MANOVA

The program performs univariate and multivariate analysis of variance and covariance for any crossed and/or nested design.

Analysis of Variance

Notation

The experimental design model (the model with covariates will be discussed later) can be expressed as

$$\underset{N \times p}{\mathbf{Y}} = \underset{N \times m}{\mathbf{W}} \underset{m \times p}{\beta} + \underset{N \times p}{\mathbf{E}}$$

where

$\mathbf{Y}$	is the observed matrix
$\mathbf{W}$	is the design matrix
β	is the matrix of parameters
$\mathbf{E}$	is the matrix of random errors
N	is the total number of observations
p	is the number of dependent variables
m	is the number of parameters

Since the rows of **W** will be identical for all observations in the same cell, the model is rewritten in terms of cell means as

$$\underset{g \times p}{\mathbf{Y}_{\bullet}} = \underset{g \times m}{\mathbf{A}} \underset{m \times p}{\beta} + \underset{g \times p}{\mathbf{E}_{\bullet}}$$

where g is the number of cells and $\mathbf{Y}_{\bullet}$ and $\mathbf{E}_{\bullet}$ denote matrices of means.

Reparameterization

The reparameterization of the model (Bock, 1975; Finn, 1977) is done by factoring **A** into

$$\underset{g\times m}{\mathbf{A}} = \underset{g\times r}{\mathbf{K}} \; \underset{r\times m}{\mathbf{L}}$$

K forms a column basis for the model and has rank *r*. **L** contains the coefficients of linear combinations of parameters and rank *r*. The contrast matrix **L** can be specified by the user. Given **L**, **K** can be obtained from $\mathbf{AL}'(\mathbf{LL}')^{-1}$. For designs with more than one factor, **L**, and hence **K**, can be constructed from Kronecker products of contrast matrices of each factor. After reparameterization, the model can be expressed as

$$\begin{aligned}\underset{g\times p}{\mathbf{Y}} &= \mathbf{A}\beta + \mathbf{E} \\ &= \mathbf{K}(\mathbf{L}\beta) + \mathbf{E} \\ &= \underset{g\times r}{\mathbf{K}} \; \underset{r\times p}{\theta} + \underset{g\times p}{\mathbf{E}}\end{aligned}$$

Parameter Estimation

An orthogonal decomposition (Golub, 1969) is performed on **K**. That is, **K** is represented as

$$\mathbf{K} = \mathbf{QR}$$

where **Q** is an orthonormal matrix such that $\mathbf{Q}'\mathbf{DQ} = \mathbf{I}$; **D** is the diagonal matrix of cell frequencies; and **R** is an upper-triangular matrix.

The normal equation of the model is

$$(\mathbf{K}'\mathbf{DK})\hat{\theta} = \mathbf{K}'\mathbf{DY}$$

or

$$\mathbf{R}\hat{\theta} = \mathbf{Q}'\mathbf{DY} = \mathbf{U}$$

This triangular system can therefore be solved forming the cross-product matrix.

Significance Tests

The sum of squares and cross-products (SSCP) matrix due to the model is

$$\hat{\theta}'\mathbf{R}'\mathbf{R}\hat{\theta} = \mathbf{U}'\mathbf{U}$$

and since $\operatorname{var}(\mathbf{U}) = \mathbf{R}\ \operatorname{var}(\theta)\mathbf{R}' = \mathbf{I} \otimes \Sigma$ the SSCP matrix of each individual effect can be obtained from the components of

$$\mathbf{U}'\mathbf{U} = (U_1, \ldots, U_k)\begin{pmatrix} U_1' \\ \vdots \\ U_k' \end{pmatrix} = U_1 U_1' + \ldots + U_k U_k'$$

Therefore the hypothesis SSCP matrix for testing $H_o : \theta_h = \mathbf{0}$ is

$$\underset{p \times p}{\mathbf{S}_H} = \underset{p \times n_h}{\mathbf{U}_h} \ \underset{n_h \times p}{\mathbf{U}_h'}$$

The default error SSCP matrix is the pooled within-groups **SSCP**:

$$\mathbf{S}_E = \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{D}\mathbf{Y}$$

if the pooled within-groups SSCP matrix does not exist, the residual SSCP matrix is used:

$$\mathbf{S}_E = \mathbf{Y}'\mathbf{Y} - \mathbf{U}'\mathbf{U}$$

Four test criteria are available. Each of these statistics is a function of the nonzero eigenvalues λ_i of the matrix $\mathbf{S}_H\mathbf{S}_E^{-1}$. The number of nonzero eigenvalues, *s*, is equal to $\min(p, n_h)$.

Pillai's Criterion (Pillai, 1967)

$$T = \sum_{i=1}^{s} \lambda_i \Big/ (1 + \lambda_i)$$

Approximate $F = (n_e - p - s)T \Big/ (b(s - T))$ with b_s and $s(n_e - p + s)$ degrees of freedom, where

$n_e =$ degrees of freedom for S_E
$b = \max(p, n_h)$

Hotelling's Trace

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

Approximate $F = 2(sn + 1)T \Big/ \left(s^2 (2m + s + 1)\right)$ with $s(2m + s + 1)$ and $2(sn + 1)$ degrees of freedom where

$m = (|n_h - p| - 1) \Big/ 2$
$n = (n_e - p - 1) \Big/ 2$

Wilks' Lambda (Rao, 1973)

$$T = \prod_{i=1}^{s} 1 \Big/ (1 + \lambda_i)$$

Approximate $F = \left(1 - T^{1/l}\right)(Ml + 1 - n_h p/2) \Big/ \left(T^{1/l} n_h p\right)$ with $n_h p$

and $(Ml + 1 - n_h p/2)$ degrees of freedom, where

$$l^2 = \left(p^2 n_h^2 - 4\right) \Big/ \left(p^2 + n_h^2 - 5\right)$$
$$M = n_e - (p + 1 - n_h)/2$$

Roy's Largest Root

$$T = \lambda_1 \big/ (1 + \lambda_1)$$

Stepdown *F* Tests

The stepdown *F* statistics are

$$F_i = \frac{\left(t^2 - t_e^2\right) \big/ n_h}{t_e^2 \big/ (n_e - i + 1)}$$

with n_h and $n_e - i + 1$ degrees of freedom, where t_e and t are the ith diagonal element of $\mathbf{T}_E$ and $\mathbf{T}$ respectively, and where

$$\mathbf{S}_E = \mathbf{T}_E' \mathbf{T}_E$$
$$\mathbf{S}_E + \mathbf{S}_H = \mathbf{T}'\mathbf{T}$$

Design Matrix

$\mathbf{K}$

Estimated Cell Means

$$\hat{\mathbf{Y}}_{\bullet} = \mathrm{K}\hat{\theta}$$

Analysis of Covariance

Model

$$\underset{g\times p}{\mathbf{Y}_\bullet} = \underset{g\times r}{\mathbf{K}} \underset{r\times p}{\theta} + \underset{g\times q}{\mathbf{X}_\bullet} \underset{q\times p}{\mathbf{B}} + \underset{g\times p}{\mathbf{E}_\bullet}$$

where *g*, *p*, and *r* are as before and *q* is the number of covariates, and $\mathbf{X}_\bullet$ is the mean of **X**, the matrix of covariates.

Parameter Estimation and Significance Tests

For purposes of parameter estimation, no initial distinction is made between dependent variables and covariates.

Let

$$\mathbf{V} = (\mathbf{Y}\,\mathbf{X})$$
$$\mathbf{V}_\bullet = (\mathbf{Y}_\bullet\,\mathbf{X}_\bullet)$$

The normal equation of the model

$$\underset{g\times(p+q)}{\mathbf{V}_\bullet} = \underset{g\times r}{\mathbf{K}} \underset{r\times(p+q)}{\theta} + \underset{g\times(p+q)}{\mathbf{E}_\bullet}$$

is

$$(\mathbf{K}'\mathbf{D}\mathbf{K})\hat{\theta} = \mathbf{K}'\mathbf{D}\mathbf{Y}_\bullet$$

or

$$\mathbf{R}\hat{\theta} = \mathbf{Q}'\mathbf{D}\mathbf{V}_\bullet = \mathbf{U}$$

or

$$\underset{r\times(p+q)}{\hat{\theta}} = \begin{pmatrix} \underset{r\times p}{\hat{\theta}_Y} & \underset{r\times q}{\hat{\theta}_X} \end{pmatrix}$$

If $\mathbf{S}_E$ and $\mathbf{S}_T$ are partitioned as

$$\mathbf{S}_E = \begin{pmatrix} \mathbf{S}_E^{(Y)} & \mathbf{S}_E^{(YX)} \\ \mathbf{S}_E^{(XY)} & \mathbf{S}_E^{(X)} \end{pmatrix}$$

$$\mathbf{S}_T = \begin{pmatrix} \mathbf{S}_T^{(Y)} & \mathbf{S}_T^{(YX)} \\ \mathbf{S}_T^{(XY)} & \mathbf{S}_T^{(X)} \end{pmatrix}$$

then the adjusted error SSCP matrix is

$$\mathbf{S}_E^* = \mathbf{S}_E^{(Y)} - \mathbf{S}_E^{(YX)} \left(\mathbf{S}_E^{(X)} \right)^{-1} \mathbf{S}_E^{(XY)}$$

and the adjusted total SSCP matrix is

$$\mathbf{S}_T^* = \mathbf{S}_T^{(Y)} - \mathbf{S}_T^{(YX)} \left(\mathbf{S}_T^{(X)} \right)^{-1} \mathbf{S}_T^{(XY)}$$

The adjusted hypothesis SSCP matrix is then

$$\mathbf{S}_H^* = \mathbf{S}_T^* - \mathbf{S}_E^*$$

The estimate of **B** is

$$\hat{\mathbf{B}} = \left(\mathbf{S}_T^{(X)} \right)^{-1} \mathbf{S}_T^{(XY)}$$

The adjusted parameter estimates are

$$\hat{\theta}^* = \hat{\theta}_Y - \hat{\theta}_X \hat{\mathbf{B}}$$

The adjusted cell means are

$$\hat{\mathbf{Y}}^* = \mathbf{K}\hat{\theta}^*$$

Repeated Measures

Notation

The following notation is used within this section unless otherwise stated:

k	Degrees of freedom for the within-subject factor
$\mathbf{SSE}^*$	Orthonormal transformed error matrix
N	Total number of observations
ndfb	Degrees of freedom for all between-subject factors (including the constant)

Statistics

Greenhouse-Geisser Epsilon

$$\text{ggeps} = \frac{\left(\text{tr}\left(\mathbf{SSE}^*\right)\right)^2}{k \times \text{tr}\left(\left(\mathbf{SSE}^*\right)^2\right)}$$

Huynh-Feldt Epsilon

$$\text{hfeps} = \frac{N \times k \times \text{ggeps} - 2}{k \times \left(N - \text{ndfb}\right) - k^2 \times \text{ggeps}}$$

if $\text{hfeps} > 1$, set $\text{hfeps} = 1$

Lower bound Epsilon

$$\text{lbeps} = \frac{1}{k}$$

Effect Size

Notation

The following notation is used within this section unless otherwise stated:

dfh	Hypothesis degrees of freedom
dfe	Error degrees of freedom
F	*F* test
W	Wilks' lambda
s	Number of non-zero eigenvalues of $\mathbf{HE}^{-1}$
T	Hotelling's trace
V	Pillai's trace

Statistic

$$\text{Partial eta-squared} = \frac{dfh \times F}{dfh \times F + dfe} = \frac{\text{SS hyp}}{\text{SS hyp} + \text{SS error}}$$

$$\text{Eta} - \text{squared}(\text{Wilks'}) = 1 - W^{1/s}$$

$$\text{Eta} - \text{squared}(\text{Hotelling's}) = \frac{T/s}{T/s + 1}$$

$$\text{Total eta-squared } = \frac{\text{sum of squares for effect}}{\text{total (corrected) sum of squares}}$$

$$\text{Hay's omega-squared} = \frac{\text{SS for effect} - \text{df}(\text{effect}) \times \text{MSE}}{\text{corrected total SS} + \text{MSE}}$$

$$\text{Pillai} = V/S$$

Power

Univariate Non-Centrality

$$\lambda = \frac{\text{SS hyp}}{\text{SS error}} \times dfe$$

Multivariate Non-Centrality

For a single degree of freedom hypothesis

$$\lambda = T \times dfe$$

where T is Hotelling's trace and dfe is the error degrees of freedom. Approximate power non-centrality based on Wilks' lambda is

$$\lambda = \frac{\text{Wilks' eta square}}{1 - \text{Wilks' eta square}} \times dfe(W)$$

where $dfe(W)$ is the error *df* from Rao's *F*-approximation to the distribution of Wilks' lambda.

Hotelling's Trace

$$\lambda = \frac{\text{Hotelling's eta square}}{1 - \text{Hotelling's eta square}} \times dfe(H)$$

where $dfe(H)$ is the error *df* from the *F*-approximation to the distribution of Hotelling's trace.

Pillai's Trace

$$\lambda = \frac{\text{Pillai's eta square}}{1 - \text{Pillai's eta square}} \times dfe(P)$$

where $dfe(P)$ is the error *df* from Pillai's *F*-approximation to the distribution of Pillai's trace.

Approximate Power

Approximate power is computed using an Edgeworth Series Expansion (Mudholkar, Chaubey, and Lin, 1976).

$$r = \upsilon_1 + \lambda$$
$$b = \lambda / r$$

$$K_1 = \left\{ \left(\frac{r}{\upsilon_1} \right)^{1/3} \left(1 - \frac{2(b+1)}{9r} - \frac{40b^2}{3^4 r^2} + \frac{80\left(1 + 3b + 33b^2 - 77b^3\right)}{3^7 r^3} + \frac{176\left(1 + 4b - 210b^2 + 2380b^3 - 2975b^4\right)}{3^9 r^4} \right) \right\}$$
$$- c^{1/3} \left\{ \left(1 - \frac{2}{9\upsilon_2} + \frac{80}{3^7 \upsilon_2^3} + \frac{176}{3^9 \upsilon_2^4} \right) \right\}$$

$$K_2 = \left\{ \left(\frac{r}{\upsilon_1} \right)^{2/3} \left(\frac{2(b+1)}{9r} + \frac{16b^2}{3^3 r^2} - \frac{8\left(13 + 39b + 405b^2 - 1025b^3\right)}{3^7 r^3} + \frac{160\left(1 + 4b - 87b^2 + 1168b^3 - 1544b^4\right)}{3^8 r^4} \right) \right\}$$
$$+ c^{2/3} \left(\frac{2}{9\upsilon_2} - \frac{104}{3^7 \upsilon_2^3} - \frac{160}{3^8 \upsilon_2^4} \right)$$

$$K_3 = \left\{ \left(\frac{-r}{\upsilon_1} \right) \left(\frac{8b^2}{27r^2} - \frac{32\left(1 + 3b + 21b^2 - 62b^3\right)}{3^6 r^3} - \frac{32\left(8 + 32b - 177b^2 + 4550b^3 - 6625b^4\right)}{3^8 r^4} \right) \right\}$$
$$- c \left(\frac{32}{3^6 \upsilon_2^3} + \frac{256}{3^8 \upsilon_2^4} \right)$$

$$K_4 = \left\{ \left(\frac{r}{\upsilon_1} \right)^{4/3} \left(\frac{16\left(1 + 3b + 12b^2 - 44b^3\right)}{3^6 r^3} + \frac{256\left(1 + 4b + 6b^2 + 247b^3 - 458b^4\right)}{3^8 r^4} \right) \right\}$$

$$- c^{4/3} \left(\frac{16}{3^6 \upsilon_2^3} + \frac{256}{3^8 \upsilon_2^4} \right)$$

$$Y = \frac{K_1}{\sqrt{K_2}}$$

$$\text{Power} = 1 - \Phi(Y) - \frac{1}{\sqrt{2\pi}} e^{-Y^2/2} \left\{ \frac{K_3}{6}\left(Y^2 - 1\right) + \frac{K_4}{24}\left(Y^3 - 3Y\right) \frac{K_1^2}{72}\left(Y^5 - 10Y^3 + 15Y\right) \right\}$$

Joint and Individual Confidence Intervals

The intervals are calculated as follows:

Lower bound = parameter estimate $-k^*$ stderr

Upper bound = parameter estimate $+k^*$ stderr

where *stderr* is the standard error of the parameter estimate, and *k* is the critical constant whose value depends upon the type of confidence interval requested.

Univariate Intervals

Individual Confidence Intervals

$$k = \sqrt{\left(F\left(a;1,ne\right)\right)}$$

where

ne is the error degrees of freedom

a is the confidence level desired

F is the percentage point of the *F* distribution

Joint Confidence Intervals

For Scheffé intervals:

$$k = \sqrt{(nh * F(a; nh, ne))}$$

where

ne is the error degrees of freedom

nh is the hypothesis degrees of freedom

a is the confidence level desired

F is the percentage point of the *F* distribution

For Bonferroni intervals:

$$k = t(a/(2 * nh), ne)$$

where

ne is the error degrees of freedom

nh is the hypothesis degrees of freedom

a is 100 minus the confidence level desired

F is the percentage point of Student's *t* distribution

Multivariate Intervals

The value of the multipliers $\underline{k}$ for the multivariate case is computed as follows:

Let

$p =$ the number of dependent variables
$nh =$ the hypothesis degrees of freedom
$ne =$ the error degrees of freedom
$a =$ the desired confidence level
$s = \min(p, nh)$
$m = (|nh - p| - 1)/2$
$n = (ne - p - 1)/2$

For Roy's largest root, define

$$c = G/(1-G)$$

where

$$\mathrm{G} = \mathrm{GCR}(a;\, s, m, n)$$, the percentage point of the largest root distribution

For Wilks' lambda, define

$$t = (p * nh)^2 - 4$$
$$b = p * p + nh * nh - 5$$
$$r = \sqrt{(t/b)} \text{ if } b \neq 0, \quad \text{else } r = 1$$
$$u = (p * nh - 2)/4$$
$$t = p * nh$$
$$b = (nh + ne - (p + nh + 1)/2) * r - 2 * u$$
$$f = (t * F(a;\, t, b))/b$$
$$W = (1/(1+c))^r$$
$$c = (1-W)/W$$

For Hotelling's trace, define

$$t = s(2m + s + 1)$$
$$b = 2(sn + 1)$$
$$T = (stF(a;\, t, b))/b$$
$$c = T$$

For Pillai's trace, define

$$t = s(\max(p, nh))$$
$$b = s(ne - p + s)$$
$$D = (F(a; t, b)t)/b$$
$$V = (sc)/(c+1)$$
$$c = V/(1-V)$$

Now for each of the above criteria, the critical value is

$$K = \sqrt{(ne * c)}$$

For Bonferroni intervals,

$$K = t(a/(2p(nh)); ne)$$

where *t* is the percentage point of the Student's *t* distribution.

Regression Statistics

Correlation between independent variables and predicted dependent variables

$$r\left(X_i, \hat{Y}_j\right) = \frac{r_{ij}}{R_j}$$

where

X_i = *i*th predictor (covariate)
$\hat{Y}_j$ = *j*th predicted dependent variable
r_{ij} = correlation between *i*th predictor and *j*th dependent variable
R_j = multiple *R* for *j*th dependent variable across all predictors

References

Bock, R. D. 1975. *Multivariate statistical methods in behavioral research.* New York: McGraw-Hill.

Finn, J. D. 1977. Multivariate analysis of variance and covariance, In: *Statistical Methods for Digital Computers*, Volume 3, K. Enslein, A. Ralston, and H. Wilf, eds. New York: John Wiley & Sons, Inc.

Golub, G. H. 1969. Matrix decompositions and statistical calculations. In: *Statistical Computation*, R. C. Milton and J. A. Nelder, eds. New York: Academic Press.

Green, P. 1977. *Analyzing multivariate data.* New York: John Wiley & Sons, Inc.

Mudholkar, G. S., Chaubrey, Y. P., and Lin, C. 1976b. Some Approximations for the noncentral F-distribution. *Technometrics,* 18:351–358.

Muller, K. E., and Peterson, B. L. 1984. Practical methods for computing power in testing the multivariate general linear hypothesis. *Computational Statistics and Data Analysis,* 2: 143–158.

Pillai, K. C. S. 1967. Upper percentage points of the largest root of a matrix in multivariate analysis. *Biometrika,* 54: 189–194.

Timm, N. H. 1975. *Multivariate analysis with applications in education and psychology*. Monterey, Calif.: Brooks and Cole.