Repeated measures by profile analysis

- More than one response *measurement* for each subject. Might be
  - measurements of the same thing at different times
  - measurements of different (but related) things
- Variation: each subject does several different treatments at different times (called *crossover design*).
- Expect measurements on same subject to be correlated, so assumptions of independence will fail.
- Called *repeated measures*. Different approaches, but *profile analysis* uses PROC GLM and looks like MANOVA.

Some fake data

- Here are some data I made up:

```
a 10 10 9 10
a 11 9 10 11
a 10 11 10 9
b 9 10 12 10
b 11 10 10 8
b 11 10 8 9
```

- 6 subjects; 2 treatments A and B, 4 (repeated) measurements of some response (at 4 different times).
- Nothing much happening:
  - no difference between the treatments (no treatment effect)
  - no trend over time (values just “jumping about randomly” for each subject).
- Expect to see no significant test results.
- Imagine plotting mean response (y-axis) vs. time (x-axis), labelling response by treatment — “profile”.

Doing a repeated measures analysis

```
data rm;
  infile "rm1.dat";
  input trt $ y1 y2 y3 y4;

proc glm;
  class trt;
  model y1 y2 y3 y4 = trt / nouni;
  repeated time;
```

- In “model”, put the multiple responses to left of =, like MANOVA.
- `nouni` suppresses univariate ANOVAs (not valid/helpful anyway).
- specify that the 4 responses are measurements at different times.
- Output contains 2 MANOVAs and a univariate ANOVA.

Output for the first analysis

Repeated Measures Level Information

Dependent Variable	y1	y2	y3	y4
Level of time	1	2	3	4

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of no time Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.60922541	0.43	3	2	0.7557
Pillai's Trace	0.39077459	0.43	3	2	0.7557
Hotelling-Lawley Trace	0.64142857	0.43	3	2	0.7557
Roy's Greatest Root	0.64142857	0.43	3	2	0.7557

- No trend over time for either treatment. (No evidence that mean responses at different times are different.)
- Next test time by treatment interaction, also non-significant: no overall difference in response over times, so that non-pattern must be same for both treatment groups.

Last ANOVA for first data set

The GLM Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	1	0.16666667	0.16666667	0.40	0.5614
Error	4	1.66666667	0.41666667		

This tests whether there is a treatment effect, by comparing mean of the 4 response variables for the treatment groups (so is ordinary ANOVA). Not significant either.

Next, change the data to produce a treatment effect but still no time trend:

Data set 2

```
a 10 10 9 10
a 11 9 10 11
a 10 11 10 9
b 11 10 13 11
b 14 12 12 11
b 15 13 9 11
```

- Now treatment B looks to have a slightly higher mean, so we might find a significant treatment effect.
- Still no apparent differences between times, same for each treatment.
- Run same code on this data set (changing only name of data file).

MANOVAs for data set 2

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of no time Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.17789982	3.08	3	2	0.2546
Pillai's Trace	0.82210018	3.08	3	2	0.2546
Hotelling-Lawley Trace	4.62114125	3.08	3	2	0.2546
Roy's Greatest Root	4.62114125	3.08	3	2	0.2546

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of no time*trt Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.23153563	2.21	3	2	0.3263
Pillai's Trace	0.76846437	2.21	3	2	0.3263
Hotelling-Lawley Trace	3.31898971	2.21	3	2	0.3263
Roy's Greatest Root	3.31898971	2.21	3	2	0.3263

No significant difference between times (or difference in pattern of responses over time for the treatments. As we guessed.

Between-subjects analysis for data set 2

The GLM Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	1	20.16666667	20.16666667	30.25	0.0053
Error	4	2.66666667	0.66666667		

Treatment effect we introduced is indeed significant.

Introducing a time effect

Now make another change to data:

```
a 10 10 11 13
a 11 9 12 14
a 10 11 12 12
b 11 10 15 15
b 10 12 14 14
b 12 13 13 15
```

This time responses at times 3 and 4 seem higher, so expect a time effect now. But pattern of responses over time still same for both treatments, so don’t expect a treatment-by-time interaction.

Run the same code again.

MANOVAs for data set 3

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of no time Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.01516477	43.29	3	2	0.0227
Pillai's Trace	0.98483523	43.29	3	2	0.0227
Hotelling-Lawley Trace	64.94230769	43.29	3	2	0.0227
Roy's Greatest Root	64.94230769	43.29	3	2	0.0227

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of no time*trt Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.31515152	1.45	3	2	0.4332
Pillai's Trace	0.68484848	1.45	3	2	0.4332
Hotelling-Lawley Trace	2.17307692	1.45	3	2	0.4332
Roy's Greatest Root	2.17307692	1.45	3	2	0.4332

- Now a significant time effect.
- Time by treatment interaction still not significant because pattern of change over time same for each treatment.

Still a significant treatment effect

The GLM Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	1	15.04166667	15.04166667	36.10	0.0039
Error	4	1.66666667	0.41666667		

because Treatment B numbers still bigger than Treatment A.

Finally. . .

Make one more change to data:

```
a 10 10 14 13
a 11 9 12 14
a 10 11 13 13
b 15 15 11 10
b 14 14 10 12
b 13 15 10 11
```

- Now the time 3 and 4 numbers are bigger for treatment A and smaller for treatment B.
- Effect of time, but different for each treatment.
- So now time by treatment interaction should be significant.

MANOVAs for data set 4

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of no time Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.44926108	0.82	3	2	0.5913
Pillai's Trace	0.55073892	0.82	3	2	0.5913
Hotelling-Lawley Trace	1.22587719	0.82	3	2	0.5913
Roy's Greatest Root	1.22587719	0.82	3	2	0.5913

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of no time*trt Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.01797044	36.43	3	2	0.0268
Pillai's Trace	0.98202956	36.43	3	2	0.0268
Hotelling-Lawley Trace	54.64692982	36.43	3	2	0.0268
Roy's Greatest Root	54.64692982	36.43	3	2	0.0268

- Interaction indeed significant: pattern of change over time depends on treatment.
- Main effect not significant because mean scores for each time (over all the data) aren't very different.

There is still a treatment effect

The GLM Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	1	4.16666667	4.16666667	25.00	0.0075
Error	4	0.66666667	0.16666667		

In summary

- Hard to understand what all the tests are showing, so manipulated data to produce results we could guess (for easier understanding).
- Test of time effect called test for “flatness” of profiles.
- Test of time by treatment(s) interaction called test of “parallelism” of profiles.
- Test of treatment effects called test of “levels”.

A more realistic example

- Do subjects from different professions differ in what they think about different leisure activities?
- 3 occupational groups, bellydancers, politicians and administrators; 5 subjects from each group.
- Each subject participates in 4 activities, reading, dancing, TV-watching, skiing; rates satisfaction with each on 10-point scale.
- Data like this. (Scores on activities as listed.)

```
bellydancer 7 10 6 5
bellydancer 8  9 5 7
bellydancer 5 10 5 8
politician  4  4 4 4
politician  6  4 5 3
politician  5  5 5 6
admin       3  1 1 2
admin       5  3 1 5
admin       4  2 2 5
```

- Profession group plays role of treatment, activity plays role of time.

Some means

Group	Reading	Dancing	TV	Skiing	Activities
Bellydancers	6.6	9.4	5.8	7.4	7.3
Politicians	5.0	4.8	5.2	5.3	5.0
Administrators	5.0	2.0	1.8	3.8	3.2
Groups	5.3	5.4	4.3	5.4	5.2

- Mean scores for each activity overall quite similar.
- Mean scores for each profession group very different.
- Bellydancers like dancing; administrators hate everything but reading.
- Are any of these differences significant?

Repeated measures code

- Code:

```
options linesize=75;
data profile;
  infile "profile.dat";
  input group $ read dance tv ski;
proc glm;
  class group;
  model read dance tv ski = group / nouni;
  repeated activity;
```

- `group` is profession group.
- “repeated” line says that the responses are all “activities”.
- “Nouni”: omit separate 1-way analyses by activity.

Output (edited)

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of no activity Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.27913735	8.61	3	10	0.0040
Pillai's Trace	0.72086265	8.61	3	10	0.0040
Hotelling-Lawley Trace	2.58246571	8.61	3	10	0.0040
Roy's Greatest Root	2.58246571	8.61	3	10	0.0040

MANOVA Test Criteria and F Approximations for
the Hypothesis of no activity*group Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.07627855	8.74	6	20	<.0001
Pillai's Trace	1.43341443	9.28	6	22	<.0001
Hotelling-Lawley Trace	5.42784967	8.73	6	11.714	0.0009
Roy's Greatest Root	3.54059987	12.98	3	11	0.0006

NOTE: F Statistic for Roy's Greatest Root is an upper bound.
NOTE: F Statistic for Wilks' Lambda is exact.

Output part 2

- Significant difference in mean scores (for all the subjects) over activities, even though overall means were not that different.
- The pattern of scores over activities is definitely different for each profession group.

Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
group	2	172.9000000	86.4500000	44.14	<.0001
Error	12	23.5000000	1.9583333		

- Those different mean scores (over activities) for each profession are very clearly significantly different.

Another example: histamine in dogs

- 8 dogs take part in experiment.
- Dogs randomized to one of 2 different drugs.
- Response: log of blood concentration of histamine 0, 1, 3 and 5 minutes after taking drug. (Repeated measures.)
- Data in dogs2.dat.

The code

```
options linesize=75;

data dogs;
  infile "dogs2.dat";
  input Drug $ x $ lh1 lh2 lh3 lh4;
  avg=(lh1+lh2+lh3+lh4)/4;

proc glm;
  class Drug;
  model lh1 lh2 lh3 lh4 = Drug / nouni;
  repeated Time;
  lsmeans Drug;

proc glm;
  class Drug;
  model avg=Drug;
  lsmeans Drug;
```

Comments on code

- Calculate mean of 4 responses (`avg`).
- Do repeated measures analysis.
- `lsmeans` convenient way to get means on 4 variables for each Drug.
- Also do ordinary ANOVA using average log-histamine level as response, and obtain means.

Output part 1

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of no Time Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.05012095	25.27	3	4	0.0046
Pillai's Trace	0.94987905	25.27	3	4	0.0046
Hotelling-Lawley Trace	18.95173763	25.27	3	4	0.0046
Roy's Greatest Root	18.95173763	25.27	3	4	0.0046

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of no Time*Drug Effect

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.10523944	11.34	3	4	0.0200
Pillai's Trace	0.89476056	11.34	3	4	0.0200
Hotelling-Lawley Trace	8.50214058	11.34	3	4	0.0200
Roy's Greatest Root	8.50214058	11.34	3	4	0.0200

Comments and drug-effect analysis

- The histamine levels do change over time, and the pattern of change differs for the 2 drugs (though latter P-value not *very* small).
- Analysis of drug effect:

The GLM Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Drug	1	11.52000000	11.52000000	3.13	0.1274
Error	6	22.10263750	3.68377292		

Averaging over time, no significant difference between drugs.

LSMEANS

The GLM Procedure
Least Squares Means

Drug	lh1 LSMEAN	lh2 LSMEAN	lh3 LSMEAN	lh4 LSMEAN
Morphine	-2.89000000	-1.16000000	-1.99750000	-2.32500000
Trimetha	-3.02250000	0.13000000	-0.17250000	-0.50750000

Both drugs show increase (to time 2) then decrease. (Time effect.) Rate of decrease smaller for Trimetha (time-drug interaction effect).

The second PROC GLM, edited

Source	DF	Squares	Mean Square	F Value	Pr > F
Drug	1	2.88000000	2.88000000	3.13	0.1274
Error	6	5.52565938	0.92094323		
Corrected Total	7	8.40565938			

The GLM Procedure
Least Squares Means

Drug	avg LSMEAN
Morphine	-2.09312500
Trimetha	-0.89312500

- P-value identical to last part of repeated measures analysis.
- Drug means look different, but not different enough to be significant.