LING82100: homework 8 solution

```
> library(lme4)
> library(lmtest)
> d <- read.table(
+     "http://wellformedness.com/courses/LING82100/Data/elp.tsv",
+     header = TRUE
+ )
```

1 Data preparation

```
> d$RT <- log(d$RT)
> d$subjID <- as.factor(d$subjID)
> d$trial <- with(d, trial - mean(trial))
> d$length <- scale(with(d, length * length))
> d$OLD <- scale(d$OLD)
> d$sbtlx.freq <- log(d$sbtlx.freq + 1)
> d$sbtlx.basefreq <- log(d$sbtlx.basefreq + 1)
> d$sbtlx.freq <- with(d, sbtlx.freq - mean(sbtlx.freq))
> d$sbtlx.basefreq <- with(d, sbtlx.basefreq - mean(sbtlx.basefreq))
> if (with(d, cor(sbtlx.freq, sbtlx.basefreq) > .5)) {
+     d$sbtlx.basefreq <- residuals(lm(sbtlx.basefreq ~ sbtlx.freq, data = d))
+     stopifnot(all.equal(with(d, cor(sbtlx.freq, sbtlx.basefreq)), 0))
+ }
```

2 Modeling

```
> r <- lmer(RT ~ trial + length + OLD + sbtlx.freq + sbtlx.basefreq +
+           (1 | subjID) + (1 | word), data = d)
> summary(r)
...
Fixed effects:
                Estimate Std. Error t value
(Intercept)    6.558e+00  1.170e-02  560.42
trial         -2.357e-05  7.069e-07  -33.34
length         2.178e-02  1.721e-03   12.66
OLD            3.588e-02  1.759e-03   20.39
```

```
sbtlx.freq     -3.665e-02  4.222e-04  -86.80
sbtlx.basefreq -1.687e-02  6.702e-04  -25.17
...
> lrtest(update(r, . ~ . - trial), r)


Likelihood ratio test

Model 1: RT ~ length + OLD + sbtlx.freq + sbtlx.basefreq + (1 | subjID) +
    (1 | word)
Model 2: RT ~ trial + length + OLD + sbtlx.freq + sbtlx.basefreq + (1 |
    subjID) + (1 | word)
  #Df LogLik Df  Chisq Pr(>Chisq)
1   8 -36295
2   9 -35754  1 1081.8  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


> lrtest(update(r, . ~ . - length), r)


Likelihood ratio test

Model 1: RT ~ trial + OLD + sbtlx.freq + sbtlx.basefreq + (1 | subjID) +
    (1 | word)
Model 2: RT ~ trial + length + OLD + sbtlx.freq + sbtlx.basefreq + (1 |
    subjID) + (1 | word)
  #Df LogLik Df  Chisq Pr(>Chisq)
1   8 -35828
2   9 -35754  1 148.13  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


> lrtest(update(r, . ~ . - OLD), r)


Likelihood ratio test

Model 1: RT ~ trial + length + sbtlx.freq + sbtlx.basefreq + (1 | subjID) +
    (1 | word)
Model 2: RT ~ trial + length + OLD + sbtlx.freq + sbtlx.basefreq + (1 |
    subjID) + (1 | word)
  #Df LogLik Df  Chisq Pr(>Chisq)
1   8 -35955
2   9 -35754  1 402.48  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

	Coef.	S.E.	χ^2	$p(\chi^2)$
(Intercept)	6.558	0.00		
Trial number	0.000	0.00	1081.80	$< .001$
Squared length	0.022	0.00	148.13	$< .001$
OLD	0.036	0.00	402.48	$< .001$
Word frequency	-0.037	0.00	6468.10	$< .001$
Base frequency	-0.017	0.00	612.32	$< .001$

Table 1: Sample results table.

```
> lrtest(update(r, . ~ . - sbtlx.freq), r)


Likelihood ratio test

Model 1: RT ~ trial + length + OLD + sbtlx.basefreq + (1 | subjID) + (1 |
    word)
Model 2: RT ~ trial + length + OLD + sbtlx.freq + sbtlx.basefreq + (1 |
    subjID) + (1 | word)
  #Df LogLik Df  Chisq Pr(>Chisq)
1   8 -38988
2   9 -35754  1 6468.1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


> lrtest(update(r, . ~ . - sbtlx.basefreq), r)


Likelihood ratio test

Model 1: RT ~ trial + length + OLD + sbtlx.freq + (1 | subjID) + (1 |
    word)
Model 2: RT ~ trial + length + OLD + sbtlx.freq + sbtlx.basefreq + (1 |
    subjID) + (1 | word)
  #Df LogLik Df  Chisq Pr(>Chisq)
1   8 -36060
2   9 -35754  1 612.32  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results are shown in Table 1; all five fixed effects are significant at $\alpha = .05$. Trial number, squared length, and neighborhood density are positively correlated with response latencies; word frequency and base frequency are negatively correlated with response latencies.