Sample Size and Use of z-tables

Estimate an appropriate sample size

Only one of the most basic calculations for sample size is shown this week. When you have multiple tests in your study, only compute the sample size for the test that requires the largest sample. It is assumed sample size and power for the most demanding test is sufficient. Other software packages give the needed sample size (such as Epi Info, Open Epi). Inputs for the sample size formulas include the desired power, the level of significance and the effect size.

For example: Run t-tests and Chi squares prior to a multiple linear regression, Then compute the sample size for the multiple linear regression and assume there is sufficient power for the t-test and Chi Squares.

Will increasing the sample size, increase the power?

The relationship between power and sample size is linear only up to about 90% of power. After that increasing sample size may not necessarily lead to increased power. On the other hand, the relationship between power and effect size is more subtle.

Statistical power- the probability of finding a particular sized effect.

Effect size is defined in terms of population means and a population standard deviation. That is $d = \text{absolute value of (mean1-mean2)/standard deviation}$ where $d$ is Cohen’s d. Power depends on effect size in such a manner that the smaller an effect the more observations you need to establish its existence. Generally power is directly related to effect size, sample size and significance level. An increase in the effect size, the sample size or the significance level will produce increased statistical power, all other factors being equal. Power is inversely related to variability. Decreasing variability will increase the power of a study. If the power of a study is relatively high and a statistically significant effect is not observed, this implies the effect, if any, is small.

©2014 Walden University Academic Skills Center
**Power** is the additive inverse of type II error (known as Beta (B), which is why power is calculated as 1-B.

**Sample size**—estimated by reviewing literature or using some softwares. Use the following link to learn more.

http://www.indiana.edu/~statmath/stat/all/power/power.html

**Important factors:**

- Power
- Significance level
- Standard deviation
- Time available for sampling
- Resources (human, financial, and other), as well as other factors.

**Other factors to consider:**

- Dropouts and missing responses
- Return rate on questionnaires
- Use of clusters in analysis

**Resource for z-value**

Follow this link to learn how to find the z value given a probability or the vice versa from the standard normal curve.

http://www.youtube.com/watch?v=rEmNUkKSpbU

©2014 Walden University Academic Skills Center
How $Z_{0.8}= 0.84$ value is derived from the Z tables

The percent you wish the power to be (80%) based on 1-.20, you would look in the body of the table on page 273 of your text for the value closest to 80. That value happens to be 0.7995.

Then look to see the row and column that the number is in. It is in row 0.8 and column 0.04.

When you combine the column and row you get 0.84. So the z value associated with 80% power is 0.84. Some of the more commonly used z-values are on pages 274 and 275, but 80% is not there.

The z of .2 can be converted to a z of .8 by subtracting from 1. So the charts can be used for both type I error and type II error with type II error converting to Power when subtracted from 1.

Standard Normal distribution

The z-values of 0.84 and 1.282 in the standard normal distribution correspond to 80% and 90% power, respectively. These can be obtained from the standard normal probability table in the back of the textbook (p. 272, etc.).

For Example: The 80% power, you need to locate the 80th percentile under the standard normal curve. Look inside the table, and try to find the area (cumulative probability) of 0.80. 0.7995 is closest to 0.80, and, so, we'll use that value; the corresponding z-value is 0.84. The 80th percentile is 0.84, or $P(Z < 0.84) = 0.80$ (approximately).

Type I Error

The numbers for power are fixed as are the z statistics associated with the type 1 error rate. You only need to know what percent and whether your test is one-tailed or two-tailed to know which fixed value to use.