Due: 5pm, Friday, 26 Apr

1. The data in enzyme.txt are measurements of the velocity of an enzymatic reaction as a series of substrate concentrations. The Michaelis-Menton model is commonly used to model this relationship.

$$\text{E } Velocity_i = \frac{V_m\, Concentration_i}{Concentration_i + K_m}, \tag{1}$$

where the two parameters are V_m, the maximum velocity, and K_m, the half-saturation constant. When $Concentration = K_m$, the expected velocity is one-half of the maximum velocity. Assume that the observations can be modeled as

$$Velocity \sim N(\text{E } Velocity, \sigma^2 \boldsymbol{I}), \tag{2}$$

where E $Velocity$ is given by equation (1).

   (a) Using the data in enzyme.txt, estimate V_m and K_m and their standard errors.

   (b) Is model (2) reasonable? Explain why or why not. Check as much as you can and is appropriate, but you do not have to check independence.

   (c) Estimate the concentration at which the mean reaction velocity is equal to 10.

   (d) Estimate the standard error of your estimate in part 1c.
       Hint: remember the delta method

   (e) Estimate an approximate 95% confidence interval for the estimate in part 1c.
       Hint: A Wald interval is acceptable

   (f) The M-M model, equation (1), can be written in many ways. One of them is to use the concentration at which the mean reaction velocity is equal to 10 as a parameter. I will call this X_{10}. The reparameterized model is:

$$Velocity_i = \frac{Vm\, Concentration_i}{Concentration_i + X_{10}(V_m - 10)/10}.$$

   Estimate the parameters of this model and their standard errors. Are the estimated V_m and its se the same as in part 1a? Are the estimated X_{10} and its se the same as in parts 1c and 1d? Explain why they are the same (or why they should not be the same).

2. The data in cefamadole.csv are pilot-study data on the pharmocokinetics of the drug cfamadole. Six healthy volunteers were injected with 15 mg/kg body weight of cefamandole. Plasma concentrations of the drug were measured at 14 time points. Because the drug was injected into the blood, there is no need for a "stomach" compartment. The only biochemical process is elimination from the blood plasma. The biologists knowledgeable about cefamandole believe there are two separate mechanisms for elimination from the blood. These two mechanisms have different rate coefficients (the rate coefficients are "rate" in the $\exp(-rate * time)$ parts

of the model). As commonly done, the log of the rate coefficients are assumed to have normal distributions. This leads to a "two-exponential" model for plasma concentration over time, which using log rates as the parameters is:

$$Concentration_{ij} = A_{1i}\exp(-\exp(lrc_{1i})*time_j) + A_{2i}\exp(-\exp(lrc_{2i})*time_j) + \varepsilon_{ij},$$

where i indexes subjects and j indexes time. The parameter values for the 6 subjects potentially differ, hence the subscript i. A_1 is the amount of cefamandole eliminated by mechanism 1; $lrc1$ is the rate of that elimination. Both vary by subject. A_2 is the amount of cefamandole eliminated by mechanism 2; $lrc2$ is the rate of that elimination. Both may vary by subject, but preliminary inspection of the data suggests that $lrc2$ may be a constant, not something that varies among subjects. In R, the SSbiexp(Time, A1, lrc1, A2, lrc2) function is a self-start function for this model. If you use SAS, you will have to code it.

(a) Fit the "two-exponential" model with A_1, lrc_1, and A_2 parameters that vary randomly over subjects. Assume that lcr_2 is the same constant for all subjects. Report the estimated mean for each of the four parameters and their se's.

(b) Fit a model that allows the lrc_2 parameter to vary among subjects. Does the data support this model or the simpler model without random variation in lrc_2? Explain why or why not.
Hint: I suggest you use AIC here, because the hypothesis that $\mathrm{Var}(lrc_{2i}) = 0$ is also a hypothesis that all covariances including lrc_{2i} are 0, which we haven't discussed.

(c) The estimated variance of the lrc_{1i} parameter seems small. Is it necessary to include random variation in this parameter in the model? Explain why or why not.
Hint: Again AIC is a useful approach, because this is not a "single parameter on the boundary" test.

(d) Is it appropriate to assume that the conditional variance, i.e. Var $Y_{ij}|A_{1i}, lrc_{1i}, A_{2i}$ is constant? Explain why or why not.

(e) In the summary output from part 2a, the variance component for $lrc1$ looks quite small. Is it appropriate to consider a model where that variance is zero? Explain why or why not.