Homework Assignment #10: Probability & Statistics Review

This Homework Assignment is structured around both the review of concepts from the course and some concepts from Lectures 20 & 21. Remember, this Homework Assignment is **not collected or graded**! But you are advised to do it anyway because the problems for Homework Quiz #10 will be closely based on these problems![1]

1. **Review of Properties of Expectation and Variance:** A professor wishes to make up a true-false exam with 20 questions. She assumes that she can design the problems in such a way that a student will answer each problem correctly with probability p, and that the answers to the various problems may be considered independent experiments.

   Let X_j be equal to 0 if a student gets problem j wrong and 1 if they get it right. We know that X_j is a Bernoulli random variable with success probability p.

   Let S be the number of problems that a student will get correct. That is,

   $$S = \sum_{i=1}^{20} X_i.$$

   (a) What is the expected value of S?

   (b) The professor wants to choose p so that $E(S) = 0.7(20)$ What does this tell us about p?

   (c) What is the variance of S in this case?

   (d) Suppose the instructor decides to weight the values of questions differently. Now the first 10 questions are worth 1 point each and the last 10 questions are worth 2 points each. That is, the total number of points received by a student is:

   $$T = \sum_{i=1}^{10} X_i + \sum_{i=11}^{20} 2X_i.$$

   Determine the mean and variance of T.

2. **Review of Joint Discrete Random Variables/Variance/Covariance/Independence.** Let X and Y each take only the values -1 and 1. That is, their entire joint probability mass function can be written as:

   $$p_{1,1} = P(X = 1, Y = 1), p_{1,-1} = P(X = 1, Y = -1)$$
   $$p_{-1,1} = P(X = -1, Y = 1), p_{-1,-1} = P(X = -1, Y = -1)$$

   Suppose that $E[X] = E[Y] = 0.$

   (a) Show that $p_{1,1} = p_{-1,-1}$ and $p_{1,-1} = p_{-1,1}$.

   (b) Let $c = 2p_{1,1}$. In terms of c, find:
   - $Var(X)$
   - $Var(Y)$
   - $Cov(X, Y)$.

   (c) Under what values of c do X and Y have 0 Covariance. Are X and Y independent in this case?

3. **Kernel Density Function (Lecture 20):** Define the function K as follows:

   $$K(u) = (\pi/4)\cos(\pi u/2) \text{ for } -1 \le u \le 1 \text{ and } 0 \text{ otherwise.}$$

   Determine if $K(u)$ satisfies the following three properties of Kernel Density Functions:

[1] Course instructors reserve the right to *slightly* modify the questions from these when they make the Homework Quiz!

(1) $K(u)$ is a probability density $K(u) \geq 0$ and $\int_{-\infty}^{\infty} K(u)du = 1$.
(2) $K(u)$ is symmetric around zero, i.e., $K(u) = K(-u)$.
(3) $K(u) = 0$ for $|u| > 1$.

4. **Empirical CDF (Lectures 20 & 21):** Suppose you are given the following histogram of data. The table below specifies both the intervals of the bins and their respective bin heights. (Notice that this is a probability histogram because the sum of all the bin heights is 1.) From this histogram, determine the value of the empirical CDF when $x = 6$.

Bin	Height
(0,2]	0.1
(2,4]	0.2
(4,6]	0.4
(6,8]	0.1
(8,10]	0.2

5. **Sample Mean (Lecture 20):** Compute the sample mean for the dataset $1, 2, \ldots, N$.

6. **Sample Variance (Lecture 20)** Compute the sample variance for the dataset:

$$-N, \ldots, -1, 0, 1, \ldots, N.$$

You might find it helpful to remember:

$$1^2 + 2^2 + \cdots + N^2 = \frac{N(N+1)(2N+1)}{6}.$$

7. **Maximum Absolute Deviation (Lecture 20):** Let N be an even number. Compute the MAD of the dataset:

$$-N, \ldots, -1, 0, 1, \ldots, N.$$