pass.lme: technical note

Prepared by Marco, Chak Yan, YU Date: 2023-06-22

Table of Content

Statistical Theory

Linear mixed effect model

Denote random vector *Y* as the interested measure following a **L**inear **M**ixed **E**ffect (**LME**) model with fixed covariates *X* and random covariates *Z* as follow:

$$Y = XB + Zu + e,\ u \sim N(0, D),\ e \sim N(0, R)$$

where *B* represents the fixed effect parameters, *u* represents the random effect following a normal distribution with mean 0, and *e* represents the random residual errors.

Let Y_i represents the random response *Y* for subject *i* with

$$Y_i \sim N(X_iB, V_i),\ V_i = Z_iDZ_i' + R_i$$

The **B**est **L**inear **U**nbiased **E**stimator (**BLUE**) of *B* is given by

$$b = \left(\sum X_i' V_i^{-1} X_i\right)^{-1} \left(\sum X_i' V_i^{-1} y_i\right) \text{ with } var(b) = \left(\sum X_i' V_i^{-1} X_i\right)^{-1} = \Sigma_{b_i}/n$$

if *V* is known.
We have $b \sim N\left(B, \Sigma_{b_i}/n\right)$.

(Please refer to [1] and [2] for details.)

Test for 1 contrast of fixed effect coefficients

Given $b \sim N\left(B, \Sigma_{b_i}/n\right)$,
we have

$$Lb \sim N\left(LB, L\Sigma_{b_i}L'/n\right).$$

For testing $H_0: LB = 0$,

Set *L* as a row vector, with all elements equal 0 except the element corresponding to the interested fixed effect coefficient as 1,

power, $1 - \beta = 1 - \Phi\left(z_{1-\alpha/2} - |LB| / \sqrt{L\Sigma_{b_i}L'/n}\right)$, and

sample size $n = \left(z_{1-\frac{\alpha}{2}} - z_\beta\right)^2 \left(L\Sigma_{b_i}L' / (LB)^2\right)$.

<u>Test for multiple contrasts of fixed effect coefficients</u>

Given $Lb \sim N\left(LB, L\Sigma_{b_i}L'/n\right)$,

we have

$$(Lb)'\left(\frac{L\Sigma_{b_i}L'}{n}\right)^{-1}(Lb) \sim \chi^2_{q,\lambda} \qquad \text{where } \lambda = (LB)'\left(\frac{L\Sigma_{b_i}L'}{n}\right)^{-1}(LB).$$

For testing $H_0: LB = 0$,

power, $1-\beta = P\left(\chi^2_{q,\lambda} > \chi^2_{q,0;\alpha}\right)$

where $\chi^2_{q,\lambda}$ is a chi-squared distribution with degree of freedom $q = rank\left(L\Sigma_{b_i}L'\right)$ and non-centrality parameter λ.

Sample size can be obtained by solving

$$1-\beta = P\left(\chi^2_{q,\lambda} > \chi^2_{q,0;\alpha}\right).$$

(Bisection method is used to solve for required sample size in the R package.)

<u>Test for multiple contrasts of fixed effect coefficients for more than 1 population</u>

Given $L_k b_k \sim N\left(L_k B_k, L_k \Sigma_{b_{k_i}} L'_k/n_k\right)$ for group *k=1,2,… and they are independent to each other*, *w*e have

$$\begin{bmatrix} L_1 b_1 \\ L_2 b_2 \\ \vdots \end{bmatrix} \sim N\left(\begin{bmatrix} L_1 B_1 \\ L_2 B_2 \\ \vdots \end{bmatrix}, \begin{bmatrix} \frac{L_1\Sigma_{b_{1_i}}L'_1}{n_1} & 0 & \cdots \\ 0 & \frac{L_2\Sigma_{b_{2_i}}L'_2}{n_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}\right).$$

Denote the above as $Lb \sim N(LB, L\Sigma_b L')$, testing for any contrasts among those population can be formulated as $H_0: CLB = d$.

We have $(CLb - d) \sim N(CLB - d, CL\Sigma_b L'C')$.

power, $1-\beta = P\left(\chi^2_{q,\lambda} > \chi^2_{q,0;\alpha}\right)$

where $q = rank(CL\Sigma_b L'C')$ and non-centrality parameter $\lambda = (CLB-d)'(CL\Sigma_b L'C')^{-1}(CLB-d)$.

Sample size can be obtained by solving

$$1-\beta = P\left(\chi^2_{q,\lambda} > \chi^2_{q,0;\alpha}\right).$$

(Bisection method is used to solve for required sample size in the R package.)

Reference list

[1] N. M. Laird and J. H. Ware, “Random-Effects Models for Longitudinal Data,” *Biometrics*, vol. 38, no. 4, p. 963, Dec. 1982.

[2] J. Jiang, *Linear and generalized linear mixed models and their applications*. New York ; London: Springer, 2007.

R coding examples

Example 1 – Sample size calculation for testing fixed-effect-coefficient-2 = 0 with significant level of 5% and power of 80% for 1-level LME model

Consider the 1-level LME model as

$$Y_i = X_iB + Z_iu_i + e_i,\ u_i \sim N(0, D),\ e_i \sim N(0, R)$$

with $X_i = Z_i = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$, where the first column represents the intercept, the second column represents the time covariate with observation at time 1, 2 and 3.

Assume the fixed effect coefficients $B = [B_1, B_2]' = [100, -0.5]'$; random effect variance $D = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$; and residual variance $R = 0.2$.

To test H_0: $B_2 = 0$, firstly we have to set the mean and calculate the variance for b_2 when we have one subject only (i.e. $Lb \sim N\left(LB, L\Sigma_{b_i}L'\right)$ such that $L = [0, 1]'$)
The following code will save the required mean and variance into 'theta'.

R code

```
B <- matrix(c(100,-0.5),2,1)
D <- matrix(c(2,1,1,2),2,2)
R <- 0.2
X <- cbind(rep(1,3),1:3)
Z <- X
theta <- lme.Lb.dist.theta(B,D,R,X,Z)
```

Note that we do not need to input L into the function as the default L input will be a column vector of 0, except the last one as 1. That means the default setting is to obtain the distribution of the last element in b.

The required sample size can be obtained by the following code:

R code

```
pass.lme.CLb.test( list(theta), alpha=0.05, power=0.8 )
```

If we have a sample and want to obtain the corresponding power, users can use the following code:

R code

```
pass.lme.CLb.test( list(theta), alpha=0.05, n=66 )
```

Example 2 – Same as Example 1 but for 3-level LME model

Consider the 3-level LME model as

$$Y_{k|j|i} = X_{k|j|i}B + Z_i u_i + Z_{j|i}u_{j|i} + Z_{k|j|i}u_{k|j|i} + e_{k|j|i},$$
$$u_i \sim N(0, D_1), \quad u_{j|i} \sim N(0, D_2), \quad u_{k|j|i} \sim N(0, D_3), \quad e_{k|j|i} \sim N(0, R)$$

with $X_{k|j|i} = Z_i = Z_{j|i} = Z_{k|j|i} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$, where the first column represents the intercept, the second column represents the time covariate with observation at time 1, 2 and 3.

Assume the fixed effect coefficients $B = [B_1, B_2]' = [100, -0.5]'$; random effect variances $D_1 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $D_2 = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, $D_3 = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix}$; residual variance $R = 0.2$; and for every i, there are 4 repeated measures of $j|i$ and 7 repeated measure of $k|(j|i)$.

To test H_0: $B_2 = 0$, firstly we have to find the mean and variance of b_2 when we have one subject only (i.e. $Lb \sim N(LB, L\Sigma_{b_i}L')$ such that $L = [0, 1]'$)
The following code will save the required mean and variance into 'theta'.

R code

```
B <- matrix(c(100,-0.5),2,1)
D <- list(matrix(c(2,1,1,2),2,2),matrix(c(3,1,1,3),2,2), matrix(c(5,1,1,5),2,2))
R <- 0.2
X <- cbind(rep(1,3),1:3)
Z <- X
m <- c(4,7)
theta <- lme.Lb.dist.theta(B,D,R,X,Z,m)
```

The required sample size can be obtained by the following code:

R code

```
pass.lme.CLb.test( list(theta), alpha=0.05, power=0.8 )
```

Example 3 – Sample size calculation for <u>comparing fixed-effect-coefficient-2 between two groups</u> with significant level of 5% and power of 80% for <u>1-level LME model</u>

Consider the 1-level LME model as

$$Y_{G,i} = X_{G,i}B_G + Z_{G,i}u_{G,i} + e_{G,i}, u_{G,i} \sim N(0, D_G), e_{G,i} \sim N(0, R_G)$$

with $X_i = Z_i = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$, where the first column represents the constant covariate and the second column represents the time covariate with observation at time 1, 2 and 3 for *G=C, T* for control and treatment group, respectively.

Assume the fixed effect coefficients

$B_C = [B_{C1}, B_{C2}]' = [100, -0.5]'$ for control group and

$B_T = [B_{T1}, B_{T2}]' = [100, -0.5 \boldsymbol{x70\%}]'$ for treatment group;

random effect variance $D_C = D_T = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$; and residual variance $R_C = R_T = 0.2$.

To test H_0: $(B_C = B_T) \Leftrightarrow (B_C - B_T = 0)$, firstly we have to set the mean and calculate the variance for b_{C2} and b_{T2} when we have one subject only (i.e. $Lb_G \sim N\left(LB_G, L\Sigma_{b_{G_i}}L'\right)$ such that $L = [0, 1]'$)

The following code will save the required mean and variance into 'theta' for b_{C2}.

R code

```
B <- matrix(c(100,-0.5),2,1)
D <- matrix(c(2,1,1,2),2,2)
R <- 0.2
X <- cbind(rep(1,3),1:3)
Z <- X
theta <- lme.Lb.dist.theta(B,D,R,X,Z)
```

And the following code will copy the mean and variance of b_{C2} and change the mean of b_{T2} as 70% of the mean of b_{C2}.

R code

```
theta2 <- theta
theta2$mu <- theta$mu *0.7
```

To calculate sample size for the test H_0: $(B_C - B_T = 0) \Leftrightarrow C[B_C \quad B_T]' = 0$, we have to define the contrast matrix $C = [1, -1]$. The required sample size can be obtained by the following code:

R code
```
C <- matrix(c(1,-1),1,2)
pass.lme.CLb.test( list(theta,theta2), C, alpha=0.05, power=0.8 )
```

If we have a sample and want to obtain the corresponding power, users can use the following code:

R code
```
C <- matrix(c(1,-1),1,2)
pass.lme.CLb.test( list(theta,theta2), C, alpha=0.05, n=1468 )
```

Example 4 – Same as Example 3 but with different sample size for the two groups

Say for example if we believe it is not likely to obtain a balanced sample, but we might be more confidence to obtain a sample in 1:2 ratio between control and treatment groups. We can add the sample size ratio into the required sample size calculation as follow:

R code
```
pass.lme.CLb.test( list(theta,theta2), C, alpha=0.05, power=0.8, n=c(1,2) )
```

If we have the two samples and want to obtain the corresponding power, users can use the following code:

R code
```
pass.lme.CLb.test( list(theta,theta2), C, alpha=0.05, n=c(1101,2202) )
```

Example 5 – Sample size calculation for <u>multiple means comparison</u> with significant level of 5% and power of 80% for <u>repeated-measures analysis of variance (ANOVA)</u>

Consider the repeated-measures ANOVA model with each subject measured twice given by

$$Y_{G,i} = X_{G,i}B_G \ + \ Z_{G,i}u_{G,i} \ + \ e_{G,i}, u_{G,i} \sim N(0, D_G), e_{G,i} \sim N(0, R_G)$$

with $X_{G,i} = Z_{G,i} = [1, 1]'$ represents the constant covariates for the overall mean and for the subject deviation with 2 measures for $G=1, 2, 3$.

Assume the mean for the three groups are $B_1 = 100$, $B_2 = 99$, $B_3 = 102$; with within-subject variance $D_1 = D_2 = D_3 = 15$ and residual variance $R_1 = R_2 = R_3 = 10$.

To test H_0: $(B_1 = B_2 = B_3)$ is the same as H_0: $\{(B_1 - B_2 = 0)$ and $(B_1 - B_3 = 0)\}$.

We first set the mean and calculate the variance for b_1, b_2 and b_3 through the following R code:

R code

```
B <- 100
D <- 15
R <- 10
X <- matrix(1,2,1)
Z <- X
theta <- lme.Lb.dist.theta(B,D,R,X,Z)
theta2 <- theta
theta3 <- theta
theta2$mu <- 99
theta3$mu <- 102
```

To calculate sample size for the test H_0: $\{(B_1 - B_2 = 0)$ and $(B_1 - B_3 = 0)\}$ ⇔ $C[B_1 \quad B_2 \quad B_3]' = 0$, we have to define the contrast matrix $C = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$. The required sample size can be obtained by the following code:

R code

```
C <- rbind(c(1,-1,0),c(1,0,-1))
pass.lme.CLb.test( list(theta,theta2,theta3), C, alpha=0.05, power=0.8 )
```

If we have the three samples and want to obtain the corresponding power, users can use the following code:

R code

```
C <- rbind(c(1,-1,0),c(1,0,-1))
pass.lme.CLb.test( list(theta,theta2,theta3), C, alpha=0.05, n=41)
```