Power Analysis: The Issue of Sample Size

Dr. Mary McHugh
Dean of Nursing

How Large a Sample does the Researcher Need?

The underlying question is: What characteristics of the sample allow it to truly reflect the behavior of the population under the experimental treatment conditions?

Critical Sample Characteristics for Generalizability

1. Similarity to the population to which findings are to be generalized
2. Homogeneity of the population
3. Sample size

Other Factors that Affect Required Sample Size

1. How strong is the effect of the experimental treatment? (Effect Size)
   - If the treatment exhibits a very strong effect, the sample size need not be large

Other Factors that Affect Required Sample Size

1. experimental treatment? (Effect Size)
   - If the treatment exhibits a very strong effect, the sample size need not be large.

The Purpose of Sampling is to Represent the Population

The point of sampling is to find a difference between the treatment and control group if such a difference truly exists “in nature”.

The Purpose of Sampling is to Represent the Population

But obtaining a sample and implementing experimental treatments are expensive research activities. Thus, sample sizes larger than needed simply waste time and resources.

To Obtain a Significant Result, Adequate Power is Essential

Power represents the likelihood that a researcher will find a significant difference in the sample if one truly exists in the population. Finding the smallest sample that will find a real difference is the goal of power analysis.

Key Components of Power

Power is a function of a variety of factors, including:

1. Homogeneity on key study characteristics of the group being studied: Population σ (SD)
2. The effect size (correlation)
3. Sample size
4. Power of the statistic (parametric vs non-parametric)

Key Components of Power

5. Chosen significance level p<.10 produces much more power than p.<01 but a much greater probability of a Type I error
6. One versus two-tailed testing. A one-tailed test has more power than a two-tailed test

Key Components of Power

7. Normal versus non-normal distribution may affect power. All inferential statistics work better with normal distribution, even the distribution free (non-parametric) statistics.

Is the Population Highly Variable?

If the population is highly variable, or heterogenous, then it will be more difficult to represent that population with a sample.

Is the Population Highly Variable?

Much larger samples are needed to represent heterogenous populations.
For example, the U.S. is politically heterogenous.
Sample sizes of 1500 to 2000 are needed for accurate political polls, and even then the polls sometimes get it wrong.

Effect Size

Strength of the effect of the experimental intervention plays a huge role in power. For very strong effects; correlations of .80 and above, a sample size of 10 may be sufficient to produce statistical significance.

Effect Size

But, in a heterogeneous population, a strong effect in a small sample may not be reflected in the full population.

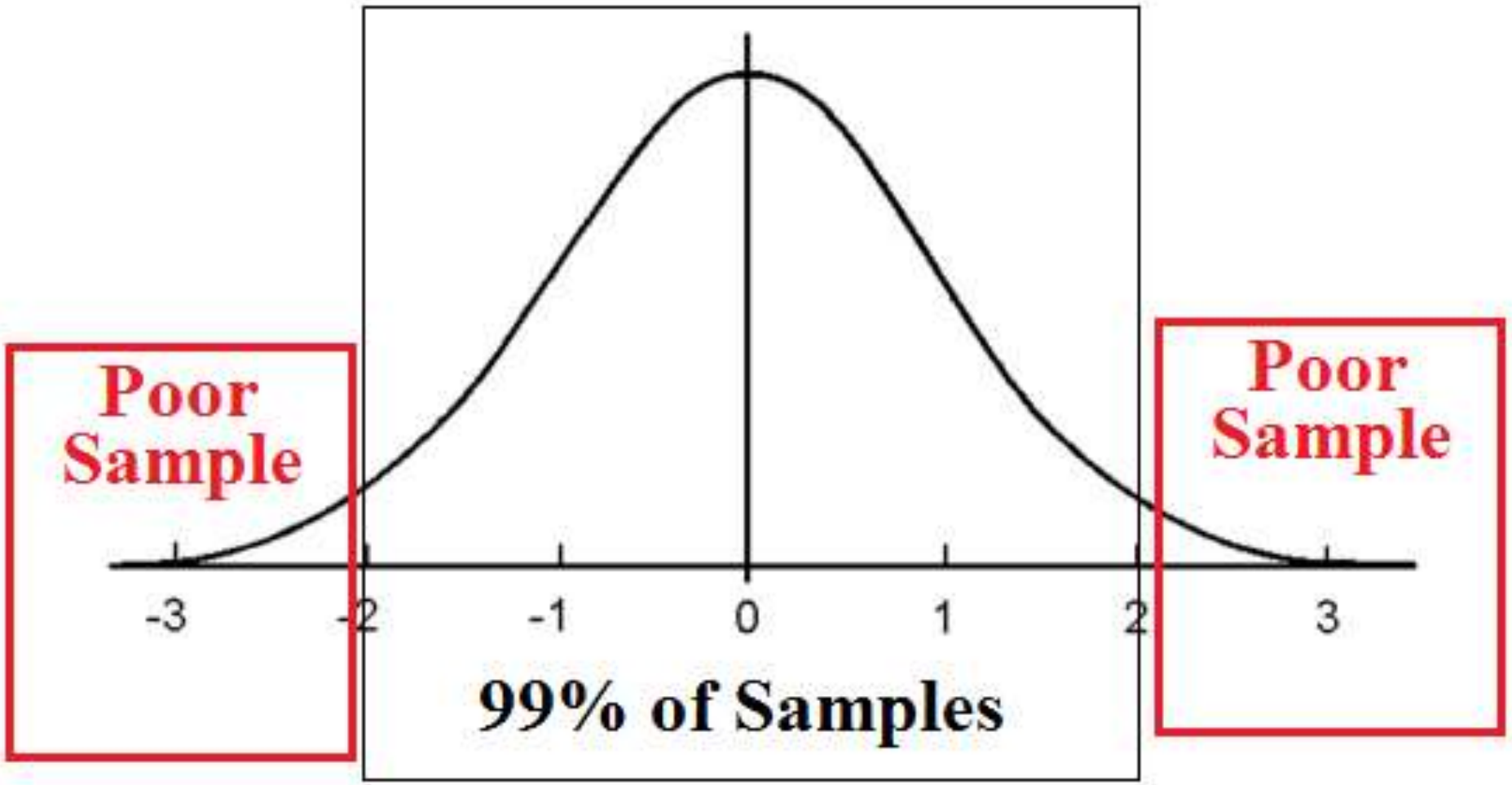
Small samples cannot represent heterogenous populations effectively because sizeable subpopulations are very likely to not show up in a sample.

Sampling

Sampling must be truly random from the population if inferential statistics are to produce reliable results, and all too often that kind of sampling is not possible.

Sampling

Random sampling generally does a good job of representing populations—assuming the sample size is adequate to represent the population variablity—but samples also follow a normal distribution with respect to their representativeness.


Poor Sample
Poor Sample
-3
-2
-1
0
1
2
3
99% of Samples

Conclusion

Selecting an appropriate sample size is not as cut and dried as some books would have it. Although calculating a power analysis is important, and is now a standard of practice in research, those statistics make certain assumptions.

Conclusion

Statists “assume” a truly random selection of subjects, they assume that the variable of interest is normally distributed in the population, and finally, they assume typical variability (i.e. normal curve) of the variable of interest in the population.

Conclusion

Often, population characteristics of interest do not meet those assumptions. In those instances, a larger or smaller sample size might be needed to find differences due to the experimental treatment than the power analysis might indicate.

Conclusion

However, the power analysis is still the researcher's best tool to determine an adequate sample size in order to have sufficient power to find differences between the experimental and control groups.

QUESTIONS???