[Homework] Variance, covariance, and correlation coefficient

1. For X and Y with $E[X] = 1$, $Var(X) = 5$, $E[Y] = 2$, $Var(Y) = 2$, and $\rho = 0.5$, find

   (a) $E[1 + 2X + 3Y]$

   (b) $Var(1 + 2X + 3Y)$.
   Sol.

2. We flips a fair coin many times and let X_i be the i-th flip record, where H and T are recorded as 1 and 0. Let Y_i be $Y_i = 2X_i - 1$, i.e., the i-th flip record where H and T are recorded as 1 and -1. Calculate the mean and variance of related random variables, i.e., fill up the blank in below table.

Random variable	Mean	Variance
Y_i	0	1
$\sum_{i=1}^{n} Y_i$	0	n
$\frac{\sum_{i=1}^{n} Y_i}{\sqrt{n}}$	0	$\sqrt{n}$

3. Let X_i be independent with same mean 0, variance 2, and let $Y_n = X_n + X_{n+1} + X_{n+2}$. For $j \geq 0$, calculate $Cov(Y_n, Y_{n+j})$.

4. Let X be the number of heads minus the number of tails obtained when a fair coin is tossed n times. Calculate the mean and variance of X.
   Sol. Let i be the number of tails. Then $n - i$ is the number of heads and $X = n - 2i$.

$$P(X = i) = \binom{n}{i}(\frac{1}{2})^i(\frac{1}{2})^{n-i} = \binom{n}{i}(\frac{1}{2})^n$$

$$\begin{aligned}
E(X) &= \sum_x x \cdot P(X = x) \\
&= \sum_{i=0}^{n}(n - 2i)\binom{n}{i}(\frac{1}{2})^n \\
&= \{n\sum_{i=0}^{n}\binom{n}{i} - 2\sum_{i=0}^{n} i \cdot \binom{n}{i}\} \cdot (\frac{1}{2})^n \\
&= \{n \cdot 2^n - 2 \cdot n \cdot 2^{n-1}\} \cdot (\frac{1}{2})^n \\
&= 0 \cdot (\frac{1}{2})^n = 0
\end{aligned}$$

$$\begin{aligned}
Var(X) &= \sum_x x^2 \cdot P(X=x) - (EX)^2 \\
&= \sum_{i=0}^{n} (n-2i)^2 \binom{n}{i} (\frac{1}{2})^n - 0^2 \\
&= \{n^2 \sum_{i=0}^{n} \binom{n}{i} - 4n \sum_{i=0}^{n} i \cdot \binom{n}{i} + 4 \sum_{i=0}^{n} i^2 \cdot \binom{n}{i}\} \cdot (\frac{1}{2})^n \\
&= \{n^2 \cdot 2^n - 4n \cdot n \cdot 2^{n-1} + 4 \cdot (n^2+n) 2^{n-2}\} \cdot (\frac{1}{2})^n \\
&= n
\end{aligned}$$

5. We flips a fair coin until heads appears 10 times. Let X be the number of tails during these flips. Calculate the mean and variance of X .

   Sol.

$$Y \sim NB(r=10, p=\frac{1}{2})$$

$$X \sim HG$$

$$\begin{aligned}
P(X=n) &= P(Y=n+10) \\
&= \binom{n+10-1}{r-1} p^r (1-p)^{(n+10)-r} \\
&= \binom{n+9}{9} (\frac{1}{2})^{(n+10)}
\end{aligned}$$

6. If X_i, $1 \leq i \leq 4$, have mean 3 and variance 2 and if the correlation ρ between two are all 0.5, compute

   (a) $Cov(X_1 + X_2, X_2 + X_3)$.

   (b) $Cov(X_1 + X_2, X_3 + X_4)$.

7. Consider a graph having n vertices labeled $1, 2, \ldots, n$, and suppose that, between each of the $\binom{n}{2}$ pairs of distinct vertices, an edge is independently present with probability p. The degree of vertex i, designated as D_i, is the number of edges that have vertex i as one of their vertices.

   (a) What is the distribution of D_i?

   (b) Find $\rho(D_i, D_j)$, the correlation between D_i and D_j.

8. A group of eighteen students contains seven business majors; the rest are liberal arts majors. Five students will be drawn from the group and let X be the number of business majors in the sample.

   (a) Calculate $P(X = 3)$.

   (b) What is the distribution of X?

   (c) What is EX?

   (d) What is $Var(X)$?

9. A pond contains 100 fish, of which 30 are carp. If 20 fish are caught, what are the mean and variance of the number of carp among the 20?

10. A box contains 5 red and 5 blue marbles. Two marbles are withdrawn randomly. If they are the same color, then you win \$1.10; if they are different colors, then you win $-$\$1.00. (That is, you lose \$1.00.) Calculate

    (a) the expected value of the amount you win.

    (b) the variance of the amount you win.

[Extra] Variance, covariance, and correlation coefficient

1. A sample of 3 items is selected at random from a box containing 20 items of which 4 are defective. Find the expectation and variance of the number of defective items in the sample.