**Stat 510 Week 10 Homework Solutions**

1. Use the glucose.csv dataset. The data are from a study in which the effects of three different foods on serum glucose levels in humans are compared. For each person, serum glucose level was measured at 15, 30, and 45 minutes after eating the food. There were four different persons per food, for a total n = 12 persons in the study.

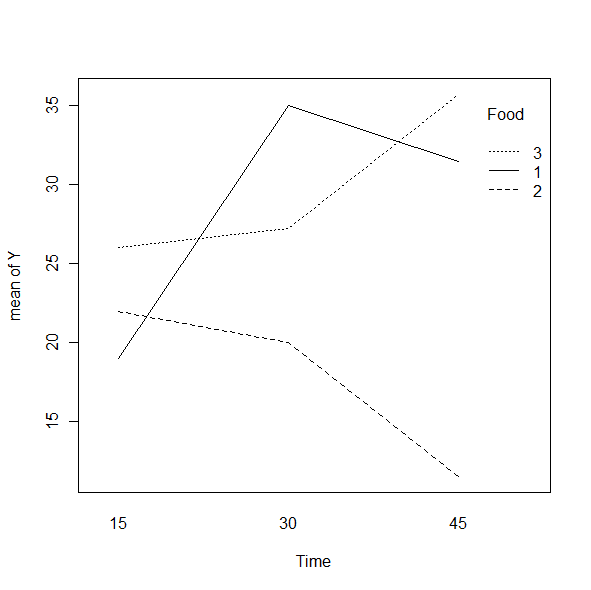
Variables in the dataset are Food (1, 2, or 3), Subject (numbered 1 to 12), Time (15, 30, 45) and Gluc (short for Glucose, the response variable).

A. Create an interaction plot that gives mean response by food type and time. Give the plot and briefly describe the Food, Time, and interaction effects.

**Overall, the average value for Foods 1 and 3 are generally above that for Food 2.**

**For foods 1 and 3, mean Y increases over time whereas for Food 2, the mean decreases.**

**There is an interaction (as described just above). The plot is**



B. Using gls in R, fit repeated measures models for each of the following assumptions about correlations among measurements at different times:

* Compound symmetry
* Completely unstructured with possibly unequal variances
* AR(1)
* AR(1) with unequal variances

Compare the AIC and BIC values for the four different models using the anova command. Give the results and explain which model looks to be the best.

**Model df AIC BIC logLik Test L.Ratio p-value**

**fit.compsym 1 11 180.8759 195.1301 -79.43794**

**fit.nostruct 2 15 180.4815 199.9191 -75.24076 1 vs 2 8.394358 0.0782**

**fit.ar1 3 11 182.0474 196.3016 -80.02371 2 vs 3 9.565902 0.0484**

**fit.ar1het 4 13 183.5032 200.3491 -78.75158 3 vs 4 2.544248 0.2802**

**On the basis of the BIC, the compound symmetry model is the best (by just a bit compared to the AR(1) errors model). But, with AIC, the unstructured errors model is best, although it’s basically a tie with the compound symmetry model.**

C. For the best model, use the anova command to determine F-test results for Food, Time, and the Food\*Time interaction. Give the results and indicate which effects are statistically significant.

**Answer will depend on whether you picked the compound symmetry model or the unstructured errors model. (AIC and BIC give different minimums.)**

**For the compound symmetry model, the ANOVA is**

**numDF F-value p-value**

**(Intercept) 1 503.0710 <.0001**

**factor(Food) 2 11.1121 3e-04**

**factor(Time) 2 11.9493 2e-04**

**factor(Food):factor(Time) 4 30.5345 <.0001**

**For the unstructured errors model, the ANOVA is**

**numDF F-value p-value**

**(Intercept) 1 543.2401 <.0001**

**factor(Food) 2 30.8966 <.0001**

**factor(Time) 2 6.7605 0.0042**

**factor(Food):factor(Time) 4 67.3480 <.0001**

**All factors are statistically significant in both models.**

2. Use the dataset termites.csv from the Datasets folder. The data are from an experiment to investigate the effects of certain tree resins on termites. A resin derived from the bark of tropical trees is dissolved in a solvent and placed on filter paper in different concentrations (5mg and 10mg). For each dosage level, eight dishes are set up with 25 termites in each dish. The Y variable is the number of termites still alive after a number of days. The y-variable is measured at 2, 4, 6, 8, 10, 12, and 14 days from the beginning of the experiment. Variable names in the dataset are Dish, Dose, Time, and Y.

A. Create an interaction plot that shows the mean response for each dose at each time. Time should be on the horizontal. Give the plot and write a brief description of the Dose, Time and Dose by Time interaction effects.

**Overall the mean is less for Dose = 10 than for Dose = 5. Over time, mean Y decreases. There may be a slight interaction as the difference between mean Y for dose = 5 and dose = 10 seems to increase a bit over time.**



B. Using gls in R, fit repeated measures models for each of the following assumptions about the correlations among observations at different times:

* Compound symmetry
* AR(1) with equal variances

Then, compare the AIC and BIC values using the anova command. Give the results and comment on which model provides the better fit.

**The AR(1) errors model gives the better fit.**

**Model df AIC BIC logLik**

**fit.compsym 1 16 549.8521 591.2115 -258.9260**

**fit.ar1 2 16 509.8885 551.2480 -238.9442**

C. Use the anova command to get ANOVA F-test results for the model with the compound symmetry assumption. Give the results and briefly describe them.

**numDF F-value p-value**

**(Intercept) 1 140.3743 <.0001**

**factor(Dose) 1 11.5200 0.001**

**factor(Time) 6 107.7644 <.0001**

**factor(Dose):factor(Time) 6 7.2400 <.0001**

**All factors are statistically significant.**

D. Use the anova command to get ANOVA F-test results for the model with the AR(1) assumption about correlations between measurements at different times. Give the results and briefly describe how they differ from the results in part C.

**numDF F-value p-value**

**(Intercept) 1 250.14551 <.0001**

**factor(Dose) 1 7.49346 0.0074**

**factor(Time) 6 42.36379 <.0001**

**factor(Dose):factor(Time) 6 8.34215 <.0001**

**All in all, there’s no substantive difference in the results as all factors are statistically significant in both models. The F values are smaller for the main effect factors (especially dose) in the AR(1) model. The F value for interaction is slightly larger in the AR(1) model.**

E. Now use a cubic orthogonal polynomial to describe the time trend along with an AR(1) assumption about correlations between measurements at different times. (Note: With seven time points it would be possible to fit a 6th degree polynomial, but powers past the 3rd power don’t seem very efficient, so we’ll use a cubic.

Use the summary command to summarize the coefficient estimates. Briefly describe what is shown about the overall time trend and how the doses differ with regard to the time trend.

**Coefficients:**

**Value Std.Error t-value p-value**

**(Intercept) 17.31932 1.374190 12.603294 0.0000**

**factor(Dose)10 -7.71294 1.943398 -3.968792 0.0001**

**poly(Time, degree = 3)1 -53.56562 6.217058 -8.615911 0.0000**

**poly(Time, degree = 3)2 2.70761 3.571091 0.758203 0.4500**

**poly(Time, degree = 3)3 3.96439 2.659799 1.490483 0.1391**

**factor(Dose)10:poly(Time, degree = 3)1 -4.82441 8.792247 -0.548712 0.5844**

**factor(Dose)10:poly(Time, degree = 3)2 27.67688 5.050286 5.480259 0.0000**

**factor(Dose)10:poly(Time, degree = 3)3 -13.50800 3.761523 -3.591099 0.0005**

**We see a statistically significant difference between Dose 10 and Dose 5.**

**The linear polynomial effect is significant as a main effect while the quadratic and cubic effects are not.**

**The breakdown of the interaction is:**

**No difference between the two doses with regard to the linear effect.**

**There is a difference between the two doses with regard to both the quadratic and the cubic effects.**