Exercise EADM Workshop Multilevel Regression

Reference

Schultze, T., & Rakotoarisoa, A. F. (2015). Effects of distance between initial estimates and advice on advice utilization. *Judgment and Decision Making, 10(2)*, 144-171.

Details Study 1

Task

60 Trials: Airline distance between European Union capitals

3 Phases: Initial Estimate (IE), Advise Estimate (AE), Final Estimate (FE)

Independent Variable

Advise Estimate: -75% to 75% (in steps of 5%) from IE

Example

IE: 600km, AD (10%) 660km, FE 630km

Dependent Variable

$$AT = \frac{IE - FE}{IE - AD} = .50$$

Range of score: restricted between 0 to 1

Read in the data

```
dataStudy1 = read.csv(file="http://journal.sjdm.org/14/141112a/ferni1.csv")

head(dataStudy1)
```

```
##     X VP trial dev        AT DC   IE   FE   AD abs_dev abs_dev_sq  log_dev
## 1 1.1  1     1  10 0.0000000  0  500  500  550      10        100 2.302585
## 2 2.1  2     1  10 0.0000000  3 1000 1000 1100      10        100 2.302585
## 3 3.1  3     1  10 0.0000000  0  700  680  770      10        100 2.302585
## 4 4.1  4     1  10 1.0000000  1 3000 3700 3300      10        100 2.302585
## 5 5.1  5     1  10 0.3333333  1  600  620  660      10        100 2.302585
## 6 6.1  6     1  10 0.7142857  0  700  750  770      10        100 2.302585
##        shift ignore
## 1 0.00000000      1
## 2 0.00000000      1
## 3 0.00000000      1
## 4 0.23333333      0
## 5 0.03333333      0
## 6 0.07142857      0
```

Install and Load Libraries

```
# install.packages("arm")
# install.packages("lattice")
# install.packages("lmerTest")
```

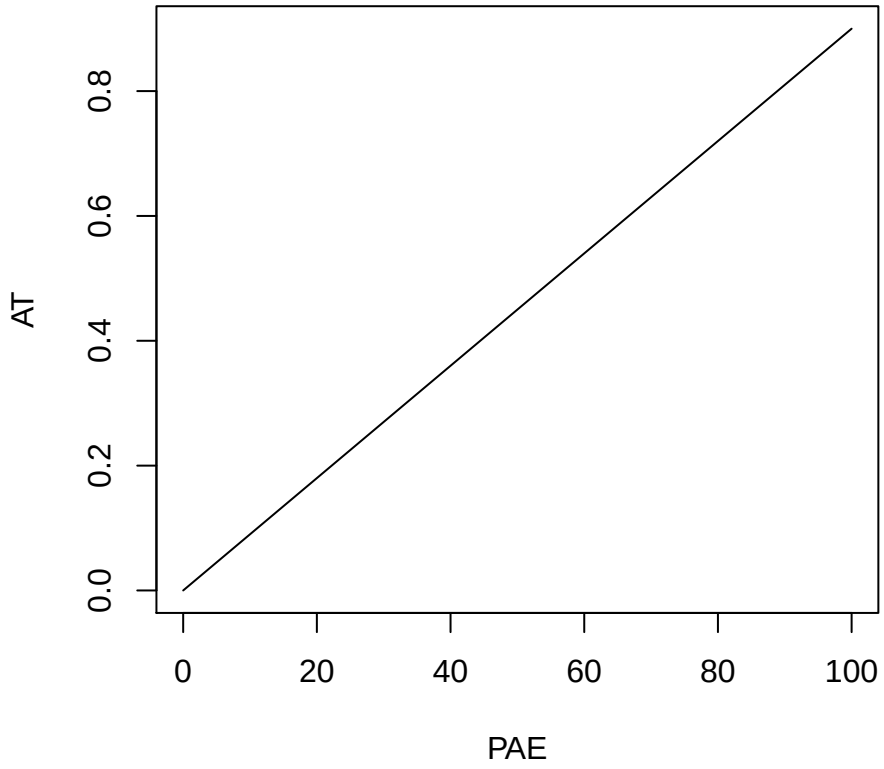
```r
library("arm")
library("lattice")
library("lmerTest")
```

Three Proposed Models

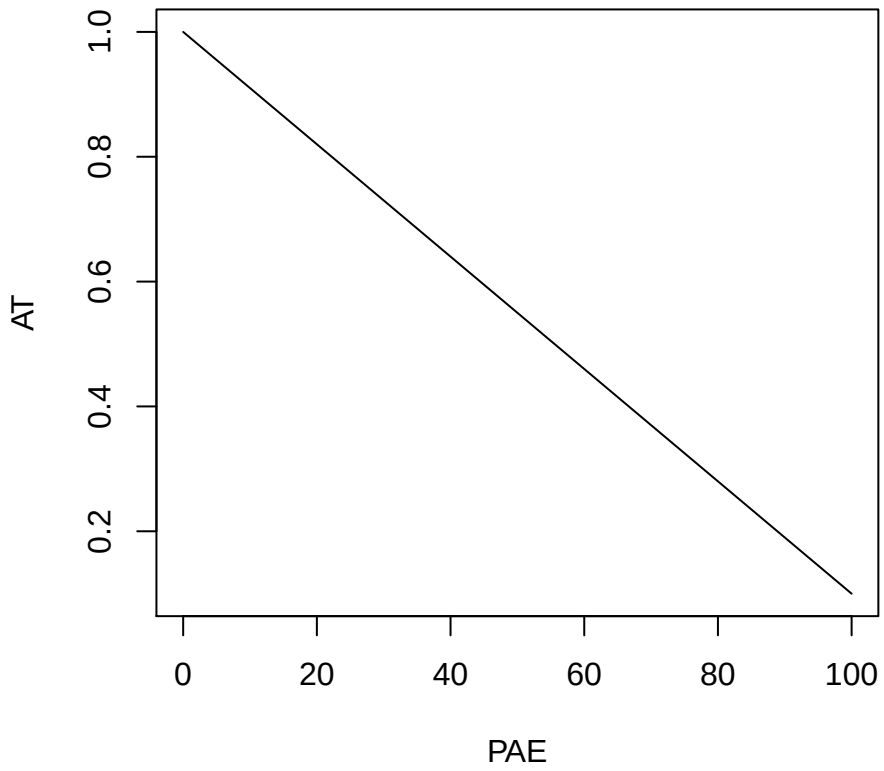
Linear model

$$AT = (\beta_{00} + u_{0j}) + (\beta_{01} + u_{1j}) \times PAE_{ij} + e_{ij}$$

```r
curve(0+.009*x,0,100,xlab="PAE",ylab="AT")
```



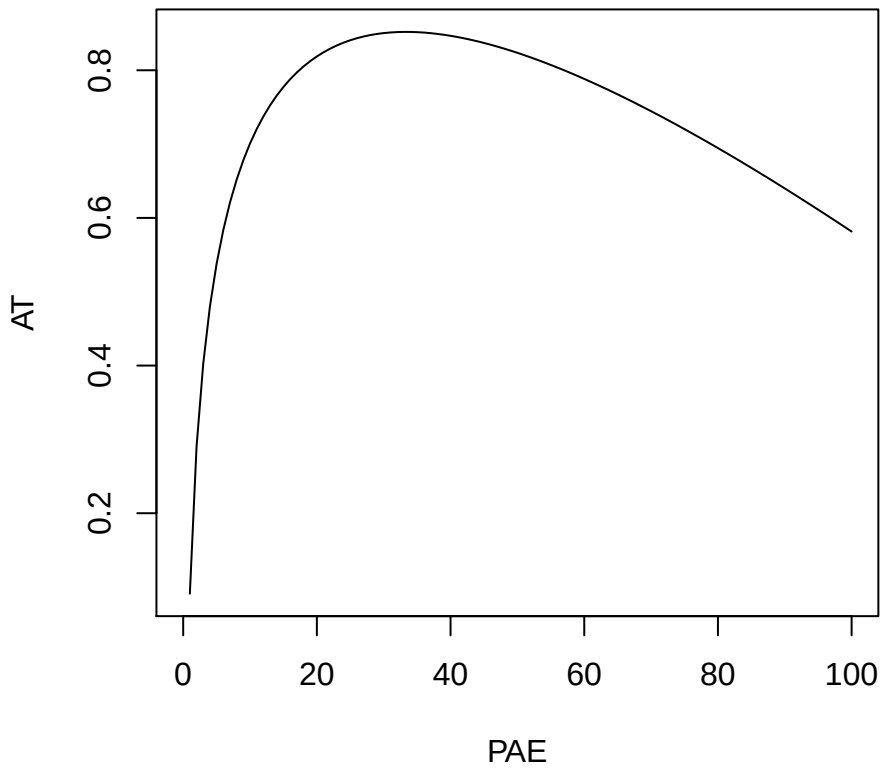

```r
curve(1-.009*x,0,100,xlab="PAE",ylab="AT")
```



Log-model

$$AT = (\beta_{00} + u_{0j}) + (\beta_{01} + u_{1j}) \times PAE_{ij} + (\beta_{02} + u_{2j}) \times log(PAE_{ij}) + e_{ij}$$

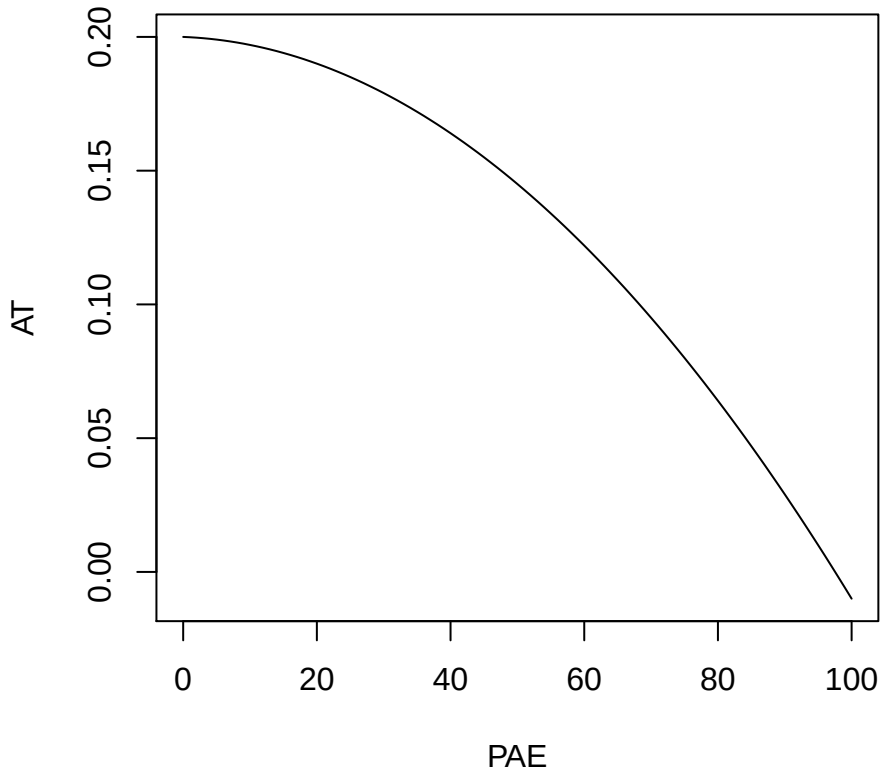
```
curve(.1-.009*x+.3*log(x),0,100,xlab="PAE",ylab="AT")
```



Squared-model

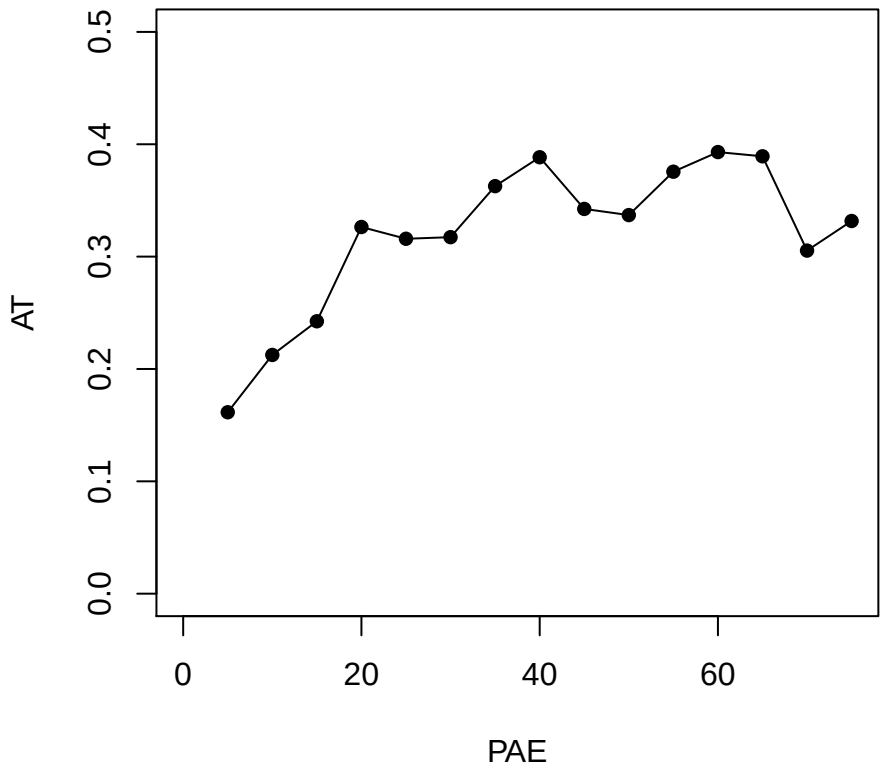
$$AT = (\beta_{00} + u_{0j}) + (\beta_{01} + u_{1j}) \times PAE_{ij} + (\beta_{02} + u_{2j}) \times (PAE_{ij})^2 + e_{ij}$$

```
curve( .2+ -.0001*x+ -.00002*(x)^2,0,100,xlab="PAE",ylab="AT")
```



Plot Data

```
stepsManip = seq(0,75,5)

totalMeans = numeric()

for(loop in stepsManip){
  totalMeans = c(totalMeans,mean(dataStudy1$AT[dataStudy1$abs_dev==loop]))
}

plot(stepsManip,totalMeans,pch=16,xlab="PAE",ylab="AT",ylim=c(0,.5))
points(stepsManip,totalMeans,type="l")
```



Run Multilevel Models

All models contain a random intercept and random slopes for all predictors.

Fit Linear Model

$$AT = (\beta_{00} + u_{0j}) + (\beta_{01} + u_{1j}) \times PAE_{ij} + e_{ij}$$

```
## Linear mixed model fit by maximum likelihood t-tests use Satterthwaite
##   approximations to degrees of freedom [lmerMod]
## Formula: AT ~ 1 + abs_dev + (1 + abs_dev | VP)
##    Data: dataStudy1
##
##      AIC      BIC   logLik deviance df.resid
##     48.7     80.6    -18.3     36.7     1505
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9857 -0.5888 -0.1293  0.6236  3.9321
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  VP       (Intercept) 3.065e-02 0.175067
```

```
##           abs_dev     4.108e-06 0.002027 0.05
##  Residual             5.523e-02 0.235020
## Number of obs: 1511, groups:  VP, 27
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept) 2.408e-01  3.602e-02 2.710e+01   6.687 3.47e-07 ***
## abs_dev     2.049e-03  4.808e-04 2.724e+01   4.262 0.000217 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr)
## abs_dev -0.147
```

Fit Log-Model

$$AT = (\beta_{00} + u_{0j}) + (\beta_{01} + u_{1j}) \times PAE_{ij} + (\beta_{02} + u_{2j}) \times 10 \times log(PAE_{ij}) + e_{ij}$$

```
log_devSc = dataStudy1$log_dev * 10
```

```
## Linear mixed model fit by maximum likelihood t-tests use Satterthwaite
##   approximations to degrees of freedom [lmerMod]
## Formula: AT ~ 1 + abs_dev + log_devSc + (1 + abs_dev + log_devSc | VP)
##    Data: dataStudy1
##
##      AIC      BIC   logLik deviance df.resid
##     -4.6     48.6     12.3    -24.6     1501
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.0785 -0.5888 -0.1595  0.5470  4.0821
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  VP       (Intercept) 1.145e-01 0.338431
##           abs_dev     2.429e-05 0.004928  0.79
##           log_devSc   2.493e-04 0.015789 -0.87 -0.92
##  Residual             5.210e-02 0.228261
## Number of obs: 1511, groups:  VP, 27
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept) -0.094774   0.083086 26.699000  -1.141 0.264130
## abs_dev     -0.002963   0.001239 26.156000  -2.392 0.024238 *
## log_devSc    0.015426   0.003813 26.167000   4.046 0.000411 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) abs_dv
## abs_dev    0.809
## log_devSc -0.906 -0.927
```

Fit Squared Model

$$AT = (\beta_{00} + u_{0j}) + (\beta_{01} + u_{1j}) \times PAE_{ij} + (\beta_{02} + u_{2j}) \times \frac{1}{100} \times (PAE_{ij})^2 + e_{ij}$$

```
abs_dev_sqResc = dataStudy1$abs_dev_sq/100
```

```
## Linear mixed model fit by maximum likelihood t-tests use Satterthwaite
##   approximations to degrees of freedom [lmerMod]
## Formula:
## AT ~ 1 + abs_dev + abs_dev_sqResc + (1 + abs_dev + abs_dev_sqResc |
##     VP)
##    Data: dataStudy1
##
##      AIC      BIC   logLik deviance df.resid
##      0.6     53.8      9.7    -19.4     1501
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1699 -0.5890 -0.1651  0.5680  4.0980
##
## Random effects:
##  Groups   Name           Variance  Std.Dev. Corr
##  VP       (Intercept)    3.164e-02 0.177877
##           abs_dev        6.266e-05 0.007916 -0.31
##           abs_dev_sqResc 7.210e-05 0.008491  0.30 -0.97
##  Residual                5.249e-02 0.229111
## Number of obs: 1511, groups:  VP, 27
##
## Fixed effects:
##                 Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)     0.130335   0.040044 27.582000   3.255 0.002998 **
## abs_dev         0.009779   0.001938 27.010000   5.047 2.68e-05 ***
## abs_dev_sqResc -0.009631   0.002186 26.883000  -4.407 0.000151 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) abs_dv
## abs_dev     -0.497
## abs_dv_sqRs  0.468 -0.972
```

Compare Models Using the Likelihood Ratio Test

Test Linear versus Log Model

```
## Data: dataStudy1
## Models:
## object: AT ~ 1 + abs_dev + (1 + abs_dev | VP)
## ..1: AT ~ 1 + abs_dev + log_devSc + (1 + abs_dev + log_devSc | VP)
##        Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## object  6 48.658 80.581 -18.329   36.658
## ..1    10 -4.562 48.643  12.281  -24.562 61.22      4  1.607e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Test Linear versus Squared Model

```
## Data: dataStudy1
## Models:
## object: AT ~ 1 + abs_dev + (1 + abs_dev | VP)
## ..1: AT ~ 1 + abs_dev + abs_dev_sqResc + (1 + abs_dev + abs_dev_sqResc |
## ..1:     VP)
##        Df    AIC    BIC   logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  6 48.658 80.581 -18.3288   36.658
## ..1    10  0.586 53.791   9.7071  -19.414 56.072      4  1.937e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Test Log versus Squared Model

Bayesian information criterion.

```
BICs = BIC(logMod,sqMod)

BICs = data.frame(BICs)

BICs
```

```
##        df      BIC
## logMod 10 48.64294
## sqMod  10 53.79102
```

Bayes factor

```
exp((BICs[2,2] - BICs[1,2])/2)
```

```
## [1] 13.11874
```

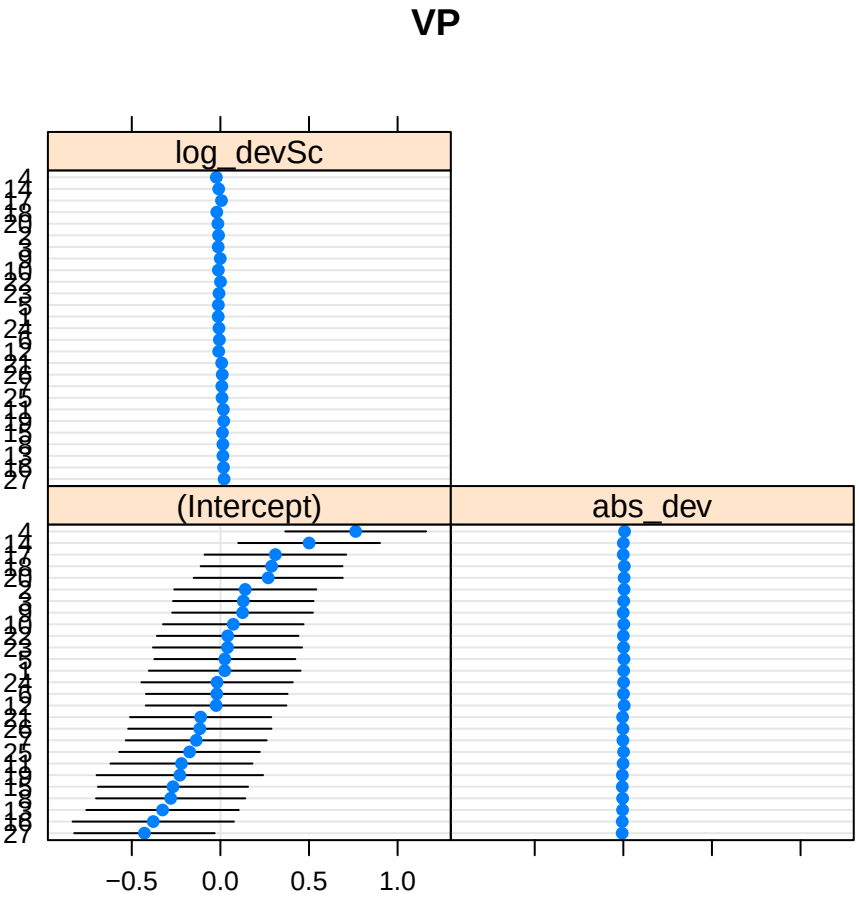
Bonus Task

Test for the Log-Model if we need all random effects (Hint: nested models; test starts with L).

Print and Plot Random Effects

```
## $VP
##    (Intercept)       abs_dev     log_devSc
## 1   0.02444579  0.0022032535 -0.0121002892
## 2   0.14013174  0.0045494865 -0.0102646530
## 3   0.12938669  0.0024550710 -0.0122413846
## 4   0.76346104  0.0071389043 -0.0229845955
## 5   0.02502868  0.0034267500 -0.0114220460
## 6  -0.02051552  0.0011046381 -0.0056433301
## 7  -0.13626241 -0.0023907587  0.0080050245
## 8  -0.28059357 -0.0039641636  0.0140944888
## 9   0.12516583 -0.0005988885 -0.0008109228
## 10  0.07285746  0.0025467721 -0.0113313189
## 11 -0.21958611 -0.0014808515  0.0169881599
## 12 -0.02431117  0.0048771166 -0.0096427238
```

```
## 13 -0.32630168 -0.0043161016  0.0141704711
## 14  0.50076592  0.0001663427 -0.0096312745
## 15 -0.26737378 -0.0057993109  0.0117668654
## 16 -0.37878783 -0.0061223484  0.0175377788
## 17  0.31013401 -0.0004429105  0.0062670763
## 18  0.28904421  0.0056165946 -0.0204093811
## 19 -0.22877269 -0.0053480723  0.0188565949
## 20  0.26971843  0.0043812687 -0.0140829491
## 21 -0.11164453 -0.0041289801  0.0072157646
## 22  0.04122889  0.0000155346  0.0006686806
## 23  0.04022186  0.0013244984 -0.0078229147
## 24 -0.01871451  0.0015998654 -0.0078883293
## 25 -0.17406334  0.0015259701  0.0093129705
## 26 -0.11614266 -0.0016904869  0.0105446406
## 27 -0.42852075 -0.0066491936  0.0208475965

## $VP
```



Bonus Task

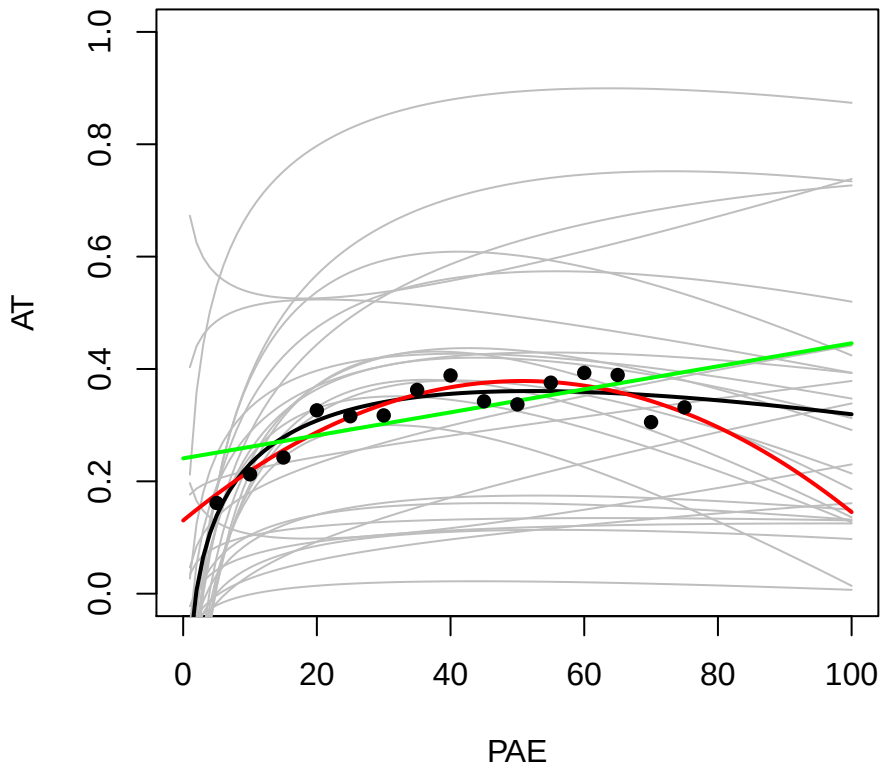
What is the predicted AT for Participant 10 in a task with 40 PAE?

Confidence Intervals for Random and Fixed Effects

```
##                    2.5 %       97.5 %
## .sig01       0.290660357  0.4306436690
## .sig02       0.634333246  0.8185529828
## .sig03      -0.881780538 -0.8513384893
## .sig04       0.002868468  0.0056161554
## .sig05      -0.937868311 -0.7437788051
## .sig06       0.015246500  0.0214966570
## .sigma       0.228127802  0.2307459632
## (Intercept) -0.265467275  0.0731725108
## abs_dev     -0.005695276 -0.0004631329
## log_devSc    0.008341204  0.0225499996
```

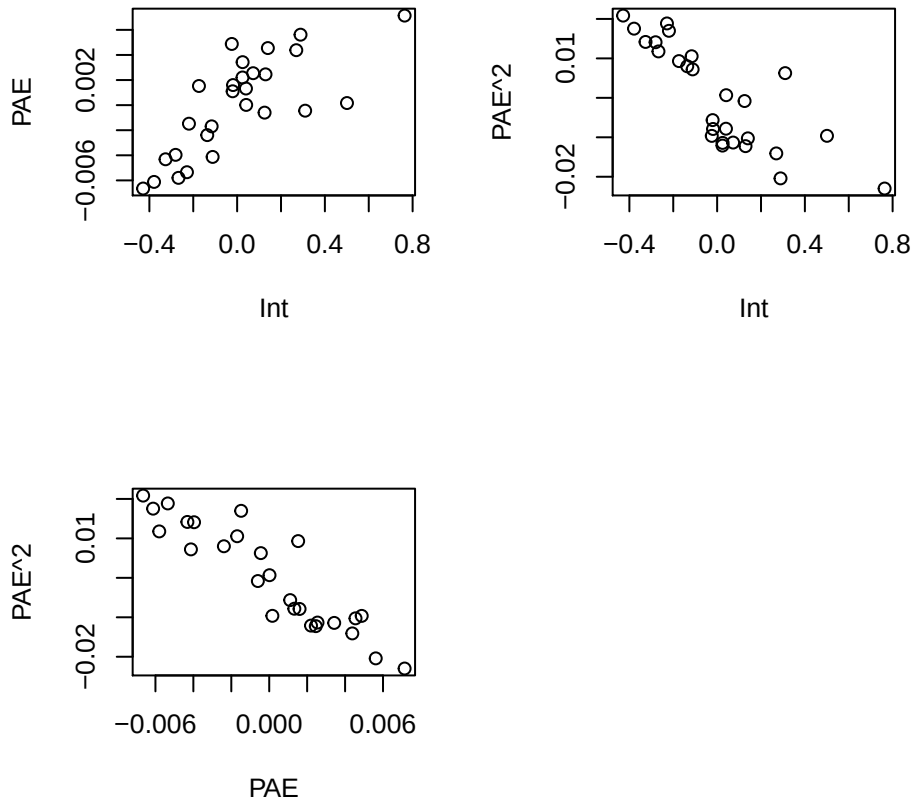
Plot Models with Data

```
for(loop in 1 : 27){

  plotF = function(x){(fixef(logMod)[1] + ranef(logMod)$VP[loop,1]) +
      (fixef(logMod)[2] + ranef(logMod)$VP[loop,2]) * (x) +
      (fixef(logMod)[3] + ranef(logMod)$VP[loop,3]) * 10*log(x)}

  if(loop == 1){
    curve(plotF,0,100,ylim=c(0,1),col="grey",
          xlab="PAE", ylab="AT")
  }else{

    curve(plotF,0,100,add=T,col="grey")

  }}

plotF = function(x){fixef(logMod)[1] +
    (fixef(logMod)[2] ) * (x) +
    (fixef(logMod)[3] ) * 10 * log(x)}

curve(plotF,0,100,add=T,col="black",lwd=2)

plotF = function(x){fixef(sqMod)[1] +
    (fixef(sqMod)[2] ) * (x) +
    (fixef(sqMod)[3] ) * (x)^2/100}

curve(plotF,0,100,add=T,lwd=2,col="red")

plotF = function(x){fixef(linearMod)[1] +
    (fixef(linearMod)[2] ) * (x) }

curve(plotF,0,100,add=T,lwd=2,col="green")

points(stepsManip,totalMeans,pch=16)
```



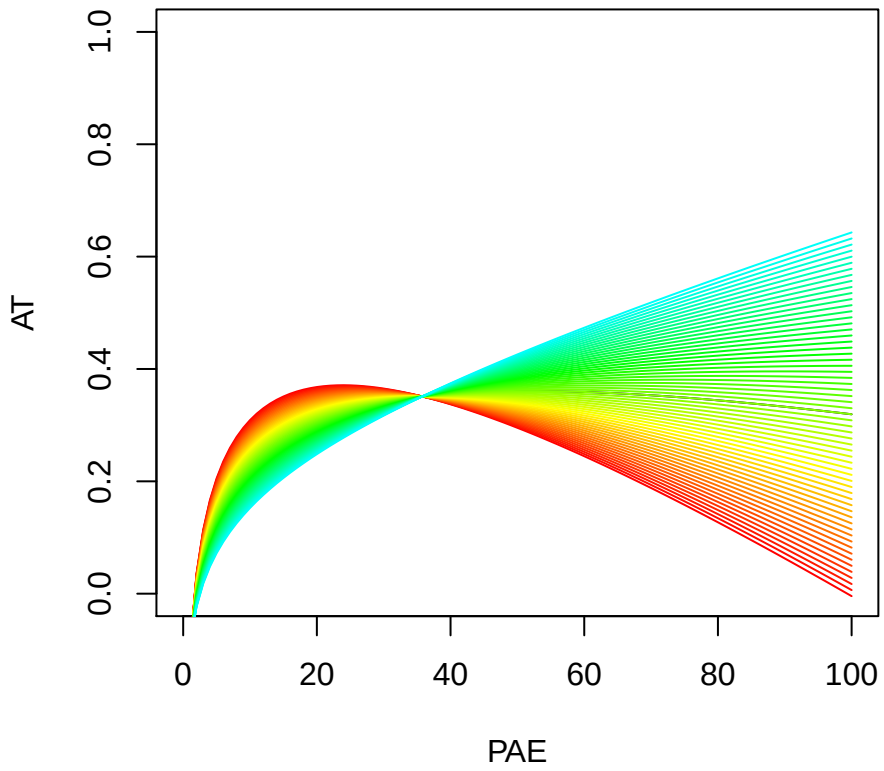
Scatterplot random effects.

```
layout(rbind(c(1,2),c(3,4)))
plot(ranef(logMod)$V[,1],ranef(logMod)$V[,2],xlab="Int",ylab="PAE")
plot(ranef(logMod)$V[,1],ranef(logMod)$V[,3],xlab="Int",ylab="PAE^2")
plot(ranef(logMod)$V[,2],ranef(logMod)$V[,3],xlab="PAE",ylab="PAE^2")
```



Visualizing Correlation of Random Slopes

```
plotF = function(x,a,b,c){fixef(logMod)[1]+a +
    (fixef(logMod)[2] +b) * (x) +
    (fixef(logMod)[3] +c) * 10 * log(x)}


curve(plotF(x,a=0,b=0,c=0),0,100,col="black",lwd=1,ylim=c(0,1),
      xlab="PAE",ylab="AT")

colPl=rainbow(121)

counter=1

for(loop in seq(-.006,.006,.0002)){

curve(plotF(x,a=0,b=loop,c=-loop),0,100,add=T,col=colPl[counter])
  counter = counter + 1
  }
```


AT
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100
PAE