Problem Set 1

ECON 340: Economic Research Methods — Instructor: Div Bhagia

1. (1 pt) Is the following statement true or false?

$$\sum_{i=1}^{n}(3X_i^2 + 2k) = 3\left(\sum_{j=1}^{n-1} X_j^2\right) + 3X_n^2 + 2nk$$

   Show the work that led you to your conclusion.

2. (1 pt) Fill in the following frequency distribution table.

X_i	Relative Frequency	Cumulative Frequency
1		0.1
2		0.3
3	0.4	
4		
Total		X

3. (2 pts) I have data on the calorie consumption of five adults as follows:

$$1400, 1800, 2000, 2200, 2600$$

   (a) What is the *mean* calorie consumption for this group?

   (b) What is the *median* calorie consumption for this group?

   Now let's say I add one more individual to the data set whose calorie consumption is 4100.

   (c) What is the *mean* calorie consumption now?

   (d) What is the *median* calorie consumption now?

(e) Which of the two is more sensitive to outliers, *median* or *mean*? Why do you think that's the case?

4. (1 pt) Here is the amount (in $) that I spent on groceries in the last three weeks:

$$100, 120, 80$$

Calculate the variance of my grocery spending during these weeks. (Write down the formula you used and then plug in the values so I can see how you calculated your answer.)

5. (3 pts) The following table is constructed from a sample of 6 students. X_i represents the number of hours an individual usually sleeps and Y_i represents the number of hours the individual typically exercises per week.

Obs	X_i	Y_i	$(X_i - \bar{X})$	$(Y_i - \bar{Y})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$
1	8	3	0.5	-1	0.25	1	-0.5
2	7	4	-0.5	0	0.25	0	0
3	6.5	2	-1	-2	1	4	2
4	7.5	4	0	0	0	0	0
5	9	6	1.5	2	2.25	4	3
6	7	5	-0.5	1	0.25	1	-0.5
Total	**45**	**24**	**0**	**0**	**4**	**10**	**4**

(a) What is the variance of X and Y?

(b) What is the standard deviation of X and Y?

(c) How many standard deviations is the fifth observation away from the average hours of sleep?

(d) What is the covariance between hours of sleep per night and hours of exercise per week?

(e) What is the correlation between X and Y?

For each of the parts, write down the formula you used and then plug in the values so I can see how you calculated your answer.

6. (2 pts) We asked 1000 individuals whether they watched the movie Barbie in the theatre. We create a variable X_i that takes value 1 if the individual watched the movie in the theatre and 0 if they did not. 200 individuals said they watched it in the theatre, while 800 individuals said they didn't. Calculate the mean and variance of X_i. (Show your work.)