VII. Multilevel Data Analysis

BIOS719 Generalized Linear Models
Department of Biostatistics and Bioinformatics
Duke University

Instructor: Hwanhee Hong, PhD

Attachment: `BIOS719_Note7_Suppl.pdf` (SAS code glimmix)

Contents

Contents

Repeated measurements

- So far, we considered cases where the outcomes Y_i, $i = 1 \ldots, n$ are assumed to be independent. However, this assumption is not plausible when the outcomes are **repeated** measurements on the **same subjects**.
- Repeated measurements over **time** on the same subjects are called **longitudinal data**.
  - The weights of the same people when they are 30, 40, 50, and 60 years old
- Repeated measurements can be obtained within **clusters**.
  - The weights of samples of women aged 40 years selected from specific locations

- We expect repeated measurements to be correlated (within-subject or within-cluster) in these examples.
- Suppose that we want to estimate the mean difference in weights between successive measurements in the first example and between two areas in the second example. If we ignore within-subject or within-cluster correlation, the standard deviation of the mean difference in weights will be estimated incorrectly.

Multilevel data

- An alternative approach to view repeated measures is to consider a hierarchical structure of the study design. This is called **hierarchical or multilevel modelling**.

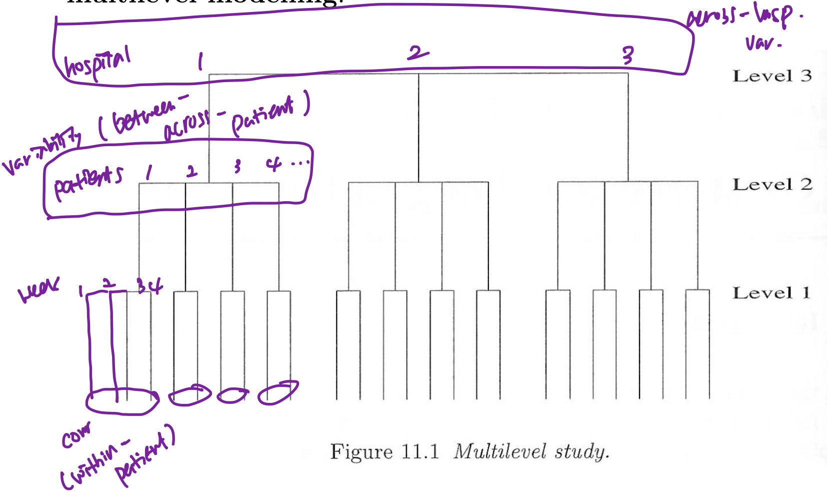


Figure 11.1 *Multilevel study.*

Stroke data example

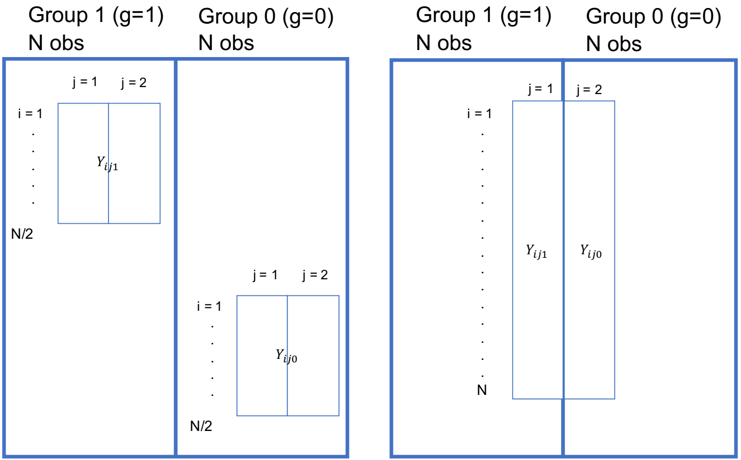
Y　　　　　　　　　　　　pop.

Table 11.1 *Functional ability scores measuring recovery from stroke for patients in three experimental groups over 8 weeks of the study.*

Subject	Group	Week 1	2	3	4	5	6	7	8
1	A	45	45	45	45	80	80	80	90
2	A	20	25	25	25	30	35	30	50
3	A	50	50	55	70	70	75	90	90
4	A	25	25	35	40	60	60	70	80
5	A	100	100	100	100	100	100	100	100
6	A	20	20	30	50	50	60	85	95
7	A	30	35	35	40	50	60	75	85
8	A	30	35	45	50	55	65	65	70
9	B	40	55	60	70	80	85	90	90
10	B	65	65	70	70	80	80	80	80
11	B	30	30	40	45	65	85	85	85
12	B	25	35	35	35	40	45	45	45
13	B	45	45	80	80	80	80	80	80
14	B	15	15	10	10	10	20	20	20
15	B	35	35	35	45	45	45	50	50
16	B	40	40	40	55	55	55	60	65
17	C	20	20	30	30	30	30	30	30
18	C	35	35	35	40	40	40	40	40
19	C	35	35	35	40	40	40	45	45
20	C	45	65	65	65	80	85	95	100
21	C	45	65	70	90	90	95	95	100
22	C	25	30	30	35	40	40	40	40
23	C	25	25	30	30	30	30	35	40
24	C	15	35	35	35	40	50	65	65

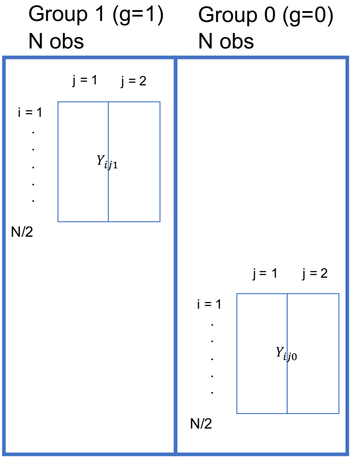
Correlated measurements: two cases

- We consider two simple cases with correlated measurements and how correlation affects estimation and inference.
- Goal: compare means in two groups



Case 1: the same covariate value within a $\boxed{\text{cluster}}$

subject.

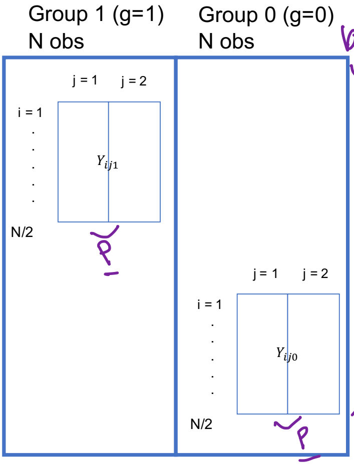


- i: subject (patient, cluster), $i = 1, \dots, N$
- j: element of a subject (repetition), $j = 1, 2$
- g: group (covariate) $g = 0, 1$
- 2 observations per subject (cluster)
- Y_{ijg}
- Y_{ij}, x_{ij}=group indicator

$x_{ij} = 1$ for $i = 1, \dots, \frac{N}{2}$

$x_{ij} = 0$ for $i = \frac{N}{2}+1, \dots, N$

Case 1: the same covariate value within a cluster



between-subject variability

- We want to obtain

$$\text{t-test stat} = \frac{\overline{Y}_1 - \overline{Y}_0}{SE(\overline{Y}_1 - \overline{Y}_0)}$$

- We denote and assume

$$\underline{Var(Y_{ijg}) = \sigma_e^2 \text{ for } g = 0, 1}$$

$$\underline{Corr(Y_{i1g}, Y_{i2g})} = \rho \text{ for } g = 0, 1$$

- We can write

$$\overline{Y}_1 = \frac{1}{N} \sum_{i=1}^{N/2} \sum_{j=1}^{2} Y_{ij1}$$

$$\overline{Y}_0 = \frac{1}{N} \sum_{i=1}^{N/2} \sum_{j=1}^{2} Y_{ij0}$$

Now, let's calculate $Var(\overline{Y}_1)$.

$$
\begin{aligned}
Var(\overline{Y}_1) &= Var\left(\frac{1}{N}\sum_{i=1}^{N/2}\sum_{j=1}^{2} Y_{ij1}\right) = \frac{1}{N^2}\sum_{i=1}^{N/2} Var(Y_{i11} + Y_{i21}) \\
&= \frac{1}{N^2}\sum_{i=1}^{N/2}\left\{ V(Y_{i11}) + Var(Y_{i21}) + 2\,Cov(Y_{i11}, Y_{i21}) \right\} \\
&= \frac{1}{N^2}\sum_{i=1}^{N/2}\left\{ \sigma_e^2 + \sigma_e^2 + 2\cdot\rho\cdot\sigma_e^2 \right\} \\
&= \frac{1}{N^2}\sum_{i=1}^{N/2} 2\cdot\sigma_e^2(1+\rho) \\
&= \frac{1}{N^{\cancel{2}}}\cdot\frac{\cancel{N}}{\cancel{2}}\cdot\cancel{2}\cdot\sigma_e^2\cdot(1+\rho) = \frac{\sigma_e^2}{N}(1+\rho)
\end{aligned}
$$

Then, similarly $Var(\overline{Y}_0) = \dfrac{\sigma_e^2}{N}(1+\rho)$

Case 1: the same covariate value within a cluster

Now, let's calculate $Var(\overline{Y}_1 - \overline{Y}_0)$.

$Var(\overline{Y}_1) = \frac{\sigma_e^2}{N}(1+\rho)$ and $Var(\overline{Y}_0) = \frac{\sigma_e^2}{N}(1+\rho)$

$$
\begin{aligned}
Var(\overline{Y}_1 - \overline{Y}_0) &= Var(\overline{Y}_1) + Var(\overline{Y}_0) \\
&= \frac{\sigma_e^2}{N}(1+\rho) + \frac{\sigma_e^2}{N}(1+\rho) \\
&= \sigma_e^2\left(\frac{1}{N}+\frac{1}{N}\right)(1+\rho)
\end{aligned}
$$

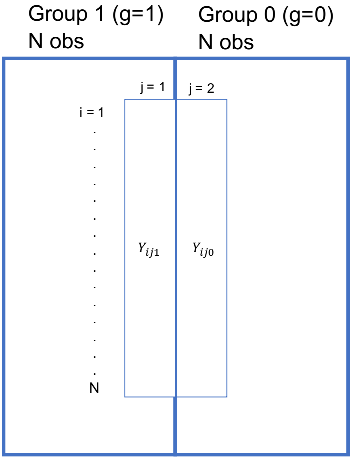
Then, $\frac{\overline{Y}_1 - \overline{Y}_0}{SE(\overline{Y}_1 - \overline{Y}_0)} = \frac{\overline{Y}_1 - \overline{Y}_0}{\sigma_e\sqrt{\left(\frac{1}{N}+\frac{1}{N}\right)(1+\rho)}} = \underbrace{\frac{\overline{Y}_1 - \overline{Y}_0}{\sigma_e\sqrt{\frac{1}{N}+\frac{1}{N}}}} \cdot \frac{1}{\sqrt{1+\rho}}$

$\rho = 0$

$\rho > 0$ $\quad \frac{1}{\sqrt{1+\rho}} < 1$ $\quad$ t-test stat $\downarrow$ $\quad$ p-val $\uparrow$

$\rho < 0$ $\quad \frac{1}{\sqrt{1+\rho}} > 1$ $\quad$ " $\uparrow$ $\quad$ " $\downarrow$

Case 2: covariate value changes within a cluster

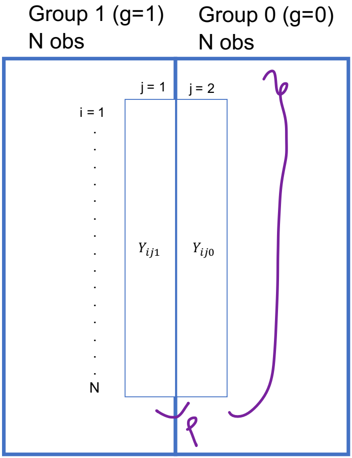


- i: subject (patient, cluster), $i = 1, \ldots, N$
- j: element of a subject (repetition), $j = 1, 2$
- g: group (covariate) $g = 0, 1$
- 2 observations per subject (cluster)
- Y_{ijg}
- Y_{ij}, x_{ij}=group indicator

if $j=1$ then $g=1$

if $j=2$ then $g=0$

Case 2: covariate value changes within a cluster



- We want to obtain

$$\frac{\overline{Y}_1 - \overline{Y}_0}{SE(\overline{Y}_1 - \overline{Y}_0)}$$

- We denote and assume

$$Var(Y_{i11}) = Var(Y_{i20}) = \sigma_e^2$$

$$Corr(Y_{i11}, Y_{i20}) = \rho$$

- We can write

$$\overline{Y}_1 = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{1} Y_{ij1} = \frac{1}{N}\sum_{i=1}^{N} Y_{i11}$$

$$\overline{Y}_0 = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=2}^{2} Y_{ij0} = \frac{1}{N}\sum_{i=1}^{N} Y_{i20}$$

Case 2: covariate value changes within a cluster

Now, let's calculate $Var(\overline{Y}_1 - \overline{Y}_0)$.

$$
\begin{aligned}
Var(\overline{Y}_1 - \overline{Y}_0) &= Var\left(\frac{1}{N}\sum_{i=1}^{N} Y_{i11} - \frac{1}{N}\sum_{i=1}^{N} Y_{i20}\right) = Var\left(\frac{1}{N}\sum_{i=1}^{N}[Y_{i11} - Y_{i20}]\right) \\
&= \frac{1}{N^2}\sum_{1}^{N} V(Y_{i11} - Y_{i20}) \quad \because \text{indep between subjects.} \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\left\{V(Y_{i11}) + V(Y_{i20}) - 2\,Cov(Y_{i11}, Y_{i20})\right\} \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\left(\sigma_e^2 + \sigma_e^2 - 2\cdot\rho\cdot\sigma_e^2\right) \\
&= \frac{1}{N^{\cancel{2}}}\cdot\cancel{N}\cdot 2\sigma_e^2(1-\rho) = \underbrace{\sigma_e^2\left(\frac{1}{N}+\frac{1}{N}\right)(1-\rho)}
\end{aligned}
$$

Then, $\dfrac{\overline{Y}_1 - \overline{Y}_0}{SE(\overline{Y}_1 - \overline{Y}_0)} = \dfrac{\overline{Y}_1 - \overline{Y}_0}{\sigma_e\sqrt{\frac{1}{N}+\frac{1}{N}}} \cdot \dfrac{1}{\sqrt{(1-\rho)}}$

$\rho = 0$

$\rho > 0$ test stat $\uparrow$

$\rho < 0$ " $\downarrow$

Case 1: Simulated data example CorrTwoOutcomes.R

N=2000

```
> library(mvtnorm)
> #### Simulated data with two groups; Generate 2000 observations in each group
> ## Case 1. Same covariate value within a cluster
> set.seed(10)
> sigma2 <- diag(c(2,2)) + 1
> sigma2  # Positive correlation (corr=1/3) within a cluster
     [,1] [,2]
[1,]    3    1
[2,]    1    3
> # group 1
> tmp.grp1 <- rmvnorm(n=1000, mean=c(4,4), sigma=sigma2, pre0.9_9994 = T)
> head(tmp.grp1)
         [,1]     [,2]
[1,] 4.339544 5.797976
[2,] 3.769256 4.434424
[3,] 1.729452 4.009016
[4,] 3.221050 5.246025
[5,] 4.437510 3.705614
[6,] 4.749875 4.606402
> # group 0
> tmp.grp0 <- rmvnorm(n=1000, mean=c(4.2,4.2), sigma=sigma2, pre0.9_9994 = T)
> head(tmp.grp0)
         [,1]     [,2]
[1,] 3.809323 4.895977
[2,] 5.562970 4.823251
```

ρ

N=1000

N=1000

Case 1: Simulated data example

1st measurement

2nd -

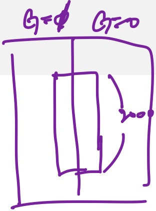
```
> # Sample mean
> apply(tmp.grp1, 2, mean)
[1] 4.023996 4.030014
> apply(tmp.grp0, 2, mean)
[1] 4.157445 4.174604
> # Sample correlation matrix
> cor(tmp.grp1)
          [,1]      [,2]
[1,] 1.0000000 0.3724683
[2,] 0.3724683 1.0000000
> cor(tmp.grp0)
          [,1]      [,2]
[1,] 1.0000000 0.3527705
[2,] 0.3527705 1.0000000
> # To estimate a common correlation
> cor(rbind(tmp.grp1, tmp.grp0))
          [,1]      [,2]
[1,] 1.0000000 0.3634918
[2,] 0.3634918 1.0000000
> rho <- cor(rbind(tmp.grp1, tmp.grp0))[1,2]
> rho
[1] 0.3634918
```

Case 1: Simulated data example

```
> # Assume all the observations are independent (incorrect assumption)
> t.test(as.vector(tmp.grp1), as.vector(tmp.grp0), var.equal=T)

Two Sample t-test

data:  as.vector(tmp.grp1) and as.vector(tmp.grp0)
t = -2.4684, df = 3998, p-value = 0.01361

> # Let's incorporate the correlation
> stat1 <- -2.4684/sqrt(1+rho)
> stat1
[1] -2.113924
> pval1 <- 2*(1-pnorm(abs(stat1)))
> pval1
[1] 0.0345218
```

Interpretation: When we account for correlation, the results become [less, more] significant (p-value=0.035). In this example, not accounting for correlation can provide smaller p-value (0.014) than warranted p-value.

Case 2: Simulated data example



```
> ## Case 2. covariate value changes within a cluster
> set.seed(10)
> sigma2 <- diag(c(2,2)) + 1
> sigma2  # Positive correlation (corr=1/3) within a cluster
     [,1] [,2]
[1,]    3    1
[2,]    1    3
> tmp.grp <- rmvnorm(n=2000, mean=c(4, 4.2), sigma=sigma2, pre0.9_9994 = T)
```

y(1) y(0)

```
> head(tmp.grp)
         [,1]     [,2]
[1,] 3.941830 3.679932
[2,] 3.907499 5.440167
[3,] 1.490912 2.818700
[4,] 2.702205 2.421971
[5,] 4.494707 4.238985
[6,] 4.353123 2.493968
> # Sample mean
> apply(tmp.grp, 2, mean)
[1] 4.018075 4.174955
> # Sample correlation matrix
> cor(tmp.grp)
          [,1]      [,2]
[1,] 1.0000000 0.3264518
[2,] 0.3264518 1.0000000
> rho <- cor(tmp.grp)[1,2]  # 0.3264518
```

Case 2: Simulated data example

```
> # Assume all the observations are independent (incorrect assumption)
> t.test(tmp.grp[,1], tmp.grp[,2], var.equal=T)

Two Sample t-test

data:  tmp.grp[, 1] and tmp.grp[, 2]
t = -2.8044, df = 3998, p-value = 0.005066

> # Correctly assume and run paired t-test
> t.test(tmp.grp[,1], tmp.grp[,2], paired=T)

Paired t-test

data:  tmp.grp[, 1] and tmp.grp[, 2]
t = -3.417, df = 1999, p-value = 0.0006458
> stat2 <- -2.8044/sqrt(1-rho)
> stat2
[1] -3.417084
```

Interpretation: When we account for correlation, the results become [less, more] significant. In this example, not accounting for correlation can provide larger p-value than warranted p-value.

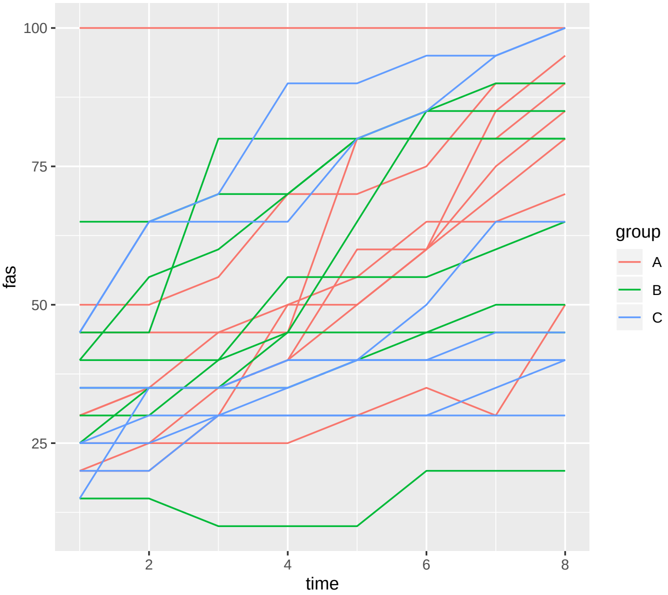
Contents

Stroke data example

Table 11.1 *Functional ability scores measuring recovery from stroke for patients in three experimental groups over 8 weeks of the study.*

		Week							
Subject	Group	1	2	3	4	5	6	7	8
1	A	45	45	45	45	80	80	80	90
2	A	20	25	25	25	30	35	30	50
3	A	50	50	55	70	70	75	90	90
4	A	25	25	35	40	60	60	70	80
5	A	100	100	100	100	100	100	100	100
6	A	20	20	30	50	50	60	85	95
7	A	30	35	35	40	50	60	75	85
8	A	30	35	45	50	55	65	65	70
9	B	40	55	60	70	80	85	90	90
10	B	65	65	70	70	80	80	80	80
11	B	30	30	40	45	65	85	85	85
12	B	25	35	35	35	40	45	45	45
13	B	45	45	80	80	80	80	80	80
14	B	15	15	10	10	10	20	20	20
15	B	35	35	35	45	45	45	50	50
16	B	40	40	40	55	55	55	60	65
17	C	20	20	30	30	30	30	30	30
18	C	35	35	35	40	40	40	40	40
19	C	35	35	35	40	40	40	45	45
20	C	45	65	65	65	80	85	95	100
21	C	45	65	70	90	90	95	95	100
22	C	25	30	30	35	40	40	40	40
23	C	25	25	30	30	30	30	35	40
24	C	15	35	35	35	40	50	65	65

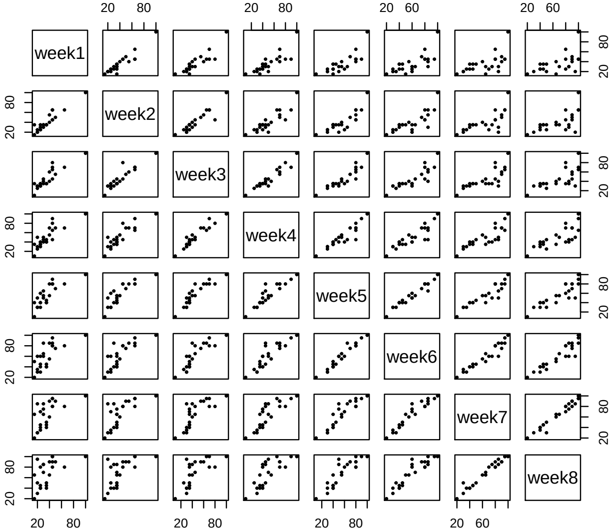
Stroke data example: spaghetti plot



Stroke data example: correlation matrix

```
> round(cor(dt.w[, paste("week",1:8,sep="")]), 2)
      week1 week2 week3 week4 week5 week6 week7 week8
week1  1.00  0.93  0.88  0.83  0.79  0.71  0.62  0.55
week2  0.93  1.00  0.92  0.88  0.85  0.79  0.70  0.64
week3  0.88  0.92  1.00  0.95  0.91  0.85  0.77  0.70
week4  0.83  0.88  0.95  1.00  0.92  0.88  0.83  0.77
week5  0.79  0.85  0.91  0.92  1.00  0.97  0.91  0.88
week6  0.71  0.79  0.85  0.88  0.97  1.00  0.96  0.93
week7  0.62  0.70  0.77  0.83  0.91  0.96  1.00  0.98
week8  0.55  0.64  0.70  0.77  0.88  0.93  0.98  1.00
```

Stroke data example: matrix plot



We will cover...

- Notations
- Correlation structure
- Random intercept model
- Random intercept and slope model

Illustration of random intercept model

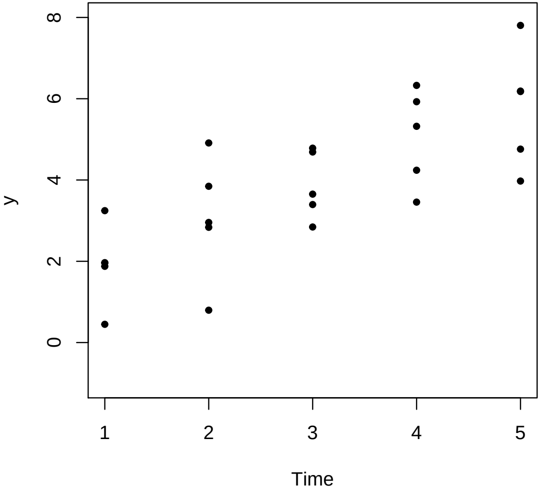


Illustration of random intercept model

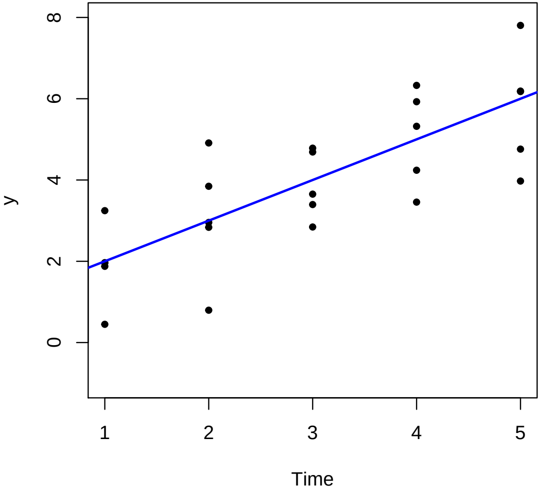


Illustration of random intercept model

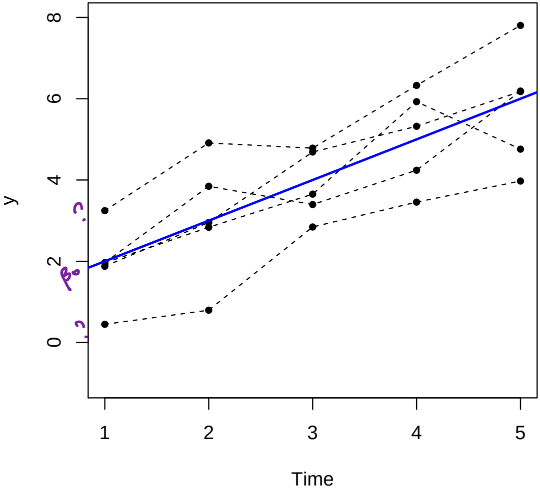


Illustration of random intercept model

i: ind
$n=4$
$\beta_{0i} \sim N(\beta_0, \tau_0^2)$

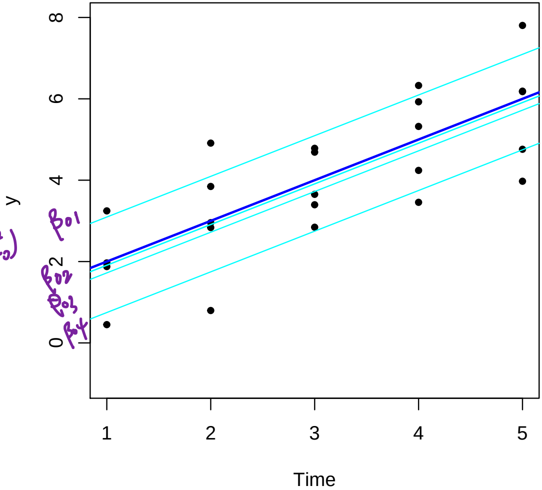


Illustration of random intercept & slope model

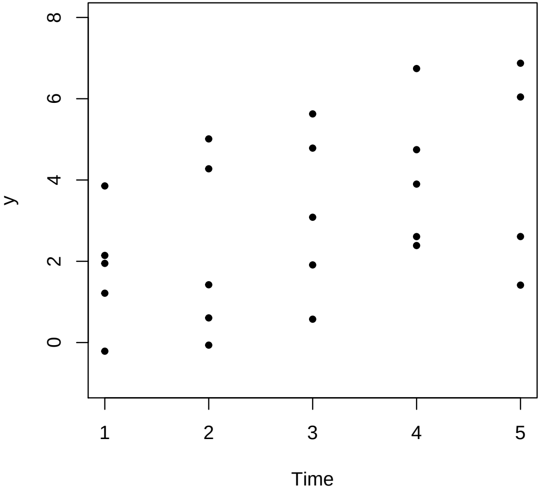


Illustration of random intercept & slope model

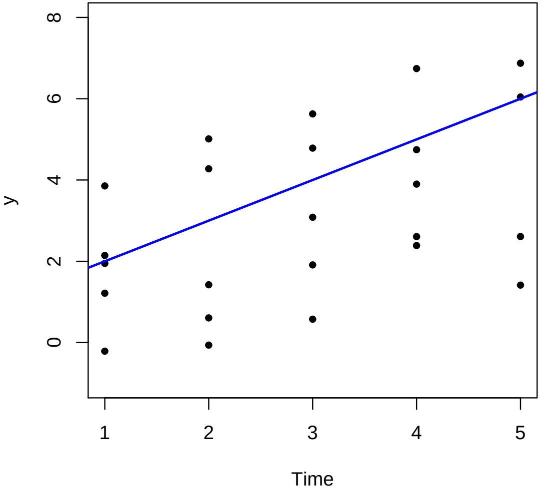


Illustration of random intercept & slope model

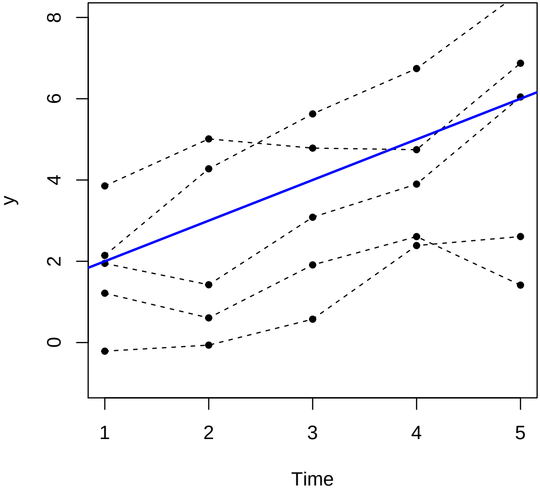
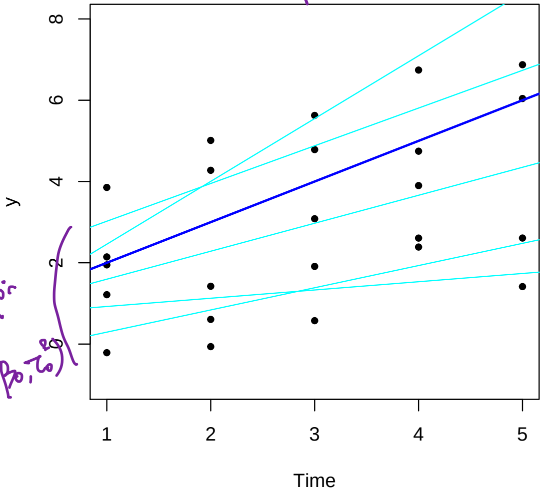


Illustration of random intercept & slope model

β_{1i}

$\beta_{1i} \sim N(\beta_1, \tau_1^2)$

$i = 1, \ldots 5$



β_{0i}

$\beta_{0i} \sim N(\beta_0, \tau_0^2)$

Notations

id	group	week	fas	time
1	A	week1	45	1
1	A	week2	45	2
1	A	week3	45	3
1	A	week4	45	4
1	A	week5	80	5
1	A	week6	80	6
1	A	week7	80	7
1	A	week8	90	8
2	A	week1	20	1
2	A	week2	25	2
2	A	week3	25	3
2	A	week4	25	4
2	A	week5	30	5
2	A	week6	35	6
2	A	week7	30	7
2	A	week8	50	8
...				
24	C	week1	15	1
24	C	week2	35	2
24	C	week3	35	3
24	C	week4	35	4
24	C	week5	40	5
24	C	week6	50	6
24	C	week7	65	7
24	C	week8	65	8

X_1 y y_1 X_2 y_2 Y_{26} y_{24}

- i: subject (unit, cluster), $i = 1, \ldots, N$
- n_i: number of obs in the i^{th} subject $n_i = 8$ for $\forall i$
- $n = \sum_i^N n_i$: total number of obs
- Y_{ij}: outcome measured at the j^{th} visit for the i^{th} subject
- $\mathbf{y}_i$, $\mathbf{X}_i$

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_N \end{pmatrix} \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}$$

$$\mathbf{y} \sim MVN(\boldsymbol{\mu}, \mathbf{V})$$

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} = \boldsymbol{\mu}$$

Variance-covariance matrix

- Assume independence **between** subjects

$n_1 \times n_1$ $\quad n_2 \times n_2$ $\quad n_N \times n_N$

$$V(\mathbf{y}) = \mathbf{V} = \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}_N \end{pmatrix}$$

- Assume dependence **within** subjects

$$V(\mathbf{y}_i) = \mathbf{V}_i = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n_i} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n_i} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n_i1} & \sigma_{n_i2} & \cdots & \sigma_{n_in_i} \end{pmatrix}$$

where $\sigma_{jj} = \sigma_j^2$. Note that $\mathbf{V}_i$ can be different for different subject i.

Working correlation matrix

$$\mathbf{V}_i = A_i^{1/2} R_i A_i^{1/2} =$$

R

$$\begin{pmatrix} \sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sigma_{n_i n_i}} \end{pmatrix} \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n_i} \\ \rho_{21} & 1 & \cdots & \rho_{2n_i} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n_i1} & \rho_{n_i2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} \sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sigma_{n_i n_i}} \end{pmatrix}$$

Independent

$$R = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Unstructured

$$R = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n_i} \\ \rho_{21} & 1 & \cdots & \rho_{2n_i} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n_i1} & \rho_{n_i2} & \cdots & 1 \end{bmatrix}$$

$$\begin{pmatrix} 1 & \rho^{|d-b|} \\ & 1 \\ & & 1 \\ & & & 1 \end{pmatrix}$$

Exchangeable (compound symmetry)

$$R = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}$$

Autoregressive (1)

$$R = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix}$$

Linear mixed effects models LMM

- For now, we consider that outcomes follow a multivariate normal distribution with the identity link.
  - Random intercept model
  - Random intercept and slope model
- For binary, count, categorical outcomes, you can use a proper link function (e.g., logit, log) and fit generalized linear mixed models (GLMM).
- Side note: multivariate regression vs. multivariable regression

univariate " vs. univariable "

Stroke data example

Table 11.1 *Functional ability scores measuring recovery from stroke for patients in three experimental groups over 8 weeks of the study.*

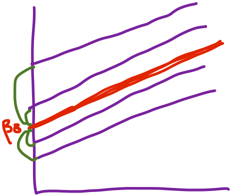
		Week							
Subject	Group	1	2	3	4	5	6	7	8
1	A	45	45	45	45	80	80	80	90
2	A	20	25	25	25	30	35	30	50
3	A	50	50	55	70	70	75	90	90
4	A	25	25	35	40	60	60	70	80
5	A	100	100	100	100	100	100	100	100
6	A	20	20	30	50	50	60	85	95
7	A	30	35	35	40	50	60	75	85
8	A	30	35	45	50	55	65	65	70
9	B	40	55	60	70	80	85	90	90
10	B	65	65	70	70	80	80	80	80
11	B	30	30	40	45	65	85	85	85
12	B	25	35	35	35	40	45	45	45
13	B	45	45	80	80	80	80	80	80
14	B	15	15	10	10	10	20	20	20
15	B	35	35	35	45	45	45	50	50
16	B	40	40	40	55	55	55	60	65
17	C	20	20	30	30	30	30	30	30
18	C	35	35	35	40	40	40	40	40
19	C	35	35	35	40	40	40	45	45
20	C	45	65	65	65	80	85	95	100
21	C	45	65	70	90	90	95	95	100
22	C	25	30	30	35	40	40	40	40
23	C	25	25	30	30	30	30	35	40
24	C	15	35	35	35	40	50	65	65

- We want to study temporal effect of outcomes (fas).
- Y_{ij}, $i = 1, \ldots, 24$ and $j = 1, \ldots 8$
- t_{ij}: time (week)

$t_{ij} = j$

Random intercept model

$$
\begin{aligned}
Y_{ij} &= \beta_0 + \beta_1 t_{ij} + a_i + e_{ij} \\
&= \underbrace{a_i + \beta_0} + \beta_1 t_{ij} + e_{ij} \quad ✓ \\
&= \underbrace{\beta_{0i}} + \beta_1 t_{ij} + e_{ij}, \quad ✓ \rightarrow \beta_{0i} \sim N(\beta_0, \sigma_a^2)
\end{aligned}
$$



- $\beta_0 + \beta_1 t_{ij}$: fixed effects
- a_i: random effects
- $a_i + \beta_0 = \beta_{0i}$: **random intercept**

We assume that

$$\underline{a_i \sim N(0, \sigma_a^2)} \text{ and } e_{ij} \sim N(0, \sigma_e^2)$$

a_i and e_{ij} are independent.

Random intercept model

$$Y_{ij} = a_i + \beta_0 + \beta_1 t_{ij} + e_{ij}, \text{ where } a_i \sim N(0, \sigma_a^2) \text{ and } e_{ij} \sim N(0, \sigma_e^2)$$

$$E(Y_{ij}) = E[\beta_0 + \beta_1 t_{ij} + a_i + e_{ij}]$$

$$= \beta_0 + \beta_1 t_{ij}$$

$$Var(Y_{ij}) = E\left[(Y_{ij} - E[Y_{ij}])^2\right] = E\left((a_i + e_{ij})^2\right)$$

$$= E[a_i^2 + e_{ij}^2 + 2a_i e_{ij}]$$

$$= E(a_i^2) + E(e_{ij}^2) + 2\underbrace{\cancel{E(a_i \cdot e_{ij})}}_{0}$$

$$= \sigma_a^2 + \sigma_e^2$$

$$\because E(a_i \cdot e_{ij}) = Cov(a_i, e_{ij}) = 0$$

Random intercept model

$$Y_{ij} = a_i + \beta_0 + \beta_1 t_{ij} + e_{ij}, \text{ where } a_i \sim N(0, \sigma_a^2) \text{ and } e_{ij} \sim N(0, \sigma_e^2)$$

$$Cov(Y_{ik}, Y_{im}) = E[(Y_{ik} - E[Y_{ik}])(Y_{im} - E[Y_{im}])]$$

between k and m
time points
within the i^th patient

$$= E[(a_i + e_{ik})(a_i + e_{im})]$$

$$= E[a_i^2 + \cancel{a_i e_{im}}^{\,0} + \cancel{a_i e_{ik}}^{\,0} + \cancel{e_{ik} \cdot e_{im}}^{\,0}]$$

$E(e_{ik} \cdot e_{im}) = 0$ $\because e_{ij}$ are iid

$$= E[a_i^2] = \sigma_a^2$$

$$Cov(Y_{ik}, Y_{\ell m}) = E[(Y_{ik} - E[Y_{ik}])(Y_{\ell m} - E[Y_{\ell m}])]$$

between k and m
time points
between two patients
i and l

$$= E[(a_i + e_{ik})(a_\ell + e_{\ell m})]$$

$$= E[\cancel{a_i a_\ell}^{\,0} + \cancel{a_i e_{\ell m}}^{\,0} + \cancel{e_{ik} a_\ell}^{\,0} + \cancel{e_{ik} e_{\ell m}}^{\,0}]$$

$$= 0$$

Random intercept model - summary

Var between subjects

Var across times within a subject

$$Y_{ij} = a_i + \beta_0 + \beta_1 t_{ij} + e_{ij}, \text{ where } a_i \sim N(0, \sigma_a^2) \text{ and } e_{ij} \sim N(0, \sigma_e^2)$$

$$Var(Y_{ij}) = \sigma_a^2 + \sigma_e^2$$
$$Cov(Y_{ik}, Y_{im}) = \sigma_a^2$$
$$Corr(Y_{ik}, Y_{im}) = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2} = \rho$$

: intra-class correlation / within-subject correlation

$\rho \to 1$ then $\sigma_e^2 \to 0$

$\sigma_e^2 \ll \sigma_a^2$

response within a subject is much more alike than between subjects

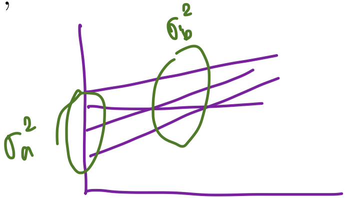
$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}_N \end{pmatrix}$$

$$\mathbf{V}_i = \begin{pmatrix} \sigma_a^2+\sigma_e^2 & & \sigma_a^2 \\ & \ddots & \\ \sigma_a^2 & & \sigma_a^2+\sigma_e^2 \end{pmatrix} = (\sigma_a^2+\sigma_e^2) \begin{pmatrix} 1 & & \rho \\ & 1 & \\ \rho & & \ddots \ 1 \end{pmatrix}$$

: exchangeable / compound symmetry

Random intercept and slope model

$$
\begin{aligned}
Y_{ij} &= \beta_0 + \beta_1 t_{ij} + a_i + b_i t_{ij} + e_{ij} \\
&= a_i + \beta_0 + (b_i + \beta_1) t_{ij} + e_{ij} \\
&= \beta_{0i} + \beta_{1i} t_{ij} + e_{ij},
\end{aligned}
$$



- $\beta_0 + \beta_1 t_{ij}$: fixed effects
- a_i, b_i: random effects
- $a_i + \beta_0 = \beta_{0i}$: **random intercept**
- $b_i + \beta_1 = \beta_{1i}$: **random slope**

We assume that

$$a_i \sim N(0, \sigma_a^2), b_i \sim N(0, \sigma_b^2) \text{ and } e_{ij} \sim N(0, \sigma_e^2)$$

Assume e_{ij} is independent from a_i and b_i.
Assume a_i and b_i are independent (potentially can be correlated)

Random intercept and slope model - summary

$$Y_{ij} = a_i + \beta_0 + (b_i + \beta_1)t_{ij} + e_{ij},$$

where $a_i \sim N(0, \sigma_a^2)$, $b_i \sim N(0, \sigma_b^2)$ and $e_{ij} \sim N(0, \sigma_e^2)$

$$Var(Y_{ij}) = Var(a_i + b_i t_{ij} + e_{ij}) = \sigma_a^2 + t_{ij}^2 \sigma_b^2 + \sigma_e^2$$

$$\begin{aligned} Cov(Y_{ik}, Y_{im}) &= E[(Y_{ik} - E[Y_{ik}])(Y_{im} - E[Y_{im}])] \\ &= E[(a_i + b_i t_{ik} + e_{ik})(a_i + b_i t_{im} + e_{im})] \\ &= E(a_i^2) + t_{ik} t_{im} E(b_i^2) \\ &= \sigma_a^2 + t_{ik} t_{im} \sigma_b^2 \end{aligned}$$

unstructured

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}_N \end{pmatrix}$$

Linear mixed effects model

- $Y_{ij} = \beta_0 + \beta_1 t_{ij} + a_i + e_{ij}$
- $Y_{ij} = \beta_0 + \beta_1 t_{ij} + a_i + b_i t_{ij} + e_{ij}$
- $\mathbf{y} = \underbrace{\mathbf{X}\boldsymbol{\beta}}_{\text{fixed effts}} + \underbrace{\mathbf{Z}\mathbf{u}}_{\text{random effects}} + \mathbf{e}$

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_N \end{bmatrix} \quad \boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \vdots \\ \boldsymbol{X}_N \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \quad \boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Z}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{Z}_N \end{bmatrix} \quad \boldsymbol{u} = \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_N \end{bmatrix} \quad \boldsymbol{e} = \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_N \end{bmatrix}$$

$$\boldsymbol{y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{in_i} \end{bmatrix} \quad \boldsymbol{u}_i = \begin{bmatrix} u_{i1} \\ u_{i2} \\ \vdots \\ u_{iq} \end{bmatrix} \quad \boldsymbol{e}_i = \begin{bmatrix} e_{i1} \\ e_{i2} \\ \vdots \\ e_{in_i} \end{bmatrix}$$

$\boldsymbol{X}_i : n_i \times p$

p: # of fixed parameters

$\boldsymbol{Z}_i : n_i \times q$

q: # of random parameters

Linear mixed effects model: define design matrices

$i = 1, \cdots, N$

$j = 1, \cdots, n_i$

$\sum_{i=1}^{N} n_i = n$

1. • $Y_{ij} = \beta_0 + \beta_1 t_{ij} + a_i + e_{ij}$
2. • $Y_{ij} = \beta_0 + \beta_1 t_{ij} + a_i + b_i t_{ij} + e_{ij}$

• $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$

$$X_i = \begin{bmatrix} 1 & t_{i1} \\ \vdots & \vdots \\ 1 & t_{in_i} \end{bmatrix} \qquad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

1

$$Z_i = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{n_i \times q} \qquad u_i = [a_i]_{q \times 1}$$

$$Z = \begin{bmatrix} Z_1 & & & 0 \\ & Z_2 & & \\ & & \ddots & \\ 0 & & & Z_N \end{bmatrix}_{n_i \cdot N \times n_i \cdot q}$$

2

$$Z_i = \begin{bmatrix} 1 & t_{i1} \\ 1 & \vdots \\ \vdots & \vdots \\ 1 & t_{in_i} \end{bmatrix}_{n_i \times q} \qquad u_i = \begin{bmatrix} a_i \\ b_i \end{bmatrix}_{q \times 1}$$

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}_{n_i \cdot q \times 1}$$

Linear mixed effects model

- $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$

$Var(\mathbf{y}) = Var(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}) = Var(\mathbf{Z}\mathbf{u} + \mathbf{e})$

Assumed that $\mathbf{u}$ and $\mathbf{e}$ are independent then

$$= Var(\mathbf{Z}\mathbf{u}) + Var(\mathbf{e})$$

$$\rightarrow \quad = \mathbf{Z}Var(\mathbf{u})\mathbf{Z}^T + Var(\mathbf{e})$$

$$\begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim N\left(0, \underbrace{\begin{pmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix}}_{=G}\right)$$

- For subject (cluster) level,
  $\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i + \mathbf{e}_i$, where $\mathbf{u}_i \sim N(0, \mathbf{G})$ and $\mathbf{e}_i \sim N(0, \mathbf{R})$

Then,

$Var(\mathbf{y}_i) = \mathbf{V}_i = \mathbf{Z}_i Var(\mathbf{u}_i)\mathbf{Z}_i^T + Var(\mathbf{e}_i) = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^T + \mathbf{R}$

- For <u>generalized</u> linear mixed effects model (GLMM)

$$E(\mathbf{y}|\mathbf{u}) = \boldsymbol{\mu} \text{ and } g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$$

Contents

Stroke data example: three models

- Model 1: $Y_{ij} = \beta_0 + a_i + \beta_1 t_{ij} + e_{ij}$, where
  - $a_i \sim N(0, \sigma_a^2)$
  - $e_{ij} \sim N(0, \sigma_e^2)$
  - a_i and e_{ij} are independent.
- Model 2: $Y_{ij} = \beta_0 + a_i + (\beta_1 + b_i)t_{ij} + e_{ij}$, where
  - $a_i \sim N(0, \sigma_a^2)$, $b_i \sim N(0, \sigma_b^2)$, a_i and b_i are independent.
  - $e_{ij} \sim N(0, \sigma_e^2)$
  - (a_i, b_i) and e_{ij} are independent.
- Model 3: $Y_{ij} = \beta_0 + a_i + (\beta_1 + b_i)t_{ij} + e_{ij}$, where
  - $a_i \sim N(0, \sigma_a^2)$, $b_i \sim N(0, \sigma_b^2)$, a_i and b_i are dependent.
  - $e_{ij} \sim N(0, \sigma_e^2)$
  - (a_i, b_i) and e_{ij} are independent.

Mixed effects models in R and SAS

R package "nlme"

- I will show R and SAS code to fit the three mixed linear models with the stroke data using R package `lme4` and SAS `glimmix`.
- More about `lme4`. Page 30 will be useful when extracting values from results
  - https://cran.r-project.org/web/packages/lme4/vignettes/lmer.pdf
- Our data example is for mixed linear model. For more information about generalized linear mixed model (e.g., for binary or count outcomes) with link functions, please check
  - https://stats.idre.ucla.edu/other/mult-pkg/introduction-to-generalized-linear-mixed-models/
  - https://stats.idre.ucla.edu/r/dae/mixed-effects-logistic-regression/
- There are many other R packages that you can use to fit mixed effects models. For comprehensive lists, reviews, and examples, please check out http://bbolker.github.io/mixedmodels-misc/glmmFAQ.html

Wide data vs. long data

wide

Table 11.1 *Functional ability scores measuring recovery from stroke for patients in three experimental groups over 8 weeks of the study.*

		Week							
Subject	Group	1	2	3	4	5	6	7	8
1	A	45	45	45	45	80	80	80	90
2	A	20	25	25	25	30	35	30	50
3	A	50	50	55	70	70	75	90	90
4	A	25	25	35	40	60	60	70	80
5	A	100	100	100	100	100	100	100	100
6	A	20	20	30	50	50	60	85	95
7	A	30	35	35	40	50	60	75	85
8	A	30	35	45	50	55	65	65	70
9	B	40	55	60	70	80	85	90	90
10	B	65	65	70	70	80	80	80	80
11	B	30	30	40	45	65	85	85	85
12	B	25	35	35	35	40	45	45	45
13	B	45	45	80	80	80	80	80	80
14	B	15	15	10	10	10	20	20	20
15	B	35	35	35	45	45	45	50	50
16	B	40	40	40	55	55	55	60	65
17	C	20	20	30	30	30	30	30	30
18	C	35	35	35	40	40	40	40	40
19	C	35	35	35	40	40	40	45	45
20	C	45	65	65	65	80	85	95	100
21	C	45	65	70	90	90	95	95	100
22	C	25	30	30	35	40	40	40	40
23	C	25	25	30	30	30	30	35	40
24	C	15	35	35	35	40	50	65	65

long

id	group	week	fas	time
1	A	week1	45	1
1	A	week2	45	2
1	A	week3	45	3
1	A	week4	45	4
1	A	week5	80	5
1	A	week6	80	6
1	A	week7	80	7
1	A	week8	90	8
2	A	week1	20	1
2	A	week2	25	2
2	A	week3	25	3
2	A	week4	25	4
2	A	week5	30	5
2	A	week6	35	6
2	A	week7	30	7
2	A	week8	50	8
...				
24	C	week1	15	1
24	C	week2	35	2
24	C	week3	35	3
24	C	week4	35	4
24	C	week5	40	5
24	C	week6	50	6
24	C	week7	65	7
24	C	week8	65	8

Wide data vs. long data (R code)

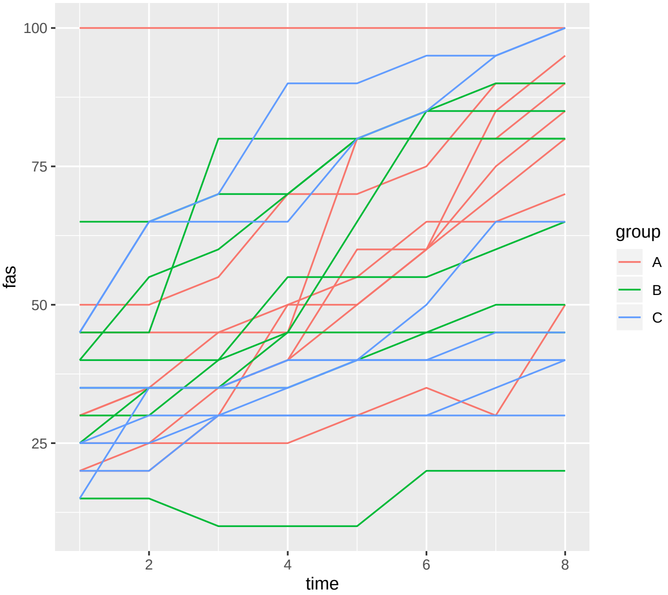
```
> #### Wide vs. long data
> dt.w <- read.csv("table11_1.csv")
> head(dt.w)
  id group week1 week2 week3 week4 week5 week6 week7 week8
1  1     A    45    45    45    45    80    80    80    90
2  2     A    20    25    25    25    30    35    30    50
3  3     A    50    50    55    70    70    75    90    90
4  4     A    25    25    35    40    60    60    70    80
5  5     A   100   100   100   100   100   100   100   100
6  6     A    20    20    30    50    50    60    85    95
> dt.l <- melt(dt.w, id.vars=c("id", "group"),
+              measure.vars=paste("week",1:8,sep=""),
+              variable.name="week",
+              value.name="fas")
> dt.l$time <- recode(dt.l$week, "week1"=1, "week2"=2, "week3"=3, "week4"=4,
+                     "week5"=5, "week6"=6, "week7"=7, "week8"=8)
> dt.l <- dt.l[order(dt.l$id),]
> head(dt.l)
                       Y
    id group  week fas time
1    1     A week1  45    1
25   1     A week2  45    2
49   1     A week3  45    3
73   1     A week4  45    4
97   1     A week5  80    5
121  1     A week6  80    6
```

Some figures (R code)

```
#### Spaghetti plot
p <- ggplot(data=dt.l, aes(x=time, y=fas, group=id, colour=group))
p + geom_line()

#### Estimated correlations
round(cor(dt.w[, paste("week",1:8,sep="")]), 2)

#### Correlation matrix plot
pairs(dt.w[,paste("week",1:8,sep="")], pch=19, cex=0.4)
```

- There are so many cool graphs to visualize your data!
- https://www.r-graph-gallery.com/
- https://www.data-to-viz.com/
- https://datavizproject.com/
- http://visualizationuniverse.com/

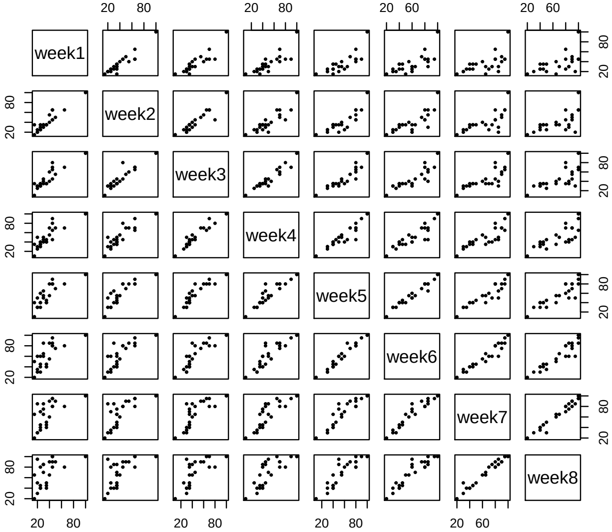
Stroke data example: spaghetti plot



Stroke data example: correlation matrix

```
> round(cor(dt.w[, paste("week",1:8,sep="")]), 2)
      week1 week2 week3 week4 week5 week6 week7 week8
week1  1.00  0.93  0.88  0.83  0.79  0.71  0.62  0.55
week2  0.93  1.00  0.92  0.88  0.85  0.79  0.70  0.64
week3  0.88  0.92  1.00  0.95  0.91  0.85  0.77  0.70
week4  0.83  0.88  0.95  1.00  0.92  0.88  0.83  0.77
week5  0.79  0.85  0.91  0.92  1.00  0.97  0.91  0.88
week6  0.71  0.79  0.85  0.88  0.97  1.00  0.96  0.93
week7  0.62  0.70  0.77  0.83  0.91  0.96  1.00  0.98
week8  0.55  0.64  0.70  0.77  0.88  0.93  0.98  1.00
```

Stroke data example: matrix plot



Model 1 (R code)

- Model 1: $Y_{ij} = \beta_0 + a_i + \beta_1 t_{ij} + e_{ij}$, where
  - $a_i \sim N(0, \sigma_a^2)$
  - $e_{ij} \sim N(0, \sigma_e^2)$
  - a_i and e_{ij} are independent.

```
> library(lme4)
> ## Model 1: Random intercept model
> fit1 <- lmer(fas ~ time + (1 | id), data=dt.l)
                          FE        RE
> summary(fit1)
Linear mixed model fit by REML ['lmerMod']
Formula: fas ~ time + (1 | id)
   Data: dt.l

Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept) 393.8    19.844     σa²
 Residual              80.3     8.961     σe²
Number of obs: 192, groups:  id, 24
Fixed effects:
            Estimate Std. Error t value
(Intercept)  30.9301     4.2941   7.203   β0
time          4.7644     0.2822  16.881   β1
```

Model 1 (R code)

```
> Zi <- cbind(rep(1,8))
> G <- matrix(393.8,1,1)
> sigmasq.e <- 80.3
> Vi <- Zi%*%G%*%t(Zi) + sigmasq.e * diag(8)
> Vi
      [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]
[1,] 474.1 393.8 393.8 393.8 393.8 393.8 393.8 393.8
[2,] 393.8 474.1 393.8 393.8 393.8 393.8 393.8 393.8
[3,] 393.8 393.8 474.1 393.8 393.8 393.8 393.8 393.8
[4,] 393.8 393.8 393.8 474.1 393.8 393.8 393.8 393.8
[5,] 393.8 393.8 393.8 393.8 474.1 393.8 393.8 393.8
[6,] 393.8 393.8 393.8 393.8 393.8 474.1 393.8 393.8
[7,] 393.8 393.8 393.8 393.8 393.8 393.8 474.1 393.8
[8,] 393.8 393.8 393.8 393.8 393.8 393.8 393.8 474.1
> round(sweep(sweep(Vi,1,sqrt(diag(Vi)),"/"),2,sqrt(diag(Vi)),"/"),3)
      [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]
[1,] 1.000 0.831 0.831 0.831 0.831 0.831 0.831 0.831
[2,] 0.831 1.000 0.831 0.831 0.831 0.831 0.831 0.831
[3,] 0.831 0.831 1.000 0.831 0.831 0.831 0.831 0.831
[4,] 0.831 0.831 0.831 1.000 0.831 0.831 0.831 0.831
[5,] 0.831 0.831 0.831 0.831 1.000 0.831 0.831 0.831
[6,] 0.831 0.831 0.831 0.831 0.831 1.000 0.831 0.831
[7,] 0.831 0.831 0.831 0.831 0.831 0.831 1.000 0.831
[8,] 0.831 0.831 0.831 0.831 0.831 0.831 0.831 1.000
> 393.8/(393.8 + 80.3)
[1] 0.8306265
```

Model 2 (R code)

- Model 2: $Y_{ij} = \beta_0 + a_i + (\beta_1 + b_i)t_{ij} + e_{ij}$, where
  - $a_i \sim N(0, \sigma_a^2)$, $b_i \sim N(0, \sigma_b^2)$, a_i and b_i are independent.
  - $e_{ij} \sim N(0, \sigma_e^2)$
  - (a_i, b_i) and e_{ij} are independent.

```
> ## Model 2: Random intercept and slope model (random intercept and slope are inde
> fit2 <- lmer(fas ~ time + (1 | id) + (0 + time | id), data=dt.l)
> summary(fit2)
Formula: fas ~ time + (1 | id) + (0 + time | id)

Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept) 392.129  19.802
 id.1     time          8.938   2.990
 Residual               26.981   5.194
Number of obs: 192, groups:  id, 24

Fixed effects:
            Estimate Std. Error t value
(Intercept)  30.9301     4.1257   7.497
time          4.7644     0.6318   7.541
```

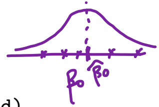
σ_a^2
σ_b^2
σ_e^2

Model 2 (R code)

```
> Zi <- cbind(rep(1,8), seq(1:8))
> G <- matrix(c(392.129,0,0,8.938),2,2)
> sigmasq.e <- 26.981
> Vi <- Zi%*%G%*%t(Zi) + sigmasq.e * diag(8)
> Vi
        [,1]    [,2]    [,3]    [,4]    [,5]    [,6]    [,7]    [,8]
[1,] 428.048 410.005 418.943 427.881 436.819 445.757 454.695 463.633
[2,] 410.005 454.862 445.757 463.633 481.509 499.385 517.261 535.137
[3,] 418.943 445.757 499.552 499.385 526.199 553.013 579.827 606.641
[4,] 427.881 463.633 499.385 562.118 570.889 606.641 642.393 678.145
[5,] 436.819 481.509 526.199 570.889 642.560 660.269 704.959 749.649
[6,] 445.757 499.385 553.013 606.641 660.269 740.878 767.525 821.153
[7,] 454.695 517.261 579.827 642.393 704.959 767.525 857.072 892.657
[8,] 463.633 535.137 606.641 678.145 749.649 821.153 892.657 991.142
> round(sweep(sweep(Vi,1,sqrt(diag(Vi)),"/"),2,sqrt(diag(Vi)),"/"),3)
      [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]
[1,] 1.000 0.929 0.906 0.872 0.833 0.792 0.751 0.712
[2,] 0.929 1.000 0.935 0.917 0.891 0.860 0.828 0.797
[3,] 0.906 0.935 1.000 0.942 0.929 0.909 0.886 0.862
[4,] 0.872 0.917 0.942 1.000 0.950 0.940 0.926 0.909
[5,] 0.833 0.891 0.929 0.950 1.000 0.957 0.950 0.939
[6,] 0.792 0.860 0.909 0.940 0.957 1.000 0.963 0.958
[7,] 0.751 0.828 0.886 0.926 0.950 0.963 1.000 0.969
[8,] 0.712 0.797 0.862 0.909 0.939 0.958 0.969 1.000
```

Model 3 (R code)

- Model 3: $Y_{ij} = \beta_0 + a_i + (\beta_1 + b_i)t_{ij} + e_{ij}$, where
  - $a_i \sim N(0, \sigma_a^2)$, $b_i \sim N(0, \sigma_b^2)$, a_i and b_i are dependent.
  - $e_{ij} \sim N(0, \sigma_e^2)$
  - (a_i, b_i) and e_{ij} are independent.




```
> ## Model 3: Random intercept and slope model (With unstructured)
> fit3 <- lmer(fas ~ time + (1 + time| id), data=dt.l)
> summary(fit3)
Formula: fas ~ time + (1 + time | id)

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 id       (Intercept) 405.101  20.127
          time          9.239   3.040   -0.35
 Residual              26.846   5.181
Number of obs: 192, groups:  id, 24

Fixed effects:
            Estimate Std. Error t value
(Intercept)  30.9301     4.1903   7.381
time          4.7644     0.6416   7.426
```

$$G = \begin{pmatrix} \sigma_a^2 & \rho\sigma_a\sigma_b \\ & \sigma_b^2 \end{pmatrix}$$

Model 3 (R code)

```
> Zi <- cbind(rep(1,8), seq(1:8))
> G <- matrix(c(405.101,(-0.35*sqrt(405.101)*sqrt(9.239)),(-0.35*sqrt(405.101)*sqrt
> sigmasq.e <- 26.846
> Vi <- Zi%*%G%*%t(Zi) + sigmasq.e * diag(8)
> Vi
         [,1]     [,2]     [,3]     [,4]     [,5]     [,6]     [,7]     [,8]
[1,] 398.3615 359.3423 347.1690 334.9958 322.8225 310.6493 298.4760 286.3028
[2,] 359.3423 383.2540 353.4738 350.5395 347.6053 344.6710 341.7368 338.8026
[3,] 347.1690 353.4738 386.6245 366.0833 372.3880 378.6928 384.9976 391.3023
[4,] 334.9958 350.5395 366.0833 408.4730 397.1708 412.7146 428.2583 443.8021
[5,] 322.8225 347.6053 372.3880 397.1708 448.7996 446.7363 471.5191 496.3018
[6,] 310.6493 344.6710 378.6928 412.7146 446.7363 507.6041 514.7798 548.8016
[7,] 298.4760 341.7368 384.9976 428.2583 471.5191 514.7798 584.8866 601.3013
[8,] 286.3028 338.8026 391.3023 443.8021 496.3018 548.8016 601.3013 680.6471
> round(sweep(sweep(Vi,1,sqrt(diag(Vi)),"/"),2,sqrt(diag(Vi)),"/"),3)
      [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]
[1,] 1.000 0.920 0.885 0.830 0.763 0.691 0.618 0.550
[2,] 0.920 1.000 0.918 0.886 0.838 0.781 0.722 0.663
[3,] 0.885 0.918 1.000 0.921 0.894 0.855 0.810 0.763
[4,] 0.830 0.886 0.921 1.000 0.928 0.906 0.876 0.842
[5,] 0.763 0.838 0.894 0.928 1.000 0.936 0.920 0.898
[6,] 0.691 0.781 0.855 0.906 0.936 1.000 0.945 0.934
[7,] 0.618 0.722 0.810 0.876 0.920 0.945 1.000 0.953
[8,] 0.550 0.663 0.763 0.842 0.898 0.934 0.953 1.000
```

Compare three models (R code)

```
> ## Compare model fit
> cbind(extractAIC(fit1), extractAIC(fit2), extractAIC(fit3))
         [,1]     [,2]     [,3]
[1,]    4.000    5.000    6.000  Penalty
[2,] 1481.583 1365.422 1364.602  AIC
## Model 1
Fixed effects:
            Estimate Std. Error t value
(Intercept)  30.9301     4.2941   7.203
time          4.7644     0.2822  16.881
## Model 2
            Estimate Std. Error t value
(Intercept)  30.9301     4.1257   7.497
time          4.7644     0.6318   7.541
## Model 3
            Estimate Std. Error t value
(Intercept)  30.9301     4.1903   7.381
time          4.7644     0.6416   7.426

> confint(fit2)
Computing profile confidence intervals ...
                2.5 %    97.5 %
(Intercept) 22.691755 39.168364
time         3.502985  6.025785
```

GLMM (R code)

```
library(lme4)
fit <- glmer(y ~ cov1 + cov2 + (1 | id), data=data, family=binomial)
```

SAS code

BIOS719_Note7_Suppl.pdf