Repeated measures analysis

The data for this exercise came from the Laboratory School of Univ. of Chicago. They consist of scores from a cohort of pupils in grades 8-11 on the vocabulary section of the Cooperative Reading Test. The scores are scaled to a common, but arbitrary origin and unit of measurement, so as to be comparable over the four grades. There is no between-S effect (e.g., gender is unknown).

Since these data cover an age range in which physical growth is beginning to decelerate, it is of interest whether a similar effect occurs in the acquisition of new vocabulary. That is, we are particularly interested in linear and quadratic trends.

1. Read the data, **`vocab.sas`** into SAS. (It is in **N:\psy6140\data**).

```
%include data(vocab);
proc print data=vocab(obs=20); run;
```

2. Set up a one-way repeated measures ANOVA / MANOVA for these data. Use the SAS help to find the syntax required to ask for orthogonal polynomial contrasts on the REPEATED statement.:

```
proc glm data=vocab;
     model grade8-grade11 = /nouni;
     repeated grade 4  ???? /
         short summary;
     title 'Multivariate Repeated Measures Analysis';
      run;
```

3. For other purposes, it is useful to restructure the data into ‘long’ format. Here, we would like to have one numeric variable, GRADE and another variable VOCAB containing the vocabulary score.

```
data vlong;
    set vocab;
    keep subject grade vocab;
    grade=8;  vocab=grade8;  output;
    grade=9;  vocab=grade9;  output;
    grade=10; vocab=grade10; output;
    grade=11; vocab=grade11; output;
    run;
proc print data=vlong(obs=20); run;
```

4. Before going further, let’s plot the data. See if you can get the meanplot program to add a linear regression line, and a quadratic regression curve.

```
%meanplot(data=vlong, var=..., class=..., interp=...);
```

5. There are several other ways to analyze this data. One is to treat grade as a (polynomial) regression effect. Compare the results to those you obtained using the repeated measures approach.

```
proc glm data=vlong;
      class subject;
      model vocab = subject grade|grade|grade / ss1;
      run;
```

What is the difference if you specify the model as follows?

```
proc glm data=vlong;
      class subject grade;
      model vocab = subject grade;
      contrast %poly(4, 1, grade);
      contrast %poly(4, 2, grade);
      contrast %poly(4, 3, grade);
      run;
```

See if you can relate the result of these analyses to the initial one in step 2.

6. Finally, the same data can be analyzed as a *longitudinal mixed model*. We won't take up this topic now in any detail, but while we're at it, here is a model you can run with PROC MIXED that specifies random intercepts and slopes for the growth trajectories of individual subjects.

```
proc mixed data=vlong;
      class subject;
      model vocab = grade / solution;
      random intercept grade / subject=subject type=un;
```

(In this application, we would also be particularly interested in modeling curvature, the effect of grade^2.)

Repeated measure analysis in R

The same data is contained in the heplots package, as VocabGrowth. In R, repeated measure analyses require a separate data frame (`idata`), describing the repeated factors. Note that here we create grade as an *ordered* factor, so that polynomial contrasts will be applied automatically. The lines below are available in the script **N:/psy6140/tutorials/vocab.R**.

```
library(heplots)
library(car)
data(VocabGrowth, package="heplots")

# make a plot
boxplot(VocabGrowth, ylab="Vocabulary score")
means <- colMeans(VocabGrowth)
points(1:4, means, pch=16, cex=1.5, col="blue")
      # linear trend
abline(lm(means ~ I(1:4)), col="blue", lwd=2)
      # quadratic model
quad.mod <- lm(means ~ poly(I(1:4),2))
quad.mod
lines(1:4, predict(quad.mod), col="red", lwd=2)

# Standard Multivariate & Univariate repeated measures analysis
Vocab.mod <- lm(cbind(grade8,grade9,grade10,grade11) ~ 1, data=VocabGrowth)
idata <-data.frame(grade=ordered(8:11))

# basic short summary
(Vocab.aov <- Anova(Vocab.mod, idata=idata, idesign=~grade, type="III"))
# detailed summary
summary(Vocab.aov)

heplot(Vocab.mod, type="III", idata=idata, idesign=~grade, iterm="grade",
      main="HE plot for Grade effect")

### doing this 'manually' by explicitly transforming Y -> Y M
# calculate Y M, using polynomial contrasts
trends <- as.matrix(VocabGrowth) %*% poly(8:11, degree=3)
colnames(trends)<- c("Linear", "Quad", "Cubic")
colMeans(trends)

# test all trend means = 0 == Grade effect
within.mod <- lm(trends ~ 1)
Anova(within.mod)
heplot(within.mod, terms="(Intercept)", col=c("red", "blue"), type="3",
  term.labels="Grade",
  main="HE plot for Grade effect")
mark.H0()
```