Sample Size Estimation for Longitudinal Studies

Comparison of two groups at a single timepoint

Number of subjects (N) in each of two groups (Fleiss, 1986):

$$N = \frac{2(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_1 - \mu_2)^2} = \frac{2(z_\alpha + z_\beta)^2}{[(\mu_1 - \mu_2)/\sigma]^2}$$

- z_α is the value of the standardized score cutting off $\alpha/2$ proportion of each tail of a standard normal distribution (for a two-tailed hypothesis test)
- z_β is the value of the standardized score cutting off the upper β proportion
- σ^2 is the assumed common variance in the two groups
- $\mu_1 - \mu_2$ is the difference in means of the two groups

Some common choices:

- $z_\alpha = 1.645, 1.96, 2.576$ for 2-tailed .10, .05, and .01 test
- $z_\beta = .842, 1.036, 1.282$ for power = .8, .85. and .90
- effect size = $(\mu_1 - \mu_2)/\sigma = .2, .5, .8$ for "small," "medium," and "large" effects (Cohen, 1988)

Example

- $z_\alpha = 1.96$ 2-tailed .05 hypothesis test
- $z_\beta = .842$ power = .8
- effect size $(\mu_1 - \mu_2)/\sigma = .5$

$$N = \frac{2(1.96 + .842)^2}{(.5)^2} = 15.7/.25 = 62.8$$

$\Rightarrow$ need 63 subjects in each group

notice $N \approx (4/\delta)^2$, where δ = effect size

Comparison of two groups across time
consistent difference across time

Number of subjects N in each of two groups (Diggle *et al.*, 2002)

$$N = \frac{2(z_\alpha + z_\beta)^2 \, (1 + (n-1)\rho)}{n[(\mu_1 - \mu_2)/\sigma]^2}$$

- σ^2 is the assumed common variance in the two groups
- $\mu_1 - \mu_2$ is the difference in means of the two groups
- n is the number of timepoints
- ρ is the assumed correlation of the repeated measures

Example

- $z_\alpha = 1.96$ 2-tailed .05 hypothesis test
- $z_\beta = .842$ power = .8
- effect size $(\mu_1 - \mu_2)/\sigma = .5$
- $n = 2$ timepoints
- $\rho = .6$ correlation of repeated measures

$$N = \frac{2(1.96 + .842)^2(1 + (2-1) \times .6)}{2 \times (.5)^2} = \frac{(15.7)(1.6)}{(2)(.25)} = 50.3$$

$\Rightarrow$ need approximately 50 subjects in each group

if $\rho = 0$ then $N = 31.4$ (cross-sectional)
if $\rho = 1$ then $N = 62.8$ (one-timepoint)

Comparing two groups across timepoints - balanced case

As in Overall and Doyle (1994), sample size of contrast Ψ_c of group population means across n timepoints:

$$N = \frac{2(z_\alpha + z_\beta)^2 \sigma_c^2}{\Psi_c^2}$$

with

$$\Psi_c = \sum_{i=1}^{n} c_i(\mu_{1i} - \mu_{2i})$$

$$\sigma_c^2 = \sum_{i=1}^{n} c_i^2 \sigma_i^2 + 2 \sum_{i<j}^{n} c_i c_j \sigma_{ij}$$

- σ_i^2 = common variance in the two groups at timepoint i
- σ_{ij} = common covariance in the two groups between timepoints i and j
- c_i = contrast applied at timepoint i

If the sample size is known and the degree of power is to be determined, the formula can be re-expressed as:

$$z_\beta = \sqrt{\frac{N\Psi_c^2}{2\sigma_c^2}} - z_\alpha = \sqrt{\frac{\Psi_c^2}{V(\hat{\Psi}_c)}} - z_\alpha$$

where the variance of the sample contrast $\hat{\Psi}_c$ equals

$$V(\hat{\Psi}_c) = \frac{2}{N}\sigma_c^2$$

Example

- $z_\alpha = 1.96$ 2-tailed .05 hypothesis test
- $z_\beta = .842$ power = .8
- $n = 2$ timepoints
- variance-covariance of repeated measures

$$V(y) = \begin{bmatrix} 1 & .6 \\ .6 & 1 \end{bmatrix}$$

I. Average group difference over time

- mean difference $\mu_1 - \mu_2 = .5$ at both $t1$ and $t2$
- time-related contrasts: $c_1 = c_2 = 1/2$ (*i.e.*, average over time)

$$\Psi_c = \frac{1}{2}(.5) + \frac{1}{2}(.5) = .5$$

$$\sigma_c^2 = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + 2\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)(.6) = .8$$

contrast effect size $\delta = \Psi_c/\sigma_c = .5/\sqrt{.8} = .56$

$$N = \frac{2(1.96 + .842)^2}{(.56)^2} = 50$$

Notice

- if $\rho = 1$, then $\sigma_c^2 = 1,\ \delta = .5,\ N = 63$ (one-timepoint)
- if $\rho = 0$, then $\sigma_c^2 = 1/2,\ \delta = 1,\ N = 16$ (cross-sectional)

where ρ is the assumed correlation of the repeated measures

II. Group difference across time

- mean difference $\mu_1 - \mu_2 = 0$ at $t1$ and $.5$ at $t2$
- time-related contrasts: $c_1 = -1$ and $c_2 = 1$

$$\Psi_c = -1(0) + 1(.5) = .5$$

$$\sigma_c^2 = (-1)^2(1)^2 + (1)^2(1)^2 + 2(-1)(1)(.6) = .8$$

contrast effect size $\delta = \Psi_c/\sigma_c = .5/\sqrt{.8} = .56$

$$N = \frac{2(1.96 + .842)^2}{(.56)^2} = 50$$

Notice

- if N was calculated based on $t2$ only, then $N = 63$
  $H_0 : \mu_{12} = \mu_{22} \quad \neq \quad H_0 : (\mu_{12} - \mu_{11}) = (\mu_{22} - \mu_{21})$
- if $\rho = 1$, then $\sigma_c^2 = 0$
- if $\rho = .9$, then $\sigma_c^2 = .2,\ \delta = 1.12,\ N = 14$
- if $\rho = 0$, then $\sigma_c^2 = 1/2,\ \delta = .25,\ N = 63$ cross-sectional

For average group effect over time

- as $\rho \uparrow$, then $N \uparrow$

since it's a between-subjects comparison of averages

$\Rightarrow$ less subjects needed if the averages are based on more independent data

For group difference across time

- as $\rho \uparrow$, then $N \downarrow$

since it's a between-subjects comparison of a within-subjects comparison

$\Rightarrow$ less subjects needed if the subject differences (i.e., pre to post) are based on more reliable data

More than 2 timepoints

- mean differences across time
- var-covar and/or correlation of repeated measures
- time-related contrast

3 timepoints

$t1$	$t2$	$t3$	
1/3	1/3	1/3	average across time
-1	0	1	linear trend
1	-2	1	quadratic trend

trend coefficients from tables of orthogonal polynomials

4 timepoints

$t1$	$t2$	$t3$	$t4$	
1/4	1/4	1/4	1/4	average across time
-3	-1	1	3	linear trend
1	-1	-1	1	quadratic trend
-1	3	-3	1	cubic trend

often investigators expect

- overall group difference, or
- group by (approximately) linear time interaction

What about Attrition?

- could use N from calculations as N for last timepoint
  - e.g., $N = 50$, retention at last timepoint = .9
    $\Rightarrow$ start the study with $50/.9 = 56$ subjects
- can build the retention rate information into the sample size formula
  - Hedeker, Gibbons, & Waternaux (1999), *JEBS*, 24:70-93
    program RMASS2 available at
    www.uic.edu/∼hedeker/works.html

Comparing two groups across timepoints
unbalanced case

Denote sample size in first group as N_{1i} and second group as N_{2i} at timepoint i ($i = 1, \ldots, n$). The variance of the sample contrast $\hat{\Psi}_c$ equals

$$V(\hat{\Psi}_c) = \sum_{i=1}^{n} c_i^2 \sigma_i^2 \left(\frac{1}{N_{1i}} + \frac{1}{N_{2i}} \right) + 2 \sum_{i<j}^{n} c_i c_j \sigma_{ij} \left(\frac{1}{\sqrt{N_{1i} N_{1j}}} + \frac{1}{\sqrt{N_{2i} N_{2j}}} \right)$$

To calculate power for the sample contrast,

$$z_\beta = \sqrt{\frac{\Psi_c^2}{V(\hat{\Psi}_c)}} - z_\alpha$$

Use sample size in first group at first timepoint (N_{11}) as a reference

- define retention rates for this group as r_{1i} for timepoints $i = 1, \ldots, n$, which indicate the proportion of N_1 subjects observed at timepoint i
(note that $r_{11} = 1$ and $N_{1i} = r_{1i} N_{11}$)
- similarly, define N_{21} and r_{2i} for group two

Then,

$$V(\hat{\Psi}_c) = \frac{1}{N_{11}} \left[\sum_{i=1}^{n} c_i^2 \sigma_i^2 \left(\frac{1}{r_{1i}} + \frac{1}{r_{2i}} \frac{N_{11}}{N_{21}} \right) + 2 \sum_{i<j}^{n} c_i c_j \sigma_{ij} \left(\frac{1}{\sqrt{r_{1i} r_{1j}}} + \frac{N_{11}}{N_{21}} \frac{1}{\sqrt{r_{2i} r_{2j}}} \right) \right]$$

and, denoting the ratio of sample sizes at the first timepoint (N_{11}/N_{21}) as $N_{.1}$, then

$$V(\hat{\Psi}_c) = \frac{1}{N_{11}} \left[\sum_{i=1}^{n} c_i^2 \sigma_i^2 \left(\frac{1}{r_{1i}} + \frac{N_{.1}}{r_{2i}} \right) + 2 \sum_{i<j}^{n} c_i c_j \sigma_{ij} \left(\frac{1}{\sqrt{r_{1i} r_{1j}}} + N_{.1} \frac{1}{\sqrt{r_{2i} r_{2j}}} \right) \right]$$

If the retention rates are equal for the two groups across time $r_{1i} = r_{2i} = r_i$, then

$$V(\hat{\Psi}_c) = \frac{N_{.1} + 1}{N_{11}} \left[\sum_{i=1}^{n} \frac{c_i^2 \sigma_i^2}{r_i} + 2 \sum_{i<j}^{n} \frac{c_i c_j \sigma_{ij}}{\sqrt{r_i r_j}} \right] = \frac{N_{.1} + 1}{N_{11}} \sigma_{rc}^2$$

where σ_{rc}^2 extends σ_c^2 given earlier, namely

$$\sigma_c^2 = \sum_{i=1}^{n} c_i^2 \sigma_i^2 + 2 \sum_{i<j}^{n} c_i c_j \sigma_{ij}$$

for the case where sample sizes vary across timepoints (although group retention rates are assumed equal)

To calculate power for any of the above variance formulations of the sample contrast,

$$z_\beta = \sqrt{\frac{\Psi_c^2}{V(\hat{\Psi}_c)}} - z_\alpha$$

In particular, for the case of common retention rates across time

$$z_\beta = \sqrt{\left(\frac{N_{11}}{N_{.1}+1}\right)\frac{\Psi_c^2}{\sigma_{rc}^2}} - z_\alpha$$

where $N_{.1}$ is the sample size ratio between groups

Re-expressing, the number of subjects needed in the first group at the first timepoint equals:

$$N_{11} = \frac{(N_{.1}+1)(z_\alpha + z_\beta)^2 \sigma_{rc}^2}{\Psi_c^2}$$

Based on the sample size ratio between groups $N_{.1}$ and retention rates r_i, required sample size at each timepoint for both groups can be calculated

Example

- $z_\alpha = 1.96$ 2-tailed .05 hypothesis test
- $z_\beta = .842$ power = .8
- $n = 2$ timepoints, retention rates $r_1 = 1$ and $r_2 = .8$
- sample size ratio $N_{.1} = 1$
- variance-covariance of repeated measures

$$V(y) = \begin{bmatrix} 1 & .6 \\ .6 & 1 \end{bmatrix}$$

I. Average group difference over time

- mean difference $\mu_1 - \mu_2 = .5$ at both $t1$ and $t2$
- time-related contrasts: $c_1 = c_2 = 1/2$

$$\Psi_c = \frac{1}{2}(.5) + \frac{1}{2}(.5) = .5$$

$$\sigma_{rc}^2 = \left(\frac{1}{2}\right)^2 + \frac{(1/2)^2}{.8} + 2\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)(.6)/\sqrt{.8} = .9$$

contrast effect size $\delta = \Psi_c/\sigma_{rc} = .5/\sqrt{.9} = .53$

$$N_{11} = \frac{2(1.96 + .842)^2}{(.53)^2} = 56.4$$

Note:if $r_2 = 1$ then $N_{11} = 50$

II. Group difference across time

- mean difference $\mu_1 - \mu_2 = 0$ at $t1$ and .5 at $t2$
- time-related contrasts: $c_1 = -1$ and $c_2 = 1$

$$\Psi_c = -1(0) + 1(.5) = .5$$

$$\sigma_{rc}^2 = (-1)^2(1)^2 + \frac{(1)^2(1)^2}{.8} + 2(-1)(1)(.6)/\sqrt{.8} = .91$$

contrast effect size $\delta = \Psi_c/\sigma_{rc} = .5/\sqrt{.91} = .525$

$$N_{11} = \frac{2(1.96 + .842)^2}{(.525)^2} = 57.1$$

Note:if $r_2 = 1$ then $N_{11} = 50$

Dichotomous outcomes

Comparison of two groups at a single timepoint

Number of subjects (N) in each of two groups (Fleiss, 1981):

$$N = \frac{\left[z_\alpha(2\bar{p}\bar{q})^{1/2} + z_\beta(p_1q_1 + p_2q_2)^{1/2}\right]^2}{(p_1 - p_2)^2}$$

- p_1 = response proportion in group 1 $\quad (q_1 = 1 - p_1)$
- p_2 = response proportion in group 2 $\quad (q_2 = 1 - p_2)$
- $\bar{p} = (p_1 + p_2)/2$
- $\bar{q} = 1 - \bar{p}$

Example

- $z_\alpha = 1.96$ 2-tailed .05 hypothesis test
- $z_\beta = .842$ power = .8
- $p_1 = .5$ and $p_2 = .7$

$$N = \frac{\left[1.96(2 \times .6 \times .4)^{1/2} + .842(.5 \times .5 + .7 \times .3)^{1/2}\right]^2}{(.5 - .7)^2}$$
$$= 93.03$$

Dichotomous outcomes - longitudinal case

The number of subjects (N) in each of two groups for a consistent difference in proportions $p_1 - p_2$ between two groups across n timepoints (Diggle *et al.*, (2002):

$$N = \frac{\left[z_\alpha(2\bar{p}\bar{q})^{\frac{1}{2}} + z_\beta(p_1q_1 + p_2q_2)^{\frac{1}{2}}\right]^2 (1 + (n-1)\rho)}{n(p_1 - p_2)^2}$$

- p_1 = response proportion in group 1 $\quad (q_1 = 1 - p_1)$
- p_2 = response proportion in group 2 $\quad (q_2 = 1 - p_2)$
- $\bar{p} = (p_1 + p_2)/2$
- $\bar{q} = 1 - \bar{p}$
- ρ is the common correlation across the n observations

Example

- $z_\alpha = 1.96$ 2-tailed .05 hypothesis test
- $z_\beta = .842$ power $= .8$
- $n = 2$ timepoints
- correlation of repeated outcomes $= .6$
- $p_1 = .5$ and $p_2 = .7$

$$N = \frac{\left[1.96(2 \times .6 \times .4)^{\frac{1}{2}} + .842(.5 \times .5 + .7 \times .3)^{\frac{1}{2}}\right]^2 (1 + (2-1).6)}{2(.5 - .7)^2}$$
$$= 74.42$$

if $\rho = 0$ then $N = 46.51$ (cross-sectional)
if $\rho = 1$ then $N = 93.03$ (one-timepoint)

Ordinal outcomes

- methods not as developed
- use methods for continuous outcomes, but adjust the detectable effect sizes by an efficiency loss (*e.g.*, 80%)
  - Armstrong & Sloan (1989, Amer Jrn of Epid) report efficiency losses between 89% to 99% comparing an ordinal to continuous outcome, depending on the number of categories and distribution within the ordinal categories
  - Strömberg (1996, Amer Jrn of Epid) report efficiency losses between 87% to 97% comparing an ordinal outcome with 3 or 4 categories to one with 5 categories