— PROGRAM ABSTRACTS/ALGORITHMS —

The Rev.: An IBM BASIC program for Bayesian test interpretation

RONALD D. FRANKLIN and DAVID B. ALLISON
*The Johns Hopkins School of Medicine
and The Kennedy Institute, Baltimore, Maryland*

The Rev. is a BASIC computer program for IBM-PC-compatible systems that provides Bayesian estimates of "true" scores from multiple scores measuring the same construct. Psychological reports often include test scores from earlier evaluations without objectively incorporating them into current findings. Using Thorndike's formulas for objectively combining test scores while providing estimates with reduced standard errors, the Rev. is an interactive program that facilitates test interpretation by combining information from many test administrations. The user provides four easily obtainable pieces of information for each test administration. The output includes an estimated "true" score and the standard error of the estimate.

The crux of Bayesian thinking is that people have a priori subjective or objective probability distributions for the likelihood of some condition being met. New information can be added later, producing a new, and presumably more accurate, posterior probability distribution. It should be noted that neither the new nor the old information is discarded in favor of the other. Rather, they combine, producing a better probability estimate than either could provide alone (Iverson, 1984).

Bayesian statistics provide an alternative measure of probability. Rather than asking the traditional statistical question, "If 60% of children with Johnny's IQ are retarded, what is the probability that Johnny is retarded?", Bayesian methods ask "If Johnny scores in the retarded range on an IQ test, what is the probability he is retarded?". Stated formally:

$$P(H|E) = [P(H) \times P(E|H)] / P(E),$$

where $P(H|E)$ = the posterior probability of the suspected result given the specified event, $P(H)$ = the outcome probability when no a priori knowledge of the event exists, $P(E|H)$ = the probability of the event within the outcome group, and $P(E)$ = probability of the event's occurrence in the base population (after Thorndike, 1986c, p. 54).

R. D. Franklin is now at South Florida State Hospital, Pembroke Pines, FL. D. B. Allison is at St. Luke's/Roosevelt Hospital, New York, NY. Requests for reprints and copies of the software should be sent to R. D. Franklin, 7958 Pines Boulevard, Suite 136, Pembroke Pines, FL 33024.

Thorndike (1986a, 1986b) extended this thinking to the interpretation of standardized tests. In an assessment situation, he proposed combining a priori with current information to form an estimate of the subject's "true" score on some variable. For example, if, when assessing a child's intelligence, knowledge of the child's score on an intelligence test administered 2 years earlier and on an achievement test administered earlier that day, can both be used as estimates of the child's "true" score on an intelligence test today?

Disregarding earlier test scores when administering later tests not only wastes information—described by Cohen (1990) as a psychologist's most precious commodity—but may ignore the more valid measure. Many psychologists subjectively combine such information from different sources, but Thorndike (1986c) provides formulas for objectively combining test scores.

Thorndike's (1986c) formulas for estimates are represented as:

"True" Score

$$= \frac{[(X_1/(SE_{\text{meas}})^2) + (X_2/(SE_{1 \text{ est}})^2) \dots + (X_n/(SE_{n \text{ est}})^2)]}{[(1/(SE_{\text{meas}})^2) + (1/(SE_{1 \text{ est}})^2) \dots + (1/(SE_{n \text{ est}})^2)]}$$

and

"True" Score SE

$$= \frac{1}{[(1/(SE_{\text{meas}})^2) + (1/(SE_{1 \text{ est}})^2) \dots + (1/(SE_{n \text{ est}})^2)]},$$

where X_1 is the original, or primary, measure of the construct and X_2 to X_n are secondary measures of the construct. SE_{meas} is the standard error of measurement defined as $SD_1 ((1-r_{11})^{-.5})$, and SE_{est} is the standard error of the estimate defined as $SD_n ((1-r_{1n})^{-.5})$; where SD_n is the standard deviation of the n th test, r_{11} is the reliability coefficient (preferable internal consistency) of the primary test, and r_{1n} is the correlation between the primary test and the n th test (the meaning of a primary test is given below). In essence, a weighted average of scores is computed, where each score is weighted by the reciprocal of its error estimate. In addition to providing a better estimate than any single score, this method can also provide estimates with substantially reduced standard errors.

The Rev. is named for Reverend Thomas Bayes, who developed the conditional probability statements that formed the foundations of Bayesian statistics. It is an interactive program that makes Thorndike's (1986c) formulas easy to apply in test interpretation. The Rev. can combine information from a large number of test administrations. The user need only provide four pieces of infor-

mation for each test administration. Except for the subject's test score, other required information is generally available in most test manuals and includes: (1) the mean score from the standardization sample, (2) the standard deviation from the standardization sample, and (3) either the reliability (internal consistency) of a primary measure of the desired construct or the correlation of the test with a primary measure of the desired construct. By primary measure, two things are implied: (1) the test is generally accepted to be a measure of some construct, and (2) the test is given at the approximate point in time that one wishes to measure the construct. (See Appendix.)

Program Description

The Rev. was written in unmaximized generic BASIC. It was interpreted and debugged on a Zenith 286 system and operates under most interpreters supporting 8080, 8086, 80286, or 80386 architectures. It requires 2.2 KB or more of RAM and runs under MS-DOS Version 2 or higher. Output is routed to the console or a standard LPT configuration. The report includes an estimated "true" score and the standard error of this estimate.

Program Availability

The program source code is available free of charge by sending a blank formatted diskette along with a return postage-paid mailer to the senior author at 7958 Pines Boulevard, Suite 136, Pembroke Pines, FL 33024; 305-432-2157.

REFERENCES

- COHEN, J. (1990). Some things I have learned (so far). *American Psychologist*, 45, 1304-1312.
- IVERSON, G. R. (1984). *Bayesian statistical inference*. Beverly Hills: Sage.
- THORNDIKE, R. L. (1986a). Bayesian concepts and test making. *Journal of Counseling & Development*, 65, 110-111.
- THORNDIKE, R. L. (1986b). Bayesian concepts in test interpretation. *Journal of Counseling and Development*, 65, 170-172.
- THORNDIKE, R. L. (1986c). The role of Bayesian concepts in test development and test interpretation. *Journal of Counseling & Development*, 65, 54-65.

APPENDIX

Program Listing

```

1 REM      The REV
2 REM      Bayesian predictor of true test scores
3 REM      Ronald D. Franklin, Ph.D.
4 REM      David B. Allison, Ph.D.
5 REM      The Kennedy Institute and
6 REM      Johns Hopkins School of Medicine
7 REM      Send inquiries to R.D. Franklin
8 REM      7958 Pines Blvd., Suite 136
9 REM      Pembroke Pines, FL 33024 (305) 431-2158
10 :
```

APPENDIX (continued)

```

101 REM A SETS THE DEFAULT ARRAY SUBSCRIPT
190 CLS
201 GOSUB 1000
210 LET A% = NT%
300 DIM NS%(A%), TS%(A%), TM%(A%), TSD%(A%), R(A%)
400 FOR A% = 1 TO NT%
410 LET NS%(A%) = ""
411 LET TS%(A%) = 0
412 LET TM%(A%) = 0
413 LET TSD%(A%) = 0
414 LET R(A%) = 0
415 CLS
419 GOSUB 1100
420 GOSUB 10000
550 NEXT A%
600 GOSUB 20000
700 GOSUB 50500
998 END

999 :
1000 PRINT "Enter the name of the test subject"
1010 INPUT TNS
1020 PRINT "How many tests will you compare "
1030 INPUT NT%
1031 CLS
1040 RETURN

1050 :
1100 PRINT "Enter the name of test number "; A%
1110 INPUT NS(A%)
1200 PRINT "Enter the examinee's Standard Score for the "; NS(A%)
1210 INPUT TS(A%)
1300 PRINT "Enter the Publisher's Mean score for the "; NS(A%)
1310 INPUT TM(A%)
1400 PRINT "Enter the Publisher's Standard Deviation for the "; NS(A%)
1410 INPUT TSD(A%)
1500 IF A%=1 THEN PRINT "Enter the Publisher's reliability for the test"
1510 IF A% > 1 THEN PRINT "Enter the Publisher's correlational coefficient between "; NS(A%); " and "; NS(1)
1520 INPUT R(A%)
1600 RETURN

1700 :
10000 STD(A%) = (((TS%(A%) - TM%(A%)) / TSD%(A%)) * TSD%(1)) + TM%(1))
10100 SE(1) = (TSD%(1)) * (SQR((1 - (R(1))^2)))
10150 IF A% > 1 THEN LET SE(A%) = TSD%(A%) * SQR(1 - R(A%))
10200 RETURN

10300 :
20000 LET PEST = 0
20010 LET TEST = 0
20050 FOR K% = 1 TO NT%
20100 PEST = PEST + (1 / (SE(K%)^2))
20200 TEST = TEST + (STD(K%) / (SE(K%)^2))
20300 NEXT K%
40000 LET EST = TEST / PEST
50100 RETURN

50200 :
50500 CLS
50600 PRINT "The REV":PRINT:PRINT:PRINT
50700 PRINT "Estimate of True test scores for "; TNS:PRINT:PRINT
50800 PRINT "Test Name      Score      Mean      SD"
50900 PRINT "Coefficient      Error"
51000 PRINT "-----"
51100 PRINT "-----"
51200 FOR K% = 1 TO NT%
51300 PRINT NS(K%);TAB(15);TS%(K%);TAB(25);TM%(K%);TAB(35);TSD%(K%);TAB(45);R(K%);TAB(55);SE(K%)
51400 PRINT
51500 NEXT K%
51700 PRINT "The estimate of the true score is ";EST
51800 PRINT "The standard error of the true score estimate is ";TSE
51900 PRINT "NOTE: Predictions are based on comparison with ";NS(1);"."
52000 RETURN
```