

Statistics 5401

17. Repeated Measures

Gary W. Oehlert
School of Statistics
313B Ford Hall
612-625-1557
gary@stat.umn.edu

Today we have two topics:

- Fancier experimental designs
- Repeated measures

We've basically talked about simple situations: completely randomized designs with one- and two-way factorial treatment structure.

We can also get fancy: multi-way factorials, split-plot designs, nesting, mixed effects, and on and on.

There is good news:

If you know how to do the univariate ANOVA, then you know how to do the multivariate ANOVA.

Numerator and denominator sums of squares for tests become hypothesis and error matrices, but which effects go where stay the same.

Example. We wish to study how testing conditions affect sensory assessments of cheddar cheese. This experiment uses 48 subjects, divided at random into 12 groups. The 12 groups are the factor/level combinations of room temperature (three levels) and room noise (four levels). Each subject tastes cheddar cheese five times and rates a sensory quality. However, each time he or she rates the cheese, a different perfume has been misted into the room. The order of the perfumes is randomized for each subject.

This is a split-plot experiment.

Subjects are whole plots. The five tasting slots for each subject are the split plots.

The whole plot treatments are the twelve combinations of temperature and noise in 4 by 3 factorial structure. Here whole-plot treatments are applied to whole plots in a completely randomized fashion with four replications.

The split plot treatments are the perfumes, with five levels.

Let M, N, and P indicate the temperature, noise, and perfume treatment factors, and let R indicate the four replications at the whole-plot level. Then the appropriate model in MacAnova is

"y = M*N + E(M.N.R) + P + M.P + N.P + M.N.P"

```
Cmd> anova("y = M*N + E(M.N.R)\n+ P + M.P + N.P + M.N.P", fstats:T)
```

Model used is y = M*N + E(M.N.R) + P + M.P + N.P + M.N.P

	DF	SS	MS	F	P-value
CONSTANT	1	85285	85285	5993.54414	0
M	2	185.82	92.911	6.52952	0.0038063
N	3	26.721	8.9069	0.62595	0.60293
M.N	6	62.4	10.4	0.73088	0.62785
ERROR1	36	512.26	14.229	14.31695	0
P	4	514.3	128.57	129.36501	0
M.P	8	4.4481	0.55602	0.55944	0.80959
N.P	12	14.349	1.1957	1.20310	0.28635
M.N.P	24	26.217	1.0924	1.09908	0.35215
ERROR2	144	143.12	0.99389		

Suppose now that we have the same experiment, but that each subject doesn't simply rate the cheese, but rates six flavor characteristics.

Now we have a vector response, so we do MANOVA instead of ANOVA.

We'll have hypothesis and error matrices, but the pairings are the same.

The **E** for M, N, and M.N is whole-plot error, and the **E** for P, M.P, N.P, and M.N.P is split-plot error.

Our choice of tests is pretty impressive:

- Wilks likelihood ratio
- Lawley-Hotelling trace
- Pillai V
- Roy's maximum root
- Bonferronized variablewise F tests
- Bonferronized T^2 tests between levels of factors

Bonferronized tests are easier to interpret, and variablewise tests only require univariate assumptions.

Bonferronized tests can be more powerful if the deviation from the null is concentrated in one variable.

Bonferronized tests can be less powerful in cases of high correlation.

```
Cmd> manova("Y = M*N + E(M.N.R)\  
+ P + M.P + N.P + M.N.P")  
Model used is Y = M*N + E(M.N.R) + P + M.P + N.P + M.N.P  
WARNING: summaries are sequential  
NOTE: SS/SP matrices suppressed because of size; use 'manova(,sssp:T)'
```

SS and SP Matrices

	DF	
CONSTANT	1	Type 'SS[1,,]' to see SS/SP matrix
M	2	Type 'SS[2,,]' to see SS/SP matrix
N	3	Type 'SS[3,,]' to see SS/SP matrix
M.N	6	Type 'SS[4,,]' to see SS/SP matrix
ERROR1	36	Type 'SS[5,,]' to see SS/SP matrix
P	4	Type 'SS[6,,]' to see SS/SP matrix
M.P	8	Type 'SS[7,,]' to see SS/SP matrix
N.P	12	Type 'SS[8,,]' to see SS/SP matrix
M.N.P	24	Type 'SS[9,,]' to see SS/SP matrix
ERROR2	144	Type 'SS[10,,]' to see SS/SP matrix

```

Cmd> vals<-releigen(SS[2,,],SS[5,,])$values;\
vals
(1)  2.3674    0.26275    0    0    0    0

Cmd> prod(1/(1+vals))
(1)      0.23517

Cmd> chi <- (36+(2-6-1)/2)*sum(log(1+vals));chi
(1)      48.489

Cmd> 1-cumchi(chi,2*6)
(1)  2.5687e-06

Cmd> vals<-releigen(SS[6,,],SS[10,,])$values;\
vals
(1) 19.059  0.029792  0.022947 0.0010347  0  0

Cmd> sum(vals)
(1)      19.113

Cmd> chi <- (144-6-1)*sum(vals);chi
(1)      2618.4

Cmd> 1-cumchi(chi,4*6)
(1)      0

```

Now let's look at a different situation. We still have the same 48 subjects doing ratings under different temperature and noise conditions. Now they rate the cheese on six different occasions.

This is a *repeated measures* design with time as a repeated or within-subject factor. It differs from the split plot in that the repeated factor is not randomized.

More generally, we could rate six characteristics on each of six occasions, and obtain a time by characteristic structure in the repeated measure.

We have

$$\mathbf{Y} = \mathbf{XB} + \boldsymbol{\varepsilon}$$

In our example, \mathbf{Y} has 48 rows (one for each subject), and either 6 or 36 columns depending on the setup. We have been looking at testing null hypotheses

$$H_0 : \mathbf{LB} = 0$$

for some $(q \times r)$ matrix \mathbf{L} .

Now consider

$$H_0 : \mathbf{LBR} = 0$$

for some $(q \times r)$ matrix \mathbf{L} and some $(p \times s)$ matrix \mathbf{R} .

We looked at hypotheses

$$H_0 : \mathbf{BR} = 0$$

when we did profile analysis. These new hypotheses

$$H_0 : \mathbf{LBR} = 0$$

allow us to combine hypotheses about both between-subjects-type treatments (through \mathbf{L}) and hypotheses about the profile or repeated-measures-type treatments (through \mathbf{R}).

Why?

In our example, we have a four by three factorial with four replications. Each response is either a six-vector (over time) or a 36-vector (time and characteristic).

Take the 36-vector case. There

$$12 \times 36 = 3 \times 4 \times 6 \times 6$$

different means. We can decompose those means into main effects for temperature, noise, time, and characteristic, two-factor interaction, three-factor interactions, and a four-factor interaction.

This is similar to a four-way factorial, but here the treatments have a two-way structure, and the response components have a two-way structure.

What should our \mathbf{L} and \mathbf{R} matrices look like?

For \mathbf{L} , we need to represent the overall mean, the main effects of the two factors, and the two-factor interaction.

For \mathbf{R} , we need to represent the overall mean, the main effects of the two within-subjects factors, and their two-factor interaction.

Let's call these $\mathbf{L}_1, \mathbf{L}_M, \mathbf{L}_N, \mathbf{L}_{MN}, \mathbf{R}_1, \mathbf{R}_O, \mathbf{R}_C$, and \mathbf{R}_{OC} .

\mathbf{L}	\mathbf{R}	effect
\mathbf{L}_1	\mathbf{R}_1	grand mean
\mathbf{L}_M	\mathbf{R}_1	main effect of temperature
\mathbf{L}_N	\mathbf{R}_1	main effect of noise
\mathbf{L}_{MN}	\mathbf{R}_1	temperature.noise interaction
\mathbf{L}_1	\mathbf{R}_O	main effect of occasion (time)
\mathbf{L}_M	\mathbf{R}_O	temperature.occasion interaction
\mathbf{L}_N	\mathbf{R}_O	noise.occasion interaction
\mathbf{L}_{MN}	\mathbf{R}_O	temperature.noise.occasion interaction
\mathbf{L}	\mathbf{R}	effect
\mathbf{L}_1	\mathbf{R}_C	main effect of characteristic
\mathbf{L}_M	\mathbf{R}_C	temperature.characteristic interaction
\mathbf{L}_N	\mathbf{R}_C	noise.characteristic interaction
\mathbf{L}_{MN}	\mathbf{R}_C	temperature.noise.characteristic interaction
\mathbf{L}_1	\mathbf{R}_{OC}	occasion.characteristic interaction
\mathbf{L}_M	\mathbf{R}_{OC}	temperature.occasion.characteristic interaction
\mathbf{L}_N	\mathbf{R}_{OC}	noise.occasion.characteristic interaction
\mathbf{L}_{MN}	\mathbf{R}_{OC}	temperature.noise.occasion.characteristic interaction

To test these hypotheses, form $\tilde{\mathbf{Y}} = \mathbf{YR}$ and then do the "left-hand side" analysis on $\tilde{\mathbf{Y}}$. In our example, the left-hand side analysis is the effects of temperature and noise.

If \mathbf{R} has one column, \mathbf{YR} has one column and we do a univariate analysis with F-tests.

If \mathbf{R} has two or more columns, \mathbf{YR} also has more than one column and we do a multivariate analysis.

For our problem, the 36-vector of responses is arranged as six characteristics for the first occasion, the same six for the second, and so on.

The **R** matrix for occasion must have five columns and make comparisons between six groups of six columns in **Y**: (1, 2, 3, 4, 5, 6), (7, 8, 9, 10, 11, 12), ..., (31, 32, 33, 34, 35, 36). There are many choices.

The **R** matrix for occasion must have five columns and make comparisons between six groups of six columns in **Y**: (1, 7, 13, 19, 25, 31), (2, 8, 14, 20, 26, 32), ..., (6, 12, 18, 24, 30, 36). There are again many choices.

The interaction matrix is usually chosen as the matrix of elementwise products of every pair of columns from the two main effects matrices.

For our example, I will use the standard contrasts used by MacAnova when doing an ANOVA model: each group is compared to the last group.

These are easy to obtain — just fake an ANOVA and use `modelinfo(xvars:T)` to get the **X** matrix; these are the columns we need for **R**.

```
Cmd> O <- factor( rep(run(6),rep(6,6)) )
```

```
Cmd> C <- factor( rep(run(6),6) )
```

```
Cmd> anova("run(36)=O*C",silent:T)
```

```
Cmd> modelinfo(colcts:T)
```

```
(1)          1          5          5          25
```

```
Cmd> R1 <- modelinfo(xvars:T)[,1]
```

```
Cmd> RO <- modelinfo(xvars:T)[,run(2,6)]
```

```
Cmd> RC <- modelinfo(xvars:T)[,run(7,11)]
```

```
Cmd> ROC <- modelinfo(xvars:T)[,run(12,36)]
```

```
Cmd> R1
```

```
(1,1)  1
(2,1)  1
(3,1)  1
(4,1)  1
(5,1)  1
(6,1)  1
(7,1)  1
(8,1)  1
(9,1)  1
(10,1) 1
(11,1) 1
(12,1) 1
(13,1) 1
(14,1) 1
(15,1) 1
(16,1) 1
(17,1) 1
(18,1) 1
(19,1) 1
(20,1) 1
```

```

(21,1) 1
(22,1) 1
(23,1) 1
(24,1) 1
(25,1) 1
(26,1) 1
(27,1) 1
(28,1) 1
(29,1) 1
(30,1) 1
(31,1) 1
(32,1) 1
(33,1) 1
(34,1) 1
(35,1) 1
(36,1) 1

```

Cmd> RO

```

(1,1) 1 0 0 0 0
(2,1) 1 0 0 0 0
(3,1) 1 0 0 0 0
(4,1) 1 0 0 0 0
(5,1) 1 0 0 0 0
(6,1) 1 0 0 0 0
(7,1) 0 1 0 0 0
(8,1) 0 1 0 0 0
(9,1) 0 1 0 0 0
(10,1) 0 1 0 0 0
(11,1) 0 1 0 0 0
(12,1) 0 1 0 0 0
(13,1) 0 0 1 0 0
(14,1) 0 0 1 0 0
(15,1) 0 0 1 0 0
(16,1) 0 0 1 0 0
(17,1) 0 0 1 0 0
(18,1) 0 0 1 0 0
(19,1) 0 0 0 1 0
(20,1) 0 0 0 1 0
(21,1) 0 0 0 1 0
(22,1) 0 0 0 1 0
(23,1) 0 0 0 1 0
(24,1) 0 0 0 1 0
(25,1) 0 0 0 0 1
(26,1) 0 0 0 0 1
(27,1) 0 0 0 0 1
(28,1) 0 0 0 0 1
(29,1) 0 0 0 0 1
(30,1) 0 0 0 0 1
(31,1) -1 -1 -1 -1 -1
(32,1) -1 -1 -1 -1 -1
(33,1) -1 -1 -1 -1 -1
(34,1) -1 -1 -1 -1 -1
(35,1) -1 -1 -1 -1 -1
(36,1) -1 -1 -1 -1 -1

```

```

Cmd> RC
(1,1)   1   0   0   0   0
(2,1)   0   1   0   0   0
(3,1)   0   0   1   0   0
(4,1)   0   0   0   1   0
(5,1)   0   0   0   0   1
(6,1)  -1  -1  -1  -1  -1
(7,1)   1   0   0   0   0
(8,1)   0   1   0   0   0
(9,1)   0   0   1   0   0
(10,1)  0   0   0   1   0
(11,1)  0   0   0   0   1
(12,1) -1  -1  -1  -1  -1
(13,1)  1   0   0   0   0
(14,1)  0   1   0   0   0
(15,1)  0   0   1   0   0
(16,1)  0   0   0   1   0
(17,1)  0   0   0   0   1
(18,1) -1  -1  -1  -1  -1
(19,1)  1   0   0   0   0
(20,1)  0   1   0   0   0
(21,1)  0   0   1   0   0
(22,1)  0   0   0   1   0
(23,1)  0   0   0   0   1
(24,1) -1  -1  -1  -1  -1
(25,1)  1   0   0   0   0
(26,1)  0   1   0   0   0
(27,1)  0   0   1   0   0
(28,1)  0   0   0   1   0
(29,1)  0   0   0   0   1
(30,1) -1  -1  -1  -1  -1
(31,1)  1   0   0   0   0
(32,1)  0   1   0   0   0
(33,1)  0   0   1   0   0
(34,1)  0   0   0   1   0
(35,1)  0   0   0   0   1
(36,1) -1  -1  -1  -1  -1

```

```

Cmd>print(ROC,format:"f2.0",labels:F)
  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
-1 -1 -1 -1 -1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  -1 -1 -1 -1 -1  0  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0
  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0

```

```

0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
-1 0 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 0
0 -1 0 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0
0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0
0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0
0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0 0 0 -1 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```
Cmd> anova("Y1 = M*N",fstats:T)
```

```
Model used is Y1 = M*N
```

	DF	SS	MS	F	P-value
CONSTANT	1	4.2275e+07	4.2275e+07	1.1469e+06	0
M	2	41002	20501	556.15809	0
N	3	64.944	21.648	0.58728	0.62728
M.N	6	426.76	71.127	1.92956	0.1025
ERROR1	36	1327	36.862		

```
Cmd> YO <- Y %*% RO
```

```
Cmd> manova("YO = M*N",print:F)
```

```
Cmd> h1 <- matrix(SS[1,,])
```

```
Cmd> hm <- matrix(SS[2,,])
```

```
Cmd> hn <- matrix(SS[3,,])
```

```
Cmd> hmn <- matrix(SS[4,,])
```

```
Cmd> e <- matrix(SS[5,,])
```

```
Cmd> DF
```

CONSTANT	M	N	M.N	ERROR1
1	2	3	6	36

```
Cmd> mvtest(h1,e,1,36) # occasion
```


	statistic	p-value
Wilks lambda	0.32868	5.2773e-07
Hotelling-Lawley trace	2.0424	6.6293e-12
Pillai trace	0.67132	0.00014969
Roy's largest root	0.67132	MISSING

```
Cmd> mvtest(hm,e,2,36) # temp.occasion
```

	statistic	p-value
Wilks lambda	0.77445	0.56169
Hotelling-Lawley trace	0.28339	0.57995
Pillai trace	0.23164	0.55097
Roy's largest root	0.20144	MISSING

```
Cmd> mvtest(hn,e,3,36) # noise.occasion
```

	statistic	p-value
Wilks lambda	0.71263	0.70244
Hotelling-Lawley trace	0.36952	0.7465
Pillai trace	0.31197	0.66637
Roy's largest root	0.20004	MISSING

```
Cmd> mvtest(hmn,e,6,36) # temp.noise.occasion
```

	statistic	p-value
Wilks lambda	0.53011	0.82145
Hotelling-Lawley trace	0.72371	0.86434
Pillai trace	0.56244	0.78883
Roy's largest root	0.30067	MISSING

```
Cmd> YC <- Y %*% RC
```

```
exmpl2.mac> manova("YC = M*N",print:F)
```

This gives us characteristic, temperature by characteristic, noise by characteristic, and temperature by noise by characteristic.

```
Cmd> YOC <- Y %*% ROC
```

```
Cmd> manova("YOC = M*N",print:F)
```

This gives us occasion by characteristic, temperature by occasion by characteristic, noise by occasion by characteristic, and temperature by noise by occasion by characteristic.