NCGS (Fitzmaurice exercise 5.1)

We consider data from the National Cooperative Gallstone Study (NCGS). In this study patients were randomly assigned to high-dose (750 mg/day) or low-dose (375 mg/day) of the drug chenondiol or to a placebo.We focus on a subset of data on patients who had floating gallstones and who were assigned to the high-dose and placebo group. These data are found in the file cholesterol.txt.

In the NCGS it was suggested that chenodiol would dissolve gallstones but in doing so might increase levels of serum cholesterol. As a result serum cholesterol (mg/dL) was measured at baseline and at 6, 12, 20, and 24 months of follow-up. Note that many cholesterol measurements are missing due to missed visits, drop out, or missing or inadequate laboratory specimens.
Note the groups: 1=high dose, 2=placebo.

1. Read the data from the file. It is in the wide format. Use proc sgscatter to plot week 0, week 6, week 12, week 24 in a matrix plot for each treatment.

2. Compute the sample means, standard deviations and variances of the serum cholesterol at each occation for each treatment group.

3. Transform data to long format and use proc sgpanel to construct two spaghetti plots displaying the raw data from each group. Use proc sgpanel to construct boxplots to compare treatments at each occasion.

4. On a single graph, construct at time plot that display the mean serum cholesterol versus time (in months) for the two treatment groups. Describe the characteristics of the time trends in each group.

In 5. and 6. baseline is modelled as a response.
5. Assuming an unstructured covariance matrix, conduct an analysis of response profiles. Determine whether the patterns of change over time differs between the groups. Use lsmestimate to estimate the treatment difference at week 24 with 95% confidence interval. Use lsmeans and ods statement to take out lsmeans. Plot lsmeans.

6. Display the estimated 5x5 covariance and correlation matrices for the five repeated measurements of serum cholesterol.

In 7. and 8. baseline is treated as a covariate
7. Write a model for mean serum cholesterol with baseline as a week specific covariate and with effects of week, trt and week*trt. Use lsmestimate to estimate lsmeans at week 24 and treatment difference at week 24. Use the store statement to store the sufficient statistics.

8. Write the model in 7. using the nested structure within week, and lsmestimate to estimate treatment difference at week 24. Use the lsmeans statement with the slice option to test for treatment difference at week 6, 12, 20 and 24.

9. Use proc plm to estimate the treatment difference at week 24, and to make an interaction plot.

10. Compare the estimated treatment differences found in 5 and 7. As extensions, change to treatment specific covariance matrices and discuss whether the assumption of equal variance matrices is adequate. Does the change of covariance matrix change the estimated treatment effect at week 24?