BIOS 755: Sample Size and Power Estimation in Longitudinal Data Analysis

Alexander McLain

April 13, 2023

Outline

- Quantities required to determine the sample size in longitudinal studies
- Review of type I error, type II error, and power
- For continuous data
  - Comparison of two groups for univariate data
  - Comparison of two groups across time
  - Comparison of two groups for the rate of change
  - Comparing Two Groups Across Timepoints (Balanced Case)
- For dichotomous outcome
  - Comparison of two groups across time
- Sample size with missing data.

Review of terms

Quantities required to determine the sample size in longitudinal studies

- Type I error rate (α).
- Type II error rate (β), power ($P = 1 - \beta$).
- Smallest meaningful difference to be detected (d) .
- Measurement variation (σ^2).
- Number of repeated observations per person (n).
- Correlation among the repeated observations (ρ) or a general correlation matrix ($\boldsymbol{R}$).

Hypothesis testing, type I and II error, and power

Let H_0 denote the null hypothesis, and H_1 denote the alternative hypothesis,

Conclusion	H0 is true	H1 is true
Reject H_0	Type I error (α)	Power (1-β)
Fail to reject H_0	★	Type II error (β)

- Choice of α: often α is specified at 0.05; $z_{\alpha/2} = 1.96$ for a two-sided test and $z_\alpha = 1.64$ for a one-sided test.
- Choice of β: power of a test is the probability of detecting a true underlying difference and depends on the alternative hypothesis. The power (1 - β) is often set to 0.8, i.e., $z_\beta = .842$.
- We want to choose the sample size to ensure the desired power for detecting the smallest meaningful difference.

Comparison of two groups for univariate continuous outcome

The number of subjects (N) needed in each of two groups is

$$N = \frac{2(z_{\alpha/2} + z_{\beta})^2 \sigma^2}{d^2} = \frac{2(z_{\alpha/2} + z_{\beta})^2}{\Delta^2},$$

where $\Delta = d/\sigma$ is referred to as the effect size and σ^2 is the assumed common variance in the two groups.

- $N \uparrow$ if $\sigma^2 \uparrow$;
- $N \uparrow$ if $1 - \beta \uparrow$ (or $\beta \downarrow$);
- $N \uparrow$ if $\alpha \downarrow$;
- $N \uparrow$ if $d \downarrow$.

Comparison of two groups for univariate continuous outcome

Example

- Suppose we have a type I error of 0.05, type II error of 0.2 (power of 80%), two-sided, effect size of 0.7, the required sample size is

$$N = \frac{2(1.96 + .842)^2}{0.7^2} = \frac{15.7}{0.49} = 32.04$$

- Note that $N \approx (4/\Delta)^2$.
- The sample size formula can also be manipulated to determine the power for a given sample size.
- The sample size formula can also be modified to allow groups of unequal size.

Comparison of two groups across time

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

$$N = \frac{2(z_{\alpha/2} + z_{\beta})^2\{1 + (n-1)\rho\}}{nd^2}$$

where

- σ^2 is the assumed common variance in the two groups
- $d = (\mu_1 - \mu_2)/\sigma$ is the standardized different between the two groups.
- n is the number of timepoints
- ρ assumed within subject correlation.

Comparison of two groups across time

Example

- Suppose we have a type I error of 0.05, type II error of 0.2 (power of 80%), effect size of $d = (\mu_1 - \mu_2)/\sigma = 0.5$, $n = 2$ timepoints, and $\rho = 0.6$ correlation.
- The required sample size is

$$N = \frac{2(1.96 + .842)^2\{1 + (2-1)0.6\}}{2 \times 0.5^2} = \frac{15.7(1.6)}{2(25)} = 50.3$$

  thus we need ≈ 50 subjects per group.
- if $\rho = 0$ then $m = 31.4$.
- if $\rho = 1$ then $m = 62.8$.

Comparison of two groups for the rate of change for continuous outcomes

- Consider a simple problem of comparing two groups, A and B with continuous outcomes. Assuming the responses depend on a single covariate as follows:

$$Y_{kij} = \beta_{0k} + \beta_{1k} x_{kij} + e_{kij}, \qquad j = 1, \ldots, n;\ i = 1, \ldots, N;\ k = A \text{ or } B.$$

- Both groups have the same number of subjects, N.
- We assume that $\mathrm{Var}(e_{kij}) = \sigma^2$ and $\mathrm{Cor}(Y_{kij}, Y_{kij'}) = \rho$ for all $j \neq j$.
- We also assume that each person has the same set of covariate so that $x_{kij} = x_j$.
- The regression coefficients β_{1A} and β_{1B} are the rate of changes in Y for groups A and B, respectively.

Comparison of two groups for the rate of change for continuous outcomes

- With n fixed and known, the number of subjects (m) in each of two groups are needed to achieve type I error rate α and power 1 - β for comparing the rate of change in a continuous response between two groups, is

$$N = \frac{2(z_{\alpha/2} + z_{\beta})^2(1-\rho)\sigma^2}{ns_x^2 d^2} = \frac{2(z_{\alpha/2} + z_{\beta})^2(1-\rho)}{ns_x^2 \Delta^2}$$

  where $d = \beta_{1B} - \beta_{1A}$, $\Delta = d/\sigma$, and $s_x^2 = \sum_j (x_j - \bar{x})^2/n$ is the within-subject variance of the x_j.
- When $\rho \uparrow$, N?
- When $s_x^2 = \sum_j (x_j - \bar{x})^2/n \uparrow$, N?
  - For fixed length of study, τ and no requirement for the spacing between repeated measurements:
  - For equally spaced measurements:

Comparison of two groups for the rate of change for continuous outcomes

Example

- Consider a hypothetical clinical trial on the effect of a new treatment in reducing blood pressure.
- There are three visits, including the baseline, are planned at years 0, 2, and 5. Thus, $n = 3$ and $s_x^2 = 4.22$.
- With type I error of 0.05, type II error of 0.2 (power of 80%), one-sided test, testing smallest meaningful difference d $=$ .5 mmHg/year.

Comparison of two groups for the rate of change for continuous outcomes

Example (cont.)

- Below are the number of subjects required for both treated and control groups for some selected values of ρ and σ^2, are

	σ^2		
ρ	100	200	300
0	391	781	1172
0.2	313	625	938
0.5	196	391	586
0.8	79	157	235

- Note for each value of σ^2, the required sample size decreases as the correlation, ρ, increases.

Comparing Two Groups Across Timepoints (Balanced Case)

- As in Overall and Doyle (1994), sample size of contrast 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

Comparing Two Groups Across Timepoints (Balanced Case)

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

$$z_\beta = \sqrt{\frac{N\Psi_c^2}{2\sigma_2^2}} - z_\alpha = \sqrt{\frac{\Psi_c^2}{V(\hat{\Psi}_c)}} - z_\alpha$$

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

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

Comparing Two Groups Across Timepoints (Balanced Case)

Example

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

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

Comparing Two Groups Across Timepoints (Balanced Case)

Example (cont.) 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 $\delta = \Psi_c/\sigma_c = .5/\sqrt{.8} = .56$

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

Comparing Two Groups Across Timepoints (Balanced Case)

Example (cont.) Note that

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

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

Comparing Two Groups Across Timepoints (Balanced Case)

Example (cont.) 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 $\delta = \Psi_c/\sigma_c = .5/\sqrt{.8} = .56$

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

Comparing Two Groups Across Timepoints (Balanced Case)

Example (cont.) Note that

- 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

Comments

- More than 2 timepoints
  - mean differences across time
  - variance-covariance 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

Comments

- 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

SAS code

```
PROC IML;
za = PROBIT(.975);
zb = PROBIT(.8);
meandiff = 0, .25, .5;
contrast = -1, 0, 1;
corrmat = 1 .5 .25,.5 1 .5 , .25 .5 1;
contdiff = T(contrast) * meandiff;
contvar = T(contrast)*corrmat*contrast;
NperGrp = ((2*(za+zb)**2) * contvar)/(contdiff**2);
PRINT NperGrp;
```

determines number per group; 3 timepoints; linear increasing effect sizes of 0 .25 .5; group by linear contrast across time; AR1 structure with rho=.5; power = .8 for a 2-tailed .05 test.

Dichotomous Outcomes

- 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):

$$N = \frac{\left[z_\alpha (2\bar{p}\bar{q})^{\frac{1}{2}} + z_\beta (p_1 q_1 + p_2 q_2)^{\frac{1}{2}}\right]^2 (1 + (n-1)\rho)}{n(p_1 - p_2)^2}$$

- $p_1 =$ 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

Dichotomous Outcomes

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 $p_2 = .7$

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

$$= 74.42$$

SAS Code

```
DATA one;
za = PROBIT(.975);          /* determines number per group;
zb = PROBIT(.8);               5 timepoints (ICC=.4);
n = 5;                         difference in proportions of .5 and .67 (OR=2);
p1 = .5; p2 = 2/3;             power = .8 for a 2-tailed .05 test; */
q1 = 1-p1; q2 = 1-p2;
pbar = (p1+p2)/2;
qbar = (q1+q2)/2;
rho = .4;
num = ((za*SQRT(2*pbar*qbar) + zb*SQRT(p1*q1 + p2*q2))**2)*(1 + (n-
   1)*rho);
den = n*((p1-p2)**2);
npergrp = num/den;
PROC PRINT;VAR npergrp;
RUN;
```

Missing Data and Sample Size

- We have assumed no missing data or attrition so far.
- The impact of missing data is difficult to quantify precisely because it depends on the patterns of missingness.
- An ad hoc approach is to inflate the required sample size to account for the assumed attrition rate. For example, if the attrition rate is assumed to be 10%, then the target sample size should be $m/0.9$.

Missing Data and Sample Size

- We have assumed no missing data or attrition so far.
- The impact of missing data is difficult to quantify precisely because it depends on the patterns of missingness.
- An ad hoc approach is to inflate the required sample size to account for the assumed attrition rate. For example, if the attrition rate is assumed to be 10%, then the target sample size should be $m/0.9$.

Missing Data and Sample Size

Further Reading:

- Chapter 2 and 8.5 in **"Analysis of Longitudinal Data"** and Chapter 15 of Fitzmaurice, Laird and Ware (2004). Both listed on syllabus.

Further Reading:

- Hedeker D, Gibbons R, and Waternaux C. (1999) Sample size estimation for longitudinal designs with attrition: comparing time-related contrasts between two groups. *Journal of Educational and Behavioral Statistics*, 24(1):70- 93.
- Heo M and Leon AC. (2008). Statistical power and sample size requirements for three level hierarchical cluster randomized trials. *Biometrics* 64(4):1256-62.
- Liu G and Liang KY. (1997). Sample size calculations for studies with correlated observations. *Biometrics* 53(3):937-47.