SAMPLE SIZE FOR CONFIDENCE INTERVAL-BASED INFERENCE, VERSION 1

Michael R. Jiroutek and Keith E. Muller

The program may be downloaded at no cost from *http://www.bios.unc.edu/~muller*

Development of the programs was supported mainly by NCI R01 CA095749-01A1, and also in part by NIBIB EB000219 and NCI P01 CA47 982-04. We also thank Michael Getter for helpful comments on the program and manual.

1. OVERVIEW

1.1 Features

The program *CISIZE01* is based on the code used to produce the example results and graphs in Jiroutek, Muller, Kupper and Stewart (2003). We use the notation in that paper, and assume familiarity with the content of the article. The program computes the probability of width and rejection given validity, denoted Pr$\{(W \cap R)|V\}$, as well as the special (limiting) cases Pr$\{W|V\}$, Pr$\{W\}$ and Pr$\{R\}$. Here Pr$\{R\}$, the probability of rejecting the null hypothesis, may be described as unconditional power. The results apply, even in small samples, to any scalar (mean) parameter in a General Linear Multivariate Model (GLMM) with Gaussian errors and fixed predictors. Muller, LaVange, Ramey and Ramey (1992) reviewed the theory of the model under the null and alternative in the context of power analysis. O'Brien and Muller (1993) gave a more extensive tutorial on power in linear models. Timm's (2002) book contains an excellent presentation of the GLMM within the greater context of multivariate data analysis.

The generality of the program stems from using matrix notation and the matrix syntax of the SAS/IML® environment. Please refer to the IML manual for the rules of creating matrices and other language syntax. Numerical integration of exact formulas are used exclusively in the program. For studies focused on confidence interval-based inferences, the criterion Pr$\{(W \cap R)|V\}$, when compared to other probability criteria, typically better aligns the sample size rule with scientific goals. The criterion simultaneously controls interval width, the validity of the interval (whether or not the interval contains the parameter of interest) and whether the interval contains the null value (rejection).

As discussed in Jiroutek, et al. (2003), only five logically consistent and theoretically plausible combinations of one- and two-sided confidence interval and hypothesis test occur. All are handled by this program. For invalid combinations, the program returns a missing value, ".", as appropriate. In addition, some combinations of inputs lead to probabilities of exactly zero (for impossible events).

We expect most users, most of the time, will need only the default combination of a two-sided confidence interval with a two-sided hypothesis test. Much of the text in this document involves the other combinations. If you are amongst the great majority interested only in the default two-sided confidence interval with two-sided hypothesis test combination, you may skip the text *related to the other cases. For convenience, we have marked all discussion involving one-sided cases with a vertical bar in both margins.*

1.2 The Simplest Program: One $\Pr\{(W \cap R)|V\}$ Value

The default is for the program to output all four probabilities of interest ($\Pr\{(W \cap R)|V\}$, $\Pr\{W|V\}$, $\Pr\{W\}$, and $\Pr\{R\}$) simultaneously, for the two-sided confidence interval, two-sided hypothesis test combination. Of course, any single probability or combination of interest can be displayed by controlling which output is specified.

The program assumes that the design matrix, $\boldsymbol{X}$, is specified in terms of an *essence* matrix, $\text{Es}(\boldsymbol{X})$, and a replication factor (REPN in the program): $\boldsymbol{X} = \text{Es}(\boldsymbol{X}) \otimes \mathbf{1}_{\text{REPN}}$, i.e. for any two matrices $\boldsymbol{G}$ and $\boldsymbol{H}$, $\boldsymbol{G} \otimes \boldsymbol{H} = \{g_{ij} \cdot \boldsymbol{H}\}$. The essence matrix (Helms, 1988) contains one and only one copy of each unique row of the original matrix. For example, cell mean coding for an independent groups t-test with 10 observations per group has, with $\mathbf{1}_{10}$ a 10×1 vector of 1's and $\mathbf{0}_{10}$ a 10×1 vector of 0's,

$$\begin{aligned} \boldsymbol{X} &= \begin{bmatrix} \mathbf{1}_{10} & \mathbf{0}_{10} \\ \mathbf{0}_{10} & \mathbf{1}_{10} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \mathbf{1}_{10} \end{aligned}$$

and

$$\text{Es}(\boldsymbol{X}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I}_2 \ .$$

Muller and Fetterman (2000, Chapters 12-16) provided detailed discussions of defining and using essence matrices in coding linear model design matrices.

Computing $\Pr\{(W \cap R)|V\}$ for a GLH with fixed predictors requires knowing only eight variables: $\boldsymbol{\Sigma}$, $\boldsymbol{X}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{U}$, α, $\boldsymbol{\Theta}_0$ and δ, where δ is the desired width of the confidence interval. The user must always specify these eight variables, including $\boldsymbol{X}$ in terms of $\text{Es}(\boldsymbol{X})$ and a replication factor.

By default, the program will compute probabilities for a two-sided confidence interval in combination with a two-sided test. Choosing `CISIDE = {"LU"};` is equivalent to the default of a two-sided confidence interval of the form $[L, U]$. Choosing `ALTHYP = {"^="};` to specify the alternative hypothesis of interest is equivalent to the default of a two-sided test of $H_0: \theta_d = 0$ versus the alternative $H_{\text{A}}: \theta_d \neq 0$, where $\theta_d = \theta - \theta_0$.

The user may specify a one-sided confidence interval or test. Choosing `CISIDE = {"L."};` specifies a confidence interval of the form $[L, \infty)$, while choosing `CISIDE = {".U"};` specifies a confidence interval of the form $(\infty, U]$. Choosing `ALTHYP = {"<0"};` specifies a one-sided test of $H_0: \theta_d \geq 0$ versus the alternative $H_{\text{A}}: \theta_d < 0$. For such a one- sided test, a critical (rejection) region occurs in the lower tail, not the upper. Finally, choosing `ALTHYP = {">0"};` specifies a one-sided test of $H_0: \theta_d \leq 0$ versus the alternative $H_{\text{A}}: \theta_d > 0$. For such a one-sided test, a critical (rejection) region occurs in the upper tail, not the lower.

Allowing three choices for CISIDE, namely {"LU"}, {"L ."} or {". U"}, and three choices for ALTHYP, namely {"$\wedge =$"}, {"<0"} or {">0"}, implies $3 \cdot 3 = 9$ combinations. However, only 5 of 9 cases are logically appealing and theoretically possible for the calculation of Pr$\{(W \cap R)|V\}$. As discussed in Jiroutek et al. (2003), a two-sided hypothesis test paired with a one-sided confidence interval has no logical appeal. Hence requesting such a combination causes the program to return a missing value for Pr$\{(W \cap R)|V\}$ (for 2 of the 9 cases). Two other cases lead to impossible events because the critical regions of the test and the confidence interval lie in opposite tails. The program returns a zero for Pr$\{(W \cap R)|V\}$ in such cases.

The program requires the user to specify the total probability of the critical (rejection) regions as ALPHA. There is a correspondence between the rejection regions of the hypothesis test and confidence interval, which in essence preserves the total alpha level. Consider the two-sided confidence interval and two-sided hypothesis case. Specifying an alpha level of 0.05 in the code will result in alpha of 0.025 in each tail of the hypothesis test. As shown in the underlying theory in Jiroutek, et al. (2003), the relationship between the hypothesis test and confidence interval results in the same 0.025 in each tail of the hypothesis test to be used in each tail of the confidence interval, resulting in a overall alpha level of 0.05, despite there being four tails of interest (two each for the hypothesis test and confidence interval). Therefore, for a two-sided test paired with a two-sided confidence interval, each of the two critical regions for the test and each of the two critical regions for the confidence interval have probability $\alpha/2$ (a total Type I error probability of α).

A similar logic holds for the other cases. For a one-sided test paired with a one-sided confidence interval, the one critical region for the test and the one critical region for the confidence interval each have probability α (a total Type I error probability of α). For a one-sided test paired with a two-sided confidence interval, the one critical region for the test has probability $\alpha/2$, while the two critical regions for the confidence interval each have probability $\alpha/2$ (again, a total Type I error probability of α). For this last case, allowing an alpha of 0.05 in the one tail of the hypothesis test would result in serious logical inconsistencies between the test and confidence interval. To avoid this, the rejection area of the test tail is set to correspond to the rejection area of each tail of the confidence intervals.

The following code suffices to compute Pr$\{(W \cap R)|V\}$, Pr$\{W|V\}$, Pr$\{W\}$, and Pr$\{R\}$ for a two-sample independent groups t-test with 25 observations per group and reference cell coding, assuming a two-sided alternative hypothesis and two-sided confidence interval. Muller and Fetterman (2002, Chapters 12-14) gave a useful introduction to ANOVA coding in one and two way designs.

```
TITLE1 "EXAMPLE01.SAS Independent Groups t Test";
PROC IML SYMSIZE=4000 WORKSIZE=4000;
%INCLUDE "..\CISIZE01.IML" / NOSOURCE2;

ALPHA={0.05};

ESSENCEX={1 0,
          1 1};
REPN={25};

C={0 1};
BETA={0,
      1};  BETASCAL={1};
U={1};

SIGMA=I(1); *Variance value, not standard deviation;
```

```
SIGSCAL={1};
DELTA={1}; DELTSCAL={1};
THETA0={0};

EVENTS={"WARGV" "WGV" "WIDTH" "REJECT"};
RUN PREVENT;

********************************************;
*Control matrices with defaults are as follows;
*CISIDE= { "LU" };
*ALTHYP= { "^=" } ;
*ROUND={3};
*MAXPDIFF={1E-5};
*OPT_ON is left undefined if wish program to print result matrix;
*assigning value of OPT_ON={NOPRINT} causes program to print nothing;
********************************************;
```

The program prints the following output.

```
                                  _HOLDPR_

           ALPHA  BETASCAL  DELTSCAL   SIGSCAL   TOTAL_N   ICISIDE

  ROW1      0.05         1         1         1        50         2


                                  _HOLDPR_

         IALTHYP     WARGV       WGV     WIDTH    REJECT

  ROW1         2     0.121     0.121     0.127     0.934
```

Please note the following:

1) The first two program statements are always required. They initialize IML, ask for extra symbol and work space, make the module available for use.

2) The next block of code (from the ALPHA= to the EVENTS= line) must always be present (in any order) to specify ALPHA, SIGMA (SIGSCAL), ESSENCEX, C, BETA (BETASCAL), DELTA (DELTSCAL), REPN, THETA0 and EVENTS.

3) The RUN PREVENT; statement is always required to execute the program (except when using the code as a parameterized module with a CALL statement, as in Section 1.6).

4) The four columns in the output matrix created, `WARGV`, `WGV`, `WIDTH`, and `REJECT`, correspond to Pr$\{(W \cap R)|V\}$, Pr$\{W|V\}$, Pr$\{W\}$, and Pr$\{R\}$, respectively.

5) For univariate models, SIGMA is the *variance*, not the standard deviation, because $\mathbf{\Sigma} = \sigma^2$ is 1×1. The potential source of confusion arises from the substantial advantages of treating the univariate case as a special case of the multivariate, which requires the same variable name (SIGMA) for both settings.

6) ICISIDE and IALTHYP in the output display the "sidedness" of the confidence interval and hypothesis test chosen. Values of -1, 1 or 2 are possible, indicating the left, right or both tails. This helps to confirm that the desired CISIDE and ALTHYP inputs were chosen.

7) See Section 1.5 of this document for further details about the required inputs.

1.3 Unequal Cell Sizes

There are various ways to allow for unequal cell sizes. As an example, consider

$$\boldsymbol{X} = \begin{bmatrix} \mathbf{1}_{10} & \mathbf{0}_{10} & \mathbf{0}_{10} \\ \mathbf{0}_{15} & \mathbf{1}_{15} & \mathbf{0}_{15} \\ \mathbf{0}_{20} & \mathbf{0}_{20} & \mathbf{1}_{20} \end{bmatrix} .$$

One way to specify this is:

```
ESSENCEX = {      1 0 0,
                  1 0 0,
                  0 1 0,
                  0 1 0,
            0 1 0,
                  0 0 1,
                  0 0 1,
            0 0 1,
            0 0 1};
REPN = 5;
```

If in doubt about how to use ESSENCEX and REPN together, the user can always specify ESSENCEX as the entire $\boldsymbol{X}$ matrix and set REPN = 1. For example, the $\boldsymbol{X}$ matrix immediately above could be given using:

```
ONE = {1 0 0}; TWO = {0 1 0}; THREE = {0 0 1};
ESSENCEX = REPEAT(ONE, 10, 1) // REPEAT(TWO, 15, 1) // REPEAT(THREE, 20, 1);
REPN = 1;
```

1.4 Controlling Printing

Given valid inputs, the program always produces the matrix _HOLDPR_ with column names _PRNM_. The user has control of whether the program prints the matrix. Merely add the statement OPT_ON={NOPRINT}; before RUN PREVENT; to disallow the program from printing.

1.5 Producing $\Pr\{(\boldsymbol{W} \cap \boldsymbol{R})|\boldsymbol{V}\}$ for a Range of Scenarios

The program makes it easy to compute $\Pr\{(W \cap R)|V\}$ for ranges of sample sizes, test sizes, mean differences, confidence interval widths, and variances. The following code generalizes the preceding example. The inputs ask for 10 values of REPN, 3 values of BETASCAL, 3 values of SIGSCAL, 3 values of DELTSCAL and 2 values of ALPHA. BETASCAL is multiplied times BETA, and could be called THETASCAL because it also multiplies $\boldsymbol{\Theta}$ by the same amount. In this t-test example, $\boldsymbol{\Theta} = \mu_1 - \mu_2$. Hence the BETASCAL statement leads to the four probabilities available in the program to be computed for $\mu_1 - \mu_2 = 0.5$, $\mu_1 - \mu_2 = 1.0$, and $\mu_1 - \mu_2 = 1.5$. Similarly, SIGSCAL values are multiplied times $\boldsymbol{\Sigma}$. Hence the SIGSCAL statement leads to the four probabilities available in the program to be computed for $\sigma^2 = 0.5$, $\sigma^2 = 1.0$, and $\sigma^2 = 1.5$. Further, DELTSCAL values are multiplied times δ. Hence the DELTSCAL statement leads to the four probabilities available in the program to be computed for $\delta = 0.5$, $\delta = 1.0$, and $\delta = 1.5$. Including the two values in ALPHA leads to the four probabilities available in the program to be computed for $\alpha = 0.025$ and $\alpha = 0.05$. Each combination of inputs leads to a value of $\Pr\{(W \cap R)|V\}$, $\Pr\{W|V\}$, $\Pr\{W\}$, and $\Pr\{R\}$, for a total of $10 \cdot 3 \cdot 3 \cdot 3 \cdot 2 = 540$ values of each of the four probabilities of interest.

```
TITLE1 "EXAMPLE02.SAS Independent Groups t Test, input lists for many values";
PROC IML SYMSIZE=4000 WORKSIZE=4000;
```

```
%INCLUDE "..\IML\CISIZE01.IML" / NOSOURCE2;

ALPHA={0.025 0.05};

ESSENCEX={1 0,
          1 1};
REPN=DO(10,100,10);

C={0 1};
BETA={0,
      1}; BETASCAL={0.5 1.0 1.5};
U={1};

SIGMA=I(1); SIGSCAL={0.5 1.0 1.5};
DELTA={1}; DELTSCAL={0.5 1.0 1.5};
THETA0={0};

EVENTS={"WARGV" "WGV" "WIDTH" "REJECT"};
RUN PREVENT;

*Next 2 lines creates a SAS data file in default library (usually WORK);
CREATE ONE FROM _HOLDPR_[COLNAME=_PRNM_];
APPEND FROM _HOLDPR_;
* sorted BY SIGSCAL BETASCAL DELTSCAL TOTAL_N ALPHA;
*End of IML code;

PROC PRINT DATA=ONE UNIFORM;
VAR  TOTAL_N
WARGV WGV WIDTH REJECT;
FORMAT WARGV WGV WIDTH REJECT 7.5;
BY SIGSCAL BETASCAL DELTSCAL;
PAGEBY SIGSCAL;
RUN;
```

Please note the following general principles, as just illustrated.

2) Some program inputs, including REPN, ALPHA, SIGSCAL, BETASCAL, DELTSCAL and EVENTS, may be assigned lists of values.

3) Such lists may be matrices which are $1 \times n$ (a matrix with one row) or $n \times 1$ (a vector).

4) Although the REPN values requested are 10, 20, 30, ..., 100, because ESSENCEX represents a two-sample independent groups t-test, the sample size values output from the program are 20, 40, 60, ..., 200 and represent the *total* sample size. The input values specify the *per group* sample size.

5) The program produces $\Pr\{(W \cap R)|V\}$, $\Pr\{W|V\}$, $\Pr\{W\}$, and $\Pr\{R\}$ for all factorial combinations of inputs, though output can be restricted to any specific probabilities of interest.

6) Creating a permanent file is easy. Merely change the temporary file name ONE in this program to the desired target filename.

7) Inspection of the output (see EXAMPLE02.LST) reveals several missing values for $\Pr\{(W \cap R)|V\}$ and $\Pr\{W|V\}$. They are due to convergence difficulties of the numerical integration QUAD function at the lower probability boundary of zero.

1.6 Calling the Program from a User Written Module

The software can also be run by using a CALL statement. This allows using the power software within other modules. Note that the *CISIZE01.IML* input matrices must all be specified to be global variables. The following incomplete code illustrates how to define a user module that calls the power software.

```
START user_defined_module ( parm1, parm2, ... parmn)
        GLOBAL(EVENTS,CISIDE,ALTHYP,ALPHA,ESSENCEX,REPN,
                      C,BETA,BETASCAL,U,
                      SIGMA,SIGSCAL,DELTA,DELTSCAL,
                      THETA0,ROUND,MAXPDIFF);

*some of your code here;
CALL _PREVENT_(_HOLDPR_, _PRNM_);
*some more of your code here;

FINISH user_defined_module;
```

1.7 Summary of Inputs Controlling The Program

Seventeen global matrices, along with OPT_ON, are the only inputs to the program. The input matrices are listed in the following table, grouped by application.

Table 1 All Possible Input Matrices By Category

Design & Hypothesis	Scenario Variations	Output & Methods Desired	Numeric
CISIDE	REPN	OPT_ON	ROUND
TESTSIDE	BETASCAL		MAXPDIFF
ESSENCEX	DELTSCAL		
BETA	SIGSCAL		
C	ALPHA		
U			
THETA0			
SIGMA			
DELTA			

Table 2 All Possible Input Matrices, Definitions and Descriptions

Name	Description	Size	Default
ALPHA	Type I error rates	1 row or col	user specified
CISIDE[1]	Specifies CI side(s) of interest	1×1	{"LU"}
ALTHYP[2]	Specifies test side(s) of interest	1×1	{" ^= "}
ESSENCEX[3]	Essence $\boldsymbol{X}$ matrix	$G \times q$	user specified
BETA	Matrix $\boldsymbol{B}$ in GLH	$q \times p$	user specified
BETASCAL	List of multipliers for BETA	1 row or col	user specified
REPN	List specifying # of times to duplicate each row of ESSENCEX	1 row or col	user specified
DELTA	CI width desired	1×1	user specified
DELTSCAL	List of multipliers for DELTA	1 row or col	user specified
SIGMA	Variance	1×1	user specified
SIGSCAL	List of multipliers for SIGMA	1 row or col	user specified
THETA0	Matrix $\boldsymbol{\Theta}_0$ in GLH	$a \times b$	user specified
C	Matrix $\boldsymbol{C}$ in GLH	$a \times q$	user specified
U	Matrix $\boldsymbol{U}$ in GLH	$p \times b, b \leq p$	user specified
ROUND	Scalar specifying how many decimal places to round output values	1×1	3
MAXPDIFF	Scalar specifying what the software considers numeric zero	1×1	1×10^{-5}
OPT_ON[4]	Turn off print option	1×1	-

[1]The only other options available for CISIDE are {"L . "} and {" . U"}.

[2]The only other options available for ALTHYP are {" < 0 "} and {" > 0 "}.

[3]Created by deleting any duplicate rows from the design matrix, $\boldsymbol{X}$ (Helms, 1988) and hence $G \times q$, with G the number of distinct groups of subjects.

[4]OPT_ON={NOPRINT}; suppresses printing, otherwise output displayed.

2. POLYNOMIAL CONTRASTS

If polynomial contrasts are desired, the user utility modules in LINMOD software can be used to help make the task easier. LINMOD may be downloaded at no cost from *http://www.bios.unc.edu/~muller*.

3. LIMITATIONS AND KNOWN DIFFICULTIES

1) The user must avoid using any of the following matrix names except as input to the software: ESSENCEX, SIGMA, BETA, C, U, THETA0, REPN, BETASCAL, SIGSCAL, ALPHA, ROUND, MAXPDIFF, CISIDE, TESTSIDE, DELTA, DELTSCAL.

2) Each new execution of the program causes any values previously stored in the following output matrix names to be lost: _HOLDPR_, _PRNM_. Simply copying the matrices to matrices with distinct names allows retaining results from two or more invocations of the program.

4. MORE EXAMPLES

The code and output listings for all the following programs are found in the "EXAMPLES" folder included in the ZIP file available for download on the web site. They can be run in an interactive or a batch environment and do not require any preexisting files.

Only one change *must be made* for the programs to run successfully. The location of the IML code must be specified correctly in the statement

```
%INCLUDE "your location here\CISIZE01.IML" / NOSOURCE2;
```

The programs that produce plots will likely require changing the FILENAME statement, as well as a few GOPTIONS, such as DEVICE, to tailor the output to the particular computer. The DISPLAY (in GOPTIONS) may also need to be controlled.

Example03.SAS

The following program calculates $\Pr\{(W \cap R)|V\}$, $\Pr\{W|V\}$, $\Pr\{W\}$, and $\Pr\{R\}$ for a paired data t-test with 25 observations, assuming a two-sided alternative hypothesis and two-sided confidence interval.

```
TITLE1 "EXAMPLE03.SAS--Paired data T-test";
TITLE2 "two sided CI [L, U], two sided test (H_a: theta_d ^= 0)";
PROC IML SYMSIZE=4000 WORKSIZE=4000;
%INCLUDE "..\IML\CISIZE01.IML" / NOSOURCE2;

ALPHA={0.05};

ESSENCEX={1};
REPN={25};

C={1};
BETA={1}; BETASCAL={1};
U={1};

SIGMA=I(1); SIGSCAL={1};
DELTA={1}; DELTSCAL={1};
THETA0={0};

EVENTS={"WARGV" "WGV" "WIDTH" "REJECT"};
RUN PREVENT;
```

Example04.SAS

The following program calculates $\Pr\{(W \cap R)|V\}$, $\Pr\{W|V\}$, $\Pr\{W\}$, and $\Pr\{R\}$ for a two-sample independent groups t-test with 25 observations per group and reference cell coding, assuming a one-sided test of the form $H_0 : \theta_d \geq 0$ versus $H_{\mathrm{A}} : \theta_d < 0$, and a two-sided confidence interval.

```
TITLE1 "EXAMPLE04.SAS--Independent groups T-test";
TITLE2 "two sided CI, one sided test (H_a: theta_d < 0)";
PROC IML SYMSIZE=4000 WORKSIZE=4000;
%INCLUDE "..\IML\CISIZE01.IML" / NOSOURCE2;
```

```
ALTHYP={"<0"};
*Note: default CISIDE={"LU"} still in effect, no need to specify;

ALPHA={0.05};

ESSENCEX={1 0,
          1 1};
REPN={25};

C={0 1};
BETA={0,
      1};  BETASCAL={-1};
U={1};

SIGMA=I(1); SIGSCAL={1};
DELTA={1}; DELTSCAL={1};
THETA0={0};

EVENTS={"WARGV" "WGV" "WIDTH" "REJECT"};
RUN PREVENT;
```

Note that BETASCAL is negative. For this particular combination of confidence interval and hypothesis test, a negative BETASCAL is required for the request to make sense. The alternative hypothesis being considered creates the requirement. A positive BETASCAL value in this program will result in a missing value returned for Pr$\{(W \cap R)|V\}$.

Example05.SAS

The following program calculates Pr$\{(W \cap R)|V\}$, Pr$\{W|V\}$, Pr$\{W\}$, and Pr$\{R\}$ for a two-sample independent groups t-test with 25 observations per group and reference cell coding, assuming a one-sided test of the form $H_0 : \theta_d \leq 0$ versus $H_\text{A} : \theta_d > 0$, and a two-sided confidence interval.

```
TITLE1 "EXAMPLE05.SAS--Independent groups T-test";
TITLE2 "two sided CI, one sided test (H_a: theta_d > 0)";
PROC IML SYMSIZE=4000 WORKSIZE=4000;
%INCLUDE "..\IML\CISIZE01.IML" / NOSOURCE2;

ALTHYP={">0"};
ALPHA={0.05};

ESSENCEX={1 0,
          1 1};
REPN={25};

C={0 1};
BETA={0,
      1};  BETASCAL={1};
U={1};

SIGMA=I(1); SIGSCAL={1};
DELTA={1}; DELTSCAL={1};
THETA0={0};

EVENTS={"WARGV" "WGV" "WIDTH" "REJECT"};
RUN PREVENT;
```

Example06.SAS

The following program calculates Pr$\{(W \cap R)|V\}$, Pr$\{W|V\}$, Pr$\{W\}$, and Pr$\{R\}$ for a two-sample independent groups t-test with 25 observations per group and reference cell coding, assuming a one-sided test of the form $H_0 : \theta_d \leq 0$ versus $H_\mathrm{A} : \theta_d > 0$, a and one-sided confidence interval of the form $[L, \infty)$.

```
TITLE1 "EXAMPLE06.SAS--Independent groups T-test";
TITLE2 "one sided CI [L, inf.), one sided test (H_a: theta_d > 0)";
PROC IML SYMSIZE=4000 WORKSIZE=4000;
%INCLUDE "..\IML\CISIZE01.IML" / NOSOURCE2;

CISIDE={"L."}; ALTHYP={">0"};
ALPHA={0.05};

ESSENCEX={1 0,
          1 1};
REPN={25};

C={0 1};
BETA={0,
      1};  BETASCAL={1};
U={1};

SIGMA=I(1); SIGSCAL={1};
DELTA={1}; DELTSCAL={1};
THETA0={0};

EVENTS={"WARGV" "WGV" "WIDTH" "REJECT"};
RUN PREVENT;
```

Example07.SAS

The following program calculates Pr$\{(W \cap R)|V\}$, Pr$\{W|V\}$, Pr$\{W\}$, and Pr$\{R\}$ for a two-sample independent groups t-test with 25 observations per group and reference cell coding, assuming a one-sided test of the form $H_0 : \theta_d \geq 0$ versus $H_\mathrm{A} : \theta_d < 0$, and a one-sided confidence interval of the form $(-\infty, U]$.

```
TITLE1 "EXAMPLE07.SAS--Independent groups T-test";
TITLE2 "one sided CI (-inf., U], one sided test (H_a: theta_d < 0)";
PROC IML SYMSIZE=4000 WORKSIZE=4000;
%INCLUDE "..\IML\CISIZE01.IML" / NOSOURCE2;

CISIDE={".U"}; ALTHYP={"<0"};
ALPHA={0.05};

ESSENCEX={1 0,
          1 1};
REPN={25};

C={0 1};
BETA={0,
      1};  BETASCAL={-1};
U={1};

SIGMA=I(1); SIGSCAL={1};
DELTA={1}; DELTSCAL={1};
```

```
THETA0={0};

EVENTS={"WARGV" "WGV" "WIDTH" "REJECT"};
RUN PREVENT;
```

Note that BETASCAL is negative. For this particular combination of confidence interval and hypothesis test, a negative BETASCAL is required for the request to make sense. The alternative hypothesis being considered creates the requirement. A positive BETASCAL value in this program will result in a missing value returned for Pr$\{(W \cap R)|V\}$.

Example08.SAS

The following program calculates Pr$\{(W \cap R)|V\}$, Pr$\{W|V\}$, Pr$\{W\}$, and Pr$\{R\}$ for a paired data t-test with 25 observations, assuming a two-sided alternative hypothesis and two-sided confidence interval. It also constructs a plot with the analog of a power curve for Pr$\{(W \cap R)|V\}$ from the output. Example08.SAS differs from Example03.SAS above only by the addition of more sample sizes and the code necessary to create the plot. The resulting plot is included below the code.

```
TITLE1 "EXAMPLE08.SAS--Paired data T-test";
TITLE2 "two sided CI [L, U], two sided test (H_a: theta_d ^= 0)";
TITLE3 "create .JPEG file containing plot of data";

PROC IML SYMSIZE=4000 WORKSIZE=4000;
%INCLUDE "..\IML\CISIZE01.IML" / NOSOURCE2;
FILENAME OUT01 "..\EXAMPLE08.JPEG";
LIBNAME INOUT01 "..\DATA\";

ALPHA={0.05};
ESSENCEX={1};
REPN=DO(5,100,1);

C={1};
BETA={1}; BETASCAL={0.0625};
U={1};

SIGMA=I(1); SIGSCAL={0.012};
DELTA={1}; DELTSCAL={0.125};
THETA0={0};

EVENTS={"WARGV" "WGV" "WIDTH" "REJECT"};
RUN PREVENT;

*Next 2 lines creates a SAS data file in a specified library;
CREATE INOUT01.EXAMPLE08 FROM _HOLDPR_[COLNAME=_PRNM_];
APPEND FROM _HOLDPR_;
*End of IML code;

*Start of plot code;
SYMBOL1 I=SPLINE V=NONE L=1 W=5 COLOR=RED;

AXIS1 ORDER=(0 TO 1 BY 0.2) WIDTH=3 MINOR=NONE MAJOR=(W=3);
AXIS2 ORDER=(5 TO 50 BY 10) WIDTH=3 MINOR=NONE MAJOR=(W=3);

GOPTIONS DISPLAY GSFMODE=REPLACE
         DEVICE=JPEG TARGETDEVICE=PSLEPSFC
         CBACK=WHITE  COLORS=(BLACK)
```

```
          HSIZE=6. IN   VSIZE=3.0 IN
          HTEXT=.16666 IN  HTITLE=.16666 IN  HBY=.16666 IN
          FTEXT="TRIPLEX" FTITLE=TRIPLEX  FBY=TRIPLEX
          HORIGIN=0 IN  VORIGIN=0 IN;

GOPTIONS GSFNAME=OUT01;
PROC GPLOT DATA=INOUT01.EXAMPLE08;
     PLOT WARGV*TOTAL_N / NOFRAME
          VZERO VAXIS=AXIS1
          HZERO HAXIS=AXIS2;
LABEL TOTAL_N  = "Total sample size"
      WARGV    = "Pr{ }";
TITLE;
RUN; QUIT; RUN;
```

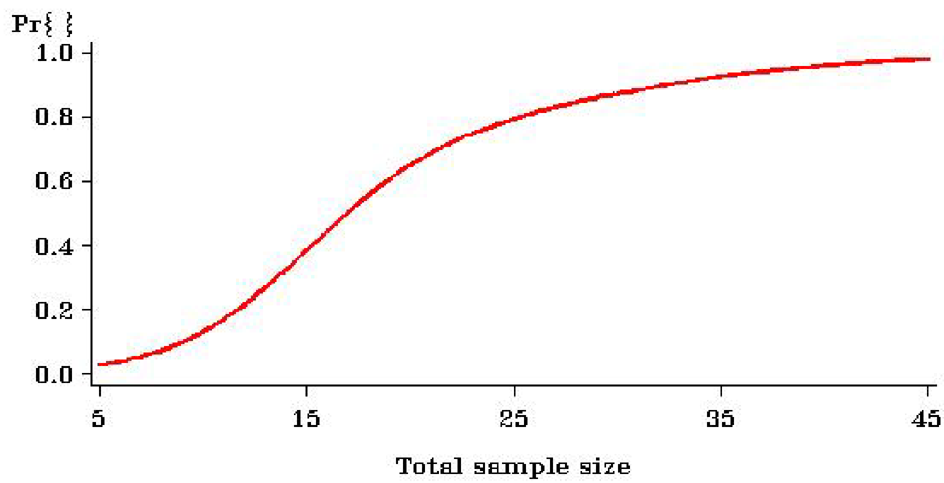

Figure 1. Example08.SAS $\Pr\{(W \cap R)|V\}$ plot.

Example09.SAS

This example utilizes the data created in Example08.SAS above to overlay $\Pr\{(W \cap R)|V\}$, $\Pr\{W|V\}$, and $\Pr\{R\}$ on one set of axes. Therefore this program assumes that a SAS file named EXAMPLE08 resides in the ../DATA/ directory. The resulting plot is included below the code.

```
TITLE1 "EXAMPLE09.SAS--Paired data T-test, EXAMPLE08 data";
TITLE2 "create .JPEG file containing overlay plot of WARGV, WGV and REJECT";

LIBNAME IN01 "..\DATA\";
FILENAME OUT01 "..\PLOTS\EXAMPLE09.JPEG";

SYMBOL1 I=SPLINE V=NONE L=1 W=5 COLOR=RED; *WARGV;
SYMBOL2 I=SPLINE V=NONE L=34  W=5 COLOR=VIOLET; *WGV;
SYMBOL3 I=SPLINE V=NONE L=41 W=5 COLOR=BLUE; *REJECT;

AXIS1 ORDER=(0 TO 1 BY 0.2) WIDTH=3 MINOR=NONE MAJOR=(W=3);
AXIS2 ORDER=(5 TO 50 BY 10) WIDTH=3 MINOR=NONE MAJOR=(W=3);

GOPTIONS DISPLAY GSFMODE=REPLACE
          DEVICE=JPEG TARGETDEVICE=PSLEPSFC
          CBACK=WHITE  COLORS=(BLACK)
```

```
          HSIZE=6. IN   VSIZE=3.0 IN
          HTEXT=.16666 IN  HTITLE=.16666 IN  HBY=.16666 IN
          FTEXT="TRIPLEX" FTITLE=TRIPLEX  FBY=TRIPLEX
          HORIGIN=0 IN  VORIGIN=0 IN;

GOPTIONS GSFNAME=OUT01;
PROC GPLOT DATA=IN01.EXAMPLE08;
     PLOT (WARGV WGV REJECT)*TOTAL_N / OVERLAY NOFRAME
          VZERO VAXIS=AXIS1
          HZERO HAXIS=AXIS2;
LABEL TOTAL_N  = "Total sample size"
      WARGV    = "Pr{ }";
TITLE;
RUN; QUIT; RUN;
```

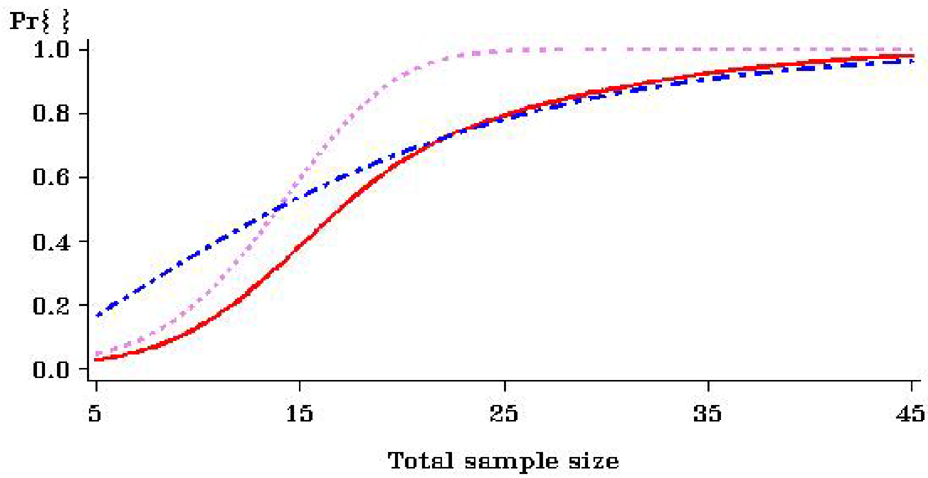

Figure 2. Example09.SAS overlay plot of Pr$\{(W \cap R)|V\}$, Pr$\{W|V\}$ and Pr$\{R\}$.
Note: Pr$\{(W \cap R)|V\}$ is red, Pr$\{W|V\}$ is violet and Pr$\{R\}$ is blue.

REFERENCES

Helms, R. W. (1988) Definitions of linear model parameters and hypotheses as functions of E(Y). *Communications in Statistics: Theory and Method*. **17**, 2715-2723.

Jiroutek, M. R., Muller, K. E., Kupper, L. L. and Stewart, P. W. (2003) A new method for choosing sample size for confidence interval-based inferences, *Biometrics*, **59**, 580-590.

Muller, K. E., and Fetterman, B. A. (2002) *Regression and ANOVA: An Integrated Approach Using SAS® Software*. Cary, NC: SAS Institute.

Muller, K. E., LaVange, L. M., Ramey, S. L., and Ramey, C. T. (1992) Power calculations for general linear multivariate models including repeated measures applications. *Journal of the American Statistical Association* **87**, 1209-1226.

O'Brien, R. G. and Muller, K. E. (1993) A unified approach to statistical power for t-tests to multivariate models. Chapter 8, *Applied Analysis of Variance in Behavioral Sciences*, L. K. Edwards, ed., 297-344. New York: Marcel Dekker.

Timm, N. H. (2002) *Applied Multivariate Analysis*. New York: Springer-Verlag.