

Basics of Probability and Statistics Week 1 Challenge

1. Why there is a difference in the formula of variance for population and sample
2. Difference between stratified and clustered sampling
3. How many different samples can be created out of population of size n
4. What is the probability of drawing 2 jacks from a pack of card
5. What is the probability that both the numbers on dice are same while rolling 2 dice
6. Why standard deviation when variance can measure the dispersion in data

Answers:

1. Why there is a difference in the formula of variance for population and sample?

We divide by N (the population size), When we calculate sample variance, we divide by $n-1$ (the sample size $- 1$).

2. Difference between stratified and clustered sampling

Cluster Sampling is a type of sampling method in which we split a population into clusters, then randomly select some of the clusters and include all members from those clusters in the sample. Stratified sampling is a type of sampling method in which we split a population into groups, then randomly select some members from each group to be in the sample.

3. How many different samples can be created out of population of size n

Lets assume Population Size = N
and sample size = n
so the total number of samples created is = $N C n$

4 . What is the probability of drawing 2 jacks from a pack of card

Total number of cards in a pack = 52
Total number of Jack in a pack = 4
The probability of drawing a Jack = $4/52$
The probability that both are jack = $(4/52) \cdot (3/51)$
 $= (1/13) \cdot (1/17)$
 $= 1/221$.

5. What is the probability that both the numbers on dice are same while rolling 2 dice

[(1, 1) (1, 2) (1, 3) (1, 4) (1, 5) (1, 6)
(2, 1) (2, 2) (2, 3) (2, 4) (2, 5) (2, 6)

(3, 1) (3, 2) (3, 3) (3, 4) (3, 5) (3, 6)
(4, 1) (4, 2) (4, 3) (4, 4) (4, 5) (4, 6)
(5, 1) (5, 2) (5, 3) (5, 4) (5, 5) (5, 6)
(6, 1) (6, 2) (6, 3) (6, 4) (6, 5) (6, 6)]

So, pairs with same numbers are (1,1) (2,2) (3,3) (4,4) (5,5) (6,6) i.e. total 6 pairs

Total outcomes = 36

Favorable outcomes = 6

Probability of getting pair with same numbers = Favorable outcomes / Total outcomes = $6 / 36 = 1/6$

So, $P(N, N) = 1/6$.

6. Why standard deviation when variance can measure the dispersion in data?

Square root of variance is equal to Standard Deviation.

$\sigma = \text{Std.Dev}$

$\sigma^2 = \text{Variance}$

Variance is a method to find or obtain the measure between the variables that how are they different from one another, whereas standard deviation shows us how the data set or the variables differ from the mean or the average value from the data set.