STAT 8004 Lecture 7

Thursday, March 19, 2015 5:10 PM

Review: Linear mixed (effect) Model

$$Y = X\beta + Zu + \varepsilon$$

- β: capturing the "fixed" effect
  e.g. known as "treatment"
  main concern

- u: capturing the "random" effect
  - cannot be controlled
  - it is known as information in the dataset
  - Affecting the statistical analysis

- impact from the random effect is to incorporate the dependence/correlation between measurements (responses)

e.g. Penicillin example
five "random" batches as random effect.

$$y_{ij} = \mu + \alpha_i + \boxed{u_j} + \varepsilon_{ij}$$

jth batch of the ith process
$j=1,\cdots,5$
$i=1,\cdots,4$

$u_j \sim N(0, \sigma_u^2)$
$\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$

$$Y = X\beta + Zu + \varepsilon$$

$$E(Y) = X\beta$$

$$\Rightarrow Var(Y) = \Sigma = \boxed{ZGZ^T + R}$$

$$\begin{pmatrix} u \\ \varepsilon \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} G & 0 \\ 0 & R \end{pmatrix} \right)$$

- Pigment example
  - only one production process
  - b barrels sampled
  - s subsample for each barrel
  - $\underline{\underline{r}}$ splitting of each sub-sample

$$y_{ijk} = \underline{\underline{\mu}} + \boxed{\beta_i + \delta_{ij}} + \underline{\underline{\varepsilon_{ijk}}}$$

k's splitting
j's sub-sample in
i's barrel.

fixed

$\beta_i \sim N(0, \sigma_\beta^2)$
$\delta_{ij} \sim N(0, \sigma_\delta^2)$
$\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$

- Inference $\begin{cases} \text{fixed effect} \begin{cases} \text{estimation (set, point)} \\ \text{testing} \end{cases} \end{cases}$

- Inference
  - fixed effect
    - estimation (set, point)
    - testing
  - variance component: inference
  - prediction of random effect

- REML for variance component
  - computational advantage
  - less biased than the MLE

$\underline{\underline{Y}}$ $\quad (\Sigma = ZGZ^T + R)$

$\underline{\underline{V}} = (V_1, V_2 \cdots V_p)$

e.g. $\rightarrow V = (\underline{\underline{\sigma_u^2}}, \underline{\underline{\sigma_\varepsilon^2}})$

$V = (\underline{\underline{\sigma_\beta^2}}, \underline{\underline{\sigma_s^2}}, \underline{\underline{\sigma_\varepsilon^2}})$

Likelihood

$\rightarrow L(Y \mid X\beta, V)$

restricted likelihood

$\rightarrow \underline{\underline{L_R(\underline{\underline{BY}} \mid V)}}$

- BY is invariant of X

$\left(B = M(I - P_X) \right.$
$\left. \text{rank}(M) = n - \text{rank}(X) \right)$

$\Downarrow$

$\hat{V}_{REML}$

$\Downarrow$

$\hat{G}, \hat{R} \Rightarrow \hat{\Sigma} = Z\hat{G}Z^T + \hat{R}$

(? testing, confidence set for V)

- Point estimation for the fixed effect

$$Y \sim N(X\beta, \Sigma)$$

GLS says for estimable function $c^T\beta$ $(c^T = a^T X)$

$$\underline{\underline{\widehat{c^T\beta}}} = c^T (X^T \Sigma^{-1} X)^- X^T \Sigma^{-1} Y \leftarrow$$

is BLUE

$$\text{var}(\widehat{c^T\beta}) = c^T (X^T \Sigma^{-1} X)^- c$$

$\widehat{c^T\beta}$ is normally distributed.

Now, Σ is unknown, we have REML estimator $\hat{\Sigma}$.

- REML for estimable function $c^T\beta$

$$\widehat{\widehat{c^T\beta}} = c^T (X^T \hat{\Sigma}^{-1} X)^- X^T \hat{\Sigma}^{-1} Y$$

$$\text{var}(\widehat{\widehat{c^T\beta}}) = \underline{c^T (X^T \Sigma^{-1} X)^- c} \quad \text{approximately}$$

$$\widehat{\widehat{c^T\beta}} \overset{AN}{\sim} \left(c^T\beta, \underbrace{c^T (X^T \Sigma^{-1} X)^- c} \right)$$

$$\widehat{Var}(\widehat{c^T\beta}) = c^T(X^T\hat{\Sigma}^{-1}X)^- c$$

$$\Rightarrow \quad \widehat{c^T\beta} \pm z_{\alpha/2}\sqrt{c^T(X^T\hat{\Sigma}^{-1}X)^- c}$$

will be an approximate $1-\alpha$ level confidence interval for $c^T\beta$.

(plug-in principle).

- Generalization to $C\beta$ for $C = \begin{pmatrix} c_1^T \\ c_2^T \\ \vdots \\ c_d^T \end{pmatrix}$, estimable vector, is also easy.

—x—

- (REML $\hat{V} \to \hat{\Sigma} \to \underline{\underline{\widehat{c^T\beta}}}$)

- random effect u,
  - prediction of u, given Y.

Under normal assumption $\begin{pmatrix} u \\ \varepsilon \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} G & 0 \\ 0 & R \end{pmatrix}\right)$

$$\begin{pmatrix} u \\ Y \end{pmatrix} = \begin{pmatrix} u \\ X\beta + Zu + \varepsilon \end{pmatrix} = \begin{pmatrix} 0 \\ X\beta \end{pmatrix} + \underbrace{\begin{pmatrix} I & 0 \\ Z & I \end{pmatrix}\begin{pmatrix} u \\ \varepsilon \end{pmatrix}}$$

$$\sim N\left(\begin{pmatrix} 0 \\ X\beta \end{pmatrix}, \begin{pmatrix} G & GZ^T \\ ZG^T & ZGZ^T + R \end{pmatrix}\right)$$

(Theorem 4.4a of R&S).

$$\underline{\underline{\hat{u}}} = \hat{E}(u|Y) = GZ^T(ZGZ^T+R)^{-1}(Y - \widehat{X\beta}) \in BLUP$$

[Approximate $X\beta$ by
$$\widehat{X\beta} = X(X^T\Sigma^{-1}X)^- X^T\Sigma^{-1}Y$$
(Theorem 4.4d of R&S)]

$$\hat{\hat{u}} = \underline{\widehat{E(u|Y)}} = \hat{G}Z^T\hat{\Sigma}^{-1}(Y - \widehat{X\beta})$$

$$\underbrace{\boxed{\widehat{Var}(\hat{u})}}_{(*)} = \hat{G}Z^T\hat{\Sigma}^{-1}(I-\hat{P})\hat{\Sigma}(I-\hat{P})\hat{\Sigma}^{-1}Z\hat{G}^T$$

$$\hat{P} = X(X^T\hat{\Sigma}^{-1}X)^- X^T\hat{\Sigma}^{-1} \quad // \quad \hat{\Omega}$$

use $(*)$, we can do inference for $\underline{\underline{u}}$ using the estimation $\hat{\hat{u}}$ & $\widehat{Var}(\hat{u})$

e.g. prediction interval for u_j

- General linear functional of the fixed & random effects
  - $\ell = c^T\beta + s^Tu$

- $\ell = c^T\beta + s^T u$

- $\widehat{\ell} = c^T\widehat{\beta} + s^T\widehat{u}$ ← point estimation

is the prediction for ℓ.

$$\underline{\widehat{var}(\widehat{\ell})} = c^T(X^T\hat{\Sigma}^{-1}X)^- c + s^T\hat{\Omega}s$$

$$- 2c^T(X^T\hat{\Sigma}^{-1}X)^- X^T\hat{\Sigma}^{-1}Z\hat{G}s \quad \text{(approximately)}$$

$var\begin{pmatrix} c^T\beta \\ \hat{u} \end{pmatrix}$ as a linear function of $\underline{Y}$

$(\widehat{\ell}, \widehat{var}(\widehat{\ell}))$ for inference for ℓ).

$(\hat{v} \to \hat{\Sigma} \to c^T\widehat{\beta} \to \widehat{\ell})$.

$\to \widehat{u}$

- Inference for $v = (\sigma_1^2, \sigma_2^2, \cdots \sigma_p^2)$

$\hat{v}_{REML} = (\hat{\sigma}_1^2, \hat{\sigma}_2^2, \cdots \hat{\sigma}_p^2)$

- General properties of MLE
  - Regularity conditions
    - parameter space ($H_0: \sigma^2 \neq 0$)
    - identifiability
    - Smoothness of the likelihood
    - Existence of expectations

- $\ell(\theta)$ ← log-likelihood function for θ

$0 = \dfrac{\partial \ell(\theta)}{\partial \theta}$ Score function

Score equation

$\hat{\theta}$ is the root of the score equation.

- $E\left\{\dfrac{\partial^2 \ell(\theta)}{\partial\theta\partial\theta^T}\right\} = I(\theta)$ information matrix.

- $\underline{\hat{\theta} \overset{AN}{\sim} N(\theta, I^{-1}(\theta))}$

$\sqrt{n}\, I^{1/2}(\theta)(\hat{\theta} - \theta) \xrightarrow{d} N(0, I)$

as $n \to \infty$ (n is sample size)

- $\hat{\theta}$ has approximate optimality.

- Result: if $\hat{\theta}$ is the MLE, then

$g(\hat{\theta})$ is the MLE for $g(\theta)$

e.g. $\hat{\beta}$ is the OLS in G-M model with normal assumptions

then $\hat{\beta}$ is MLE

$= (\hat{\beta}_1, \hat{\beta}_2, \cdots \hat{\beta}_k)$

$c^T\hat{\beta}$, $\hat{\beta}_1/\hat{\beta}_2$, $g(\hat{\beta})$ .. are MLE

By delta method:

$$\underset{q\times 1}{g(\hat{\theta})} = g(\theta + \hat{\theta} - \theta)$$

$$= \underset{q\times 1}{g(\theta)} + \underline{\underline{g'(\theta)}}(\hat{\theta}-\theta)\{1+o_p(1)\}$$

$$g'(\theta) = \underset{q\times k}{\frac{\partial g(\theta)}{\partial \theta^T}}$$

$$g(\hat{\theta}) \overset{AN}{\sim} \left(g(\theta), \underbrace{g'(\theta) I^{-1}(\theta)(g'(\theta))^T}\right)$$

$$\sqrt{n}\left(g'(\theta) I^{-1}(\theta)(g'(\theta))^T\right)^{-1/2}(g(\hat{\theta}) - g(\theta))$$

$$\xrightarrow{d} N(0, I)$$

as $n \to \infty$.

- use this for inference of $g(\theta)$.

($\underbrace{\hat{\sigma}_i^2 \pm z_{\alpha/2}\sqrt{\hat{V}(\hat{\sigma}_i^2)}}$ is an approximate $1-\alpha$ level C.I for σ_i^2, $i, \cdots, p$).

- Based on the MLE, we can do testing

Wald's $\quad \hat{\theta} - \theta \overset{AN}{\sim} N(0, I^{-1}(\theta))$

$$\underbrace{(I^{1/2})(\hat{\theta}-\theta) \overset{AN}{\sim} N(0, I)}$$

$$(\hat{\theta}-\theta)^T I^{1/2} I^{1/2} (\hat{\theta}-\theta) \sim \chi^2_{\dim(\theta)}$$

$$\|$$

$$W = (\hat{\theta} - \theta)^T I(\theta)(\hat{\theta}-\theta).$$

$H_0: \theta = \theta_0$. $\quad$ using Wald's

reject if $W > \chi^2_{\dim(\theta), 1-\alpha}$.

$C_\alpha = \{\theta : W(\theta) \le \chi^2_{\dim(\theta), 1-\alpha}\}$

is an approximate $1-\alpha$ confidence set for θ.

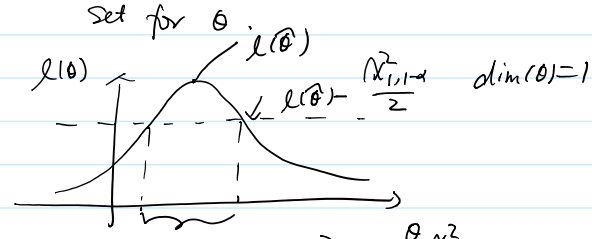
- likelihood ratio test:

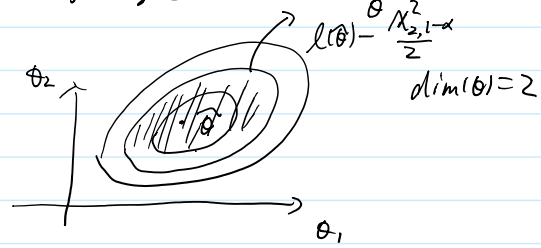
$$2(\ell(\hat{\theta}) - \ell(\theta))$$

$$\sim \chi^2_{\dim(\theta)}$$

$$C_{2,\alpha}(\theta) = \{ \theta : \ 2(\ell(\hat\theta) - \ell(\theta)) \le \chi^2_{\dim(\theta), 1-\alpha} \}$$

is also an approximate $1-\alpha$ confidence set for θ.





- $\theta = (\theta_1^T, \theta_2^T)^T$, suppose the interest is on θ_1

$$\ell(\hat\theta_2(\theta_1)) \leftarrow \text{MLE for } \theta_2 \text{ with given } \theta_1$$

$$\sim \chi^2_{\dim(\theta_1)}$$

profile likelihood.

- Application of the MLE theory for inference of V

$$V = (\sigma_1^2, \cdots \sigma_p^2) \leftarrow \text{variance component}$$
$$\| \quad (e^{\gamma_1}, \cdots e^{\gamma_2})$$

$$\gamma = (\gamma_1, \cdots \gamma_p)$$
$$\gamma_i = \log(\sigma_i^2)$$

Instead of working on $\ell(V)$, one can work on the likelihood $\ell(\gamma)$

$$\underline{\underline{\hat\gamma}} = (\log \hat\sigma_1^2, \cdots \log \hat\sigma_p^2)$$

is the MLE

$$\ell_R(\gamma) = \ell_R(e^{\gamma_1}, \cdots, e^{\gamma_p})$$

$$\to I(\gamma)$$

let $Q = I^{-1}(\gamma)$ is the approximate variance-covariance matrix for γ

$$\hat\gamma_i \pm z_{\alpha/2}\sqrt{q_{ii}}, \quad q_{ii} \text{ is the } i^{th} \text{ diagonal component of } Q$$

then $\left(e^{\hat{\gamma}_i - z_{\alpha/2}\sqrt{\tau_{ii}}}, \quad e^{\hat{\gamma}_i + z_{\alpha/2}\sqrt{\tau_{ii}}} \right)$

is an approximate $1-\alpha$ confidence interval for σ_i^2

- Don't have to worry about negative variance by working on γ.

- Small sample performance is better.

$\left(V \to \gamma \to \hat{\gamma} \to \text{interval for } \sigma_i^2 \ (i=1,\cdots,p) \right)$

$\hat{\Sigma}, \ \widehat{\mathrm{var}}\hat{\gamma}, \ \widehat{\ell}$

$Y = X\beta + Zu + \varepsilon$ (Lmel

—x—

Printout

Thursday, March 19, 2015 5:18 PM

Linear Mixed Models: R Examples

STAT 8004 *2015* *1/18*

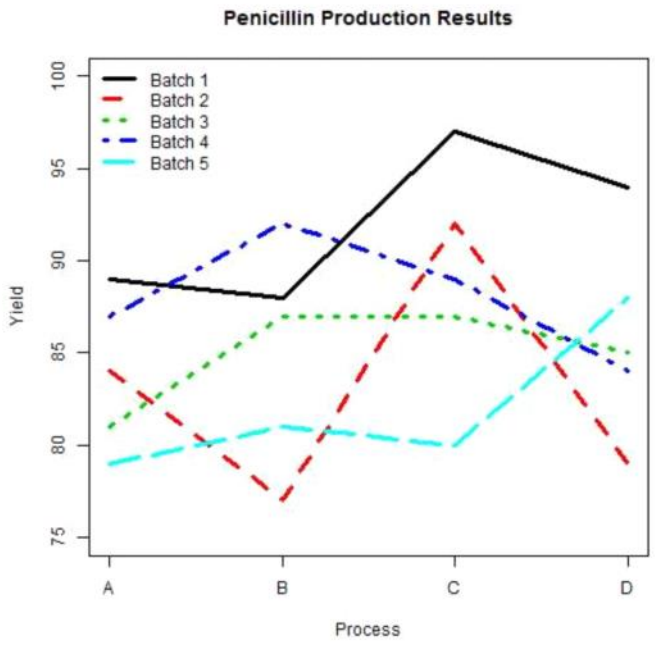
Penicillin Example

$$y_{ij} = \mu + \alpha_i + u_j + \varepsilon_{ij}$$

nlme lme( )

lme4 lmer( )

```
> penclln
   Batch Process Yield
1      1       A    89
2      1       B    88
.....
19     5       C    80
20     5       D    88
```


Penicillin Production Results
Batch 1
Batch 2
Batch 3
Batch 4
Batch 5
Yield
100
95
90
85
80
75
A
B
C
D
Process

Page 4

Thursday, March 19, 2015 5:18 PM

model.matrix( )

```
options(contrasts=c("contr.treatment",
    "contr.poly"))

penclln.lme <- lme(Yield ~ Process,
    random= ~ 1|Batch, data=penclln,
    method=c("REML"))
```

fixed effect

batch is random

Page 5

Thursday, March 19, 2015 5:18 PM

```
> model.matrix(penclln.lme)
   (Intercept) ProcessB ProcessC ProcessD
1            1        0        0        0
2            1        1        0        0
3            1        0        1        0
4            1        0        0        1
5            1        0        0        0
6            1        1        0        0
7            1        0        1        0
8            1        0        0        1
9            1        0        0        0
10           1        1        0        0
11           1        0        1        0
12           1        0        0        1
13           1        0        0        0
14           1        1        0        0
15           1        0        1        0
16           1        0        0        1
17           1        0        0        0
18           1        1        0        0
19           1        0        1        0
20           1        0        0        1
```

```
> summary(penclln.lme)
Linear mixed-effects model fit by REML
 Data: penclln
       AIC      BIC    logLik
  115.8299 120.4655 -51.91497

Random effects:
 Formula: ~1 | Batch
        (Intercept) Residual
StdDev:      3.4339 4.339739
```

σ_u σ_ε

Page 7

Thursday, March 19, 2015 5:18 PM

```
Fixed effects: Yield ~ Process
            Value Std.Error DF  t-value p-value
(Intercept)    84  2.474874 12 33.94113  0.0000
ProcessB        1  2.744692 12  0.36434  0.7219
ProcessC        5  2.744692 12  1.82170  0.0935
ProcessD        2  2.744692 12  0.72868  0.4802
 Correlation:
         (Intr) PrcssB PrcssC
ProcessB -0.555
ProcessC -0.555  0.500
ProcessD -0.555  0.500  0.500

Standardized Within-Group Residuals:
       Min         Q1        Med         Q3        Max
-1.4151575 -0.5017350 -0.1643840  0.6829939  1.2836503

Number of Observations: 20
Number of Groups: 5
```

$\widehat{Var}(\widehat{c^T\beta})$

lm( )

VCOV for fixed effect

Page 8

Thursday, March 19, 2015 5:18 PM

M

```
> names(penclln.lme)
 [1] "modelStruct"  "dims"         "contrasts"
 [4] "coefficients" "varFix"       "sigma"
 [7] "apVar"        "logLik"       "numIter"
[10] "groups"       "call"         "terms"
[13] "method"       "fitted"       "residuals"
[16] "fixDF"        "na.action"    "data"
```

Page 9

Thursday, March 19, 2015 5:18 PM

```
> ranef(penclln.lme)
  (Intercept)
1   4.2878788
2  -2.1439394
3  -0.7146465
4   1.4292929
5  -2.8585859
```

✓ prediction for u

$\hat{Var}(\hat{u}-u)$?

Page 10

Thursday, March 19, 2015 5:18 PM

```
> penclln.lme$coef
$fixed
(Intercept)    ProcessB    ProcessC    ProcessD
         84           1           5           2

$random
$random$Batch
  (Intercept)
1   4.2878788
2  -2.1439394
3  -0.7146465
4   1.4292929
5  -2.8585859
```

Page 11

Thursday, March 19, 2015 5:18 PM

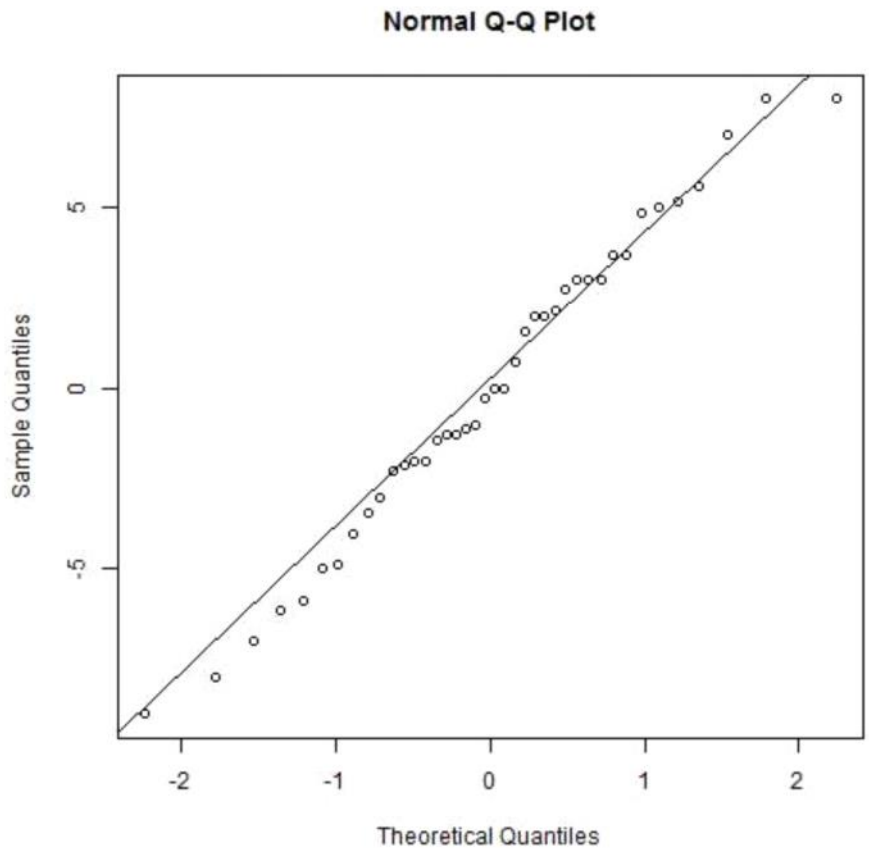
```
> intervals(penclln.lme)
Approximate 95% confidence intervals

 Fixed effects:
                 lower est.    upper
(Intercept) 78.6077134   84 89.39229
ProcessB    -4.9801698    1  6.98017
ProcessC    -0.9801698    5 10.98017
ProcessD    -3.9801698    2  7.98017
attr(,"label")
[1] "Fixed effects:"

 Random Effects:
  Level: Batch
                  lower   est.    upper
sd((Intercept)) 1.285273 3.4339 9.174443

 Within-group standard error:
   lower     est.    upper
2.908794 4.339739 6.474620
```

σ_u

σ_ε

Page 12

Thursday, March 19, 2015 5:18 PM



Pigment example

$$y_{ijk} = \mu + \beta_i + \delta_{ij} + \varepsilon_{ijk}$$

```
> pigment
   Batch Sample Test  Y
1      1      1    1 40
2      1      1    2 39
3      1      2    1 30
4      1      2    2 30
5      2      1    1 26
....
56    14      2    2 25
57    15      1    1 39
58    15      1    2 37
59    15      2    1 26
60    15      2    2 28
```

Page 14

Thursday, March 19, 2015 5:18 PM

```
library(lme4)
pmt.lme2<-lmer(Y~1+
 (1|Batch)+(1|Batch:Sample),data=pigment)

pmt.lme <- lme(Y ~ 1,
    random=~1|Batch/Sample, data=pigment,
    method=c("REML"))
```

Page 15

Thursday, March 19, 2015 5:18 PM

```
> summary(pmt.lme2)
Linear mixed model fit by REML ['lmerMod']
Formula: Y ~ 1 + (1 | Batch) + (1 | Batch:Sample)
   Data: pigment

REML criterion at convergence: 292.3

Scaled residuals:
     Min       1Q   Median       3Q      Max
-1.73260 -0.46499  0.00975  0.52279  1.58243

Random effects:
 Groups       Name        Variance Std.Dev.
 Batch:Sample (Intercept) 28.5333  5.3417
 Batch        (Intercept)  7.1280  2.6698
 Residual                  0.9167  0.9574
Number of obs: 60, groups:  Batch:Sample, 30; Batch, 15

Fixed effects:
            Estimate Std. Error t value
(Intercept)   26.783      1.201   22.31
```

Σ

$\sigma^2_{\delta_2}$

σ^2_{β}

σ^2_{ε}

$\hat{\mu}$

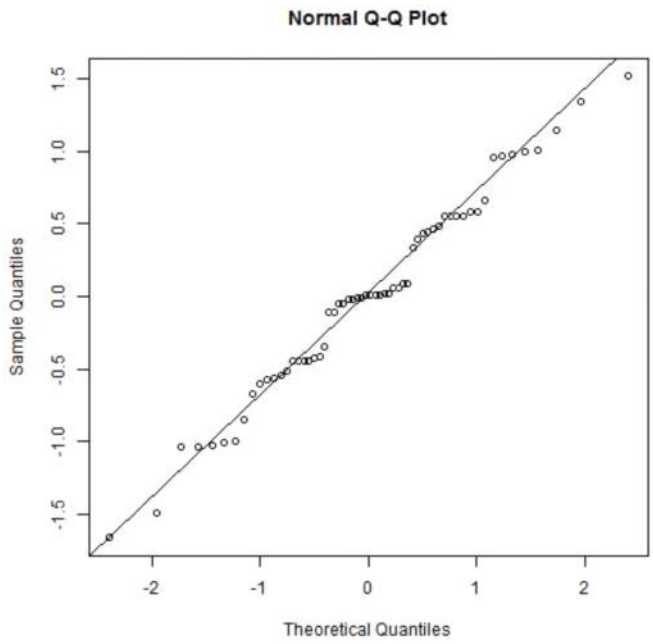
Page 16

Thursday, March 19, 2015 5:18 PM

```
> pmt.lme2<-lmer(Y~1+(1|Batch)+(1|Batch:Sample),data=pigment)
> pmt.lme3<-lmer(Y~1+(1|Batch),data=pigment)
> anova(pmt.lme3,pmt.lme2)
refitting model(s) with ML (instead of REML)
Data: pigment
Models:
pmt.lme3: Y ~ 1 + (1 | Batch)
pmt.lme2: Y ~ 1 + (1 | Batch) + (1 | Batch:Sample)
         Df    AIC    BIC  logLik deviance  Chisq Chi Df
pmt.lme3  3 376.81 383.09 -185.41   370.81
pmt.lme2  4 302.43 310.81 -147.22   294.43 76.378      1
         Pr(>Chisq)
pmt.lme3
pmt.lme2  < 2.2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

$\widehat{Var}(\hat{u} - u)$?

```
> confint(pmt.lme2)
Computing profile confidence intervals ...
                 2.5 %    97.5 %
.sig01       3.8512670  7.606216
.sig02       0.0000000  5.659430
.sigma       0.7581481  1.262118
(Intercept) 24.3562433 29.210423
```

Page 18

Thursday, March 19, 2015 5:18 PM