Sampling Distribution

Siddharth Kalla45.9K reads

Sampling distributions are an important part of study for a variety of reasons. In most cases, the feasibility of an experiment dictates the sample size. Sampling distribution is the probability distribution of a sample of a population instead of the entire population.

In simpler words, suppose from a given population you take all possible samples of size n and compute a statistic (say mean) of all these samples. If you then prepare a probability distribution of this statistic, you will get a sampling distribution.

The properties of sampling distribution can vary depending on how small the sample is as compared to the population. The population is assumed to be normally distributed as is generally the case. If the sample size is large enough, the sampling distribution will also be nearly normal.

If this is the case, then the sampling distribution can be totally determined by two values - the mean and the standard deviation. These two parameters are important to compute for the sampling distribution if we are given the normal distribution of the entire population.

Sampling Distribution of the Mean and Standard Deviation

Sampling distribution [1] of the mean is obtained by taking the statistic under study of the sample to be the mean. The say to compute this is to take all possible samples of sizes n from

the population of size N and then plot the probability distribution. It can be shown that the mean of the sampling distribution is in fact the mean of the population.

The standard deviation however is different for the sampling distribution as compared to the population. If the population is large enough, this is given by:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Where ? is the mean of the population and ?x? is the population mean.

Other Distributions

These formulas are only valid when the population [2] is normally distributed [3]. If this is not the case, then the mean and standard deviation of the sampling distribution will be different and will depend on the type of distribution of the population.

The normal distribution is one of the simplest probability distributions and so it is quite easy to study and analyze. We can easily find mathematical formulas for the sampling distribution statistics that we want to find.

However, when the distribution is not normal, it can be quite complicated and such easy mathematical formulations might be hard to find or even impossible in some cases. In those cases, we use approximate methods because finding the exact value will entail studying every single sample of size n taken from the population, which is very hard and time consuming.

Links
[1] http://stattrek.com/sampling/sampling-distribution.aspx
[2] https://explorable.com/research-population
[3] https://explorable.com/normal-probability-distribution