Quantitative Understanding in Biology
Problem Set 3: Model Fitting and ANOVA

Jason Banfelder and Luce Skrabanek

Due: 20th October, 2020

1 Model Comparison

In class, we fit data for oxygen binding to myoglobin using a two-parameter Michaelis-Menten-like model that captures the idea that oxygen ultimately saturates myoglobin molecules' binding sites:

$$V = \frac{V_{max}[S]}{K_m + [S]} \tag{1}$$

This model can be extended to include additional terms that represent non-specific binding, and background signal:

$$V = \frac{V_{max}[S]}{K_m + [S]} + M_{ns}[S] + V_{bk} \tag{2}$$

Using the techniques for model comparison covered in this course, evaluate whether a more complex model that incorporates either or both of these additional terms better explains the data.

P_{O_2} (torr)	1.1	1.5	1.6	2.3	3.4	5.3	7.5	8.4	14.1
$[O_2]$ (mL/dL)	1.49	1.79	1.79	2.11	2.83	3.42	3.79	3.97	4.08

Explain the method you used and your reasoning for doing so.

2 ANOVA

The purpose of this exercise is to appreciate the difference between an ANOVA analysis and a series of serial t-tests, and to reinforce the appropriate use of multiple hypothesis testing corrections. You'll use the same data as we used in class, measuring the effect of four drugs on the time it takes for subjects to solve a puzzle:

	ctrl	drgA	drgB	drgC	drgD
1	7.16	9.62	6.19	5.64	10.42
2	6.85	8.05	6.57	6.55	10.85
3	8.10	8.27	7.21	6.79	10.75
4	7.40	9.33	5.97	7.24	10.86
5	7.22	9.49	6.93	7.33	11.61
6	7.83	8.44	6.09	7.24	11.68
7	7.76	9.82	6.91	6.71	11.14
8	6.02	8.19	6.78		11.29
9	8.06		7.16		
10	6.96				

1. **Compare the effect of each drug relative to control using an independent t-test. Without any multiple hypothesis testing corrections, which drugs are deemed to be statistically significant?**
2. **Perform an appropriate multiple hypothesis testing correction. Do any of your conclusions change?**
3. **Re-analyze the data using an omnibus ANOVA analysis followed by appropriate post-tests. Do any of your conclusions change?**
4. **Which method do you think is the most appropriate way to analyze this data?**

© *Copyright 2018-2020 J Banfelder, L Skrabanek, Weill Cornell Medical College*

Logistics

Answers are due by 11:59 PM on Tuesday, 20th October, 2020. E-mail your solution to jbanfelder@rockefeller.edu, las2017@med.cornell.edu, and arb4007@med.cornell.edu

This Problem Set should be completely, and submitted, individually.

Your submission should be generated using R Markdown. Include both the script needed to reproduce your results and generate any figures in your writeup, as well as the compiled HTML file.

© *Copyright 2018-2020 J Banfelder, L Skrabanek, Weill Cornell Medical College*