Advanced Regression in R

Tiffany Himmel
tiffany@stat.ucla.edu

August 19th, 2010
Outline

1. Generalized Linear Models
2. Mixed Effects Models
3. Resources
Objective

Give an overview of different generalized linear models as well mixed models. On the way, we will also discuss

- Statistical significance
- Diagnostics
- Interpretation in context

R will be used, however most statistical softwares contain comparable packages (STATA, SAS, SPS)
1 Generalized Linear Models
   • Poisson Regression
   • Logistic Regression
   • Multinomial Logistic

2 Mixed Effects Models

3 Resources

Tiffany Himmel  tiffany@stat.ucla.edu
Generalized Linear Models

Objective

To model the relationship between our predictors $X_1, X_2, \ldots, X_p$ and a binary, count or other non-normal variable $Y$ which cannot be modeled with standard regression.

Possible Applications

- Medical Data
- Internet Traffic
- Survival Data
Poisson Regression

Example: Randomized Controlled Trial (Taken from Dobson, 1990)

- Outcome is in the form of counts rather than a continuous numeric variable
- Explore the relationship between two categorical variables

<table>
<thead>
<tr>
<th>Treatment</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>13</td>
<td>12</td>
</tr>
</tbody>
</table>
### Poisson Regression

**Example: Randomized Controlled Trial (Taken from Dobson, 1990)**

```r
1 counts <- c(18, 17, 15, 20, 10, 20, 25, 13, 12)
2 outcome <- gl(3, 1, 9)
3 treatment <- gl(3, 3)
4 print(d.AD <- data.frame(treatment, outcome, counts))
```

<table>
<thead>
<tr>
<th>treatment</th>
<th>outcome</th>
<th>counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>12</td>
</tr>
</tbody>
</table>

Tiffany Himmel  tiffany@stat.ucla.edu

Regression in R II  UCLA SCC
Exploratory Plots

Example: Randomized Controlled Trial (Taken from Dobson, 1990)

```r
contin.table <- xtabs(counts ~ treatment + outcome)
mosaicplot(contin.table, main = "Mosaic Plot")
dimnames(contin.table)$outcome <- c("Out_1", "Out_2", "Out_3")
dimnames(contin.table)$treatment <- c("Treat_1", "Treat_2", "Treat_3")
dotchart(contin.table, main = "Dot Chart")
```
Fitting a GLM

Example: Randomized Controlled Trial (Taken from Dobson, 1990)

```r
1 glm.D93 <- glm(counts ~ outcome + treatment,
                   family = poisson())
2 anova(glm.D93)
```

Analysis of Deviance Table

Model: poisson, link: log

Response: counts

Terms added sequentially (first to last)

<table>
<thead>
<tr>
<th>Df</th>
<th>Deviance</th>
<th>Resid. Df</th>
<th>Resid. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>NULL</td>
<td>8</td>
<td>10.5814</td>
<td></td>
</tr>
<tr>
<td>outcome</td>
<td>2</td>
<td>5.4523</td>
<td>6</td>
</tr>
<tr>
<td>treatment</td>
<td>2</td>
<td>8.882e-15</td>
<td>4</td>
</tr>
</tbody>
</table>
Fitting a GLM II

Example: Randomized Controlled Trial (Taken from Dobson, 1990)

```r
summary(glm.D93)
```

Coefficients:
```
                Estimate Std. Error  z value Pr(>|z|)
(Intercept)  3.045e+00  1.709e-01  17.815  <2e-16 ***
outcome2    -4.543e-01  2.022e-01  -2.247  0.0246 *
outcome3    -2.930e-01  1.927e-01  -1.520  0.1285
```

---

Signif. codes:  0 ***  0.001 **  0.01 *  0.05 .  0.1  1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 10.5814 on 8 degrees of freedom
Residual deviance: 5.1291 on 4 degrees of freedom
AIC: 56.761
Chi-Squared Test

Example: Randomized Controlled Trial (Taken from Dobson, 1990)

```r
anova(glm.D93, test = "Chisq")
```

Analysis of Deviance Table

Model: poisson, link: log

Response: counts

Terms added sequentially (first to last)

| Term      | Df | Deviance | Resid. Df | Resid. Dev | P(>|Chi|) |
|-----------|----|----------|-----------|------------|---------|
| NULL      |    | 10.5814  | 8         | 10.5814    |         |
| outcome   | 2  | 5.4523   | 6         | 5.1291     | 0.0655  |
| treatment | 2  | 8.882e-15| 4         | 5.1291     | 1.0000  |
Logistic Regression

- Like Poisson, outcomes are not continuous
- Outcome is in the form of binary variables
  - yes/no, dead/alive, male/female
  - Makes estimates based on % yes/no
  - x-covariate continuous (usually)
- Very popular in medical research

Tiffany Himmel  tiffany@stat.ucla.edu
Logistic Regression

Example: Insecticide Concentration (Taken from Faraway, 2006)

```r
library(faraway)
data(bliss)
bliss

dead alive conc
1  2  28  0
2  8  22  1
3 15  15  2
4 23  7  3
5 27  3  4
```
Exploratory Plots

Example: Insecticide Concentration (Taken from Faraway, 2006)

Thinking percentages...

1. `attach(bliss)`
2. `survival <- (alive / (alive + dead))`
3. `bliss <- data.frame(bliss, survival)`
4. `plot(survival ~ conc, main = "Insecticide")`
Fitting a Logistic Model
Example: Insecticide Concentration (Taken from Faraway, 2006)

```r
modl <- glm(cbind(dead, alive) ~ conc, family = binomial, data = bliss)
summary(modl)
```

Call:
glm(formula = cbind(dead, alive) ~ conc, family = binomial, data = bliss)

Deviance Residuals:

```
  1     2     3     4     5
-0.4510 0.3597 0.0000 0.0643 -0.2045
```

Coefficients:

```
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.3238    0.4179  -5.561 2.69e-08 ***
conc          1.1619    0.1814   6.405 1.51e-10 ***
```

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 64.76327 on 4 degrees of freedom
Residual deviance: 0.37875 on 3 degrees of freedom
AIC: 20.854

Number of Fisher Scoring iterations: 4
Other Link Functions

Example: Insecticide Concentration (Taken from Faraway, 2006)

1. \[ \text{modp} <- \text{glm}( \text{cbind(dead, alive)} \sim \text{conc}, \text{family} = \text{binomial(link=probit)}, \text{data} = \text{bliss}) \]

2. \[ \text{modc} <- \text{glm}( \text{cbind(dead, alive)} \sim \text{conc}, \text{family} = \text{binomial(link=cloglog)}, \text{data} = \text{bliss}) \]
Other Link Functions

Example: Insecticide Concentration (Taken from Faraway, 2006)

```r
1 cbind(fitted(modl), fitted(modp), fitted(modc))
```

```
[,1]    [,2]    [,3]
1 0.08917177 0.08424186 0.1272700
2 0.23832314 0.24487335 0.2496909
3 0.50000000 0.49827210 0.4545910
4 0.76167686 0.75239612 0.7217655
5 0.91082823 0.91441122 0.9327715
```
Other Link Functions

Example: Insecticide Concentration (Taken from Faraway, 2006)

```r
x <- seq(-2, 8, .2)
pc <- 1 - exp(-exp((modc$coef[1] + modc$coef[2] * x)))
plot(x, pl, type = "l", ylab = "Probability", xlab = "Dose")
lines(x, pp, lty = 2)
lines(x, pc, lty = 5)
```
But choose carefully!

Example: Insecticide Concentration (Taken from Faraway, 2006)

```r
matplot(x, cbind(pp/pl, (1-p)/pl), type="l",
        xlab="Dose", ylab="Ratio")
```

![Graph showing the relationship between dose and ratio](image)
But choose carefully!

Example: Insecticide Concentration (Taken from Faraway, 2006)

```R
matplot(x, cbind(pc/pl, (1-pc)/(1-pl)), type="l",
       xlab="Dose", ylab="Ratio")
```

![Graph showing dose vs. ratio for insecticide concentration example](image)
Multinomial Logistic

- Like Logistic + Poisson, outcomes are not continuous
- Outcome is in the form of a categorical variable with 3 or more options
  - black/white/other, married/unmarried/divorced
  - Our example, politics: Democrat/Republican/Independent
# Multinomial Logistic

Example: American National Election Study (Taken from Faraway, 2006)

```r
1  data(nes96)
2  head(nes96)
```

<table>
<thead>
<tr>
<th>popul</th>
<th>TVnews</th>
<th>selfLR</th>
<th>ClinLR</th>
<th>DoleLR</th>
<th>PID</th>
<th>age</th>
<th>educ</th>
<th>income</th>
<th>vote</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>7</td>
<td>extCon</td>
<td>extLib</td>
<td>Con</td>
<td>strRep</td>
<td>36</td>
<td>HS $3Kminus</td>
<td>Dole</td>
</tr>
<tr>
<td>2</td>
<td>190</td>
<td>1</td>
<td>sliLib</td>
<td>sliLib</td>
<td>sliCon</td>
<td>weakDem</td>
<td>20</td>
<td>Coll $3Kminus</td>
<td>Clinton</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>7</td>
<td>Lib</td>
<td>Lib</td>
<td>Con</td>
<td>weakDem</td>
<td>24</td>
<td>BAdeg $3Kminus</td>
<td>Clinton</td>
</tr>
<tr>
<td>4</td>
<td>83</td>
<td>4</td>
<td>sliLib</td>
<td>Mod</td>
<td>sliCon</td>
<td>weakDem</td>
<td>28</td>
<td>BAdeg $3Kminus</td>
<td>Clinton</td>
</tr>
<tr>
<td>5</td>
<td>640</td>
<td>7</td>
<td>sliCon</td>
<td>Con</td>
<td>Mod</td>
<td>strDem</td>
<td>68</td>
<td>BAdeg $3Kminus</td>
<td>Clinton</td>
</tr>
<tr>
<td>6</td>
<td>110</td>
<td>3</td>
<td>sliLib</td>
<td>Mod</td>
<td>Con</td>
<td>weakDem</td>
<td>21</td>
<td>Coll $3Kminus</td>
<td>Clinton</td>
</tr>
</tbody>
</table>
Exploratory Plots

Example: American National Election Study (Taken from Faraway, 2006)

```r
1 PartyAffil <- nes96$PID
2 levels(PartyAffil) <- c("Dem", "Dem", "Ind", "Ind", "Ind", "Rep", "Rep")
3 income.categories <- c(1.5, 4, 6, 8, 9, 10.5, 11.5, 12.5, 13.5, 14.5, 16, 18.5, 21, 23.5, 27.5, 32.5, 37.5, 42.5, 47.5, 55, 67.5, 82.5, 97.5, 115)
4 Income.numeric <- income.categories[unclass(nes96$income)]
5 hist(Income.numeric, main="")
```
Useful Tables

Example: American National Election Study (Taken from Faraway, 2006)

1. \( \text{Income.Split} \leftarrow \text{cut}(\text{Income.numeric},7) \)
2. \( \text{PartyProp.by.Income} \leftarrow \text{prop.table}(\text{table}(\text{Income.Split,PartyAffil}),1) \)

<table>
<thead>
<tr>
<th>Income.Split</th>
<th>Dem</th>
<th>Ind</th>
<th>Rep</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1.39,17.6]</td>
<td>0.5402299</td>
<td>0.2356322</td>
<td>0.2241379</td>
</tr>
<tr>
<td>(17.6,33.9]</td>
<td>0.5000000</td>
<td>0.1932773</td>
<td>0.3067227</td>
</tr>
<tr>
<td>(33.9,50.1]</td>
<td>0.4409938</td>
<td>0.2236025</td>
<td>0.3354037</td>
</tr>
<tr>
<td>(50.1,66.4]</td>
<td>0.2900000</td>
<td>0.3400000</td>
<td>0.3700000</td>
</tr>
<tr>
<td>(66.4,82.6]</td>
<td>0.2820513</td>
<td>0.2628205</td>
<td>0.4551282</td>
</tr>
<tr>
<td>(82.6,98.9]</td>
<td>0.1914894</td>
<td>0.3191489</td>
<td>0.4893617</td>
</tr>
<tr>
<td>(98.9,115]</td>
<td>0.2058824</td>
<td>0.3823529</td>
<td>0.4117647</td>
</tr>
</tbody>
</table>
Exploratory Plots II
Example: American National Election Study (Taken from Faraway, 2006)

1. \texttt{Income.Split.Mids} <- \texttt{c(8,26,42,58,74,90,107)}
2. \texttt{matplot(Income.Split.Mids,PartyProp.by.Income,}
   lty=\texttt{c(1,2,3)}, \texttt{col=5\texttt{c(5,3,4)}}, \texttt{lwd=2}, \texttt{type="l"},
   \texttt{xlab="Income in 10s of Thousands"}, \texttt{ylab="Party Proportion by Income"})
3. \texttt{legend(45,.54, c("Democrat","Independent","Republican"), col=\texttt{c(5,3,4)}, lty=\texttt{c(1,2,3)})}

```
Income.Split.Mids <- c(8,26,42,58,74,90,107)
matplot(Income.Split.Mids, PartyProp.by.Income, lty=c(1,2,3), col=c(5,3,4), lwd=2, type="l", xlab="Income in 10s of Thousands", ylab="Party Proportion by Income")
legend(45,.54, c("Democrat","Independent","Republican"), col=c(5,3,4), lty=c(1,2,3))
```
Fitting a Multinomial Logistic Model
Example: American National Election Study (Taken from Faraway, 2006)

1. `library(nnet)`
2. `multinomial.model <- multinom(PartyAffil ~ Income.numeric)`
3. `summary(multinomial.model)`

Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>(Intercept)</th>
<th>Income.numeric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ind</td>
<td>-1.173592</td>
<td>0.01606080</td>
</tr>
<tr>
<td>Rep</td>
<td>-0.949674</td>
<td>0.01765140</td>
</tr>
</tbody>
</table>

Std. Errors:

<table>
<thead>
<tr>
<th></th>
<th>(Intercept)</th>
<th>Income.numeric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ind</td>
<td>0.1535317</td>
<td>0.002848357</td>
</tr>
<tr>
<td>Rep</td>
<td>0.1416261</td>
<td>0.002651135</td>
</tr>
</tbody>
</table>

Residual Deviance: 1985.485
AIC: 1993.485
Making Predictions

Example: American National Election Study (Taken from Faraway, 2006)

```r
1 predict(multinomial.model, data.frame(Income.numeric=Income.Split.Mids))
2 Party.Prob.Predictions <- predict(multinomial.model, data.frame(Income.numeric=Income.Split.Mids), type="probs")
```

<table>
<thead>
<tr>
<th></th>
<th>Dem</th>
<th>Ind</th>
<th>Rep</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5564220</td>
<td>0.1956684</td>
<td>0.2479096</td>
</tr>
<tr>
<td>2</td>
<td>0.4803742</td>
<td>0.2255528</td>
<td>0.2940730</td>
</tr>
<tr>
<td>3</td>
<td>0.4133828</td>
<td>0.2509706</td>
<td>0.3356467</td>
</tr>
<tr>
<td>4</td>
<td>0.3494149</td>
<td>0.2742923</td>
<td>0.3762928</td>
</tr>
<tr>
<td>5</td>
<td>0.2904121</td>
<td>0.2947737</td>
<td>0.4148142</td>
</tr>
<tr>
<td>6</td>
<td>0.2377033</td>
<td>0.3119688</td>
<td>0.4503278</td>
</tr>
<tr>
<td>7</td>
<td>0.1893241</td>
<td>0.3264816</td>
<td>0.4841943</td>
</tr>
</tbody>
</table>

Tiffany Himmel  tiffany@stat.ucla.edu

Regression in R II  UCLA SCC
Visualizing Our Model Predictions

Example: American National Election Study (Taken from Faraway, 2006)

```r
     \texttt{Predictions}, \texttt{lty=c}(1,2,3), \texttt{col=c}(5,3,4), \texttt{lwd}
     =2, \texttt{type="l"}, \texttt{xlab="Income in 10s of}
     \texttt{Thousands"}, \texttt{ylab="Party Proportion by Income"})

2 \texttt{legend}(45, .54, \texttt{c("Democrat","Independent","Republican")},
     \texttt{col=c}(5,3,4), \texttt{lty=c}(1,2,3))
```
Second Order Terms

Example: American National Election Study (Taken from Faraway, 2006)

```r
multinomial.model.II <- multinom(PartyAffil ~ Income.numeric + I(Income.numeric^2))
```

Call:
multinom(formula = PartyAffil ~ Income.numeric + I(Income.numeric^2))

Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>(Intercept)</th>
<th>Income.numeric</th>
<th>I(Income.numeric^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ind</td>
<td>-1.038220</td>
<td>0.00993433</td>
<td>4.752663e-05</td>
</tr>
<tr>
<td>Rep</td>
<td>-1.181813</td>
<td>0.02888725</td>
<td>-9.569798e-05</td>
</tr>
</tbody>
</table>

Std. Errors:

<table>
<thead>
<tr>
<th></th>
<th>(Intercept)</th>
<th>Income.numeric</th>
<th>I(Income.numeric^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ind</td>
<td>0.0001707178</td>
<td>0.004581055</td>
<td>5.391936e-05</td>
</tr>
<tr>
<td>Rep</td>
<td>0.0001556944</td>
<td>0.004204987</td>
<td>5.067708e-05</td>
</tr>
</tbody>
</table>

Residual Deviance: 1982.148
AIC: 1994.148
Visualizing Our Model Predictions II

Example: American National Election Study (Taken from Faraway, 2006)

1. `predict(multinomial.model.II, data.frame(Income.numeric=Income.Split.Mids))`


4. `legend(45,.54, c("Democrat", "Independent", "Republican"), col=c(5,3,4), lty=c(1,2,3))`
Multinomial Logistic ANOVA

Example: American National Election Study (Taken from Faraway, 2006)

```
1  anova(multinomial.model, multinomial.model.II)

           Model Resid. df Resid. Dev  Test Df  LR stat.  Pr(Chi)
1 Income.numeric      1884     1985.485 NA   NA      NA       NA
2 Income.numeric + I(Income.numeric^2)  1882    1982.148  1 vs 2  2  3.336716  0.1885564
```
<table>
<thead>
<tr>
<th>Preliminaries</th>
<th>Generalized Linear Models</th>
<th>Mixed Effects Models</th>
<th>Resources</th>
</tr>
</thead>
</table>

1. **Generalized Linear Models**

2. **Mixed Effects Models**
   - Definitions
   - Model Fitting
   - Assumption Validation

3. **Resources**

Tiffany Himmel  tiffany@stat.ucla.edu

Regression in R II  UCLA SCC
What is a Mixed Effects Model?

- **Fixed effect:** $\beta$ is the same over all groups/patients
- **Random Effect:** $\nu_i \sim N(0, \sigma^2_\nu)$ where each group has their own coefficient

Often used in medical or survey data

$$y_{ij} = \beta_0 + \ldots + \beta_l + \nu_{1i} + \ldots + \nu_{mi} + \epsilon_{ij}$$

subject $i = 1 \ldots n \leftarrow \# \text{ of subjects}$
observation $j = 1 \ldots n_i \leftarrow \# \text{ of observations for that subject}$
Exploratory Plots

Example: Sleep Study (Taken from Belenky et al., 2003)

1. `library(lme4)`
2. `data(sleepstudy)`
3. `xyplot(Reaction ~ Days | Subject, sleepstudy, type = c("g","p","r"), index = function(x,y) coef(lm(y ~ x))[1], xlab = "Days of sleep deprivation", ylab = "Average reaction time (ms)", aspect = "xy")`
Fitting a Mixed Effects Model
Example: Sleep Study (Taken from Belenky et al., 2003)

```r
(fm1 <- lmer(Reaction ~ Days + (Days | Subject), sleepstudy))
```

AIC BIC logLik deviance REMLdev
1756 1775 -871.8 1752 1744

Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>(Intercept)</td>
<td>612.092</td>
<td>24.7405</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Days</td>
<td>35.072</td>
<td>5.9221</td>
<td>0.066</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>654.941</td>
<td>25.5918</td>
<td></td>
</tr>
</tbody>
</table>

Number of obs: 180, groups: Subject, 18

Fixed effects:

<table>
<thead>
<tr>
<th>(Intercept)</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>251.405</td>
<td>6.825</td>
<td>36.84</td>
</tr>
<tr>
<td>Days</td>
<td>10.467</td>
<td>1.546</td>
<td>6.77</td>
</tr>
</tbody>
</table>

Correlation of Fixed Effects:

<table>
<thead>
<tr>
<th>(Intr)</th>
<th>Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days</td>
<td>-0.138</td>
</tr>
</tbody>
</table>

Tiffany Himmel  tiffany@stat.ucla.edu
Regression in R II UCLA SCC
Fitting a Mixed Effects Model II

Example: Sleep Study (Taken from Belenky et al., 2003)

```r
(fm2 <- lmer(Reaction ~ Days + (1|Subject) + (0+Days|Subject), sleepstudy))
```

AIC  BIC logLik deviance REMLdev
1754 1770 -871.8 1752 1744

Random effects:
- Groups Name Variance Std.Dev.
  - Subject (Intercept) 627.568 25.0513
  - Subject Days 35.858 5.9882
  - Residual 653.584 25.5653

Number of obs: 180, groups: Subject, 18

Fixed effects:
- Estimate Std. Error t value
  - (Intercept) 251.405 6.885 36.51
  - Days 10.467 1.559 6.71

Correlation of Fixed Effects:
- (Intr)
- Days -0.184

Tiffany Himmel  tiffany@stat.ucla.edu
Anova with Nested Models

Example: Sleep Study (Taken from Belenky et al., 2003)

```r
anova(fm1, fm2)
```

<table>
<thead>
<tr>
<th>Df</th>
<th>AIC</th>
<th>BIC</th>
<th>logLik</th>
<th>Chisq</th>
<th>Chi Df</th>
<th>Pr(&gt;Chisq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>fm2</td>
<td>5</td>
<td>1762.0</td>
<td>1778.0</td>
<td>-876.02</td>
<td>0.0609</td>
<td>1</td>
</tr>
<tr>
<td>fm1</td>
<td>6</td>
<td>1764.0</td>
<td>1783.1</td>
<td>-875.99</td>
<td>0.0609</td>
<td>1</td>
</tr>
</tbody>
</table>
Intra-Class Correlation

Example: Sleep Study (Taken from Belenky et al., 2003)

\[
\text{ICC}_1 = \frac{612.092 + 35.072}{612.092 + 35.072 + 654.941}
\]

\[
\text{ICC}_2 = \frac{627.568 + 35.858}{627.568 + 35.858 + 653.584}
\]

\[1\] 0.497

\[1\] 0.5037
Random Effects by Subject with Errors

Example: Sleep Study (Taken from Belenky et al., 2003)

```r
1 rr1 <- ranef(fm1, postVar=TRUE)
2 dotplot(rr1, scales = list(x = list(relation = 'free')))[["Subject"]]
```
Random Effects QQ Plot

Example: Sleep Study (Taken from Belenky et al., 2003)

```r
rand_eff <- ranef(fm2, drop=TRUE)[[1]]
qqmath(~ '(Intercept)' + Days, rand_eff, xlab="Normal quantiles", ylab="Random Effect Estimates", auto.key=T, outer = TRUE, scales = list(y = list(relation = "free")), layout = c(2,1), aspect = 1, type = c("g", "p"))
```

Tiffany Himmel  tiffany@stat.ucla.edu

Regression in R II  UCLA SCC
Simulated Data
Example: Sleep Study (Taken from Belenky et al., 2003)

```
1  set.seed(483)
2  simfe <- data.frame(t(apply(simulate(fm2, nsim = 100), 2, function(y) fixef(refit(fm2, y))), check.names = FALSE)
```

**Fixed effects:**

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>251.187</td>
<td>8.261</td>
<td>36.51</td>
</tr>
<tr>
<td>Days</td>
<td>10.239</td>
<td>1.346</td>
<td>6.71</td>
</tr>
</tbody>
</table>

Tiffany Himmel  tiffany@stat.ucla.edu
Regression in R II   UCLA SCC
1 Generalized Linear Models

2 Mixed Effects Models

3 Resources
   - Online Resources for R
   - References

Tiffany Himmel  tiffany@stat.ucla.edu

Regression in R II

UCLA SCC
Online Resources for R

Download R: http://cran.stat.ucla.edu

Search Engine for R: rseek.org

R Reference Card: http://cran.r-project.org/doc/contrib/Short-refcard.pdf

UCLA Statistics Information Portal:
http://info.stat.ucla.edu/grad/

UCLA Statistical Consulting Center http://scc.stat.ucla.edu
References I

- **P. Daalgard**

- **B.S. Everitt and T. Hothorn**

- **J.J. Faraway**
References II

J. Maindonald and J. Braun
Data Analysis and Graphics using R – An Example-Based Approach,

[Sheather, 2009] S.J. Sheather
A Modern Approach to Regression with R,
DOI: 10.1007/978-0-387-09608-7-3,