# Colour experiment

## Introduction

This analysis looks at the sign variants used in a colour naming game between signers of different sign languages meeting after 1 week of interaction and after 3 weeks of interaction. The data was collected by Kang Suk Byun (Kang-Suk.Byun@mpi.nl).

The analysis tries to predict the relative frequency of each variant within a colour category in week 3, based on measures from week 1.

## Load libraries

```
library(ggplot2)
library(lme4)
library(party)
```

## Load data

```
variants = read.csv('../data/processedData/variants_summary.csv', stringsAsFactors = F)
```

There were only 2 variants for 'white', and one of those only appeared once. Therefore, we remove it from this statistical analysis.

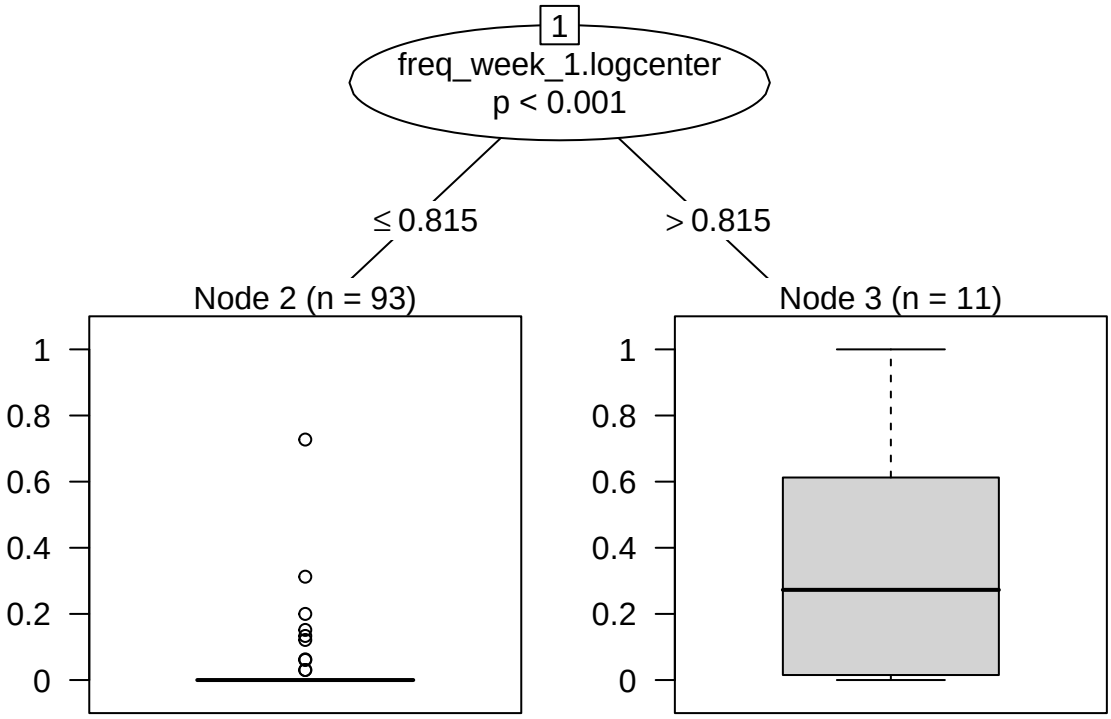
```
variants = variants[variants$colourName!='white',]
```

Transform some variables.

```
variants$Teach = variants$Teach >0

variants$check.any = variants$check>0

variants$freq_week_1.logcenter = log(variants$freq_week_1 + 1)
variants$freq_week_1.logcenter =
  variants$freq_week_1.logcenter - mean(variants$freq_week_1.logcenter)

# cut TryMarking into two categories
variants$TryMarked.cat = cut(variants$TryMarked,
                             c(-Inf,3,Inf),
                             labels = c("Low",'High'))

variants$averageLength_week_1.logcenter = log(variants$averageLength_week_1)
variants$averageLength_week_1.logcenter =
  variants$averageLength_week_1.logcenter -
    mean(variants$averageLength_week_1.logcenter)
```

## Ctree

```
ct =  ctree(freq_week_4_withinColour ~
            factor(indexical) +
            (factor(Teach) +factor(TryMarked.cat)) +
            freq_week_1.logcenter +
            averageLength_week_1.logcenter+
            factor(check.any), data=variants)
plot(ct)
```




## LMER models

```
m0 = lmer(freq_week_4_withinColour ~
            1
          + (1 | colourName),
          data=variants)
m1 = lmer(freq_week_4_withinColour ~
            (indexical) +
            + (1 | colourName),
          data=variants)
m2 = lmer(freq_week_4_withinColour ~
            (indexical) +
            (Teach)
          + (1 | colourName),
          data=variants)

m3 = lmer(freq_week_4_withinColour ~
            (indexical) +
```

```
            (Teach) + (TryMarked.cat)
          + (1 | colourName),
          data=variants)

m4 = lmer(freq_week_4_withinColour ~
            (indexical) +
            Teach * TryMarked.cat
          + (1 | colourName),
          data=variants)

m5 = lmer(freq_week_4_withinColour ~
            (indexical) +
            (Teach * TryMarked.cat) +
            freq_week_1.logcenter
          + (1 | colourName),
          data=variants)

m6 = lmer(freq_week_4_withinColour ~
            (indexical) +
            (Teach * TryMarked.cat) +
            freq_week_1.logcenter +
            averageLength_week_1.logcenter
          + (1 | colourName),
          data=variants)

m7 = lmer(freq_week_4_withinColour ~
            (indexical) +
            (Teach * TryMarked.cat) +
            freq_week_1.logcenter +
            averageLength_week_1.logcenter+
            check.any
          + (1 | colourName),
          data=variants)

m8 = lmer(freq_week_4_withinColour ~
            (indexical) +
            (Teach * TryMarked.cat) +
            freq_week_1.logcenter +
            averageLength_week_1.logcenter+
            check.any + inventedBy
          + (1 | colourName),
          data=variants)
```

# Results

```
anova(m0,m1,m2,m3,m4,m5, m6,m7,m8)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: variants
## Models:
## m0: freq_week_4_withinColour ~ 1 + (1 | colourName)
```

```
## m1: freq_week_4_withinColour ~ (indexical) + +(1 | colourName)
## m2: freq_week_4_withinColour ~ (indexical) + (Teach) + (1 | colourName)
## m3: freq_week_4_withinColour ~ (indexical) + (Teach) + (TryMarked.cat) +
## m3:     (1 | colourName)
## m4: freq_week_4_withinColour ~ (indexical) + Teach * TryMarked.cat +
## m4:     (1 | colourName)
## m5: freq_week_4_withinColour ~ (indexical) + (Teach * TryMarked.cat) +
## m5:     freq_week_1.logcenter + (1 | colourName)
## m6: freq_week_4_withinColour ~ (indexical) + (Teach * TryMarked.cat) +
## m6:     freq_week_1.logcenter + averageLength_week_1.logcenter +
## m6:     (1 | colourName)
## m7: freq_week_4_withinColour ~ (indexical) + (Teach * TryMarked.cat) +
## m7:     freq_week_1.logcenter + averageLength_week_1.logcenter +
## m7:     check.any + (1 | colourName)
## m8: freq_week_4_withinColour ~ (indexical) + (Teach * TryMarked.cat) +
## m8:     freq_week_1.logcenter + averageLength_week_1.logcenter +
## m8:     check.any + inventedBy + (1 | colourName)
##    Df      AIC     BIC logLik deviance   Chisq Chi Df Pr(>Chisq)
## m0  3  -72.634 -64.701 39.317  -78.634
## m1  5  -78.372 -65.150 44.186  -88.372  9.7379      2  0.0076814 **
## m2  6  -76.554 -60.688 44.277  -88.554  0.1820      1  0.6696992
## m3  7  -86.814 -68.304 50.407 -100.814 12.2606      1  0.0004626 ***
## m4  8  -88.303 -67.148 52.151 -104.303  3.4885      1  0.0617963 .
## m5  9 -112.663 -88.863 65.332 -130.663 26.3600      1  2.833e-07 ***
## m6 10 -111.601 -85.157 65.800 -131.601  0.9379      1  0.3328193
## m7 11 -110.103 -81.014 66.051 -132.103  0.5016      1  0.4787842
## m8 14 -105.424 -68.403 66.712 -133.424  1.3219      3  0.7239414
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(m7)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## freq_week_4_withinColour ~ (indexical) + (Teach * TryMarked.cat) +
##     freq_week_1.logcenter + averageLength_week_1.logcenter +
##     check.any + (1 | colourName)
##    Data: variants
##
## REML criterion at convergence: -88
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.7920 -0.5589  0.0296  0.2283  4.7524
##
## Random effects:
##  Groups     Name        Variance  Std.Dev.
##  colourName (Intercept) 0.0006052 0.0246
##  Residual               0.0175678 0.1325
## Number of obs: 104, groups:  colourName, 6
##
## Fixed effects:
##                               Estimate Std. Error t value
## (Intercept)                    0.06140    0.02376   2.584
## indexicalYes                   0.05931    0.06311   0.940
```

```
## indexicalYes-body               0.05158    0.03678   1.402
## TeachTRUE                      -0.05149    0.04810  -1.070
## TryMarked.catHigh               0.11954    0.07712   1.550
## freq_week_1.logcenter           0.17886    0.03396   5.267
## averageLength_week_1.logcenter -0.01703    0.01638  -1.039
## check.anyTRUE                  -0.02277    0.03126  -0.729
## TeachTRUE:TryMarked.catHigh    -0.18552    0.09334  -1.988
##
## Correlation of Fixed Effects:
##             (Intr) indxcY indxY- TcTRUE TryM.H fr__1. aL__1. c.TRUE
## indexicalYs -0.173
## indxclYs-bd -0.408  0.059
## TeachTRUE   -0.279  0.072  0.068
## TryMrkd.ctH -0.256 -0.004 -0.035  0.134
## frq_wk_1.lg  0.383  0.099 -0.224 -0.184 -0.436
## avrgLng__1. -0.031 -0.173  0.153 -0.045  0.019 -0.138
## chck.nyTRUE -0.508  0.020  0.241  0.101  0.121 -0.325  0.087
## TcTRUE:TM.H  0.158 -0.042 -0.042 -0.490 -0.662  0.027 -0.021 -0.084
```

## Summary

There was a significant main effect of try marking ( beta = 0.12 , std.err = 0.077 , Wald t = 1.6 ; log likelihood difference = 6.1 , df = 1 , Chi Squared = 12.26 , p = 0.00046 ).

There was a marginal interaction between try marking and teaching ( beta = -0.19 , std.err = 0.093 , Wald t = -2 ; log likelihood difference = 1.7 , df = 1 , Chi Squared = 3.49 , p = 0.062 ).

There was a significant main effect of frequency in week 1 ( beta = 0.18 , std.err = 0.034 , Wald t = 5.3 ; log likelihood difference = 13 , df = 1 , Chi Squared = 26.36 , p = 2.8e-07 ).

There was a significant main effect of indexicality ( beta = 0.059 , std.err = 0.063 , Wald t = 0.94 ; log likelihood difference = 4.9 , df = 2 , Chi Squared = 9.74 , p = 0.0077 ).

There was no significant main effect of sign length ( beta = -0.017 , std.err = 0.016 , Wald t = -1 ; log likelihood difference = 0.47 , df = 1 , Chi Squared = 0.94 , p = 0.33 ).