

ESTIMATING POWER FOR TESTS OF THE MULTIVARIATE GENERAL LINEAR HYPOTHESIS
 Bercedis L. Peterson and Keith E. Muller, University of North Carolina, Chapel Hill

ABSTRACT

Although methods for exact power calculations for tests of the multivariate general linear hypothesis do not exist, convenient accurate approximations are available for three commonly used multivariate test statistics: Wilks' Lambda, Hotelling-Lawley Trace, and Pillai's Trace (Muller and Peterson, 1983). In this paper two computer programs are presented which calculate these approximate powers. One program performs "prospective" power analyses in which powers are calculated for tests on hypothesized models. The other program performs "retrospective" power calculations on observed test statistics.

INTRODUCTION

The two power programs documented in this manual perform power calculations for tests of the multivariate general linear hypotheses. For a detailed presentation of the theory of power analysis for tests of the multivariate linear hypotheses see Muller and Peterson (1983). See Cohen (1977), Koele (1982), and O'Brien (1982) for the best discussion of univariate power analysis. The two power programs will be discussed separately below, but first a clear specification of the multivariate hypothesis is required.

The general multivariate model is:

$$Y = XB + \epsilon$$

$N \times p$ $(N \times r)$ $(r \times p)$ $N \times p$

Here N is the number of observations, p the number of dependent variables and r the number of columns of the design matrix, X , which is a set of known constants. It is convenient to assume $N > p + r$. The matrix of unknown parameters is B . Each row of ϵ is assumed i.i.d. multivariate normal with mean vector zero and covariance matrix Σ . Important special cases of this model include multivariate multiple regression, canonical correlation, discriminant analysis, fixed effect MANOVA and MANCOVA, and all univariate special cases. For a general treatment of this material see Timm (1975), Finn (1974), Morrison (1967), Rao (1973), Kshirsagar (1972), Harris (1975), and Anderson (1974), among others.

The general linear hypothesis (GLH) may be stated as

$$H_0: CBU - \theta_0 = \theta = 0$$

$(a \times r)$ $(r \times p)$ $(p \times b)$ $a \times b$ $a \times b$ $a \times b$

For testing purposes, $rk(C) = asr$ and $rk(U) = bsp$ will be required. The matrix of secondary parameters, θ , contains linear combinations of the matrix of primary parameters, B . The matrix C , the usual contrast matrix in univariate linear models, computes linear combinations of rows of B , which correspond to different predictors in the design matrix. The matrix U computes linear combinations of columns of B , which correspond to different dependent variables. It is useful and accurate to think of U as transforming the dependent variables.

The model contains two parameter matrices to be estimated: B and Σ . Under the

assumptions stated, estimates are

$$\hat{B} = (X'X)^{-1}X'Y$$

$$\hat{\Sigma} = Y'[I - X(X'X)^{-1}X']Y/[N - rk(X)].$$

In turn, testing the GLH requires estimating θ :

$$\hat{\theta} = C\hat{B}U - \theta_0$$

C and U must be such that $\hat{\theta}$ is estimable (identically equal to some linear function of the expected value of Y), which guarantees invariance with respect to a particular B .

The three multivariate test statistics for which power is calculated in these two programs are based on the s eigenvalues of $H\hat{E}^{-1}$ where:

$$s = \min(rk(C), rk(U))$$

$$H = \hat{\theta}'[C(X'X)^{-1}C']^{-1}\hat{\theta}$$

$$E = U'\hat{\Sigma}U[N - rk(X)].$$

If c_i is allowed to denote the i -th eigenvalue of $H\hat{E}^{-1}$, where $c_1 \geq c_2 \geq \dots \geq c_s > 0$, then the three multivariate test statistics can be defined as below:

Hotelling-Lawley Trace $HLT = \sum_{i=1}^s c_i$

Pillai's Trace $PB = \sum_{i=1}^s \frac{c_i}{1+c_i}$

Wilks' Lambda $W = \prod_{i=1}^s \frac{1}{1+c_i}$

Alternatively, the three multivariate test statistics can be thought of as functions of the eigenvalues of $H(H+E)^{-1}$. If λ_i denotes the i -th eigenvalue of $H(H+E)^{-1}$, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_s > 0$, then the three statistics are:

$$HLT = \sum_{i=1}^s \frac{\lambda_i}{1-\lambda_i}$$

$$PB = \sum_{i=1}^s \lambda_i$$

$$W = \prod_{i=1}^s (1-\lambda_i)$$

POWER1: PROSPECTIVE

This program performs "prospective" power analysis on any number of linear hypotheses of the form $CBU - \theta_0 = 0$, corresponding to the general linear model $Y = XB + \epsilon$. By "prospective" it is meant that the program calculates

the approximate power of a planned multivariate test based upon a hypothetical matrix of regression coefficients.

POWER1 requires that the user supply an estimate of the variance-covariance matrix $\hat{\Sigma}$, an hypothesized beta matrix β , an "essence" X matrix (defined below), C , U , and alpha. Finally, the user must provide a scalar named REPN which determines the sample size to be used in the power calculations. Any one or more of these inputs may be varied within the program so as to obtain several powers within one job. For example, the user may want to vary sample size so as to determine the smallest for which satisfactory power is achieved. As another example, the user may want to vary β for a fixed sample size so as to compare the powers of important alternative hypotheses.

The entire program is written in PROC MATRIX in a form which allows it to be LINKed to the user. The program's first step is to call the CHECK module. This module checks that the matrices conform, that the sample size is large enough, and that the variance-covariance matrix is symmetric and positive semi-definite. If any check is failed, a message is printed and the program terminates immediately.

The second step of the program is to calculate the H and E matrices, which requires an "essence" X matrix [$Es(X)$] provided by the user. For equal cell sizes, $Es(X)$ has exactly one row for every unique row in X (Helms, 1980). For example, if the full X matrix represents a 3-group cell mean model, $Es(X) = I_3$, a 3x3 identity matrix. REPN specifies the number of times each row of $Es(X)$ is replicated to give \tilde{X} . If REPN = 10 for the cell mean model above, powers would be calculated for a sample size of 30 ($3*10$).

$Es(X)$ can also be used in some situations requiring unequal cell sizes. Suppose X represents a 3-group cell mean model for which cells 2 and 3 are $1\frac{1}{2}$ and 2 times the size of cell 1, respectively. Then

$$[Es(X)]' = \begin{pmatrix} 1 & & & \\ & 1 & 1 & 1 \\ & & 1 & 1 & 1 & 1 \end{pmatrix}$$

If REPN were specified to be, say 4, then power would be calculated for cell sizes of 8, 12, and 16.

Some situations disallow the use of an essence matrix. This is true, for the example, if the required cell sizes are 10, 11, and 12. In such situations, $Es(X) = X$ and REPN must be specified by the user to be 1. Another important example is a design using a continuous predictor, such as an ANCOVA.

To calculate H , an "essence" H matrix is first created by using the formula given previously for H except that the "essence" X matrix is used instead of the X matrix. The H matrix is then created by multiplying "essence" H by REPN.

The calculation of the E matrix requires a variance-covariance matrix to be computed as $U'\hat{\Sigma}U$. $U'\hat{\Sigma}U$ is a unit SSCP matrix, since it represents the variance-covariance of one element. The E matrix is then created by

multiplying $U'\hat{\Sigma}U$ by the error degrees of freedom, $N-r$. \tilde{N} is calculated by multiplying REPN times the number of rows of "essence" X . The scalar r is, of course, just the number of columns of "essence" X .

The eigenvalues of HE^{-1} are then calculated and used as input in three different modules which compute the powers of three of the most common multivariate test statistics: Hotelling-Lawley Trace, Pillai's Trace, and Wilks' Lambda. Each of these statistics is a simple function of the eigenvalues of HE^{-1} . The three modules will be described later. Powers may either be printed out or output to a SAS dataset.

Use of POWER1

The user must supply, in order, the PROC MATRIX statement, an INCLUDE statement, the seven essential matrices listed below (and two optional matrices), and one or more LINK statements.

| Matrix Name | Description | Size of Matrix |
|-------------------|---|----------------|
| ESSENCEX | Essence X matrix, described herein | r columns |
| SIGMA | Estimate of covariance matrix Σ | pxp |
| C | C matrix of contrasts across rows of β | axr, asr |
| U | U matrix of contrasts across columns of β | pxb, bsp |
| G (optional) | $\hat{\Sigma}_a$ subtracted from CBU. Default = 0. Once the default is changed, the user must always specify G. | axb |
| BETA | Hypothesized coefficients, β | rxp |
| REPN | # of times each row of ESSENCEX is duplicated | 1×1 |
| ALPHA | Alpha level, or size, of the test | 1×1 |
| T2FLAG (optional) | = 1 if exact powers are wanted for a multivariate 1- or 2-sample t test (i.e., $b=1, a=1$, and t test). Default = 0. Once the default is changed, the user must always specify T2FLAG. | 1×1 |

The INCLUDE statement attaches the POWER1 program onto the user's own program. The code inside POWER1 is then executed via LINK statements provided by the user. Of course, the user does not LINK until after supplying the seven required matrices. Three LINKs are available to the user:

- (1) LINK POWCOMP will cause the three multivariate powers to be calculated for the current values of the seven user-supplied matrices.
- (2) LINK PRINTALL will print a matrix named RESULTS in which all powers calculated within the entire job are stored.
- (3) LINK PRINTONE will print a vector named RESULTS1 which is the last row of RESULTS, i.e., that row containing the powers of current multivariate test.

The user can cause any number of power analyses to be performed by redefining one or more of the seven matrices and using LINK POWCOMP repeatedly. For example, in the job below five power analyses will be performed, one for each of the five implied sample sizes.

```
// EXEC SAS
//POWERPROG DD DSN= ... ,DISP=OLD
//SYSIN DD *
TITLE1 EXAMPLE USING THE POWER1 PROGRAM;
PROC MATRIX;
%INCLUDE POWERPROG/ NOSOURCE2;
ALPHA = .05;
ESSENCEX = I.(4);
```

```
SIGMA = (119.75  84.71  10.35 /
          84.71  88.23  19.10 /
          10.35  19.10  67.5) #/12;
```

```
U = I.(3);
C = 1  -1  0  0/
    1  0  -1  0/
    1  0  0  -1;
```

```
BETA = 6  4.5  2.5/
        6  4.5  2.5/
        6  6    6 /
        6  6    6 ;
```

```
DO REPN = 5 TO 25 BY 5;
LINK POWCOMP;
END;
```

```
LINK PRINTALL;
```

The sample sizes are determined by the number of rows of ESSENCEX (4 in this case) and by REPN. Thus the total sample sizes will be 20 = (4*5), 40 = (4*10), 60 = (4*15), 80 = (4*20), and 100 = (4*25). A partial listing of the output is shown below.

| RESULTS | N | HLT_POWR | PT_POWER | WLK_POWR |
|---------|-----|----------|----------|----------|
| | 20 | 0.462468 | 0.397913 | 0.41832 |
| | 40 | 0.920186 | 0.809258 | 0.870731 |
| | 60 | 0.994201 | 0.960999 | 0.984097 |
| | 80 | 0.999732 | 0.994113 | 0.998672 |
| | 100 | 0.999991 | 0.999289 | 0.999916 |

In the more complex job below both REPN and BETA are varied. After calculating powers for two sample sizes for the first BETA, this BETA and its corresponding powers are printed. At this point the matrix RESULTS is FREEd so that it can be built up anew for the second specification of BETA. After powers are calculated for this second BETA, it is printed along with its corresponding powers.

```
// EXEC SAS
//POWERPROG DD DSN= ... ,DISP=OLD
//SYSIN DD *
TITLE1 MORE COMPLEX EXAMPLE USING POWER1;
PROC MATRIX ;
%INCLUDE POWERPROG/ NOSOURCE2;
T2FLAG = 1;
ESSENCEX = I.(2);
```

```
SIGMA = (119.75  84.75/
          84.75  88.75) #/12;
```

```
U = I.(2);
C = 1  -1;
```

```
BETA = 6  4.5/
        6  6 ;
```

```
DO REPN = 15 TO 20 BY 5;
LINK POWCOMP;
END;
```

```
PRINT BETA;
LINK PRINTALL;
FREE RESULTS;
```

```
BETA = 4  4.5/
        4  4 ;
```

```
DO REPN = 15 TO 20 BY 5;
LINK POWCOMP;
END;
```

```
PRINT BETA;
LINK PRINTALL;
```

A partial listing of the output from this job is given below.

| | BETA | COL1 | COL2 |
|------|------|------|------|
| ROW1 | 6 | 6 | |
| ROW2 | 6 | 6 | |

| RESULTS | N | HLT_POWR | PT_POWER | WLK_POWR |
|---------|----|----------|----------|----------|
| | 30 | 0.605835 | 0.605835 | 0.605835 |
| | 40 | 0.752878 | 0.752878 | 0.752878 |

| | BETA | COL1 | COL2 |
|------|------|------|------|
| ROW1 | 4 | 4.5 | |
| ROW2 | 4 | 4 | |

| RESULTS | N | HLT_POWR | PT_POWER | WLK_POWR |
|---------|----|----------|----------|----------|
| | 30 | 0.10655 | 0.10655 | 0.10655 |
| | 40 | 0.129357 | 0.129357 | 0.129357 |

Although not illustrated in either of the two examples above, the OUTPUT statement can be used to output RESULTS to a SAS dataset. By specifying the option COLNAME = COLNAME, one obtains the proper column names.

POWER2: RETROSPECTIVE

This program performs "retrospective" power analyses on any number of multivariate linear hypotheses of the form $CBU = 0$. By "retrospective" it is meant that the hypotheses have already been tested on a sample of data, and observed test statistics have been obtained. The results from such a completed analysis can be used as input into POWER2 to obtain the approximate powers of three of the most common multivariate test statistics. As an example of the use of this program, suppose an analysis had yielded a nonsignificant result. Then POWER2 could be used by varying sample size so as to determine how large the sample should have been to achieve satisfactory power.

The entire POWER2 program is written in PROC matrix. The first part of the program defines matrices and scalars based on the user-provided input. The program's next

step is to call the CHECK module via the LINK statement. This module checks, for example, that the user-supplied eigenvalues are non-negative and that the sample size and ranks of the matrices are positive integers. If the input associated with a multivariate test contains any invalid values, all three powers for that test are set to missing, and a warning message is printed. The third step in the program is to LINK to three modules, described below, which take as input the eigenvalues of HE^{-1} and which produce powers. The approximate powers thus produced are those for Hotelling-Lawley Trace, Pillai's Trace, and Wilks' Lambda. At the end of the program, these powers are printed out, along with the user-supplied parameters associated with them.

Use of POWER2

The user must supply, in order, a PROC MATRIX statement, an INCLUDE statement, a matrix named POWERDAT (described below) and a LINK POWCOMP statement. The INCLUDE statement is used to attach the POWER2 program onto the user's own program. The matrix named POWERDAT may either be created directly or by FETCHing a SAS dataset. The LINK POWCOMP statement will cause the POWER2 code to be executed.

The matrix POWERDAT has one row for each multivariate for which power is required. The columns of POWERDAT must be (in order):

Variable
Position

- | Variable Position | Variable |
|-------------------|---|
| 1 | A scalar which is equal to 1 if exact powers are wanted for a multivariate one- or two-sample t test (i.e., Rank(U)>1, Rank(C)=1, and t test). Otherwise, this variable is equal to 0. |
| 2 | Total sample size |
| 3 | Rank of the X matrix |
| 4 | Rank of the C matrix |
| 5 | Rank of the U matrix |
| 6 | Alpha level of the test |
| 7+ | The s eigenvalues of HE^{-1} , where $s = \min(\text{Rank}(C), \text{Rank}(U))$. For $s=1$, the eigenvalue may be obtained from the observed F statistic (which has freedom DF1 and DF2) as: $F*(DF1/DF2)$. |

Note that if the number of eigenvalues is not the same for all tests, the user must fill out the last eigenvalue positions with zeros. This is illustrated in the two examples below.

In the first example POWERDAT is created within PROC MATRIX. In the second example, POWERDAT is created by FETCHing in a SAS dataset.

Example 1:

```
// EXEC SAS
//POWRPROG DD DSN= . . .,DISP=OLD
//SYSIN DD *
TITLE1 EXAMPLE USING THE POWER2 PROGRAM;
TITLE2 A MATRIX IS CREATED DIRECTLY;
```

```
PROC MATRIX;
%INCLUDE POWRPROG/NOSOURCE2;
POWERDAT =
      (1 48 2 1 3 .05 .111 0 0/
       0 32 4 3 3 .05 .625 .131 .008);
LINK POWCOMP;
```

Example 2:

```
// EXEC SAS
//POWRPROG DD DSN= . . .,DISP=OLD
//DATAIN DD *
1 48 2 1 3 .05 .111 0 0
D 32 4 3 3 .05 .625 .131 .008
//SYSIN DD *
TITLE1 EXAMPLE USING THE POWER2 PROGRAM;
TITLE2 A DATASET IS INPUT;
DATA POWERDAT;
INFILE DATAIN;
INPUT FLAG N X C U ALPHA E1 E2 E3;

PROC MATRIX;
%INCLUDE POWRPROG/NOSOURCE2;
FETCH POWERDAT DATA=POWERDAT;
LINK POWCOMP;
```

For both of the above examples, the output from POWER2 will include:

| RESULTS | HLT_POWR | PT_POWER | WLK_POWR | N |
|---------|----------|----------|----------|----|
| | 0.414279 | 0.414279 | 0.414279 | 48 |
| | 0.834971 | 0.796516 | 0.812892 | 32 |

If the user wishes, the OUTPUT statement can be used to output _RESULTS to a SAS dataset. By specifying the option COLNAME = _COLNAME, one obtains the proper column names.

THE _HLTLOOP, _PTLOOP, AND _WLKLOOP MODULES

These modules calculate asymptotic powers of the three multivariate test statistics. All three modules follow an algorithm based on two important results. One, approximations to the central distributions of the three test statistics have been suggested using F distributions. All approximations were based on equating first and second moments. Recently Nagarsenker and Suniaga (1983) demonstrated that the F approximation for Wilks' lambda was asymptotically correct. Two, the noncentral distributions of the three statistics are a suggestion of Muller and Peterson (1983). They suggest using non-central F distributions with the same degrees of freedom as for the central case and with noncentrality parameter equal to the observed approximate F statistic times its numerator degrees of freedom.

The modules operate by first calculating a multivariate test statistic as a function of the eigenvalues. Then the degrees of freedom and the noncentrality parameter of the F-statistic are calculated. A critical value from the central F-distribution is obtained by using the SAS function FINV. This critical value along with the degrees of freedom and the noncentrality parameter are used in the SAS function FPROB to get power.

A listing of these two power programs is available from the authors.

REFERENCES

- Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, 1974, 203-210.
- Cohen, J. Statistical Power Analysis for the Behavioral Sciences. New York: Academic Press, 1977 (rev. ed).
- Finn, J. D. A General Model for Multivariate Analysis. New York: Holt, Rinehart and Winston, 1974.
- Harris, R. J. A Primer of Multivariate Statistics. New York: Academic Press, 1975, 314-315.
- Helms, R. W. Definitions of primary parameters in linear models. Paper presented at the meeting of the American Statistical Association. Proceedings of the Statistical Computing Section, 1980, 105-113.
- Koele, P. Calculating power in analysis of variance. Psychological Bulletin, 1982, 92, 513-516.
- Kshirsagar, A. M. Multivariate Analysis. New York: Marcel Dekker, 1972, 291-304.
- Morrison, D. F. Multivariate Statistical Methods. New York: McGraw-Hill, 1967.
- Muller, K. E. and Peterson, B. L. Practical methods for computing power in testing the multivariate general linear hypothesis. Paper presented at the meeting of the American Statistical Association, Proceedings of the Statistical Computing Section, 1983, (in press).
- Nagarsenker, B. N. and Suniaga, J. Distribution of a class of statistics useful in multivariate analysis. Journal of the American Statistical Association, 1983, 78, 472-475.
- O'Brien, R. G. Performing power sensitivity analyses on general linear model hypotheses. Paper presented at the meeting of the American Statistical Association, Proceedings of the Statistical Computing Section, 1982, 114-118.
- Rao, C. R. Linear Statistical Inference and Its Applications (2nd ed). New York: John Wiley and Sons, 1973. 533-543; 555-556.