A BASIC program to determine regions of significance using the Johnson-Neyman technique

CHARLES T. SCIALFA

The Pennsylvania State University, University Park, Pennsylvania

In many experimental designs, the researcher's intent is to determine the relationship between some criterion and two independent factors, one of which is continuous and the second categorical (e.g., representing independent groups or treatment conditions). The analysis of data from such designs involves the comparison of regression slopes across levels of the categorical variable, where the full model being tested allows for group differences in both slopes and intercepts (Huitema, 1980; Pedhazur, 1982; Rogosa, 1980). This test is referred to as a test of homogeneity of regression, or the aptitude $\times$ treatment interaction (ATI). As with the more common factorial model involving two discrete independent factors, after obtaining a significant ATI, one would typically conduct follow-up tests, analogous to tests of simple effects, in order to specify the nature of the interaction. One applicable test of significance involves establishing regions of significance using the Johnson-Neyman technique (Johnson & Neyman, 1936).

Assume, for example, that an investigator has examined the relationship between classroom performance and need for achievement (nAch), measured as continuous variables, under both traditional and self-directed teaching styles. Data from this hypothetical study are shown in Figure 1, where performance is plotted as a function of nAch for both treatment groups. It can be seen that there exists a disordinal ATI, wherein slopes relating nAch to performance differ for the two conditions. Assuming that the interaction is significant, one might wish to determine points along the nAch continuum where there are significant performance differences between the two groups.

The Johnson-Neyman technique accomplishes this goal by establishing regions of significance beyond which the groups display criterion differences. In Figure 1, Region 1 indicates the nAch level at which performance in the traditional group is superior, whereas in Region 3 the self-directed group is superior. Obviously then, in Region 2 performance does not differ significantly be-

tween groups. If one or both regions of significance are in the range of theoretical or applied interest, then the researcher can make more explicit statements concerning the person-in-situation (Cronbach, 1957) relationship manifest in his/her data.

Although Borich, Godbout, and Wunderlich (1976) have developed mainframe programs for conducting the Johnson-Neyman tests, to date none of the commonly used packages (e.g., SPSS-X, SAS, BMDP) include algorithms for establishing regions of significance. The BASIC program described here (see Appendix) allows the user to quickly determine the point of intersection, as well as regions of significance, from ATI data. The user is required only to provide (for each of two groups or conditions) the predictor mean, standard deviation, sample size, residual sums of squares, and least squares estimates for slope and intercept parameters. These data can be most easily obtained by conducting simple regression analyses for each group, using either large-scale software or any of a variety of microcomputer-based statistical packages (e.g., HSD Resources Regress II, MINITAB). The only additional information necessary at input is the tabled critical F with 1 and $N-4$ df, where N is the combined sample size.

At output, the user is provided with three values. The first is the point of intersection, or the value along the covariate continuum at which the two groups exhibit equivalent performance. Also included at output are the two regions of significance ("pick-a-point" or simultaneous, depending on user preference), at which there exist significant group differences for that covariate value. Again, if one or both of these regions are within a range that is of theoretical or applied interest, the user is better able to determine person-in-situation relationships.

The regions of significance established using this technique are nonsimultaneous regions. That is, these regions are appropriate when one wishes to make inferences con-

Figure 1. Performance as a function of nAch for traditional and self-directed teaching conditions, with regions of significance.

This work was supported in part by a grant from the Gerontological Education, Research, and Services Program (GERAS), The University of Notre Dame, Notre Dame, IN 46556, and in part by National Institute on Aging Grant AG00110-02.

The author wishes to express many thanks to Scott Maxwell for providing the motivation for this project as well as insightful comments concerning the manuscript. Reprint requests should be addressed to the author, The Pennsylvania State University, College of Human Development, Amy Gardner House, University Park, PA 16802.

cerning group differences at a specific point along the predictor continuum. However, it does not follow that one can state with 95% confidence that there exist group criterion differences at all points within the regions. Potthoff (1964) described a modification of the Johnson-Neyman technique for establishing simultaneous regions of significance. His modification can be easily implemented with the present program by inputting a critical F of $2F(2, N-4)$.

Finally, the program is restricted to the analysis of ATI data where there are only two groups and one continuous predictor. This is not a limitation of the technique, because extensions to more complex independent-groups designs are available (Abelson, 1953). The author is presently working to develop BASIC programs that will treat data from more complex ATI designs.

REFERENCES

- ABELSON, R. P. (1953). A note on the Neyman-Johnson technique. *Psychometrika*, **18**, 213-218.
- BORICH, G. D., GODBOUT, R. C., & WUNDERLICH, K. W. (1976). *The analysis of aptitude-treatment interactions: Computer programs and calculations*. Chicago: International Educational Services.
- CRONBACH, L. J. (1957). The two disciplines in scientific psychology. *American Psychologist*, **12**, 671-684.
- HUITEMA, B. E. (1980). *Analysis of covariance and alternatives*. New York: Wiley.
- JOHNSON, P. O., & NEYMAN, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs*, **1**, 57-93.
- PEDHAZUR, E. J. (1982). *Multiple regression in behavioral research* (2nd ed.). New York: Holt, Rinehart & Winston.
- POTTHOFF, R. F. (1964). On the Johnson-Neyman technique and some extensions thereof. *Psychometrika*, **29**, 241-256.
- ROGOSA, D. (1980). Comparing non-parallel regression lines. *Psychological Review*, **88**, 307-321.

APPENDIX

Listing for Program to Determine Regions of Significance Using the Johnson-Neyman Technique

```

100 REM *****
110 REM JOHNSON-NEYMAN TECHNIQUE PROGRAM
120 REM *****
130 REM
140 REM THIS PROGRAM TAKES USER INPUT
150 REM AND CALCULATES POINT OF INTERSECTION
160 REM FOR 2 REGRESSION LINES.
170 REM THEN DETERMINES REGIONS OF SIGNIFICANCE
180 REM USING JOHNSON-NEYMAN TECHNIQUE
190 REM IBM AND APPLE USERS SHOULD
200 REM ADJUST CODE AS INDICATED IN
210 REM SUBROUTINES 2000, AND 2200
220 REM *****
230 GOSUB 2000
240 REM SUBROUTINE AT 2000 ALLOWS IBM OR APPLE SCREEN CLEAR
250 PRINT " ";
260 PRINT "*****"
270 PRINT " ";
280 PRINT " *";
290 PRINT " ";
300 PRINT " * JOHNSON-NEYMAN *";
310 PRINT " ";
320 PRINT " * TECHNIQUE *";
330 PRINT " ";
340 PRINT " *";
350 PRINT " ";
360 PRINT "*****"
370 FOR I = 1 TO 3000: NEXT I: REM PAUSE TO READ FRAME
380 GOSUB 2000
390 PRINT "THIS PROGRAM ASKS YOU FOR "
400 PRINT "INFORMATION CONTAINED ON THE "
410 PRINT "OUTPUT FROM REGRESSION ANALYSES"
420 PRINT "IN MOST STATISTICAL PACKAGES"
430 PRINT "(E.G., SPSSX,SAS,BMDP), AND"
440 PRINT "DETERMINES THE POINT OF INTERSECTION"
450 PRINT "AND REGIONS OF SIGNIFICANCE USING"
460 PRINT "THE JOHNSON-NEYMAN TECHNIQUE."
470 PRINT
480 PRINT "THE PROGRAM IS RESTRICTED TO THE"
490 PRINT "TWO-GROUP, SIMPLE REGRESSION CASE."
500 PRINT "IN ORDER TO USE THE PROGRAM"
510 PRINT "YOU NEED TO PROVIDE INFORMATION"
520 PRINT "THAT CAN BE MOST EASILY OBTAINED"
530 PRINT "BY CONDUCTING SEPARATE SIMPLE"
540 PRINT "REGRESSION ANALYSES FOR BOTH GROUPS."
550 PRINT
560 PRINT "PRESS ANY KEY TO CONTINUE."
570 GOSUB 2200: REM IBM OR APPLE KEYSTROKE CHECK
580 GOSUB 620: REM INPUT DATA ROUTINE
590 GOSUB 1420: REM CALCULATIONS ROUTINE
600 GOSUB 1620: REM OUTPUT ROUTINE
610 END

```

APPENDIX (Continued)

```

620 REM *****
630 REM INPUT DATA ROUTINE
640 REM *****
650 GOSUB 2000
660 PRINT " ";
670 PRINT "*****"
680 PRINT " ";
690 PRINT "* FOR GROUP 1  *"
700 PRINT " ";
710 PRINT "*****"
720 PRINT
730 INPUT "WHAT IS THE SAMPLE SIZE? ";N1
740 IF N1 < = 0 THEN PRINT "ILLEGAL VALUE!": GOTO 730
750 INPUT "WHAT ARE THE RESIDUAL SUMS OF SQUARES? ";S1
760 IF S1 < = 0 THEN PRINT "ILLEGAL VALUE!": GOTO 750
770 INPUT "WHAT IS THE PREDICTOR MEAN? ";M1
780 INPUT "WHAT IS THE PREDICTOR STANDARD DEVIATION? ";V1
790 IF V1 < = 0 THEN PRINT "ILLEGAL VALUE!": GOTO 780
800 INPUT "WHAT IS THE REGRESSION WEIGHT(UNSTANDARDIZED)? ";B1
810 INPUT "WHAT IS THE INTERCEPT? ";I1
820 GOSUB 2000
830 PRINT "INPUT DATA FOR GROUP 1"
840 PRINT
850 PRINT "SAMPLE SIZE = ";N1
860 PRINT "RESID. SUMS OF SQUARES = ";S1
870 PRINT "PREDICTOR MEAN = ";M1
880 PRINT "PREDICTOR S.D. = ";V1
890 PRINT "REGRESSION WEIGHT = ";B1
900 PRINT "INTERCEPT = ";I1
910 PRINT : PRINT
920 PRINT "IS THIS INFORMATION CORRECT?"
930 PRINT "ENTER 'YES OR 'N'O"
940 PRINT
950 GOSUB 2200
960 IF A$ = "N" OR A$ = "n" THEN PRINT "TRY AGAIN": FOR I = 1 TO 3000:
NEXT I: GOTO 650
970 IF A$ < > "Y" OR A$ < > "y" THEN GOTO 940
1000 GOSUB 2000
1010 PRINT " ";
1020 PRINT "*****"
1030 PRINT " ";
1040 PRINT "* FOR GROUP 2  *"
1050 PRINT " ";
1060 PRINT "*****"
1070 PRINT
1080 INPUT "WHAT IS THE SAMPLE SIZE? ";N2
1090 IF N2 < = 0 THEN PRINT "ILLEGAL VALUE!": GOTO 1080
1100 INPUT "WHAT ARE THE RESIDUAL SUMS OF SQUARES? ";S2
1110 IF S2 < = 0 THEN PRINT "ILLEGAL VALUE!": GOTO 1100
1120 INPUT "WHAT IS THE PREDICTOR MEAN? ";M2
1130 INPUT "WHAT IS THE PREDICTOR STANDARD DEVIATION? ";V2
1140 IF V2 < = 0 THEN PRINT "ILLEGAL VALUE!": GOTO 1130
1150 INPUT "WHAT IS THE REGRESSION WEIGHT? ";B2
1160 INPUT "WHAT IS THE INTERCEPT? ";I2
1170 GOSUB 2000
1180 PRINT "INPUT DATA FOR GROUP 2"
1190 PRINT
1200 PRINT "SAMPLE SIZE = ";N2
1210 PRINT "RESID. SUMS OF SQUARES = ";S2
1220 PRINT "PREDICTOR MEAN = ";M2
1230 PRINT "PREDICTOR S.D. = ";V2
1240 PRINT "REGRESSION WEIGHT = ";B2
1250 PRINT "INTERCEPT = ";I2
1260 PRINT
1270 PRINT
1280 PRINT "IS THIS INFORMATION CORRECT?"
1290 PRINT "TYPE 'YES OR 'N'O"
1300 PRINT
1310 GOSUB 2200
1320 IF A$ = "N" OR A$ = "n" THEN PRINT "TRY AGAIN": FOR I = 1 TO 3000:
NEXT I: GOTO 1000
1340 IF A$ < > "Y" OR A$ < > "y" THEN GOTO 1300
1360 GOSUB 2000
1370 PRINT "FINALLY, WHAT IS THE CRITICAL F"
1380 PRINT "FOR ALPHA=.05 AND D.F.="
1390 INPUT "1 AND TOTAL N-4? ";F
1400 IF F < = 0 THEN PRINT "ILLEGAL VALUE!": GOTO 1390
1410 RETURN
1420 REM *****
1430 REM CALCULATIONS ROUTINE
1440 REM *****
1450 CROSS = (I1 - I2) / (B2 - B1): REM POINT OF INTERSECTION
1460 SS = S1 + S2: REM POOLED SUMS OF SQUARES RESIDUAL
1470 N = (N1 + N2): REM TOTAL N
1480 V1 = ((V1 ^ 2) * (N1 - 1)): V2 = ((V2 ^ 2) * (N2 - 1)): REM PREDICTOR SUMS
OF SQUARES

```

APPENDIX (Continued)

```

1490 A = (1 / V1) + (1 / V2)
1500 A = A * SS * (- F / (N - 4))
1510 A = A + ((B1 - B2) ^ 2)
1520 B = ((M1 / V1) + (M2 / V2))
1530 B = B * SS * (F / (N - 4))
1540 B = B + ((I1 - I2) * (B1 - B2))
1550 C = ((N / (N1 * N2)) + ((M1 ^ 2) / V1) + ((M2 ^ 2) / V2))
1560 C = C * SS * (- F / (N - 4))
1570 C = C + ((I1 - I2) ^ 2)
1580 Z = (SQR ((B ^ 2) - (A * C)))
1590 Z1ANS = (- B + Z) / A
1600 Z2ANS = (- B - Z) / A
1610 RETURN
1620 REM *****
1630 REM OUTPUT ROUTINE
1640 REM *****
1650 GOSUB 2000
1660 PRINT "THE POINT OF INTERSECTION FOR"
1670 PRINT "THE TWO FUNCTIONS = ";CROSS
1680 PRINT : PRINT
1690 PRINT "THE GROUPS ARE SIGNIFICANTLY DIFFERENT"
1700 PRINT "ABOVE ";Z1ANS;" AND BELOW ";Z2ANS
1710 RETURN
2000 REM APPLE USERS DELETE LINE 2030
2010 REM IBM USERS DELETE LINE 2020
2020 HOME
2030 CLS
2040 RETURN
2200 REM APPLE USERS DELETE LINE 2230
2210 REM IBM USERS DELETE LINE 2220
2220 GET A$: IF A$ = "" THEN GOTO 2220
2230 INKEY$: IF A$ = "" THEN GOTO 2230
2240 RETURN

```

(Revision accepted for publication October 2, 1986.)