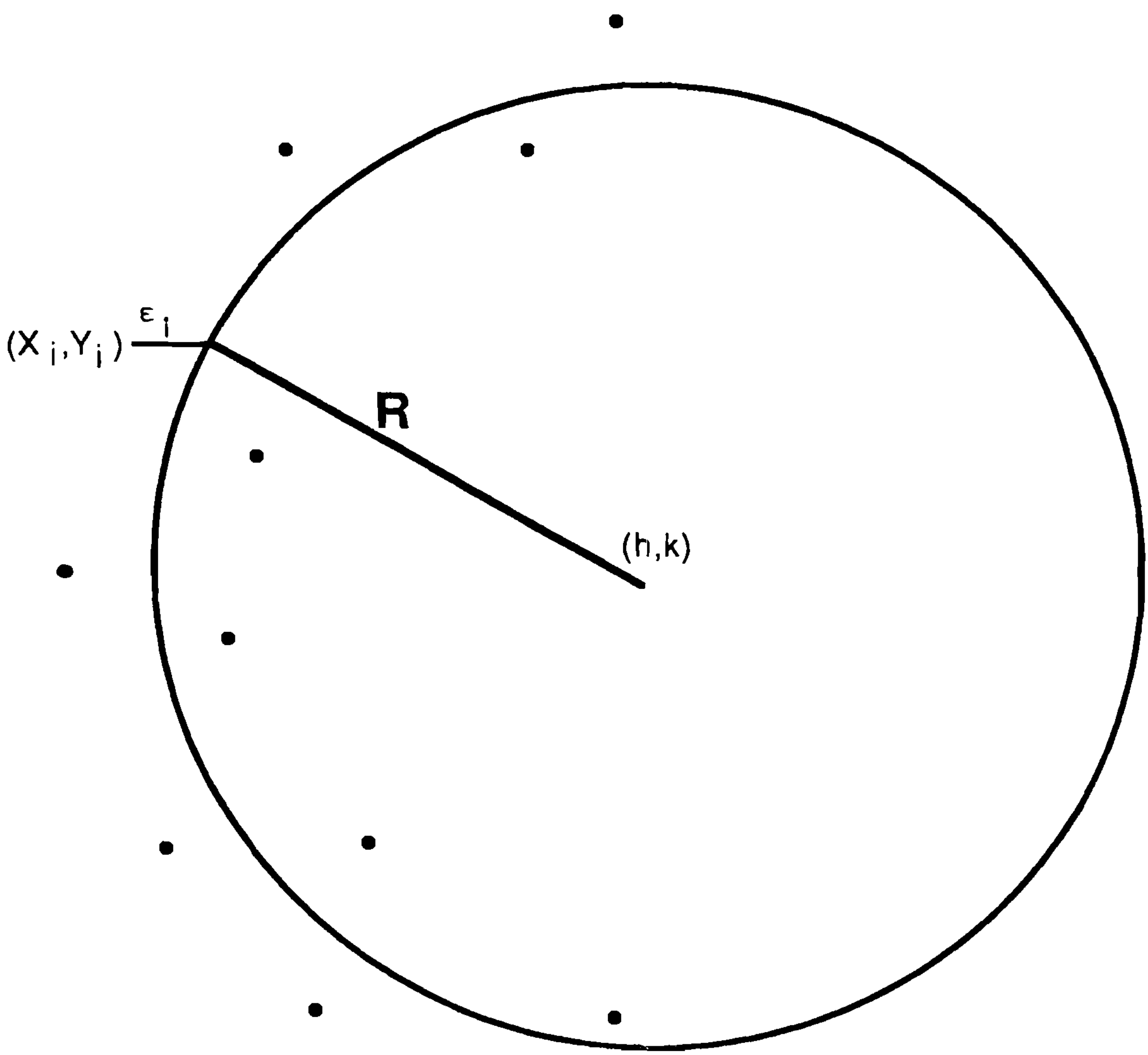

ILLUSTRATION 1

Throughout this paper when conditions are imposed on the n observed points, the assumption is made that $X_i \leq 0$ (i = 1, 2, ..., n). However, if some of the observed values for X_i are positive, then these results would apply after a translation or rotation of axes.

Applications

A statistician is generally concerned with summarizing data. With the assumption that the observations come from a semi-circle, one such way of summarizing the data is to estimate its radius. In the medical field a researcher may be interested in the change of the curvature of a woman's lower back during pregnancy. Observations can be taken with a measuring device. If the values of Y are measured with negligible error, then the researcher can use the methods discussed in this thesis, to estimate the radius of the circle from which the observations are assumed to be taken. The estimate of R can provide an indication of the change in curvature; as R decreases the curvature would increase. Looking at this illustration, a variety of other possible applications can be seen. For instance, taking observations from a fixed circular arch one could estimate the radius of the circle from which the arch originated. Within the reader's surroundings, there are many possible applications for estimating the radius of a circle from which n points are observed.

CHAPTER 2

CIRCUMSCRIBED CIRCLE

From geometry, one knows that three noncollinear points determine a circle, better known as the circumscribed circle of the triangle formed by the three noncollinear points. Taking n ($n \geq 3$) observations the method to be discussed in this chapter will involve averaging the R_i 's ($i = 1,2,\ldots,t$), where each R_i is the radius of the circumscribed circle associated with a single combination of three noncollinear points and $t \leq {}_nC_3$. The discussion of why $t \leq {}_nC_3$ and some basic assumptions will be considered later. The writer will first look at two approaches for finding the radius of the circumscribed circle, one being an algebraic/geometric approach, and the other using systems of equations and matrix theory.

The Algebraic/Geometric Approach

Given three noncollinear points, say $N(x_1, y_1)$, $Q(x_2, y_2)$ and $P(x_3, y_3)$, calculate the midpoint of line segment NQ, also denoted as $\overline{NQ}$, and the slope associated with $\overline{NQ}$. The midpoint $M_1 = (\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2})$ and the slope $m_1 = \frac{y_2 - y_1}{x_2 - x_1}$. Having these two pieces of information one can then calculate the equation of the line that is perpendicular to $\overline{NQ}$ and passes through the midpoint M_1. Using the point-slope formula, the equation of the line is

$$(y - \frac{y_1 + y_2}{2}) = \frac{x_1 - x_2}{y_2 - y_1} (x - \frac{x_1 + x_2}{2})$$

Thus,

$$Y = \frac{x_1 - x_2}{y_2 - y_1} X + \frac{x_2^2 - x_1^2}{2(y_2 - y_1)} + \frac{y_1 + y_2}{2}, \text{ and hence}$$

$$= \frac{x_1 - x_2}{y_2 - y_1} X + \frac{(x_2^2 + y_2^2) - (x_1^2 + y_1^2)}{2(y_2 - y_1)} . \qquad \text{(Eq. 2.1)}$$

Similarly, using $\overline{QP}$

$$Y = \frac{x_2 - x_3}{y_3 - y_2} X + \frac{(x_3^2 + y_3^2) - (x_2^2 + y_2^2)}{2(y_3 - y_2)} . \qquad \text{(Eq. 2.2)}$$

Since N, Q, and P are noncollinear, the slope of Eq. 2.1 does not equal the slope of Eq. 2.2 thus the two lines must intersect at some point, say C(h, k) where

$$h = \frac{(y_3 - y_2)[(x_2^2 + y_2^2) - (x_1^2 + y_1^2)] + (y_1 - y_2)[(x_3^2 + y_3^2) - (x_2^2 + y_2^2)]}{2[(x_2 - x_1)(y_3 - y_2) + (x_3 - x_2)(y_1 - y_2)]} \qquad \text{(Eq. 2.3)}$$

and

$$k = \frac{(x_2 - x_3)[(x_2^2 + y_2^2) - (x_1^2 + y_1^2)] + (x_2 - x_1)[(x_3^2 + y_3^2) - (x_2^2 + y_2^2)]}{2[(y_2 - y_1)(x_2 - x_3) + (y_3 - y_2)(x_2 - x_1)]} \qquad \text{(Eq. 2.4)}$$

Point C is the center of the circumscribed circle of triangle NQP. To find the radius of the circumscribed circle one may use the distance formula, $d(P_1, P_2) = [(x_2 - x_1)^2 + (y_2 - y_1)]^{\frac{1}{2}}$ where $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$. Thus the radius of a circumscribed circle

$$R = d(C, T) = [(x - h)^2 + (y - k)^2]^{\frac{1}{2}} \qquad \text{(Eq. 2.5)}$$

where T is either N, Q, or P.

For the geometer the following is a proof that three noncollinear points must lie on a circle. Illustration 2 is to graphically aid in the proof.

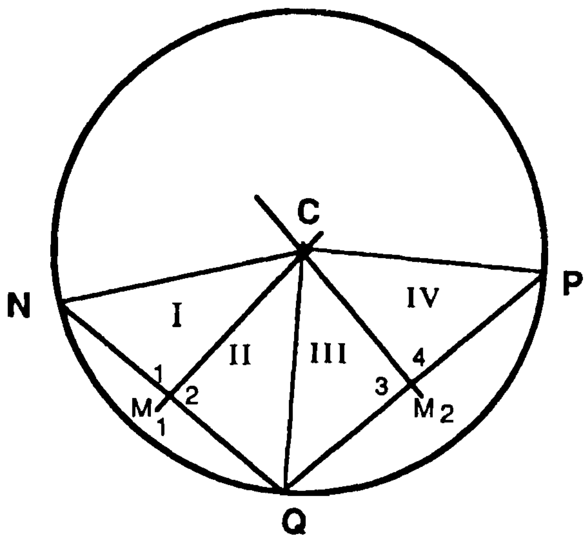

ILLUSTRATION 2

Proof

Given three noncollinear points N, Q, and P. Construct the perpendicular bisectors of $\overline{NQ}$ and $\overline{QP}$. Since N, Q, and P are noncollinear the perpendicular bisectors must meet at a point, say C. Furthermore, by the definition of a perpendicular bisector, the bisector must pass through the midpoint of the line segment, for $\overline{NQ}$ name the midpoint M_1 and for QP name it M_2. Also, <1, <2, <3, and <4 are right angles. Thus, <1 = <2 and <3 = <4. $NM_1 = M_1Q$ and $QM_2 = M_2P$ by the definition of a midpoint. By the reflexive property of equality $M_1C = M_1C$ and $M_2C = M_2C$. Hence, by the side-angle-side theorem $\Delta I = \Delta II$ and $\Delta III = \Delta IV$. Thus NC = QC and QC = PC, since corresponding parts of congruent triangles are congruent. Therefore, NC = QC = PC which implies that points N, Q, and P lie on a circle whose center is C and whose radius can be taken as $\overline{NC}$, $\overline{QC}$, or $\overline{PC}$.

The following is an example of the method described above. Given points N(-2, 2), Q(0, 0) and P(2, 2). First calculate m and M ,

$$m_1 = \frac{0 - 2}{0 + 2} = -1 \qquad \text{and } M_1 = \left(\frac{-2 + 0}{2}, \frac{2 + 0}{2}\right) = (-1, 1)$$

and m_2 and M_2,

$$m_2 = \frac{2 - 0}{2 - 0} = 1 \qquad \text{and } M_2 = \left(\frac{0 + 2}{2}, \frac{0 + 2}{2}\right) = (1, 1).$$

Next find the equations of the perpendicular bisectors of $\overline{NQ}$ and $\overline{QP}$. The equation of the perpindicular bisector of $\overline{NQ}$ is

$$(y - 1) = 1(x + 1) \quad ==> \quad y = x + 2$$

and the equation of the perpendicular bisector of $\overline{QP}$ is

$$(y - 1) = -1(x - 1) \quad ==> \quad y = -x + 2$$

Solving the two equations by addition, the solution is the ordered pair (0, 2). One knows from the previous discussion that (0, 2) is the center of the circumscribed circle of triangle NQP. Now all that is needed is to find the radius, thus $R = d(C, Q) = [(0 - 0)^2 + (0 - 2)^2]^{\frac{1}{2}} = 2$.

However, there is no need to go through all the work mentioned above since Eq. 2.3 and Eq. 2.4 have been derived to find the abscissa and ordinate of the center. All that is needed is to compute the center and then calculate the radius. This is easily illustrated by using Eq. 2.3, Eq. 2.4 and Eq. 2.5 and the points previously used. Let C(h, k) be the center of the circumscribed circle.

From Eq. 2.3

$$h = \frac{(2-0)[(0^2+0^2)-((-2)^2+(2)^2)]+(2-0)[2^2+2^2)-(0^2+0^2)]}{2[(0+2)(2-0)+(2-0)(2-0)]}$$

$$= \frac{2(-8)+2(8)}{2(8)}$$

$$= 0 .$$

From Eq. 2.4

$$k = \frac{(0-2)[(0^2+0^2)-((-2)^2+2^2)]+(0+2)[(2^2+2^2)-(0^2+0^2)]}{2[(0-2)(0-2)+(2-0)(0+2)]}$$

$$= \frac{2(-8)+2(8)}{2(8)}$$

$$= 2 .$$

Hence, the center is (0, 2).

From Eq. 2.5

$$R = d(C, Q) = [(0-0)^2 + (0-2)^2]^{\frac{1}{2}} = 2 .$$

Thus through this example one can see that all that is needed is to calculate the center from Eq. 2.3 and Eq. 2.4, and then calculate R by equation Eq. 2.5.

Using Systems of Equation and Matrix Theory.

In many algebra and geometry books the general equation of a circle is given in the following form,

$x^2 + y^2 + cx + dy + e = 0$ where c, d, e $\in$ Reals.

Given three noncollinear points (x_1, y_1), (x_2, y_2), and (x_3, y_3) one can set up a system of three equations with three unknown. The system of equations would be as follows

$$\begin{Bmatrix} x_1^2 + y_1^2 + cx_1 + dy_1 + e = 0 \\ x_2^2 + y_2^2 + cx_2 + dy_2 + e = 0 \\ x_3^2 + y_3^2 + cx_3 + dy_3 + e = 0 \end{Bmatrix} \sim \begin{Bmatrix} x_1c + y_1d + e = -x_1^2 - y_1^2 \\ x_2c + y_2d + e = -x_2^2 - y_2^2 \\ x_3c + y_3d + e = -x_3^2 - y_3^2 \end{Bmatrix}$$

Using Guassian elemination on the following matrix,

$$\begin{bmatrix} x_1 & y_1 & 1 & (-x_1^2 - y_1^2) \\ x_2 & y_2 & 1 & (-x_2^2 - y_2^2) \\ x_3 & y_3 & 1 & (-x_3^2 - y_3^2) \end{bmatrix}$$

one can determine the parameters c, d and e.

After computing the parameters, one can then use completing the square to convert the general equation into the form, $(x - h)^2 + (y - k)^2 = R^2$, where (h, k) is the center of the circle and R is the radius. Thus the equation is $(x + \frac{c}{2})^2 + (y + \frac{d}{2})^2 = \frac{c^2 + d^2 - 4e}{4}$. Hence,

$$R = [\frac{c^2 + d^2 - 4e}{4}]^{\frac{1}{2}} = \frac{[c^2 + d^2 - 4e]^{\frac{1}{2}}}{2}. \qquad \text{(Eq. 2.6)}$$

The following is an example of the above method using points (-2, 2), (0, 0) and (2, 2). The system of equations would be:

$$\begin{Bmatrix} 4 + 4 - 2c + 2d + e = 0 \\ 0 + 0 - 0c + 0d + e = 0 \\ 4 + 4 + 2c + 2d + e = 0 \end{Bmatrix} \sim \begin{Bmatrix} -2c + 2d + e = -8 \\ 0c + 0d + e = 0 \\ 2c + 2d + e = -8 \end{Bmatrix}.$$

The matrix associated with this system of equations would be:

$$\begin{bmatrix} -2 & 2 & 1 & -8 \\ 0 & 0 & 1 & 0 \\ 2 & 2 & 1 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Thus $c = 0$, $d = -4$, $e = 0$.

Hence, by Eq. 2.6 $R = [\frac{0^2 + (-4)^2 - 4(0)}{4}]^{\frac{1}{2}} = [4]^{\frac{1}{2}} = 2$.

In the following sections of this chapter a discussion of two approaches dealing with the R_i's to estimate R will be presented.

Estimating R by Finding the Mean of the R_i's

Taking any set of three noncollinear points one can obtain an equation of a circle, thus, the radius R of the circle can be calculated. Therefore, given n ($n \geq 3$) points one can calculate at most ${}_nC_3$ R_i's. The reason for emphasizing finding at most ${}_nC_3$ R_i's is that there may be combinations of three points that are collinear. Having calculated all possible R_i one can then calculate the mean of the R_i's to be used as an estimate of R, that is, $\hat{R} = \sum_{i=1}^{t} R_i / t$ where t is the number of R_i's and $t \leq {}_nC_3$.

Example

Assume $n = 4$ and the following points are observed A(-4.27, 9), B(-12.59, 3), C(-10.32, -3), and D(-4.49, -9). Using ABC as a combination of three points $R_1 = 6.95755$. For the combination ABD $R_2 = 9.00521$. Likewise, ACD yields $R_3 = 9.52693$. $R_4 = 18.2247$ when using BCD as a combination of three points. Therefore,

$\hat{R} = \sum_{i=1}^{4} R_i / 4 = 10.9286.$

Finding R by Restricting the Selection of Points

In the previous, method all possible combinations of three points were considered in calculating the R_i's except when the three points were collinear. However, in the following method some conditions will be placed on the selection of points that are used to determine R. Again, any combination of three points must be noncollinear. Furthermore, assuming $y_1 > y_2 > y_3$ then $x_1 > x_2$ and $x_3 > x_2$. Any other combination of three points will not be considered because they do not conform to the known orientation of the circle. As in the previous method $\hat{R} = \sum_{i=1}^{t} R_i/t$ where $t \leq {}_nC_3$, but in this method one is concerned only with the R_i's that are obtained with the combination of three points meeting the conditions imposed.

Example

Assume $n = 4$ and the observed points are those in the previous example. Notice that the combination BCD does not meet the stated conditions, thus BCD is not used in the calculations of $\hat{R}$. Nevertheless, ABC, ABD, and ACD are used in calculating R. For the combination ABC $R_1 = 6.95755$. Using ABD $R_2 = 9.00521$. $R_3 = 9.52693$ when using ACD as a combination of three points. Thus, $\hat{R} = \sum_{i=1}^{3} R_i/3 = 8.49656$.

As can be seen by the two previous examples that both methods yield different estimates of R. The question becomes what method should be used. This question is not easily answered. If the investigator knows something about the orientation of the curve from which the n points are observed, then the second method may be more appropriate. If nothing is known about the curve, a further investigation is needed to determine which method is appropriate. In Chapter 5, the usefulness of these methods will be discussed as well as a comparison of these methods to the other methods mentioned in Chapter 1.

CHAPTER 3

THE RADIUS OF CURVATURE

Most people are first introduced to the notion of curvature in a beginning calculus course. In this chapter, two approaches dealing with conic sections and radii of curvature at points on the conic sections will be discussed. Furthermore, examples will be provided to illustrate the use of these different approaches.

Any set of five points, no three of which are collinear, determines the equation of a conic section. Implicitly differentiating this equation to find $\frac{dy}{dx}$ or y' and $\frac{d^2y}{dx^2} = y''$ one can then use the radius of curvature, ρ , to determine R. Observing n points there would be at most ${}_nC_5$ different equations of conic sections and at most $5({}_nC_5)$ values for ρ . A further discussion of this notion will be involved with the development of the assumptions that are necessary for this model. However, first, there is a need to discuss and illustrate how to find the equation of a conic section given five points no three being collinear. Secondly, it will be necessary to define curvature and the radius of curvature, and to derive the equation to calculate the curvature of any nondegenerate conic section at a known point.

Finding the Equation of a Conic Section

The general equation of a conic section is

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \qquad \text{(Eq 3.1)}$$

Given five points, no three being collinear, one can substitute each point into the general equation resulting in the following system of equations,

$$x_1^2A + x_1y_1B + y_1^2C + x_1D + y_1E + F = 0$$
$$x_2^2A + x_2y_2B + y_2^2C + x_2D + y_2E + F = 0$$
$$x_3^2A + x_3y_3B + y_3^2C + x_3D + y_3E + F = 0$$
$$x_4^2A + x_4y_4B + y_4^2C + x_4D + y_4E + F = 0$$
$$x_5^2A + x_5y_5B + y_5^2C + x_5D + y_5E + F = 0$$

Thus, there are five equations with six unknowns so one of the unknowns must be assigned some arbitrary value. Generally the value of one is used for easier computations.

Provided that the graph of the conic section does not pass through the origin the constant term F may be assigned an arbitrary value. However, if the conic section passes through the origin the assignment of an arbitrary value may be given to a parameter other than that of the constant term or the constant term may be assigned the value of zero. If the conic does not lie on the origin, the following system of equations can be used to find the other parameters,

$$x_1^2A + x_1y_1B + y_1^2C + x_1D + y_1E = -1$$
$$x_2^2A + x_2y_2B + y_2^2C + x_2D + y_2E = -1$$
$$x_3^2A + x_3y_3B + y_3^2C + x_3D + y_3E = -1$$
$$x_4^2A + x_4y_4B + y_4^2C + x_4D + y_4E = -1$$
$$x_5^2A + x_5y_5B + y_5^2C + x_5D + y_5E = -1.$$

The associated matrix,

$$\begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & -1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & -1 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 & -1 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 & -1 \\ x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 & -1 \end{bmatrix}$$

when reduced to a row echelon matrix will give the values for the other parameters. An example of how to find the equation of a conic section given five points, such that no three are collinear is as follows. Given five points, say (-1, -1), (-1, 1), (2, -1), (2, 0), and (0, 2), the following system of equations may be obtained:

$$\begin{aligned} A + B + C + D + E + F &= 0 \\ A - B + C - D + E + F &= 0 \\ 4A - 2B + C + 2D - E + F &= 0 \\ 4A + 2D + F &= 0 \\ 4C + 2E + F &= 0 \end{aligned}$$

Assigning F = 1 the associated matrix is:

$$\begin{bmatrix} 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 4 & -2 & 1 & 2 & -1 & -1 \\ 4 & 0 & 0 & 2 & 0 & -1 \\ 0 & 0 & 4 & 0 & 2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -\frac{6}{14} \\ 0 & 1 & 0 & 0 & 0 & -\frac{1}{14} \\ 0 & 0 & 1 & 0 & 0 & -\frac{3}{14} \\ 0 & 0 & 0 & 1 & 0 & \frac{5}{14} \\ 0 & 0 & 0 & 0 & 1 & -\frac{1}{14} \end{bmatrix}$$

From the row echelon matrix it can be seen that when F = 1, $A = -\frac{6}{14}$, $B = -\frac{1}{14}$, $C = -\frac{3}{14}$, $D = \frac{5}{14}$, and $E = -\frac{1}{14}$. Thus, the equation of the conic section is:

$$-\frac{6}{14}x^2 - \frac{1}{14}xy - \frac{3}{14}y^2 + \frac{5}{14}x - \frac{1}{14}y + 1 = 0$$

or

$$6x^2 + xy + 3y^2 - 5x + y - 14 = 0$$

Curvature

The curvature of a graph can be intuitively considered as the rate at which the curve bends as one travels along the curve. At each point on the curve, the numeric value of curvature, can be calculated. The more the curve bends at a point the greater the numeric value for curvature. For instance, a straight line will have a curvature of zero for each point on that line, since it does not bend at all. However, a circle has a uniform rate at which the curve bends and thus each point on the circle will have the same curvature . The larger the circle, the less the circle bends at each point. Consequently, the curvature of a larger circle is less than that of a smaller circle.

Given an equation of a curve described in terms of rectangular coordinates and that equation being twice differentiable, the formula for curvature at a point on the curve is given by:

$$K = \frac{|y''|}{[1 + (y')^2]^{3/2}} \qquad \text{(Eq 3.2)}$$

Theorem The curvature at any point on a circle of radius r is 1/r.

Proof

There is no loss in generality in assuming the circle has a center at the origin, so the equation $x^2 + y^2 = r^2$ will be used. Implicitly differentiating this equation results in the following equation,

$$2x^2 + 2yy' = 0 \Rightarrow y' = -x/y$$

Implicitly differentiating $y' = -x/y$ yields:

$$y'' = -\frac{(y - xy')}{y^2} = -\frac{(y - x(-x/y))}{y^2}$$

$$= -\frac{y^2 + x^2}{y^3}$$

$$= -\frac{r^2}{y^3}$$

Substituting into Eq 3.2

$$K = \frac{\frac{-r^2}{y^3}}{[1 + (-x/y)^2]^{3/2}} = \frac{r^2 (y^2)^{3/2}}{|y^3| [y^2 + x^2]^{3/2}} = \frac{r^2}{(r^2)^{3/2}} = \frac{1}{r}$$

This leads one to talk about the radius of curvature, ρ. The radius of curvature at a point P on a curve is the radius of a circle that "fits" the curve there better than any other circle (7,733). The radius of curvature can be calculated by the following equation,

$\rho = 1/K$

provided $K \neq 0$.

Derivation of the Formulas to Calculate K and ρ

To derive the formulas for K and one must first implicitly differentiate Eq 3.1, which yields,

$$2Ax + By + Bxy' + 2Cyy' + D + Ey' = 0$$

Thus, $y' = -\dfrac{2Ax + By + D}{Bx + 2Cy + E}$ Eq (3.3)

Implicitly differentiating Eq 3.3 one obtains,

$$y'' = -\frac{(2A + By')(Bx + 2Cy + E) - (2Ax + By + D)(B + 2Cy')}{(Bx + 2Cy + E)^2}$$

Substituting y' with $-\dfrac{2Ax + By + D}{Bx + 2Cy + E}$

$$y'' = \frac{-[(2A - B\left(\frac{2Ax + By + D}{Bx + 2Cy + E}\right))(Bx + 2Cy + E) - (2Ax + By + D)(B - 2C\left(\frac{2Ax + By + D}{Bx + 2Cy + E}\right))]}{(Bx + 2Cy + E)^2}$$

which simplifies to:

$$y'' = -\frac{2A(Bx + 2Cy + E)^2 - 2B(2Ax + By + D)(Bx + 2Cy + E) + 2C(2Ax + By + D)^2}{(Bx + 2Cy + E)^3}$$

Substituting y' and y'' into Eq 3.1 results in finding a general equation for K, that is,

$$K = \frac{\left|\dfrac{2A(Bx + 2Cy + E)^2 - 2B(2Ax + By + D)(Bx + 2Cy + E) + 2C(2Ax + By + D)^2}{(Bx + 2Cy + E)^3}\right|}{\left[1 + \dfrac{(2Ax + By + D)^2}{(Bx + 2Cy + E)^2}\right]^{3/2}}$$

Thus, $K = \dfrac{|2A(Bx + 2Cy + E)^2 - 2B(2Ax + By + D)(Bx + 2Cy + E) + 2C(2Ax + By + D)^2|}{[(Bx + 2Cy + E)^2 + (2Ax + By + D)^2]^{3/2}}$

(Eq. 3.4)

Hence, the radius of curvature is given by

$$\rho = \frac{[(Bx + 2Cy + E)^2 + (2Ax + By + D)^2]^{3/2}}{|2A(Bx + 2Cy + E)^2 - 2B(2Ax + By + D)(Bx + 2Cy + E) + 2C(2Ax + By + D)^2|}$$

(Eq. 3.5)

Using $6x^2 + xy + 3y^2 - 5x + y - 14 = 0$, the equation from the previous example, K and ρ can be found at the point (2, 0). From Eq 3.4

$$K = \frac{|2(6)(1(2) + 2(3)(0) + 1)^2 - 2(1)(2(6)(2) + (1)(0) - 5)(1(2) + 2(3)(0)) + 2(3)(2(6)(2) + 1(0) - 5)^2|}{[(1(2) + 2(3)(0) + 1)^2 + (2(6)(2) + 1(0) - 5)^2]^{3/2}}$$

$= .3035$

Thus, $\rho = 1/K = 3.295$. However, the method of determing R in this chapter will only involve ρ, which means that

Eq 3.5 can be used to calculate ρ.

In the next two sections of this chapter, the writer will discuss various approaches for estimating R using ρ.

Estimating R

Taking combinations of five points, no three being collinear, one can calculate the equation of a conic section. Thus, after obtaining the equation, calculations of ρ at each of the five points is possible. Given n points one can calculate at most ${}_nC_5$ equations of conic sections. Hence, one has at most $5({}_nC_5)$ values for ρ. One approach is to calculate the mean of the ρ_{ij}'s, that is, $\hat{R} = \sum_{i=1}^{r}\sum_{j=1}^{5}\rho_{ij}/5r$, where $r \leq {}_nC_5$ (r is the number of equations obtained by combinations of 5 points with the conditions imposed).

Example

Assume n = 6 and the following points are observed A(-4.29708, 9), B(-10.1553, 6), C(-10.0884, 3), D(-10.3031, -3), E(-9.10405, -6), and F(-4.24907, -9).

Combination of points	Eq.	ρ_{iA}	ρ_{iB}	ρ_{iC}	ρ_{iD}	ρ_{iE}	ρ_{iF}
ABCDE	1)	12.595	393.35	154.12	14.533	111.34	----
ABCDF	2)	141.37	86.815	76.582	102.56	----	148.70
ABCEF	3)	57.054	68.340	73.799	----	43.902	34.061
ABDEF	4)	4.7668	12.683	----	15.609	10.577	3.1648
ACDEF	5)	4.9348	----	10.891	11.034	8.2651	4.8906
BCDEF	6)	----	1013.4	364.55	2.0398	2.9421	291.54

1) $.0127x^2 - .00472xy - .000215y^2 + .227x - .0470y + 1 = 0$

2) $.0207x^2 - .00063xy - .000729y^2 + .309x - .0023y + 1 = 0$

3) $.0279x^2 + .00269xy + .001270y^2 + .377x + .0119y + 1 = 0$

4) $.1970x^2 + .02590xy + .060500y^2 + 2.22x + .1120y + 1 = 0$

5) $-.0348x^2 + .00265xy - .018900y^2 - .274x + .0114y + 1 = 0$

6) $.0064x^2 + .00841xy + .000243y^2 + .164x + .0846y + 1 = 0$

Therefore, $\hat{R} = \sum_{i=1}^{6}\sum_{j=1}^{5} \rho_{ij}/30 = 109.346$.

Finding R by Restricting the Selection of Points.

In the previous approach there were no restrictions on the combinations of five points selected from the n observed points except that no three points could be collinear. However, in the following method for determining R, some conditions are needed. First, as before, no three points can be collinear. Secondly, assuming $Y_1 > Y_2 > Y_3 > Y_4 > Y_5$, then $X_1 > X_2 > X_3$ and $X_5 > X_4 > X_3$. No other combination of five points will not be considered because of the known orientation of the circle. $\hat{R} = \sum_{i=1}^{r}\sum_{j=1}^{5} \rho_{ij}/5r$ for $r \leq {}_nC_5$.

Example

Assume n = 6 and A, B, C, D, E, and F are the points used in the previous example. The following table provides the combinations of points that satisfied the given conditions, and the ρ_{ij} that are associated with such combinations.

Combination of points	Eq.	ρ_{iA}	ρ_{iB}	ρ_{iC}	ρ_{iD}	ρ_{iE}	ρ_{iF}
ABDEF	1)	4.7668	12.683	----	15.609	10.577	3.1648
ACDEF	2)	4.9348	----	10.891	11.034	8.2651	4.8906

1) $.1970x^2 + .02590xy + .060500y^2 + 2.22x + .1120y + 1 = 0$

2) $-.0348x^2 + .00265xy - .018900y^2 - .274x + .0114y + 1 = 0$

Thus $\hat{R} = \sum_{i=1}^{2} \sum_{j=1}^{5} \rho_{ij} /10 = 8.68152.$

As can be seen from the examples, the two approaches can lead to two estimates of R that are drastically different. One may ask why the second approach might be used instead of the first approach. The reason for using the second approach is that, in general, the investigator knows the overall shape of the curve from which the n points are observed, and when the conditions imposed in second example appear to be reasonable assumptions. However, if nothing is known about the curve a further investigation would be needed to determine which method, if either, is appropriate.

In both methods there appears to be many disadvantages. One such disadvantage is the enormous amount of effort that is needed to compute R by hand. On the other hand, with the use of computers, the computations are somewhat effortless after the initial writing of the computer program. After looking at the examples more closely what appears to happen at many of the points is that the conic section obtained by these points has a small curvature, thus resulting in a large radius of curvature. What one would wish is that the averaging of the ρ_{ij} would minimize this problem. However, the problem does not appear to be corrected by employing this averaging technique.

From a computational standpoint, the second approach generally requires fewer computations, hence making it a more attractive approach. Through computer simulations, questions such as the feasibility of using the radius of curvature to estimate R and which method is more appropriate to use, can be answered. These questions will be considered in some detail in Chapter 5.

CHAPTER 4

THE SQUARED RESIDUALS

In statistics, one method of estimation is based on the minimization of the sum of the squared residuals. For example, in regression analysis one is concerned with the minimization of the squared residual. Where the residual is the difference in the observed value of the dependent variable and the predicted value of the dependent variable. In this chapter, the author will be looking at an approach that minimizes the sum of the squared residuals. The concept of squared residuals will be defined later.

Linssen and Banens (4,307) suggest a least squares estimator $\hat{R}$ for R assuming that the center of the circle is at the origin. Furthermore, they assume that there is measurement error in both the x and y directions. The Least Square estimator for R is $\hat{R} = \sum_{i=1}^{n} R_i / n$ with $R_i = (x_i^2 + y_i^2)^{\frac{1}{2}}$, where (x_i, y_i) is an observed point. However, the original problem defined at the beginning of this paper, made no assumptions of knowing the center of the circle, nor where any made relating to a measurement error in the y direction. In this chapter, on the other hand, the writer will suggest a way to calculate at most ${}_nC_3$ different centers from which one could then calculate an estimate of R from each center. Nevertheless, one will only be concerned with the estimate that minimizes the sum of the squared residuals. After further development, the residual will be defined.

Finding the Different Centers

In the previous chapter the discussion dealt with calculating at most ${}_nC_3$ different circles from n points. Since each circle has a center, there are at most ${}_nC_3$ different centers associated with these circles. This collection of centers may be used to calculate estimates of R. Using all combinations of three noncollinear points and Eq. 2.3 and Eq. 2.4, one can calculate this set of centers. Throughout the remainder of this chapter the notation (h_j, k_j); $j \leq {}_nC_3$ has reference to the jth center of the collection of centers previously defined.

Calculating an Estimate of R

Having found a center (h_j, k_j) one could use a method similar to Linssen and Banens (4,307) to calculate an estimator R* for R. That is, using the jth center, $R^*_j = \sum_{i=1}^{n} R_i/n$ with $R_i = [(x_i - h_j)^2 + (y_i - k_j)^2]^{\frac{1}{2}}$.

One would thus have a collection of R*, each being associated with one of the centers from the set of centers discussed in the previous section. Therefore, there is at most ${}_nC_3$ different R*'s.

How to Estimate R

A question one might ask is how can this collection of R*'s be used to find a "good" estimate of R. One approach may be to calculate the mean of the R*'s to provide an estimate of R. However, the author considered a second approach which minimizes the sum of the squared residuals.

From this collection of R*'s one is only concerned with the R^*_j that minimizes $\sum_{i=1}^{n}(R_i - R^*_j)^2$: $R^*_j = \sum_{i=1}^{n} R_i$ /n and $R = [(x_i - h_j)^2 + (y_i - k_j)^2]^{\frac{1}{2}}$. The R* that minimizes this sum of the squared residuals is itself the estimate of R; that is, $\hat{R} = R^*$.

Example

Assume n = 4 and the following points are observed A(-3.98, 9), B(-10.04, 3), C(-9.25, -3) and D(-3.34, -9). Using ABC as a combination of three points the center $(h_1, k_1) = (-2.05592, .999018)$, $R^*_1 = \sum_{i=1}^{4}[(x - h_1)^2 + (y_i - k_1)^2] /4 = 8.69221$, and $\sum_{i=1}^{4}(R - R^*_1)^2 = 2.56983$. For the combination ABD the center $(h_2, k_2) = (-1.15532, 0.0898925)$, $R^*_2 = 9.17641$, and $\sum_{i=1}^{4}(R_i - R^*_2)^2 = .35165$. Likewise, ACD yields $(h_3, k_3) = (.0731166, .128759)$, $R^*_3 = 9.8632$, and $\sum_{i=1}^{n}(R_i - R^*_3)^2 = .343348$. For BCD the center $(h_4, k_4) = (1.2518, 1.43454)$, $R^*_4 = 10.84$, and $\sum_{i=1}^{4}(R_i - R^*_4)^2 = 3.62693$. Thus, $\hat{R} = R^*_3 = 9.8632$.

As with most of the approaches mentioned in this paper the computations by hand are quite extensive. However, the use of a computer would aid in the tedious computations, especially for larger values of n. The performance of this method using computer simulations will be discussed in Chapter 5, along with comparing this method to other methods previously mentioned.

CHAPTER 5

SIMULATIONS

After building a model, the question becomes how do the estimators perform. The ideal approach to answering this question is with sound mathematical proofs, however, many times the mathematics is very difficult or impossible.

Simulations as defined by Graybeal and Pooch (3,1) is "the process of designing a computerized model of a system (or process) and conducting experiments with this model for the purpose either of understanding the behavior of the system or of evaluating various strategies for the operation of a system." The traditional model-building approach to problem solving linked with simulation is a very important tool. This chapter will deal with checking the performance of the estimators previously discussed.

Advantages and Disadvantages

As with any type of modeling, simulations have distinct advantages and disadvantages. Graybeal and Pooch (3,10) list the following advantages of simulations.

1.) It permits controlled experimentation. A simulation experiment can be run a number of times with varying input parameters to test the behavior of the system under a variety of situations and conditions.

2.) It permits time compression. Operation of the situations and conditions over extended periods of time can be simulated in only minutes with ultrafast computers.

3.) It permits sensitivity analysis by manipulation of input variables.

They also list the following disadvantages to simulations.

1.) Extensive development time may be encountered. Most simulation models are quite large and, like any large programming projects, take time.

2.) Simulations may become expensive in terms of computer time.

3.) Hidden critical assumptions may cause the model to diverge from reality.

Simulations have been proven to be an effective approach to problem solving, but the investigator must be aware of the disadvantages.

Properties of Estimators

There are a number of desirable properties that any estimate of a population parameter should have. The following is a list of such properties.

1.) Unbiasedness - An estimate $\hat{\theta}$ of a parameter θ is said to be unbiased provided $E(\hat{\theta}) = \theta$ (6,110).

2.) Minimum variance - An estimate $\hat{\theta}$ of some parameter θ is said to be a minimum variance estimate provided $\sigma^2_{\hat{\theta}} \leq \sigma^2_{\theta^*}$ for any other estimate θ^* (6,115).

3.) Consistency - As the sample size increases, if the estimate $\hat{\theta}$ approaches the value of θ, then the estimate $\hat{\theta}$ is said to be consistent (6,110).

This list is only some of the properties of estimators. However, these will be used in this thesis to evaluate the performance of the approaches previously discussed.

Simulation Language

The SAS (Stastical Analysis System) was chosen for writing the simultations due to the flexibility and simplicity of PROC MATRIX. A representative listing of the simulation programs are in APPENDIX A. By repeatedly changing the parameters in the model, an enormous amount of output is generated. For the different simulations, this output is in tabular form and located in APPENDIX B. The following section will use partial tables from APPENDIX B to provide evidence that some of the metods should not be used to estimate R.

Evaluation of Methods Used to Estimate R.

In the simulations a radius of ten was chosen. However, if the variance in the error term was three, equivalent result could be obtained having a radius of one hundred with the variance in the error term being three hundred. Throughout the evaluation of the individual methods, the reader needs to keep in mind that the true value for R is 10.

TABLE 1

SIMULATION RESULTS USING THE RADIUS OF CURVATURE TO ESTIMATE R USING ALL FEASIBLE COMBINATIONS OF FIVE POINTS.

n = Number of unique Y's	σ^2	$\bar{\hat{R}}$	VAR($\hat{R}$)	Number of Simulations	Values of Y
7	3	951.583		100	*
7	1	1016.32		100	*
7	.5	1043.87		100	*
7	.1	75.9853		100	*
5	1	200.845		100	+
5	.5	36.1984		100	+
5	.1	11.9531		100	+

* y= -9,-6,-3,0,3,6,9 + y= -9,-4.5,0,4.5,9

TABLE 2

SIMULATION RESULTS USING THE RADIUS OF CURVATURE TO ESTIMATE R USING THE IMPOSED CONDITIONS.

n = Number of unique Y's	σ^2	$\bar{\hat{R}}$	VAR($\hat{R}$)	Number of Simulations	Values of Y
7	3	422.305		100	*
7	1	994.002		100	*
7	.5	565.84		100	*
7	.1	84.816		100	*
5	3	118.68		100	+
5	1	155.002		100	+
5	.5	31.786		100	+
5	.1	11.9531		100	+

* y = -9, -6, -3, 0, 3, 6, 9 + y = -9, -4.5, 0, 4.5, 9

The results of the simulations involving curvature are listed in detail in TABLE 1 and TABLE 2. Looking at these tables, it can be seen that the radius of curvature is a biased estimator of R. Furthermore, the methods using the radius of curvature do not improve as the number of unique y's increase. Thus, there is evidence to suggest a further study of using the radius of curvature is needed. At the present time, it appears that using the radius of curvature to estimate R is inappropriate.

TABLE 3

SIMULATION RESULTS USING THE RADIUS OF ALL POSSIBLE CIRCLES.

n = Number of Unique Y's	σ^2	$\bar{\hat{R}}$	VAR($\hat{R}$)	Number of Simulations	Values of Y
7	3	26.5385		100	*
7	1	27.8108		100	*
7	.5	32.0948		100	*
7	.1	13.4527		100	*
4	2	22.3971	3493.13	1000	**
4	.5	11.9949	384.459	1000	**
3	2	10.9975	8.88369	1000	***
3	1	10.4686		100	***
3	.5	10.2015	.623521	1000	***
3	.1	10.0281		100	***

* y = -9, -6, -3, 0, 3, 6, 9

** y = -9, -3, 3, 9

*** y = -9, 0, 9

A partial listing of the simulations results using the mean of radii of all possible circles to estimate R is contained in TABLE 3. As can be seen from TABLE 3, R is also a biased estimator of R. Nevertheless, if n = 3 and the spread in y's is equally spaced about zero and close to the end points of the interval [-10, 10] a reasonable estimate of R can be achieved provided the variance in the error terms is small. A somewhat surprising result is that as the number of distinct y's increase the bias of the estimate is not reduced.

TABLE 4

SIMULATION RESULTS USING THE RADIUS OF CIRCLES WHERE $Y_1 > Y_2 > Y_3$ AND $X_1 > X_2$ AND $X_3 > X_2$.

n = Number of Unique Y's	σ^2	$\bar{\hat{R}}$	VAR($\hat{R}$)	Number of Simulations	Values of Y
9	3	8.33541		1000	*
9	2	8.69646	4.2556	1000	*
9	1	9.30024		1000	*
9	.5	9.59886	2.95244	1000	*
9	.1	9.94725		1000	*
7	3	9.30576		1000	**
7	2	9.40000	3.6845	1000	**
7	1	9.60765		1000	**
7	.5	9.90503	2.960903	1000	**
7	.1	9.99178		1000	**
7	.5	4.70076	.964667	1000	+*
7	.5	4.08815	3.34535	1000	+**

TABLE 4 CONTINUED

SIMULATION RESULTS USING THE RADIUS OF CIRCLES WHERE $Y_1 > Y_2 > Y_3$ AND $X_1 > X_2$ AND $X_3 > X_2$.

n = Number of Unique Y's	σ^2	$\bar{\hat{R}}$	VAR($\hat{R}$)	Number of Simulations	Value of Y
7	.5	10.485	1.0197	1000	++
7	.5	8.02883	14.003	1000	++*
7	.5	9.18329	4.00019	1000	++**
5	2	9.75276	3.59142	1000	***
5	1	9.77383		1000	***
5	.5	9.89979	2.727	1000	***
5	.1	10.2214		1000	***
4	3	10.5384		1000	****
4	2	10.2625	12.8887	1000	****
4	1	9.82344		1000	****
4	.5	9.69888	.464514	1000	****
4	.1	9.77921		1000	****

* y = -8, -6, -4, -2, 0, 2, 4, 6, 8

** y = -9, -6, -3, 0, 3, 6, 9

+* y = 0, 1.75, 3.25, 4.75, 6.9, 8.5, 9.75

+** y = -3, -1.25, -.75, 0, 1, 2.75, 3.5

++ y = -9.99, -8.25, -6.95, 0, 7, 8.75, 9.5

++* y = -5.99, -2.25, -1.75, 0, 1, 3.75, 6.5

++** y = -9.99, -4.25, -2.75, 0, 2, 5.75, 8.5

*** y = -9, -4.5, 0, 4.5, 9

**** y = -9, -3, 3, 9

TABLE 4 provides a representative listing of the simulation results involving the mean of the radii of circles with conditions imposed on the selection of combinations of points to be used in calculations. It can be seen that when σ^2 is small, good estimates of R are achieved when the spread in y's are equally spaced about zero and covers the interval [-10, 10]. However, in general, this method appears to be biased, as can be seen from TABLE 4. When σ^2 is large, fair estimates of R are achieved when n = 7 and the y's are equally spaced about zero covering the interval [-10, 10]. Thus, there is evidence that the estimator may be biased by the number of distinct y's, or the spread in y's, or other possible factors that are not detectible in these simulations.

TABLE 5

SIMULATION RESULTS USING THE RESIDUAL APPROACH.

n = Number of unique Y's	σ^2	$\bar{\hat{R}}$	VAR($\hat{R}$)	Number of Simulations	Values of Y
9	2	9.83033	6.31518	100	*
9	.5	9.88901	.937487	100	*
7	3	10.3748		100	**
7	3	10.3429		1000	**
7	2	10.132	3.51527	100	**
7	1	10.205		100	**
7	1	10.045		100	**
7	.5	10.039	.419528	100	**

TABLE 5 CONTINUED

SIMULATION RESULTS USING THE RESIDUAL APPROACH.

n = Number of unique Y's	σ^2	$\bar{\hat{R}}$	VAR($\hat{R}$)	Number of Simulations	Values of Y
7	.5	10.0044		1000	**
7	.1	9.99301		1000	**
5	2	10.8199	8.49873	100	***
5	.5	10.2119	.513119	100	***
4	2	10.183	6.10877	100	****
4	.5	9.8531	.674463	100	****

* y = -8,-6,-4,-2,0,2,4,6,8 ** y = -9,-6,-3,0,3,6,9

*** y = -9, -4.5, 0, 4.5, 9 **** y = -9, -3, 3, 9

The results of the simulations involving the minimization of the sum of the squared residuals are listed in TABLE 5. As can be seen from this table, the residual approach provides a reasonable estimate for R. Looking at the VAR($\hat{R}$) = $\sum_{i=1}^{n} \frac{(R_i - \bar{\hat{R}})^2}{n}$ where n = 7 produces the smallest VAR($\hat{R}$) with values of $\bar{\hat{R}}$ very close to the true value of R. These facts suggest that the method creates a biased estimate of R. That is, the number of distinct y's appear to bias the estimator. Like the other methods, the residual approach yields an estimator that does not appear to improve as the number of distinct y's increases.

A biased estimator is many times preferable over an unbiased estimator. If the unbiased estimator has a large variance and the biased estimator has a very small variance, the biased estimator may be prefered over the unbiased estimator provided a function can be found to calculate the bias. Considering the residual approach it would be nice if the researcher could find a function to calculate the biased in terms of the number of distinct y's. The researcher needs to keep in mind that a bias estimator is not always undesirable.

Comparing Estimators

In this section a comparison between the residual approach and the method involving the mean of the radii of circles with conditions imposed on the selection of combinations of points to be used in the calculation will be discussed. As previously stated, the other three methods do not appear to produce reasonable estimates of R. First, a discussion on the comparison of two estimators seem to be in order. Bratley, Fox, and Schrage (1,27) state that

> The observations from which the sample variance is calculated are often correlated. In such cases, the usual variance estimators are biased estimators of the mean squared error, it is not always the case that the estimator with the smallest theoretical mean squared error will give the smallest sample variance in any particular situation. In general, however, when no better criterion is available we advocate comparison of candidate estimators using their observed sample variances.

Since the $VAR(\hat{R})$ is the only such criterion that has been calculated for the comparison of models it will be used to select a model.

From TABLE 4 and TABLE 5 it can be seen that when the number of distinct y's was 7, both estimators provide a good estimate for R. In general, when the variance in the error term was small, the $VAR(\hat{R})$ was much smaller using the residual approach, thus making it the better method. However, when the variance in the error term was large the $VAR(\hat{R})$ for the two estimators were comparable for n = 7 or 9. In this case the value of $\bar{\hat{R}}$ was close to R using the residual approach. Thus, these simulations suggest that the residual approach provides a better method for estimating R.

CHAPTER 6

CONCLUSION

In this paper the writer has examined a variety of methods to estimate the radius of a circle when observing n points that lie on a semi-circle but are measured with error. Computer simulations were used to check the performance of each method. The methods involving the radius of curvature were found to be inappropriate in estimating R. Furthermore, using the mean of the radii of all possible circles appeared to produce a biased estimator. However, two methods, one using the mean of the radii of all possible circles where conditions were imposed on the selection of points to be used in calculations, and the other using a residual approach, provided good estimates for R when the variance in the error term was small. Often when the variance in the error term was large and the number of unique y's were small, no calculations were carried out because of the conditions imposed. Overall, the evidence provided by the simulations suggested that the method involving the minimization of the sum of the squared residuals yielded the best estimate of those methods mentioned in this paper. There is one major drawback to all of the methods presented. Without the use of a computer, the calculations needed to carry out the estimation of R were extremely long and tedious. However, with the use of a computer the calculations became trival after the initial writing of the computer program.

Further Studies

As the researcher seeks answers to the problems that confronts him, many questions arise that are closely related to the problem but are not within the scope of this study. Looking at the residual approach, one such question may be, what is the actual minimum value of the sum of the squared residuals? Using the centers of all possible circles, can one calculate another point to be used as the center that will reduce the sum of the squared residuals? Furthermore, how does the centers of the centers behave? If the number of distinct y's is equal to four, then there are at most four different centers that can be computed from combinations of three points. Doing this in an iterative manner the collection of centers can be seen as a sequence. The question that follows is, in what cases does this sequence converge? If the sequence converges to a point, does this point when used as the center produce the minimum valve for the sum of the squared residuals? Another question that arises is, can a function of the bias be found in terms of the factors that bias the estimate (e.g. the number of distinct y's)?

The above questions deal with possible improvements of the methods presented in this paper. The next step may be to backup and look at the experimental design. Before starting an experiment, the main question is, what is the optimal design? Should one take a single observation of x for each y, but make several replications, and then average the results of the replications before estimating the radius? What number of distinct y's is the best? Is there a need for a large number of distinct y's with each having a single observation for x or a need for a small number of unique y's with each having multiple observations of x? Once the design is chosen, then one can investigate whether the estimator is consistent. In other words, does the value of the estimate approach the true value of R as the sample size increase? These questions pose only a handful of possible avenues for further study.

Sherlock Holmes, a detective with great problem solving skills, in A Study in Scarlet makes the following statement which summarizes a problem solving stratagy in both mathematics and statistics.

> In solving a problem of this sort, the grand thing is to be able to reason backward. That is a very useful accomplishment, and a very easy one, but people do not practice it much....Most people, if you describe a train of events to them, will tell you what the result would be....There are few people, however, who, if you told them a result, would be able to evolve from their own inner consciousness what the steps were which led up to that result. This power is what I mean when I talk of reasoning backward.

BIBLIOGRAPHY

[1] Bartley, Paul; Fox, Bennent L.; and Schrage, Linus E. A Guide to Simulation. New York : Springer-Verlag New York Inc., 1983.

[2] Fraleigh, John. Calculus with Analtic Geometry. 2nd ed. Reading: Addison-Wesley Publishing Company, Inc., 1985.

[3] Graybill, Wanye T., and Pooch, Udo W. Simulation: Principles and Methods. Boston: Little, Brown and Company, Inc., 1980.

[4] Linssen, H. N., and Banens, P. J. A. "Estimation of the Radius of a Circle when the Coordinates of a Number of Points Lie on its Circumference are Observed : An Example of Bootstrapping," Statistics and Probability Letters. 6 (1983); 307-11.

[5] Simmons, George. Calculus and Analytic Geometry. New York: McGraw-Hill Book Company, 1985.

[6] Tucker, Howard G. An Intoduction to Probability and Mathematical Statistics. New York: Academic Press, 1962.

[7] Zill, Dennis. Calculus and Analytic Geometry. Boston: Prindle, Weber, & Schmidt, 1984.

APPENDIX A

COMPUTER SIMULATION USING THE RADIUS OF CURVATURE TO ESTIMATE R USING ALL FEASIBLE COMBINATIONS OF FIVE POINTS.

```
ITLE THESIS PROJECT MARVIN HARRELL CURVATURE ALL POSSIBLE CHOICES;
ATA ONE;
=10;
O K=1 TO 100;
=12;
O J=1 TO 7;
=Y-3;
ERROR=SQRT(.1)*RANNOR(1613218064);
= -SQRT(R**2-(Y)**2)+XERROR;
=1;
SQ=X**2;
SQ=Y**2;
Y=X*Y;
UTPUT;
ND;
ND;
ATA NOW;
ET ONE;
ROP X Y  XERROR XSQ YSQ XY R F J K;
ISQ=XSQ;
IYI=XY;
ISQ=YSQ;
I=X;
I=Y;
ATA C;
NPUT NEGF;
ARDS;
1
1
1
1
1
ROC MATRIX;
ETCH SIM DATA=NOW;
ETCH CONST DATA=C;
=0;
OTR=0;
O TLOOP =1 TO 700 BY 7;
1=TLOOP;
2=TLOOP+1;
3=TLOOP+2;
4=TLOOP+3;
5=TLOOP+4;
6=TLOOP+5;
7=TLOOP+6;
=SIM(P1,)//SIM(P2,)//SIM(P3,)//SIM(P4,)//SIM(P5,)//SIM(P6,)//SIM(P7,
OUNT=0;
TOT=0;
O I=1 TO (NROW(T)-4);
O J=I+1 TO (NROW(T)-3);
O K=J+1 TO (NROW(T)-2);
O L=K+1 TO (NROW(T)-1);
O M=L+1 TO NROW(T);
=T(I,)//T(J,)//T(K,)//T(L,)//T(M,);
=SOLVE(A,CONST);
1=X(1,1);B=X(2,1);C=X(3,1);D=X(4,1);E=X(5,1);
1=(B*A(1,4)+2*C*A(1,5)+E)**2;D2=(2*A1*A(1,4)+B*A(1,5)+D)**2;
I=ABS(2*A1*D1-2*B*D1**.5*D2**.5+2*C*D2);
I=(D1+D2)**1.5;
I=TI*/NI;
```

COMPUTER SIMULATION USING THE RADIUS OF CURVATURE TO ESTIMATE R' USING ALL FEASIBLE COMBINATIONS OF FIVE POINTS (CONTINUED).

```
RTOT=RTOT+RI;
COUNT=COUNT+1;
D1=(B*A(2,4)+2*C*A(2,5)+E)**2;D2=(2*A1*A(2,4)+B*A(2,5)+D)**2;
NJ=ABS(2*A1*D1-2*B*D1**.5*D2**.5+2*C*D2);
TJ=(D1+D2)**1.5;
RJ=TJ*/NJ;
RTOT=RTOT+RJ;
COUNT=COUNT+1;
D1=(B*A(3,4)+2*C*A(3,5)+E)**2;D2=(2*A1*A(3,4)+B*A(3,5)+D)**2;
NK=ABS(2*A1*D1-2*B*D1**.5*D2**.5+2*C*D2);
TK=(D1+D2)**1.5;
RK=TK*/NK;
RTOT=RTOT+RK;
COUNT=COUNT+1;
D1=(B*A(4,4)+2*C*A(4,5)+E)**2;D2=(2*A1*A(4,4)+B*A(4,5)+D)**2;
NL=ABS(2*A1*D1-2*B*D1**.5*D2**.5+2*C*D2);
TL=(D1+D2)**1.5;
RL=TL*/NL;
RTOT=RTOT+RL;
COUNT=COUNT+1;
D1=(B*A(5,4)+2*C*A(5,5)+E)**2;D2=(2*A1*A(5,4)+B*A(5,5)+D)**2;
NM=ABS(2*A1*D1-2*B*D1**.5*D2**.5+2*C*D2);
TM=(D1+D2)**1.5;
RM=TM*/NM;
RTOT=RTOT+RM;
COUNT=COUNT+1;
END;
END;
END;
END;
END;
R=RTOT*/COUNT;
IF TLOOP=1 THEN ALR=R;
            ELSE ALR=ALR||R;
TOTR=TOTR+R;
END;
SIMR=TOTR*/100;
PRINT SIMR ALR;
```

COMPUTER SIMULATION USING THE RADIUS OF CURVATURE TO ESTIMATE R USING THE IMPOSED CONDITIONS.

```
DATA ONE;
TITLE THESIS PROJECT MARVIN HARRELL CURVATURE X1>X2>X3 & X5>X4>X3;
R=10;
DO K=1 TO 100;
Y=12;
DO J=1 TO 7;
Y=Y-3;
XERROR=SQRT(.1)*RANNOR(161321806 4);
X= -SQRT(R**2-(Y)**2)+XERROR;
F=1;
XSQ=X**2;
YSQ=Y**2;
XY=X*Y;
OUTPUT;
END;
END;
DATA NOW;
SET ONE;
DROP X Y XERROR XSQ YSQ XY R F J K;
XISQ=XSQ;
XIYI=XY;
YISQ=YSQ;
XI=X;
YI=Y;
DATA C;
INPUT NEGF;
CARDS;
-1
-1
-1
-1
-1
PROC MATRIX;
FETCH SIM DATA=NOW;
FETCH CONST DATA=C;
R=0;
TOTR=0;
DO TLOOP =1 TO 700 BY 7;
P1=TLOOP;
P2=TLOOP+1;
P3=TLOOP+2;
P4=TLOOP+3;
P5=TLOOP+4;
P6=TLOOP+5;
P7=TLOOP+6;
T=SIM(P1,)//SIM(P2,)//SIM(P3,)//SIM(P4,)//SIM(P5,)//SIM(P6,)//SIM(P7,
COUNT=0;
RTOT=0;
DO I=1 TO (NROW(T)-4);
DO J=I+1 TO (NROW(T)-3);
DO K=J+1 TO (NROW(T)-2);
DO L=K+1 TO (NROW(T)-1);
DO M=L+1 TO NROW(T);
A=T(I,)//T(J,)//T(K,)//T(L,)//T(M,);
IF (A(1,4)>A(2,4)&A(2,4)>A(3,4) & A(5,4)>A(4,4)&A(4,4)>A(3,4)) THEN D
X=SOLVE(A,CONST);
A1=X(1,1);B=X(2,1);C=X(3,1);D=X(4,1);E=X(5,1);
D1=(B*A(1,4)+2*C*A(1,5)+E)**2;D2=(2*A1*A(1,4)+B*A(1,5)+D)**2;
NI=ABS(2*A1*D1-2*B*D1**.5*D2**.5+2*C*D2);
TI=(D1+D2)**1.5;
```

COMPUTER SIMULATION USING THE RADIUS OF CURVATURE TO ESTIMATE R USING THE IMPOSED CONDITIONS (CONTINUED).

```
RI=TI*/NI;
RTOT=RTOT+RI;
COUNT=COUNT+1;
D1=(B*A(2,4)+2*C*A(2,5)+E)**2;D2=(2*A1*A(2,4)+B*A(2,5)+D)**2;
NJ=ABS(2*A1*D1-2*B*D1**.5*D2**.5+2*C*D2);
TJ=(D1+D2)**1.5;
RJ=TJ*/NJ;
RTOT=RTOT+RJ;
COUNT=COUNT+1;
D1=(B*A(3,4)+2*C*A(3,5)+E)**2;D2=(2*A1*A(3,4)+B*A(3,5)+D)**2;
NK=ABS(2*A1*D1-2*B*D1**.5*D2**.5+2*C*D2);
TK=(D1+D2)**1.5;
RK=TK*/NK;
RTOT=RTOT+RK;
COUNT=COUNT+1;
D1=(B*A(4,4)+2*C*A(4,5)+E)**2;D2=(2*A1*A(4,4)+B*A(4,5)+D)**2;
NL=ABS(2*A1*D1-2*B*D1**.5*D2**.5+2*C*D2);
TL=(D1+D2)**1.5;
RL=TL*/NL;
RTOT=RTOT+RL;
COUNT=COUNT+1;
D1=(B*A(5,4)+2*C*A(5,5)+E)**2;D2=(2*A1*A(5,4)+B*A(5,5)+D)**2;
NM=ABS(2*A1*D1-2*B*D1**.5*D2**.5+2*C*D2);
TM=(D1+D2)**1.5;
RM=TM*/NM;
RTOT=RTOT+RM;
COUNT=COUNT+1;
END;
END;
END;
END;
END;
END;
R=RTOT*/COUNT;
IF TLOOP=1 THEN ALR=R;
           ELSE ALR=ALR||R;
TOTR=TOTR+R;
END;
SIMR=TOTR*/100;
PRINT SIMR ALR COUNT;
```

COMPUTER SIMULATION USING THE RADIUS OF ALL POSSIBLE CIRCLES.

```
DATA ONE;
TITLE THESIS PROJECT MARVIN HARRELL CIRCLES ALL POSSIBLE CHOICES;
R=10;
DO K=1 TO 100;
Y=12;
DO J=1 TO 7;
Y=Y-3;
XERROR=SQRT(3)*RANNOR(1613218064);
X= -SQRT(R**2-(Y)**2)+XERROR;
E=1;
XSQ=X**2;
YSQ=Y**2;
XY=X*Y;
OUTPUT;
END;
END;
DATA NOW;
SET ONE;
DROP X Y XERROR XY XSQ YSQ  R E J K ;
XI=X;
YI=Y;
EI=E;
CONSTANT=-XSQ-YSQ;
PROC MATRIX;
FETCH SIM DATA=NOW;
R=0;
TOTR=0;
DO TLOOP =1 TO 700 BY 7;
P1=TLOOP;
P2=TLOOP+1;
P3=TLOOP+2;
P4=TLOOP+3;
P5=TLOOP+4;
P6=TLOOP+5;
P7=TLOOP+6;
T=SIM(P1,)//SIM(P2,)//SIM(P3,)//SIM(P4,)//SIM(P5,)//SIM(P6,)//SIM(P7,
COUNT=0;
RTOT=0;
DO I=1 TO (NROW(T)-2);
DO J=I+1 TO (NROW(T)-1);
DO K=J+1 TO NROW(T);
A=T(I,1 2 3)//T(J,1 2 3)//T(K,1 2 3);
CONST=T(I,4)//T(J,4)//T(K,4);
X=SOLVE(A,CONST);
C=X(1,1);D=X(2,1);E=-X(3,1);
RI=SQRT((C#2+D#2+4#E)#/4);
RTOT=RTOT+RI;
COUNT=COUNT+1;
END;
END;
END;
R=RTOT#/COUNT;
IF TLOOP=1 THEN ALR=R;
            ELSE ALR=ALR||R;
TOTR=TOTR+R;
END;
SIMR=TOTR#/100;
PRINT SIMR ALR;
```

COMPUTER SIMULATION USING THE RADIUS OF CIRCLES WHERE $Y_1 > Y_2 > Y_3$ AND $X_1 > X_2$ AND $X_3 > X_2$.

```
DATA ONE;
TITLE THESIS PROJECT MARVIN HARRELL CIRCLES-X1>X2 & X3>X2;
R=10;
DO K=1 TO 1000;
DO J=1 TO 7;
IF J=1 THEN Y=8.5;
IF J=2 THEN Y=5.75;
IF J=3 THEN Y=2;
IF J=4 THEN Y=0;
IF J=5 THEN Y=-2.75;
IF J=6 THEN Y=-4.25;
IF J=7 THEN Y=-9.99;
XERROR=SQRT(2)*RANNOR(45054871 3);
X= -SQRT(R**2-(Y)**2)+XERROR;
E=1;
XSQ=X**2;
YSQ=Y**2;
XY=X*Y;
OUTPUT;
END;
END;
DATA NOW;
SET ONE;
DROP X Y XERROR XSQ YSQ  R  E J K XY;
XI=X;
YI=Y;
EI=E;
CONSTANT=-XSQ-YSQ;
PROC MATRIX;
FETCH SIM DATA=NOW;
R=0;
TOTR=0;
ALR=0;
DO TLOOP =1 TO 7000 BY 7;
P1=TLOOP;
P2=TLOOP+1;
P3=TLOOP+2;
P4=TLOOP+3;
P5=TLOOP+4;
P6=TLOOP+5;
P7=TLOOP+6;
T=SIM(P1,)//SIM(P2,)//SIM(P3,)//SIM(P4,)//SIM(P5,)//SIM(P6,)//SIM(P7,
COUNT=0;
RTOT=0;
DO I=1 TO (NROW(T)-2);
DO J=I+1 TO (NROW(T)-1);
DO K=J+1 TO NROW(T);
A=T(I,1 2 3)//T(J,1 2 3)//T(K,1 2 3);
CONST=T(I,4)//T(J,4)//T(K,4);
IF A(1,1)>A(2,1) & A(3,1)>A(2,1) & A(1,2)>A(2,2)&A(2,2)>A(3,2) THEN D
X=SOLVE(A,CONST);
C=X(1,1);D=X(2,1);E=-X(3,1);
RI=SQRT((C**2+D**2+4*E)*/4);
RTOT=RTOT+RI;
COUNT=COUNT+1;
END;
END;
END;
END;
R=RTOT#/COUNT;
```

COMPUTER SIMULATION USING THE RADIUS OF CIRCLES WHERE $Y_1 > Y_2 > Y_3$ AND $X_1 > X_2$ AND $X_3 > X_2$ (CONTINUED).

```
IF ALR(1,1)<>0 & R<>0 THEN ALR=ALR||R;
IF ALR=0 & R<>0 THEN ALR=R;
END;
SIMR=SUM(ALR)#/NCOL(ALR);
PRINT SIMR ;
SUMSQD=0;
DO I=1 TO NCOL(ALR);
SUMSQD=SUMSQD+(ALR(1,I)-SIMR)##2;
END;
VAR=SUMSQD#/(NCOL(ALR)-1);
PRINT VAR ALR;
```

COMPUTER SIMULATION USING THE RESIDUAL APPROACH.

```
DATA ONE;
TITLE THESIS PROJECT MARVIN HARRELL RESIDUAL APPROACH;
R=10;
DO K=1 TO 100;
Y=15;
DO J=1 TO 4;
Y=Y-6;
XERROR=SQRT(.5)*RANNOR(746392763);
X= -SQRT(R**2-(Y)**2)+XERROR;
E=1;
XSQ=X**2;
YSQ=Y**2;
XY=X*Y;
OUTPUT;
END;
END;
DATA NOW;
SET ONE;
DROP X Y XERROR XY XSQ YSQ  R E J K;
XI=X;
YI=Y;
EI=E;
CONSTANT=-XSQ-YSQ;
PROC MATRIX;
FETCH SIM DATA=NOW;
R=0;
TOTR=0;
DO TLOOP =1 TO 400 BY 4;
P1=TLOOP;
P2=TLOOP+1;
P3=TLOOP+2;
P4=TLOOP+3;
T=SIM(P1,)//SIM(P2,)//SIM(P3,)//SIM(P4,);
COUNT=0;
RTOT=0;
DO I=1 TO (NROW(T)-2);
DO J=I+1 TO (NROW(T)-1);
DO K=J+1 TO NROW(T);
A=T(I,1 2 3)//T(J,1 2 3)//T(K,1 2 3);
CONST=T(I,4)//T(J,4)//T(K,4);
X=SOLVE(A,CONST);
C=X(1,1);D=X(2,1);E=-X(3,1);
RI=SQRT((C##2+D##2+4#E)#/4);
RTOT=RTOT+RI;
COUNT=COUNT+1;
CTRABS=(-C)#/2;
CTRORD=(-D)#/2;
 IF COUNT=1 THEN CENTERS=CTRABS||CTRORD;
                ELSE DO;
                     CTR=CTRABS||CTRORD;
                     CENTERS=CENTERS//CTR;
                     END;
 IF COUNT=1 THEN ALRI=RI;
                ELSE ALRI=ALRI||RI;
 END;
 END;
 END;
 R=RTOT#/COUNT;
 DO I=1 TO NROW(CENTERS);
 RH=0;
```

COMPUTER SIMULATION USING THE RESIDUAL APPROACH (CONTINUED):

```
DO J=1 TO NROW(T);
H=CENTERS(I,1);
K=CENTERS(I,2);
RHAT=SQRT((H-T(J,1))**2+(K-T(J,2))**2);
RH=RH+RHAT;
END;
ESTR=RH#/NROW(T);
IF I=1 THEN ESTIMR=ESTR;
       ELSE ESTIMR=ESTIMR||ESTR;
END;
DO I=1 TO NROW(CENTERS);
SUMSQ=0;
DO J=1 TO NROW(T);
H=CENTERS(I,1);
K=CENTERS(I,2);
ESTR=ESTIMR(1,I);
RHAT=SQRT((H-T(J,1))**2+(K-T(J,2))**2);
RES=RHAT-ESTR;
SUMSQ=SUMSQ+RES**2;
END;
IF I=1 THEN SUMRESSQ=SUMSQ;
       ELSE SUMRESSQ=SUMRESSQ||SUMSQ;
END;
DO I=1 TO NROW(CENTERS);
IF SUMRESSQ(1,I)=MIN(SUMRESSQ) THEN Q=I;
END;
IF TLOOP=1 THEN REST=ESTIMR(1,Q);
           ELSE REST=REST||ESTIMR(1,Q);
IF TLOOP=1 THEN ALR=R;
           ELSE ALR=ALR||R;
END;
NEWR=SUM(REST)#/NCOL(REST);
SUMSQD=0;
DO I= 1 TO NCOL(REST);
SUMSQD=SUMSQD+(REST(1,I)-NEWR)##2;
END;
VAR=SUMSQD#/(NCOL(REST)-1);
PRINT NEWR VAR REST;
```

APPENDIX B

TABLE 1

SIMULATION RESULTS USING THE RADIUS OF CURVATURE TO ESTIMATE R USING ALL FEASIBLE COMBINATIONS OF FIVE POINTS.

n = Number of unique Y's	σ^2	$\bar{\hat{R}}$	$VAR(\hat{R})$	Number of Simulations	Values of Y
7	3	951.583		100	*
7	1	1016.32		100	*
7	.5	1043.87		100	*
7	.1	75.9853		100	*
5	1	200.845		100	+
5	.5	36.1984		100	+
5	.1	11.9531		100	+

* y = -9, -6, -3, 0, 3, 6, 9
+ y = -9, -4.5, 0, 4.5, 9

TABLE 2

SIMULATION RESULTS USING THE RADIUS OF CURVATURE TO ESTIMATE R USING THE IMPOSED CONDITIONS.

n = Number of unique Y's	σ^2	$\bar{\hat{R}}$	VAR($\hat{R}$)	Number of Simulations	Values of Y
7	3	422.305		100	*
7	1	994.002		100	*
7	.5	565.84		100	*
7	.1	84.816		100	*
5	3	118.68		100	+
5	1	155.00238		100	+
5	.5	31.786		100	+
5	.1	11.9531		100	+

* y = -9, -6, -3, 0, 3, 6, 9
\+ y = -9, -4.5, 0, 4.5, 9

TABLE 3

SIMULATION RESULTS USING THE RADIUS OF ALL POSSIBLE CIRCLES.

n = Number of unique Y's	σ^2	$\overline{\hat{R}}$	VAR($\hat{R}$)	Number of Simulations	Values of Y
7	3	26.5385		100	*
7	1	27.8108		100	*
7	.5	32.0948		100	*
7	.1	13.4527		100	*
4	2	22.3971	3493.13	1000	**
4	.5	11.9848	384.459	1000	**
3	3	12.7617		100	***
3	1	10.4686		100	***
3	.5	10.1964		100	***
3	.1	10.0281		100	***
3	1	10.4686		100	***
3	.5	10.0452		100	***
3	2	10.9975	8.88369	1000	***
3	.5	10.2015	.623521	1000	***

* y = -9, -6, -3, 0, 3, 6, 9
** y = -9, -3, 3, 9
*** y = -9, 0, 9

TABLE 4

SIMULATION RESULTS USING THE RADIUS OF CIRCLES WHERE $Y_1 > Y_2 > Y_3$ AND $X_1 > X_2$ AND $X_3 > X_2$.

n = Number of unique Y's	σ^2	$\bar{\hat{R}}$	$VAR(\hat{R})$	Number of Simulations	Values of Y
9	3	8.33541		1000	*
9	2	8.7353	5.94636	1000	*
9	2	8.69646	4.2556	1000	*
9	1	9.30024		1000	*
9	.5	9.59886	2.95244	1000	*
9	.5	9.58272		1000	*
9	.1	9.94725		1000	*
7	3	9.55826		100	**
7	3	9.30576		1000	**
7	3	9.32069		1000	**
7	2	9.4	3.6845	1000	**
7	1	9.64991		100	**
7	1	9.61095		100	**
7	1	9.60765		1000	**
7	.5	9.98138		100	**
7	.5	9.92176		1000	**
7	.5	9.95295		1000	**
7	.5	9.90503	2.960903	1000	**
7	.1	10.08		100	**
7	.1	10.0077		1000	**
7	.1	9.99178		1000	**
7	2	4.44426	1.31113	1000	+*
7	.5	4.70076	.964667	1000	+*
7	2	3.02089	.78478	1000	+**

TABLE 4 CONTINUED

SIMULATION RESULTS USING THE RADIUS OF CIRCLES WHERE $Y_1 > Y_2 > Y_3$ AND $X_1 > X_2$ AND $X_3 > X_2$.

n = Number of unique Y's	σ^2	$\overline{\hat{R}}$	$VAR(\hat{R})$	Number of Simulations	Values of Y
7	.5	4.08815	3.34535	1000	+**
7	2	10.813	4.77087	1000	++
7	.5	10.485	1.01097	1000	++
7	2	6.65472	7.40619	1000	++*
7	.5	8.02883	14.003	1000	++*
7	2	8.5831	3.68755	1000	++**
7	.5	9.18329	4.00019	1000	++**
5	3	9.86853		1000	***
5	2	9.75276	3.59142	1000	***
5	1	9.77383		1000	***
5	.5	10.496		1000	***
5	.5	9.89979	2.727	1000	***
5	.1	10.2214		1000	***
4	3	10.5384		1000	****
4	2	10.2625	12.8887	1000	****
4	1	9.82344		1000	****
4	.5	9.7139		1000	****
4	.5	9.69888	.464514	1000	****
4	.1	9.77821		1000	****

*	y = -8, -6, -4, -2, 0, 2, 4, 6, 8
**	y = -9, -6, -3, 0, 3, 6, 9
+*	y = 0, 1.75, 3.25, 4.75, 6.9, 8.5, 9.75
+**	y = -3, -1.25, -.75, 0, 1, 2.75, 3.5
++	y = -9.99, -8.25, -6.95, 0, 7, 8.75, 9.5
++*	y = -5.99, -2.25, -1.75, 0, 1, 3.75, 6.5
++**	y = -9.99, -4.25, -2.75, 0, 2. 5.75, 8.5
***	y = -9, -4.5, 0, 4.5, 9
****	y = -9, -3, 3, 9

TABLE 5

SIMULATION RESULTS USING THE RESIDUAL APPROACH.

n = Number of unique Y's	σ^2	$\bar{\hat{R}}$	VAR($\hat{R}$)	Number of Simulations	Values of Y
9	2	9.83033	6.31518	100	*
9	.5	9.88901	.937487	100	*
7	3	10.6352		100	**
7	3	10.3748		100	**
7	3	10.3429		1000	**
7	2	10.132	3.51527	100	**
7	1	10.2527		100	**
7	1	10.205		100	**
7	1	10.045		1000	**
7	.5	10.1114		100	**
7	.5	10.039	.419528	100	**
7	.5	10.1436		100	**
7	.5	10.0044		1000	**
7	.1	10.0434		1000	**
7	.1	10.0322		100	**
7	.1	9.99301		1000	**
5	2	10.8199	8.49873	100	***
5	.5	10.2119	.513119	100	***
4	2	10.183	6.10877	100	****
4	.5	9.8531	.674463	100	****

* y = -8, -6, -4, -2, 0, 2, 4, 6, 8
** y = -9, -6, -3, 0, 3, 6, 9
*** y = -9, -4.5, 0, 4.5, 9
**** y = -9, -3, 3, 9