A REVIEW OF STATISTICAL POWER ANALYSIS SOFTWARE

Len Thomas

Centre for Applied Conservation Biology, Faculty of Forestry, University of British Columbia, #270-2357 Main Mall, Vancouver, British Columbia, V6T 1Z4, Canada. lthomas@unixg.ubc.ca

Charles J. Krebs

Department of Zoology, University of British Columbia, 6270 University Blvd, Vancouver, British Columbia, V6T 1Z4, Canada.   krebs@zoology.ubc.ca

Although ecologists have become increasingly sophisticated in applying tests for statistical significance, few are aware of the power of these tests.   Statistical power is the probability of getting a statistically significant result given that there is a biologically real effect in the population being studied.   If a particular test is not statistically significant, is it because there is no effect or because the study design makes it unlikely that a biologically real effect would be detected?   Power analysis can distinguish between these alternatives, and is therefore a critical component of designing experiments and testing results   (Toft and Shea 1983, Rotenberry and Weins 1985, Peterman 1990, Fairweather 1991, Taylor and Gerrodette 1993, Thomas and Juanes 1996).

Discussions with colleagues suggest that one major obstacle to the use of power analysis is the perceived lack of computer software.   Our paper is designed to remove this obstacle, by reviewing numerous packages that calculate power or sample size and determining which are likely to be of use to researchers (see also Goldstein 1989).   We do not deal with software for planning the precision of studies, although this is an important consideration if the results are to be analyzed using confidence intervals rather than *p*-values (Greenland 1988, Krebs 1989).   Because of software updates and new program releases, our program evaluations will quickly become out-of-date.   Nevertheless, we highlight some important and timeless issues to consider when evaluating power software: scope, flexibility, accuracy, ease of use, and ability to deal with prospective and retrospective analyses.

Readers interested only in "the bottom line" should skip straight to the *Conclusions* by way of the summary evaluation in Table 2.   Those less experienced with power analysis will find the *Background*, *Stand-alone power and sample size software*, and *Discussion* sections of most use.   Experienced power analysts will be most interested in the section on *Programming power analysis using general-purpose statistical software* and Table 4.

BACKGROUND

The concepts of statistical power are covered in detail in a number of texts (Kraemer and Thiemann 1987, Cohen 1988, Lipsey 1990; see also a particularly clear paper by Muller and Benignus 1992).   Briefly, the power of a test is the probability of rejecting the null hypothesis given that the alternative hypothesis is true.   Power depends on the type of test, increases with increasing sample size, effect size, and higher α-level, and declines with increasing sampling variance.   Effect size is the difference between the null and alternative hypotheses, and can be measured either using raw or standardized values.   Raw measures, such as the slope in a regression analysis or difference between means in a *t*-test, are closer to the measurements that researchers take and so are easier to visualize and interpret.   Standardized measures, such as the correlation coefficient or *d*-value (difference in means divided by the standard deviation), are dimensionless and incorporate the sampling variance implicitly, removing the need to specify variance when calculating power.

Power analysis is most useful when planning a study.   Such "prospective" power analyses are usually exploratory in nature, investigating the relationship between the range of sample sizes that are deemed feasible, effect sizes thought to be biologically important, levels of variance that could exist in the population (usually taken from the literature or from pilot data), and desired levels of α and statistical power.   The result is a decision about the sample size and α-level that will be used in the study, and the target effect size that will be "detectable" with the given level of statistical power.

After the study is completed and the results analyzed, a "retrospective" power analysis can also be useful if a statistically non-significant result was obtained (e.g., Thomas and Juanes 1996).   Here the actual sample size and α-level are known, and the variance observed in the sample provides an estimate of the variance in the population.   These values are used to calculate power at the minimum effect size thought to be of biological significance, or alternatively the effect size detectable with the minimum desired level of power.   Note that it is rarely useful to calculate power using the effect size observed in the sample: such analyses tell us nothing about the ability of the test to detect biologically important results (Thomas 1997).

Power calculations can be done using the tables or charts provided in many articles and texts (e.g., Kraemer and Thiemann 1987, Cohen 1988, Lipsey 1990, Zar 1996).   However these often require some hand calculations before they can be used, including interpolation between tabled values, and can give inaccurate results in some situations (e.g., see Bradley et al. 1996 regarding the accuracy of Cohen's ANOVA tables).   Computer software has the potential to make power analysis more accurate, interactive and easy to perform.

Ideally, power analysis should be integrated within the general purpose statistical software researchers use for their regular analyses.   This means having a comprehensive "study planning" module for prospective power analysis, and options to produce estimates of

retrospective power or detectable effect size for each test performed.   Failing this, researchers could use stand-alone software designed solely for power analysis.   In either case an ideal program should: (1) cover the test situations most commonly encountered by researchers; (2) be flexible enough to deal with new or unusual situations; (3) produce accurate results; (4) calculate power, sample size, and detectable effect size; (5) allow easy exploration of multiple values of input parameters; (6) take a wide variety of effect size measures as input, both raw and standardized; (7) allow estimation of sampling variance from pilot data and from the sampling variability statistics commonly reported in the literature; (8) give easy-to-interpret output; (9) produce presentation quality tables and graphs for inclusion in reports; (10) allow easy transfer of results to other applications; and (11) be well documented.

The third point, on accuracy, requires further explanation.   For some tests, such as z-tests and those based on discrete distributions like the binomial, both the p-value and statistical power are calculated using the same distribution function.   Algorithms for computing these functions to are well known, which means that accurate power calculations are easy to perform using a computer.   Most t, F, and χ^2 tests, however, require calculation of a different distribution function for power than for significance tests, called a "non-central" distribution function. Efficient algorithms for computing non-central distribution functions have been developed only recently, and it was previously common to calculate power using approximate methods.   Some programs still use approximations, a topic we will return to in the discussion.

METHODS

We compiled a list of software capable of performing power or sample size calculations by searching the published literature and the internet, word-of-mouth, and requests to internet newsgroups.   From this preliminary list, we excluded a few specialized programs with limited scope that we considered unlikely to be of use to ecologists (CRCSIZ, EPIINFO, ECHIP, HYPERSTAT, SSIZE, R2; more information about these programs is available at our world-wide web site http://www.interchg.ubc.ca/cacb/power/).   We also excluded general-purpose statistical software where power capabilities are not built-in but can be programmed, although we briefly discuss these programs later in the paper.   For all remaining software, we contacted the vendors asking for a review copy, or downloaded the software if it was available over the internet.   We received replies from all vendors except three: Epicenter Software (program POWER), Lionheart Press (NONCDIST and SMPLSIZE), and Statistics and Epidemiology Research Corporation (EGRET).   Of those that replied, all agreed to send review copies, although SPSS Inc. (SPSS and SYSTAT DESIGN) failed to deliver anything.

We reviewed three aspects of each program: scope (points 1-3 above), ease of use (4-10), and ease of learning (11; also the intuitiveness of the program layout, and the clarity and information content of the on-line help).   For scope, we compiled a list of the test situations explicitly mentioned in the program documentation or on screen.   We also recorded whether the

program uses exact or approximate methods to calculate power.   We summarized ease of use and ease of learning on a subjective scale from excellent through very good, good, fair, and poor. Note that ease of use is more important than ease of learning for most users, because all the programs can be learnt in a day or two.

After our initial review, we selected the four most promising packages and asked a class of 19 graduate students to evaluate them as part of a two week "power module" (lectures and seminars) in an advanced ecology class.   Students in the class had a wide range of statistical expertise, but none had prior practical experience of power analysis.   During the module, students were asked to complete six power analysis problems using each of the four programs, and then to try a problem from their own research.   The six questions covered a t-test, 2-way ANOVA with contrasts, trend analysis (regression), comparison of proportions, non-parametric test (Wilcoxon) and survival analysis.   Some were framed in terms of study design, others as retrospective analyses.   As part of their report, students independently completed an evaluation form and ranked the packages on the basis of which they would recommend to their colleagues.

We sent a draft of this paper to the vendors or authors of the program asking them to inform us of factual inaccuracies and to tell us of planned updates to their software, but making it clear that we would not alter any of the opinions stated in the paper.   We received 15 replies and made minor changes to the text and to Tables 1 and 3.   The paper was also reviewed by five independent researchers (see acknowledgments).

Program reviews were carried out on a number of 486 and Pentium PCs running DOS, Microsoft Windows 3.1, Windows 95, IBM OS/2 v4, and on a Power Macintosh running system 7.5.3.   All programs ran well using the operating system for which they were intended (except for one, NQUERY ADVISOR, which crashed on a number of occasions).   All 32-bit Windows programs ran under Windows 95 without problems, but none of them worked with OS/2 v4.

COMPARISON OF PROGRAMS

We reviewed a total of 29 programs (Table 1): 13 stand-alone power and sample size programs, 11 general-purpose statistics packages with built-in power capabilities, two programs that deal only with determining required sample size, and three specialized power programs of interest to ecologists.   In the following sections, we compare the programs within each of these four categories, finishing with a section on programming power analysis using general-purpose statistical software.

Stand-alone power and sample size software

We begin this section by considering the five general purpose power analysis packages that we found the easiest to use (Table 2): GPOWER, PASS, POWER AND PRECISION, NQUERY ADVISOR, and STAT POWER (see Ellison 1997 for a more detailed review of NQUERY ADVISOR). All of these programs cover a wide range of tests (Table 3) and have intuitive interfaces with a

number of helpful utilities (described below).   All except GPOWER have well written, useful documentation.   GPOWER covers slightly fewer tests, has fewer utilities and minimal documentation, but has the advantage of being free.

All five programs have a similar interface, whereby the user picks a test situation from an initial window or menu and a form appears for the specified test.   In NQUERY ADVISOR this form is laid out like a spreadsheet, with rows for the input and output parameters, and columns for different scenarios.   The user-selected output parameter is updated interactively as the input fields are filled.   This works quite well, although the conventions for cutting, pasting, and moving around the spreadsheet differ from those usually used in Windows applications (such as MS EXCEL), and the maximum number of columns (20) limits the size of a sensitivity analysis. PASS takes a different approach, with one field for each input parameter; multiple values for that parameter are specified by separating them with commas or by specifying a range and step size (x to y by z).   When the "Run" button is pressed the results are presented in the "output" window as tables and/or graphs.   Again, there are some idiosyncrasies: e.g., the user must enter a value for the parameter to be calculated even though this value is ignored.   GPOWER and POWER AND PRECISION take a similar "fill-in-the-box" approach, except that the resulting power is displayed on the same screen.   Only one result can be displayed at once, although POWER AND PRECISION includes some capabilities for producing tables.   STAT POWER's forms are mostly of the spreadsheet variety (and can even accept formulae in the spreadsheet cells), although only one result can be displayed at once for the multi-way tests (2-way ANOVA and contingency tables).

All of the programs have utilities that allow input of various types of effect size measure. For example, for a 1-way fixed effect ANOVA design the user can enter either the hypothesized cell means (all programs), the upper and lower cell means and pattern of dispersion (POWER AND PRECISION), the variance of cell means (NQUERY ADVISOR, PASS, and POWER AND PRECISION), Cohen's standardized f-value (GPOWER, POWER AND PRECISION), the η^2- (i.e., r^2-) value (GPOWER, STAT POWER), or the non-centrality parameter (all except GPOWER).   NQUERY ADVISOR also has a tool to help estimate the sampling variance from the information commonly included in publications such as the range of values, coefficient of variation and percentiles. NQUERY ADVISOR and PASS have tools for calculating the standard deviation from pilot data, STAT POWER has functions for calculating pooled standard deviations and variances, and NQUERY ADVISOR and STAT POWER can calculate confidence limits about estimates of sampling variance (which can be used to place approximate confidence limits about power estimates in prospective analyses that use pilot data or in retrospective power analyses).

For graphics production, PASS is the most flexible, allowing plots of any variable against any other, grouped by a third variable.   STAT POWER can plot power or sample size against any other variable with grouping for most tests, while GPOWER can plot most variables against one another but with only one line per graph.   NQUERY ADVISOR and POWER AND PRECISION only

plot power against sample size, although a grouping variable can be included.   Generally, the Windows-based programs (NQUERY ADVISOR, PASS, and POWER AND PRECISION) produce better quality graphs for publication than the DOS programs (GPOWER and STAT POWER).   PASS and POWER AND PRECISION use the same graphics controller (by Bits per Second Ltd.) and are the easiest to configure (PASS in "interactive graphics" mode).   However, none of the graphics utilities are as flexible and interactive as those in many widely used Windows data management or statistics programs.

Not surprisingly, the Windows-based programs also provide generally better interfaces with other applications, such as word-processors or graphics packages.   In most cases, transferring program output is simply a matter of cutting and pasting.   The output window in PASS and the report window in POWER AND PRECISION can be saved as a Rich Text Files (RTF), making importing into word-processors particularly easy.   For users unsure how to present the results, both NQUERY ADVISOR and POWER AND PRECISION can produce grammatically correct statements for pasting into grant applications, etc., although the reports from POWER AND PRECISION are rather wordy.   The two DOS programs, GPOWER and STAT POWER, can produce simple log files that could be imported into other applications, although both would require some editing before they could be made into tables or used to generate graphs.

All of the packages have manuals that clearly explain how to use the software.   The manual for GPOWER (Erdfelder et al. 1996) is very brief, and relies heavily on the user being familiar with Cohen's (1988) treatment of power analysis[1].   The other manuals assume no prior knowledge and give basic introductions to power and sample size calculations, although the NQUERY ADVISOR manual only discusses study planning and is mainly directed towards medical scientists.   The User's Guides for PASS and STAT POWER give examples for each test available, although in PASS many of these are of inappropriate retrospective analyses where the observed effect size is used to calculate power (see Background).   On-line help is very good in NQUERY ADVISOR, PASS, and POWER AND PRECISION.   All have on-screen guide windows that are displayed alongside the input menus, as well as an extensive help system and on-line examples. In PASS the entire manual is on-line, although the formulae were unreadable on some of our machines.

The four packages we asked the graduate students to evaluate were NQUERY ADVISOR, PASS, STAT POWER and GPOWER.   We excluded POWER AND PRECISION because our review copy was a pre-release version with many of the features not fully implemented.   Seventeen out of 19 students rated PASS as the package they would recommend first to their colleagues (mean rank $\pm$ SD: 1.1 $\pm$ 0.3).   NQUERY ADVISOR was ranked second on average (2.1 $\pm$ 0.8), with STAT POWER third (3.1 $\pm$ 0.7), and GPOWER a close fourth (3.2 $\pm$ 0.7).

---

[1]   We reviewed the DOS version of GPOWER.   An extensive hypercard stack is apparently available for the Macintosh version, which makes it more helpful for beginners.   In addition, a world-wide web version of the documentation is in preparation (see Table 1 for web site address).

For the remainder of this section, we briefly describe the other stand alone power software we reviewed.   Five programs have a reasonably wide scope of analyses, but are not as easy to use as the packages discussed above.   PC-SIZE is a simple shareware package (actually two separate programs) that calculates power or sample size for a range of basic tests (Table 3), but has a relatively rudimentary user interface with no graphics and limited tables.   The POWPAL suite of freeware programs have a similar interface, and can convert between a number of different standardized measures of effect size as well as performing power analysis for generic t, F and χ^2 tests and correlations.   POWERPACK is unique among the stand-alone programs we reviewed because it is the only one that is command-driven and can be programmed.   This makes it very powerful, but at the same time significantly harder to learn than menu-driven or question-and-answer style programs (see O'Brien 1988 and Goldstein 1989 for more complete reviews).   POWER PLANT is a Windows-based freeware package that contains many interesting features (such as interactive power curve plotting) and has the potential to be very useful, but is currently in public beta-testing and difficult to use.   There is no manual[2] and the on-line help is partially operational.   In addition, only a few of the ANOVA options and the statistical calculator work.   STATISTICAL POWER ANALYSIS is the official companion to Cohen's (1988) book.   It offers all the tests described in the text (with their inherent inexactness, e.g., Bradley et al. 1996), as well as some excellent tools and simulations to help users understand the concepts of statistical power and the meaning of effect size measures used (see Goldstein 1989 for a longer review).

Lastly, three programs are more limited in scope, offering power analysis for one- and two- sample tests.   PS and STPLAN are both simple freeware programs aimed mostly at medical researchers.   N is a commercial program that has few tests built-in (two-sample t- tests, proportions and McNemar's test), but many examples in the manual of how other problems can be reformulated so as to be analyzed with the program (see Goldstein 1989 for a longer review).

General-purpose statistical software with built-in power analysis capabilities

We were pleased to discover that more than 10 regular statistics packages have power capabilities built-in, but were disappointed to see how limited these capabilities are.   In some cases the power capabilities are very poorly documented, in others the documentation is confusing (specifically regarding retrospective power tests).   Although it is possible to do a useful power analysis for some tests using these programs, most researchers will find it easier to use the stand-alone software discussed in the previous section.   We emphasize, however, that power analysis is only a small part of these packages' capabilities and that our comments are not relevant for any other aspects of the programs (see Ellison 1992 and subsequent updates in the Bulletin for a more complete review of some of these packages).

---

[2] The POWER PLANT package now includes a file containing the manual.

We first discuss the five user friendly, menu-driven programs that are intended mainly for users with beginner to intermediate levels of statistical expertise: JMP, NCSS, SIGMASTAT, STATISTICA and TRUE EPISTAT.   Of these, the first four are general purpose data analysis packages while TRUE EPISTAT is intended for medical scientists.   All have large user manuals with many examples.   None offer anything like the same scope for power analysis as they do for statistical hypothesis testing.   This means that a researcher relying solely on one of these packages could analyze a study without being able to plan it properly or check its power. SIGMASTAT offers the widest scope for planning power or sample size (Table 3), although the ANOVA option doesn't go beyond one-way.   JMP is the most comprehensive in this respect, offering power calculations, tables, and power vs. sample size graphs for almost any fixed effect ANOVA design[3] (but for nothing else).   TRUE EPISTAT can calculate approximate power and sample size for a limited assortment of tests and can produce graphs and tables of power vs. effect size.   In STATISTICA, power and sample size can be calculated for some one-sample tests using the Process Analysis module, and OC (power vs. sample size) curves can be produced.

Three of the programs (JMP, NCSS, and SIGMASTAT) have options for the user to calculate the power of a test after it is performed (retrospective power analysis; indeed this is the only type of power analysis NCSS will do).   Unfortunately, NCSS and SIGMASTAT will only calculate power using the observed effect size and in JMP this is the default, making the results uninformative (see Background).   JMP can calculate bias-corrected power and approximate confidence limits about the power estimate, but the confidence limits are usually too wide to be helpful (see Thomas 1997 for an example).   JMP also has options to compute the Least Significant Difference (LSD) and Least Significant Number (LSN), which are basically the raw effect size and sample size required for a power of approximately 0.5 and so are not very useful. It is, however, the only of the three to allow calculation of retrospective power at user-specified effect sizes and so can potentially be used to perform useful retrospective analyses.

We now discuss four command-driven programs, which are generally more appropriate for intermediate or advanced users.   DATASIM is an inexpensive simulation program with limited data analysis procedures, but has the most comprehensive power analysis capabilities of any program in this section (Table 3).   Unfortunately these are not documented anywhere except in the on-line help and there is no search capability[4].   The built-in commands ESP and ESN allow the construction of statistical tables of the type found in Cohen (1988) using Cohen's standardized effect size measures, except that these cover a wider range of situations and give exact results.   There are also power and sample size commands (POWER and NSAMP) that cover a wide range of tests but require specification of effect size in the form of the non-

---

[3] It is possible to make JMP calculate power for random effects in the ANOVA, but the results are incorrect.   We have alerted the vendor to this problem.

[4] The current manual for Datasim is out of date, and is due to be rewritten soon.   Some of the power analysis capabilities of Datasim are illustrated in Bradley et al. (1996).

centrality parameter (a function of the standardized effect size, the sample size, and sample design). MACANOVA is a freeware analysis package based on the language S, and has similar commands for fixed-effect *F*-tests (power, power2 and samplesize). The disadvantage of using the non-centrality parameter is that one must go through some intermediate calculations in order to translate it into an effect size measure that is interpretable in terms of the biology of the test situation. SHAZAM is an econometrics program, with a similar command for the non-central *F* distribution, although the manual shows how the INCOEF command can be used to calculate the non-centrality parameter using hypothesized cell means. Lastly, STATA is a powerful and extensible data analysis program that, surprisingly, has only one sample size and power command (sampsi) which is restricted power analyses of one- and two- sample comparisons of means and proportions using the *z*-distribution. All four of these programs are designed to be extended by programming extra procedures as necessary. In all four cases, a useful power analysis, including sensitivity analysis and presentation quality graphic output, is best achieved by writing short programs (see later section).

Two packages remain to be covered in this section. MSUSTAT is a simple command-line and question-and-answer style program intended mostly for introductory level biostatistics. It has a limited selection of power and sample size options for biostatistics (Table 3), some of which give only approximate results. WDIST is a shareware statistical probability calculator that includes a small but eclectic collection of hypothesis testing and power routines (Table 3). Neither of these programs are particularly easy to use for power analysis (Table 2).

Stand-alone sample size software

Programs that are restricted to calculating required sample size given the desired level of power cannot be used for retrospective power analysis, but they could be useful when planning experiments if they offer some functions not available in the power and sample size programs.

EX-SAMPLE is a comprehensive survey sample size determination program aimed mostly at social scientists (reviewed in more detail by Goldstein 1989). It is unique in being the only package we reviewed that explicitly incorporates factors such as the cost of collecting data, time constraints, multiple response measures and sample design (stratified, clustered, etc.) in determining sample size. Note, however, that some of the sample size calculations are based on so-called "classic" methods, which do not take account of the desired level of power and which we do not recommend.

POWER AND EFFECT calculates required sample size and various measures of effect size based on Cohen (1988). It has some interesting animated tutorials, but is basically similar in function to some of the programs mentioned in the previous section, such as POWPAL.

Specialized power analysis software

We reviewed three single-purpose programs that are of potential interest to ecologists. Two freeware programs (MONITOR and TRENDS) are designed to calculate the power to detect trends in species abundance over time (although both could be used for any data that will be analyzed using linear regression). TRENDS takes an analytical approach, and is limited to situations in which monitoring occurs at regular intervals at one site. MONITOR estimates power using Monte-Carlo simulations and can handle unequal sampling intervals and multiple sites. One advantage of these programs over the regression modules in general purpose power software is that they use trend as the effect size measure rather than a standardized measure such as the r^2. Both programs can also estimate power for exponential trends, and for different variance structures (e.g., variance proportional to mean), although neither takes account of the autocorrelation between data points through time. Potential users should note that there has been some debate regarding appropriate methods for power analysis of trends (Gerrodette 1987, Link and Hatfield 1990, Gerrodette 1991), and for determining trends at multiple sites (Thomas 1996). [5]

We also reviewed one commercial program (NSURV) that specializes in power analysis for survival tests (also called failure-time analysis). It is similar to the survival modules in NQUERY ADVISOR, PASS and STPLAN (Table 3). As with trend analysis, appropriate methods are still being developed for survival analysis and especially for calculating power (G. Fox and G. White, pers. comms.). Our informal test runs showed large differences in results between packages for what appeared to be the same problem[6]. We therefore advise caution when interpreting the results of power analysis for survival tests. Because most survival packages are medically oriented, ecologists may need some translation to understand the terms they use.

Programming power analysis using general-purpose statistical software

For more advanced users, writing programs to do power analyses using regular general-purpose statistical software has three advantages: no new software is required, the program can be customized for the analysis at hand, and new advances in statistical theory can be quickly implemented. We know of seven packages that have macro languages and full complement of distribution functions (including non-central distributions) built-in, and can therefore be programmed to calculate exact power for any test where the appropriate formula is known: DATASIM, GAUSS, MACANOVA, SAS, SPSS, S-PLUS, and SST. In addition, two programs have

---

[5] Note also that the "Plot variance" requested by MONITOR is the residual variance in counts over multiple years after the trend line has been fit (e.g., residual mean square from a regression of pilot data against time), and not the within-year variance in counts as implied by the manual (J. Gibbs, pers. comm.).

[6] The survival module in PASS produces incorrect results when accrual time is set to 0. If all individuals are marked at the same time, then accrual time should be set to a small number, such as 0.001. This problem will be corrected in the next version of the program. We could not account for the other differences in results between programs.

almost all of the requisite functions: SHAZAM is missing the non-central χ^2, and STATA does not have non-central F or t distributions.   We found published programs for 3 of these packages (Table 4), but in many cases the code could be easily ported to the others.

Several other programmable packages, such as GLIM, MINITAB, STATISTICA, and XLISP-STAT do not have non-central distribution functions built-in, but have other probability distribution functions including the central t, F, and χ^2 distributions.   These packages can be used to calculate exact power for some tests and approximate power for others, as demonstrated by the programs for SAMG for GLIM and power.lsp for XLISP-STAT (Table 4).

A plug-in module enabling MS EXCEL to calculate non-central distribution probabilities has been written by R. Lenth (Table 4).   The sample workbooks that come with the package demonstrate the ease with which user-friendly spreadsheet applications can be built.   Another alternative for those without access to one of the above programs is POWERPACK (Table 1), a free power analysis package with a built-in command language.   Note that MACANVOA and XLISP-STAT are also freeware.

Power analysis may also be performed using Monte-Carlo simulations, either using specialized simulation software (e.g., DATASIM, SIMSTAT) or general-purpose software that contains pseudo-random number functions (all of the programs listed above).   An example of this approach is Power-Sim for XLISP-STAT (Table 4); other examples are given by Bradley et al. (1992) and Thomas and Juanes (1996).   Simulations are particularly useful to assess the power of nonparametric tests given different plausible underlying distributions (a discussion of other approaches to power analysis for nonparametric tests is in the POWERPACK User's Guide).

GENERAL COMMENTS

Accuracy

For most parametric tests commonly used by ecologists, the statistical theory of power analysis is well developed and algorithms are available for calculating the appropriate distribution functions to the desired level of accuracy.   Despite this, a number of programs use approximate methods (Table 1).   We did not attempt to evaluate rigorously the accuracy of these programs, but tests by Goldstein (1989) showed that some programs gave power estimates that were wrong by as much as 0.3 for a t-test.   Some may argue that such inaccuracies are relatively small compared with those due to the uncertainty in input parameters such as sampling variance. However inaccuracies due to the use of approximate algorithms are unnecessary, and should be avoided.

In some areas of statistics, appropriate methods of analysis are still being developed (e.g., survival analysis, trend analysis) or are controversial (e.g., contingency table analysis: Zar 1996 p.489).   For others (e.g., most multivariate statistics) appropriate tests are available but current methods of power analysis produce only approximate results.   In both cases, users should pay

close attention to the particular method implemented by the package they use and its underlying assumptions and limitations.   Results should be treated as rough guides to supplement the researcher's intuition.   This also highlights the importance of specifying the program and version number used when reporting results.

Surprisingly, some packages do not document the calculation methods used (NCSS, POWER AND PRECISION[7], and SIGMASTAT).   Most merely reference the methods, while only a third (DATASIM, EX-SAMPLE, JMP, MSUSTAT, NQUERY ADVISOR, PASS, STATA, STAT POWER, TRENDS, and TRUE EPISTAT) show the actual formulae used.   Few manuals mention accuracy, and only GPOWER, POWERPACK, and STAT POWER specify the level of accuracy that can be expected in the results.   Many manuals say that the program has been validated, but only GPOWER, N, NSURV and PC-SIZE state *exactly* how the validation was done and give the results of the tests.   Perhaps vendors are concerned by the large amount of extra paper required to properly report validation tests in the manual, knowing that few will read them.   One solution is implemented by the vendors of NQUERY ADVISOR, who have posted their validation report on the internet (http://www.statsol.ie/valid.htm; this is not yet documented in the User's Guide). We encourage others to do likewise.

Flexibility

Many of the programs we have reviewed can deal with the vast majority of power tests required by ecologists.   Nevertheless, there may be some occasions when the built-in tests are not enough.   In this event, the "generic" t-, F- and χ^2 tests offered by some programs (Table 3) allow for a wider range of situations by letting the user specify the appropriate degrees of freedom for the test.   For example, the "Other F tests" option in GPOWER can be used to calculate power for a planned comparison of means (contrasts), or the fixed effect in a randomized complete block design.   Even more flexible are the probability calculators included with some packages (Table 3), which typically calculate cumulative and inverse probabilities for central and non-central t-, F- and χ^2 distributions, as well as many others (z, binomial, negative binomial, exponential, etc.).   These require that the user know the appropriate formula for calculating power, and are laborious to use when many values are to be calculated.   JMP and NCSS have the built-in ability to automate and save the formulae used, although the use of the distribution functions for power analysis are not documented in their manuals.   Similar functionality is available via Lenth's plug-in macro for MS EXCEL (Table 4).   For those willing to write short programs, either POWERPACK or one of the statistics packages listed in the previous section on programming power could be used to automate the process.

---

[7] The methods used by POWER AND PRECISION will apparently be documented in the final version of the manual.

Retrospective power analysis

There appears to be some confusion about the appropriate approach to retrospective power analysis. At least one manual (PASS) gives examples where the actual effect size found in the test is used to compute power, and all three general statistical programs that have options to calculate retrospective power automatically (JMP, NCSS, SIGMASTAT) do so using this unhelpful approach by default. Most of the other manuals ignore retrospective power completely; the only ones that give useful examples (see Background, and Thomas 1997) are N, SHAZAM, STAT POWER and TRUE EPISTAT.

The inclusion of inappropriate methods in the general purpose packages is distressing because these tend to be used by statistically naive researchers. These packages should be modified so that the programs ask users for the alpha level of the test and either: (1) calculate power for a specified effect size (or range of effect sizes) thought to be the minimum for biological significance, or (2) calculate detectable effect size for a specified minimum desired level of power. Another useful addition would be the calculation of confidence intervals about power and detectable effect size estimates (Thomas 1997).

CONCLUSIONS

Contrary to popular perception, there are many software packages available to help researchers perform statistical power analysis (Table 1). The packages vary in the scope of their tests, their accuracy, ease of use, and ease of learning.

For beginner to intermediate level use we recommend one of the commercial general purpose power packages NQUERY ADVISOR, PASS, or STAT POWER. Of these our graduate student class preferred PASS, although both NQUERY ADVISOR and STAT POWER perform some tests not available in PASS (Table 3). Note also that the software is constantly under development: a new version of NQUERY ADVISOR is due to be released in the spring of 1997, as is the promising new program POWER AND PRECISION. Users with more specialized needs may consider other commercial packages, such as EX-SAMPLE for survey design.

For those unable or unwilling to purchase commercial software, there are a large number of freeware or shareware programs available. GPOWER is a reasonably comprehensive free program, and POWER PLANT also has the potential to be very useful once the final version is released. Other free software may also be helpful in some situations, for example STPLAN for survival analysis and other medical statistics, MONITOR and TRENDS for population trend analysis.

The power analysis capabilities built into general purpose statistical software are not currently adequate for most users. Of the programs we reviewed, SIGMASTAT can handle the most tests but has no built-in features to assist with the analysis (such as tools to help calculate effect size or sampling variance, produce tables or graphs, or export the results to other

applications).   JMP is relatively useful for fixed-effect ANOVA designs, but does nothing else. Unfortunately all of the programs that calculate retrospective power (JMP, NCSS, and SIGMASTAT) use the observed effect size to calculate power by default, which produces uninformative results.   In all three programs the manuals do not clearly explain the implications of this approach.

More advanced users who already work with programmable statistics packages, especially those that have built-in non-central distribution functions, should consider using one of the programs already available (Table 4) or writing their own.   For those without access to these packages, a free alternative is POWERPACK.

At present, power analyses are still rare in the ecological literature.   Given the diversity of software available there is no longer any reason to ignore statistical power when planning studies and analyzing their results.   In the near future we expect to see an explosive increase in the use of power analyses fueled by the kinds of high quality statistical software we have reviewed here.

ACKNOWLEDGMENTS

LITERATURE CITED

Bradley, D. R., R. L. Helmstreet, and S. T. Zeigenhagen. 1992. A simulation laboratory for statistics. Behaviour Research Methods, Instruments, and Computers **24**:190-204.

Bradley, D. R., R. L. Russell, and C. P. Reeve. 1996. Statistical power in complex experimental designs. Behaviour Research Methods, Instruments, and Computers **28**:319-326.

Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd Edition. Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey.

Ellison, A. M. 1992. Statistics for PCs. Bulletin of the Ecological Society of America **73**:74-87.

---. 1997. Review of nQuery Advisor, version 1.0. Bulletin of the Ecological Society of America **78**. (In Press)

Erdfelder, E., F. Faul, and A. Buchner. 1996. GPOWER: A general power analysis program. Behaviour Research Methods, Instruments, and Computers **28**:1-11.

Fairweather, P. G. 1991. Statistical power and design requirements for environmental monitoring. Australian Journal of Marine and Freshwater Research **42**:555-567.

Gerrodette, T. 1987. A power analysis for detecting trends. Ecology **68**:1364-1372.

---. 1991. Models for power of detecting trends - a reply to Link and Hatfield. Ecology **72**:1889-1892.

Goldstein, R. 1989. Power and sample size via MS/PC-DOS computers. The American Statistician **43**:253-260.

Greenland, S. 1988. On sample-size and power calculations for studies using confidence intervals. American Journal of Epidemiology **128**:231-237.

Kraemer, H. C., and S. Thiemann. 1987. How Many Subjects? Sage Publications, London, UK.

Krebs, C. J. 1989. Ecological Methodology. Harper Collins, New York, New York.

Link, W. A., and J. S. Hatfield. 1990. Power calculations and model selection for trend analysis: a comment. Ecology **71**:1217-1220.

Lipsey, M. W. 1990. Design Sensitivity: Statistical Power for Experimental Research. Sage, Newbury Park, California.

Muller, K. E., and V. A. Benignus. 1992. Increasing scientific power with statistical power. Neurotoxicology and Teratology **14**:211-219.

Nemac, A. F. L. 1991. Power Analysis Handbook for the Design and Analysis of Forestry Trials. Biometrics Information Handbook **2**. B.C. Ministry of Forests, Victoria.

O'Brien, R. G. 1988. Statistical computing software reviews: PowerPack (Version 2.22). The American Statistician **42**:266-270.

Peladeau, N., and Y. Lacouture. 1993. SIMSTAT: Bootstrap computer simulation and statistical program for IBM personal computers. Behaviour Research Methods, Instruments, and Computers **25**:410-413.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fish and Aquatic Sciences **47**:2-15.

Rotenberry, J. T., and J. A. Wiens. 1985. Statistical power analysis and community-wide patterns. The American Naturalist **125**:164-168.

Taylor, B. L., and T. Gerrodette. 1993. The uses of statistical power in conservation biology: the Vaquita and Northern Spotted Owl. Conservation Biology **7**:489-500.

Thomas, L. 1996. Monitoring long-term population change: why are there so many analysis methods? Ecology **77**:49-58.

---. 1997. Retrospective power analysis. Conservation Biology **11**. (In Press)

Thomas, L., and F. Juanes. 1996. The importance of statistical power analysis: an example from Animal Behaviour. Animal Behaviour **52**:856-859.

Toft, C. A., and P. J. Shea. 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. The American Naturalist **122**:618-625.

Zar, J. H. 1996. Biostatistical analysis, 3rd Edition. Prentice-Hall, Upper Saddle River, New Jersey.

Table 1 - Statistical power analysis software reviewed [1]

Software	Version	Vendor or Principal author	Address	Phone	internet URL [2]	Cost US$ [3]	Operating systems [4]	Calculation methods [5]
A) Stand-alone power and sample size software								
GPOWER	2.0	Edgar Erdfelder (D) / Axel Buchner (M)	(D) Dept. of Psychology, Univ. Bonn, Roemerstrasse 164, D53111 Bonn, Germany / (M) Dept. of Psychology, Univ. Trier, D-54286 Trier, Germany	-	http://www.psychologie.uni-trier.de:8000/projects/gpower.html	free	D, M	E and A
N	2.2	IDV-Datenanalyse und Versuchsplanung	Wessobrunner Str. 6, D-82131 Gauting, Germany	(49-89) 850-3666	-	149	D	E
NQUERY ADVISOR	1.0	Statistical Solutions	Omni Office Plus, 60 State St., Ste. 700, Boston, MA 02109.	(800) 262-1171	http://www.statsolUSA.com/	475	W	E
PASS	6.0	NCSS Statistical Software	329 North 1000 East, Kaysville, UT 84037	(800) 898-6109	http://www.ncss.com/pass.html	249	W	E
PC-SIZE (/CONSULTANT)	2.13 / 1.01	Gerrard E. Dallal	54 High Plain Road, Andover, MA 01810	-	ftp://ftp.coast.net/SimTel/msdos/statistic/st-size.zip	15	D	E
POWER AND PRECISION	pre-release	Michael Borenstein [6]	1421 Hudson Road, Teaneck, NJ, 07666	(201) 692-8155	-	TBA	W95	E
POWPAL	5.0	Bernard S. Gorman	Dept. of Psychology, Nassau Community College, Garden City, NY 11530	(516) 572-7459	-	free	D	A
POWERPACK	2.24	Russell V. Lenth	9 Wellesley Way, Iowa City, IA 52245	(319) 337-8549	ftp://ftp.stat.uiowa.edu/pub/rlenth/PowerPack/	free (25)	D	E
POWER PLANT	public beta	CSIRO Biometrics Unit	Private Bag, PO. Wembley, Western Australia 6014, Australia.	-	ftp://ftp.per.its.csiro.au/csiro-wa/biometrics/	free	W	E
PS	2.0	W. D. Dupont	Vanderbilt Univ. School of Medicine, Nashville, TN 37232-2637	(615) 322-2001	ftp://ftp.vanderbilt.edu/pub/biostat/ps.zip	free	D	A
STATISTICAL POWER ANALYSIS	1.0	Lawrence Erlbaum Associates Inc.	10 Industrial Avenue, Mahwah, NJ 07430	(800) 926-6579	http://www.erlbaum.com/1190.htm	140	D	A
STAT POWER	3.01	QEI Systems / James Bavry	PO Box 12734, Portland, OR 97212	(503) 287-8072	-	180	D	E
STPLAN	4.1	Barry W. Brown	Dept. of Biomathematics, Univ. of Texas, Box 237, 1515 Holcome Blvd, Houston, TX 77030	-	http://odin.mdacc.tmc.edu/anonftp/	free	D, M	A
B) General-purpose statistical software with built-in power analysis capabilities								
DATASIM	1.2	Drake R. Bradley	Dept. of Psychology, Bates College, Lewiston, ME 04240	-	-	55 (M), 35 (D)	M, D	E and S
JMP	3.1	SAS Institute Inc.	SAS Campus Drive, Cary, NC 27513	(919) 677-8000	http://www.sas.com/otherprods/jmp/Home.html	695 / 230	W32, M	E
MACANOVA	4.02	Gary W. Oehlert	School of Statistics, University of Minnesota, 1994 Buford Ave., St. Paul, MN 55108	-	http://www.stat.umn.edu/~gary/macanova/macanova.home.html	free (8)	W32, D, M, U	E
MSUSTAT	5.25	Research & Development Institute, Inc.	Montana State University, Bozeman, MO 59717-0002	(406) 994-3271	-	286	D	E or A

Name	Version	Company/Author	Address	Phone	URL	Price	OS	Code
NCSS	6.0	NCSS Statistical Software	329 North 1000 East, Kaysville, UT 84307	(801) 546-0445	http://www.ncss.com/ncss	399	W	E
SHAZAM	7.0	Ken White	Dept. of Economics, University of British Columbia, Vancouver, BC V6T 1Z1, Canada	(604) 822-5062	http://shazam.econ.ubc.ca	395	W, D, O, M, U	E
SIGMASTAT	2.0	Jandel Scientific Software Corp.	2591 Kerner Boulevard, San Rafael, CA 94901	(800) 452-6335	http://www.jandel/com/	495	W32	?
STATISTICA	5.1	Statsoft	2300 East 14th St., Tulsa, OK 74104	(918) 749-1119	http://www.statsoftinc.com/	995	W	-
STATA	4.0	Stata Corp.	702 University Drive East, College Station, TX 77840	(800) 782-8272	http://www.stata.com/	945 / 395	W, D, M, U, W95	-
TRUE EPISTAT 5.0		Epistat Services	2011 Cap Rock Circle, Richardson, TX 75080-3417	(800) 326-1488	-	498	D	A
WDIST	2.3	Mark Von Tress	P.O. Box 1711173, Arlington TX 76003	-	ftp://ftp.winsite.com/pub/pc/win3/m isc/wdist23.zip	20	W	E
C) Stand-alone sample size software								
EX-SAMPLE	3.0	Idea Works, Inc.	607 Jackson St, Columbia, MO 65203	(800) 537-4866	http://www.ideaworks.com/	150	D, W, M	A
POWER AND EFFECT	2.0	Glenn E. Meyer	Dept. of Psychology, Trinity University, 715 Stadium Drive, San Antonio, TX 78212-7200	(210) 736-8323	http://www.trinity.edu/~gmeyer/power.html	50	W, M	A
D) Specialized power analysis software								
MONITOR	6.2	James P. Gibbs	Dept. of Earth and Environmental Sciences, Columbia University, 1200 Amsterdam Ave., New York, NY 10027	(212) 854-8056	ftp://ftp.im.nbs.gov/pub/software/monitor/	free	D	S
NSURV	2.2	IDV-Datenanalyse und Versuchsplanung	Wessobrunner Str. 6, D-82131 Gauting, Germany	(49-89) 850-3666	-	149	D	E
TRENDS	1.0	Tim Gerrodette	Southwest Fisheries Science Center, National Oceanic and Atmospheric Administration, PO Box 271, La Jolla, CA 92038	-	ftp://ftp.im.nbs.gov/pub/software/CS E/wsb21515/trends.zip	free	D	E

[1] Further information, including e-mail addresses and references to software descriptions are available at http://www.interchg.ubc.ca/cacb/power/

[2] For commercial packages, URL is WWW page. For free software, URL is anonymous FTP or WWW site for downloading the program.

[3] First number is recommended retail price, second is academic price (neither include shipping). Most commercial packages offer multi-user discounts, networking options, etc. For freeware, the number in brackets is the price of the printed manual, if available.

[4] D = DOS; M = Macintosh; O = OS/2; U = Unix; W = Windows 3.x; W32 = Windows 32s, W95 = Windows 95. The operating system listed first is the one used to review the program.

[5] E = exact non-central distributions used when appropriate; A = approximations to non-central distribtuions; - = does not perform tests where non-central distributions are required; S = simulations.

[6] A version of POWER AND PRECISION will be distributed by SPSS, Inc.

Table 2 - Summary evaluation of power analysis software

Software	Ease of use [1]	Ease of learning [1]
A) Stand-alone power and sample size software		
GPOWER	***	**
N	**	***
NQUERY ADVISOR	***	****
PASS	****	****
PC-SIZE	*	***
POWER AND PRECISION	***	****
POWPAL	*	**
POWERPACK	**	**
POWER PLANT	**	*
PS	*	**
STPLAN	*	**
STATISTICAL POWER ANALYSIS	**	****
STAT POWER	***	****
B) General-purpose statistical software with built-in power analysis capabilities [2]		
DATASIM	**	*
JMP	**	***
MACANOVA	*	**
MSUSTAT	*	**
NCSS	*	***
SHAZAM	*	**
SIGMASTAT	**	***
STATISTICA	*	*
STATA	*	***
TRUE EPISTAT	**	****
WDIST	*	**
C) Stand-alone sample size software		
EX-SAMPLE	**	***
POWER AND EFFECT	**	***
D) Specialized power analysis software		
MONITOR	**	**
NSURV	**	***
TRENDS	*	**

[1] ***** = excellent, **** = very good, *** = good, ** = fair, * = poor.
[2] Evaluation based solely on the power analysis components of these programs

Table 3 – A selection of situations where the calculation of power (P), sample size (S) or detectable effect size (D) is discussed in the documentation of the programs reviewed.[1,2]

Situation:[3]	z-test			t-test			Fixed-effects ANOVA				Random-effects ANOVA		Repeated measures		Regression			Logistic	Correlation	
Software	1-samp	2-samp	Gen-eric[4]	1-samp	2-samp	Gen-eric[4]	1-way	1-way contrasts	2-way	3-way	1-way	2-way	Uni-variate	Multi-variate	Simple	Compare 2 slopes	Multiple	Logistic	Pearson	Diff. in corrs.
A) Stand-alone power and sample size software																				
GPOWER	PS	PS	P	PS	PS	P	PS	PS									PS		PS	
N			PSD	PSD																
NQUERY ADVISOR	PSD	PSD		PSD	PSD		PSD	PSD	PS				PSD		PSD		PSD	PS	PSD	PSD
PASS			PS	PSD	PSD	PS	PSD	PSD	PS	PS			PS		PS	PS	PS	PS	PSD	PSD
PC-SIZE			PS	PS	PS	PS	PS	PS	PS				PS		PS			PS	PS	
POWER AND PRECISION	PS	PS	P	PS	PS	P	PS	PS	PS	PS			PS		PS	PS	PS	PS	PS	PS
POWPAL			PS			PS														PS
POWERPACK	P	P	P	P	P	PSD	PSD	PSD			PSD		PSD	PS					PSD	
POWER PLANT						P	P	P				P								
PS			PSD	PSD	PSD								PSD	PSD	PSD	PSD			PSD	PSD
STPLAN			PSD	PSD	PSD															
STATISTICAL POWER ANALYSIS			PS	PS		PS	PS	PS					PS		PS		PS		PS	PS
STAT POWER	S	PS	PS	PS	PS	PS	PS	PS	PS	PS			PS	PS	PS		PS		PS	PS
B) General-purpose statistical software with built-in power analysis capabilities																				
DATASIM			PS		PS	P	P						P	PS	PS				PS	
JMP						P	P	P	P											
MACANOVA			PS	PS	PS	P								S						
MSUSTAT			S	S	S	S								S			S			
NCSS			P	P	P	P	P	P						P	P		P			
SHAZAM			P																	
SIGMASTAT	PS	PS	PS	PS	PS	PS	P	P					P	P	PS		P	P	PS	
STATISTICA	S																			
STATA	PS	PS																		
TRUE EPISTAT	PS	PS	PS	PS	PS										S				S	
WDIST	PS	PS	PS	PS	S		S	S												
C) Stand-alone sample size software																				
EX-SAMPLE			S	S	S	S	S	S	S				S	S	S	S	S	S	S	S
POWER AND EFFECT			S	S	S	S	S	S					S	S	S		S	S	S	S

[1] A complete listing of test situations is available at http://www.interchg.ubc.ca/cacb/power/

[2] Because of their narrow scope, the three specialized power analysis packages reviewed (MONITOR, NSURV and TRENDS) are not listed. MONITOR and TRENDS calculate P, S, or D for trend analysis; NSURV calculates P, S ,or D for survival analysis.

[3] Note some situations are special cases of others: for example a program that can perform power analysis for an M x N contingency table will also work for 2 x N and 2 x 2 tables, and 2 and N category goodness-of-fit tests.

[4] For "Generic" tests, standardized effect size measures are used and the user inputs the appropriate degrees of freedom for the test.

Table 3 (contd.)

Situation:[3] Program	Proportions: Prop. = p	Proportions: Diff. in props.[4]	χ²: Gen-eric[4]	Goodness-of-fit: 2 cats.	Goodness-of-fit: N cats.	Contingency tables: 2 x 2	Contingency tables: with Yates	Contingency tables: Fisher's	Contingency tables: 2 x N	Contingency tables: M x N	Survival: Log-rank	Survival: Exponential	Non-parametric: Sign test	Non-parametric: Wilcoxon	Non-parametric: Spearman	Probability calculator[5]
A) Stand-alone power and sample size software																
GPOWER			PS													
N	PSD	PSD														
NQUERY ADVISOR	PS	PS	PS			PS		PS	PS		PS	PS				PD
PASS	PSD	PSD	PSD			PS	PS	PSD	PSD		PSD	PS		PSD		P[6]
PC-SIZE	PS	PS	PS													
POWER AND PRECISION	PS	PS	P			PS				PS						
POWPAL			PS	P												
POWERPACK	P			P		P				P			P	PS		P
POWER PLANT											PSD					P
PS	PSD	PSD														
STPLAN	PSD	PSD	PS	P		PS	PS	PSD	PSD		PSD	PSD				PSD[7]
STATISTICAL POWER ANALYSIS		PS	PS			PS	PS									
STAT POWER	PS	PS	PS	P			P			PS			PS			P
B) General-purpose statistical software with built-in power analysis capabilities																
DATASIM			PS													P
JMP																P
MACANOVA																P
MSUSTAT																
NCSS																P
SHAZAM																
SIGMASTAT		PS	PS			P	P	P								
STATISTICA	S															
STATA	PS	PS														
TRUE EPISTAT	PS			S		PS					S	S		P		
WDIST	PS	PS	P	P				P		P						P
C) Stand-alone sample size software																
EX-SAMPLE	S	S		S	S					S	S		S	S	S	
POWER AND EFFECT	S		S	S		S		P		S	S			P	S	P

Table 4 - Published programs that enable general-purpose statistical software to perform power or sample size calculations.

Software [1]	Program / Macro	Principal Author	Source / Reference
DATASIM	not named	D. R Bradley	Bradley et al. 1992 (simulations). Bradley et al. 1996 (distribution functions)
MS EXCEL	PiFace	R. V. Lenth	http://www.stat.uiowa.edu/ftp/rlenth/PiFace/
GLIM	SAMG	N. Andrews	http://www.nag.co.uk/stats/GDGE/sample/sample.html
SHAZAM	not named	K. White	User's reference manual, v7.0 (p345-348)
SAS	AOVPOWER.SAS	SAS Institute Inc.	SAS sample library
	POWER.SAS	R. G. O'Brien	SAS sample library
	POWERLIB.SAS	K. E. Muller	ftp://ftp.uga.edu/pub/sas/contrib/cntb0014
	UNIFYPOW.SAS	R. G. O'Brien	ftp://ftp.bio.ri.ccf.org/UnifyPow.all
	FPOWER.SAS	M. Friendly	ftp://hotspur.psych.yorku.ca/pub/sas/macros/fpower.sas
	several programs	L. Thomas	http://www.interchg.ubc.ca/cacb/people/lthomas/download.html
	several programs	A. F. Nemec	Nemec 1991
SIMSTAT	not named	N. Peledau	Peladeau and Lacoutre 1993
S-PLUS	POWER	D. F. Heitjan	ftp://lib.stat.cmu.edu/s/power/
	asypower	B. W. Brown	ftp://odin.mdacc.tmc.edu/pub/S/asypow/
XLISP-STAT	power.lsp	J. Bond	http://www.stat.ucla.edu/code/statistics/distributions
	Power-Sim	J. T. Behrens	http://seamonkey.ed.asu.edu/~behrens/asu/reports/power/

[1] A list of vendors and contact information for these packages is available at http://www.interchg.ubc.ca/cacb/power/