On the Methodological Issue of Sample Size*

J.C. Sharma
Executive Editor
BVIMSR's Journal of Management Research

In our editorials we have been discussing issues relating to research in management. In this issue the problem of sample size is being discussed. At times we receive papers which are based on a small sample of 50 or 60 respondents. The results of such small samples can't be relied on.

Sampling is the foundation of all scientific Investigations. Reliable sample helps researcher make decisions with confidence. A small sample has large sampling error and hence gives less reliable results. Larger the sample size, more reliable is the result. Though it is not always true, beyond a point increase in sample size does not lead to a similar increase in the reliability of results.

A common notion that 'larger the population more the sample size required' is not correct. In fact sample size has very little relation with the population size, especially for large populations. Another notion that the sample size is decided as some fixed percentage of population is also unfounded.

The sample size 'n' primarily depends on three factors viz., (i) the variance of the variable being investigated, (ii) amount of error acceptable in the result, and (iii) level of confidence required. The population size 'N' may also play a role in the determination of sample size if the population size is not large enough.

Statistical formulae are available in the text books for determination of sample size. The formula depends on the objective and the hypothesis. For example to determine how much sample is required to estimate the 'percentage of population using organic food' will be estimated by a formula based on Binomial distribution test for testing the hypothesis that P=some percentage. A small percentage will need a large sample.

Similarly to estimate 'how much on an average does a family spend on medical care' will be estimated by a formula based on t-distribution or Z-distribution for testing the hypothesis that mean=some value. Higher the SD (large variance in the variable) more will be the sample size required.

Another example could be for estimating the 'correlation between Motivation and Productivity' the sample size will be based on a formula for testing of hypothesis that correlation r=some value. Smaller the correlation larger will be the sample size required.

Given below is a formula for estimating sample size to estimate percentage of population (P) using organic foods.

$n_0 = Z^2 * (P)*(100-P) / d^2$

Where n_0 = sample size required (considering population size as infinite)
Z = value of Normal deviate for given level of confidence
P = approximate percentage of population using organic food

d = amount of error permissible in estimate of P

The value of 'Z' for 95% confidence is 1.96.
The approximate value of 'P' will be obtained from previous studies or through pilot study. That is why Literature review and Pilot Study are important in any research.
The amount of error'd' is some percentage of 'P', e.g. if P=20% then d=(10% of P)=2.

If the population is not large enough then additionally a second formula may be used.

Final sample size $n = n_0 / (1 + n_0 / N)$

Where N = size of population.
The second formula may be applied only if the ratio $(n_0 / N) > 0.05$.

The researchers should look at the theory of determining sample size and give proper justification for the sample size chosen for the study in the article.

J.C. Sharma
Executive Editor

* Originally published in BVIMSR's Journal of Management Research (Vol.-5 No.-1)