Homework 1

1. (Basic Concepts) A random sample of house prices (in thousands) in a city is listed below.

   212.32, 327.56, 876.54, 745.23, 654.86, 189.07, 264.44

   a. Find the ranks and order statistics for this random sample.

      Solution: The ranks are 2 4 7 6 5 1 3 and the order statistics are {189.07 212.32 264.44 327.56 654.86 745.23 876.54}.

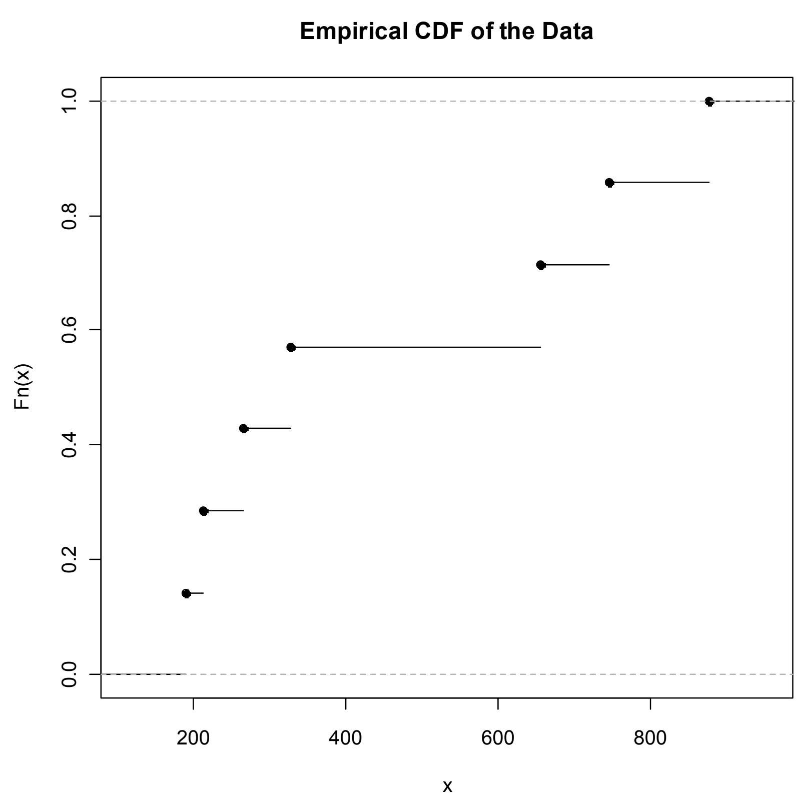
   b. Compute the mean and standard deviation

      Solution: The sample mean is 467.1457and the sample standard deviation is 283.7273.

   c. Find the median, first quartile, and third quartile.

      Solution: They are 238.380 327.560 700.045, respectively.

   d. Compute the ECDF of the data and plot.



Note: Try it out in R

```
x <- c(212.32, 327.56, 876.54, 745.23,
       654.86, 189.07, 264.44)
# Part (a)
rank(x); sort(x)
# Part (b)
mean(x); sd(x)
# Part (c)
median(x); quantile(x, probs = c(.25, .50, 0.75))
# Part (d)
ecdf(x)
plot.ecdf(x)
```

2. (*Discrete Uniform Distribution*) Suppose we have a random sample of size six observations, denoted as $\{X_1, \ldots, X_6\}$. Let R_1 be the random variable denoting the rank corresponding to the first observation X_1.
   a. What distribution does R_1 follow?

      Solution: We have P(R_1 = 1) = P(R_1 = 2) = 1/6 etc. In fact, P($R_1 = k$) = 1/6 for all k between 1 and 6. Hence we have a uniform distribution.
   b. Find the mean of R_1.

      Solution: Since $R_1 \sim$ Discrete Unif$\{1, 2, \ldots, 6\}$, we have

      $$E(R_1) = (1 + 2 + \cdots + 6)/6 = \frac{6 \times 7}{2} \cdot \frac{1}{6} = 3.5$$

3. (Binomial Distribution) The probability that a student is accepted to a prestigeous college is 0.3. If 5 students from the same school apply, what is the probability that at most 2 are accepted?

   Solution: To solve this problem, we compute 3 individual probabilities, using the binomial formula. The sum of all these probabilities is the answer we seek. Thus,

   $$b(x \le 2; 5, 0.3) = b(x = 0; 5, 0.3) + b(x = 1; 5, 0.3) + b(x = 2; 5, 0.3)$$
   $$= 0.1681 + 0.3601 + 0.3087 = 0.8369$$

4. (Normal Distribution) Suppose scores on an IQ test are normally distributed. If the test has a mean of 100 and a standard deviation of 10, what is the probability that a person who takes the test will score between 90 and 110?

Solution: Here, we want to know the probability that the test score falls between 90 and 110. The "trick" to solving this problem is to realize the following:

$$P(90 < X < 110) = P(X < 110) - P(X < 90)$$

We use the Normal Distribution Table or R to compute both probabilities on the right side of the above equation.

- To compute P(X < 110), we enter the following inputs into the calculator: The value of the normal random variable is 110, the mean is 100, and the standard deviation is 10. We find that P(X < 110) is 0.84.
- To compute P(X < 90), we enter the following inputs into the calculator: The value of the normal random variable is 90, the mean is 100, and the standard deviation is 10. We find that P(X < 90) is 0.16.

We use these findings to compute our final answer as follows:

$$P(90 < X < 110) = P(X < 110) - P(X < 90) = 0.84 - 0.16 = 0.68$$

Thus, about 68% of the test scores will fall between 90 and 110.

5. (*Hypergeometric Distribution*) Suppose we randomly select 5 cards without replacement from an ordinary deck of playing cards. What is the probability of getting exactly 2 red cards (i.e., hearts or diamonds)?

   Solution: This is a hypergeometric experiment in which we know the following:
   - N = 52; since there are 52 cards in a deck.
   - k = 26; since there are 26 red cards in a deck.
   - n = 5; since we randomly select 5 cards from the deck.
   - x = 2; since 2 of the cards we select are red.

   We plug these values into the hypergeometric formula as follows:

   $h(x; N, n, k) = [{}_{k}C_{x}] [{}_{N-k}C_{n-x}] / [{}_{N}C_{n}]$

   $h(2; 52, 5, 26) = [{}_{26}C_{2}] [{}_{26}C_{3}] / [{}_{52}C_{5}] = [325] [2600] / [2,598,960] = 0.32513$

   Thus, the probability of randomly selecting 2 red cards is 0.32513.