Week 13 homework problems

1. Do problem **7.8**. In part b., the textbook contains a significant mistake. It should say, "WLS should be used with variance function Var(`Weight`|`Age`)=$\texttt{SD}^2\sigma^2/n$." The point that the author is trying to make is that the applicable weights are $n_i/\texttt{SD}^2$. Skip parts **7.8.4** and **7.8.5**.

2. Use the `salarygov` data referenced in problem **7.4**. Although the response (`MaxSalary`) is a maxmimum of, rather than a mean of, sub-observations, fit the WLS model with weights that represent rows' differing sample sizes. Your model should include the predictor `Female_dominated`, the spline bases for `Score`, and the interaction terms between these. A description of how to create `Female_dominated` is given in **5.9.3**. For the splines, use B-splines with 3 degrees of freedom. Once the model is fit, do the following:

   a. Report the fitted model.

   b. Perform a partial $F$-test on the interaction terms to determine if female-dominated occupations require different spline coefficients than other occupations.

   c. Give the residuals-vs.-fitted-values plot from the model that includes the interaction terms and interpret the plot in terms of model assumptions.

3. The `Blackmore` data set in `alr4` provides the number of hours of exercise performed each week by 236 teenage girls at five different ages. It also provides a categorical indicator of whether the subject was hospitalized for an eating disorder. Do the following:

   a. Fit a mixed model that controls for age and group as fixed effects and has a random intercept for subject. Give the estimated variance component for `subject` and interpret it.

   b. Test that the variance component for `subject` is equal to 0 using a likelihood ratio test. Report a test statistic, $p$-value, and your conclusion. (Hint: when testing between two models in the `anova()` function, we have heretofore always put the *simpler* model first. But in cases like this, where the simple model is just a regular fixed-effects model and the more complex model is a mixed-effects model, the function gives a strange error unless you put the *mixed-effect* model first.

   c. Produce a normal probability plot of the predicted random effects for `subject`. Interpret the plot and what it says about your model.