manova — Multivariate analysis of variance and covariance

Description

The manova command fits multivariate analysis-of-variance (MANOVA) and multivariate analysis-of-covariance (MANCOVA) models for balanced and unbalanced designs, including designs with missing cells, and for factorial, nested, or mixed designs, or designs involving repeated measures.

The mvreg command (see [MV] **mvreg**) will display the coefficients, standard errors, etc., of the multivariate regression model underlying the last run of manova.

See [R] **anova** for univariate ANOVA and ANCOVA models. See [MV] **mvtest covariances** for Box's test of MANOVA's assumption that the covariance matrices of the groups are the same, and see [MV] **mvtest means** for multivariate tests of means that do not make this assumption.

Quick start

One-way MANOVA model of y1 and y2 for factor a

```
manova y1 y2 = a
```

Two-way full-factorial MANOVA model for factors a and b

```
manova y1 y2 = a##b
```

Add continuous covariate x1 for a MANCOVA model

```
manova y1 y2 = a##b c.x1
```

MANOVA model with factor b nested within a

```
manova y1 y2 = a / b|a /
```

Menu

Statistics > Multivariate analysis > MANOVA, multivariate regression, and related > MANOVA

Syntax

`manova` *depvarlist* = *termlist* [*if*] [*in*] [*weight*] [, *options*]

where *termlist* is a factor-variable list (see **[U] 11.4.3 Factor variables**) with the following additional features:

- Variables are assumed to be categorical; use the `c.` factor-variable operator to override this.
- The `|` symbol (indicating nesting) may be used in place of the `#` symbol (indicating interaction).
- The `/` symbol is allowed after a *term* and indicates that the following *term* is the error term for the preceding *terms*.

options	Description
Model	
`noconstant`	suppress constant term
`dropemptycells`	drop empty cells from the design matrix

`bootstrap`, `by`, `collect`, `jackknife`, and `statsby` are allowed; see **[U] 11.1.10 Prefix commands**.

Weights are not allowed with the `bootstrap` prefix; see **[R] bootstrap**.

`aweight`s are not allowed with the `jackknife` prefix; see **[R] jackknife**.

`aweight`s and `fweight`s are allowed; see **[U] 11.1.6 weight**.

See **[U] 20 Estimation and postestimation commands** for more capabilities of estimation commands.

Options

Model

`noconstant` suppresses the constant term (intercept) from the model.

`dropemptycells` drops empty cells from the design matrix. If `c(emptycells)` is set to `keep` (see **[R] set emptycells**), this option temporarily resets it to `drop` before running the MANOVA model. If `c(emptycells)` is already set to `drop`, this option does nothing.

Remarks and examples

Remarks are presented under the following headings:

Introduction

MANOVA is a generalization of ANOVA allowing multiple dependent variables. Several books discuss MANOVA, including Anderson (2003); Mardia, Kent, and Taylor (2024); Morrison (2005a); Rencher (1998); Rencher and Christensen (2012); Seber (1984); and Timm (1975). Introductory articles are provided by Pillai (1985) and Morrison (2005b). Pioneering work is found in Wilks (1932), Pillai (1955), Lawley (1938), Hotelling (1951), and Roy (1939).

Four multivariate statistics are commonly computed in MANOVA: Wilks's lambda, Pillai's trace, Lawley–Hotelling trace, and Roy's largest root. See *Methods and formulas* for details.

Why four statistics? Arnold (1981), Rencher (1998), Rencher and Christensen (2012), Morrison (2005b), Pillai (1985), and Seber (1984) provide guidance. All four tests are admissible, unbiased, and invariant. Asymptotically, Wilks's lambda, Pillai's trace, and the Lawley–Hotelling trace are the same, but their behavior under various violations of the null hypothesis and with small samples is different. Roy's largest root is different from the other three, even asymptotically.

None of the four multivariate criteria appears to be most powerful against all alternative hypotheses. For instance, Roy's largest root is most powerful when the null hypothesis of equal mean vectors is violated in such a way that the mean vectors tend to lie in one line within p-dimensional space. For most other situations, Roy's largest root performs worse than the other three statistics. Pillai's trace tends to be more robust to nonnormality and heteroskedasticity than the other three statistics.

The `#` symbol indicates interaction. The `|` symbol indicates nesting (`a|b` is read "`a` is nested within `b`"). A `/` between *terms* indicates that the *term* to the right of the slash is the error term for the *terms* to the left of the slash.

One-way MANOVA

A one-way MANOVA is obtained by specifying the dependent variables followed by an equal sign, followed by the categorical variable defining the groups.

▷ Example 1: One-way MANOVA with balanced data

Rencher and Christensen (2012, 183–186) presents an example of a balanced one-way MANOVA by using data from Andrews and Herzberg (1985, 357–360). The data from eight trees from each of six apple tree rootstocks are from table 6.2 of Rencher and Christensen (2012). Four dependent variables are recorded for each tree: trunk girth at 4 years (mm × 100), extension growth at 4 years (m), trunk girth at 15 years (mm × 100), and weight of tree above ground at 15 years (lb × 1000). The grouping variable is `rootstock`, and the four dependent variables are `y1`, `y2`, `y3`, and `y4`.

```
. use https://www.stata-press.com/data/r19/rootstock
(Table 6.2. Rootstock data, Rencher and Christensen (2012))

. describe

Contains data from https://www.stata-press.com/data/r19/rootstock.dta
 Observations:            48                  Table 6.2. Rootstock data,
                                                Rencher and Christensen (2012)
    Variables:             5                  30 Aug 2024 14:00
                                              (_dta has notes)
--------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
--------------------------------------------------------------------------------
rootstock       byte    %9.0g
y1              float   %4.2f                 Trunk girth at 4 years (mm x 100)
y2              float   %5.3f                 Extension growth at 4 years (m)
y3              float   %4.2f                 Trunk girth at 15 years (mm x 100)
y4              float   %5.3f                 Weight of tree above ground at 15
                                                years (lb x 1000)
--------------------------------------------------------------------------------
Sorted by:

. list in 7/10

     +-----------------------------------------+
     | rootst~k     y1      y2     y3      y4  |
     |-----------------------------------------|
  7. |        1   1.11   3.211   3.98   1.209  |
  8. |        1   1.16   3.037   3.62   0.750  |
  9. |        2   1.05   2.074   4.09   1.036  |
 10. |        2   1.17   2.885   4.06   1.094  |
     +-----------------------------------------+
```

There are six rootstocks and four dependent variables. We test to see if the four-dimensional mean vectors of the six rootstocks are different. The null hypothesis is that the mean vectors are the same for the six rootstocks. To obtain one-way MANOVA results, we type

```
. manova y1 y2 y3 y4 = rootstock

                           Number of obs =         48

                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df   F(df1,      df2) =   F   Prob>F
             ------------+----------------------------------------------------------
               rootstock | W   0.1540          5     20.0     130.3     4.94 0.0000 a
                         | P   1.3055                20.0     168.0     4.07 0.0000 a
                         | L   2.9214                20.0     150.0     5.48 0.0000 a
                         | R   1.8757                 5.0      42.0    15.76 0.0000 u
                         |----------------------------------------------------------
                Residual |                    42
             ------------+----------------------------------------------------------
                   Total |                    47
             -----------------------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

All four multivariate tests reject the null hypothesis, indicating some kind of difference between the four-dimensional mean vectors of the six rootstocks.

Let's examine the output of `manova`. Above the table, it lists the number of observations used in the estimation. It also gives a key indicating that `W` stands for Wilks's lambda, `P` stands for Pillai's trace, `L` stands for Lawley–Hotelling trace, and `R` indicates Roy's largest root.

The first column of the table gives the source. Here we are testing the `rootstock` term (the only term in the model), and we are using residual error for the denominator of the test. Four lines of output are presented for `rootstock`, one line for each of the four multivariate tests, as indicated by the `W`, `P`, `L`, and `R` in the second column of the table.

The next column gives the multivariate statistics. Here Wilks's lambda is 0.1540, Pillai's trace is 1.3055, the Lawley–Hotelling trace is 2.9214, and Roy's largest root is 1.8757. Some authors report λ_1 and others (including Rencher and Christensen) report $\theta = \lambda_1/(1+\lambda_1)$ for Roy's largest root. Stata reports λ_1.

The column labeled "df" gives the hypothesis degrees of freedom, the residual degrees of freedom, and the total degrees of freedom. These are just as they would be for an ANOVA. Because there are six rootstocks, we have 5 degrees of freedom for the hypothesis. There are 42 residual degrees of freedom and 47 total degrees of freedom.

The next three columns are labeled "F(df1, df2) = F ", and for each of the four multivariate tests, the degrees of freedom and F statistic are listed. The following column gives the associated p-values for the F statistics. Wilks's lambda has an F statistic of 4.94 with 20 and 130.3 degrees of freedom, which produces a p-value small enough that 0.0000 is reported. The F statistics and p-values for the other three multivariate tests follow on the three lines after Wilks's lambda.

The final column indicates whether the F statistic is exactly F distributed, is approximately F distributed, or is an upper bound. The letters `e`, `a`, and `u` indicate these three possibilities, as described in the footer at the bottom of the table. For this example, the F statistics (and corresponding p-values) for Wilks's lambda, Pillai's trace, and the Lawley–Hotelling trace are approximate. The F statistic for Roy's largest root is an upper bound, which means that the p-value is a lower bound.

Examining some of the underlying matrices and values used in the calculation of the four multivariate statistics is easy. For example, you can list the sum of squares and cross products (SSCP) matrices for error and the hypothesis that are found in the `e(E)` and `e(H_m)` returned matrices, the eigenvalues of $\mathbf{E}^{-1}\mathbf{H}$ obtained from the `e(eigvals_m)` returned matrix, and the three auxiliary values (`s`, `m`, and `n`) that are returned in the `e(aux_m)` matrix.

```
. mat list e(E)

symmetric e(E)[4,4]
            y1          y2          y3          y4
y1   .31998754
y2   1.6965639    12.14279
y3   .55408744   4.3636123   4.2908128
y4   .21713994   2.1102135   2.4816563   1.7225248

. mat list e(H_m)

symmetric e(H_m)[4,4]
            y1          y2          y3          y4
y1   .07356042
y2   .53738525   4.1996621
y3   .33226448   2.3553887   6.1139358
y4   .20846994   1.6371084   3.7810439   2.4930912

. mat list e(eigvals_m)

e(eigvals_m)[1,4]
            c1          c2          c3          c4
r1   1.8756709   .79069412   .22904906   .02595358
```

```
. mat list e(aux_m)

e(aux_m)[3,1]
    value
s       4
m       0
n    18.5
```

The values s, m, and n are helpful when you do not want to rely on the approximate F tests but instead want to look up critical values for the multivariate tests. Tables of critical values can be found in many multivariate texts, including Rencher (1998) and Rencher and Christensen (2012).

See example 1 in [MV] **manova postestimation** for an illustration of using test for Wald tests on expressions involving the underlying coefficients of the model and lincom for displaying linear combinations along with standard errors and confidence intervals from this MANOVA example.

See examples 1–5 in [MV] **discrim lda postestimation** for a descriptive linear discriminant analysis of the rootstock data. Many researchers use linear discriminant analysis as a method of exploring the differences between groups after a MANOVA model.

◁

▷ Example 2: One-way MANOVA with unbalanced data

Table 4.5 of Rencher (1998) presents data reported by Allison, Zappasodi, and Lurie (1962). The dependent variables y1, recording the number of bacilli inhaled per tubercle formed, and y2, recording tubercle size (in millimeters), were measured for four groups of rabbits. Group one (unvaccinated control) and group two (infected during metabolic depression) have seven observations each, whereas group three (infected during heightened metabolic activity) has 5 observations, and group four (infected during normal activity) has only 2 observations.

```
. use https://www.stata-press.com/data/r19/metabolic
(Table 4.5. Metabolic comparisons of rabbits, Rencher (1998))
. list
```

	group	y1	y2
1.	1	24	3.5
2.	1	13.3	3.5
3.	1	12.2	4
4.	1	14	4
5.	1	22.2	3.6
6.	1	16.1	4.3
7.	1	27.9	5.2
8.	2	7.4	3.5
9.	2	13.2	3
10.	2	8.5	3
11.	2	10.1	3
12.	2	9.3	2
13.	2	8.5	2.5
14.	2	4.3	1.5
15.	3	16.4	3.2
16.	3	24	2.5
17.	3	53	1.5
18.	3	32.7	2.6
19.	3	42.8	2
20.	4	25.1	2.7
21.	4	5.9	2.3

The one-way MANOVA for testing the null hypothesis that the two-dimensional mean vectors for the four groups of rabbits are equal is

```
. manova y1 y2 = group
```

Number of obs = 21

W = Wilks' lambda L = Lawley-Hotelling trace
P = Pillai's trace R = Roy's largest root

Source	Statistic		df	F(df1,	df2) =	F	Prob>F	
group	W	0.1596	3	6.0	32.0	8.02	0.0000	e
	P	1.2004		6.0	34.0	8.51	0.0000	a
	L	3.0096		6.0	30.0	7.52	0.0001	a
	R	1.5986		3.0	17.0	9.06	0.0008	u
Residual			17					
Total			20					

e = exact, a = approximate, u = upper bound on F

All four multivariate tests indicate rejection of the null hypothesis. This indicates that there are one or more differences among the two-dimensional mean vectors for the four groups. For this example, the F test for Wilks's lambda is exact because there are only two dependent variables in the model.

manovatest tests terms or linear combinations of the model's underlying design matrix. Example 2 of [MV] **manova postestimation** continues this example and illustrates manovatest.

◁

Reporting coefficients

The mvreg command (see [MV] **mvreg**) is used as a coefficient displayer after manova. Simply type mvreg to view the coefficients, standard errors, t statistics, p-values, and confidence intervals of the multivariate regression model underlying the previous manova.

▷ Example 3: Reporting coefficients by using mvreg

Continuing with example 2, we now use mvreg to display the coefficients underlying our MANOVA.

```
. mvreg
Equation             Obs   Parms        RMSE    "R-sq"          F        P>F
----------------------------------------------------------------------------
y1                    21       4    8.753754    0.5867   8.045716     0.0015
y2                    21       4    .6314183    0.6108   8.891362     0.0009

------------------------------------------------------------------------------
             | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
y1           |
       group |
          2  |  -9.771429   4.679078    -2.09   0.052    -19.64342    .1005633
          3  |   15.25143   5.125673     2.98   0.008     4.437203    26.06565
          4  |  -3.028571   7.018617    -0.43   0.672    -17.83656    11.77942
             |
       _cons |   18.52857   3.308608     5.60   0.000     11.54802    25.50912
-------------+----------------------------------------------------------------
y2           |
       group |
          2  |  -1.371429   .3375073    -4.06   0.001    -2.083507   -.6593504
          3  |  -1.654286   .3697207    -4.47   0.000    -2.434328   -.8742432
          4  |  -1.514286   .5062609    -2.99   0.008    -2.582403   -.4461685
             |
       _cons |   4.014286   .2386537    16.82   0.000      3.51077    4.517801
------------------------------------------------------------------------------
```

◁

mvreg options allowed on replay, such as level(), vsquish, and base, may also be specified to alter what is displayed.

Two-way MANOVA

You can include multiple explanatory variables with the manova command, and you can specify interactions by placing '#' between the variable names.

▷ Example 4: Two-way MANOVA with unbalanced data

Table 4.6 of Rencher (1998) presents unbalanced data from Woodard (1931) for a two-way MANOVA with three dependent variables (y1, y2, and y3) measured on patients with fractures of the jaw. y1 is age of patient, y2 is blood lymphocytes, and y3 is blood polymorphonuclears. The two design factors are gender (1 = male, 2 = female) and fracture (indicating the type of fracture: 1 = one compound fracture, 2 = two compound fractures, and 3 = one simple fracture). gender and fracture are numeric variables with value labels.

```
. use https://www.stata-press.com/data/r19/jaw
(Table 4.6. Two-way unbalanced data for fractures of the jaw, Rencher (1998))

. describe

Contains data from https://www.stata-press.com/data/r19/jaw.dta
 Observations:            27                  Table 4.6. Two-way unbalanced
                                                data for fractures of the jaw,
                                                Rencher (1998)
    Variables:             5                  20 Apr 2024 14:53
                                              (_dta has notes)
--------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
--------------------------------------------------------------------------------
gender          byte    %9.0g      gender
fracture        byte    %22.0g     fractype
y1              byte    %9.0g                 Age
y2              byte    %9.0g                 Blood lymphocytes
y3              byte    %9.0g                 Blood polymorphonuclears
--------------------------------------------------------------------------------
Sorted by:

. list in 19/22
```

	gender	fracture	y1	y2	y3
19.	Male	One simple fracture	55	32	60
20.	Male	One simple fracture	30	34	62
21.	Female	One compound fracture	22	56	43
22.	Female	Two compound fractures	22	29	68

The two-way factorial MANOVA for these data is

```
. manova y1 y2 y3 = gender fracture gender#fracture
                           Number of obs =         27

                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df    F(df1,      df2) =    F   Prob>F
         ----------------+--------------------------------------------------------
                   Model | W   0.2419          5     15.0      52.9     2.37 0.0109 a
                         | P   1.1018                15.0      63.0     2.44 0.0072 a
                         | L   1.8853                15.0      53.0     2.22 0.0170 a
                         | R   0.9248                 5.0      21.0     3.88 0.0119 u
                         |--------------------------------------------------------
                Residual |                    21
         ----------------+--------------------------------------------------------
                  gender | W   0.7151          1      3.0      19.0     2.52 0.0885 e
                         | P   0.2849                 3.0      19.0     2.52 0.0885 e
                         | L   0.3983                 3.0      19.0     2.52 0.0885 e
                         | R   0.3983                 3.0      19.0     2.52 0.0885 e
                         |--------------------------------------------------------
                fracture | W   0.4492          2      6.0      38.0     3.12 0.0139 e
                         | P   0.6406                 6.0      40.0     3.14 0.0128 a
                         | L   1.0260                 6.0      36.0     3.08 0.0155 a
                         | R   0.7642                 3.0      20.0     5.09 0.0088 u
                         |--------------------------------------------------------
         gender#fracture | W   0.5126          2      6.0      38.0     2.51 0.0380 e
                         | P   0.5245                 6.0      40.0     2.37 0.0472 a
                         | L   0.8784                 6.0      36.0     2.64 0.0319 a
                         | R   0.7864                 3.0      20.0     5.24 0.0078 u
                         |--------------------------------------------------------
                Residual |                    21
         ----------------+--------------------------------------------------------
                   Total |                    26
         ---------------------------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

For MANOVA models with more than one term, the output of `manova` shows test results for the overall model, followed by results for each term in the MANOVA.

The interaction term, `gender#fracture`, is significant at the 0.05 level. Wilks's lambda for the interaction has an exact F that produces a p-value of 0.0380.

Example 3 of [MV] **manova postestimation** illustrates how the `margins` postestimation command can be used to examine details of this significant interaction. It also illustrates how to obtain residuals by using `predict`.

◁

N-way MANOVA

Higher-order MANOVA models are easily constructed using # to indicate the interaction terms.

▷ Example 5: MANOVA with interaction terms

Data on the wear of coated fabrics is provided by Box (1950) and is presented in table 6.20 of Rencher and Christensen (2012, 249). Variables `y1`, `y2`, and `y3` are the wear after successive 1,000 revolutions of an abrasive wheel. Three factors are also recorded. `treatment` is the surface treatment and has two levels. `filler` is the filler type, also with two levels. `proportion` is the proportion of filler and has three levels (25%, 50%, and 75%).

```
. use https://www.stata-press.com/data/r19/fabric
(Table 6.20. Wear of coated fabrics, Rencher and Christensen (2012))

. describe

Contains data from https://www.stata-press.com/data/r19/fabric.dta
 Observations:            24                  Table 6.20. Wear of coated
                                                fabrics, Rencher and Christensen
                                                (2012)
    Variables:             6                  30 Aug 2024 14:01
                                              (_dta has notes)
--------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
--------------------------------------------------------------------------------
treatment       byte    %9.0g                 Surface treatment
filler          byte    %9.0g                 Filler type
proportion      byte    %9.0g      prop       Proportion of filler
y1              int     %9.0g                 First 1000 revolutions
y2              int     %9.0g                 Second 1000 revolutions
y3              int     %9.0g                 Third 1000 revolutions
--------------------------------------------------------------------------------
Sorted by:

. label list prop
prop:
           1 25%
           2 50%
           3 75%

. list

     +------------------------------------------------------+
     | treatm~t   filler   propor~n    y1    y2    y3 |
     |------------------------------------------------------|
  1. |        0        1        25%   194   192   141 |
  2. |        0        1        50%   233   217   171 |
  3. |        0        1        75%   265   252   207 |
  4. |        0        1        25%   208   188   165 |
  5. |        0        1        50%   241   222   201 |
     |------------------------------------------------------|
  6. |        0        1        75%   269   283   191 |
  7. |        0        2        25%   239   127    90 |
  8. |        0        2        50%   224   123    79 |
  9. |        0        2        75%   243   117   100 |
 10. |        0        2        25%   187   105    85 |
     |------------------------------------------------------|
 11. |        0        2        50%   243   123   110 |
 12. |        0        2        75%   226   125    75 |
 13. |        1        1        25%   155   169   151 |
 14. |        1        1        50%   198   187   176 |
 15. |        1        1        75%   235   225   166 |
     |------------------------------------------------------|
 16. |        1        1        25%   173   152   141 |
 17. |        1        1        50%   177   196   167 |
 18. |        1        1        75%   229   270   183 |
 19. |        1        2        25%   137    82    77 |
 20. |        1        2        50%   129    94    78 |
     |------------------------------------------------------|
 21. |        1        2        75%   155    76    92 |
 22. |        1        2        25%   160    82    83 |
 23. |        1        2        50%    98    89    48 |
 24. |        1        2        75%   132   105    67 |
     +------------------------------------------------------+
```

proportion is a numeric variable taking on values 1, 2, and 3, and is value-labeled with labels 25%, 50%, and 75%. treatment takes on values of 0 and 1, whereas filler is either 1 or 2.

First, we examine these data, ignoring the repeated-measures aspects of y1, y2, and y3. In example 12, we will take it into account.

```
. manova y1 y2 y3 = proportion##treatment##filler

                           Number of obs =         24

                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df   F(df1,      df2) =   F   Prob>F
         ----------------+--------------------------------------------------------
                   Model | W    0.0007        11    33.0      30.2     10.10 0.0000 a
                         | P    2.3030              33.0      36.0      3.60 0.0001 a
                         | L   74.4794              33.0      26.0     19.56 0.0000 a
                         | R   59.1959              11.0      12.0     64.58 0.0000 u
                         |--------------------------------------------------------
                Residual |                    12
         ----------------+--------------------------------------------------------
              proportion | W    0.1375         2     6.0      20.0      5.65 0.0014 e
                         | P    0.9766               6.0      22.0      3.50 0.0139 a
                         | L    5.4405               6.0      18.0      8.16 0.0002 a
                         | R    5.2834               3.0      11.0     19.37 0.0001 u
                         |--------------------------------------------------------
               treatment | W    0.0800         1     3.0      10.0     38.34 0.0000 e
                         | P    0.9200               3.0      10.0     38.34 0.0000 e
                         | L   11.5032               3.0      10.0     38.34 0.0000 e
                         | R   11.5032               3.0      10.0     38.34 0.0000 e
                         |--------------------------------------------------------
             proportion# | W    0.7115         2     6.0      20.0      0.62 0.7134 e
               treatment | P    0.2951               6.0      22.0      0.63 0.7013 a
                         | L    0.3962               6.0      18.0      0.59 0.7310 a
                         | R    0.3712               3.0      11.0      1.36 0.3055 u
                         |--------------------------------------------------------
                  filler | W    0.0192         1     3.0      10.0    170.60 0.0000 e
                         | P    0.9808               3.0      10.0    170.60 0.0000 e
                         | L   51.1803               3.0      10.0    170.60 0.0000 e
                         | R   51.1803               3.0      10.0    170.60 0.0000 e
                         |--------------------------------------------------------
       proportion#filler | W    0.1785         2     6.0      20.0      4.56 0.0046 e
                         | P    0.9583               6.0      22.0      3.37 0.0164 a
                         | L    3.8350               6.0      18.0      5.75 0.0017 a
                         | R    3.6235               3.0      11.0     13.29 0.0006 u
                         |--------------------------------------------------------
        treatment#filler | W    0.3552         1     3.0      10.0      6.05 0.0128 e
                         | P    0.6448               3.0      10.0      6.05 0.0128 e
                         | L    1.8150               3.0      10.0      6.05 0.0128 e
                         | R    1.8150               3.0      10.0      6.05 0.0128 e
                         |--------------------------------------------------------
             proportion# | W    0.7518         2     6.0      20.0      0.51 0.7928 e
        treatment#filler | P    0.2640               6.0      22.0      0.56 0.7589 a
                         | L    0.3092               6.0      18.0      0.46 0.8260 a
                         | R    0.2080               3.0      11.0      0.76 0.5381 u
                         |--------------------------------------------------------
                Residual |                    12
         ----------------+--------------------------------------------------------
                   Total |                    23
         ----------------+--------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

The MANOVA table indicates that all the terms are significant, except for proportion#treatment and proportion#treatment#filler.

◁

❑ Technical note

MANOVA uses the same design matrix as ANOVA. manova saves the full variance–covariance matrix and coefficient vector. These need a dimension equal to the dimension of the design matrix times the number of variables in the *depvarlist*.

❑

MANCOVA

MANCOVA models are specified by including the covariates as *terms* in the manova preceded by the c. operator to indicate that they are to be treated as continuous instead of categorical variables.

▷ Example 6: MANCOVA

Table 4.9 of Rencher (1998) provides biochemical measurements on four weight groups. Rencher extracted the data from Brown and Beerstecher (1951) and Smith, Gnanadesikan, and Hughes (1962). Three dependent variables and two covariates are recorded for eight subjects within each of the four groups. The first two groups are underweight, and the last two groups are overweight. The dependent variables are modified creatinine coefficient (y1), pigment creatinine (y2), and phosphate in mg/mL (y3). The two covariates are volume in ml (x1) and specific gravity (x2).

```
. use https://www.stata-press.com/data/r19/biochemical
(Table 4.9. Biochemical measurements, Rencher (1998))

. describe

Contains data from https://www.stata-press.com/data/r19/biochemical.dta
 Observations:            32                  Table 4.9. Biochemical
                                                measurements, Rencher (1998)
    Variables:             6                  22 Apr 2024 21:48
                                              (_dta has notes)
--------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
--------------------------------------------------------------------------------
group           byte    %9.0g
y1              float   %9.0g                 Modified creatinine coefficient
y2              float   %9.0g                 pigment creatinine
y3              float   %9.0g                 Phosphate (mg/ml)
x1              int     %9.0g                 Volume (ml)
x2              byte    %9.0g                 Specific gravity
--------------------------------------------------------------------------------
Sorted by:
```

Rencher performs three tests on these data. The first is a test of equality of group effects adjusted for the covariates. The second is a test that the coefficients for the covariates are jointly equal to zero. The third is a test that the coefficients for the covariates are equal across groups.

```
. manova y1 y2 y3 = group c.x1 c.x2
                           Number of obs =         32

                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                Source |  Statistic          df   F(df1,      df2) =   F   Prob>F
          -------------+--------------------------------------------------------
                 Model | W   0.0619           5    15.0      66.7     7.73 0.0000 a
                       | P   1.4836                15.0      78.0     5.09 0.0000 a
                       | L   6.7860                15.0      68.0    10.25 0.0000 a
                       | R   5.3042                 5.0      26.0    27.58 0.0000 u
                       |--------------------------------------------------------
              Residual |                     26
          -------------+--------------------------------------------------------
                 group | W   0.1491           3     9.0      58.6     7.72 0.0000 a
                       | P   0.9041                 9.0      78.0     3.74 0.0006 a
                       | L   5.3532                 9.0      68.0    13.48 0.0000 a
                       | R   5.2872                 3.0      26.0    45.82 0.0000 u
                       |--------------------------------------------------------
                    x1 | W   0.6841           1     3.0      24.0     3.69 0.0257 e
                       | P   0.3159                 3.0      24.0     3.69 0.0257 e
                       | L   0.4617                 3.0      24.0     3.69 0.0257 e
                       | R   0.4617                 3.0      24.0     3.69 0.0257 e
                       |--------------------------------------------------------
                    x2 | W   0.9692           1     3.0      24.0     0.25 0.8576 e
                       | P   0.0308                 3.0      24.0     0.25 0.8576 e
                       | L   0.0318                 3.0      24.0     0.25 0.8576 e
                       | R   0.0318                 3.0      24.0     0.25 0.8576 e
                       |--------------------------------------------------------
              Residual |                     26
          -------------+--------------------------------------------------------
                 Total |                     31
          ----------------------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

The test of equality of group effects adjusted for the covariates is shown in the MANCOVA table above. Rencher reports a Wilks's lambda value of 0.1491, which agrees with the value shown for the `group` term above. `group` is found to be significant.

The test that the coefficients for the covariates are jointly equal to zero is obtained using `manovatest`.

```
. manovatest c.x1 c.x2
                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                Source |  Statistic          df   F(df1,      df2) =   F   Prob>F
          -------------+--------------------------------------------------------
                 x1 x2 | W   0.4470           2     6.0      48.0     3.97 0.0027 e
                       | P   0.5621                 6.0      50.0     3.26 0.0088 a
                       | L   1.2166                 6.0      46.0     4.66 0.0009 a
                       | R   1.1995                 3.0      25.0    10.00 0.0002 u
                       |--------------------------------------------------------
              Residual |                     26
          ----------------------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

Wilks's lambda of 0.4470 agrees with the value reported by Rencher. With a p-value of 0.0027, we reject the null hypothesis that the coefficients for the covariates are jointly zero.

To test that the coefficients for the covariates are equal across groups, we perform a MANCOVA that includes our covariates (`x1` and `x2`) interacted with `group`. We then use `manovatest` to obtain the combined test of equal coefficients for `x1` and `x2` across groups.

```
. manova y1 y2 y3 = group c.x1 c.x2 group#c.x1 group#c.x2
                           Number of obs =         32
                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df    F(df1,     df2) =    F   Prob>F
              -----------+-----------------------------------------------------------
                   Model | W    0.0205        11     33.0     53.7      4.47 0.0000 a
                         | P    1.9571               33.0     60.0      3.41 0.0000 a
                         | L   10.6273               33.0     50.0      5.37 0.0000 a
                         | R    7.0602               11.0     20.0     12.84 0.0000 u
                         |-----------------------------------------------------------
                Residual |                    20
              -----------+-----------------------------------------------------------
                   group | W    0.4930         3      9.0     44.0      1.65 0.1317 a
                         | P    0.5942                9.0     60.0      1.65 0.1226 a
                         | L    0.8554                9.0     50.0      1.58 0.1458 a
                         | R    0.5746                3.0     20.0      3.83 0.0256 u
                         |-----------------------------------------------------------
                      x1 | W    0.7752         1      3.0     18.0      1.74 0.1947 e
                         | P    0.2248                3.0     18.0      1.74 0.1947 e
                         | L    0.2900                3.0     18.0      1.74 0.1947 e
                         | R    0.2900                3.0     18.0      1.74 0.1947 e
                         |-----------------------------------------------------------
                      x2 | W    0.8841         1      3.0     18.0      0.79 0.5169 e
                         | P    0.1159                3.0     18.0      0.79 0.5169 e
                         | L    0.1311                3.0     18.0      0.79 0.5169 e
                         | R    0.1311                3.0     18.0      0.79 0.5169 e
                         |-----------------------------------------------------------
                group#x1 | W    0.4590         3      9.0     44.0      1.84 0.0873 a
                         | P    0.6378                9.0     60.0      1.80 0.0869 a
                         | L    0.9702                9.0     50.0      1.80 0.0923 a
                         | R    0.6647                3.0     20.0      4.43 0.0152 u
                         |-----------------------------------------------------------
                group#x2 | W    0.5275         3      9.0     44.0      1.47 0.1899 a
                         | P    0.5462                9.0     60.0      1.48 0.1747 a
                         | L    0.7567                9.0     50.0      1.40 0.2130 a
                         | R    0.4564                3.0     20.0      3.04 0.0527 u
                         |-----------------------------------------------------------
                Residual |                    20
              -----------+-----------------------------------------------------------
                   Total |                    31
              -----------------------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
. manovatest group#c.x1 group#c.x2
                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df    F(df1,     df2) =    F   Prob>F
       ------------------+-----------------------------------------------------------
       group#x1 group#x2 | W    0.3310         6     18.0     51.4      1.37 0.1896 a
                         | P    0.8600               18.0     60.0      1.34 0.1973 a
                         | L    1.4629               18.0     50.0      1.35 0.1968 a
                         | R    0.8665                6.0     20.0      2.89 0.0341 u
                         |-----------------------------------------------------------
                Residual |                    20
       --------------------------------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

Rencher reports 0.3310 for Wilks's lambda, which agrees with the results of manovatest above. Here we fail to reject the null hypothesis.

◁

MANOVA for Latin-square designs

▷ Example 7: MANOVA with Latin-square data

Exercise 5.11 from Timm (1975) presents data from a multivariate Latin-square design. Two dependent variables are measured in a 4 × 4 Latin square. W is the student's score on determining distances within the solar system. B is the student's score on determining distances beyond the solar system. The three variables comprising the square are machine, ability, and treatment, each at four levels.

```
. use https://www.stata-press.com/data/r19/solardistance
(Multivariate Latin square, Timm (1975), exercise 5.11 #1)

. describe

Contains data from https://www.stata-press.com/data/r19/solardistance.dta
 Observations:            16                  Multivariate Latin square, Timm
                                                (1975), exercise 5.11 #1
    Variables:             5                  23 Apr 2024 03:27
                                              (_dta has notes)
--------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
--------------------------------------------------------------------------------
machine         byte    %9.0g                 Teaching machine
ability         byte    %9.0g                 Ability tracks
treatment       byte    %9.0g                 Method of measuring astronomical
                                                distances
W               byte    %9.0g                 Solar system distances (within)
B               byte    %9.0g                 Solar system distances (beyond)
--------------------------------------------------------------------------------
Sorted by:

. list

     +-----------------------------------------+
     | machine   ability   treatm~t    W    B  |
     |-----------------------------------------|
  1. |       1         1          2   33   15  |
  2. |       1         2          1   40    4  |
  3. |       1         3          3   31   16  |
  4. |       1         4          4   37   10  |
  5. |       2         1          4   25   20  |
     |-----------------------------------------|
  6. |       2         2          3   30   18  |
  7. |       2         3          1   22    6  |
  8. |       2         4          2   25   18  |
  9. |       3         1          1   10    5  |
 10. |       3         2          4   20   16  |
     |-----------------------------------------|
 11. |       3         3          2   17   16  |
 12. |       3         4          3   12    4  |
 13. |       4         1          3   24   15  |
 14. |       4         2          2   20   13  |
 15. |       4         3          4   19   14  |
     |-----------------------------------------|
 16. |       4         4          1   29   20  |
     +-----------------------------------------+
```

```
. manova W B = machine ability treatment
                           Number of obs =          16
                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root
                  Source | Statistic          df   F(df1,      df2) =   F   Prob>F
             ------------+----------------------------------------------------
                   Model | W    0.0378         9     18.0      10.0     2.30 0.0898 e
                         | P    1.3658               18.0      12.0     1.44 0.2645 a
                         | L   14.7756               18.0       8.0     3.28 0.0455 a
                         | R   14.0137                9.0       6.0     9.34 0.0066 u
                         |----------------------------------------------------
                Residual |                     6
             ------------+----------------------------------------------------
                 machine | W    0.0561         3      6.0      10.0     5.37 0.0101 e
                         | P    1.1853                6.0      12.0     2.91 0.0545 a
                         | L   12.5352                6.0       8.0     8.36 0.0043 a
                         | R   12.1818                3.0       6.0    24.36 0.0009 u
                         |----------------------------------------------------
                 ability | W    0.4657         3      6.0      10.0     0.78 0.6070 e
                         | P    0.5368                6.0      12.0     0.73 0.6322 a
                         | L    1.1416                6.0       8.0     0.76 0.6199 a
                         | R    1.1367                3.0       6.0     2.27 0.1802 u
                         |----------------------------------------------------
               treatment | W    0.4697         3      6.0      10.0     0.77 0.6137 e
                         | P    0.5444                6.0      12.0     0.75 0.6226 a
                         | L    1.0988                6.0       8.0     0.73 0.6378 a
                         | R    1.0706                3.0       6.0     2.14 0.1963 u
                         |----------------------------------------------------
                Residual |                     6
             ------------+----------------------------------------------------
                   Total |                    15
             -----------------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

We find that machine is a significant factor in the model, whereas ability and treatment are not.

◁

MANOVA for nested designs

Nested terms are specified using a vertical bar. A|B is read as A nested within B. A|B|C is read as A nested within B, which is nested within C. A|B#C is read as A nested within the interaction of B and C. A#B|C is read as the interaction of A and B, which is nested within C.

Different error terms can be specified for different parts of the model. The forward slash is used to indicate that the next term in the model is the error term for what precedes it. For instance, manova y1 y2 = A / B|A indicates that the multivariate tests for A are to be tested using the SSCP matrix from B|A in the denominator. Error terms (terms following the slash) are generally not tested unless they are themselves followed by a slash. The residual-error SSCP matrix is the default error-term matrix.

For example, consider T_1 / T_2 / T_3, where T_1, T_2, and T_3 may be arbitrarily complex terms. manova will report T_1 tested by T_2 and T_2 tested by T_3. If we add one more slash on the end to form T_1 / T_2 / T_3 /, then manova will also report T_3 tested by the residual error.

When you have nested terms in your model, we recommend using the dropemptycells option of manova or setting c(emptycells) to drop; see [R] **set emptycells**. See the technical note at the end of the *Nested designs* section of [R] **anova** for details.

▷ Example 8: MANOVA with nested data

A chain of retail stores produced two training videos for sales associates. The videos teach how to increase sales of the store's primary product. The videos also teach how to follow up a primary sale with secondary sales of the accessories that consumers often use with the primary product. The company trainers are not sure which video will provide the best training. To decide which video to distribute to all their stores to train sales associates, they selected three stores to use one of the training videos and three other stores to use the other training video. From each store, two employees (sales associates) were selected to receive the training. The baseline weekly sales for each of these employees was recorded and then the increase in sales over their baseline was recorded for 3 or 4 different weeks. The `videotrainer` data are described below.

```
. use https://www.stata-press.com/data/r19/videotrainer
(Video training)

. describe

Contains data from https://www.stata-press.com/data/r19/videotrainer.dta
 Observations:            42                  Video training
    Variables:             5                  9 May 2024 12:50
--------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
--------------------------------------------------------------------------------
video           byte    %9.0g                 Training video
store           byte    %9.0g                 Store (nested in video)
associate       byte    %9.0g                 Sales associate (nested in store)
primary         float   %9.0g                 Primary sales increase
extra           float   %9.0g                 Secondary sales increase
--------------------------------------------------------------------------------
Sorted by: video  store  associate
```

In this fully nested design, `video` is a fixed factor, whereas the remaining terms are random factors.

```
. manova primary extra =  video / store|video / associate|store|video /,
> dropemptycells

                           Number of obs =         42

                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df   F(df1,     df2) =   F   Prob>F
              -----------+-------------------------------------------------------
                   Model | W   0.2455         11    22.0      58.0     2.68 0.0014 e
                         | P   0.9320               22.0      60.0     2.38 0.0042 a
                         | L   2.3507               22.0      56.0     2.99 0.0005 a
                         | R   1.9867               11.0      30.0     5.42 0.0001 u
                         |-------------------------------------------------------
                Residual |                    30
              -----------+-------------------------------------------------------
                   video | W   0.1610          1     2.0       3.0     7.82 0.0646 e
                         | P   0.8390                2.0       3.0     7.82 0.0646 e
                         | L   5.2119                2.0       3.0     7.82 0.0646 e
                         | R   5.2119                2.0       3.0     7.82 0.0646 e
                         |-------------------------------------------------------
             store|video |                     4
              -----------+-------------------------------------------------------
             store|video | W   0.3515          4     8.0      10.0     0.86 0.5775 e
                         | P   0.7853                8.0      12.0     0.97 0.5011 a
                         | L   1.4558                8.0       8.0     0.73 0.6680 a
                         | R   1.1029                4.0       6.0     1.65 0.2767 u
                         |-------------------------------------------------------
        associate|store| |                     6
                   video |
              -----------+-------------------------------------------------------
        associate|store| | W   0.5164          6    12.0      58.0     1.89 0.0543 e
                   video | P   0.5316               12.0      60.0     1.81 0.0668 a
                         | L   0.8433               12.0      56.0     1.97 0.0451 a
                         | R   0.7129                6.0      30.0     3.56 0.0087 u
                         |-------------------------------------------------------
                Residual |                    30
              -----------+-------------------------------------------------------
                   Total |                    41
              -------------------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

There appears to be a difference in the videos (with significance levels just a bit above the standard 5% level). There also appears to be a sales associate effect but not a store effect.

See example 4 of [MV] **manova postestimation** for a continuation of this example. It illustrates how to test pooled terms against nonresidual error terms by using the `manovatest` postestimation command. In that example, `store` is pooled with `associate` from the original fully specified MANOVA. Another way of pooling is to refit the model, discarding the higher-level terms. Be careful in doing this to ensure that the remaining lower-level terms have a numbering scheme that will not mistakenly consider different subjects as being the same. The `videotrainer` dataset has `associate` numbered uniquely, so we can simply type

```
. manova primary extra = video / associate|video /, dropemptycells
                            Number of obs =         42
                            W = Wilks' lambda      L = Lawley-Hotelling trace
                            P = Pillai's trace     R = Roy's largest root
                  Source |  Statistic          df   F(df1,      df2) =   F   Prob>F
         ----------------+---------------------------------------------------------
                   Model | W   0.2455          11    22.0      58.0     2.68 0.0014 e
                         | P   0.9320                22.0      60.0     2.38 0.0042 a
                         | L   2.3507                22.0      56.0     2.99 0.0005 a
                         | R   1.9867                11.0      30.0     5.42 0.0001 u
                         |---------------------------------------------------------
                Residual |                     30
         ----------------+---------------------------------------------------------
                   video | W   0.4079           1     2.0       9.0     6.53 0.0177 e
                         | P   0.5921                 2.0       9.0     6.53 0.0177 e
                         | L   1.4516                 2.0       9.0     6.53 0.0177 e
                         | R   1.4516                 2.0       9.0     6.53 0.0177 e
                         |---------------------------------------------------------
         associate|video |                     10
         ----------------+---------------------------------------------------------
         associate|video | W   0.3925          10    20.0      58.0     1.73 0.0546 e
                         | P   0.7160                20.0      60.0     1.67 0.0647 a
                         | L   1.2711                20.0      56.0     1.78 0.0469 a
                         | R   0.9924                10.0      30.0     2.98 0.0100 u
                         |---------------------------------------------------------
                Residual |                     30
         ----------------+---------------------------------------------------------
                   Total |                     41
         --------------------------------------------------------------------------
                             e = exact, a = approximate, u = upper bound on F
```

and get the same results that we obtained using `manovatest` to get a pooled test after the full MANOVA; see example 4 of [MV] **manova postestimation**.

With `store` omitted from the model, `video` now has a significance level below 5%. The increase from 4 to 10 denominator degrees of freedom for the test of `video` provides a more powerful test.

The `margins` command provides a predictive marginal mean increase in sales based on the two videos. We could request the marginal means for primary sales increase or for extra sales increase, or we can use the `expression()` option to obtain the marginal means for combined primary and secondary sales increase. By default, the predicted means are constructed taking into account the number of observations in each cell.

```
. margins, within(video) expression(predict(eq(primary))+predict(eq(extra)))
Predictive margins                                      Number of obs = 42
Expression:  predict(eq(primary))+predict(eq(extra))
Within:      video
Empty cells: reweight

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       video |
          1  |   883.1395   30.01873    29.42   0.000     824.3039    941.9752
          2  |   698.0791   30.01873    23.25   0.000     639.2434    756.9147
------------------------------------------------------------------------------
```

Alternatively, we can examine the adjusted marginal mean increase in sales letting each cell have equal weight (regardless of its sample size) by using the `asbalanced` option of the `margins` command.

```
. margins, within(video) expression(predict(eq(primary))+predict(eq(extra)))
> asbalanced
Adjusted predictions                                    Number of obs = 42
Expression:  predict(eq(primary))+predict(eq(extra))
Within:      video
Empty cells: reweight
At: 1.video
        associate   (asbalanced)
    2.video
        associate   (asbalanced)
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
video						
1	876.8818	30.32981	28.91	0.000	817.4365	936.3271
2	695.315	30.32981	22.93	0.000	635.8697	754.7603

Though the values are different between the two tables, the conclusion is the same. Using training video 1 leads to increased primary and secondary sales.

◁

MANOVA for mixed designs

▷ Example 9: Split-plot MANOVA

reading2.dta has data from an experiment involving two reading programs and three skill-enhancement techniques. Ten classes of first-grade students were randomly assigned so that five classes were taught with one reading program and another five classes were taught with the other. The 30 students in each class were divided into six groups with 5 students each. Within each class, the six groups were divided randomly so that each of the three skill-enhancement techniques was taught to two of the groups within each class. At the end of the school year, a reading assessment test was administered to all the students. Two scores were recorded. The first was a reading score (score), and the second was a comprehension score (comprehension).

Example 13 of [R] **anova** uses reading.dta to illustrate mixed designs for ANOVA. reading2.dta is the same as reading.dta, except that the comprehension variable is added.

```
. use https://www.stata-press.com/data/r19/reading2
(Reading experiment data)
. describe
Contains data from https://www.stata-press.com/data/r19/reading2.dta
 Observations:           300                  Reading experiment data
    Variables:             6                  24 Apr 2024 08:31
                                              (_dta has notes)
```

Variable name	Storage type	Display format	Value label	Variable label
score	byte	%9.0g		Reading score
comprehension	byte	%9.0g		Comprehension score
program	byte	%9.0g		Reading program
class	byte	%9.0g		Class nested in program
skill	byte	%9.0g		Skill enhancement technique
group	byte	%9.0g		Group nested in class and skill

```
Sorted by:
```

In this split-plot MANOVA, the whole-plot treatment is the two reading programs, and the split-plot treatment is the three skill-enhancement techniques.

For this split-plot MANOVA, the error term for `program` is `class` nested within `program`. The error term for `skill` and the `program` by `skill` interaction is the `class` by `skill` interaction nested within `program`. Other terms are also involved in the model and can be seen below.

```
. manova score comp = pr / cl|pr sk pr#sk / cl#sk|pr / gr|cl#sk|pr /,
> dropemptycells

                           Number of obs =        300

                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df    F(df1,      df2) =    F   Prob>F
  -----------------------+--------------------------------------------------
                   Model | W   0.5234         59    118.0     478.0     1.55 0.0008 e
                         | P   0.5249               118.0     480.0     1.45 0.0039 a
                         | L   0.8181               118.0     476.0     1.65 0.0001 a
                         | R   0.6830                59.0     240.0     2.78 0.0000 u
                         |--------------------------------------------------
                Residual |                   240
  -----------------------+--------------------------------------------------
                 program | W   0.4543          1      2.0       7.0     4.20 0.0632 e
                         | P   0.5457                 2.0       7.0     4.20 0.0632 e
                         | L   1.2010                 2.0       7.0     4.20 0.0632 e
                         | R   1.2010                 2.0       7.0     4.20 0.0632 e
                         |--------------------------------------------------
           class|program |                     8
  -----------------------+--------------------------------------------------
                   skill | W   0.6754          2      4.0      30.0     1.63 0.1935 e
                         | P   0.3317                 4.0      32.0     1.59 0.2008 a
                         | L   0.4701                 4.0      28.0     1.65 0.1908 a
                         | R   0.4466                 2.0      16.0     3.57 0.0522 u
                         |--------------------------------------------------
           program#skill | W   0.3955          2      4.0      30.0     4.43 0.0063 e
                         | P   0.6117                 4.0      32.0     3.53 0.0171 a
                         | L   1.5100                 4.0      28.0     5.29 0.0027 a
                         | R   1.4978                 2.0      16.0    11.98 0.0007 u
                         |--------------------------------------------------
     class#skill|program |                    16
  -----------------------+--------------------------------------------------
     class#skill|program | W   0.4010         16     32.0      58.0     1.05 0.4265 e
                         | P   0.7324                32.0      60.0     1.08 0.3860 a
                         | L   1.1609                32.0      56.0     1.02 0.4688 a
                         | R   0.6453                16.0      30.0     1.21 0.3160 u
                         |--------------------------------------------------
       group|class#skill|                     30
                 program |
  -----------------------+--------------------------------------------------
       group|class#skill| W   0.7713         30     60.0     478.0     1.10 0.2844 e
                 program | P   0.2363                60.0     480.0     1.07 0.3405 a
                         | L   0.2867                60.0     476.0     1.14 0.2344 a
                         | R   0.2469                30.0     240.0     1.98 0.0028 u
                         |--------------------------------------------------
                Residual |                   240
  -----------------------+--------------------------------------------------
                   Total |                   299
  --------------------------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

The `program#skill` interaction is significant.

◁

MANOVA with repeated measures

One approach to analyzing repeated measures in an ANOVA setting is to use correction factors for terms in an ANOVA that involve the repeated measures. These correction factors attempt to correct for the violated assumption of independence of observations; see [R] **anova**. In this approach, the data are in long form; see [D] **reshape**.

Another approach to repeated measures is to use MANOVA with the repeated measures appearing as dependent variables, followed by tests involving linear combinations of these repeated measures. This approach involves fewer assumptions than the repeated-measures ANOVA approach.

The simplest possible repeated-measures design has no between-subject factors and only one within-subject factor (the repeated measures).

▷ Example 10: MANOVA with repeated-measures data

Here are data on five subjects, each of whom took three tests.

```
. use https://www.stata-press.com/data/r19/nobetween
. list

     +-----------------------------------+
     | subject   test1   test2   test3   |
     |-----------------------------------|
  1. |       1      68      69      95   |
  2. |       2      50      74      69   |
  3. |       3      72      89      71   |
  4. |       4      61      64      61   |
  5. |       5      60      71      90   |
     +-----------------------------------+
```

`manova` must be tricked into fitting a constant-only model. To do this, you generate a variable equal to one, use that variable as the single *term* in your `manova`, and then specify the `noconstant` option. From the resulting MANOVA, you then test the repeated measures with the `ytransform()` option of `manovatest`; see [MV] **manova postestimation** for syntax details.

```
. generate mycons = 1
. manova test1 test2 test3 = mycons, noconstant

                           Number of obs =          5

                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df   F(df1,     df2) =   F   Prob>F
              -----------+--------------------------------------------------------
                  mycons | W    0.0076          1      3.0       2.0    86.91 0.0114 e
                         | P    0.9924                 3.0       2.0    86.91 0.0114 e
                         | L  130.3722                 3.0       2.0    86.91 0.0114 e
                         | R  130.3722                 3.0       2.0    86.91 0.0114 e
                         |--------------------------------------------------------
                Residual |                      4
              -----------+--------------------------------------------------------
                   Total |                      5
              --------------------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

```
. mat c = (1,0,-1\0,1,-1)

. manovatest mycons, ytransform(c)

 Transformations of the dependent variables
 (1)    test1 - test3
 (2)    test2 - test3

                          W = Wilks' lambda      L = Lawley-Hotelling trace
                          P = Pillai's trace     R = Roy's largest root

                  Source |  Statistic          df   F(df1,      df2) =   F   Prob>F
              -----------+--------------------------------------------------
                  mycons | W    0.2352          1     2.0       3.0      4.88 0.1141 e
                         | P    0.7648                2.0       3.0      4.88 0.1141 e
                         | L    3.2509                2.0       3.0      4.88 0.1141 e
                         | R    3.2509                2.0       3.0      4.88 0.1141 e
                         |--------------------------------------------------
                Residual |                      4
              -----------+--------------------------------------------------
                          e = exact, a = approximate, u = upper bound on F
```

The test produced directly with `manova` is not interesting. It is testing the hypothesis that the three test score means are zero. The test produced by `manovatest` is of interest. From the contrasts in the matrix `c`, you produce a test that there is a difference between the `test1`, `test2`, and `test3` scores. Here the test produces a p-value of 0.1141, and you fail to reject the null hypothesis of equality between the test scores.

You can compare this finding with the results obtained from a repeated-measures ANOVA,

```
. reshape long test, i(subject) j(testnum)

. anova test subject testnum, repeated(testnum)
```

which produced an uncorrected p-value of 0.1160 and corrected p-values of 0.1181, 0.1435, and 0.1665 by using the Huynh–Feldt, Greenhouse–Geisser, and Box's conservative correction, respectively.

◁

▷ Example 11: Randomized block design with repeated measures

Milliken and Johnson (2009) demonstrate using `manova` to analyze repeated measures from a randomized block design used in studying the differences among varieties of sorghum. Table 27.1 of Milliken and Johnson (2009) provides the data. Four sorghum varieties were each planted in five blocks. A leaf-area index measurement was recorded for each of 5 weeks, starting 2 weeks after emergence.

The tests of interest include a test for equal variety marginal means, equal time marginal means, and a test for the interaction of variety and time. The MANOVA below does not directly provide these tests. `manovatest` after the `manova` gives the three tests of interest.

```
. use https://www.stata-press.com/data/r19/sorghum, clear
(Leaf area index on 4 sorghum varieties, Milliken & Johnson (2009))
. manova time1 time2 time3 time4 time5 = variety block
                           Number of obs =         20

                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source │ Statistic          df   F(df1,     df2) =    F   Prob>F
              ───────────┼──────────────────────────────────────────────────────
                   Model │ W    0.0001         7     35.0      36.1     9.50 0.0000 a
                         │ P    3.3890               35.0      60.0     3.61 0.0000 a
                         │ L  126.2712               35.0      32.0    23.09 0.0000 a
                         │ R  109.7360                7.0      12.0   188.12 0.0000 u
                         ├──────────────────────────────────────────────────────
                Residual │                    12
              ───────────┼──────────────────────────────────────────────────────
                 variety │ W    0.0011         3     15.0      22.5    16.11 0.0000 a
                         │ P    2.5031               15.0      30.0    10.08 0.0000 a
                         │ L   48.3550               15.0      20.0    21.49 0.0000 a
                         │ R   40.0068                5.0      10.0    80.01 0.0000 u
                         ├──────────────────────────────────────────────────────
                   block │ W    0.0047         4     20.0      27.5     5.55 0.0000 a
                         │ P    1.7518               20.0      44.0     1.71 0.0681 a
                         │ L   77.9162               20.0      26.0    25.32 0.0000 a
                         │ R   76.4899                5.0      11.0   168.12 0.0000 u
                         ├──────────────────────────────────────────────────────
                Residual │                    12
              ───────────┼──────────────────────────────────────────────────────
                   Total │                    19
              ──────────────────────────────────────────────────────────────────
                           e = exact, a = approximate, u = upper bound on F
```

Two matrices are needed for transformations of the `time#` variables. `m1` is a row vector containing five ones. `m2` provides contrasts for `time#`. The `manovatest, showorder` command lists the underlying ordering of columns for constructing two more matrices used to obtain linear combinations from the design matrix. Matrix `c1` provides contrasts on `variety`. Matrix `c2` is used to collapse to the overall margin of the design matrix to obtain time marginal means.

```
. matrix m1 = J(1,5,1)
. matrix m2 = (1,-1,0,0,0 \ 1,0,-1,0,0 \ 1,0,0,-1,0 \ 1,0,0,0,-1)
. manovatest, showorder
 Order of columns in the design matrix
      1: (variety==1)
      2: (variety==2)
      3: (variety==3)
      4: (variety==4)
      5: (block==1)
      6: (block==2)
      7: (block==3)
      8: (block==4)
      9: (block==5)
     10: _cons
. matrix c1 = (1,-1,0,0,0,0,0,0,0,0\1,0,-1,0,0,0,0,0,0,0\1,0,0,-1,0,0,0,0,0,0)
. matrix c2 = (.25,.25,.25,.25,.2,.2,.2,.2,.2,1)
```

The test for equal variety marginal means uses matrix m1 to obtain the sum of the time# variables and matrix c1 to provide the contrasts on variety. The second test uses m2 to provide contrasts on time# and matrix c2 to collapse to the appropriate margin for the test of time marginal means. The final test uses m2 for contrasts on time# and c1 for contrasts on variety to test the variety-by-time interaction.

```
. manovatest, test(c1) ytransform(m1)
 Transformation of the dependent variables
 (1)    time1 + time2 + time3 + time4 + time5

 Test constraints
 (1)    1.variety - 2.variety = 0
 (2)    1.variety - 3.variety = 0
 (3)    1.variety - 4.variety = 0

                         W = Wilks' lambda      L = Lawley-Hotelling trace
                         P = Pillai's trace     R = Roy's largest root

               Source |  Statistic          df     F(df1,      df2) =    F    Prob>F
          ------------+---------------------------------------------------------------
           manovatest | W    0.0435          3        3.0       12.0     88.05 0.0000 e
                      | P    0.9565                   3.0       12.0     88.05 0.0000 e
                      | L   22.0133                   3.0       12.0     88.05 0.0000 e
                      | R   22.0133                   3.0       12.0     88.05 0.0000 e
                      |---------------------------------------------------------------
             Residual |                     12
          ----------------------------------------------------------------------------
                         e = exact, a = approximate, u = upper bound on F
. manovatest, test(c2) ytransform(m2)
 Transformations of the dependent variables
 (1)    time1 - time2
 (2)    time1 - time3
 (3)    time1 - time4
 (4)    time1 - time5

 Test constraint
 (1)    .25*1.variety + .25*2.variety + .25*3.variety + .25*4.variety +
        .2*1.block + .2*2.block + .2*3.block + .2*4.block + .2*5.block + _cons
        = 0

                         W = Wilks' lambda      L = Lawley-Hotelling trace
                         P = Pillai's trace     R = Roy's largest root

               Source |  Statistic          df     F(df1,      df2) =    F    Prob>F
          ------------+---------------------------------------------------------------
           manovatest | W    0.0050          1        4.0        9.0    445.62 0.0000 e
                      | P    0.9950                   4.0        9.0    445.62 0.0000 e
                      | L  198.0544                   4.0        9.0    445.62 0.0000 e
                      | R  198.0544                   4.0        9.0    445.62 0.0000 e
                      |---------------------------------------------------------------
             Residual |                     12
          ----------------------------------------------------------------------------
                         e = exact, a = approximate, u = upper bound on F
```

```
. manovatest, test(c1) ytransform(m2)
 Transformations of the dependent variables
 (1)    time1 - time2
 (2)    time1 - time3
 (3)    time1 - time4
 (4)    time1 - time5

 Test constraints
 (1)    1.variety - 2.variety = 0
 (2)    1.variety - 3.variety = 0
 (3)    1.variety - 4.variety = 0

                     W = Wilks' lambda      L = Lawley-Hotelling trace
                     P = Pillai's trace     R = Roy's largest root

           Source |  Statistic          df   F(df1,     df2) =   F    Prob>F
       -----------+--------------------------------------------------
       manovatest |W    0.0143           3     12.0      24.1     8.00 0.0000 a
                  |P    2.1463                 12.0      33.0     6.91 0.0000 a
                  |L   12.1760                 12.0      23.0     7.78 0.0000 a
                  |R    8.7953                  4.0      11.0    24.19 0.0000 u
                  |--------------------------------------------------
         Residual |                     12
       -----------+--------------------------------------------------
                     e = exact, a = approximate, u = upper bound on F
```

All three tests are significant, indicating differences in `variety`, in `time`, and in the `variety`-by-`time` interaction.

◁

▷ Example 12: MANOVA and dependent-variable effects

Recall the fabric-data example from Rencher and Christensen (2012, 249) that we used in example 5 to illustrate a three-way MANOVA. Rencher and Christensen have an additional exercise to test the period effect (the `y1`, `y2`, and `y3` repeated-measures variables) and the interaction of period with the other factors in the model. The `ytransform()` option of `manovatest` provides a method to do this; see [MV] **manova postestimation**. Here are the tests of the period effect interacted with each term in the model. We create the matrix `c` with rows corresponding to the linear and quadratic contrasts for the three dependent variables.

```
. quietly manova y1 y2 y3 = proportion##treatment##filler
. matrix c = (-1,0,1 \ -1,2,-1)
. manovatest proportion, ytransform(c)
 Transformations of the dependent variables
 (1)    - y1 + y3
 (2)    - y1 + 2*y2 - y3

                     W = Wilks' lambda      L = Lawley-Hotelling trace
                     P = Pillai's trace     R = Roy's largest root

           Source |  Statistic          df   F(df1,     df2) =   F    Prob>F
       -----------+--------------------------------------------------
       proportion |W    0.4749           2      4.0      22.0     2.48 0.0736 e
                  |P    0.5454                  4.0      24.0     2.25 0.0936 a
                  |L    1.0631                  4.0      20.0     2.66 0.0630 a
                  |R    1.0213                  2.0      12.0     6.13 0.0147 u
                  |--------------------------------------------------
         Residual |                     12
       -----------+--------------------------------------------------
                     e = exact, a = approximate, u = upper bound on F
```

```
. manovatest treatment, ytransform(c)
 Transformations of the dependent variables
 (1)     - y1 + y3
 (2)     - y1 + 2*y2 - y3
                      W = Wilks' lambda      L = Lawley-Hotelling trace
                      P = Pillai's trace     R = Roy's largest root

              Source | Statistic          df    F(df1,      df2) =   F    Prob>F
           ----------+----------------------------------------------------------
           treatment | W    0.1419         1       2.0      11.0    33.27 0.0000 e
                     | P    0.8581                 2.0      11.0    33.27 0.0000 e
                     | L    6.0487                 2.0      11.0    33.27 0.0000 e
                     | R    6.0487                 2.0      11.0    33.27 0.0000 e
                     |----------------------------------------------------------
            Residual |                    12
           ----------+----------------------------------------------------------
                      e = exact, a = approximate, u = upper bound on F
. manovatest proportion#treatment, ytransform(c)
 Transformations of the dependent variables
 (1)     - y1 + y3
 (2)     - y1 + 2*y2 - y3
                      W = Wilks' lambda      L = Lawley-Hotelling trace
                      P = Pillai's trace     R = Roy's largest root

              Source | Statistic          df    F(df1,      df2) =   F    Prob>F
 --------------------+----------------------------------------------------------
proportion#treatment | W    0.7766         2       4.0      22.0     0.74 0.5740 e
                     | P    0.2276                 4.0      24.0     0.77 0.5550 a
                     | L    0.2824                 4.0      20.0     0.71 0.5972 a
                     | R    0.2620                 2.0      12.0     1.57 0.2476 u
                     |----------------------------------------------------------
            Residual |                    12
 --------------------+----------------------------------------------------------
                      e = exact, a = approximate, u = upper bound on F
. manovatest filler, ytransform(c)
 Transformations of the dependent variables
 (1)     - y1 + y3
 (2)     - y1 + 2*y2 - y3
                      W = Wilks' lambda      L = Lawley-Hotelling trace
                      P = Pillai's trace     R = Roy's largest root

              Source | Statistic          df    F(df1,      df2) =   F    Prob>F
           ----------+----------------------------------------------------------
              filler | W    0.0954         1       2.0      11.0    52.17 0.0000 e
                     | P    0.9046                 2.0      11.0    52.17 0.0000 e
                     | L    9.4863                 2.0      11.0    52.17 0.0000 e
                     | R    9.4863                 2.0      11.0    52.17 0.0000 e
                     |----------------------------------------------------------
            Residual |                    12
           ----------+----------------------------------------------------------
                      e = exact, a = approximate, u = upper bound on F
```

```
. manovatest proportion#filler, ytransform(c)
 Transformations of the dependent variables
 (1)     - y1 + y3
 (2)     - y1 + 2*y2 - y3
                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source  Statistic        df    F(df1,     df2) =   F   Prob>F
       
       proportion#filler  W   0.6217        2       4.0      22.0     1.48 0.2436 e
                          P   0.3870                4.0      24.0     1.44 0.2515 a
                          L   0.5944                4.0      20.0     1.49 0.2439 a
                          R   0.5698                2.0      12.0     3.42 0.0668 u
                       
                Residual                   12
       
                           e = exact, a = approximate, u = upper bound on F
. manovatest treatment#filler, ytransform(c)
 Transformations of the dependent variables
 (1)     - y1 + y3
 (2)     - y1 + 2*y2 - y3
                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source  Statistic        df    F(df1,     df2) =   F   Prob>F
       
        treatment#filler  W   0.3867        1       2.0      11.0     8.72 0.0054 e
                          P   0.6133                2.0      11.0     8.72 0.0054 e
                          L   1.5857                2.0      11.0     8.72 0.0054 e
                          R   1.5857                2.0      11.0     8.72 0.0054 e
                       
                Residual                   12
       
                           e = exact, a = approximate, u = upper bound on F
. manovatest proportion#treatment#filler, ytransform(c)
 Transformations of the dependent variables
 (1)     - y1 + y3
 (2)     - y1 + 2*y2 - y3
                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source  Statistic        df    F(df1,     df2) =   F   Prob>F
       
             proportion#  W   0.7812        2       4.0      22.0     0.72 0.5857 e
        treatment#filler  P   0.2290                4.0      24.0     0.78 0.5518 a
                          L   0.2671                4.0      20.0     0.67 0.6219 a
                          R   0.2028                2.0      12.0     1.22 0.3303 u
                       
                Residual                   12
       
                           e = exact, a = approximate, u = upper bound on F
```

The first test, `manovatest proportion, ytransform(c)`, provides the test of `proportion` interacted with the period effect. The F tests for Wilks's lambda, Pillai's trace, and the Lawley–Hotelling trace do not reject the null hypothesis with a significance level of 0.05 (p-values of 0.0736, 0.0936, and 0.0630). The F test for Roy's largest root is an upper bound, so the p-value of 0.0147 is a lower bound.

The tests of `treatment` interacted with the period effect, `filler` interacted with the period effect, and `treatment#filler` interacted with the period effect are significant. The remaining tests are not.

To test the period effect, we call `manovatest` with both the `ytransform()` and `test()` options. The `showorder` option guides us in constructing the matrix for the `test()` option.

```
. manovatest, showorder
 Order of columns in the design matrix
      1: (proportion==1)
      2: (proportion==2)
      3: (proportion==3)
      4: (treatment==0)
      5: (treatment==1)
      6: (proportion==1)*(treatment==0)
      7: (proportion==1)*(treatment==1)
      8: (proportion==2)*(treatment==0)
      9: (proportion==2)*(treatment==1)
     10: (proportion==3)*(treatment==0)
     11: (proportion==3)*(treatment==1)
     12: (filler==1)
     13: (filler==2)
     14: (proportion==1)*(filler==1)
     15: (proportion==1)*(filler==2)
     16: (proportion==2)*(filler==1)
     17: (proportion==2)*(filler==2)
     18: (proportion==3)*(filler==1)
     19: (proportion==3)*(filler==2)
     20: (treatment==0)*(filler==1)
     21: (treatment==0)*(filler==2)
     22: (treatment==1)*(filler==1)
     23: (treatment==1)*(filler==2)
     24: (proportion==1)*(treatment==0)*(filler==1)
     25: (proportion==1)*(treatment==0)*(filler==2)
     26: (proportion==1)*(treatment==1)*(filler==1)
     27: (proportion==1)*(treatment==1)*(filler==2)
     28: (proportion==2)*(treatment==0)*(filler==1)
     29: (proportion==2)*(treatment==0)*(filler==2)
     30: (proportion==2)*(treatment==1)*(filler==1)
     31: (proportion==2)*(treatment==1)*(filler==2)
     32: (proportion==3)*(treatment==0)*(filler==1)
     33: (proportion==3)*(treatment==0)*(filler==2)
     34: (proportion==3)*(treatment==1)*(filler==1)
     35: (proportion==3)*(treatment==1)*(filler==2)
     36: _cons
```

We create a row vector, `m`, starting with 1/3 for three columns (corresponding to `proportion`), followed by 1/2 for two columns (corresponding to `treatment`), followed by 1/6 for six columns (for `proportion#treatment`), followed by 1/2 for two columns (for `filler`), followed by 1/6 for six columns (for `proportion#filler`), followed by four columns of 1/4 (for `treatment#filler`), followed by 1/12 for 12 columns (corresponding to the `proportion#treatment#filler` term), and finally, a 1 for the last column (corresponding to the constant in the model). The test of period effect then uses this `m` matrix and the `c` matrix previously defined as the basis of the test for the period effect.

```
. matrix m = J(1,3,1/3), J(1,2,1/2), J(1,6,1/6), J(1,2,1/2), J(1,6,1/6),
> J(1,4,1/4), J(1,12,1/12), (1)

. manovatest, test(m) ytrans(c)

 Transformations of the dependent variables
 (1)    - y1 + y3
 (2)    - y1 + 2*y2 - y3

 Test constraint
 (1)    .3333333*1.proportion + .3333333*2.proportion + .3333333*3.proportion +
        .5*0.treatment + .5*1.treatment + .1666667*1.proportion#0.treatment +
        .1666667*1.proportion#1.treatment + .1666667*2.proportion#0.treatment +
        .1666667*2.proportion#1.treatment + .1666667*3.proportion#0.treatment +
        .1666667*3.proportion#1.treatment + .5*1.filler + .5*2.filler +
        .1666667*1.proportion#1.filler + .1666667*1.proportion#2.filler +
        .1666667*2.proportion#1.filler + .1666667*2.proportion#2.filler +
        .1666667*3.proportion#1.filler + .1666667*3.proportion#2.filler +
        .25*0.treatment#1.filler + .25*0.treatment#2.filler +
        .25*1.treatment#1.filler + .25*1.treatment#2.filler +
        .0833333*1.proportion#0.treatment#1.filler +
        .0833333*1.proportion#0.treatment#2.filler +
        .0833333*1.proportion#1.treatment#1.filler +
        .0833333*1.proportion#1.treatment#2.filler +
        .0833333*2.proportion#0.treatment#1.filler +
        .0833333*2.proportion#0.treatment#2.filler +
        .0833333*2.proportion#1.treatment#1.filler +
        .0833333*2.proportion#1.treatment#2.filler +
        .0833333*3.proportion#0.treatment#1.filler +
        .0833333*3.proportion#0.treatment#2.filler +
        .0833333*3.proportion#1.treatment#1.filler +
        .0833333*3.proportion#1.treatment#2.filler + _cons = 0

                           W = Wilks' lambda      L = Lawley-Hotelling trace
                           P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df   F(df1,      df2) =   F   Prob>F
             ------------+----------------------------------------------------------
              manovatest | W    0.0208         1      2.0      11.0    259.04 0.0000 e
                         | P    0.9792                2.0      11.0    259.04 0.0000 e
                         | L   47.0988                2.0      11.0    259.04 0.0000 e
                         | R   47.0988                2.0      11.0    259.04 0.0000 e
                         |----------------------------------------------------------
                Residual |                    12
             ------------+----------------------------------------------------------
                           e = exact, a = approximate, u = upper bound on F
```

This result agrees with the answers provided by Rencher and Christensen (2012).

◁

In the previous three examples, one factor has been encoded within the dependent variables. We have seen that the `ytransform()` option of `manovatest` provides the method for testing this factor and its interactions with the factors that appear on the right-hand side of the MANOVA.

More than one factor could be encoded within the dependent variables. Again the `ytransform()` option of `manovatest` allows us to perform multivariate tests of interest.

▷ Example 13: MANOVA and multiple dependent-variable effects

Table 6.14 of Rencher and Christensen (2012) provides an example with two within-subject factors represented in the dependent variables and one between-subject factor.

```
. use https://www.stata-press.com/data/r19/table614
(Table 6.14. Repeated measures experiment, Rencher and Christensen (2012))
. list in 9/12, noobs compress
```

c	sub~t	ab11	ab12	ab13	ab21	ab22	ab23	ab31	ab32	ab33
1	9	41	32	23	37	51	39	27	28	30
1	10	39	32	24	30	35	31	26	29	32
2	1	47	36	25	31	36	29	21	24	27
2	2	53	43	32	40	48	47	46	50	54

There are 20 observations. Factors a and b are encoded in the names of the nine dependent variables. Variable name ab23, for instance, indicates factor a at level 2 and factor b at level 3. Factor c is the between-subject factor.

We first compute a MANOVA by using the dependent variables and our one between-subject term.

```
. manova ab11 ab12 ab13 ab21 ab22 ab23 ab31 ab32 ab33 = c
                          Number of obs =         20
                          W = Wilks' lambda      L = Lawley-Hotelling trace
                          P = Pillai's trace     R = Roy's largest root
```

Source	Statistic		df	F(df1,	df2) =	F	Prob>F	
c	W	0.5330	1	9.0	10.0	0.97	0.5114	e
	P	0.4670		9.0	10.0	0.97	0.5114	e
	L	0.8762		9.0	10.0	0.97	0.5114	e
	R	0.8762		9.0	10.0	0.97	0.5114	e
Residual			18					
Total			19					

```
                          e = exact, a = approximate, u = upper bound on F
```

This approach provides the basis for computing tests on all terms of interest. We use the ytransform() and test() options of manovatest with the following matrices to obtain the tests of interest.

```
. mat a = (2,2,2,-1,-1,-1,-1,-1,-1 \ 0,0,0,1,1,1,-1,-1,-1)
. mat b = (2,-1,-1,2,-1,-1,2,-1,-1 \ 0,1,-1,0,1,-1,0,1,-1)
. forvalues i = 1/2 {
  2.         forvalues j = 1/2 {
  3.                 mat g = nullmat(g) \ vecdiag(a[`i',1...]'*b[`j',1...])
  4.         }
  5. }
. mat list g
g[4,9]
    c1  c2  c3  c4  c5  c6  c7  c8  c9
r1   4  -2  -2  -2   1   1  -2   1   1
r1   0   2  -2   0  -1   1   0  -1   1
r1   0   0   0   2  -1  -1  -2   1   1
r1   0   0   0   0   1  -1   0  -1   1
```

```
. mat j = J(1,9,1/9)
. mat xall = (.5,.5,1)
```

Matrices a and b correspond to factors a and b. Matrix g is the elementwise multiplication of each row of a with each row of b and corresponds to the a#b interaction. Matrix j is used to average the dependent variables, whereas matrix xall collapses over factor c.

Here are the tests for a, b, and a#b.

```
. manovatest, test(xall) ytrans(a)
 Transformations of the dependent variables
 (1)    2*ab11 + 2*ab12 + 2*ab13 - ab21 - ab22 - ab23 - ab31 - ab32 - ab33
 (2)    ab21 + ab22 + ab23 - ab31 - ab32 - ab33

 Test constraint
 (1)    .5*1.c + .5*2.c + _cons = 0
                            W = Wilks' lambda      L = Lawley-Hotelling trace
                            P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df     F(df1,      df2) =    F   Prob>F
              -----------+--------------------------------------------------------------
              manovatest |W    0.6755          1       2.0      17.0     4.08 0.0356 e
                         |P    0.3245                  2.0      17.0     4.08 0.0356 e
                         |L    0.4803                  2.0      17.0     4.08 0.0356 e
                         |R    0.4803                  2.0      17.0     4.08 0.0356 e
                         |--------------------------------------------------------------
                Residual |                    18
              -----------+--------------------------------------------------------------
                            e = exact, a = approximate, u = upper bound on F
. manovatest, test(xall) ytrans(b)
 Transformations of the dependent variables
 (1)    2*ab11 - ab12 - ab13 + 2*ab21 - ab22 - ab23 + 2*ab31 - ab32 - ab33
 (2)    ab12 - ab13 + ab22 - ab23 + ab32 - ab33

 Test constraint
 (1)    .5*1.c + .5*2.c + _cons = 0
                            W = Wilks' lambda      L = Lawley-Hotelling trace
                            P = Pillai's trace     R = Roy's largest root

                  Source | Statistic          df     F(df1,      df2) =    F   Prob>F
              -----------+--------------------------------------------------------------
              manovatest |W    0.3247          1       2.0      17.0    17.68 0.0001 e
                         |P    0.6753                  2.0      17.0    17.68 0.0001 e
                         |L    2.0799                  2.0      17.0    17.68 0.0001 e
                         |R    2.0799                  2.0      17.0    17.68 0.0001 e
                         |--------------------------------------------------------------
                Residual |                    18
              -----------+--------------------------------------------------------------
                            e = exact, a = approximate, u = upper bound on F
```

```
. manovatest, test(xall) ytrans(g)
 Transformations of the dependent variables
 (1)    4*ab11 - 2*ab12 - 2*ab13 - 2*ab21 + ab22 + ab23 - 2*ab31 + ab32 + ab33
 (2)    2*ab12 - 2*ab13 - ab22 + ab23 - ab32 + ab33
 (3)    2*ab21 - ab22 - ab23 - 2*ab31 + ab32 + ab33
 (4)    ab22 - ab23 - ab32 + ab33
 Test constraint
 (1)    .5*1.c + .5*2.c + _cons = 0
                          W = Wilks' lambda      L = Lawley-Hotelling trace
                          P = Pillai's trace     R = Roy's largest root

              Source | Statistic          df   F(df1,     df2) =   F   Prob>F
        -------------+-----------------------------------------------------------
          manovatest | W   0.2255          1      4.0     15.0    12.88 0.0001 e
                     | P   0.7745                 4.0     15.0    12.88 0.0001 e
                     | L   3.4347                 4.0     15.0    12.88 0.0001 e
                     | R   3.4347                 4.0     15.0    12.88 0.0001 e
                     |-----------------------------------------------------------
            Residual |                    18
        -------------+-----------------------------------------------------------
                          e = exact, a = approximate, u = upper bound on F
```

Factors a, b, and a#b are significant with p-values of 0.0356, 0.0001, and 0.0001, respectively. The multivariate statistics are equivalent to the T^2 values Rencher and Christensen report using the relationship $T^2 = (n_1 + n_2 - 2) \times (1 - \Lambda)/\Lambda$ that applies in this situation. For instance, Wilks's lambda for factor a is reported as 0.6755 (and the actual value recorded in r(stat) is 0.67554286) so that $T^2 = (10+10-2) \times (1-0.67554286)/0.67554286 = 8.645$, as reported by Rencher and Christensen.

We now compute the tests for c and the interactions of c with the other terms in the model.

```
. manovatest c, ytrans(j)
 Transformation of the dependent variables
 (1)    .1111111*ab11 + .1111111*ab12 + .1111111*ab13 + .1111111*ab21 +
        .1111111*ab22 + .1111111*ab23 + .1111111*ab31 + .1111111*ab32 +
        .1111111*ab33
                          W = Wilks' lambda      L = Lawley-Hotelling trace
                          P = Pillai's trace     R = Roy's largest root

              Source | Statistic          df   F(df1,     df2) =   F   Prob>F
        -------------+-----------------------------------------------------------
                   c | W   0.6781          1      1.0     18.0     8.54 0.0091 e
                     | P   0.3219                 1.0     18.0     8.54 0.0091 e
                     | L   0.4747                 1.0     18.0     8.54 0.0091 e
                     | R   0.4747                 1.0     18.0     8.54 0.0091 e
                     |-----------------------------------------------------------
            Residual |                    18
        -------------+-----------------------------------------------------------
                          e = exact, a = approximate, u = upper bound on F
```

```
. manovatest c, ytrans(a)
 Transformations of the dependent variables
 (1)    2*ab11 + 2*ab12 + 2*ab13 - ab21 - ab22 - ab23 - ab31 - ab32 - ab33
 (2)    ab21 + ab22 + ab23 - ab31 - ab32 - ab33

                          W = Wilks' lambda      L = Lawley-Hotelling trace
                          P = Pillai's trace     R = Roy's largest root

               Source |  Statistic          df   F(df1,      df2) =   F   Prob>F
          ------------+---------------------------------------------------------
                    c | W   0.9889           1     2.0      17.0     0.10 0.9097 e
                      | P   0.0111                 2.0      17.0     0.10 0.9097 e
                      | L   0.0112                 2.0      17.0     0.10 0.9097 e
                      | R   0.0112                 2.0      17.0     0.10 0.9097 e
                      |---------------------------------------------------------
             Residual |                     18
          ----------------------------------------------------------------------
                          e = exact, a = approximate, u = upper bound on F
. manovatest c, ytrans(b)
 Transformations of the dependent variables
 (1)    2*ab11 - ab12 - ab13 + 2*ab21 - ab22 - ab23 + 2*ab31 - ab32 - ab33
 (2)    ab12 - ab13 + ab22 - ab23 + ab32 - ab33

                          W = Wilks' lambda      L = Lawley-Hotelling trace
                          P = Pillai's trace     R = Roy's largest root

               Source |  Statistic          df   F(df1,      df2) =   F   Prob>F
          ------------+---------------------------------------------------------
                    c | W   0.9718           1     2.0      17.0     0.25 0.7845 e
                      | P   0.0282                 2.0      17.0     0.25 0.7845 e
                      | L   0.0290                 2.0      17.0     0.25 0.7845 e
                      | R   0.0290                 2.0      17.0     0.25 0.7845 e
                      |---------------------------------------------------------
             Residual |                     18
          ----------------------------------------------------------------------
                          e = exact, a = approximate, u = upper bound on F
. manovatest c, ytrans(g)
 Transformations of the dependent variables
 (1)    4*ab11 - 2*ab12 - 2*ab13 - 2*ab21 + ab22 + ab23 - 2*ab31 + ab32 + ab33
 (2)    2*ab12 - 2*ab13 - ab22 + ab23 - ab32 + ab33
 (3)    2*ab21 - ab22 - ab23 - 2*ab31 + ab32 + ab33
 (4)    ab22 - ab23 - ab32 + ab33

                          W = Wilks' lambda      L = Lawley-Hotelling trace
                          P = Pillai's trace     R = Roy's largest root

               Source |  Statistic          df   F(df1,      df2) =   F   Prob>F
          ------------+---------------------------------------------------------
                    c | W   0.9029           1     4.0      15.0     0.40 0.8035 e
                      | P   0.0971                 4.0      15.0     0.40 0.8035 e
                      | L   0.1075                 4.0      15.0     0.40 0.8035 e
                      | R   0.1075                 4.0      15.0     0.40 0.8035 e
                      |---------------------------------------------------------
             Residual |                     18
          ----------------------------------------------------------------------
                          e = exact, a = approximate, u = upper bound on F
```

The test of c is equivalent to an ANOVA using the sum or average of the dependent variables as the dependent variable. The test of c produces an F of 8.54 with a p-value of 0.0091, which agrees with the results of Rencher and Christensen (2012, 229–230).

The tests of a#c, b#c, and a#b#c produce p-values of 0.9097, 0.7845, and 0.8035, respectively.

In summary, the factors that are significant are a, b, a#b, and c.

◁

Stored results

manova stores the following in e():

Scalars

e(N)	number of observations
e(k)	number of parameters
e(k_eq)	number of equations in e(b)
e(df_m)	model degrees of freedom
e(df_r)	residual degrees of freedom
e(df_#)	degrees of freedom for term #
e(rank)	rank of e(V)

Macros

e(cmd)	manova
e(cmdline)	command as typed
e(depvar)	names of dependent variables
e(indepvars)	names of the right-hand-side variables
e(term_#)	term #
e(errorterm_#)	error term for term # (defined for terms using nonresidual error)
e(wtype)	weight type
e(wexp)	weight expression
e(r2)	R^2 for each equation
e(rmse)	RMSE for each equation
e(F)	F statistic for each equation
e(p_F)	p-value for F test for each equation
e(properties)	b V
e(estat_cmd)	program used to implement estat
e(predict)	program used to implement predict
e(marginsnotok)	predictions disallowed by margins
e(marginsdefault)	default predict() specification for margins
e(asbalanced)	factor variables fvset as asbalanced
e(asobserved)	factor variables fvset as asobserved

Matrices

e(b)	coefficient vector (a stacked version of e(B))
e(B)	coefficient matrix
e(E)	residual-error SSCP matrix
e(xpxinv)	generalized inverse of $X'X$
e(H_m)	hypothesis SSCP matrix for the overall model
e(stat_m)	multivariate statistics for the overall model
e(eigvals_m)	eigenvalues of $E^{-1}H$ for the overall model
e(aux_m)	s, m, and n values for the overall model
e(H_#)	hypothesis SSCP matrix for term #
e(stat_#)	multivariate statistics for term # (if computed)
e(eigvals_#)	eigenvalues of $E^{-1}H$ for term # (if computed)
e(aux_#)	s, m, and n values for term # (if computed)
e(V)	variance–covariance matrix of the estimators

Functions

e(sample)	marks estimation sample

Methods and formulas

Let $\mathbf{Y}$ denote the matrix of observations on the left-hand-side variables. Let $\mathbf{X}$ denote the design matrix based on the right-hand-side variables. The last column of $\mathbf{X}$ is equal to all ones (unless the `noconstant` option was specified). Categorical right-hand-side variables are placed in $\mathbf{X}$ as a set of indicator (sometimes called dummy) variables, whereas continuous variables enter as is. Columns of $\mathbf{X}$ corresponding to interactions are formed by multiplying the various combinations of columns for the variables involved in the interaction.

The multivariate model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

leads to multivariate hypotheses of the form

$$\mathbf{C}\boldsymbol{\beta}\mathbf{A}' = \mathbf{0}$$

where $\boldsymbol{\beta}$ is a matrix of parameters, $\mathbf{C}$ specifies constraints on the design matrix $\mathbf{X}$ for a particular hypothesis, and $\mathbf{A}$ provides a transformation of $\mathbf{Y}$. $\mathbf{A}$ is often the identity matrix.

An estimate of $\boldsymbol{\beta}$ is provided by

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

The error sum of squares and cross products (SSCP) matrix is

$$\mathbf{E} = \mathbf{A}(\mathbf{Y}'\mathbf{Y} - \mathbf{B}'\mathbf{X}'\mathbf{X}\mathbf{B})\mathbf{A}'$$

and the SSCP matrix for the hypothesis is

$$\mathbf{H} = \mathbf{A}(\mathbf{C}\mathbf{B})'\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-}\mathbf{C}'\}^{-1}(\mathbf{C}\mathbf{B})\mathbf{A}'$$

The inclusion of weights, if specified, enters the formulas in a manner similar to that shown in *Methods and formulas* of [R] **regress**.

Let $\lambda_1 > \lambda_2 > \cdots > \lambda_s$ represent the nonzero eigenvalues of $\mathbf{E}^{-1}\mathbf{H}$. $s = \min(p, \nu_h)$, where p is the number of columns of $\mathbf{YA}'$ (that is, the number of y variables or number of resultant transformed left-hand-side variables), and ν_h is the hypothesis degrees of freedom.

Wilks's (1932) lambda statistic is

$$\Lambda = \prod_{i=1}^{s} \frac{1}{1+\lambda_i} = \frac{|\mathbf{E}|}{|\mathbf{H}+\mathbf{E}|}$$

and is a likelihood-ratio test. This statistic is distributed as the Wilks's Λ distribution if $\mathbf{E}$ has the Wishart distribution, $\mathbf{H}$ has the Wishart distribution under the null hypothesis, and $\mathbf{E}$ and $\mathbf{H}$ are independent. The null hypothesis is rejected for small values of Λ.

Pillai's (1955) trace is

$$V = \sum_{i=1}^{s} \frac{\lambda_i}{1+\lambda_i} = \text{trace}\left\{(\mathbf{E}+\mathbf{H})^{-1}\mathbf{H}\right\}$$

and the Lawley–Hotelling trace (Lawley 1938; Hotelling 1951) is

$$U = \sum_{i=1}^{s} \lambda_i = \text{trace}(\mathbf{E}^{-1}\mathbf{H})$$

and is also known as Hotelling's generalized T^2 statistic.

Roy's largest root is taken as λ_1, though some report $\theta = \lambda_1/(1+\lambda_1)$, which is bounded between zero and one. Roy's largest root provides a test based on the union-intersection approach to test construction introduced by Roy (1939).

Tables providing critical values for these four multivariate statistics are found in many of the books that discuss MANOVA, including Rencher (1998) and Rencher and Christensen (2012).

Let p be the number of columns of $\mathbf{YA}'$ (that is, the number of y variables or the number of resultant transformed y variables), ν_h be the hypothesis degrees of freedom, ν_e be the error degrees of freedom, $s = \min(\nu_h, p)$, $m = (|\nu_h - p| - 1)/2$, and $n = (\nu_e - p - 1)/2$. Transformations of these four multivariate statistics to F statistics are as follows.

For Wilks's lambda, an approximate F statistic (Rao 1951) with df_1 and df_2 degrees of freedom is

$$F = \frac{(1-\Lambda^{1/t})\mathrm{df}_2}{(\Lambda^{1/t})\mathrm{df}_1}$$

where

$$\mathrm{df}_1 = p\nu_h \qquad \mathrm{df}_2 = wt + 1 - p\nu_h/2$$

$$w = \nu_e + \nu_h - (p + \nu_h + 1)/2$$

$$t = \left(\frac{p^2\nu_h^2 - 4}{p^2 + \nu_h^2 - 5}\right)^{1/2}$$

t is set to one if either the numerator or the denominator equals zero. This F statistic is exact when p equals 1 or 2 or when ν_h equals 1 or 2.

An approximate F statistic for Pillai's trace (Pillai 1954, 1956b) with $s(2m+s+1)$ and $s(2n+s+1)$ degrees of freedom is

$$F = \frac{(2n+s+1)V}{(2m+s+1)(s-V)}$$

An approximate F statistic for the Lawley–Hotelling trace (Pillai 1954, 1956a) with $s(2m+s+1)$ and $2sn+2$ degrees of freedom is

$$F = \frac{2(sn+1)U}{s^2(2m+s+1)}$$

When p or ν_h are 1, an exact F statistic for Roy's largest root is

$$F = \lambda_1 \frac{\nu_e - p + 1}{p}$$

with $|\nu_h - p| + 1$ and $\nu_e - p + 1$ degrees of freedom. In other cases, an upper bound F statistic (providing a lower bound on the p-value) for Roy's largest root is

$$F = \lambda_1 \frac{\nu_e - d + \nu_h}{d}$$

with d and $\nu_e - d + \nu_h$ degrees of freedom, where $d = \max(p, \nu_h)$.

Samuel Stanley Wilks (1906–1964) was born in Texas. He gained degrees in architecture, mathematics, and statistics from North Texas Teachers' College and the universities of Texas and Iowa. After periods in Columbia and England, he moved to Princeton in 1933. Wilks published various widely used texts, was founding editor of the *Annals of Mathematical Statistics*, and made many key contributions to multivariate statistics. Wilks's lambda is named for him.

References

Allison, M. J., P. Zappasodi, and M. B. Lurie. 1962. The correlation of biphasic metabolic response with a biphasic response in resistance to tuberculosis in rabbits. *Journal of Experimental Medicine* 115: 881–890. https://doi.org/10.1084/jem.115.5.881.

Anderson, T. W. 1965. Samuel Stanley Wilks, 1906–1964. *Annals of Mathematical Statistics* 36: 1–23. https://doi.org/10.1214/aoms/1177700267.

———. 2003. *An Introduction to Multivariate Statistical Analysis*. 3rd ed. New York: Wiley.

Andrews, D. F., and A. M. Herzberg, eds. 1985. *Data: A Collection of Problems from Many Fields for the Student and Research Worker*. New York: Springer. https://doi.org/10.1007/978-1-4612-5098-2.

Arnold, S. F. 1981. *The Theory of Linear Models and Multivariate Analysis*. New York: Wiley.

Box, G. E. P. 1950. Problems in the analysis of growth and wear curves. *Biometrics* 6: 362–389. https://doi.org/10.2307/3001781.

Brown, J. D., and E. Beerstecher. 1951. "Metabolic patterns of underweight and overweight individuals". In *Biochemical Institute Studies IV, No. 5109*. Austin, TX: University of Texas Press.

Hotelling, H. 1951. A generalized t^2 test and measurement of multivariate dispersion. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* 1: 23–41.

Lawley, D. N. 1938. A generalization of Fisher's z-test. *Biometrika* 30: 180–187. https://doi.org/10.2307/2332232.

Mardia, K. V., J. T. Kent, and C. C. Taylor. 2024. *Multivariate Analysis*. 2nd ed. Hoboken, NJ: Wiley.

Mehmetoglu, M., and T. G. Jakobsen. 2022. *Applied Statistics Using Stata: A Guide for the Social Sciences*. 2nd ed. Thousand Oaks, CA: Sage.

Milliken, G. A., and D. E. Johnson. 2009. *Designed Experiments*. Vol. 1 of *Analysis of Messy Data*, 2nd ed. Boca Raton, FL: CRC Press.

Morrison, D. F. 2005a. *Multivariate Statistical Methods*. 4th ed. Belmont, CA: Duxbury.

———. 2005b. "Multivariate analysis of variance". In *Encyclopedia of Biostatistics*, edited by P. Armitage and T. Colton, vol. 5: 3467–3472. Chichester, UK: Wiley. https://doi.org/10.1002/0470011815.b2a13045.

Pillai, K. C. S. 1954. "On some distribution problems in multivariate analysis". In *Mimeograph Series No. 88*. Institute of Statistics, University of North Carolina, Chapel Hill.

———. 1955. Some new test criteria in multivariate analysis. *Annals of Mathematical Statistics* 26: 117–121. https://doi.org/10.1214/aoms/1177728599.

———. 1956a. Some results useful in multivariate analysis. *Annals of Mathematical Statistics* 27: 1106–1114. https://doi.org/10.1214/aoms/1177728076.

———. 1956b. On the distribution of the largest or the smallest root of a matrix in multivariate analysis. *Biometrika* 43: 122–127. https://doi.org/10.1093/biomet/43.1-2.122.

———. 1985. "Multivariate analysis of variance (MANOVA)". In *Encyclopedia of Statistical Sciences*, edited by S. Kotz, N. L. Johnson, and C. B. Read, vol. 6: 20–29. New York: Wiley.

Rao, C. R. 1951. An asymptotic expansion of the distribution of Wilks' criterion. *Bulletin of the International Statistical Institute* 33: 177–180.

Rencher, A. C. 1998. *Multivariate Statistical Inference and Applications*. New York: Wiley.

Rencher, A. C., and W. F. Christensen. 2012. *Methods of Multivariate Analysis*. 3rd ed. Hoboken, NJ: Wiley. https://doi.org/10.1002/9781118391686.

Roy, S. N. 1939. p-statistics or some generalizations in analysis of variance appropriate to multivariate problems. *Sankhyā* 4: 381–396.

Seber, G. A. F. 1984. *Multivariate Observations*. New York: Wiley. https://doi.org/10.1002/9780470316641.

Smith, H., R. Gnanadesikan, and J. B. Hughes. 1962. Multivariate analysis of variance (MANOVA). *Biometrics* 18: 22–41. https://doi.org/10.2307/2527708.

Timm, N. H. 1975. *Multivariate Analysis with Applications in Education and Psychology*. Pacific Grove, CA: Brooks/Cole.

Wilks, S. S. 1932. Certain generalizations in the analysis of variance. *Biometrika* 24: 471–494. https://doi.org/10.2307/2331979.

Woodard, D. E. 1931. Healing time of fractures of the jaw in relation to delay before reduction, infection, syphilis and blood calcium and phosphorus content. *Journal of the American Dental Association* 18: 419–442. https://doi.org/10.14219/jada.archive.1931.0096.

Also see

[MV] **manova postestimation** — Postestimation tools for manova

[MV] **mvreg** — Multivariate regression

[MV] **mvtest** — Multivariate tests

[D] **encode** — Encode string into numeric and vice versa

[D] **reshape** — Convert data from wide to long form and vice versa

[R] **anova** — Analysis of variance and covariance

Stata Structural Equation Modeling Reference Manual

[U] 13.5 Accessing coefficients and standard errors

[U] 20 Estimation and postestimation commands