BST 140.752
Problem Set 5

1 LMM

1. Let $Y_{ij} = \mu + u_i + \epsilon_{ij}$ for $u_i \sim N(0, \sigma_u^2)$ and $\epsilon_{ij} \sim N(0, \sigma^2)$. Calculate the BLUP for u_i
2. Let $Y = X\beta + Zu + \epsilon$ for $u \sim N(0, \Sigma_u)$ and $\epsilon \sim N(0, \sigma^2 I)$. Calculate the BLUP for u.
3. Load the Rail data set in R. Fit a mixed model of the form from question 1. Compare a the estimates of the mean for each rail with the empirical mean.
4. Load the pixel data set in R. Fit a linear mixed effect model where you have $Y_{ijk} = \beta_0 + \beta_1 x_k + u_i + u_{ij} + \epsilon_{ijk}$ where Y_{ijk} is pixel, i is dog, j is side and k is day index and x_k is day. Fit the model and interpret the results.
5. Consider the model $Y_i = \mu + \epsilon_{ij}$. Consider putting a so-called "flat" prior on μ. That is acting like a distribution that is 1 from $-\infty$ to $+\infty$ is a valid density. Calculate the distribution marginalized over μ and show that it is the same likelihood used to obtain the REML estimates.

2 FFT

1. Consider a time series Y_t for $t = 1, \ldots, n$. Give the correspondence between the terms of the discrete Fourier transform of the data and the fit using linear models with a trigonometric basis.
2. Write an R program that takes in a chord (the addition of three notes) and guesses the notes. The frequencies for the various notes can be found here `http://www.phy.mtu.edu/ suits/notefreqs.`
3. Consider the Haar wavelet basis for a time series Y_t for $t = 1 ldots, 2^k$. Derive the correspondence between the wavelet coefficients and binned means of the time series.

3 Principal components

1. Formally prove that the population principal components explain the maximum variability subject to being linear combinations of the data and orthogonal to the others.