Analysis of variance

- Analysis of variance used with:
  - counted/measured response
  - categorical explanatory variable(s)
  - that is, data divided into groups, and see if response significantly different among groups
  - or, see whether knowing group membership helps to predict response.
- Typically two stages:
  - F-test to detect *any* differences among/due to groups
  - if F-test significant, do *multiple comparisons* to see which groups significantly different from which.
  - Need special multiple comparisons method because just doing (say) two-sample t-tests on each pair of groups gives too big a chance of finding “signficant” differences by accident.

Example: jumping rats

- Link between exercise and healthy bones: exercise stresses bones and helps them grow stronger.
- Study assessed effect of jumping on bone density of rats. Rats randomly assigned to one of 3 treatment groups:
  - no jumping (control)
  - low-jump (30 cm)
  - high-jump (60 cm)
- 8 jumps/day, 5 days/week, measure bone density (response) at end.
- PROC GLM to analyze (or PROC ANOVA, only works for balanced designs).

The data

- Some of the data (10 rats in each group). Data separated by tabs.

```
Control     1       603
Control     1       569
...
Lowjump     2       635
Lowjump     2       605
...
Highjump    3       643
Highjump    3       650
```

Code

- Code below. Note format for reading tab-separated data.

```
options linesize=70;

data jumping;
  infile "jumping.dat" delimiter='09'x;
  input group $ g density;

proc means;
  var density;
  class group;

proc glm;
  class group;
  model density=group;
  means group / tukey;
  means group / bon;
```

Comments

- "Straightforward" one-way ANOVA.
- Get table of group means and SDs. Assumption: population SD in each group the same, so sample SDs should be "not too different".
- Tukey's method asks: "how far apart might lowest and highest sample group means be, if population means all same?". Anything larger than that declared significantly different.
- Bonferroni's method allows for number of paired comparisons, in general for n groups is $n(n-1)/2$, here 3: divide α by 3 for each test (eg. $0.05/3 = 0.0167$). More "conservative" than Tukey.

Output part 1

Analysis Variable : density

group	N Obs	N	Mean	Std Dev	Minimum	Maximum
Control	10	10	601.1000000	27.3636011	554.0000000	653.0000000
Highjump	10	10	638.7000000	16.5935061	622.0000000	674.0000000
Lowjump	10	10	612.5000000	19.3290225	588.0000000	638.0000000

Dependent Variable: density

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	7433.86667	3716.93333	7.98	0.0019
Error	27	12579.50000	465.90741		
Corrected Total	29	20013.36667			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
group	2	7433.866667	3716.933333	7.98	0.0019

Source	DF	Type III SS	Mean Square	F Value	Pr > F
group	2	7433.866667	3716.933333	7.98	0.0019

Notes

- Sample SDs not too different. (Argue that rats were randomly assigned to groups, so population SDs necessarily same.)
- F-tests for model as a whole and for groups (same) significant: there is effect of jumping on bone density. Use multiple comparisons to see what: Tukey then Bonferroni.

Tukey

```
Tukey's Studentized Range (HSD) Test for density

Minimum Significant Difference        23.934

Means with the same letter are not significantly different.

Tukey Grouping          Mean       N    group

             A       638.700      10    Highjump

             B       612.500      10    Lowjump
             B
             B       601.100      10    Control
```

High jumping has a significantly different (better) effect on bone density; no signficant difference between low jumping and control.

Bonferroni

```
Bonferroni (Dunn) t Tests for density

Minimum Significant Difference   24.639

Means with the same letter are not significantly different.

 Bon Grouping          Mean      N    group

            A       638.700     10    Highjump

            B       612.500     10    Lowjump
            B
            B       601.100     10    Control
```

- Here, same conclusions as before.
- But note min sig difference, 24.639, larger than Tukey (23.934).
- Bonferroni has harder job finding significant differences if they exist.

Another example: scaffolds

- Repair serious wounds by inserting material as “scaffold” for body’s repair cells to use as template for new tissue.
- Scaffolds made from extracellular material (ECMs) promising (made from biological material).
- Study: use mice to compare effects of 6 types of material.
- Response: % glucose phosphated isomerase (GPI) cells in region of wound: higher better.
- GPI measured 2, 4, 8 weeks after tissue repair.
- 3 mice for each combo of material (6) and weeks (3): 54 total.
- Data: material, weeks, GPI.
- See whether GPI depends on either/both of material and weeks or their interaction.

Data

```
ecm1  2   70
ecm1  2   75
ecm1  2   65
ecm1  4   55
ecm1  4   70
ecm1  4   70
ecm1  8   60
ecm1  8   65
ecm1  8   65
ecm2  2   60
...
mat3  8   5
mat3  8   15
mat3  8   10
```

Code

```
options linesize=75;

data scaffold;
  infile "scaffold.dat";
  input material $ weeks gpi;

proc glm;
  class material weeks;
  model gpi=material|weeks;
```

- Declare “weeks” as a categorical variable too (look for any differences among weeks), then fit model saying GPI depends on both and interaction too.
- The | between material and weeks means “fit interaction as well as main effects”.
- (Looking to see whether interaction significant first, then decide what to do next.)

ANOVA output

```
                    The GLM Procedure
Dependent Variable: gpi
                                     Sum of
 Source                     DF      Squares   Mean Square  F Value  Pr > F

 Model                      17  37609.25926    2212.30937    86.88  <.0001
 Error                      36    916.66667      25.46296
 Corrected Total            53  38525.92593
...
 Source                     DF    Type I SS   Mean Square  F Value  Pr > F
 material                    5  35659.25926    7131.85185   280.09  <.0001
 weeks                       2    867.59259     433.79630    17.04  <.0001
 material*weeks             10   1082.40741     108.24074     4.25  0.0006

 Source                     DF  Type III SS   Mean Square  F Value  Pr > F
 material                    5  35659.25926    7131.85185   280.09  <.0001
 weeks                       2    867.59259     433.79630    17.04  <.0001
 material*weeks             10   1082.40741     108.24074     4.25  0.0006
```

Look at interaction test (bottom line) first: significant, so don't do any other tests. GPI depends on weeks in different way according to materials.

Doing Tukey for interactions

Requires a trick: make new variable that is material-week combination, then do *1*-way ANOVA on that, looking only at Tukey output:

```
data scaffold;
  infile "scaffold.dat";
  input material $ weeks gpi;
  mw=cat(material,"-",weeks);

proc glm;
  class mw;
  model gpi=mw;
  means mw / tukey;
```

If you check, “model SS” same for this analysis as for original one.

Tukey output

Tukey Grouping			Mean	N	mw	
	A		73.333	3	ecm3	-8
	A		73.333	3	ecm3	-4
	A		71.667	3	ecm3	-2
	A		70.000	3	ecm1	-2
	A		65.000	3	ecm1	-4
	A		65.000	3	ecm2	-2
B	A		63.333	3	ecm1	-8
B	A		63.333	3	ecm2	-8
B	A		63.333	3	ecm2	-4
B			48.333	3	mat1	-2
	C		26.667	3	mat3	-2
D	C		23.333	3	mat1	-4
D	C	E	21.667	3	mat1	-8
D	C	E	11.667	3	mat3	-4
D		E	10.000	3	mat2	-2
D		E	10.000	3	mat3	-8
		E	6.667	3	mat2	-8
		E	6.667	3	mat2	-4

Interpretation

- Complicated, because of overlapping lines.
- No sig. differences among ECMs.
- ECMs all better than MATs except mat1 at 2 weeks.
- Other MATs worse, with complicated pattern of significant differences.
- No consistent pattern of which #weeks best for each material (explains significant interaction).
- Next step should be: MAT materials no good, so do another experiment on just ECMs.
- We cheat — extract data for just ECMs!

Just the ECMs: code

First do the same analysis again, checking for significant interaction:

```
data scaffold;
  infile "scaffold2.dat";
  input material $ weeks gpi;

proc glm;
  class material weeks;
  model gpi=weeks|material;
```

Interaction test

```
                         The GLM Procedure

Dependent Variable: gpi
                                         Sum of
 Source                      DF         Squares   Mean Square  F Value  Pr > F

 Model                        8      468.518519     58.564815     1.62  0.1874
 Error                       18      650.000000     36.111111
 Corrected Total             26     1118.518519

 Source                      DF       Type I SS   Mean Square  F Value  Pr > F
 weeks                        2      24.0740741    12.0370370     0.33  0.7209
 material                     2     385.1851852   192.5925926     5.33  0.0152
 material*weeks               4      59.2592593    14.8148148     0.41  0.7989

 Source                      DF     Type III SS   Mean Square  F Value  Pr > F
 weeks                        2      24.0740741    12.0370370     0.33  0.7209
 material                     2     385.1851852   192.5925926     5.33  0.0152
 material*weeks               4      59.2592593    14.8148148     0.41  0.7989
```

No significant interaction (very bottom line), so re-run analysis without (and do Tukey accordingly).

Revised code

Read data as before, and then this:

```
proc glm;
  class material weeks;
  model gpi=weeks material;
  means material weeks / tukey;
```

- Note lack of `|` in `model` line, and back to regular Tukey.
- No interaction means effect of weeks on GPI same for each material, and effect of material on GPI same for each number of weeks.
- So get separate Tukeys to see which materials best, which #weeks best.

The ANOVA

The GLM Procedure

Dependent Variable: gpi

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	409.259259	102.314815	3.17	0.0335
Error	22	709.259259	32.239057		
Corrected Total	26	1118.518519			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
weeks	2	24.0740741	12.0370370	0.37	0.6927
material	2	385.1851852	192.5925926	5.97	0.0085

Source	DF	Type III SS	Mean Square	F Value	Pr > F
weeks	2	24.0740741	12.0370370	0.37	0.6927
material	2	385.1851852	192.5925926	5.97	0.0085

Significant effect of materials, but not of #weeks.

Tukey

```
Minimum Significant Difference        6.7238

Means with the same letter are not significantly different.

Tukey Grouping          Mean      N    material

             A        72.778      9    ecm3
             A
        B    A        66.111      9    ecm1
        B
        B             63.889      9    ecm2
```

- Means more than 6.72 different significantly different.
- ecm3 better than ecm2.
- ecm1 in curious middle ground: not sig. worse than ecm3, not sig. better than ecm2.
- Not enough data to resolve this (ecm1 and ecm3 “almost” sig. different).

Tukey for weeks

No sig. difference due to weeks, so shouldn't really even look at Tukey, but results not surprising:

```
               Minimum Significant Difference        6.7238

        Means with the same letter are not significantly different.

         Tukey Grouping          Mean      N    weeks

                      A        68.889      9    2
                      A        67.222      9    4
                      A        66.667      9    8
```