Notes: You might not finish all the sections—that’s okay! Feel free to choose what problems you find most interesting. I’ve also posted solution code on Wordpress, so you can see how things work out for any parts you don’t finish.

Model Setup and Candidate Models

Compartmental Model of Avastatin. Vervalsing et al. (2010) recently discovered a new anti-angiogenic cancer drug, Avastatin, which been shown to be highly effective on small-cell lung cancer cells in rats. A preliminary study in a small number of patients yielded some data in response to a 150 ug injection of Avastatin, shown in Figure 1.

![Figure 1: Blood plasma concentration of avastatin after 150 µg initial injection.](image)

They are beginning more extensive clinical trials in humans, but do not yet have a pharmacokinetic model of Avastatin distribution in the body. In particular, there remain open questions as to whether Avastatin is eliminated via the urine only (represented here as filtration/elimination from the blood plasma), or if it is metabolized in other tissues as well (represented here as elimination from the tissue compartment in Models 2 and 3 below).

Figure 2 shows three possible models of Avastatin distribution and elimination in humans. The equations for each are:
Model 1: 1-compartment model

\[ \frac{dx_1}{dt} = -k_{01}x_1 \]

Model 2: 2-compartment model, elimination from blood plasma

\[ \frac{dx_1}{dt} = k_{12}x_2 - (k_{01} + k_{21})x_1 \]
\[ \frac{dx_2}{dt} = k_{21}x_1 - k_{12}x_2 \]

Model 3: 2-compartment model, elimination from both plasma and tissue

\[ \frac{dx_1}{dt} = k_{12}x_2 - (k_{01} + k_{21})x_1 \]
\[ \frac{dx_2}{dt} = k_{21}x_1 - (k_{02} + k_{12})x_2 \]

In each model, \( x_1 \) is the mass of the drug in the blood plasma, and \( x_2 \) is the mass of the drug in tissue. The \( k \)'s represent either exchange of drug between plasma and tissue (\( k_{12} \) and \( k_{21} \)) or removal/elimination of the drug from a compartment (\( k_{01} \) and \( k_{02} \)).

**Measurement Model.** For all three underlying models, we will consider the measurement to be the concentration of avastatin in the blood plasma, \( y = x_1/V \) (where \( x_1 \) is the mass of the drug in plasma and \( V \) is the plasma volume).

**Initial Conditions and Parameter Values.** For initial conditions, we’ll use \( x_1(0) = 150 \) \( \mu \)g (representing the 150 \( \mu \)g dose given at \( t = 0 \)), and let \( x_2(0) = 0 \). As initial parameter values, we will start with \( k_{01} = k_{21} = k_{12} = k_{02} = 0.5 \), and \( V = 50 \) dL.

We will examine each model in the sections below. Code to for ODEs for each model is given on Wordpres, and you can also follow the markdown template given on the Wordpres site to help you write the main code.
Part 1: Model Simulation and Parameter Estimation

1) Simulate the models. Using the initial conditions and parameter values given above, simulate all three candidate models and plot the measured concentration output for each model \((y)\) together with the data (i.e. all in one plot).

2) Parameter estimation with ordinary least squares. Estimate the model parameters using maximum likelihood, assuming:

- Normally distributed observation error with a constant \(\sigma = 0.0777\), i.e. we assume the data is given by a normal distribution with mean \(y\) and standard deviation 0.777.

Note this is equivalent to ordinary least squares (although writing it this way changes the cost function value by a scaling factor). When you estimate the parameters, also record the final (best-fit) negative log likelihood value.

- Plot the the fits to the data for each model (i.e. plot \(y\) for each model, as well as the data, all on one plot). Do some models fit better than others? Are there any that you think can be eliminated? Do the final negative log likelihood values for each model make sense for the visual fit that you see?

- Compare the parameter estimates across models—how well do they match?

3) Parameter estimation with weighted least squares. Next, let's try a different assumption about the likelihood. Change the cost function so that we are assuming normally distributed observation error, but now with \(\sigma = 0.1 \times \text{data}\), in other words the standard deviation for each data point is assumed to be 10% of the data value. Re-fit the model and compare the results to the ones you got in 2). Are the parameter estimates different? Do the fits change visually? If so, why do you think this is?

Part 2: Model Comparison

Now that we’ve fitted each of our candidate models, let’s use the Akaike Information Criterion (AIC) to compare them.

1) Intro to AICs. Recall that the AIC is given by \(2(p - LL)\), where \(p\) is the number of estimated parameters and \(-LL\) is the negative log likelihood. Calculate the AICs for each of your models for the weighted least squares case from Part 1. Which model has the lowest (best) AIC? Which model or models would you choose to represent Avastatin pharmacokinetics? Are there any models you can eliminate?

2) AICs for the ordinary least squares case. Now make the same comparison using your parameter estimates from the ordinary least squares case from Part 1.

3) Alternative OLS assumptions. Try fitting your model again with ordinary least squares, but now assuming \(\sigma = 1\). Note that this should not change your parameter estimates or fits at all, since the cost function is minimized at the same point in parameter space. However, now evaluate the AICs again—which model has the lowest AIC now? Why do you think this is?
Part 3: Extra Problems

These are more open-ended, optional problems, to do for fun or if you have extra time.

π) Alternative information criteria. Try out some of the other information criteria, such as the Bayesian Information Criterion (BIC) or Corrected AIC (cAIC). Do they change the ranking of the models?

π + 1) Alternative starting parameters. Try fitting your models from a range of different starting parameter values (particularly for Model 3)—do parameters converge to same solution? What might this suggest?

π + 2) Parameter uncertainty and identifiability. Use likelihood profiles (or MCMC, etc) to evaluate the uncertainty on the model parameters. You may use any of the various techniques we’ve learned so far to evaluate how the likelihood changes as you adjust the parameters, e.g. if you generate likelihood profiles for each parameter, are they flat? Bowl shaped? How does their shape vary across our candidate models?