Question No 1:

(To be solved using R) Let X1, X2, X3, X4, X5 be independent U (0, 1) random variables. Let X = X1 + X2 + X3 and Y = X3 + X4 + X5. Use the runif() function to simulate 1000 trials of each of these variables. Use these to estimate Cov (X, Y).

Solution:

Steps:

- Generated five different vectors with 1000 numbers in each with U(0,1). Used set.seed() to ensure that they are all different.
- Calculated mean of X1 and X5 (picked at random) which came out to be 0.4997 and 0.5015 which is a good approximation of 0.5. (this step is superfluous and has been done just to see how random a randomly generated dataset is)
- Created X and Y using simple vector addition
- Considering that both X and Y are additions of randomly generated variables, they should not have any significant linear relationship between them. Checking that by using cov() function.
- The covariance comes out to be 0.09 which is not a very approximation of 0. Further, the correlation coefficient (which is a more standard measure since it scales the relation to -1 to 1) is 0.349 which cannot just be ignored.
- The reason behind this non-insignificant linear relation between X and Y can be the common element of X3 between both.
- For trying out, another random variable X6 is created.
- The covariance between X1+X2+X3 and X4+X5+X6 comes out to be 0.004 which is significantly less.
- The covariance between X1+**X2+X3** and **X2+X3**+X4 comes out to be 0.176 which is significantly more than before. Checking the correlations also confirms the same.
- To summarize, the linearity between seemingly random variables increased with increasing the strength of a common additive in the variables.

R Code:

```
> # Creating 5 vectors (X1 to X5) having 1000 uniform random numbers on [0
,1]
> set.seed(1)
> X1 = runif(1000)
> set.seed(2)
> X2 = runif(1000)
> set.seed(3)
> X3 = runif(1000)
> set.seed(4)
> X4 = runif(1000)
> set.seed(5)
> X5 = runif(1000)
>
> # calculate mean of the variables just to check
> mean(X1)
[1] 0.4996917
> mean(X5)
[1] 0.5014789
>
>
> # creating X and Y as instructed in the problem statement
```

```
> X = X1 + X2 + X3
> Y = X3 + X4 + X5
>
> # calculating covariance
> cov(X,Y)
[1] 0.09055103
> # calculating the correlation
> cor(X,Y)
[1] 0.3498653
> # creating another variable X6
> set.seed(6)
> X6 = runif(1000)
> # calculating covariance with no common element
> cov(X1+X2+X3,X4+X5+X6)
[1] 0.004226472
> # calculating covariance with 2 common elements
> cov(X1+X2+X3,X2+X3+X4)
[1] 0.1765867
> #Checking the correlation coefficients
> cor(X1+X2+X3,X2+X3+X4)
[1] 0.6797308
> cor(X1+X2+X3,X4+X5+X6)
[1] 0.01656714
```

Question No 2

The random variable X takes values -1, 0, 1 with probabilities 1/8, 2/8, 5/8 respectively.
(a) Compute E(X).
(b) Give the pmf of Y = X^2 and use it to compute E(Y). [PMF - Probability Mass Function]
(c) Instead, compute E(X^2) directly from an extended table.
(d) Compute Var(X).

Solution (done with pen & paper i.e. MS Word 😊) :

(a) The expected value of a discrete random variable is the expected value of a discrete random variable is the probability-weighted average of all possible values. So,
E(X) = -1 *(1/8) + 0*(2/8) + 1 * (5/8) = 4/8 = ½

(b) Here Y is a function of X as Y = f(X) = X^2
E[f(X)] = ∑ f(x)P(X = x)

So, E(Y) = $(-1)^2$ *(1/8) + $(0)^2$*(2/8) + $(1)^2$ * (5/8) = 6/8 = ¾

(c) Extended table:

X	P(X)	XP(X)	X^2	X^2P(X)
-1	1/8	-1/8	1	1/8
0	2/8	0	0	0
1	5/8	5/8	1	5/8
Sum	**1**	**1/2**		**3/4**

From the above, E(X^2) = ¾

(d) $Var(X) = E(X^2) - [E(X)]^2$

$= \frac{3}{4} - (1/2)^2$

$= \frac{3}{4} - \frac{1}{4}$

$= 2/4 = \mathbf{1/2}$