METASTATS:
Behavioral science statistics for Microsoft Windows and the HP49G programmable calculator

ERIC K. FOSTER
Temple University, Philadelphia, Pennsylvania

METASTATS is an organized suite of approximately 150 programs designed for meta-analysis, power, effect sizes and their confidence intervals, significance test conversions, and other general statistics. This freeware works either on the HP49G calculator or preinstalled on a PC emulator under Windows. Many of the programs address gaps in commercially available software. The suite has a user-friendly interface: the screen prompts for input, and labels all output. Extensive documentation includes step-by-step examples, a fully annotated catalogue of the programs, and academic references for the formulas. Both versions of the programs (calculator and PC), documentation, and installation instructions are freely available from the author's Web site, http://www.netaxs.com/~efoster/.

Recently, the Task Force on Statistical Inference (TFSI), a committee of the American Psychological Association's (APA) Board of Scientific Affairs, announced recommendations for better data analysis and reporting (Wilkinson & Task Force on Statistical Inference, 1999). Regular reporting of power analyses, effect sizes, and confidence intervals were emphasized. Minimal analyses were sanctioned because they can "provide elegant and sufficient answers to important questions" (Wilkinson et al., 1999, p. 598).

A new suite of programs—METASTATS—comprises a convenient technology to serve the TFSI's recommendations and statistical instruction in general. For quick results with relatively small amounts of data and often used formulas, the suite is practical, powerful, and accurate. It is designed to work on either the HP49G graphing calculator or a PC emulator of the calculator (see Figure 1), and it is downloadable as freeware from my Web site. Calculator owners already have the PC-to-calc transfer software and serial cable; for those who do not own the calculator, the emulator runs under Windows (9x, ME, NT, 2000, and XP) but not Macintosh. Full instructions for download and installation are included in the extensive METASTATS documentation.

Applications for METASTATS

Two situations, common in psychological research, generate statistical applications for METASTATS. Relatively small *N*s are practical input. At the graduate and professional level in most social science fields, many studies involve samples under 30 observations. Also, in many statistics courses, textbook and question writers use small sample sizes for demonstration purposes. Power analysis, effect sizes, correlations and partial correlations, confidence intervals, regression equations, and probabilities are conveniently produced by METASTATS for small *N*s.

The other ideal application for METASTATS regards working with other people's data (Rosnow & Rosenthal, 1996). Calculating and scribbling power, effect sizes, and confidence intervals in the margins of articles can be quite helpful in getting a sense of the meaning and importance of research results. Further, statistical mistakes in publications are not uncommon (Wilkinson et al., 1999), and on-the-spot recomputing can often ferret them out. Instructive exercises of this nature could be incorporated into any statistics course, or any course that requires students to examine published data and statistical results.

Meta-analysts particularly can benefit from METASTATS, because it contains a panoply of conversion, combining, and polynomial formulas for use in this area. METASTATS readily provides precise *p* values, along with effect size calculations, conversions, and confidence intervals—in many instances as simultaneous output.

Some Examples

To illustrate METASTATS operation, here is a common statistical formula in the social sciences, the *t* test for comparing two groups:

$$t = \frac{M_1 - M_2}{\sqrt{\frac{SD_1^2 * (n_1 - 1) + SD_2^2 * (n_2 - 1)}{n_1 + n_2 - 2} * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}},$$

where *M* is mean, *SD* is standard deviation, and *n* is group sample size. This formula is nested in a METASTATS program that calculates the *t* value. Each discrete value is entered only once, regardless of the number of times it appears in the formula, and no operators or parentheses are

Correspondence concerning this article should be addressed to E. K. Foster, 191 Presidential Blvd. PH6, Bala, PA 19004 (e-mail: efoster@netaxs.com).
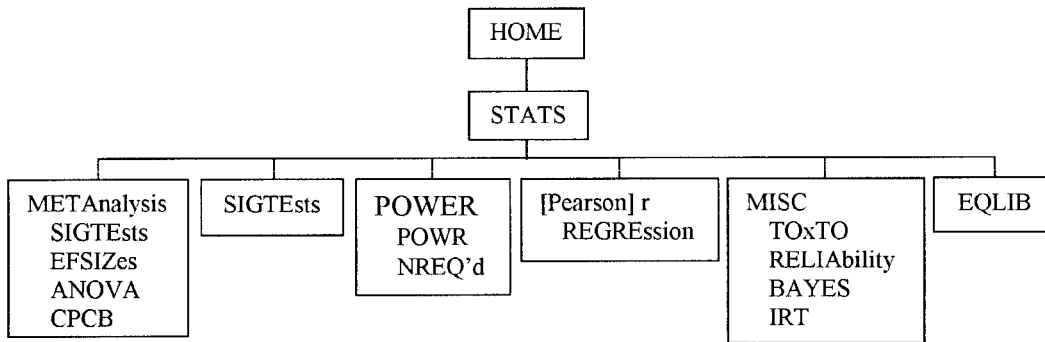
Figure 1. Screen shot of METASTATS on the PC Emulator version.

required. The program plugs everything in and performs the logic automatically. Input and output screens for this test in METASTATS are shown in Figures 2 and 3.

Note in these figures that a descriptor line guides the user with each input and that output lines are labeled. The p value and an effect size for the test statistic are also returned in this case. The opportunity for error is dramatically less than it would be with the use of a conventional calculator, and the speed of data input is much greater.

Like the t test above, most METASTATS programs return multiple related results. Consider the 2×2 contingency table of counts, frequently encountered in psychology research and medical trials. The formula for the effect size (phi, ϕ) from a 2×2 is

$$\phi = \frac{AD - BC}{\sqrt{(A+B)(C+D)(A+C)(B+D)}},$$

where the letters refer to each quadrant of the 2×2. With 14 total keystrokes (assuming two-digit counts in each cell), METASTATS returns ϕ along with seven additional statistics: the odds ratio, relative risk, risk difference, and χ^2 and Z with their associated p values. For descriptives of any list of numbers, another program returns the sum, mean, variance, standard deviation, total sum of squares, and sample variance. Another program produces the correlation between two lists, along with the 90%, 95%, and 99% confidence intervals. A regression-related program, with three correlations from the matrix, outputs R^2, β_1, β_2,

partial correlation, semipartial correlation, and a test for the suppression effect.

Locating programs in memory is facilitated by an orderly DOS-like hierarchy of directories (see Figure 4) and

Figure 2. Input screen for t test of independent means using summary data.

Figure 3. Output screen for t test of independent means.

Figure 4. Directory tree structure of METASTATS.

thorough documentation that lists all the programs detailing their inputs and outputs. METASTATS also includes a complete library of formulas for retrieval in a blackboard-like display, should one wish to view them. An example of an equation viewed in this fashion is shown in Figure 5.

Each directory has its own scrollable menu for the programs contained within it. Figure 6 shows the menu for the meta-analysis directory "Compare and Combine [r effect sizes, Z scores, or p values]" (CPCB in Figure 4). Note the "smart scroll" bar to the right.

Operations on lists of numbers are easily performed in the meta-analytic routines. Any number of rs, Zs, or ps and their weights (including polynomial weights) may be compared and combined. Figures 7 and 8 show input and output screens, respectively, for combining and weighting eight Pearson product–moment correlations.

Several dozen METASTATS programs convert various standard effect sizes and significance tests. Power programs, both "power obtained" and "N required," are included for tests of proportions and for independent and dependent ts and rs, paralleling the tables of Cohen's (1988) power analysis text. Interpolations are unnecessary: METASTATS calculates exact power, including adjustments for unequal sample sizes. Four conventional significance levels (.005, .01, .025, .05) and four power levels (.5, .6, .7, .8) are specifiable.

Accuracy

I have extensively tested all routines for accuracy. All programming results were compared with published examples, either in top-tier journals or seminal texts. For instance, dozens of power analyses were performed and the results compared with the tables in Cohen's (1988) book. No result diverged more than one point from these tables. The p values match *Biometrika Tables for Statisticians* (Pearson & Hartley, 1966) up to 10 decimals. Converting Pearson r to Fisher's Z_r matches *Biometrika*'s four decimal tables in every instance, including extreme values to .998. Many meta-analysis problems were reproduced from published examples with perfect precision. Virtually all programs were thus benchmarked. Users may be advised to perform similar comparisons with published examples for the type of work that they do.

The only known accuracy problem in the suite occurs with the single program for the estimation of the critical t-test statistic when one is entering very small degrees of freedom (about 7 or below). In this case, a warning to this effect appears in the program when the user executes it. With df of 13, accuracy is good to two decimals; at 20 and above, to three decimals or better. In general, however, interim algebraic calculations, as in an electronic spreadsheet, are performed by the calculator well beyond any practical concern, to 15 decimal places.

Device Information

The HP49G and its emulator for the PC use 2.5 MB of total memory. This is miniscule, of course, by PC standards, but quite large by calculator standards. METASTATS, one of the larger suites available for the HP49G, requires only about 130K of the 1.5 MB available to the user. An additional 25K is required for certain utilities. Thus, there is ample room for more programs and archiving. META-

Figure 5. An example from the equation library.
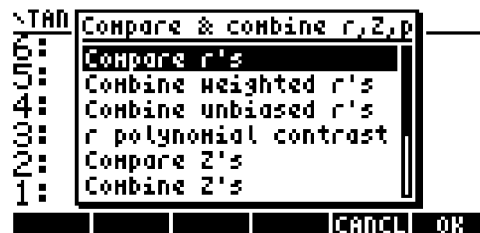
Figure 6. Menu for the meta-analysis directory.

Figure 7. Combining and weighting *r* effect sizes.

Figure 8. Output of combined weighted *r* effect sizes.

STATS includes a user routine that compresses the programs and dates a backup object so that multiple backups may be archived (and restored) in the event of accidental deletion or corruption (rare). There is no noticeable performance difference with METASTATS and its utilities loaded. Backups survive complete battery charge loss because they are stored as FLASHROM, as is system software, which makes it upgradeable and secure. Approximately 1 MB is available for the operating system; the latest one,

which METASTATS and the PC emulator require, uses about one third of the space available to it.

Conclusion

Whether for applied or for pedagogical use, METASTATS is a useful statistical package to complement commercial software. It is a helpful companion when one is reading and interpreting scientific literature. Because the suite reduces wearisome computation and look-up tasks, it can also help creative users conceptualize statistical concepts and interrelationships by enabling easily generated "what if" scenarios.

METASTATS may be appended or customized if the user learns HP programming; additional programs and directories are simply added to the existing ones. Hundreds of other kinds of programs and utilities are also freely available for download on the Web. (To my knowledge, this is the only one designed specifically for behavioral research statistics.)

METASTATS is downloadable, in either the HP calculator or the PC emulator versions, with installation instructions and extensive documentation, from http://www.netaxs.com/~efoster/.

REFERENCES

COHEN, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

PEARSON, E. S., & HARTLEY, H. O. (EDS.) (1966). *Biometrika tables for statisticians* (3rd ed., Vol. 1). Cambridge: Cambridge University Press.

ROSNOW, R. L., & ROSENTHAL, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Psychological Methods*, **1**, 331-340.

WILKINSON, L., & TASK FORCE ON STATISTICAL INFERENCE (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, **54**, 594-604.