Considerable thought, research, and concern has been expanded in an effort to determine whether the assumption of a quadratic relation between a single predictor and a criterion violated the assumptions which Johnson and Neyman (1936) state for calculating regions of significance about interacting regressions. In particular, there has been special concern for the assumption of linearity. Debate has ranged from whether linearity referred to the functional relation of the criterion and predictor to whether it was in the context of a linear statistical model, if not both. This paper extends the Johnson-Neyman technique to the curvilinear case and then illustrates this extension by reanalyzing data from two previously published research studies which have considered only a linear relationship between a covariable criterion. A computer program is included in the appendix. (Author/RC)
The Research and Development Center for Teacher Education

UT R&D

The University of Texas at Austin
The Research and Development Center for Teacher Education was established on the campus of the University of Texas at Austin in 1965 to design, build, and test effective products to prepare teachers for careers in the nation’s schools.

A staff of more than 100 are engaged in projects ranging from basic research into effective teaching behavior, through development of special counselor training strategies, to the development, implementation, and evaluation of a complete and radically different undergraduate teacher education program.

The Center’s major program, the Personalized Teacher Education Program, has its roots in teacher personality research dating back to the mid-Fifties. This early research, which demonstrated how teachers’ personalities and classroom behavior correlate with success in their teaching careers, has led to the development of a large group of products which help education facilities become aware of student teachers’ individual needs. The program also has produced products for student teachers’ use to help them build on their strengths.

The completely modularized program is currently in field test and/or use at more than a dozen important teacher education institutions nationally.

In addition to the PTEP, the Center also supports other projects in educational evaluation, development of strategies for implementing institutional change, and in consultation techniques for helping teachers plan individualized programs for children.

The Center’s work is supported by the National Institute for Education and by the University of Texas System, as well as through contract research and development programs for public agencies.
Curvilinear Extensions to Johnson–Neyman
Regions of Significance and Some
Applications to Educational
Research

Kenneth W. Wunderlich and Gary D. Borich

The University of Texas


The research reported herein was supported in part by U. S. Office of Education Contract OE-6-10-108 and National Institute of Education Contract NE-C-00-3-0066, Research and Development Center for Teacher Education. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education or the National Institute of Education and no official endorsement by those offices should be inferred.
Curvilinear Extensions to Johnson-Neyman Regions of Significance and Some Applications to Educational Research

Kenneth W. Wunderlich and Gary D. Borich
The University of Texas

Consider an experiment in which there are two treatments, one criterion and one covariable, and that the researcher would like to detect a covariable by treatment interaction. In order to detect this interaction, the researcher can (1) divide the covariable into any number of discrete categories (i.e., blocks) and perform a Treatment x Blocks analysis of variance (Edwards, 1968) or (2) leave the covariable a continuous measure and test for the homogeneity of group regressions (Walker and Lev, 1953). For the homogeneity of group regressions analysis the researcher employs each covariable value as a discrete unit of measurement and thereby avoids losing information by assigning different covariable values to the same block.

Cronbach and Snow (1973) have shown that for the case in which there is a moderately strong interaction, the statistical power of the homogeneity of group regressions test is superior to blocking at the median, blocking at the 33rd and 67th percentiles, or similar configurations that may be employed in a Treatment x Blocks design. Classification schemes such as these discard power by treating dissimilar data the same, causing the risk of accepting a false null hypothesis to increase beyond that which can be expected when the homogeneity of group regressions test is applied.

When group regressions are heterogeneous, i.e., E chooses not to employ a factorial design, Johnson and Neyman's (1936) procedure can be
used to determine the region(s) of covariable values for which treatments are significantly different. These regions allow covariable scores to be used to determine the treatment for which each S is best suited or if, for any given S, there is no best treatment. This technique is used generally in the study of aptitude-treatment or trait-treatment interactions and is considered the preferred methodology for such studies (Berliner & Cahen, 1972; Cronbach & Snow, 1973).

Considerable thought has been expended in an effort to determine whether the assumption of a quadratic relation between a single predictor and a criterion violates the assumptions which Johnson and Neyman (1936) state for calculating regions of significance about interacting regressions. In particular, there has been special concern for the assumption of linearity. One may ask whether "linearity" refers to the functional relation of the criterion and predictor or whether it is in the context of a linear statistical model, if not both. The computational procedure, however, reveals no need for the former assumption of linearity, thereby providing the foundation for extending the Johnson-Neyman procedure for determining regions of significance to the curvilinear case.

Johnson and Neyman (1936, p. 73) in discussing the difference, \( \theta \), between the regression equations for two experimental groups as \( \theta(x,y) = f_1(x,y) - f_2(x,y) \) where \( f_1 \) and \( f_2 \) are functions of predictor variables x and y, comment:

The functions \( f_1 \) and \( f_2 \) may be chosen arbitrarily, according to the conditions of the particular problem, with the only theoretical restriction that both functions \( f_1 \) and \( f_2 \) must be linear with regard to the unknown constants they involve. Thus we could assume that, e.g., \( f_1(x,y) = A_0 + A_1 \cos xy + A_2 e^{\sin x} \) but our method would fail if it were necessary to assume \( f_1(xy) = \cos A_1 x + A_2 y \) since here the dependence upon \( A_1 \) and \( A_2 \) is not linear.
Most crucial is the position of the coefficient of the variable term, \( \cos xy \). When placed before the variable term as a multiplicative factor, this term makes the equation consistent with what is known mathematically as a linear equation. The fitting of curvilinear regression equations defined by polynomials of the form \( Y = b_0 + b_1 x + b_2 x^2 + \ldots + b_p x^p \) is in principle no different from the fitting of multiple regression equations.

**Development of Curvilinear Formulae**

**Group Regressions**

Given groups 1 and 2 with regression equations of the form \( \hat{Y} = A_1 + B_1 x + C_1 x^2 \) and \( \hat{Y}_2 = A_2 + B_2 x + C_2 x^2 \) where the subscripts 1 and 2 refer to groups 1 and 2 respectively, the first consideration in the series of hypotheses related to the Johnson-Neyman procedure is that of homogeneity of criterion variance. The second consideration is that of heterogeneity of regression. Johnson and Jackson (1956), in their two-predictor case, provide the formula for computing the variance of the actual scores about the predicted scores and the variance of the observed scores about a regression line with a common slope in order to compute an F-ratio for homogeneity of group regressions. If the probability level associated with this F-ratio is found to be significant, then the Johnson-Neyman test for regions of significance is recommended.

The points at which the regression curves intersect, indicating the points of no difference, can also be determined. The computational forms to be employed in this series of steps is as follows.
Regions of Significance

- The difference in regression curves is expressed as \((A_1 - A_2) + (B_1 - B_2)X + (C_1 - C_2)X^2 = D\). Setting \(D = 0\) yields a quadratic equation \((A_1 - A_2) + (B_1 - B_2)X + (C_1 - C_2)X^2 = 0\) which can be solved. Application of the quadratic formula, \(-B \pm \sqrt{B^2 - 4AC} \over 2A\), yields two zeroes assuming the term \(B^2 - 4AC\) is non-negative. If this term were negative, there would be no point at which the two curves intersect. Such an occurrence is feasible as shown in Figure 1.

Insert Figures 1, 2 and 3

If \(B^2 - 4AC = 0\), then Figure 2 is most likely and, if \(B^2 - 4AC > 0\), then Figure 3 is likely.

To determine regions of significance there are several possibilities based on which of the situations outlined above is considered. Given Figure 1 or 2 one would most likely find a left and right boundary for the region of non-significance as indicated by the shaded portion. To determine these points along the \(X\)-axis, consider the formula:

\[
\left( \frac{D^2}{P + Q} \right)^{1/2} \left( \frac{S^2}{N - 6} \right)
\]

where \(S^2_a\) is the best estimate of the error variance obtained by pooling the variances about the group regression lines and \(\frac{D^2}{P + Q}\) is the variance for the difference in regression lines where \(D = (A_1 - A_2) + (B_1 - B_2)X + (C_1 - C_2)X^2\) and where \(P + Q\) is a "scaling" factor used to determine the variance for the difference between the regression equations for the two groups. The value of \((P + Q)\) depends upon the values of the basic characteristics of matching the \(x'\)'s and \(y'\)'s. Where \(x'\) and \(y'\) lie near the population means of \(x\) and \(y\), the value of \((P + Q)\) becomes small since \((P + Q)\) becomes \(\frac{1}{N_1} + \frac{1}{N_2}\) when \(\bar{x}_1 = x' = \bar{x}_2\) and \(\bar{y}_1 = y' = \bar{y}_2\). \(D^2\) is of
the general form $AX^4 + BX^3 + CX^2 + DX + E$ so that the equation
\[
\left( \frac{\frac{D^2}{P+Q}}{\frac{a^2}{N-6}} \right) = F_{.05}(1,N-6) \text{ expressible as } D^2 - F_{.05} \frac{[P+Q]a^2}{N-6} = 0 \]
is a fourth order (quartic) equation. This is consistent with the illustrations presented earlier. If four distinct solutions exist then two regions of non-significance are defined as indicated in Figure 3. To find the solutions to this equation, a method suggested by Standard Mathematical Tables (1965) is available which requires finding a solution of a resolvent cubic equation first but which in turn can be used to obtain all four roots of the above fourth order equation. The following formulae are used in finding these values.

Given the quartic equation (1) $x^4 + ax^3 + bx^2 + cx + d = 0$, a form which the equation obtained from $D^2$ can assume, a resolvent cubic equation of the form (2) $y^3 - by^2 + (ac - 4d)y - a^2d + 4bd - c^2 = 0$ is obtained. A root of this equation is obtained by reducing equation (2) to the form (3) $z^3 + fz + g = 0$; the "second" order term is eliminated by substituting $(z + \frac{b}{3})$ for $y$ in (2) where $f = \frac{1}{3} \left[ 3(ac - 4d) - b^2 \right]$ and $g = \frac{1}{27} \left[ 2b^3 - 9b(ac - 4d) + 27(-a^2d + 4bd - c^2) \right]$. Let $A = \sqrt[3]{\frac{-g}{2} + \frac{f^2}{4}} + \sqrt[3]{\frac{f^3}{27}}$ and $B = \sqrt[3]{\frac{-g}{2} - \frac{f^2}{4}} + \frac{f^3}{27}$; then $A + B = M$ is a root of equation (2) from which the roots for (1) can be obtained. Consider $R = \sqrt[4]{\frac{a^2}{4}} - b + M$, if $R \neq 0$,

$D = \sqrt{\frac{3a^2}{4} - R^2 - 2b + \frac{4ab - 8c}{4R}}$ and $E = \sqrt{\frac{3a^2}{4} - R^2 - 2b - \frac{4ab - 8c}{4R}}$. Otherwise $D = \sqrt{\frac{3a^2}{4} - 2b + 2M^2 - 4d}$ and $E = \sqrt{\frac{3a^2}{4} - 2b - 2M^2 - 4d}$. The four roots of the fourth order equation are then $x = \frac{-a}{4} + R + \frac{D}{2}$ and $x = \frac{-a}{4} - R + \frac{E}{2}$. These roots are real dependent upon the arguments for the square roots above being greater than or equal to zero.
Research Applications

The above computations may be simplified through the use of a computer program (Wunderlich and Borich, 1973) which, after solving the fourth order equation for determining regions of significance, plots the data points, within-group regressions and regions of significance for the case in which there are two treatments, one criterion and one predictor. Before applying the program, however, the investigator must consider the full range of calculations, both linear and nonlinear, that may be required by any given data set. These calculations are diagrammed sequentially in Figure 4.

Insert Figure 4 about here

Note that the sequence begins by testing the assumption of curvilinearity and that criterion-covariable relationships may be curvilinear for groups separately or combined. In one case the combined groups may exhibit a single underlying curvilinear relationship such that the two groups are samples from the same population, i.e., regression slopes for the groups are homogeneous. In a second case the relations between the covariable and criterion may be curvilinear for each group separately, i.e., regression slopes for the groups are heterogeneous. Only for the latter case are regions of significance tenable. Whereas the foregoing discussion addresses both homogeneity of curvilinear regressions and regions of significance, it does not cover a third opportunity to detect a curvilinear relationship. In a third case regressions that are both linear and non-intersecting (ordinal) may mask an underlying curvilinear relationship in which linear slopes of both groups form a single curvilinear trend. The continuity of separate linear slopes can take the form of a single curvilinear regression that is easily missed when the efficiency of linear
and curvilinear models are not compared at the outset of the research. This initial comparison and the sequence of steps representing the left portion of Figure 4 are calculated by the curvilinear program. When one or both within-group regressions are linear, i.e., a curvilinear model fails to significantly improve prediction, the researcher is referred to a linear program by Borich (1971) for the case in which there is one predictor or to a linear program by Borich and Wunderlich (1973) for the case in which there are two predictors.

To illustrate the above computations, a research study conducted by Hughes and published in the *American Educational Research Journal* (1973, 10, 21-37) was reanalyzed with the Wunderlich and Borich (1973) curvilinear program.

**Background.** Hughes set out to establish that differences in residual posttest achievement are a function of pupils responding to teachers' questions (a) randomly, (b) systematically and (c) in a self-selected style in which the pupil has the choice of whether or not to respond. With three schools available for experimentation, E chose a 3 x 3 factorial design with schools as the first factor and type of responding as the second. Criterion performance was established with a 222-item posttest based upon the content of the treatment, a wildlife lesson directed to seventh- and eighth-grade pupils. In addition, E collected scores on nine covariables with the foresight that these may confound posttest performance. Among these E included the most likely predictor—that of pretest achievement on the criterion instrument. E chose to remove posttest variance attributable to differences on the pretest by using as the criterion the difference between predicted and actual posttest scores (i.e., the residuals).
Hughes' data provided the opportunity to uncover nine aptitude-treatment interactions. For our illustration, however, the reanalysis was limited to one of the nine covariables, that of the pretest and to only one pair of treatments, that involving the self-selected and systematic response groups.

Results. These data, while previously reanalyzed by Borich (1974) with the assumption that covariable-criterion relationships were linear, were reanalyzed a second time assuming that relationships between covariable and criterion were curvilinear. Application of the curvilinear program to these data, however, revealed no significant improvement in prediction from that obtained from a linear model for either the systematic response or the self-selected response groups and therefore the investigator was referred to the more suitable linear program (Borich, 1971). In the interest of providing a real data illustration of the curvilinear technique, the program then was forced subsequent to the model comparisons test to continue as though curvilinear regressions had been found in one or more treatments. The program, therefore, calculated the homogeneity of group regressions test for curvilinear regressions, the test for a common intercept and regions of significance rather than "exit" as would be expected after the model comparison test for lack of a significant improvement in predictive efficiency with the curvilinear model. The resulting plot from the curvilinear analysis of these data is reported in Figure 5 and the resulting plot from the corresponding linear analysis by Borich (1974) is reported in Figure 6.

Insert Figures 5 and 6 about here
Note that for both analyses a region of significance is indicated to the left of the data mass. While the region of significance for the curvilinear plot begins at the covariable value of 14.20 (taken from the printout) and the region of significance for the linear plot at the covariable value of 13.20, both regions lie in approximately the same area and encompass approximately the same range of covariable values. These ranges represent those covariable values for which the treatments are significantly different at or beyond the .05 level. For this example the curvilinear plot lends little new information to the study and, if employed, might needlessly complicate its interpretation. Using the linear plot, we might conclude that those subjects scoring below 13.20 on the pretest should receive the self-selected response treatment while those falling above this value are likely to achieve equally well from either treatment and therefore should receive the least costly of the two treatments.

Other configurations. Lest the reader receive the impression from this example that the present data represent the only configuration of curvilinear regressions and corresponding regions of significance, other hypothetical distributions have been analyzed and plotted and are reported in an appendix to this report. A listing of the program used to generate these plots is also provided.
References


curvilinear Are group relationships curvilinear?
(Model Comparison Test)

Are curvilinear slopes different?
(Test homogeneity of regressions)

yes

Are there regions of significance?
(Johnson-Neyman)

no

Are intercepts different?
(ANCOVA)

Are intercepts different?
(ANCOVA)

yes

Is the interaction ordinal?
(Inspect regression lines)

no

Are intercepts different?
(ANCOVA)

Are there regions of significance?
(Johnson-Neyman)

Is there a single curvilinear regression?
Model Comparison Test

no

Are there regions of significance?
(Johnson-Neyman)

yes

STOP
Region of Significance

self-selected response group

systematic response group

PREDICTOR VARIABLE

CRITERION VARIABLE

Fig 5
1.3 as
Region of Significance
systematic response group
self-selected response group

Fig 6
Figure Captions

Figure 1. $B^2 - 4AC < 0$.
Figure 2. $B^2 - 4AC = 0$.
Figure 3. $B^2 - 4AC > 0$.
Figure 4. Sequence of linear and curvilinear calculations for ATI analyses.
Figure 5. Region of significance ($< 14.21$) and nonsignificance ($> 14.20$) for Hughes (curvilinear analysis).
Figure 6. Region of significance ($< 13.21$) and nonsignificance ($> 13.20$) for Hughes (linear analysis).
Appendix

a. A positive linear regression and a moderately strong curvilinear regression with small left and right regions of significance

b. Two slightly curvilinear positive regressions with a moderately sized left and large right region of significance

c. Two moderately strong curvilinear regressions in opposite directions with a small left and a moderately sized right region of significance

d. A slightly curvilinear regression and a moderately strong curvilinear regression in opposite directions with a small left and a large right region of significance
PROGRAM CURVATI( INPUT, OUTPUT, PLOT )

(PRELIMINARY VERSION: 4/5/74)

CONTROL CARDS AND THE DATA ARE ARRANGED IN THE FOLLOWING ORDER

FIRST CARD:  COLS, 1-5 = NUMBER OF SUBJECTS IN GROUP 1

COLS, 6-10 = NUMBER OF SUBJECTS IN GROUP 2

COLS, 15 = MISSING DATA OPTION

IF ZERO, ALL DATA IS INCLUDED IN THE ANALYSIS

IF ONE, BLANKS ARE NOT INCLUDED BUT ZEROES ARE

IF TWO, BLANKS AND ZEROES ARE NOT PROCESSED

COLS, 16-28 = ALPHA (P) LEVEL USED TO DECIDE WHETHER A LINEAR OR CURVILINEAR MODEL IS APPROPRIATE

SECOND CARD:  COLS, 1-5 = F-VALUE ASSOCIATED WITH P, ABOVE (DF=1 AND N=3)

THIRD CARD:  GROUP 1 FORMAT (PREDICTOR MUST PRECEDE CRITERION)

INSERT GROUP 1 DATA HERE

FOURTH CARD:  GROUP 2 FORMAT

INSERT GROUP 2 DATA HERE

PROGRAMMED BY KEN W. UNDERLICH AND GARY D. BORICH, RESEARCH AND DEVELOPMENT CENTER FOR TEACHER EDUCATION, THE UNIVERSITY OF TEXAS AT AUSTIN

DIMENSION Y1(230), X1(230), Y2(230), X2(230), X12(230), Y(450)

READ 1, N1, N2, MISS, PBLV

READ 2, (FRMT1(I), I=1,8)

READ 2( (FRMT2(I), I=1,8)

READ AMISDAT (X1, Y1, N1, MISS)

DO 3 K=1, N1

X12(K)=X1(K)*X1(K)

CALL SUMDATA(Y1, X1, X12, N1)

READ 2( (FRMT2(I), I=1,8)

READ AMISDAT (X2, Y2, N2, MISS)

DO 4 K=1, N2

X22(K)=X2(K)*X2(K)

CALL SUMDATA(Y2, X2, X22, N2)

N=N1+N2

DO 5 K=1, N

X(K)=X1(K)

Y(K)=X2(K)

5 X02(K)=X12(K)

DO 6 K=1, N2

X(N1+K)=X22(K)

5 X02(N1+K)=X22(K)

CALL SUMDATA(Y, X, X, X02, Y, XM, X02M, YSD, XSD, X02SD, Y88, X88, X0288, RX, RY

1X02, RX02, RSYX, RSYX02, R88X02, N)

PRINT 7

PRINT 7 (I11)

PRINT 8, N1, N2, N, Y1M, Y2M, Y8, Y8D, X1M, X2M, XM, X1SD, X2SD, XSD,

1YRXY, XRY, X2Y


2X, F8, 3, 6X, F8, 3, //, 5X, *CRITERION ST, DEV=9X, F8, 3, 7X, F8, 3, 6X, F8, 3, //, 5X, *PREDICTOR MEAN=9X, F8, 3, 7X, F8, 3, 6X, F8, 3, //, 5X, *PREDICTOR ST,
C MODEL COMPARISON

RSQ1 = (RYX1*RYX1 + RYX12*RYX12)/(1.0 + RXX12*RYX12)
RSQ2 = (RYX2*RYX2 + RYX22*RYX22)/(1.0 + RXX22*RYX22)
R1 = RYX1*RYX1 
R2 = RYX2*RYX2 
VAYXX1 = Y1SD*Y1SD*(1.0 + RSQ1) 
VAYXX2 = Y2SD*Y2SD*(1.0 + RSQ2) 
VAYX1 = Y1SD*Y1SD*(1.0 + RYX1*RYX1) 
VAYX2 = Y2SD*Y2SD*(1.0 + RYX2*RYX2) 
FGP1 = (RSQ1 + R1)/(1.0 + RSQ1)/ANDF1 
FGP2 = (RSQ2 + R2)/(1.0 + RSQ2)/ANDF2 
ANDF1 = N1*2.0 
ANDF2 = N2*2.0 
SLMOD1 = PRBF(1.0, ANDF1, FGP1) 
SLMOD2 = PRBF(1.0, ANDF2, FGP2) 
IF(SLMOD1.LE.PBLV.OR.SLMOD2.LE.PBLV) 17, 16
16 PRINT 16, PBLV
18 FORMAT(//, 5X, A COMPARISON OF A LINEAR AND QUADRATIC MODEL FOR EACH GROUP 1 AND 2 SUGGESTS NO SIGNIFICANT DIFFERENCE AT THE 0.05 * F, 2 *, LEVEL AND THUS USE OF A LINEAR MODEL AND AN ANALYSIS IS 3 SUGGESTED. SEE BORICH, EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT 47, */, 5X, 1971, 31, 251, 253, * )
GO TO 101

17 BX1 = (Y1SD/X1SD)*(RYX1 + RYX12*RXX12)/(1.0 + RXX12*RYX12)
BX2 = (Y2SD/X2SD)*(RYX2 + RYX22*RXX22)/(1.0 + RXX22*RYX22)
A01 = Y1M + BX1*X1M 
A02 = Y2M + BX2*X2M 
PRINT 12, A01, BX1, BX2, A02, BX1, BX2 
12 FORMAT(//, 5X, THE REGRESSION EQUATION FOR GROUP 1 IS Y = *F8.4* + *F8.4* X + *F8.4* XX, */, 5X, THE REGRESSION EQUATION FOR GROUP 2 IS Y = *F8.4* + *F8.4* X + *F8.4* XX, * )
PT1 = PT2 = 0.0 
ERROR = VAYXX1 + VAYXX2 
SYY = Y1SS + Y2SS 
SXX = X1SS + X2SS 
SXX2 = X12SS + X22SS 
RYXSS = RSSYX1 + RSSYX2 
RYX2SS = RSSYX2 + RSSYX22 
RX22SS = RSSX12 + RSSX22 
VAREGRS = SYY**2 + (RYXSS**2 + SXX2 + SXX*RYX2SS**2)/(SXX**2 + RSSX**2) 
ANO = ANDF/2.0 
FREGRES = ANDF*(VAREGRS - ERROR)/ERROR 
PREGRES = PRBF(2.0, ANDF, FREGRES) 
PRINT 14, FREGRES, PREGRES
14 FORMAT(//, 5X, THE F RATIO FOR A TEST OF HOMOGENEITY OF REGRESSION IS *F8.3/, 5X, WITH AN ASSOCIATED PROBABILITY LEVEL OF *F8.3*)
VINTER = YSD + YSD*((1.0 + RYX*RYX + RYX2*RYX2 + RXX02*RXX02 + 2.0*RYX*RYX2 + RXX02)/(1.0 + RXX02*RXX02)) 
ANO = N = 3
FINTER=(ANDF/3,0)*(VINTERERROR)/ERROR
PINTER= PRUF(3,0,ANDF,FINTER)
PRINT 15, FINTER,PINTER

15 FORMAT(///,5X,*THE HYPOTHESIS OF COMMON REGRESSION CONSTANTS IS TES
ITED BY AN F-RATIO OF *F8,3/5X,* WHICH HAS AN ASSOCIATED PROBABILIT
2Y LEVEL OF *F8,3)*
AD=A01=A02
BXD=BX1=BX2
BX2D=BX21=BX22
QUAD=BXD*BXD=4,0*B*X2D*AD
IF(QUAD,GE,0,0)20,19

19 PRINT 21
21 FORMAT(///,5X,*THE REGRESSION CURVES FOR GROUPS 1 AND 2 DO NOT INTE
RSECT,*)
GO TO 22
20 IF(QUAD,EQ,0,0)23,24
23 PT1= =BXD/(2,0*BX2D)
PRINT 25,PT1
25 FORMAT(///,5X,*THE REGRESSION CURVES INTERSECT AT A POINT WHERE X I
S EQUAL TO *F6,3)
GO TO 22
24 PT1=(SQRT(QUAD)-BXD)/(2,0*BX2D)
PT2= =-BXD=(SQRT(QUAD)))/(2,0*BX2D)
PRINT 26,PT1,PT2
26 FORMAT(///,5X,*THE REGRESSION CURVES INTERSECT AT THE POINTS WHERE* 1/X IS EQUAL TO *F6,3* AND WHERE X IS EQUAL TO *F6,3)

22 R1=1,0/(1,0=RXX12*RXX12)
R2=1,0/(1,0=RXX22*RXX22)
S1=2,0*RXX12/SQRT(X1SS*X12SS)
S2=2,0*RXX22/SQRT(X2SS*X22SS)
FACTOR=SL*ERROR/ANDF
X4COEF #BX2D*B*X2D= FACTOR*(R1 / X12SS + R2/X22SS)
X3COEF =2,0#BXD*A*XODD +FACTOR*( R1*S1 + R2*S2)
X2COEFF =BXD*BXD*2,0*AD*B*X2D=FACTOR*( R1/X1SS + R2/X25SS = R1*S1*X1
1M = R2+S2*X2M = 2,0*X12M*R1/X12SS = 2,0*B*X22M*R2/X22SS)
X1COEF =2,0*AD*A*XDD =FACTOR*(R1*S1*X12M +R2*S2*X22M=2,0*R1*X1M/X1SS
1=2,0*R2*X2M/X22SS)
CONSTA= AD=AD = FACTOR *( 1,0/AN1 +1,0/AN2 +X1M*X1M*R1/X1SS + R2*
1X2M*X22M/X2SS = R1*S1*X1M*X12M = R2*S2*X22M + R1*X12M*X12M/X12S
2S + R2*X22M*X22M/X22SS)
AA(1)= X4COEF
AA(2)= X3COEF
AA(3)= X2COEF
AA(4)= X1COEF
AA(5)= CONSTA
CALL ZPOLYR (AA,4,Z,Z)
PRINT 55,Z

55 FORMAT (8(2X,F8,10))
NST1=NST2=33
CALL AMinMax(X,N,XMIN,XMAX)
PRINT 2000, XMIN,XMAX

2000 FORMAT (5X,*XMIN = *F8,3*, XMAX = *F8,3*)
28 CALL PARABBX1,BX1,A01,XPTLN1,YPTLN1,XMIN,XMAX,AA1,NST1)
31 CALL PARABBX22,BX22,AA2,XPTLN2,YPTLN2,XMIN,XMAX,AA2,NST2)
33 CALL BGNplt
CALL PLT(1,0,1,0,=3)
CALL SCALE (X(1),10,0,N,1)
CALL SCALE (Y(1), 9,0,N,1)
Qi1(3)=QX2(3)=QX3(5)=QX4(3)=XPTLN1(NST1+1)=XPTLN2(NST2+1)=X(N+1)
Qi1(4)=QX2(4)=QX3(4)=QX4(4)=XPTLN1(NST1+2)=XPTLN2(NST2+2)=X(N+2)
Qi1(3)=QY2(3)=QY3(3)=QY4(3)=YPTLN1(NST1+1)=YPTLN2(NST2+1)=Y(N+1)
Qi1(4)=QY2(4)=QY3(4)=QY4(4)=YPTLN1(NST1+2)=YPTLN2(NST2+2)=Y(N+2)
CALL AXIS (0,0,0,0,18HPREDICTOR VARIABLE,18,10,0,0,X(N+1),X(N+2))
CALL AXIS (0,0,0,0,18HCITERION VARIABLE,18,9,0,90,0,Y(N+1),Y(N+2))
X(N+1)=X2(N+1)=X(N+2)
X(N+2)=X2(N+2)=X(N+2)
Y(N+1)=Y2(N+1)=Y(N+1)
Y(N+2)=Y2(N+2)=Y(N+2)
DO 902 K = 1,N
PRINT 903,X1(K),Y1(K),K
903 FORMAT (2X,F6,5X,F8,2X)
902 CONTINUE
DO 904 K = 1,N
PRINT 905,X2(K),Y2(K),K
905 FORMAT (2X,F8,5X,F8,4X)
904 CONTINUE
PRINT 910
910 FORMAT (* THIS IS GROUP ONE DATA*)
911 FORMAT(* THIS IS GROUP TWO DATA*)
CALL LINE(X1,Y1,N1,1,1,11,1)
CALL LINE(X2,Y2,N2,1,1,5)
PRINT 900,(XPTLN1(J),YPTLN1(J),J=1,NST1)
900 FORMAT (5(2X,F8,5X,F8,3))
PRINT 900,(XPTLN2(J),YPTLN2(J),J=1,NST2)
CALL LINE(XPTLN1(NST1),YPTLN1(NST1),NST1,1,0,0)
CALL LINE(XPTLN2(NST2),YPTLN2(NST2),NST2,1,0,0)
SX1=(XPTLN1(NST1)-X(N+1))/X(N+1)
SY1=(YPTLN1(NST1)-Y(N+1))/Y(N+1)
SX2=(XPTLN2(NST2)-X(N+1))/X(N+1)
SY2=(YPTLN2(NST2)-Y(N+1))/Y(N+1)
IF (AA1,GE,0,0)34,35
34 ANU1=90,0
GO TO 36
35 ANG1=270,0
36 IF (AA2,GE,0,0)37,38
37 ANG2=90,0
GO TO 39
38 ANG2=270,0
39 CALL SYMBOL(SX1,SY1,07,9HGROUP ONE,ANG1,9)
CALL SYMBOL(SX2,SY2,07,9HGROUP TWO,ANG2,9)
TT=10,0*X(N+2)+X(N+1)
TY=9,0*Y(N+2)+Y(N+1)
PRINT 912,TT,TY
912 FORMAT (5X,*TT=*F8,4,*TY=*F8,4)
QX1(1)=QX2(1)=QX3(1)=QX4(1)=X(N+1)
QX1(2)=QX2(2)=QX3(2)=QX4(2)=X(N+1)
QY1(1)=QY2(1)=QY3(1)=QY4(1)=Y(N+1)
QY1(2)=QY2(2)=QY3(2)=QY4(2)=Y(N+1)
DO 70 JL = 1,4
PX(JL)=0,0
IF (AIMAG(Z(JL)) ,EQ, 0,0) 71,72
71 PX(JL)=REAL(Z(JL))
72 CONTINUE
70 CONTINUE
PRINT 913,(PX(JL),JL=1,4)
913 FORMAT (2X,4(2X,F8,3))
IF (PX(1),GE,X(N+1),AND,PX(1),LE,TT)40,41
AN PX(1)=PX(1)
40 QX1(1)=QX1(2)=PX(1)
Y1(1)=Y(N+1)
Y1(2)=TY
41 IF (PX(2),GE,X(N+1),AND,PX(2),LE,TT)42,43

42 $QX2(2) = QX2(1) = PX(2)$
$QY2(1) = Y(N+1)$
$QY2(2) = TY$

43 IF (PX(3) $\geq$ X(N+1), AND, PX(3) $\leq$ TT) THEN 44, 45

44 $QX3(1) = QX3(2) = PX(3)$
$QY3(1) = Y(N+1)$
$QY3(2) = TY$

45 IF (PX(4) $\geq$ X(N+1), AND, PX(4) $\leq$ TT) THEN 46, 47

46 $QX4(1) = QX4(2) = PX(4)$
$QY4(1) = Y(N+1)$
$QY4(2) = TY$

47 CONTINUE

PRINT 901, (QX1(1), QY1(1), I=1, 2)
PRINT 901, (QX2(1), QY2(1), I=1, 2)
PRINT 901, (QX3(1), QY3(1), I=1, 2)
PRINT 901, (QX4(1), QY4(1), I=1, 2)

901 FORMAT (/2(SX,F8,3))

CALL LINE(QX1, QY1, 2, 1, 0, 0)
CALL LINE(QX2, QY2, 2, 1, 0, 0)
CALL LINE(QX3, QY3, 2, 1, 0, 0)
CALL LINE(QX4, QY4, 2, 1, 0, 0)
CALL ENDPLT

101 CONTINUE

SUBROUTINE AMISDAT (A, 3, NS, MISDATA)
DIMENSION A(200), B(200), ILL(200), AA(200), BB(200)
IF (MISDATA, EQ, 0) RETURN
IJK = 0
DO 750 NQ = 2, 200
  ILL(NQ) = AA(NQ) = BB(NQ) = 0, 0
DO 200 N = 1, NS
  IF (A(N), EQ, 0, 0, OR, B(N), EQ, 0, 0) THEN 202, 200
202 IF (MISDATA, EQ, 2) THEN 203, 204
203 IJK = IJK + 1
  ILL(IJK) = N
  GO TO 200
204 IF (.NOT., A(N), OR, .NOT., B(N)) THEN 200, 205
205 IJK = IJK + 1
  ILL(IJK) = N
200 CONTINUE
IF (ILL(1), EQ, 0) THEN 222, 211
211 ICOONT = 0
DO 206 JC = 1, NS
  AA(JC) = ICOONT
DO 207 JV = 1, IJK
  IF (JC, EQ, ILL(JV)) THEN 208, 207
207 CONTINUE
  AA(JC) = ICOONT
  BB(JC) = ICOONT
  GO TO 206
208 ICOONT = ICOONT + 1
206 CONTINUE
NS = NS + IJK
DO 209 JT = 1, NS
  A(JT) = AA(JT)
209 B(JT) = BB(JT)
222 CONTINUE
RETURN
END

SUBROUTINE SUMDATA (A, B, C, AMEAN, BMEAN, CMEAN, ASD, BBD, CSD, SSA, SSB, SSSRC, RAB, RAC, RBC, RRSAB, RRSAC, RRSBC, M)
DIMENSION A(400), B(400), C(400)
AMEAN = BMEAN = CMEAN = SSA = SSB = SSC = RAC = RBC = RSSAB = RSSBC = RSSAC = 0.0

DO 2 J = 1, M
AMEAN = AMEAN + A(J)
BMEAN = BMEAN + B(J)
CMEAN = CMEAN + C(J)

AMEAN = AMEAN / M
BMEAN = BMEAN / M
CMEAN = CMEAN / M

DO 3 K = 1, M
SSA = SSA + (A(K) - AMEAN)**2
SSB = SSB + (B(K) - BMEAN)**2
SSC = SSC + (C(K) - CMEAN)**2

RSAB = RSSAB + (A(K) - AMEAN)*(B(K) - BMEAN)
RSAC = RSSAC + (A(K) - AMEAN)*(C(K) - CMEAN)
RSBC = RSSBC + (B(K) - BMEAN)*(C(K) - CMEAN)

ASD = SQRTF(SSA/(AM**2))
BSD = SQRTF(SSB/(AM**2))
CSD = SQRTF(SSC/(AM**2))

RAB = (RSAB/AM) / (ASD*BSD)
RBC = (RSBC/AM) / (BSD*CSD)
RAC = (RSAC/AM) / (ASD*CSD)

RETURN

FUNCTION PRBF (DA, DB, FR)
PRBF = 1.0
IF (DA*DB*FR.EQ.0.0) RETURN
IF (FR.LT.1.0) GO TO 5

A = DA
B = DB
F = FR
GO TO 10

5 A = DB
B = DA
F = 1.0/FR

10 AA = 2.0/(9.0**4)
BB = 2.0/(9.0**4)
Z = ABS ((1.0 + BB)*F**0.333 + 33.3333 + 1.0 + AA) / SQRT (BB*F**0.66666/4.0)
IF (F.LT.4.0) Z = Z*(1.0 + Z*2.0 + Z**4/BB**3)
PRBF = 0.57*(1.0 + Z*(0.196854 + Z*(0.00344 + Z*0.0019527)))

RETURN

END

SUBROUTINE AMINMAX(A, M, AMIN, AMAX)
DIMENSION A(400)
AMIN = A(1)
AMAX = A(1)
DO 2 J = 2, M
IF (A(J).LT.AMIN) AMIN = A(J)
IF (A(J).GT.AMAX) AMAX = A(J)
2 CONTINUE
RETURN
END

SUBROUTINE PARA8(AA, BB, CC, XPT, YPT, AMIN, AMAX, AAA, NST)
DIMENSION XPT(33), YPT(33)
FORMAT(5, 'THE COEFFICIENT OF X-SQUARED IS ZERO,*')
RETURN
3 H=(BB)/(AA*2,0)
  AK=(AA)*((BB*BB=4,0*AA*CC)/(4,0*AA*AA)
  AAA=1,0/(4,0*AA)
YPT (1)=YPT (33)=AK+6,0*AAA
YPT (2)=YPT (32)=AK+5,5*AAA
YPT (3)=YPT (31)=AK+5,0*AAA
YPT (4)=YPT (30)=AK+4,5*AAA
YPT (5)=YPT (29)=AK+4,0*AAA
YPT (6)=YPT (28)=AK+3,5*AAA
YPT (7)=YPT (27)=AK+3,0*AAA
YPT (8)=YPT (26)=AK+2,5*AAA
YPT (9)=YPT (25)=AK+2,0*AAA
YPT (10)=YPT (24)=AK+1,5*AAA
YPT (11)=YPT (23)=AK+1,0*AAA
YPT (12)=YPT (22)=AK+0,83*AAA
YPT (13)=YPT (21)=AK+0,67*AAA
YPT (14)=YPT (20)=AK+0,50*AAA
YPT (15)=YPT (19)=AK+0,33*AAA
YPT (16)=YPT (18)=AK+0,17*AAA
YPT (17)=AK
DO 6 L=1,16
  FT= SQRT(4,0*AAA*YPT(L)=AK)
  LN =34*L
  XPT(L)=H+FT
  XPT(LN)=H+FT
XPT(17)=H
  IC=JC=0
  DO 7 K = 1,16
    LK=34*K
    IF(XPT(K),LT,AMIN)IC=IC+1
    IF(XPT(LK),GT,AMAX)JC=JC+1
    IF(1C9GT00)9,10
  9 NST=33=IC
    DO 8 M=1,33
      JKL=M+IC
      XPT(M)=XPT(JKL)
    8 YPT(M)=YPT(JKL)
  10 IF (JC9GT0)11,12
  11 NST=NST=JC
  12 CONTINUE
RETURN
END