

MECOR

Sample Size and Power: An Intuitive Approach

Stephen McCurdy, M.D., M.P.H.
Division of Environmental & Occupational
Health
Department of Public Health Sciences
U.C. Davis School of Medicine



Sample Size and Power

- A critical aspect of study planning involves estimating how many subjects are needed to answer your study question--validly and reliably.
- These estimations necessarily involve oversimplification and assumptions. Hence, the estimates should be viewed as a guide. One is always better off with yet more subjects!

Sample Size and Power

- Studies typically involve a comparison between two (or more) groups. Differences are observed between the groups, and we ask whether these differences are
 - *Small enough to be explained by chance ($p \geq 0.05$)
 - or
 - *So large they are unlikely to be explained by chance, and thus likely represent a true difference between the groups ($p < 0.05$)

Sample Size and Power

- The study power is the likelihood that your study and its sample size will show a true underlying difference between groups and yield a p value of less than 0.05.
- Think intuitively: a study with high power is likely to show you the true difference in height between men and women and yield $p < 0.05$.
- A study with low power may well show you the same true difference in height, but with a p value of > 0.05 . (This would lead to erroneous acceptance of the null hypothesis of no difference in height between men and women--a Type II error.)

Sample Size and Power

- Think intuitively: There are four factors that affect the power of the study. (Only the first three are of practical importance.)
 1. Group difference: Intuition tells us that it will be easier to detect a large difference than a small one.
 2. Sample size: Intuition tells us that a larger study is more likely to detect a true underlying difference between groups.
 3. Dispersion: Intuition tells us that if there is a tight dispersion of data such that there is little overlap between two groups, it will be easier to see that the groups are different.
 4. P-value required for rejecting the null hypothesis: Intuition tells us that if we demand a difference so large that it could only happen less than 5% of the time in order to reject the null hypothesis, we are less likely to reject the null hypothesis than if we only demand a (smaller) difference that could happen by chance 10% of the time.

Sample Size and Power

- Studies are not worth doing unless they have adequate power to answer the question. As a practical matter, we aim for a power of 80%.
- A study with 80% power has an 80% chance of detecting a true difference and yielding a p value <0.05 .
- Therefore, the investigator wrestles with the question of how many subjects must be in the study to achieve power of at least 80%.

Sample Size and Power

- Example: An investigator wants to determine whether CD4 counts in HIV patients presenting with opportunistic infections (OI) are different from CD4 counts in HIV patients presenting with neoplasms.
- 24 patients with HIV and OI: Mean CD4: 75 SD:45
- 18 patients with HIV and neoplasm: Mean CD4: 100 SD:47
- $P=0.09$

Sample Size and Power

- Frustrated, the investigator wants to know the power of the study and how many subjects are needed to achieve a power of 80%.
- Power of the present study: 41%
- Number of subjects needed in each group to achieve power of 80%: 54 in each group (i.e., total of 108 patients)

Sample Size and Power

- Example: An investigator can manage to find 60 subjects from families who cook outdoors and 45 subjects from families who cook with indoor fires. The investigator would like to see if there is increased prevalence for asthma among those cooking indoors.
- What further information do you need?
- Is there enough power for this study?
- If not, how many subjects are needed for adequate power?

Sample Size and Power

- Take-home messages:
- Differences between groups are more likely to be detected as statistically significant when:
 1. The true underlying difference between groups is large.
 2. Each group has a “tight” distribution (small SD) with little overlap into the other group.
 3. Sample size is large.
 4. Estimated sample size is a guide--more subjects are always helpful.