Power and sample size determination using Stata

The Stata function ***sampsi*** can be used to estimate the required sample size or power of tests in one- and two-sample problems. It can be used for comparing either means or proportions. It can also be used in repeated measures studies. Repeated measures analysis is not covered in the Biostat 511-513 sequence and will not be discussed here.

The commands can be written with the following options (repeated measures options have been omitted):

sampsi #1 #2, alpha(#) power(#) n1(#) n2(#) ratio(#) sd1(#) sd2(#) onesample onesided

where

#1 is the value you specify for the null hypothesis, μ_0.
#2 is the value you specify for the alternative hypothesis μ_1.
alpha is the significance level (default is 0.05).
power is the power of the test (default is 0.90).
n1 is the sample size for the first sample.
n2 is the sample size for the second sample.
ratio is equal to ***n2(#)/n1(#)*** (default is 1.0, see below for details).
sd1 is the standard deviation for the first sample.
sd2 is the standard deviation for the second sample.
onesample perform computations for a one-sample test (default is two-sample test).
onesided perform a one-sided hypothesis test (default is two-sided test).

sampsi computes sample size or power for four types of tests:

1. Two-sample comparisons of means.

   The value for μ_1 is ***#1***.
   The value for μ_2 is ***#2***.
   The postulated standard deviations are ***sd1(#)*** and ***sd2(#)***.

2. One-sample comparisons of null mean to a hypothesized value.

   The ***onesample*** option must be specified.
   The value for μ_0 is ***#1***.
   The value for μ_1 is ***#2***.
   The postulated standard deviation is ***sd1(#)***.

3. Two sample comparisons of proportions.

   The value for p_1 is ***#1***.
   The value for p_2 is ***#2***.

4. One sample comparisons of null proportion to a hypothesized value.

   The ***onesample*** option must be specified.
   The value for p_0 is ***#1***.
   The value for p_1 is ***#2***.

Power and sample size determination using Stata

Additional notes:

1. ***n1(#)*** specifies the size of the first (or only) sample and ***n2(#)*** specifies the size of the second sample. If specified, *sampsi* reports the power calculations. If not specified, *sampsi* computes sample size.
2. ***ratio(#)*** used in two-sample tests, allows one to compute the sample size when the sample sizes for the two groups are designed to be unequal. This would be used, for example, when cases and controls are not equally selected (e.g., 2 controls for every case - ***ratio(2)***).
3. ***sd1(#)*** and ***sd2(#)*** are the standard deviations for the comparisons of means tests. If not specified, a comparison of proportions is made. In two-sample cases, if only ***sd1(#)*** is specified, ***sd2(#)*** is assumed to equal to ***sd1(#)***.
4. ***onesample*** indicates a one-sample test. The default is a two-sample test.
5. ***onesided*** indicates a one-sided test. The default is a two-sided test.

Examples:

1. Determining the sample size (per group) in a two-sample comparison of means.

   Assume: H_o: $\mu_1=\mu_2$ vs. H_1: $\mu_1 \neq \mu_2$ (two-sided test).

   $\alpha = 0.05$, $\sigma_1=12$, $\sigma_2=18$.

   We would like our test to have at least 80% power for detecting a difference in the means of 5 units. That is, Power = $1-\beta = 0.80$, and $\Delta = |\mu_1-\mu_2|=5$. The Stata command is:

   sampsi 0 5, alpha(0.05) power(.80) sd1(12) sd2(18).

   We picked the two values ***0*** and ***5*** for #1 ($=\mu_1$) and #2 ($=\mu_2$) so the absolute value of their difference is ***5***. Any two values such that the absolute value of their difference is ***5*** will be sufficient in this example (e.g., ***109 114***).

2. Computing power in a two-sample test with unequal sample sizes.

   Assume: H_o: $\mu_1=\mu_2$ vs. H_1: $\mu_1 > \mu_2$ (one-sided test).

   $\alpha = 0.01$, $\sigma_1=15$, $\sigma_2=11$, $n_1=55$, $n_2=40$,

   We would like to estimate the power for detecting a difference in the means of 10 units ($\Delta = |\mu_1-\mu_2|=10$). The Stata command is: ***sampsi 10 20, alpha(0.01) n1(55) n2(40) sd1(15) sd2(11) onesided.***

3. Estimating sample size in a two-sample test where twice as many subjects will be sampled from group 1 as group two. That is, the ratio n2/n1=0.5.

   Assume: H_o: $\mu_1=\mu_2$ vs. H_1: $\mu_1 \neq \mu_2$ (two-sided test).

   $\alpha = 0.01$, $\sigma_1=8$, $\sigma_2=5$, ratio=n_2/n_1=0.50

   We would like our test to have 99% power for detecting a difference in the means of 7 units. Power = 0.99 and $\Delta = |\mu_1-\mu_2|=7$. The appropriate Stata command is:

   sampsi 10 17, alpha(0.01) power(0.99) sd1(8) sd2(5) ratio(0.5).

4. Estimating sample size in a one-sample test.

   Assume: H_o: $\mu=210$ ($=\mu_0$) vs. H_1: $\mu > 210$ (one-sided test).

$\alpha = 0.02$, $\sigma_1=24$, $\sigma_2=24$, the alternative mean $\mu_1=224$ yielding $\Delta = |\mu_0-\mu_1|=|210-224|=14$

We would like our test to have at least 85% power for detecting a difference in the means of 14 units. The Stata command is: ***sampsi 210 224, alpha(0.02) power(.85) sd1(24) onesided onesample.***

5. Computing power in a one-sample test.

Assume: H_o: $\mu=240$ ($=\mu_0$) vs. H_1: $\mu \neq 240$ ($=\mu_0$) (two-sided test).

$\alpha = 0.01$, $\sigma=30$, $n=12$, $\mu_1=270$

The Stata command is: ***sampsi 240 270, alpha(0.01) n1(12) sd1(30) onesample.***

6. Sample size determination in a two-sample comparison of proportions.

Assume: H_o: $p_1= p_2$ vs. H_1: $p_1 \neq p_2$ (two-sided test).

$\alpha = 0.10$, $n_1= n_2$, $p_1=0.25$, $p_2=0.40$, and power=75%

The Stata command is: ***sampsi 0.25 0.40, alpha(0.10) ratio(1) power(0.75).***

7. Computing power in a two-sample test of proportions.

Assume: H_o: $p_1= p_2$ vs. H_1: $p_1 > p_2$ (one-sided test).

$\alpha = 0.05$, $n_1=400$, $n_2=500$, $p_1=0.25$, $p_2=0.35$

The Stata command is: ***sampsi 0.25 0.35, alpha(0.05) n1(400) n2(500) onesided.***

8. Sample size estimation for a one-sample comparison of a proportion.

Assume: H_o: $p=0.20$ ($=p_0$) vs. H_1: $p \neq 0.20$ (two-sided test).

$\alpha = 0.05$, Power=0.95, $p_1=0.25$

The Stata command is: ***sampsi 0.20 0.25, alpha(0.05) power(0.95) onesample.***