1. The multivariate normal we've been discussing can be written as

$$Y_i = X_i\alpha + \delta_i, \quad \text{where} \quad \delta_i \sim N(0, \Sigma), \text{ i.i.d for each } i.$$

   Using this model, consider the dental dataset from our text, described in section A.4, and for the subject with id 1, write out what Y_i and X_i are in the following models, and describe what each element of α represents.

   (a) distance depends only on age, linearly

   (b) distance depends on age linearly, with a different line for each gender (that is, an age/gender interaction)

   (c) distance depends only on age, but age is treated as a categorical variable

   (d) distance depends on age, where age is again treated as a categorical variable, and there is also an additive effect of gender

2. In the model with independent errors, where $\delta_i \sim N(0, \sigma^2 I)$, i.i.d. for each i, what is the correlation between two observations from the same individual? Between two observations from different individuals? How can you tell what those correlations are from the notation I used here? How do your answers change when you consider the model as written in question 1 instead?

3. Fit the four models described in question 1, using a random intercepts by id to induce the covariance structure. Remember to use method `ML` as these models are comparing fixed effects.

   (a) For each model, show the estimates for α and use them to find the predicted values for a boy and a girl at age 10.

   (b) Compare the AIC and BIC values for these models. Which is best using these criteria? Do they agree here?

   (c) Compare models (a) vs (b), (c) vs (d), and (b) vs (d) using the likelihood ratio test. Which model is best? Why are these valid comparisons, and not, say (b) vs (c)?

   (d) Do the three criteria agree on a preferred model? If not, which one would you prefer, and why?

4. For the best model you found in question 3, fit four different covariance structures: independent errors, AR(1), random intercept by id, and random intercept and slope by id. Compare them using AIC and BIC. Which do you prefer?