

MARKETING
RESEARCH
ESSENTIALS

McDANIEL
GATES

7TH
EDITION

Chapter Twelve

Sample Size Determination

Chapter Twelve Objectives

- To learn the financial and statistical issues in the determination of the sample size.
- To discover methods for determining the sample size.
- To gain an appreciation of a normal distribution of data.
- To understand population, sample, and sampling distributions.
- To distinguish between point and interval estimates.
- To recognize problems involving sampling means and proportions.

Sample Size for Probability Samples

Budget Available:

- *What can we afford?*
- *How much do we want to spend?*
- *How much time are we allotting for each respondent*



Arbitrary / Rule of Thumb:

- *Applies some industry accepted “rule of thumb”*
- *Generally better for smaller populations*
- *Picking “x” percent of the population to be in the sample*

Sample Size for Probability Sampling

Number of Subgroups Analyzed:

- *In any sample size determination problem, consideration must be given to the number and anticipated size of various subgroups of the total sample that must be analyzed*

Traditional Statistical Methods:

- *Variance, SD, confidence interval play a key role*

Sample Size for Nonprobability Sampling

Judgment:

- *Best guess of “experts”*
- *Draw on your experience to determine sample size*

Conventional:

- *What have others done?*
- *See what the sample size has been for similar studies*

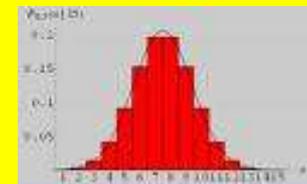
The Normal Distribution

Central Limit Theorem:

- The idea that a distribution of a large number of sample means or sample proportions will approximate a normal distribution - regardless of the distribution of the population from which they were drawn.

Normal Distribution:

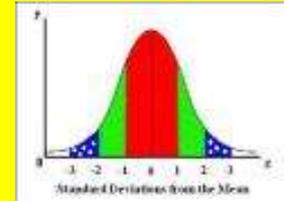
- The continuous distribution that is bell shaped and symmetrical about the mean. The mean, median, and mode are equal. About 68% of the observations are within +/- one standard deviation, 96% are within two standard deviations, and 99+% are within three standard deviations of the mean.



The Normal Distribution

Proportionate Properties:

- A feature that the number of observations falling between the mean and a given number of standard deviations from the mean is the same for all normal distributions.



Standard Normal Distribution:

- Normal distribution with a mean of zero and a standard deviation of one.

The Normal Distribution

Standard Deviation:

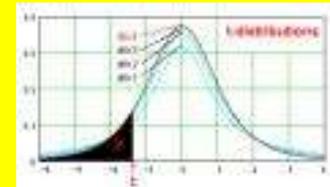
- The measure of dispersion calculated by subtracting the mean of the series from each value in a series, squaring each result, summing the results, dividing the sum by the number of observations minus 1, and taking the square root of this value.

$$\text{Standard Deviation} = \sqrt{\frac{\text{sum } (X_i - \bar{X})^2}{(N-1)}}$$

Population and Sampling Distributions

Population Distributions:

- The frequency distribution of all the elements of a population.



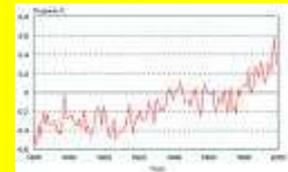
Sampling Distributions:

- The frequency distribution of all the elements of an individual sample.

Sampling Distribution of the Mean

Sampling Distribution of the Mean:

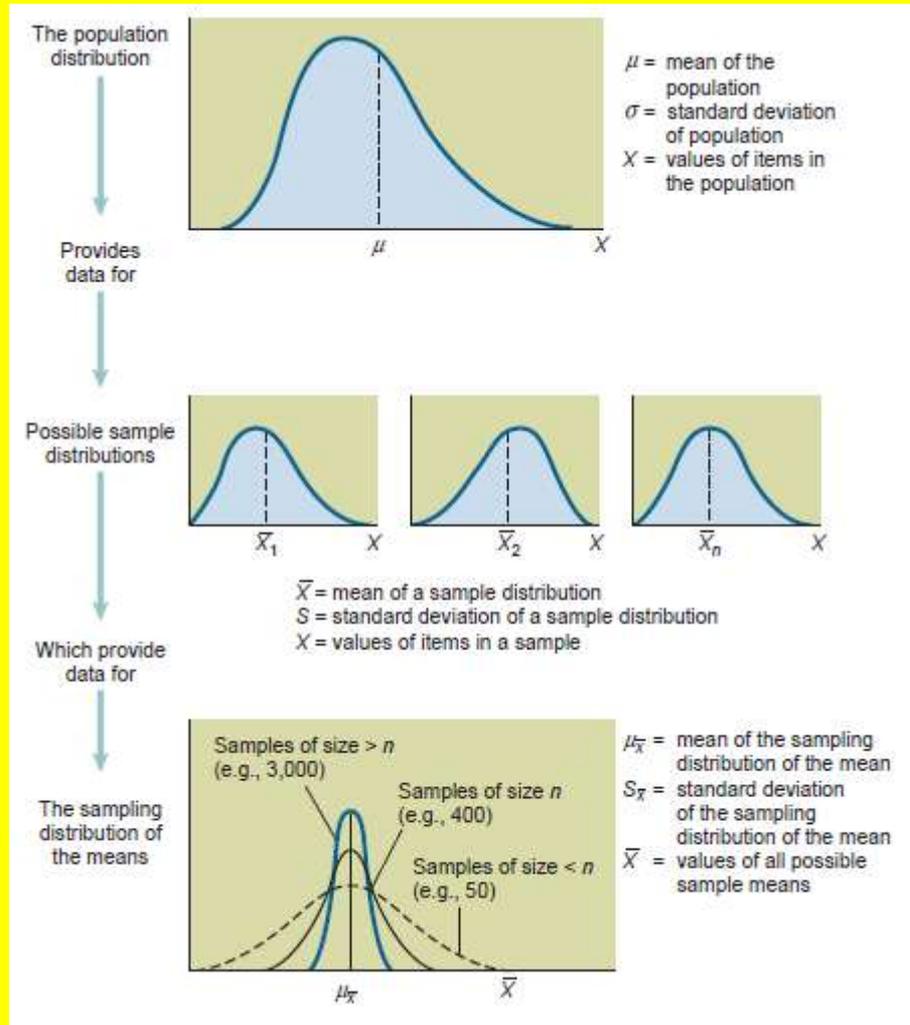
- The theoretical frequency distribution of the means of all possible samples of a given size drawn from a particular population; it is normally distributed.



Standard Error of the Mean:

- Standard deviation of a distribution of sample means.

Distributions Summary



Point and Interval Estimates

Point Estimate:

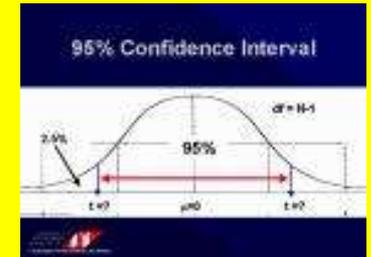
- The particular estimate of a population value.

Interval Estimate:

- The interval or range of values within which the true population value is estimated to fall.

Point and Interval Estimates

Confidence Level:



- The probability that a particular interval will include the true population value - also called the *confidence coefficient*.

Confidence Interval:

- The interval that, at the specified confidence level, includes the true population value.

Sampling Distribution of the Proportion



Sampling Distribution of the Proportion:

- The relative frequency distribution of the sample proportions of many random samples of a given size drawn from a particular population; It is normally distributed.

Determining Sample Size

**Problems
Involving
Means:**



$$N = \frac{Z^2 * \sigma^2}{e^2}$$

N = Sample Size

Z = level of confidence desired in the results. A 95% confidence interval would make $Z=1.96$. In other words, if we conduct this survey 100 times, at least 95 of those time the true population average would fall within out interval estimate.

σ = Population standard deviation.

e = error rate - a management decision (*ex. plus or minus 3%*)

Determining the Sampling Size

Allowable Sampling Error:

- Amount of sampling error the researcher is willing to accept.

Population Standard Deviation:

- Standard deviation of a variable for the entire population.

Determining Sample Size

Problems Involving Proportions:



$$N = \frac{z^2 \{P(1-P)\}}{e^2}$$

N = Sample Size

Z = level of confidence desired in the results. A 95% confidence interval would make $Z=1.96$. In other words, we would be 95% confident that the average results in the whole population (*were we to survey the whole population*) would be within 1.96 standard deviations from the mean.

p = variance (how different you predict the population is), $q = 100-p$

e = error rate - a management decision (*ex. plus or minus 3%*)

Population Size and Sampling Size

Independence Assumption:

Assumption that sample elements are drawn independently.

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

For a sample that is 5 percent or more of the population, the independence assumption is dropped, producing the following formula:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Determining Sample Size

The Finite Population Correction Factor (FPC):

- An adjustment to the required sample size that is made in cases where the sample is expected to be equal or more of the total population.

$$n' = (nN) / (N+n-1)$$

Where n' =revised sample size; n =original sample size; N =population size

Statistical Power: The probability of not making a type II error.

Determining Sample Size

Over Sampling:

You need to over sample since it's unlikely that everyone you contact will agree to answer the questionnaire. If you determine that you need 800 respondents in the survey (RSS), and a typical response rate for your type of survey is 30%, then use the calculations to below to get the number of people you would actually have to contact to get 800 completed responses.

$$O = \text{RSS}/.30$$

or

$$O = 800/.3$$

$$O = 2,667$$

Thus your new sample size is 2,667 - with 30% response you can expect 800 people to respond.



Key Sampling Considerations

Consider:



Time to Generate Sample
Scope of the Research
Budget Available
Experience with Sampling
Level of Accuracy Desired
Your Knowledge of the Population