Homework 1

BIOE 498/598 PJ

Spring 2022

Due Monday, 2/21 before 5pm. Submit answers on Gradescope.

The file `GeneEditing.csv` contains the results of a factorial experiment. The experimenters varied five factors (A–E) and measured the editing efficiency on cultured cells.

1. Load the CSV file into 'R' and look at the data frame. Is this a full factorial or fractional factorial experiment? How can you tell?

2. If the experiment is a fractional factorial, what generator was used?

3. Plot the response ('efficiency') verses the run number. Was there any systematic change during the experiment? How do you know?

4. Create a factor and response plot (farplot) for the data. Based on the plot, which factors do you expect to be active? (There is no wrong answer—we just want your opinion and a justification.)

5. Fit a linear model to the data and show the **effect sizes** in order of decreasing magnitude. Why are some of the effects listed as 'NA'?

6. Create a half-normal plot of the effects. (Remember to omit the 'NA''s.) Which effects are active?

7. If the design is a fractional factorial, compute the confounding pattern for the active effects. Are you still confident that the effects you named above are the active effects? Justify your answer.

8. What factor settings would you use for maximum efficiency gene editing based on your analysis?