Power for linear models of longitudinal data with applications to Alzheimer's Disease Phase II study design

Michael C. Donohue, Steven D. Edland, Anthony C. Gamst
Division of Biostatistics and Bioinformatics
University of California, San Diego

May 16, 2016

1 Introduction

We will discuss power and sample size estimation for randomized placebo controlled studies in which the primary inference is based on the interaction of treatment and time in a linear mixed effects model (Laird and Ware, 1982). We will demonstrate how the sample size formulas of (Liu and Liang, 1997) for marginal or model fit by generalized estimating equation (GEE) (Zeger and Liang, 1986) can be adapted for mixed effects models. Finally, using mixed effects model estimates based on data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), we will give examples of sample size calculations for models with and without baseline covariates which may help explain heterogeneity in cognitive decline and improve power.

2 Power calculations

2.1 Exchangeable correlation and random intercept models

Suppose we wish to estimate the required sample size for inference regarding the interaction of treatment and time in a longitudinal, placebo controlled study. Such calculations are relatively straightforward when the inference is based on a GEE model in which the correlation structure is assumed to be "exchangeable." An exchangeable correlation structure specifies that all observations from within the same cluster, or repeated measures on the same subject, are equally correlated. This is exactly equivalent to a random effects model which includes a random intercept for each cluster of correlated observations. Sample sizes for study designs using these models can be calculated using a simple formula such as that in (Diggle *et al.*, 1994), page 29. The formula requires the number visits, the interval between visits, the estimated model variance (σ^2), the within subject correlation (ρ), and of course the usual sample size calculation inputs (power, significance level, and effect size).

To translate the formula of Diggle *et al.* (1994) to the random effects setting, let us first consider the details of the assumed error structure of the GEE framework. The GEE model assumes that the response for subject i at time t_{ij}, denoted Y_{ij}, is the group mean, dependent on time and treatment, plus an error term ε_{ij}. Or, borrowing notation from Diggle *et al.* (1994), for group A:

$$Y_{ij} = \beta_{0A} + \beta_{1A} t_{ij} + \varepsilon_{ij}, \quad i = 1, \ldots, m; j = 1, \ldots, n.$$

and similarly for Group B. The null hypothesis is $H_0 : d = \beta_{1A} - \beta_{1B} = 0$. Under an exchangeable correlation structure $\text{var}(Y_{ij}) = \text{var}(\varepsilon_{ij}) = \sigma^2$ and $\text{corr}(Y_{ij}, Y_{ik}) = \text{corr}(\varepsilon_{ij}, \varepsilon_{ik}) = \rho$, for all subjects, i, and time points j, k.

In the mixed effects framework we can assume a random intercept model which is equivalent to the GEE model with exchangeable correlation structure. In this case we believe $\varepsilon_{ij} = \alpha_i + \varepsilon^*_{ij}$, where α_i is the random intercept term shared by all observations and ε^*_{ij} are independent and identically distributed (iid) error terms. We see that $\text{var}(Y_{ij}) = \text{var}(\varepsilon_{ij}) = \text{var}(\alpha_i) + \text{var}(\varepsilon^*_{ij})$ and $\text{corr}(Y_{ij}, Y_{ik}) = E[(\alpha_i + \varepsilon^*_{ij})(\alpha_i + \varepsilon^*_{ik})]/\sigma^2 =$

$\text{var}(\alpha_i)/\sigma^2$. The variance of the random intercept, $\text{var}(\alpha_i)$, and the residual variance, $\text{var}(\varepsilon_{ij})$, are easily obtainable from the output of mixed effects fitting software so that one might fit a random effects model to pilot data to educate a power calculation using the GEE formula of Diggle *et al.* (1994). Assuming equal numbers in the placebo and active groups, a common visit schedule for all subjects ($t_{ij} = t_{kj}$ for all i, j, k), and a random intercept model; the number of subjects per group is:

$$m = \frac{2(z_\alpha + z_Q)^2(\text{var}(\alpha_i) + \text{var}(\varepsilon^*_{ij}))^2(1 - \text{var}(\alpha_i)/\sigma^2)}{ns_x^2 d^2}$$

where z_p is the pth standard normal quantile, Q is $1 - P$, P is the specified power, and $s_x^2 = n^{-1}\sum_j (t_j - \bar{x})^2$.

2.2 General correlation and random slope models

The random intercept model is not equipped to handle variations in the rate of change from subject to subject. In many diseases, such as Alzheimer's disease, the rate of improvement or decline will vary greatly within the treatment group, regardless of treatment. This variation can be modeled with a random slope term. That is, we assume:

$$Y_{ij} = \beta_{0A} + \beta_{1A} t_{ij} + \alpha_{0i} + \alpha_{1i} t_{ij} + \varepsilon^*_{ij},$$

where we use ε^*_{ij} again to denote iid error and reserve ε_{ij} for possibly correlated error. If we derive the correlation structure of $\varepsilon_{ij} = \alpha_{0i} + \alpha_{1i} t_{ij} + \varepsilon^*_{ij}$, which is necessary in order to use GEE-based sample size formulas, we find that we no longer have an exchangeable correlation structure. In fact $\text{var}(Y_{ij}) = \text{var}(\varepsilon_{ij}) = \text{var}(\alpha_{0i}) + t_{ij}^2 \text{var}(\alpha_{1i}) + 2t_{ij}\text{cov}(\alpha_{0i}, \alpha_{1i}) + \text{var}(\varepsilon^*_{ij})$ and $\text{cov}(Y_{ij}, Y_{ik}) = \text{cov}(\varepsilon_{ij}, \varepsilon_{ik}) = \text{var}(\alpha_{0i}) + t_{ij}t_{ik}\text{var}(\alpha_{1i}) + (t_{ij} + t_{ik})\text{cov}(\alpha_{0i}, \alpha_{1i})$. For the common visit schedule case, the covariance matrix for the vector of correlated errors, $\boldsymbol{\varepsilon_i} = (\varepsilon_{i1}, \ldots, \varepsilon_{in})'$, is of the form:

$$\Sigma = [(\text{var}(\alpha_0) + t_j t_k \text{var}(\alpha_1) + (t_j + t_k)\text{cov}(\alpha_0, \alpha_1))]_{jk} + \text{diag}(\text{var}(\varepsilon^*_j))$$

With this specification of the covariance matrix, one can use the sample size formula of Liu and Liang (1997) for linear GEE models (page 941). (Warning: The formula given on the bottom page 29 of Diggle *et al.* (1994) for general correlation matrices, R, is wrong).

The formula for linear models provided by Liu and Liang (1997) is useful for testing $H_0 : \boldsymbol{\psi} = \mathbf{0}$ for any linear model of the form:

$$Y_{ij} = \mathbf{x}'_{ij}\boldsymbol{\psi} + \mathbf{z}'_{ij}\boldsymbol{\lambda} + \varepsilon_{ij}$$

where $\boldsymbol{\varepsilon_i} \sim N(\mathbf{0}, \sigma^2 R)$ and the covariates for individual i, $\mathbf{x}_i = (\mathbf{x}'_{i1}, \ldots, \mathbf{x}'_{i1})'_{n\times p}$ and $\mathbf{z}_i = (\mathbf{z}'_{i1}, \ldots, \mathbf{z}'_{i1})'_{n\times q}$, arise from a known discrete distribution. For our placebo controlled longitudinal study, the fully specified model is of the form:

$$Y_{ij} = \beta_0 + \beta_1\{\text{Group}_i = A\} + \beta_2 t_{ij} + \beta_3 t_{ij}\{\text{Group}_i = A\}.$$

That is, the parameter of interest for the interaction of treatment and time is $\psi = \beta_3$ and nuisance parameter is $\boldsymbol{\lambda} = (\beta_0, \beta_1, \beta_2)'$. The covariates are distributed as $\mathbf{x}_i = \mathbf{t} = (t_1, \ldots, t_n)'$ and $\mathbf{z}_j = [\mathbf{1}\,\mathbf{1}\,\mathbf{t}]_{n\times 3}$ with probability 1/2 (Group A); and $\mathbf{x}_i = \mathbf{0}$ and $\mathbf{z}_j = [\mathbf{1}\,\mathbf{0}\,\mathbf{t}]_{n\times 3}$ with probability 1/2 (Group B).

The Liu and Liang's formula for linear models can be coded In R as:

```
library(longpower)
```

```
liu.liang.linear.power

function (N = NULL, delta = NULL, u = NULL, v = NULL, sigma2 = 1,
    R = NULL, R.list = NULL, sig.level = 0.05, power = NULL,
    Pi = rep(1/length(u), length(u)), alternative = c("two.sided",
        "one.sided"), tol = .Machine$double.eps^2)
```

```
{
    if (sum(sapply(list(N, delta, sigma2, power, sig.level),
        is.null)) != 1)
        stop("exactly one of 'N', 'sigma2', 'delta', 'power', and 'sig.level' must be NULL")
    if (!is.null(sig.level) && !is.numeric(sig.level) || any(0 >
        sig.level | sig.level > 1))
        stop("'sig.level' must be numeric in [0, 1]")
    alternative <- match.arg(alternative)
    if (sum(c(!is.null(R), !is.null(R.list))) != 1)
        stop("Exactly one of R or R.list must be specified.")
    if (sum(Pi) != 1)
        stop("Pi must sum to 1.")
    if (!is.null(R)) {
        R.list <- lapply(1:length(u), function(i) R)
    }
    Rinv <- lapply(1:length(R.list), function(i) {
        R <- R.list[[i]]
        if (is.null(dim(R)) & length(R) == 1 & length(u[[i]]) >
            1) {
            R <- matrix(R, length(u[[i]]), length(u[[i]])) +
                diag(1 - R, length(u[[i]]))
        }
        else if (is.null(dim(R)) & length(R) == 1 & length(u[[i]]) ==
            1) {
            R <- matrix(R, length(u[[i]]), length(u[[i]]))
        }
        return(solve(R))
    })
    n.body <- quote({
        Ipl <- 0
        for (i in 1:length(u)) Ipl <- Ipl + Pi[i] * t(u[[i]]) %*%
            Rinv[[i]] %*% v[[i]]
        Ipl <- Ipl/sigma2
        Ill <- 0
        for (i in 1:length(u)) Ill <- Ill + Pi[i] * t(v[[i]]) %*%
            Rinv[[i]] %*% v[[i]]
        Illinv <- solve(Ill/sigma2)
        Sigma1 <- 0
        for (i in 1:length(u)) Sigma1 <- Sigma1 + Pi[i] * (t(u[[i]]) -
            Ipl %*% Illinv %*% t(v[[i]])) %*% Rinv[[i]] %*% (u[[i]] -
            v[[i]] %*% Illinv %*% t(Ipl))
        Sigma1 <- Sigma1/sigma2
        (qnorm(1 - ifelse(alternative == "two.sided", sig.level/2,
            sig.level)) + qnorm(power))^2/(delta %*% Sigma1 %*%
            delta)[1, 1]
    })
    if (is.null(N))
        N <- eval(n.body)
    else if (is.null(sig.level))
        sig.level <- uniroot(function(sig.level) eval(n.body) -
            N, c(1e-10, 1 - 1e-10), tol = tol, extendInt = "yes")$root
    else if (is.null(power))
        power <- uniroot(function(power) eval(n.body) - N, c(0.001,
```

```
            1 - 1e-10), tol = tol, extendInt = "yes")$root
    else if (is.null(delta))
        delta <- uniroot(function(delta) eval(n.body) - N, sqrt(sigma2) *
            c(1e-07, 1e+07), tol = tol, extendInt = "downX")$root
    else if (is.null(sigma2))
        sigma2 <- uniroot(function(sigma2) eval(n.body) - N,
            delta * c(1e-07, 1e+07), tol = tol, extendInt = "yes")$root
    else stop("internal error", domain = NA)
    METHOD <- "Longitudinal linear model power calculation (Liu & Liang, 1997)"
    structure(list(N = N, n = N * Pi, delta = delta, sigma2 = sigma2,
        sig.level = sig.level, power = power, alternative = alternative,
        R = R, note = "N is total sample size and n is sample size in each group.",
        method = METHOD), class = "power.longtest")
}
<environment: namespace:longpower>
```

The parameters include `d`, the effect size (possibly vector); `u`, the list of covariate vectors or matrices associated with the parameter of interest; `v`, the respective list of covariate vectors or matrices associated with the nuisance parameter; `sigma2`, the error variance; `R`, the correlation structure; and `Pi` the proportion of covariates of each type (`u`, `v`, and `Pi` are expected to be the same length and sorted with respect to each other).

For example, we can reproduce the table exchangeable correlations on page 29 of Diggle *et al.* (1994) for the case of $\boldsymbol{t} = (0, 2, 5)'$, $\alpha = 0.05$, power=0.80, and $d = 0.5$ as follows via the `diggle.linear.power` function:

```
n = 3
t = c(0,2,5)
rho = c(0.2, 0.5, 0.8)
sigma2 = c(100, 200, 300)
tab.diggle = outer(rho, sigma2,
      Vectorize(function(rho, sigma2){
        ceiling(diggle.linear.power(
          d=0.5,
          t=t,
          sigma2=sigma2,
          R=rho,
          alternative="one.sided",
          power=0.80)$n)}))
colnames(tab.diggle) = paste("sigma2 =", sigma2)
rownames(tab.diggle) = paste("rho =", rho)
tab.diggle

          sigma2 = 100 sigma2 = 200 sigma2 = 300
rho = 0.2          313          625          938
rho = 0.5          196          391          586
rho = 0.8           79          157          235
```

or via the `liu.liang.linear.power` function:

```
u = list(u1 = t, u2 = rep(0,n))
v = list(v1 = cbind(1,1,t),
         v2 = cbind(1,0,t))
tab.ll = outer(rho, sigma2,
      Vectorize(function(rho, sigma2){
```

```
        ceiling(liu.liang.linear.power(
          delta=0.5, u=u, v=v,
          sigma2=sigma2,
          R=rho, alternative="one.sided",
          power=0.80)$N/2)}))
colnames(tab.ll) = paste("sigma2 =", sigma2)
rownames(tab.ll) = paste("rho =", rho)
tab.ll
```

```
          sigma2 = 100 sigma2 = 200 sigma2 = 300
rho = 0.2          313          625          938
rho = 0.5          196          391          586
rho = 0.8           79          157          235
```

As a second example, consider an Alzheimer's disease trial in which assessments are taken every three months for 18 months (7 visits). We assume an smallest detectable effect size of 1.5 points on the cognitive portion of the Alzheimer's Disease Assessment Scale (ADAS-Cog). This is a 70 point scale with great variability among sick individuals. We assume the random intercept to have a variance of 55, the random slope to have a variance of 24, and a residual variance of 10. The correlation between random slope term and random intercept term is 0.8. We can estimate the necessary sample size by first generating the correlation structure. Since $\varepsilon = \text{var}(Y_{ij})$ is not constant over time in this model, we fix `sigma2=1` and set `R` equal to the covariance matrix for $\boldsymbol{\varepsilon_i}$:

```
# var of random intercept
sig2.i = 55
# var of random slope
sig2.s = 24
# residual var
sig2.e = 10
# covariance of slope and intercep
cov.s.i <- 0.8*sqrt(sig2.i)*sqrt(sig2.s)

cov.t <- function(t1, t2, sig2.i, sig2.s, cov.s.i){
        sig2.i + t1*t2*sig2.s + (t1+t2)*cov.s.i
}

t = seq(0,1.5,0.25)
n = length(t)
R = outer(t, t, function(x,y){cov.t(x,y, sig2.i, sig2.s, cov.s.i)})
R = R + diag(sig2.e, n, n)
u = list(u1 = t, u2 = rep(0,n))
v = list(v1 = cbind(1,1,t),
         v2 = cbind(1,0,t))

liu.liang.linear.power(d=1.5, u=u, v=v, R=R, sig.level=0.05, power=0.80)
```

```
     Longitudinal linear model power calculation (Liu & Liang, 1997)

              N = 414.6202
              n = 207.3101, 207.3101
          delta = 1.5
         sigma2 = 1
      sig.level = 0.05
          power = 0.8
```

```
    alternative = two.sided

 NOTE: N is total sample size and n is sample size in each group.

 R:
         [,1]      [,2]      [,3]      [,4]      [,5]     [,6]      [,7]
[1,] 65.00000  62.26636  69.53272  76.79908  84.06544  91.3318  98.59817
[2,] 62.26636  81.03272  79.79908  88.56544  97.33180 106.0982 114.86453
[3,] 69.53272  79.79908 100.06544 100.33180 110.59817 120.8645 131.13089
[4,] 76.79908  88.56544 100.33180 122.09817 123.86453 135.6309 147.39725
[5,] 84.06544  97.33180 110.59817 123.86453 147.13089 150.3972 163.66361
[6,] 91.33180 106.09817 120.86453 135.63089 150.39725 175.1636 179.92997
[7,] 98.59817 114.86453 131.13089 147.39725 163.66361 179.9300 206.19633
```

So the study would require about 207 subjects per arm to achieve 80% power, with a two-tailed $\alpha = 0.05$.

The simple formula provided in Diggle *et al.* (1994) suggests the required number of subjects can be found by $2(z_\alpha + 2_Q)\xi/d^2$, where

$$\xi_{\text{WRONG}} = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_n \end{pmatrix} R^{-1} \begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Executing this for our Alzheimer's example, we get a sample size of:

```
x = (rbind(1,t)%*%solve(R)%*%cbind(1,t))[2,2]
x*2*(qnorm(1-0.05/2) + qnorm(0.80))^2/1.5^2

[1] 0.3592744
```

which is clearly wrong. In fact, there is a typo in Diggle *et al.* (1994). The correct formula for ξ is:

$$\xi = \begin{pmatrix} 0 & 1 \end{pmatrix} \left[\begin{pmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_2 \end{pmatrix} (\sigma^2 R)^{-1} \begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix} \right]^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1}$$

Applying the correct formula, we get

```
x = solve(rbind(1,t)%*%solve(R)%*%cbind(1,t))[2,2]
x*2*(qnorm(1-0.05/2) + qnorm(0.80))^2/1.5^2

[1] 207.3101
```

Similarly, using Liu and Liang (1997), we attempt to derive the correct closed form formula for this specific linear model. The required sample size per group is given as

$$m = \nu/(\psi_1' \tilde{\Sigma}_1 \psi_1)$$

where

$$\tilde{\Sigma}_1 = \sigma^{-2} \sum_{l=1}^{m} \pi_l (\mathbf{u}_l' - I_{\psi\lambda} I_{\lambda\lambda}^{-} 1 \mathbf{v}_l') R^{-1} (\mathbf{u}_l' - \mathbf{v}_l I_{\lambda\lambda}^{-} 1 I_{\psi\lambda}'),$$

$$I_{\psi\lambda} = \sigma^{-2} \sum_{i=1}^{m} \pi_l \mathbf{u}_l' R^{-1} \mathbf{v}_l,$$

and

$$I_{\lambda\lambda} = \sigma^{-2} \sum_{i=1}^{m} \pi_l \mathbf{v}_l' R^{-1} \mathbf{v}_l.$$

Again, in our case the probability of each of the two covariate values is $\pi_1 = \pi_2 = 1/2$; and $\mathbf{u}_1 = (t_1, \ldots, t_n)'$, $\mathbf{v}_1 = [\mathbf{1}\,\mathbf{0}\,\mathbf{x}_i]_{n\times 3}$, $\mathbf{u}_2 = \mathbf{0}$, and $\mathbf{v}_2 = [\mathbf{1}\,\mathbf{0}\,\mathbf{x}_i]_{n\times 3}$. We have

$$\begin{aligned} I_{\psi\lambda} &= \sigma^{-2}/2\mathbf{u}_1' R^{-1}\mathbf{v}_1 \\ I_{\lambda\lambda} &= \sigma^{-2}/2[\mathbf{v}_1' R^{-1}\mathbf{v}_1 + \mathbf{v}_2' R^{-1}\mathbf{v}_2] = 1/2X] \\ I_{\psi\lambda} I_{\lambda\lambda}^{-1} &= \mathbf{u}_1' R^{-1}\mathbf{v}_1 X^{-1} \\ I_{\lambda\lambda}^{-1} I_{\psi\lambda}' &= X^{-1}\mathbf{v}_1' R^{-1}\mathbf{u}_1 \end{aligned}$$

$$\begin{aligned} \tilde{\Sigma}_1 &= \sigma^{-2}/2[(\mathbf{u}_1 - \mathbf{u}_1' R^{-1}\mathbf{v}_1 X^{-1}\mathbf{v}_1') R^{-1}(\mathbf{u}_1 - \mathbf{v}_1 X^{-1}\mathbf{v}_1' R^{-1}\mathbf{u}_1) \\ &\quad + \mathbf{u}_1' R^{-1}\mathbf{v}_1 X^{-1}\mathbf{v}_2' R^{-1}\mathbf{v}_2 X^{-1}\mathbf{v}_1 R^{-1}\mathbf{u}_1 \\ &= \sigma^{-2}/2[\mathbf{u}_1 R^{-1}\mathbf{u} - \mathbf{u}_1' R^{-1}\mathbf{v}_1 X^{-1}\mathbf{v}_1' R^{-1}\mathbf{u}_1] \end{aligned}$$

Applying this to our working example:

```
X = t(v[[1]])%*%solve(R)%*%v[[1]] +
    t(v[[2]])%*%solve(R)%*%v[[2]]

Sigma1 = ((t(u[[1]])%*%solve(R)%*%t -
           t(u[[1]])%*%solve(R)%*%v[[1]]%*%solve(X)%*%t(v[[1]])%*%solve(R)%*%t)/2)

(qnorm(1-0.05/2) + qnorm(0.80))^2/(Sigma1*(1.5)^2)/2

         [,1]
[1,] 207.3101
```

References

Diggle P, Liang KY, Zeger SL (1994). *Analysis of longitudinal data*. Clarendon Press ; Oxford University Press, Oxford; New York. ISBN 0198522843 : 9780198522843. ID: 29877424.

Laird NM, Ware JH (1982). "Random-effects models for longitudinal data." *Biometrics*, **38**(4), 963–74. ID: 113328916.

Liu G, Liang KY (1997). "Sample size calculations for studies with correlated observations." *Biometrics*, **53**(3), 937–47. ID: 119956120.

Zeger SL, Liang KY (1986). "Longitudinal data analysis for discrete and continuous outcomes." *Biometrics*, **42**(1), 121–30. ID: 115936579.