

STA610 Lab09

Hun Kang

2024-11-01

- Write down your answers in any blank sheet and submit your work in paper during the lab.
- Your work will not be graded. As long as you submit, you will get a full credit.
- For those who missed the lab today, you can submit it via email to me for half credit.

Logistic Regression with Random Intercept

```
library(lme4)
library(tidyverse)
```

```
dat = read.csv("basketball0910.csv") %>% select(game, hometeam, visitor, foul.home, foul.diff)
head(dat)
```

```
##   game hometeam visitor foul.home foul.diff
## 1     1      MN      IA         0         0
## 2     1      MN      IA         1        -1
## 3     1      MN      IA         1         0
## 4     1      MN      IA         0         1
## 5     1      MN      IA         0         0
## 6     1      MN      IA         0        -1
```

- `game` unique game identifier number
- `hometeam` home team abbreviation
- `visitor` visiting team abbreviation
- `foul.home` indicator if foul was called on the home team
- `foul.diff` the difference in fouls before the current foul was called (home - visitor)

Q1: Draw a plot that can visualize a relationship between ‘foul.home’ and ‘foul.diff’.

Q2: Write the formula of the logistic regression model with ‘foul.diff’ as the only predictor and fit it for each ‘game’. Plot the resulting intercepts and slopes in scatter plot.

Q3: Expand the model in Q2 to with random intercept and slope with a grouping factor ‘game’.

Q4: Do you think the random slope is statistically significant? How can we test it?