STAT 700

Problem: Orthodontic Growth Curves (cont.): We will use the R dataset Orthodont (in the `nlme` library).

(a) Following the Rat Pup R code, make a new variable sex1 which is a 1 for Female and 0 for Male. Let's center the ages, so use `I(age-11)` in your R function calls. Fit a l inear mixed model using REML, that has fixed effects that include sex1, I(age-11), sex1:I(age-11) interaction, an intercept and a random effect for just the intercept. Call this `fit1`.

Note: You can read about centering in your textbook Section 2.9.5 Centering Covariates.

(b) Use the LRT to test $H_0 : \sigma^2_{intercept} = 0$ against $H_1 : \sigma^2_{intercept} \neq 0$. State your conclusion and state which is the "best" model. Include and examine the plot of the residuals and the Q-Q plot of the residuals for the best model. Does it appear that the residuals are iid $N(0, \sigma^2)$?

(c) Examine the summary of the model `fit1` from part (a). Is there evidence to suggest that boys and girls have significantly different growth patterns? Explain.

(d) Make a plot of the standardized residuals versus fitted values by gender. (Hint: see the R code for the `ergoStool` data analysis and use `resid(., type=''p'')` option). Does it look like Males and Females have the same residual variance, σ^2?

(e) You can test part (d) as in the Rat Pup Lab, but for this homework just use your model from part (a) with the `VarIdent` option to allow for different residual variances for each gender. Call this model `fit1b`. Make a scatter plot of the standardized residuals versus fitted values for your heteroscedastic fit by gender. Now, does it look like Males and Females have the same residual variance?

(f) Which model `fit1` or `fit1b` in part (e) is the "best" model based on AIC? Plot the predicted values over time for all the subjects from the best model.