# M469 Spring 2020, Assignment 1, due Friday Jan. 24

**Suggested reading.** *Introductory science and mathematics education for 21st-century biologists*, by W. Bialek and D. Botstein, in *Science* **303** (2004) 788-790. This non-technical article contains a short discussion of the quantitative nature of modern research in the biological sciences, the difficulties this poses for researchers with inadequate preparation in mathematics, and suggestions for overcoming this problem by changing the standard life sciences undergraduate curriculum. The crux of the matter is summarized in the following excerpt: "Even though most biology students take several years of prerequisite courses in mathematics and the physical sciences, these students have too little education and experience in quantitative thinking and computation to prepare them to participate in the new world of quantitative biology. At the same time, advanced physical science students who become interested in biological phenomena can find it surprisingly difficult to master the complex and apparently unconnected information that is the working knowledge of every biologist."

This article is available electronically from the Evans library. While you're getting a copy, you may as well get (from the same volume) next week's suggested reading: *Uses and abuses of mathematics in biology*, by R. M. May, in *Science* **303** (2004) 790-793.

1. [10 pts] Show how each of the following relations can be put into a form that is linear with regard to (possibly transformations of) the parameters. Note that the values $p_1, p_2, \ldots$ are the parameters.

a. $y = (p_1 x + p_2)^2$.

b. $y = p_1(1 - x_1)^{p_2} p_3^{\ln x_2}$.

c. $y = p_1 x_1^{(p_2 x_2)} (p_3 x_3)^{p_4}$.

2. [10 pts] Suppose a set of $N$ data points $\{(x_k, y_k)\}_{k=1}^N$ appears to satisfy the relationship

$$y = ax + \frac{b}{x},$$

for some constants $a$ and $b$. Find the least squares values for $a$ and $b$.

3. [10 pts] We know from class that if least squares regression is used to fit $N$ data points $\{(x_k, y_k)\}_{k=1}^N$ to a line

$$y = mx + b,$$

the values obtained for $m$ and $b$ can be expressed as

$$m = \frac{N \sum_{k=1}^N x_k y_k - \sum_{k=1}^N x_k \sum_{k=1}^N y_k}{N \sum_{k=1}^N x_k^2 - (\sum_{k=1}^N x_k)^2}$$

$$b = \frac{1}{N}\Big( \sum_{k=1}^N y_k - m \sum_{k=1}^N x_k \Big).$$

Suppose the same data is again fit to a line with least squares regression, but that the fit is based on horizontal distances rather than vertical distances. Find expressions for $m$ and $b$ in this case.

4. [10 pts] Use Problem 3 to find the least squares line for the three points $(1, 2)$, $(3, 3)$, $(4, 5)$, first with vertical distances and second with horizontal distances. Are the same slope and intercept obtained by both fits?