Psychology 310
Homework 1
Fall, 2008

Note: Due in 10 days. You should work on these independently. You may use R to complete or check any part of this assignment. If you do, you should include the R code that you used.

1. (10 points) Given the following data set:

X	*Y*
4	5
1	7
1	3
4	6
9	4

Find

$$\sum_{i=1}^{4} X_i$$

$$\sum_{i=1}^{5} (3X_i + 7)$$

$$\sum_{i=1}^{5} X_i Y_i$$

$$\sum_{i=1}^{2} X_i^2 Y_i$$

$$\sum_{i=1}^{5} (X_i + 3) Y_i$$

2) (8 points) Suppose that you have a set of data that are at a *ratio* level of measurement.

a. What would the level of measurement be if you doubled all the numbers?
b. What would the level of measurement be if you added 4 to all the numbers?
c. What condition would the numbers have to satisfy so that if you squared them, the level of measurement would be *ordinal?*
d. What would the level of measurement be if you divided all the numbers by 4?

3. (12 points). Given the following course grades for 5 students. 78,64,77,91,52. If these grades are lineary scaled to have a mean of 70 and a standard deviation of 12, what grade will the student with a 91 receive? (Show all work, or show all R code used to solve the problem.)

4. (10 points) *X* has a mean of 70 and a standard deviation of 10. *Y* has a mean of 75 and a standard deviation of 12. *X* and *Y* correlate .60. Assume that *X* and *Y* have a bivariate normal distribution.

a. What is the linear regression equation for predicting *Y* from *X?*
b. What is the covariance between *X* and *Y*?
c. Among people who score 80 on *X*, what percentage score above 80 on *Y*?

5. (5 points) A student received a grade of 76 in a course where the class average was 70, and the standard deviation 10. If the class distribution was approximately normal in shape, what was the student's approximate percentile rank?

6. (5 points) If SAT scores have a mean of 500 and a standard deviation of 100, approximately what percentage of students obtain SAT scores between 550 and 680?

7. (10 points) Suppose that the class distribution in a large course is almost exactly normal in shape. Joe got an 88 and had a percentile rank of 79.8, while Felicia got a 75 and had a percentile rank of 40.1. From this information, estimate the mean and standard deviation of the class distribution.

8. (10 points). You have 9 numbers with a mean of 10 and a variance of 100. If you add a 10th number to this group, and this number is 15, what will be the mean and variance of the new list of numbers?

9. (10 points). You have two groups of numbers. The first group has a sample size of 20, a mean of 14, and a variance of 49. The second group has a sample size of 30, a mean of 43, and a variance of 57. If you combine these two groups into one large group of 50 observations, what will the mean and variance of the new group be?

10. (20 points). Write R functions to do the following:
a. For an input list of numbers *x*, linearly rescale the numbers to have a mean of *a* and a standard deviation of *b* and return the rescaled list of numbers.
b. Given the means, variances, and sample sizes of *J* lists of numbers, return the mean and variance that would result if all the numbers were combined into one group. Hint: your input should be a list of means, a list of variances, and a list of sample sizes. The program itself should figure out how many groups there are and proceed to do the calculations, based on the formula in Glass and Hopkins.)

11. (20 point EXTRA CREDIT).

a. Suppose you have two groups, each composed of *n* pairs of observations on the variables *X* and *Y*. Suppose that, within each group, *X* and *Y* have standard deviations of exactly 1, and correlate precisely zero. However, the first group has means of zero on both *X* and *Y*, while the second group has means of 1 on both *X* and *Y*.

a. Derive a formula for the correlation between *X* and *Y* if you combine these two groups into a single group of size 2*n*.
b. Using your knowledge of statistical theory, and preferably using *R*, create two groups of 10 observations each with the exact characteristics described above, and demonstrate that the correlation between *X* and *Y* agrees precisely with your theory.

HINT. Remember that the correlation can be computed solely from a knowledge of $\sum_{i=1}^{n} X_i$, $\sum_{i=1}^{n} X_i^2$, $\sum_{i=1}^{n} X_i Y_i$, $\sum_{i=1}^{n} Y_i$, $\sum_{i=1}^{n} Y_i^2$, and *n* in each group. And, when you stack the groups, sums are added. Remember too that these sums occur in formulas for the variance and mean, and that, when both variables have standard deviations of 1, correlation and covariance are the same. If you choose to attack this problem, and you get stumped, you may come to me for a hint.