LINEAR MIXED MODELS FOR LONGITUDINAL DATA
Multilevel Data and Other Topics

Outline:

- Review two-level model
  - Within/between- subject representation
- Nested random effects
- Crossed random effects
- Time-varying covariates
- Residual Diagnostics
- General guidelines for model building for longitudinal data analysis

Advantages of Mixed-effects Models

- Explicitly models individual change across time
- More flexible in terms of repeated measures
  – need not have same number of obs per subject
  – time can be continuous, rather than a fixed set of points
- Flexible specification of the covariance structure among repeated measures
  $\Rightarrow$ methods for testing specific determinants of this structure
- Can be extended to higher-level models
  $\Rightarrow$ repeated observations within individuals within clusters
- Generalizations for non-normal data

- Recall, 2-level model for longitudinal data

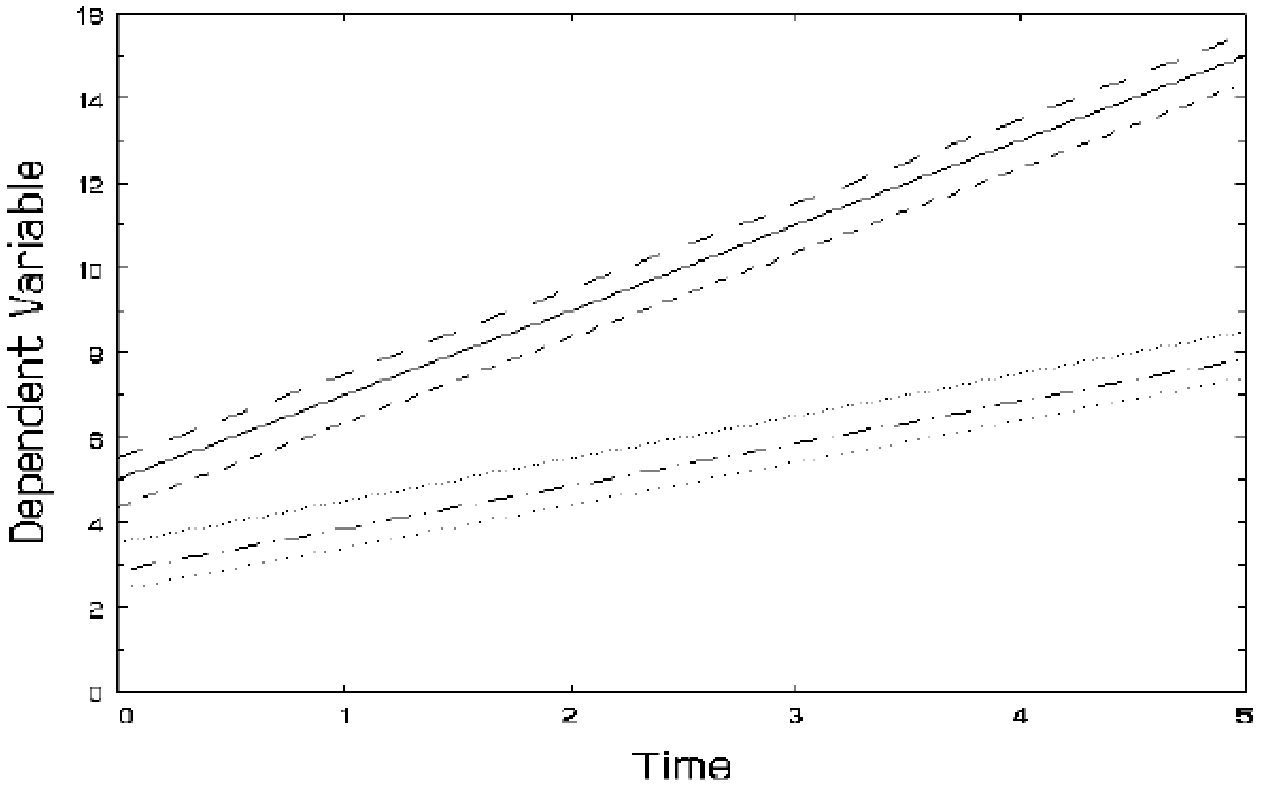
$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + D_i\boldsymbol{U}_i + \boldsymbol{Z}_i$$

where

$$\boldsymbol{U}_i \sim N(0, G) \quad \text{and} \quad \boldsymbol{Z}_i \sim N(0, R_i)$$

$i = 1, \ldots, m$ individuals

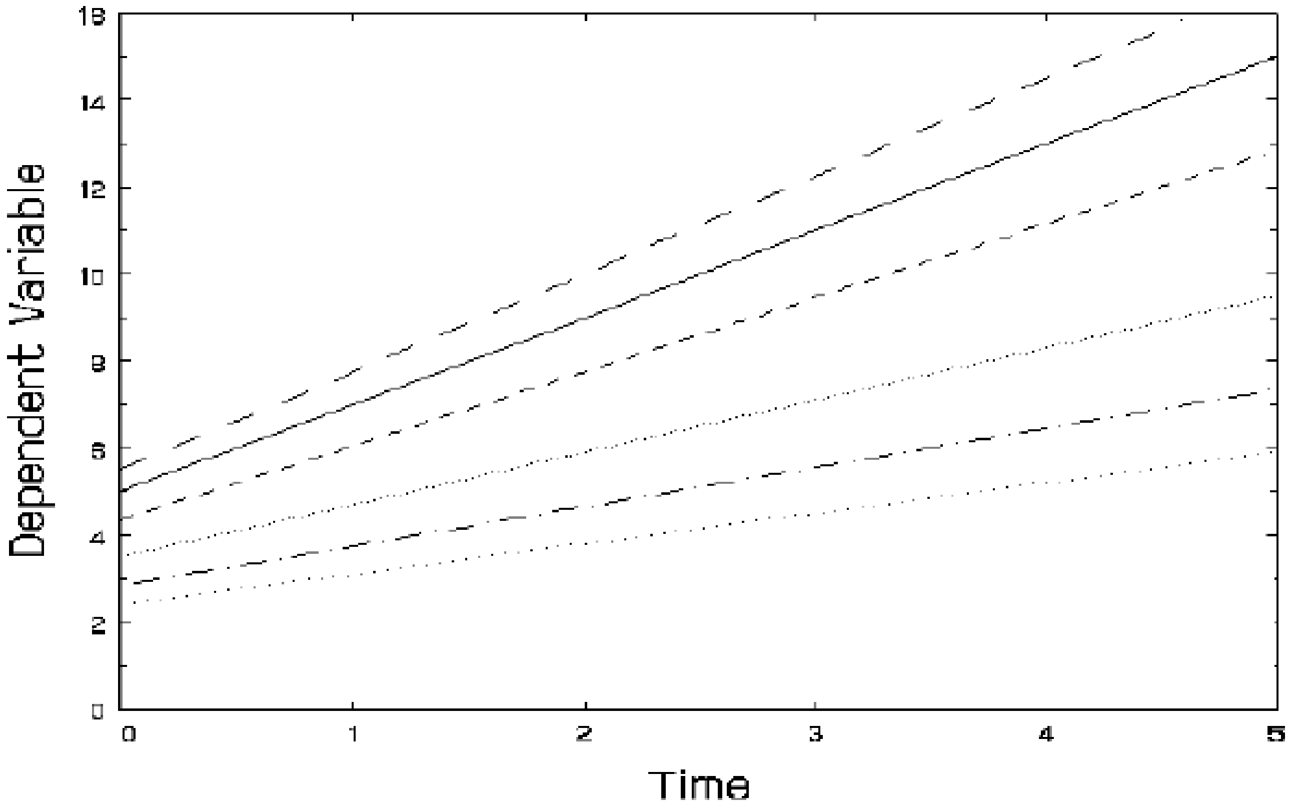
- Example: **Random intercept Model**
  each subject is parallel to their group trend



$$Y_{ij} = \beta_0 + \beta_1 \texttt{time}_{ij} + \beta_2 \texttt{group}_i + \beta_3 \texttt{group}_i \times \texttt{time}_{ij} + U_{0i} + Z_{ij}$$

$U_{0i} \sim N(0, \nu^2)$ and $Z_{ij} \sim N(0, \tau^2)$ are independent

- Example: **Random intercept and slope (time trend) Model**
  subjects deviate in terms of both intercept and slope



$$Y_{ij} = \beta_0 + \beta_1 \texttt{time}_{ij} + \beta_2 \texttt{group}_i + \beta_3 \texttt{group}_i \times \texttt{time}_{ij} + U_{0i} + U_{1i} \times \texttt{time}_{ij} + Z_{ij}$$

$(U_{0i}, U_{1i})' \sim N(0, G)$ and $Z_{ij} \sim N(0, \tau^2)$ are independent

Hierarchical linear model
Within/Between-Subject representation

- (**Level 1**) Within-subject model ($j = 1, \ldots, n_i$):

$$Y_{ij} = b_{0i} + b_{1i} X^W_{1ij} + b_{2i} X^W_{2ij} + \cdots + Z_{ij}$$

  where X^W_{1ij}, X^W_{2ij}, $\ldots$ are within-subject covariates (eg, time, drug plasma level).

- (**Level 2**) Between-subject model ($i = 1, \ldots, m$) ("slopes as outcomes" model):
  Subject-specific intercept:

$$b_{0i} = \beta_0 + \beta_{01} X^B_{1i} + \beta_{02} X^B_{2i} + \cdots + U_{0i}$$

  where X^B_{1i}, X^B_{2i}, $\ldots$ are between-subject covariates (eg, group, sex)

Subject-specific slopes:

$$b_{1i} = \beta_1 + \beta_{11} X_{1i}^B + \beta_{12} X_{2i}^B + \cdots + U_{1i}$$

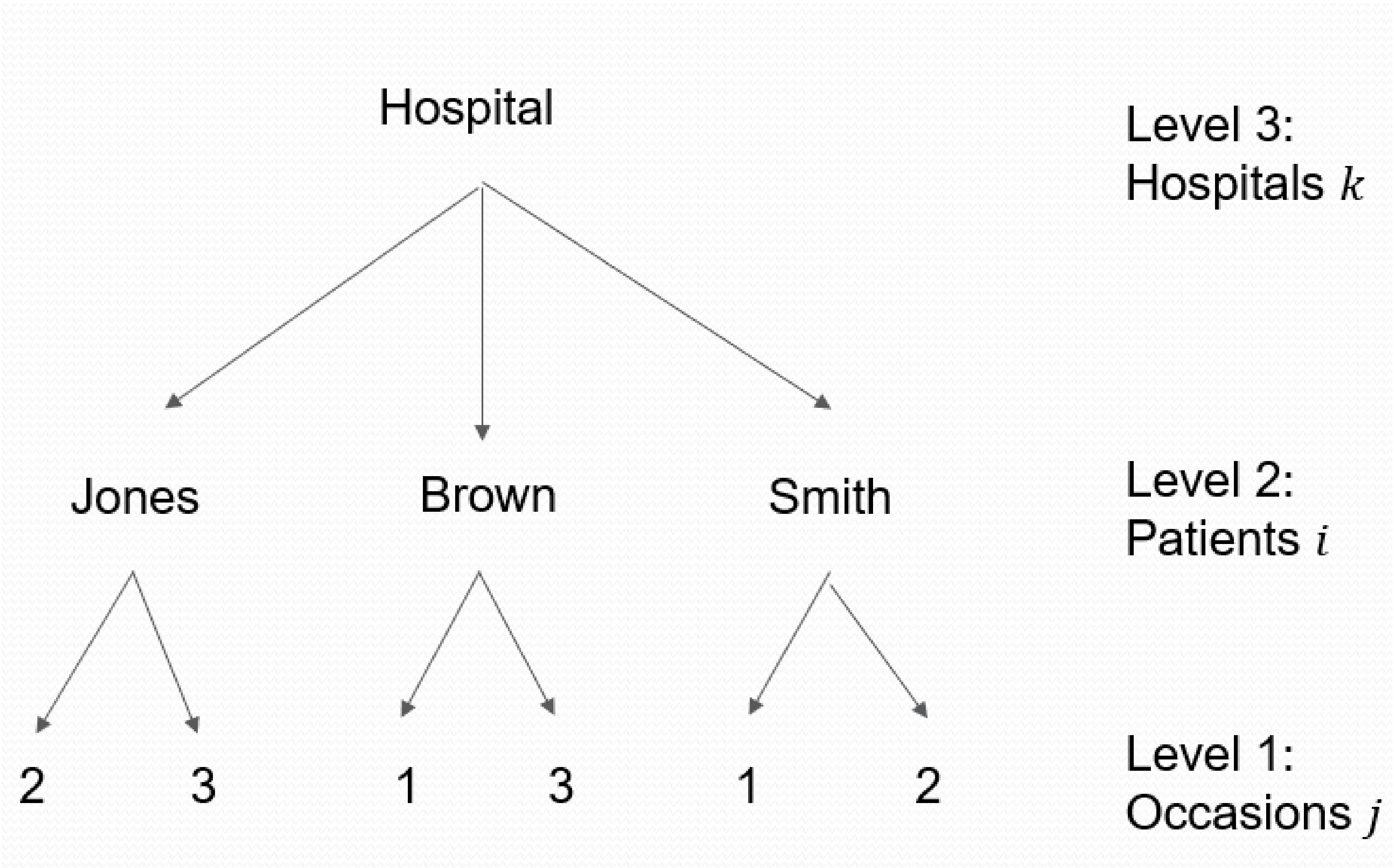
$$\cdots$$

Or, some slope could be common slope without U, eg:

$$b_{2i} = \beta_2 + \beta_{21} X_{1i}^B + \beta_{22} X_{2i}^B + \cdots$$

Nested random effects

- We have seen two-level models:
  Repeated measurements are nested in individuals (cluster).
- In three-level models, clusters themselves are nested in superclusters, forming a hierarchical structure.
- Eg, repeated measurements (level 1) for patients (level 2) who are clustered in hospitals (level 3).


Hospital
Level 3:
Hospitals k
Jones
Brown
Smith
Level 2:
Patients i
2
3
1
3
1
2
Level 1:
Occasions j

- Suppose we have random intercepts only for each level of clusters and supercluster, our mixed effect model is:

$$Y_{kij} = \boldsymbol{\beta X}_{kij} + U_{ki}^{(2)} + U_{k}^{(3)} + Z_{kij},$$

$U_{ki}^{(2)} \sim N(0, \nu_2^2)$, $U_k^{(3)} \sim N(0, \nu_3^2)$, $Z_{kij} \sim N(0, \tau^2)$ are independent

- $k = 1, \ldots, K$ indicates the supercluster (eg, hospital)
- $i = 1, \ldots, m$ indicates the level 2 cluster (eg, patient)
- $j = 1, \ldots, n_{ik}$ indicates the repeated measurements
- $\boldsymbol{X}_{kij}$ represents any fixed effect, can be of any level
- $U_{ki}^{(2)}$ is random intercept for level 2 clusters (eg, patient)
- $U_k^{(3)}$ is random intercept for level 3 supercluster (eg, hospital)
- **A more general model**: can include random slope in model
- Multilevel model can be generalized to an arbitrary number of levels.

- For three-level models, we can consider several types of intraclass correlations (ICC) for pairs of responses.

- For measurements from different clusters i and i' within the same supercluster k, we obtain

$$\rho_k = \mathrm{corr}(Y_{kij}, Y_{ki'j'}) = \frac{\nu_3^2}{\nu_2^2 + \nu_3^2 + \tau^2}$$

- For measurements within the same cluster i and supercluster k, we obtain

$$\rho_{ik} = \mathrm{corr}(Y_{kij}, Y_{kij'}) = \frac{\nu_2^2 + \nu_3^2}{\nu_2^2 + \nu_3^2 + \tau^2}$$

- Hence $\rho_{ik} > \rho_k$
  – Measurements from the same cluster are more correlated than measurements from different clusters, but within the same supercluster.

Example: Peak Expiratory Flow

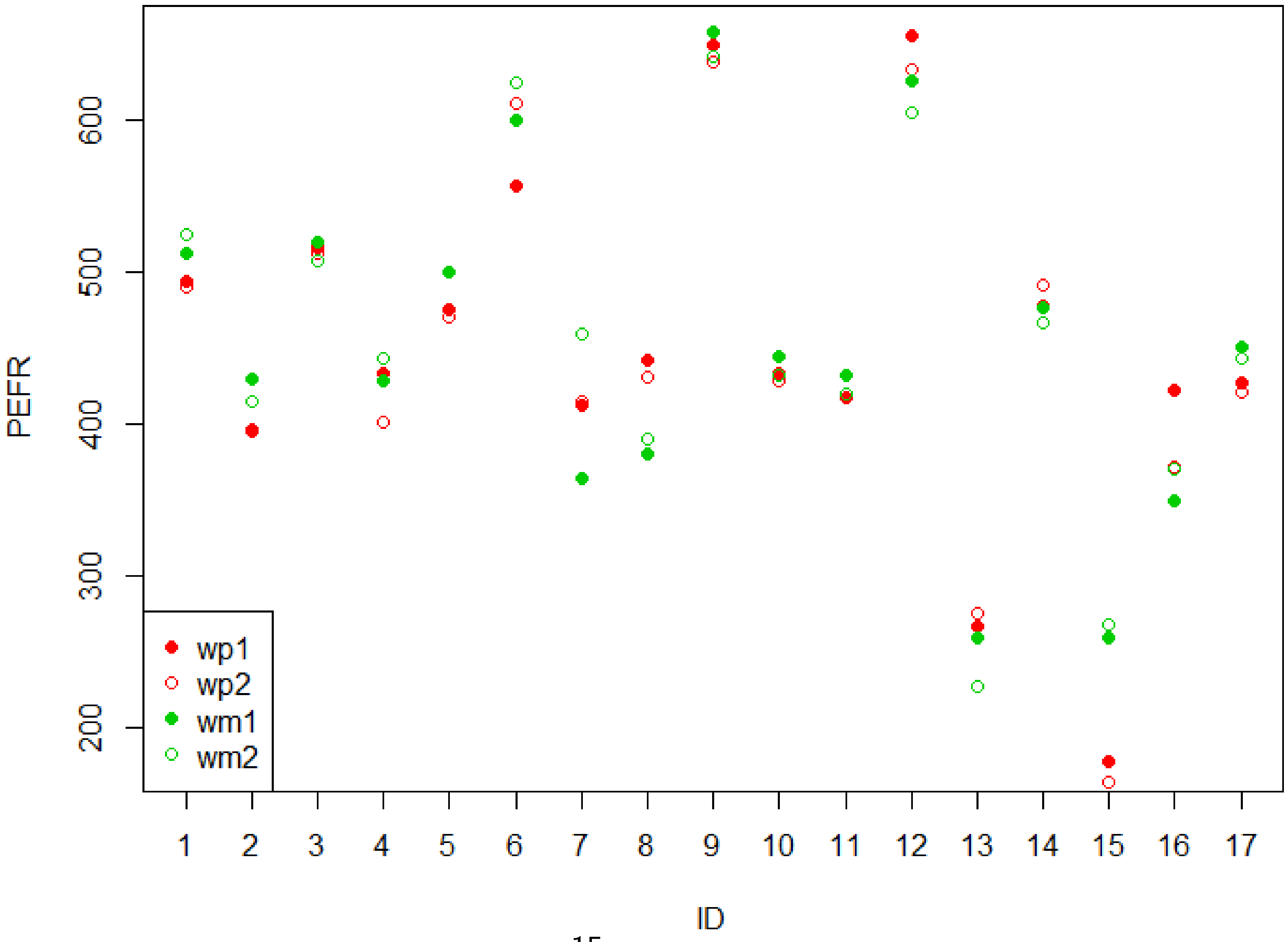
- A reliability study
- 17 subjects/persons
- Objective: illustrate a way of assessing the quality of two instruments for measuring people's peak-expiratory-flow rate (PEFR)
  – a person's maximum speed of expiration
- PEFR was measured using two methods, each on two occasions:
  – Standard Wright peak flow meter
  – Mini Wright peak flow meter

- Data:

```
   id wp1 wp2 wm1 wm2
1   1 494 490 512 525
2   2 395 397 430 415
3   3 516 512 520 508
4   4 434 401 428 444
5   5 476 470 500 500
6   6 557 611 600 625
7   7 413 415 364 460
8   8 442 431 380 390
9   9 650 638 658 642
10 10 433 429 445 432
11 11 417 420 432 420
12 12 656 633 626 605
13 13 267 275 260 227
14 14 478 492 477 467
15 15 178 165 259 268
16 16 423 372 350 370
17 17 427 421 451 443
```

- Plot by ID in R to explore data:

```
data=read.csv("pefr.csv")

plot(data$id,data$wp1,col=2,pch=19,xlab="ID",ylab="PEFR",xaxt=’n’)
axis(side = 1, at=1:17)
points(data$id,data$wp2,col=2)
points(data$id,data$wm1,col=3,pch=19)
points(data$id,data$wm2,col=3)
legend("bottomleft",c("wp1","wp2","wm1","wm2"),pch=c(19,1,19,1),col=c(2,2,3,3))
```



PEFR
200
300
400
500
600
ID
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
● wp1
○ wp2
● wm1
○ wm2

- For a given subject, the measurements using the same method tend to resemble each other more than measurements using the other method.

- For some subjects, the Wright peak flow meter measurements are lower, whereas for other subjects, the Mini Wright meter measurements are lower.

- The difference between methods is large for some subjects, and small for others.

- We can accommodate the between-method within-subject heterogeneity by including a random intercept for each combination of method and subject.

- We can fit a three level model with:
  – a level 2 random intercept for method nested in person
  – a level 3 random intercept for person (subject)

- SAS code:

```
proc import datafile="Yourpath/pefr.csv"
out=pefr1 dbms=csv replace;
getnames=yes;
run;

/* Transform the data into long format */
data pefr2;
  set pefr1;
  array vars4 WP1 WP2 WM1 WM2; /* define the array */
  do i = 1  to 2;
    w = varsi;
        method="p";
        occasion=i;
    output;
  end;
  do i = 3  to 4;
        w=varsi;
        method="m";
        occasion=i-2;
    output;
  end;
  drop WP1 WP2 WM1 WM2 i;
run;

proc print data=pefr2 (obs=10);
run;

                 Obs    ID     w     method    occasion

                   1     1    494      p           1
                   2     1    490      p           2
                   3     1    512      m           1
                   4     1    525      m           2
                   5     2    395      p           1
                   6     2    397      p           2
                   7     2    430      m           1
                   8     2    415      m           2
                   9     3    516      p           1
                  10     3    512      p           2
```

```
/* Nested random effects:
Random effect for method is nested within the random effect for id*/

proc mixed noclprint covtest data=pefr2;
  class id method;
  model w=method/s;
  random intercept/ g subject = id;
  random intercept/g subject=method(id);
  title1 ‘Three-level model, method as a fixed effect’;
run;
```

```
  Three-level model, method as a fixed effect

  Covariance Parameter Estimates

                                            Standard         Z
 Cov Parm      Subject        Estimate         Error     Value       Pr > Z

 Intercept     id                12542       4532.62      2.77       0.0028
 Intercept     method(id)       393.57        198.62      1.98       0.0238
 Residual                       315.37       76.4879      4.12       <.0001
```

```
                                          Fit Statistics

                               -2 Res Log Likelihood          675.6
                               AIC (Smaller is Better)        681.6
                               AICC (Smaller is Better)       682.0
                               BIC (Smaller is Better)        684.1



Solution for Fixed Effects

                                     Standard
  Effect      method   Estimate         Error     DF   t Value   Pr > |t|

Intercept                447.88       27.7519      16      16.14      <.0001
method       m           6.0294        8.0532      16       0.75      0.4649
method       p                0             .       .         .          .
```

- Between-subject variance is estimated as $\hat{\nu}_3^2 = 12542$
- Between-methods within-subjects variance is $\hat{\nu}_2^2 = 393.57$.
- Residual variance between occasions, within methods and subjects, is estimated as $\hat{\tau}^2 = 315.37$.

- Intraclass Correlations (ICC):
  - Estimated intraclass correlation between measurements on the same individual using different methods is

$$\hat{\rho}_k = \widehat{\mathrm{corr}}(Y_{kij}, Y_{ki'j'}) = \frac{\hat{\nu}_3^2}{\hat{\nu}_2^2 + \hat{\nu}_3^2 + \hat{\tau}^2} = \frac{12542}{393.57 + 12542 + 315.37} = 0.95$$

  - Estimated intraclass correlation between measurements on the same individual using the same method is

$$\hat{\rho}_{ik} = \widehat{\mathrm{corr}}(Y_{kij}, Y_{kij'}) = \frac{\hat{\nu}_2^2 + \hat{\nu}_3^2}{\hat{\nu}_2^2 + \hat{\nu}_3^2 + \hat{\tau}^2} = \frac{393.57 + 12542}{393.57 + 12542 + 315.37} = 0.98$$

- Test the null hypothesis: variance component for methods is zero (ie, $\mathrm{var}(U_{ki}^{(2)}) = \nu_2^2 = 0$)
  - use a likelihood ratio test
  - This is a test for **uncorrelated** random effects:
    the model assumes independence between $U_{ki}^{(2)}$ and $U_k^{(3)}$
    $\rightarrow$ always have $\mathrm{cov}(U_{ki}^{(2)}, U_k^{(3)}) = 0$
    $\rightarrow H_0 : \mathrm{var}(U_{ki}^{(2)}) = \nu_2^2 = 0$ vs $H_A : \mathrm{var}(U_{ki}^{(2)}) = \nu_2^2 > 0$
  - Null hypothesis is on the boundary of the parameter space
  - The asymptotic distribution for testing k **uncorrelated** random effects vs $k+1$ uncorrelated random effects is $0.5\chi_0^2 + 0.5\chi_1^2$.
    – equivalent to using χ_1^2 and divide p-value by 2

- Fit reduced model in SAS:

```
/* No nested random effects for method*/
proc mixed noclprint covtest data=pefr2;
  class id method;
  model w=method/s;
  random intercept/g subject = id;
  title1 'Two-level model,  method as a fixed effect only';
run;
```

```
                                   -2 Res Log Likelihood           684.9
```

- The Likelihood ratio test statistics is

$$-2LogL_{reduced} - (-2LogL_{full}) = 684.9 - 675.6 = 9.3$$

- So p-value$=0.5\,\Pr(\chi^2_1 > 9.3) = 0.001$
- Hence, we reject the null hypothesis and concludes that a random effect for methods is required.

- Summary of results:
  - There is no evidence of systematic bias between the methods
  - Since the estimated fixed effect for method is small and not significant at the 5% level, we can remove it in our final model.
  - There is some evidence for a subject by method interaction, or subject-specific bias.
  - The methods appear to have good test-retest reliability.

- We can also obtain Empirical Bayes prediction for each subject (for final model without fixed effect for method).
- SAS code:

```
proc mixed noclprint covtest data=pefr2;
  class id method;
  model w=/s;
  random intercept/ g subject = id solution;
  random intercept/g subject=method(id);
  title1 'Three-level model, no fixed effect, Empirical Bayes prediction';
run;
```

- Note for SAS: Empirical Bayes prediction of random effects can be obtained by SOLUTION option in any of RANDOM statements

- Results:

```
   Solution for Random Effects

                                             Std Err
  Effect      method  id   Estimate             Pred      DF    t Value     Pr > |t|

Intercept             1     53.2143          31.3939      34       1.70       0.0992
Intercept     m       1     10.1636          15.5792      34       0.65       0.5185
Intercept     p       1     -8.5551          15.5792      34      -0.55       0.5865
Intercept             2    -40.7746          31.3939      34      -1.30       0.2027
Intercept     m       2      8.7431          15.5792      34       0.56       0.5783
Intercept     p       2     -9.9756          15.5792      34      -0.64       0.5263
```

- Interpretation for subject 1:

  – Mini Wright meter appears to be positively biased compared with the Wright peak flow meter

  – Overall this subject has a higher PEFR compared to other subjects.

Intraclass Correlation vs Pearson Correlation

- Example:

  Suppose we only concentrate on two Mini Wright peakflow measurements at two occasions:

  – Whether Mini Wright peakflow measurements is reliable (similar if measured twice)?

- Now we only look at Mini Wright peakflow measurements

- We can fit a two-level model with one random intercept for individual

- ICC between two Mini Wright peakflow measurements within same subject is

$$\hat{\rho} = \widehat{\mathrm{corr}}(Y_{ij}, Y_{ij'}) = \frac{\hat{\nu}^2}{\hat{\nu}^2 + \hat{\tau}^2} = \frac{12188}{12188 + 396.44} = 0.968$$

- The Pearson Correlation is

$$r = \frac{\frac{1}{m-1}\sum_{j=1}^{m}(y_{1j} - \bar{y}_1)(y_{2j} - \bar{y}_2)}{s_{y1}s_{y2}} = 0.967,$$

– m is number of subjects

– y_{1j} and y_{2j} are Mini Wright peakflow measurements at occasion 1 and 2

– s_{y1} and s_{y2} are standard deviation of measurements at occasion 1 and 2, respectively.

- SAS code for above results:

```
proc mixed noclprint covtest data=pefr2(where=(method eq "m"));
  class id;
  model w=/s;
  random intercept/ g subject = id;
run;

* Person’s correlation;
proc corr data=pefr1;
var wm1 wm2;
run;
```

```
                  Covariance Parameter Estimates

                                        Standard          Z
Cov Parm      Subject     Estimate        Error       Value       Pr > Z

Intercept     id             12188      4379.54        2.78       0.0027
Residual                    396.44       135.98        2.92       0.0018
```

```
Pearson Correlation Coefficients, N = 17
       Prob > |r| under H0: Rho=0

                    wm1            wm2
     wm1        1.00000        0.96703
                                <.0001

     wm2        0.96703        1.00000
                 <.0001
```

- Suppose we add 100 to the 2nd Mini Wright peakflow measurements.

- The intraclass correlation between the two measurements is now

$$\hat{\rho} = \widehat{\text{corr}}(Y_{ij}, Y_{ij'}) = \frac{\hat{\nu}^2}{\hat{\nu}^2 + \hat{\tau}^2} = \frac{9543.39}{9543.39 + 5684.68} = 0.627$$

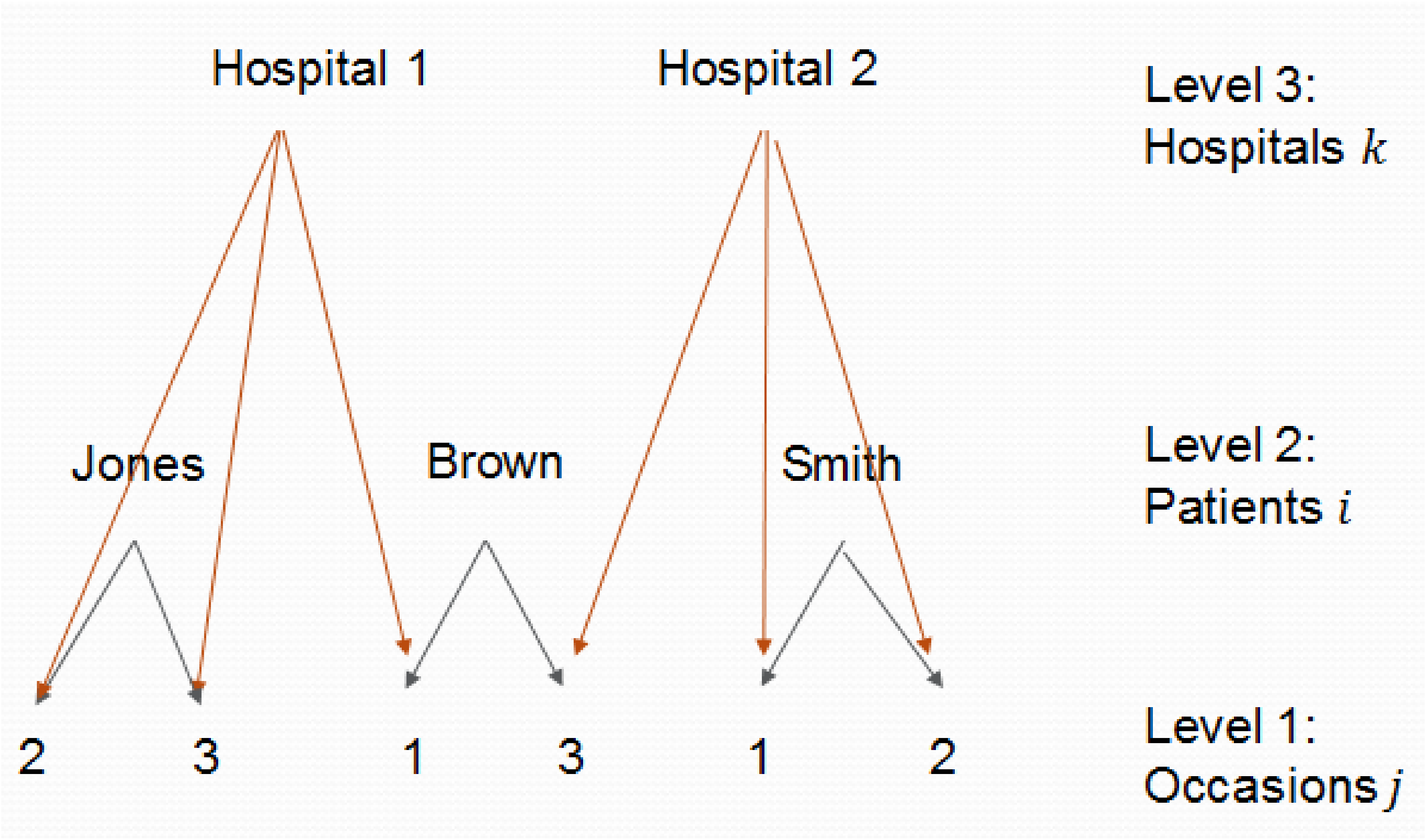
- The Pearson Correlation is still

$$r = \frac{\frac{1}{m-1}\sum_{j=1}^{m}(y_{1j} - \bar{y}_1)(y_{2j} - \bar{y}_2)}{s_{y1}s_{y2}} = 0.967,$$

- Pearson correlation is based on deviations of the first and second measurements from **their respective means**.
- ICC is based on deviations from the **overall or pooled mean**.
- ICC is useful when the measurements within a unit are exchangeable
- Pearson Correlation is only defined for pairs of variables
- ICC summarizes dependence for clusters of size ≥ 2, and clusters of variable sizes

Crossed random effects

- In some studies, the units can be cross-classified by two or more factors.
- Eg, patients can have repeated visits, however, they may go to different hospitals at different visits.
- Hence, patients and hospitals constitute crossed random effects.


Hospital 1
Hospital 2
Level 3:
Hospitals k
Jones
Brown
Smith
Level 2:
Patients i
2
3
1
3
1
2
Level 1:
Occasions j

- Suppose we have random intercepts for each of the crossed random effects, our mixed effect model can be written as:

  $$Y_{kij} = \boldsymbol{\beta} \boldsymbol{X}_{kij} + U_k^{Clus1} + U_i^{Clus2} + Z_{kij},$$

  $U_k^{Clus1} \sim N(0, \nu^2_{Clus1})$, $U_i^{Clus2} \sim N(0, \nu^2_{Clus2})$, $Z_{kij} \sim N(0, \tau^2)$ are independent

- U_k^{Clus1} $(k = 1, \ldots, K)$ indicates random intercept for one of the crossed clusters (eg, hospital)

- U_i^{Clus2} $(i = 1, \ldots, m)$ indicates random intercept for another crossed cluster (eg, individual)

- $j = 1, \ldots, n_{ik}$ indicates the repeated measurements

- $\boldsymbol{X}_{kij}$ represents any fixed effect

- For cross-effect models, we can consider several types of ICC for pairs of responses
- For measurements from different clusters i and i' within the same cluster k, we obtain

$$\text{corr}(Y_{kij}, Y_{ki'j'}) = \frac{\nu^2_{Clus1}}{\nu^2_{Clus2} + \nu^2_{Clus1} + \tau^2}$$

- For measurements within the same cluster i but different clusters k and k', we obtain

$$\text{corr}(Y_{kij}, Y_{k'ij'}) = \frac{\nu^2_{Clus2}}{\nu^2_{Clus2} + \nu^2_{Clus1} + \tau^2}$$

- For measurements within the same cluster i and the same clusters k, we obtain

$$\mathrm{corr}(Y_{kij}, Y_{kij'}) = \frac{\nu^2_{Clus2} + \nu^2_{Clus1}}{\nu^2_{Clus2} + \nu^2_{Clus1} + \tau^2}$$

- Hence, $\mathrm{corr}(Y_{kij}, Y_{kij'}) > \mathrm{corr}(Y_{kij}, Y_{k'ij'})$ and $\mathrm{corr}(Y_{kij}, Y_{kij'}) > \mathrm{corr}(Y_{kij}, Y_{ki'j'})$
  – measurements from the same crossed clusters are more correlated than the measurements from different clusters.

Example: Investment

- Grunfeld (1958) analyzed data on 10 large American corporations collected annually from 1935 to 1954
- Study question: How investment depends on capital stock of firms, etc.
- The variables in dataset are:
  - fn (firm identifier),
  - firmname,
  - yr (year),
  - I (annual gross investment in million dollars),
  - F (market value of firm in million dollars),
  - C (real value of capital stock in million dollars).

- Investment behavior of corporations depends on firms and years
  – We want to have random intercepts for firms and years respectively

- This model differs from the models considered previously
  – Two random intercepts for firm and year are crossed instead of nested.

- Random intercept for firm i is shared across all years for a given firm i

- Random intercept for year k is shared by all firms in a given year k.

- SAS code:

```
proc import datafile="Yourpath/grunfeld.csv"
out=grunfeld dbms=csv replace;
getnames=yes;
run;

proc print data=grunfeld (obs=10);
run;
```

```
  Obs    FN     YR        I          F           C       FIRMNAME

   1      1    1935    317.600    3078.50      2.800    General Motors
   2      1    1936    391.800    4661.70     52.600    General Motors
   3      1    1937    410.600    5387.10    156.900    General Motors
   4      1    1938    257.700    2792.20    209.200    General Motors
   5      1    1939    330.800    4313.20    203.400    General Motors
   6      1    1940    461.200    4643.90    207.200    General Motors
   7      1    1941    512.000    4551.20    255.200    General Motors
   8      1    1942    448.000    3244.10    303.700    General Motors
   9      1    1943    499.600    4053.70    264.100    General Motors
  10      1    1944    547.500    4379.30    201.600    General Motors
```

- SAS code for mixed-effects model:

```
proc mixed noclprint covtest data=grunfeld;
  class fn yr;
  model I=F C/s;
  random intercept/subject=fn;
  random intercept/subject=yr solution;
  title1 'Crossed Random Effect Model';
run;
```

```
                 Covariance Parameter Estimates

                                        Standard         Z
Cov Parm      Subject    Estimate          Error     Value       Pr > Z

Intercept     FN          7406.77        3576.75      2.07       0.0192
Intercept     YR          29.0444         124.01      0.23       0.4074
Residual                  2752.84         305.40      9.01       <.0001
```

```
                   Solution for Fixed Effects

                         Standard
Effect       Estimate       Error      DF    t Value    Pr > |t|

Intercept    -58.8334     29.5049       9      -1.99      0.0773
F              0.1101     0.01060     169      10.38      <.0001
C              0.3106     0.01745     169      17.80      <.0001


                       Solution for Random Effects

                                         Std Err
 Effect       FN    YR      Estimate        Pred      DF    t Value    Pr > |t|

 Intercept    1             -11.3559     43.5012     169      -0.26      0.7944
 Intercept    2               157.76     30.7586     169       5.13      <.0001
<snip>
 Intercept          1935      3.2790      5.1559     169       0.64      0.5257
 Intercept          1936      1.7722      5.1501     169       0.34      0.7312
<snip>
```

- The market value (F) and real value of capital stock (C) of a firm both have positive effects on investment.
- Estimated variance between firms (ie, variance of random intercept for firm) is 7406.77
- Estimated variance between years (ie, variance of random intercept for year) is only 29.04
- The remaining residual variance, not due to firms and years, is estimated as 2752.84.
- ICC for firms (ie, correlation within firm) is estimated as

$$\hat{\rho}_{firm} = \widehat{\text{corr}}(Y_{k'ij}, Y_{kij'}) = \frac{7406.77}{7406.77 + 29.04 + 2752.84} = 0.73$$

where i is for firm, k is for year.

- ICC for years (ie, correlation within year) is estimated as

$$\hat{\rho}_{year} = \widehat{\text{corr}}(Y_{kij}, Y_{ki'j'}) = \frac{29.04}{7406.77 + 29.04 + 2752.84} = 0.003$$

- There is a high correlation over years within firms and a small correlation over firms within years, given the covariates.

Time-Varying Covariates for Longitudinal Models

- Recall, time-varying covariates: covariates whose values change over time
  – e.g. time since baseline, current smoking status, environmental exposures

- We can divide the time-varying covariates into two classes.

  1. Covariates that vary systematically over time but are fixed by design of the study
     – treatment group indicator in a crossover trial
     – time since baseline

2. Covariates that vary randomly over time, i.e. stochastic variables which can be random
   – current blood glucose level
   – current smoking status or cumulative pack years
   – blood pressure
   – exposure to environmental pollutants

- Care has to be taken when fitting a longitudinal model with stochastic covariates.

- **Example**: Consider a longitudinal study designed to examine the effects of physical exercise on reducing blood glucose levels in patients with type 2 diabetes mellitus
  – Subjects with elevated blood glucose levels may exercise more
  – But once their blood glucose levels are normal, they stop exercise

- If we fit a model using exercise level as covariate, and blood glucose level as outcome variable:
  – seems that exercise will elevate the blood glucose level
- The wrong relationship is caused by the fact that the stochastic covariate, exercise level, is a function of (i.e. depends on) the outcome variable (blood glucose level).
- When covariates are time-varying and stochastic, the regression parameters **do not necessarily have the implied causal interpretations**.

Time-Varying Covariates - WS and BS effects

- Example: Riesby Depression study
- Study goal: relationship between Imipramine (IMI) and Desipramine (DMI) plasma levels and outcomes (Riesby and others, 1977)
- Subjects were assessed on week 0, 1, . . . , 5
- Main outcome: Hamilton depression score, measured at each time
- Time-varying covariates: imipramine (IMI) drug-plasma level and desipramine (DMI) drug-plasma level
  – No drug was administered during weeks 0 and 1
  – measurement start from week 2
- Note: In dataset 'RIESBYT4.raw": Week 2 is coded as 0 in this analysis of the last four study timepoints

- Might be ok to have causal interpretations between plasma levels and outcomes
  – because change of drug plasma level is by design in this study
  – Observational study: patients may take drug due to high depression
- Consider model with a random intercept U_{i1} and a random slope U_{i2}

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1 \texttt{week}_{ij} + \beta_2 \texttt{lnIMI}_{ij} + \beta_3 \texttt{lnDMI}_{ij} + U_{i0} + U_{i1} \texttt{week}_{ij} + Z_{ij} \\ &= (\beta_0 + U_{i0}) + (\beta_1 + U_{i1}) \texttt{week}_{ij} + \beta_2 \texttt{lnIMI}_{ij} + \beta_3 \texttt{lnDMI}_{ij} + Z_{ij} \end{aligned}$$

$\beta_0 =$ average week 2 depression score for drug-free patients
$\beta_1 =$ average depression score weekly improvement
$\beta_2 =$ average depression score difference for 1-unit change in `lnIMI`
$\beta_3 =$ average depression score difference for 1-unit change in `lnIMI`
$U_{i0} =$ individual intercept deviation
$U_{i1} =$ individual slope (w.r.t. week) deviation

- SAS code:

```
DATA one;
INFILE "Yourpath/riesbyt4.raw";
INPUT id hamdelt intcpt week sex endog lnimi lndmi;
run;

PROC SORT; BY id;run;

* calculate between-subjects component (average within each subject);
PROC MEANS NOPRINT;
CLASS id;
VAR lnimi lndmi;
OUTPUT OUT = two MEAN = mlnimi mlndmi;
run;

* calculate within-subjects component;
DATA three;
MERGE one two;
BY id;
lnidev = lnimi - mlnimi;
lnddev = lndmi - mlndmi;
run;
```

```
ods output solutionR=riesby.randeff3;
PROC MIXED data=three METHOD=ML COVTEST;
  CLASS id;
  MODEL hamdelt = week lnimi lndmi /SOLUTION
    outp=riesby.predict3 outpm=riesby.popmean3;
  RANDOM INTERCEPT week /SUB=id TYPE=UN G GCORR solution;
run;
```

- Results:

```
                    Fit Statistics

     -2 Log Likelihood                1498.8
     AIC (Smaller is Better)          1514.8
     AICC (Smaller is Better)         1515.4
     BIC (Smaller is Better)          1532.4
```

```
                   Solution for Fixed Effects

                         Standard
Effect       Estimate       Error       DF    t Value    Pr > |t|
Intercept      1.5214      3.7426       65       0.41      0.6857
week          -1.9669      0.2850       65      -6.90      <.0001
lnimi          0.6301      0.8211      116       0.77      0.4444
lndmi         -1.9666      0.6025      116      -3.26      0.0014
```

- Model with a time-varying covariate X_{ij}

$$Y_{ij} \quad = \quad (\beta_0 + U_{i0}) + (\beta_1 + U_{i1})\texttt{week}_{ij} + \beta_2 X_{ij} + Z_{ij}$$

- Is the effect of X_{ij} (eg, IMI level) purely within-subjects?
- What about decomposition

$$X_{ij} = \bar{X}_i + (X_{ij} - \bar{X}_i)$$

$\bar{X}_i$ is between-subjects component of X
$X_{ij} - \bar{X}_i$ is within-subjects component of X

- Model with decomposition of a time-varying covariate X_{ij}

  $$Y_{ij} \quad = \quad (\beta_0 + \beta_{BS}\bar{X}_i + U_{i0}) + (\beta_1 + U_{i1})\texttt{week}_{ij} + \beta_{WS}(X_{ij} - \bar{X}_i) + Z_{ij}$$

  Note effect of X is now

  $$\beta_{BS}\bar{X}_i + \beta_{WS}(X_{ij} - \bar{X}_i)$$

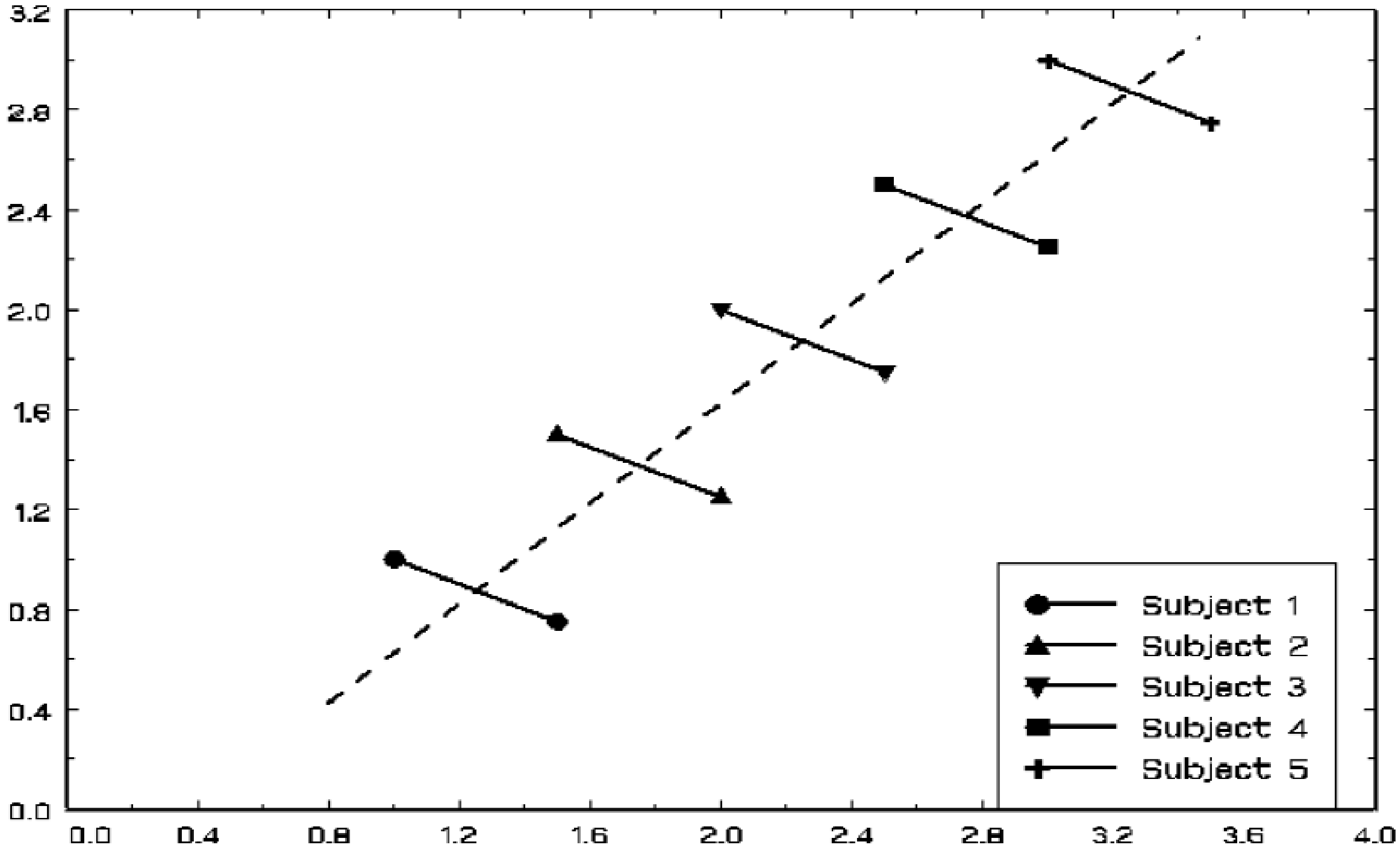
- β_{BS} = effect of $\bar{X}_i$ on $\bar{Y}_i$ (Between-subject or “cross-sectional”)
- β_{WS} = effect of $(X_{ij} - \bar{X}_i)$ on $(Y_{ij} - \bar{Y}_i)$ (Within-subject or “longitudinal”)

- Our previous model with only X_{ij} assumes equal BS and WS effects ($\beta_{BS} = \beta_{WS} = \beta^*$):
  Effect of X_{ij} is $\beta^* \bar{X}_i + \beta^*(X_{ij} - \bar{X}_i) = \beta^* X_{ij}$

- Equal WS and BS effects of X_{ij}?
  – can be tested by LR test (equivalent to nested models)

$$\beta_{BS}\bar{X}_i + \beta_{WS}(X_{ij} - \bar{X}_i) = (\beta_{BS} - \beta_{WS})\bar{X}_i + \beta_{WS}X_{ij}$$

  – there is no guarantee that β_{BS} and β_{WS} even agree on sign



Time-varying covariate effects: opposite sign WS and BS effects

- SAS code with β_{BS} and β_{WS} drug effects:

```
PROC MIXED data=three METHOD=ML COVTEST;
  CLASS id;
  MODEL hamdelt = week mlnimi mlndmi lnidev lnddev /SOLUTION;
  RANDOM INTERCEPT week /SUB=id TYPE=UN G GCORR;
TITLE2 'relaxing bs=ws drug effects';
RUN;
```

```
                          Fit Statistics

              -2 Log Likelihood               1495.8
              AIC (Smaller is Better)         1515.8
              AICC (Smaller is Better)        1516.7
              BIC (Smaller is Better)         1537.7


                     Solution for Fixed Effects

                          Standard
Effect        Estimate       Error      DF    t Value    Pr > |t|
Intercept      7.2669      5.0388      64      1.44      0.1541
week          -2.0238      0.2917      65     -6.94      <.0001
mlnimi        -0.3129      1.0037     115     -0.31      0.7558
mlndmi        -2.3671      0.7963     115     -2.97      0.0036
lnidev         2.4434      1.4561     115      1.68      0.0960
lnddev        -1.7963      0.9987     115     -1.80      0.0747
```

- LR test:
  χ^2 = 1498.8 - 1495.8 = 3 with DF=2
  $\Rightarrow$ p-value = 0.22
  $\Rightarrow$ Accept $H_0 : \beta_{BS} = \beta_{WS}$

  R code:

  ```
  > 1-pchisq(3,2)
  [1] 0.2231302
  ```

- Remark: You may further consider a more complicated mean model (eg, include diagnosis group, sex) for this study

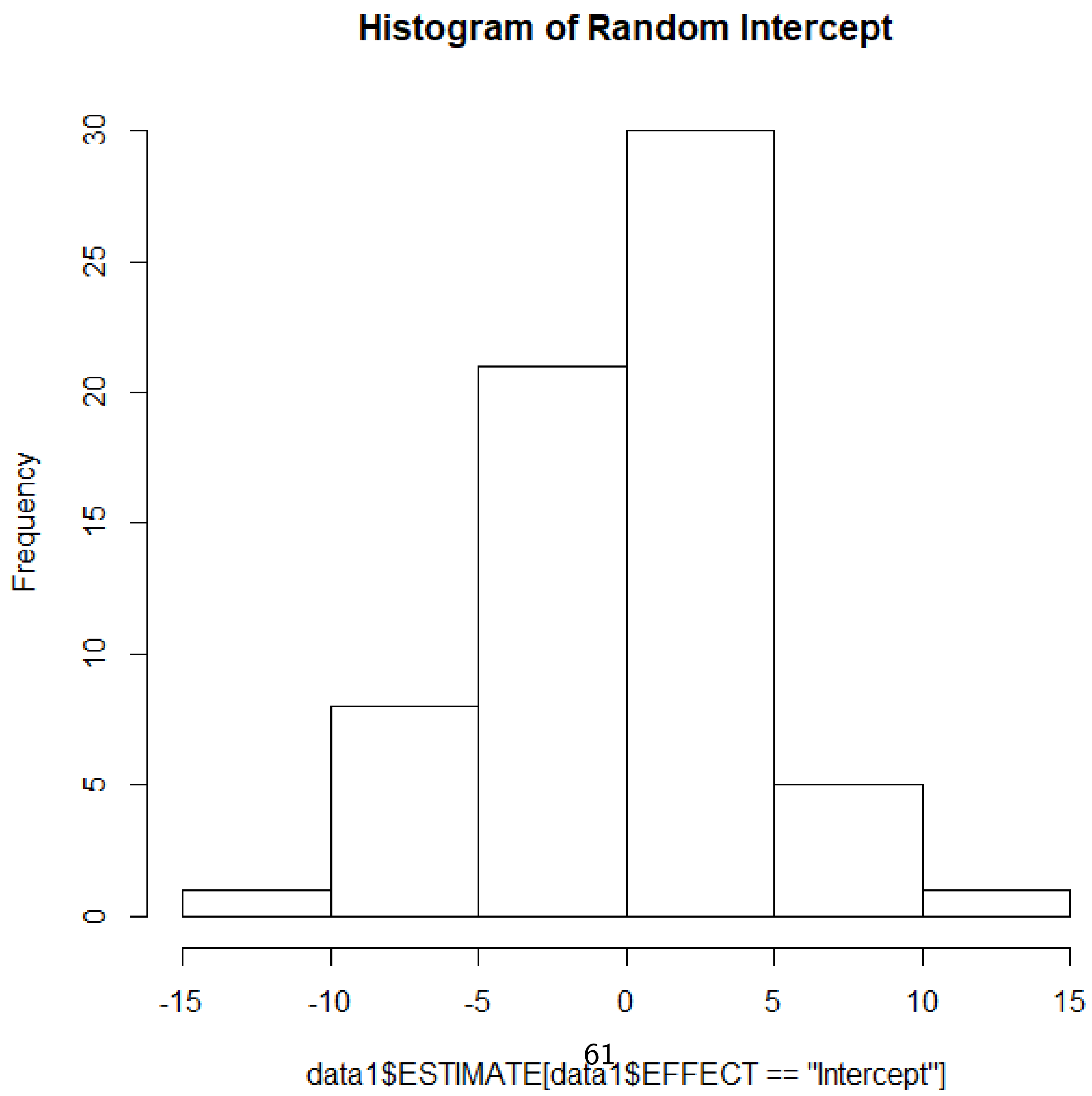
Residual Diagnostics

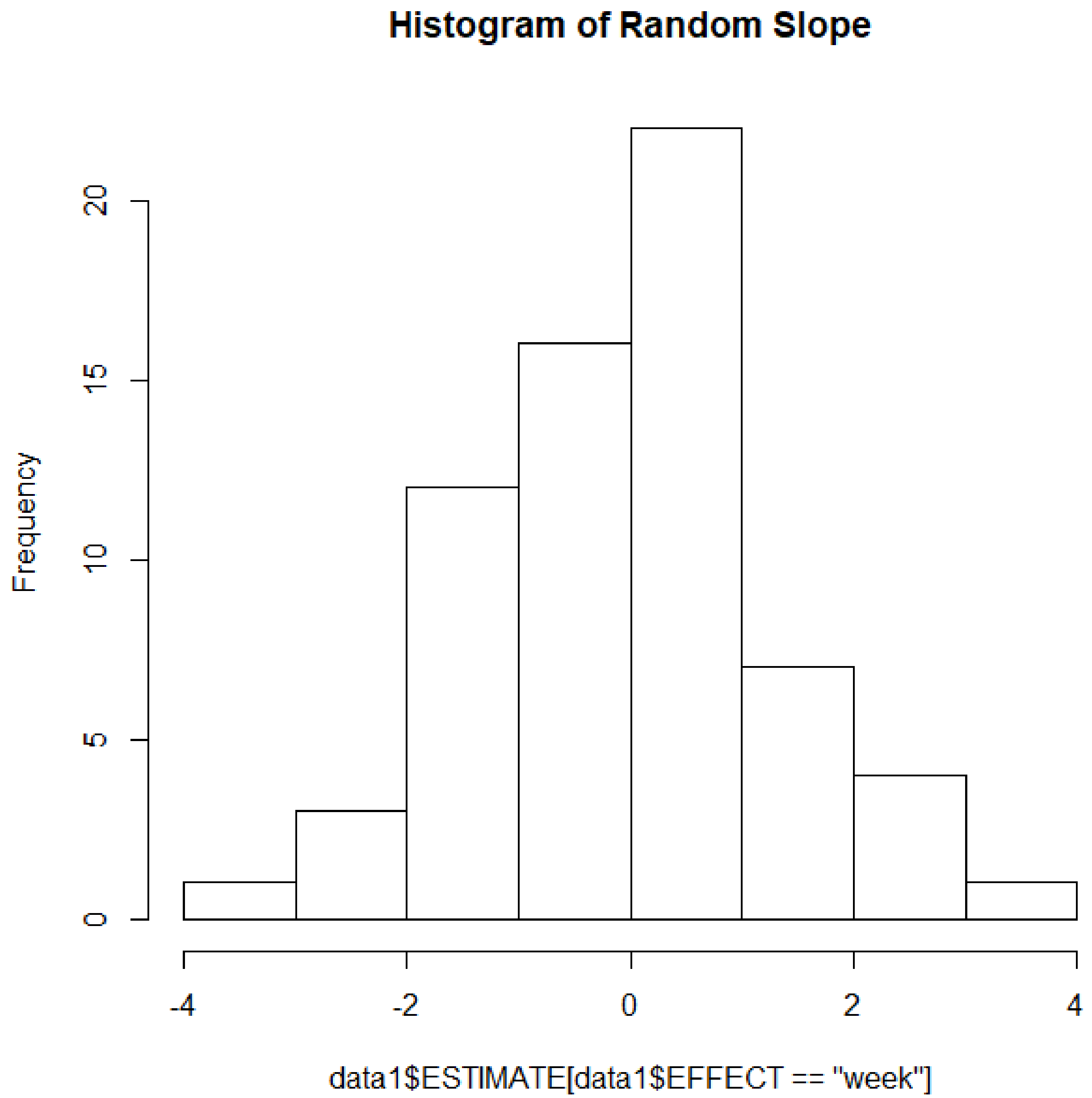
- Can use Empirical Bayes' Estimator $\widehat{U}_i$ for diagnostic purpose
  - check whether $U_i \sim \ N(0, G)$
  - check whether residual $Z_{ij} \sim \ N(0, \tau^2)$

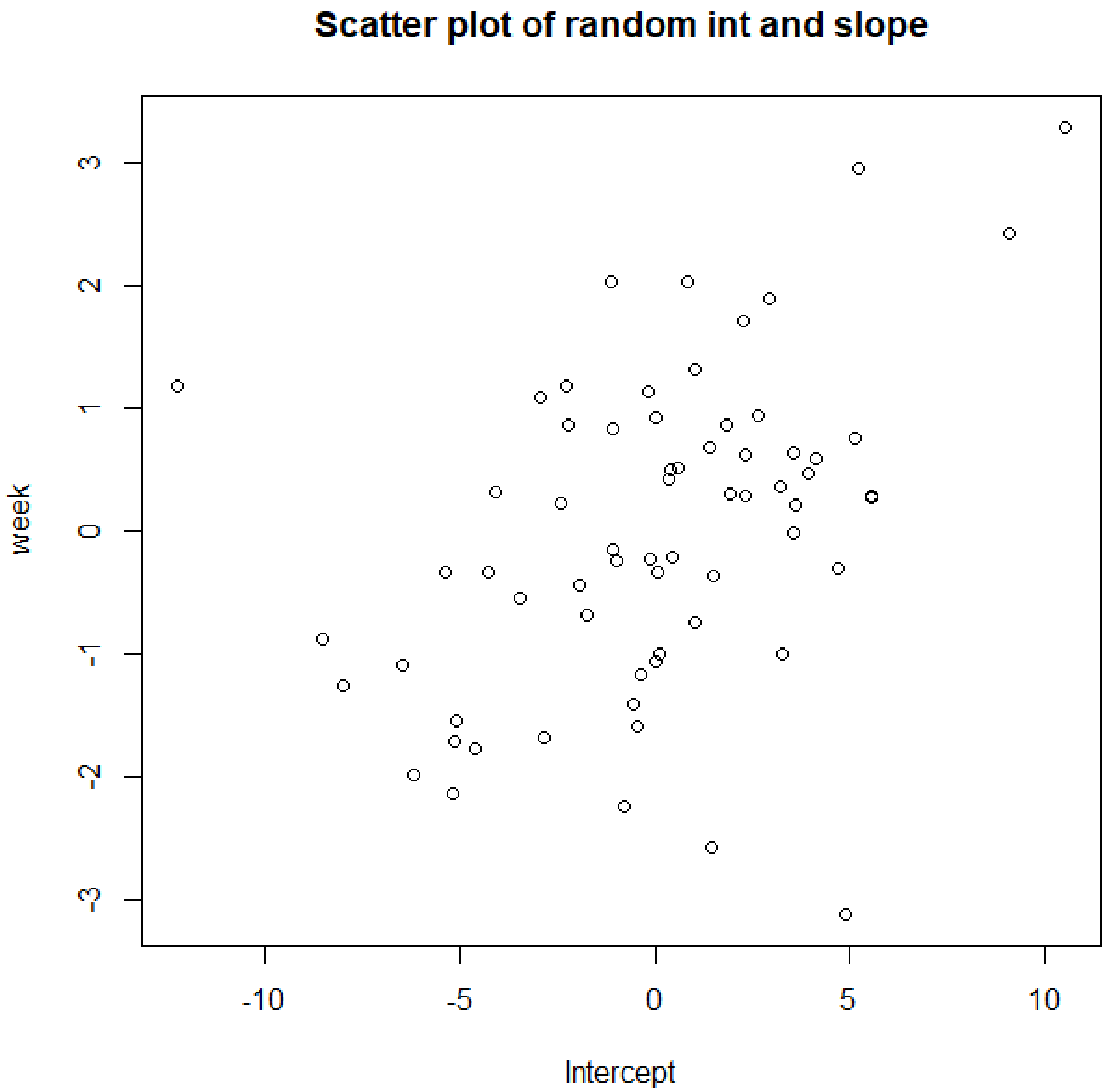
- R Code for generating residual plots (use results from model with $\beta_{BS} = \beta_{WS}$):

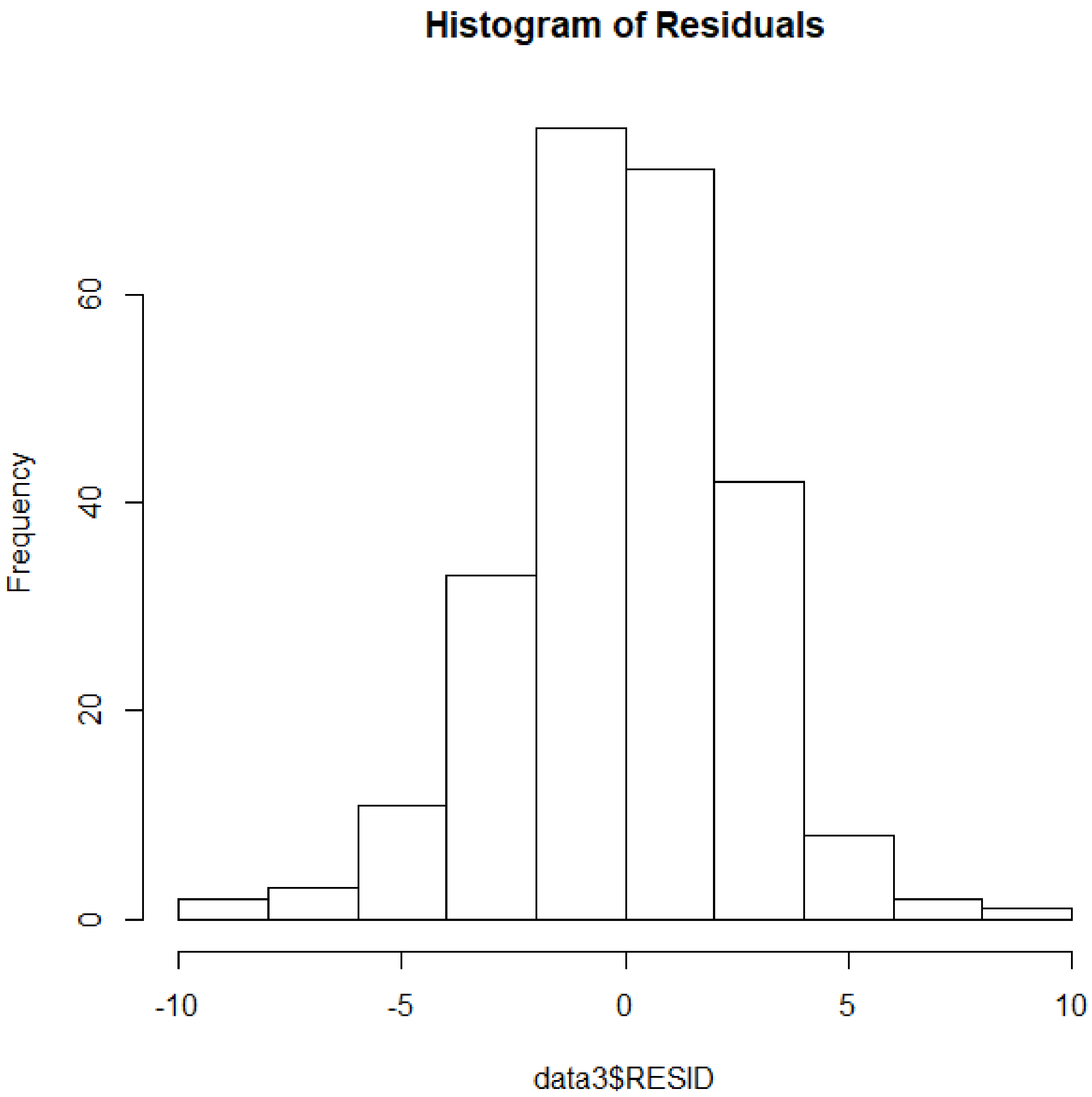
```
library(ggplot2)

#Input SAS data sets
library(foreign)
sashome <- "C:/Program Files/SASHome/SASFoundation/9.4"
datapath="Yourpath"
data1<-read.ssd(datapath,"randeff3",sascmd=file.path(sashome,"sas.exe"))
data2<-read.ssd(datapath,"popmean3",sascmd=file.path(sashome,"sas.exe"))
data3<-read.ssd(datapath,"predict3",sascmd=file.path(sashome,"sas.exe"))
```

```
#Plots for EB estimates of random intercept and slopes, and residuals
hist(data1$ESTIMATE[data1$EFFECT=="Intercept"],
    main="Histogram of Random Intercept")
hist(data1$ESTIMATE[data1$EFFECT=="week"],
    main="Histogram of Random Slope")
plot(data1$ESTIMATE[data1$EFFECT=="Intercept"],
    data1$ESTIMATE[data1$EFFECT=="week"],xlab="Intercept",
    ylab="week", main="Scatter plot of random int and slope")
hist(data3$RESID, main="Histogram of Residuals")
```


Histogram of Random Intercept
Frequency
0
5
10
15
20
25
30
-15
-10
-5
0
5
10
15
data1$ESTIMATE[data1$EFFECT == "Intercept"]


Histogram of Random Slope
Frequency
0
5
10
15
20
-4
-2
0
2
4
data1$ESTIMATE[data1$EFFECT == "week"]


Scatter plot of random int and slope
week
3
2
1
0
-1
-2
-3
-10
-5
0
5
10
Intercept



Histogram of Residuals
Frequency
0
20
40
60
-10
-5
0
5
10
data3$RESID

General Guidelines for Model Building for Longitudinal Data Analysis

- Step 1 (Exploratory analysis): Select a flexible mean structure $X\beta$, and use ordinary least squares (OLS) method
  – can use an over-elaborated (very flexible) model for the mean profile.

- Step 2 (find best Covariance structure): Use OLS residuals $Y - X\hat{\boldsymbol{\beta}}$ to help decide covariance structure, including random effects and residual covariance structure.
  – candidate models based on exploratory analysis
  – select final Covariance structure, eg, LRT (based on ReML/ML for mixed-effects model)/AIC/BIC/etc,

- Step 3 (find best Mean structure): Once final covariance structure has been selected, we fix the covariance structure, and select best mean structure $X\beta$.
  – eg, F-test, LRT (must fit by ML)

- Step 4: Perform residual analysis, etc.

- Step 5: Improve the final model if necessary.

- **Remark**: Fit Nested and Crossed Random Effects in R: https://www.r-bloggers.com/nested-vs-non-nested-crossed-random-effects-in-r/