TMA4315: Project 3

jototlan@stud.ntnu.no (10018), martigtu@stud.ntnu.no (10037)

Load data:

```
long <- read.csv("https://www.math.ntnu.no/emner/TMA4315/2020h/eliteserie.csv", colClasses = c("factor"
head(long)
```

```
##               attack            defence home goals
## 1              Molde Sandefjord_Fotball  yes     5
## 2 Sandefjord_Fotball              Molde   no     0
## 3      Stroemsgodset            Stabaek  yes     2
## 4            Stabaek      Stroemsgodset   no     2
## 5                Odd          Haugesund  yes     1
## 6          Haugesund                Odd   no     2
```

a)

We consider the model

```
library(glmmTMB)
```

```
## Warning in checkMatrixPackageVersion(): Package version inconsistency detected.
## TMB was built with Matrix version 1.3.4
## Current Matrix version is 1.3.2
## Please re-install 'TMB' from source using install.packages('TMB', type = 'source') or ask CRAN for a
```

```
mod <- glmmTMB(goals ~ home + (1|attack) + (1|defence), poisson, data=long, REML=TRUE)
```

If we consider the number of goals scored by team i in their j'th match, we can euivalently state the model as

$$y_{ij} = \beta_h x_{ij} + \gamma^{\text{attack}}_{k(i,j)} + \gamma^{\text{defence}}_{l(i,j)} + \varepsilon_{ij}.$$

Here, y_{ij} is the number of goals scored, β_h is the effect of playing home, $\gamma^{\text{attack}}_{k(i,j)}$ is the effect of team $k(i,j)$ attacking, $\gamma^{\text{defence}}_{l(i,j)}$ is the effect of team $l(i,j)$ defending, and $\varepsilon_{ij} \sim$ Poisson is the error term. The distributional assumptions on the errors is reasonable, since the number of goals are discrete, and one could imagine that 'unexpected' goals occur as a result of e.g. a 'keeper-blunder' or similar. hvilke paramerte er feilen poissonfordelt med? Er fordelingen til gammaene normal eller poisson?

b)

```
summary(mod)
```

```
##  Family: poisson  ( log )
## Formula:          goals ~ home + (1 | attack) + (1 | defence)
## Data: long
##
```

```
##      AIC      BIC   logLik deviance df.resid
##   1147.2   1163.1   -569.6   1139.2      382
##
## Random effects:
##
## Conditional model:
##  Groups  Name        Variance Std.Dev.
##  attack  (Intercept) 0.007478 0.08647
##  defence (Intercept) 0.016383 0.12800
## Number of obs: 384, groups:  attack, 16; defence, 16
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.12421    0.07809   1.591    0.112
## homeyes      0.40716    0.08745   4.656 3.22e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
ranef(mod)
```

```
## $attack
##                     (Intercept)
## BodoeGlimt         -0.036781062
## Brann               0.012026209
## Haugesund           0.011223106
## Kristiansund       -0.011367328
## Lillestroem        -0.049915996
## Molde               0.078390643
## Odd                 0.003654179
## Ranheim_TF          0.023375599
## Rosenborg           0.050622609
## Sandefjord_Fotball -0.058333079
## Sarpsborg08         0.026946364
## Stabaek            -0.026801293
## Start              -0.060500163
## Stroemsgodset       0.024556017
## Tromsoe             0.005756700
## Vaalerenga          0.007147494
##
## $defence
##                     (Intercept)
## BodoeGlimt         -0.042616090
## Brann              -0.123934761
## Haugesund          -0.061931278
## Kristiansund        0.008112432
## Lillestroem         0.030699257
## Molde              -0.036630979
## Odd                -0.052013600
## Ranheim_TF          0.062209734
## Rosenborg          -0.152631173
## Sandefjord_Fotball  0.133164228
## Sarpsborg08         0.006574064
## Stabaek             0.085376126
## Start               0.081958112
## Stroemsgodset       0.040486666
```

```
## Tromsoe            -0.009852817
## Vaalerenga          0.031030079
```

Marginal variance and intraclass covariance probit model via pmvnorm

```
#install.packages("mvtnorm")
library(mvtnorm) # to use pmvnorm()
```

Power of correct mixed vs misspecified fixed effect model vs pseudoreplication

Numerical computation of the critical value for LRT test of random slope