NPTEL DATA SCIENCE FOR ENGINEERS

ASSIGNMENT-3-SOLUTIONS

1. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Since A and B are mutually exclusive,

$P(A \cap B) = 0$

Therefore, $P(A \cup B) = P(A) + P(B)$

Also, A and B are mutually exclusive.

Therefore $A \cup B = S$, or the sample space

Which means that $P(A \cup B) = P(A) + P(B) = 1$

2. Statement I is TRUE and statement II is FALSE. Please refer to Random Phenomena in Statistical modelling video.

3. $E(10X + 5Y) = E(10X) + E(5Y) = 10E(X) + 5E(Y) = 80 + 20 = 100$

4. $Var(10*(X+Y)) = 10^2 (Var(X+Y)) = 100(Var(X) + Var(Y)) = 100(6+4) = 1000$

5. The mean and variance of a binomial distribution are np and np(1-p) respectively.

Now, $Y = aX + b$

Therefore, $E(Y) = E(aX+b) = a(E(X)) + b = anp + b$

$Var(Y) = Var(aX + b) = a^2 Var(X) = a^2 np (1-p)$

6. The sample variance s is an unbiased estimator for the population variance. This can be calculated in R using the command var(X).

```
> X=c(58,59,63,60,60,63,60,57,58,59)
> var(X)
[1] 4.011111
```

7. For any distribution, the sample mean **is an unbiased estimation** of the population mean.

NPTEL DATA SCIENCE FOR ENGINEERS

ASSIGNMENT-3-SOLUTIONS

8. Test statistic, $Z = (\bar{x} - \mu)/(\sigma/\sqrt{n}) = (1850 - 1800)/(100/\sqrt{50}) = 3.54$

The z-value at 1% significance = 2.58, from the standard normal probability table.

Therefore, the z-value from the test statistic is greater than the z-value at 1% significance.

Hence, the null hypothesis is rejected.

9. t.test(x) is the function that is used conduct a t-test

10. First **Quartile**(Q1)=$((n+1)/4)^{th}$ Term also known as the **lower quartile**.This can be calculated in R using the command var(X).

```
Z =c (57,58,67,60,60,68,62,53,56,57,66,61,62,60,57,68,62,50,63,57)
summary(Z)
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
50.00   57.00   60.00   60.20   62.25   68.00
```