

manova - 多変量分散分析 【 評価版 】

manova は多変量分散分析 (MANOVA)、あるいは多変量共分散分析 (MANCOVA) の機能を提供するコマンドです。単変量の分散/共分散分析機能については [R] anova (*mwp-042*) をご参照ください。

1. 多変量分散分析	
2. 一元配置 MANOVA	Example 1
	Example 2
3. 基盤となる回帰モデル	Example 3
4. 二元配置 MANOVA	Example 4
5. N 元配置 MANOVA	Example 5
6. MANCOVA	Example 6
7. ラテン方格デザイン	Example 7
8. ネスト型デザイン	Example 8
9. 混合型デザイン	Example 9
10. 反復測定 MANOVA	Example 10
	Example 11
	Example 12
	Example 13

1. 多変量分散分析

manova は anova コマンドを多変量に拡張したもので、多変量分散分析 (MANOVA: multivariate analysis-of-variance)、及び多変量共分散分析 (MANCOVA: multivariate analysis-of-covariance) の機能を提供します。

MANOVA で算出される標準的な統計量には次の 4 種類があります。

- (i) Wilks のラムダ
- (ii) Pillai のトレース
- (iii) Lawley-Hotelling トレース
- (iv) Roy の最大根

いずれの検定も決定則に基づくものであり (admissible)、不偏性 (unbiased) と不変性 (invariant) を備えています。(i), (ii), (iii) は漸近的には同一のものとなりますが、帰無仮説が成り立たないときや小標本に対する挙動は微妙に異なります。(iv) は漸的にみても (i), (ii), (iii) とは性質を異にします。

2. 一元配置 MANOVA

▷ Example 1: balanced data

[MV] `manova` の Example 1 には Example データセット `rootstock.dta` を用いた一元配置 MANOVA の例が紹介されています。

```
. use http://www.stata-press.com/data/r16/rootstock.dta *1
(Table 6.2 Rootstock Data, Rencher and Christensen (2012))
```

このデータセットには `rootstock`, `y1`, `y2`, `y3`, `y4` という 5 つの変数が記録されています。`rootstock` はリンゴの木の根茎を特定するためのカテゴリ変数で 1 から 6 までの値を取ります。それぞれの根茎から得られた 8 本の木について次の 4 種類のデータが記録されています。

変数	意味
<code>y1</code>	4 年時の幹周 (trunk girth)
<code>y2</code>	4 年時の伸長成長 (extension growth)
<code>y3</code>	15 年時の幹周
<code>y4</code>	15 年時の地上部重量

全体では 48 件のデータが存在するわけですが、それぞれのグループに含まれるデータ件数は一律 8 件であるため、バランスの取れた構成となっています。次はデータの一部をリスト出力したものです。

```
. by rootstock, sort: generate flag = 1 if _n == 1
. list rootstock y1 y2 y3 y4 if flag == 1, abbreviate(9) separator(0) *2
```

	rootstock	y1	y2	y3	y4
1.	1	1.11	2.569	3.58	0.760
9.	2	1.05	2.074	4.09	1.036
17.	3	1.07	2.505	3.76	0.912
25.	4	1.22	2.838	3.89	0.944
33.	5	0.91	1.532	4.04	1.084
41.	6	1.11	2.813	3.76	0.800

*1 メニュー操作 : File ▷ Example Datasets ▷ Stata 16 manual datasets と操作、Multivariate Statistics Reference Manual [MV] の `manova` の項よりダウンロードする。

*2 メニュー操作 : Data ▷ Describe data ▷ List data

この場合、y1, y2, y3, y4 はいずれも従属変数となるため、平均値の等値性を検定するには MANOVA を使用することになります。すなわち、根茎グループ i ($i = 1, \dots, 6$) の平均値を $\bar{y}_{1i}, \bar{y}_{2i}, \bar{y}_{3i}, \bar{y}_{4i}$ とした場合、4次元の平均値ベクトル $(\bar{y}_{1i}, \bar{y}_{2i}, \bar{y}_{3i}, \bar{y}_{4i})$ が $i = 1, \dots, 6$ まですべて等しいと言えるかどうかを検定しようというわけです。MANOVA の実行に際しては次のように操作します。

- Statistics ▸ Multivariate analysis ▸ MANOVA, multivariate regression, and related ▸ MANOVA と操作
- Model タブ: Dependent variables: y1 y2 y3 y4
Model: rootstock

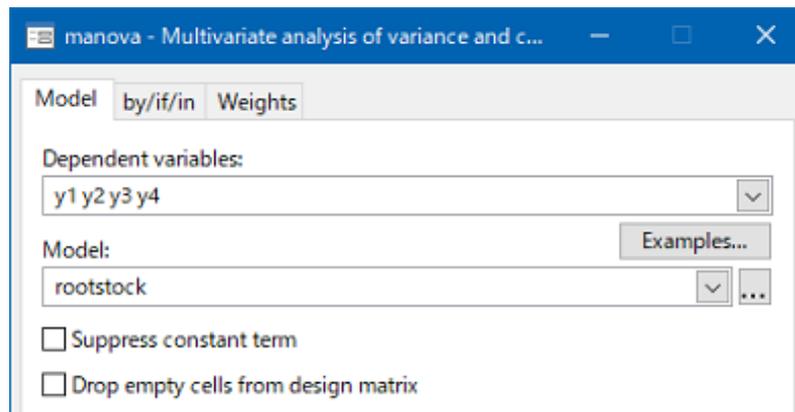


図 1 manova ダイアログ- Model タブ

```
. 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

出力テーブルの見方について解説を加えておきます。Statistic と書かれたカラムに示されているのはそれぞれ

- Wilks のラムダ
- Pillai のトレース
- Lawley-Hotelling トレース
- Roy の最大根

の値です。また df と書かれたカラムには自由度 (degrees of freedom) の値が示されています。一方、F(df1, df2) というカラムには F 統計量に関連した自由度の値が、F と表示されたカラムには F 統計量の値が示されています。最後の Prob>F と表示されたカラムが示す値は F 統計量に対応する p 値です。

テーブル上、W, P, L, R と記された 4 種類の検定結果に関する情報が集約された形で示されているわけですが、この例では p 値がいずれも 0.0000 であるため、6 個の平均値ベクトルに差はないとする帰無仮説は棄却されることとなります。なお、欄外の e, a, u というコードの意味については最終行に注釈が付されているわけですが、e は分布が正確な F 分布であることを、a は近似的に F 分布であることを、u はレポートされた F 値が上限値であることを意味します。

MANOVA モデルをフィットさせると今度はどのグループ間に違いがあるのかという点が問題となるわけですが、これについては [MV] discrim lda (*mwp-116*) の項をご参照ください。 ◀

▷ Example 2: unbalanced data

評価版では割愛しています。

3. 基盤となる回帰モデル

▷ Example 3: mvreg

評価版では割愛しています。

4. 二元配置 MANOVA

▷ Example 4: unbalanced data

評価版では割愛しています。

5. N 元配置 MANOVA

▷ Example 5: 三元配置 MANOVA

評価版では割愛しています。

6. MANCOVA

▷ Example 6: MANCOVA

評価版では割愛しています。

7. ラテン方格デザイン

▷ Example 7: ラテン方格デザイン

評価版では割愛しています。

8. ネスト型デザイン

▷ Example 8: ネスト型デザイン

評価版では割愛しています。

9. 混合型デザイン

▷ Example 9: 分割区画 MANOVA

評価版では割愛しています。

10. 反復測定 MANOVA

▷ Example 10: 反復測定データ

評価版では割愛しています。

▷ Example 11: 乱塊法

評価版では割愛しています。

▷ Example 12: 従属変数効果 (1)

評価版では割愛しています。

▷ Example 13: 従属変数効果 (2)

評価版では割愛しています。

