Exercise Set 1
Math 350
Fall 2009

Be sure to go through the Matlab handout first.

1. In Matlab, run the "Sales simulator" and print the histogram out to a file. The code is given below. Lines beginning with a percent sign are comments.

```
%Sales simulation:  Look at the profit given by the model
% L=leads, C=cost of single lead, R=percent of leads resulting in sales
% P=sales price, H=overhead (fixed)
%All variables are stored in the order:  LCRPH

H=800;
AllV=[1200 1800;0.2 0.8;0.01 0.05;47 53];
N=2000;  %Number of simulations

P=zeros(1,N);  %Storage for profit values

for j=1:N
    k=rand(4,1);  %Get 4 random numbers between 0 and 1
    Vals=AllV(:,1)+(AllV(:,2)-AllV(:,1)).*k;
    Income=Vals(1)*Vals(3)*Vals(4);
    Cost=H+Vals(1)*Vals(2);
    Profit(j)=Income-Cost;
end
```

2. In Matlab, solve the following problem using simulation:

   If the probability of rain is p percent for each day of the week, what is the probability of getting three rainy days in a row for the week?

   Here is the main simulation program:

```
function [y,X]=rainyweek(p)
%Simulation:  Given that the probability of rain is p percent for each day
%of the week, what is the probability of having three rainy days in a row?
```

```
%Perform a simulation to answer this question.
%SAMPLE USEAGE:  [y,X]=rainyweek(0.5);

y=0; C=0;  %y will be either 1 (yes there were three days in a row) or 0.
           %C is used as a counter to count the number of rainy days in a
           %row.
X=zeros(1,7);  %Storage for 7 days (X is a row vector)
for t=1:7
    xx=rand;
    if xx<p
        X(t)=1;  %Rainy day!  Notice by default, it is not raining (line 9)
    end
    if X(t)==1   %Why is there a double equal sign??
        C=C+1;
    else
        C=0;
    end
    if C>=3
        y=1;
    end
end
```

Here is a script file that will perform N simulations (for N weeks) and report the probability:

```
N=100;  %Number of weeks to simulate
p=0.5;  %Percent chance of rain each day.
Y=zeros(1,N);  %Keep track of results

for j=1:N
    [Y(j),X]=rainyweek(p);
end
Success=round(100*(sum(Y)/N));
fprintf('Probability of success: %d\n',Success);
```

First, run the script file to get a sense of what the probability is. Next, modify the script so that we run what is given above 50 times, keeping track of the success each time. Then have Matlab give you a histogram of the results.

3. Experimentally (with Matlab) verify three ways of double-centering a matrix of data- Use a random 6×8 matrix A using the command `rand`:

   - Subtract the row mean of A, then compute and subtract the column mean.
   - Subtract the column mean of A, then compute and subtract the row mean.
   - Compute the row mean and column mean and overall mean of the matrix A. Subtract row means, then column means, then add in the overall mean.

4. By hand, compute the mean and variance of the following set of data. Verify with Matlab.

$$1, 2, 9, 6, 3, 4, 3, 8, 4, 2$$

5. Obtain a sampling of 1000 points using the uniform distribution: and 1000 points using the normal distribution:

```
x=rand(1000,1);
y=randn(1000,1);
```

   Compare the distributions using Matlab's *hist* command: `hist([x y],100)` and print the results. You'll note that the histograms have not been scaled so that the areas sum to 1, but we do get an indication of the nature of the data.

6. Compute the covariance (in Matlab) between the following data sets:

x	-1.0	-0.7	-0.4	-0.1	0.2	0.5	0.8
y	-1.3	-0.7	-0.1	0.5	1.1	1.7	2.3

(1)

7. Let $\mathbf{x}$ be a vector of data with mean $\bar{x}$, and let a, b be scalars. What is the mean of $a\mathbf{x}$? What is the mean of $\mathbf{x} + b$? What is the mean of $a\mathbf{x} + b$? (NOTE: Formally, the addition of a vector and a scalar is not defined. Here, we are utilizing Matlab notation: The result of a vector plus a scalar is addition done componentwise. This is only done with scalars- for example, a matrix added to a vector is still not defined, while it is valid to add a matrix and a scalar).

8. Let $\mathbf{x}$ be a vector of data with variance s^2, and let a, b be scalars. What is the variance of $a\mathbf{x}$? What is the variance of $\mathbf{x} + b$? What is the variance of $a\mathbf{x} + b$?

9. Show that, for data in vectors $\mathbf{x}$, $\mathbf{y}$ and a real scalar a,

$$\text{Cov}(ax, y) = a\text{Cov}(x, y) \quad \text{Cov}(x, by) = b\text{Cov}(x, y)$$

10. Show that, for data in $\mathbf{x}$ and a vector consisting only of the scalar a,

$$\text{Cov}(x, a) = 0$$

11. Show that, for a and b fixed scalars, and data in vectors $\mathbf{x}, \mathbf{y}$,

$$\text{Cov}(x + a, y + b) = \text{Cov}(x, y)$$

12. If the data sets X and Y are the same, what is the covariance? What is the correlation coefficient? What if $Y = mX$? What if $Y = mX + b$?

13. Let X be a $p \times n$ matrix of data, where we n columns of p data points (you may assume each column has zero mean). Show that the $(i, j)^{\text{th}}$ entry of $\frac{1}{p}X^T X$ is the covariance between the i^{th} and j^{th} columns of X. HINT: It might be convenient to write X in terms of its columns,

$$X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$$

Also show that $\frac{1}{p}X^T X$ is a symmetric matrix.