TMATH 390 Spring 2015 **Homework Set 3**

Chapter 2 questions

1. Problem 2.45 on page 95 of the text.

Chapter 3 questions

2. A study to assess the capability of subsurface flow wetland systems to remove biochemical oxygen demand and various other chemical constituents resulted in the accompanying data on $x =$ BOD mass loading (kg/ha/d) and $y =$ BOD mass removal (kg/ha/d) ("Subsurface Flow Wetlands: A Performance Evaluation," Water Envir. Res., 1995: 244-247):

x:	3	8	10	11	13	16	27	30	35	37	38	44	103	142
y:	4	7	8	8	10	11	16	26	21	9	31	30	75	90

   (a) Construct boxplots of both mass loading and mass removal, and comment on any interesting features.

   (b) Construct a scatter plot of the data and comment on any interesting features.

   (c) Make the qqplots for x and y, and comment on whether x and/or y could have come from a Normal distribution.

3. Start with

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{s_x}\right)\left(\frac{y_i - \overline{y}}{s_y}\right).$$

   (a) Start with the formula above, and show that it is equal to

$$r = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2}\sqrt{\sum (y_i - \overline{y})^2}}.$$

   (b) Start from part (a), and show that it is equal to

$$r = \frac{S_{xy}}{\sqrt{S_{xx}\, S_{yy}}}$$

   where S_{xx}, S_{xy}, S_{yy} are defined on page 110.

4. 3.11 in your book.

5. 3.16 in your book.

6. Compute the least squares line for data in problem 2. Use the command "lm($y \sim x$)" in R.

7. 3.25 in your book.

8. Let

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \alpha - \beta\, x_i)^2\,.$$

   Set $\dfrac{\partial}{\partial\alpha}\text{MSE} = 0$, and derive the equation

$$\overline{y} - \alpha - \beta\,\overline{x} = 0.$$

9. In homework 1, you collected data which included data on 2 continuous variables. Call them x and y, depending on which variable you want to predict from the other.
   (a) Produce the scatterplot of x vs. y, and interpret.
   (b) Compute the correlations coefficient between x and y, and interpret.
   (c) Perform linear regression to estimate the regression coefficients, and interpret them.
   (d) Draw the regression line on the scatterplot of part (a). Does it look right?
   (e) Compute R^2 and interpret.