Concussion data analysis

PBS has gathered data on NFL concussions for 2012-2014 and made them available at

http://www.pbs.org/wgbh/pages/frontline/concussion-watch/

Our objective is to use these data to determine if the rate of concussions varies by team and/or year.

1. Load the data

```
dat   <- read.csv("http://www4.stat.ncsu.edu/~reich/ST590/code/Concussion
team  <- dat[,1]
Y2012 <- dat[,2]
Y2013 <- dat[,3]
Ytot  <- Y2012+Y2013
dat
```

```
##                          X X2012 X2013
## 1       Arizona Cardinals    4     2
## 2        Atlanta Falcons     2     3
## 3       Baltimore Ravens     9     3
## 4          Buffalo Bills     3     0
## 5       Carolina Panthers    3     2
## 6          Chicago Bears     5     1
## 7      Cincinnati Bengals    4    12
## 8        Cleveland Browns   10     7
## 9          Dallas Cowboys    7     1
## 10        Denver Broncos     4     5
## 11         Detroit Lions     5     6
## 12     Green Bay Packers     6     5
## 13        Houston Texans     4     5
## 14     Indianapolis Colts    9     7
## 15 Jacksonville Jaguars     9     9
## 16    Kansas City Chiefs     9     4
## 17        Miami Dolphins     2     3
## 18     Minnesota Vikings     5     7
## 19 New England Patriots     5     6
## 20     New Orleans Saints    4     4
## 21        New York Giants    4     3
## 22          New York Jets    7     5
## 23        Oakland Raiders   12     7
## 24    Philadelphia Eagles    5     4
## 25    Pittsburgh Steelers    5     5
## 26       Saint Louis Rams    4     6
## 27     San Diego Chargers    2     6
## 28    San Francisco 49ers    6     4
## 29       Seattle Seahawks    3     6
## 30 Tampa Bay Buccaneers     4     4
## 31       Tennessee Titans    4     5
## 32   Washington Redskins    6     5
```

```
team  <- team[order(Ytot)] # order the teams by total concussions
Y2012 <- Y2012[order(Ytot)]
Y2013 <- Y2013[order(Ytot)]
Ytot  <- Ytot[order(Ytot)]
```

2. Compare years

There are 32 teams and each plays 16 games, for a total of $N = 32 * 16 = 512$ team games. To estimate the rate in year t, we'll use the total number of concussions that year, Y_t, and fit the model

$$Y_t | \lambda_t \sim \text{Poisson}(N\lambda_t) \quad \lambda_t \sim \text{Gamma}(a, b).$$

```
N       <- 512
a       <- 0.01
b       <- 0.01

a2012 <- sum(Y2012)+a
b2012 <- N+b

a2013 <- sum(Y2013)+a
b2013 <- N+b

grid  <- seq(0.2,0.45,.001)

p2012 <- dgamma(grid,a2012,b2012)
p2013 <- dgamma(grid,a2013,b2013)

plot(grid,p2013,type="l",xlab="Concussion rate (per game)",ylab="Posteri
lines(grid,p2012,lty=2)

legend("topright",c("2012","2013"),lty=2:1,inset=0.05)
```
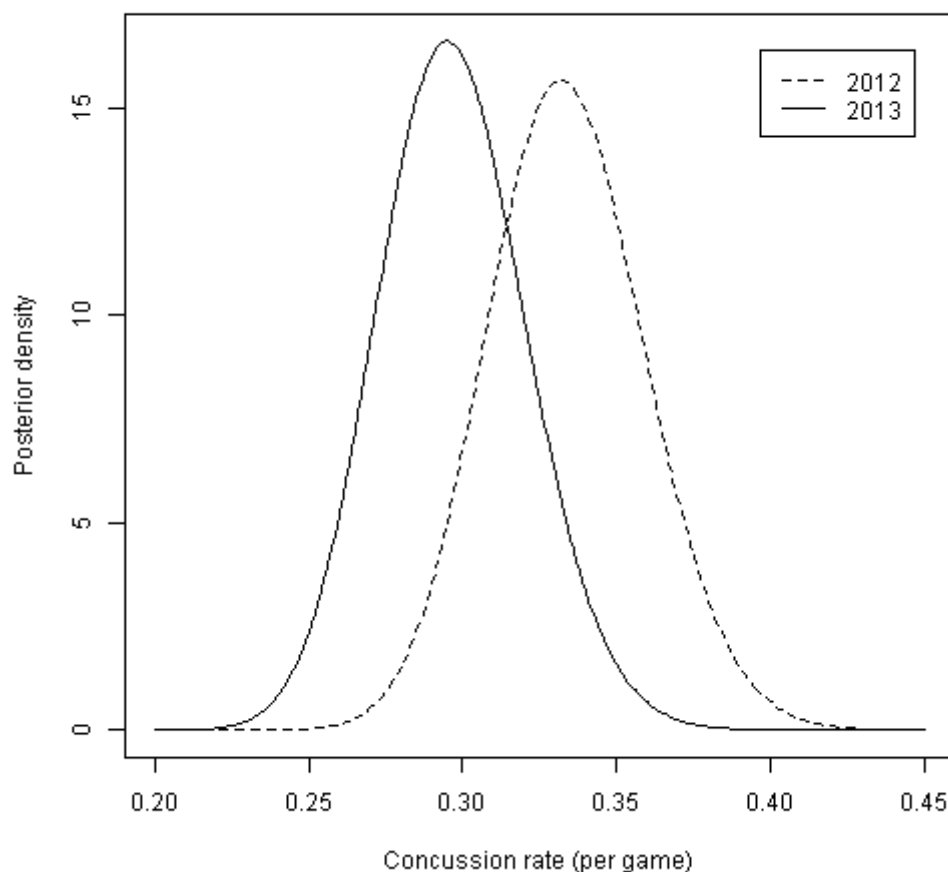

We'll use Monte Carlo sampling to compute the posterior probability that the rate in 2013 (λ_{2013}) is smaller than the rate in 2012 (λ_{2012}).

```
lambda_2012 <- rgamma(1000000,a2012,b2012)
lambda_2013 <- rgamma(1000000,a2013,b2013)
mean(lambda_2013<lambda_2012)
```

```
## [1] 0.854708
```

Therefore, the probability that the concussion rate reduced from 2012 to 2013 is 0.86.

3. Compare teams

Does the concussion rate vary by team? Each team plays 16 games and we will combine the two seasons so $N = 32$.
The analysis below plots the 90% posterior interval for the rate (λ) for each team.

```
a    <- Ytot  + .01
b    <- 32 + .01
low  <- qgamma(0.05,a,b)
high <- qgamma(0.95,a,b)

par(mar=c(4,10,4,4))
plot(NA,xlim=c(0,1),ylim=c(1,32),axes=FALSE,
     ylab="",xlab="Posterior 90% interval",
     main="Concusions per game in the NFL in 2012-2013")

for(j in 1:32){
  lines(c(low[j],high[j]),c(j,j))
  lines(c(low[j],low[j]),c(j-.2,j+.2))
  lines(c(high[j],high[j]),c(j-.2,j+.2))
}
axis(1)
axis(2,at=1:32,labels=team,las=2,cex.axis=.75)
```

Concusions per game in the NFL in 2012-2013

Posterior 90% interval
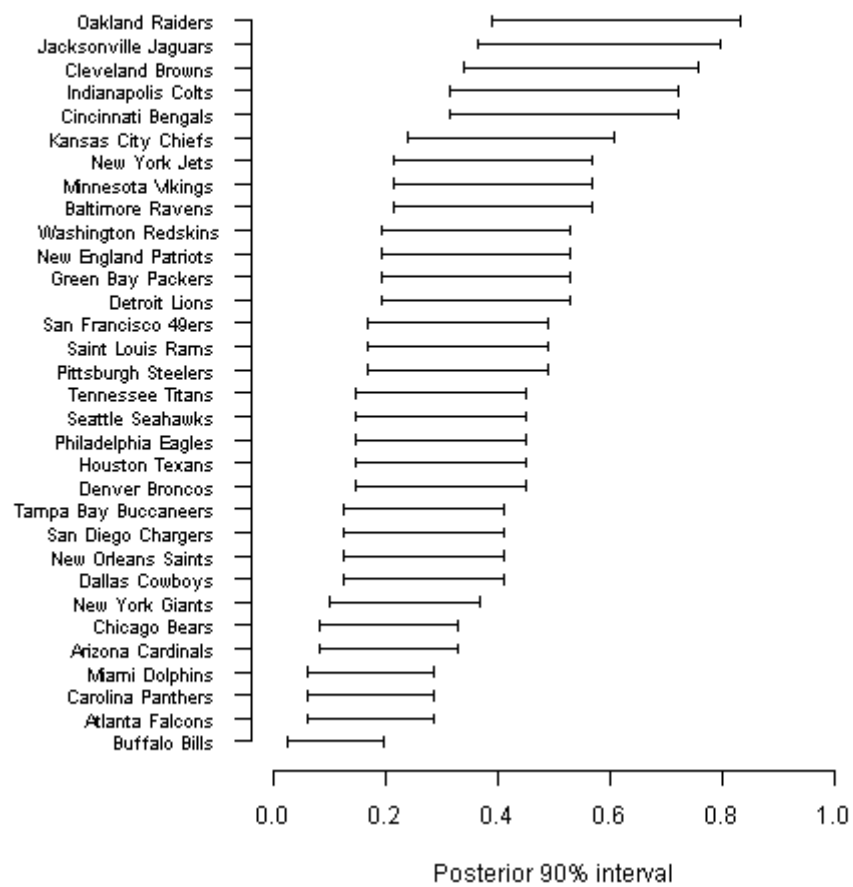
There is some evidence (non-overlapping intervals) that the rate varies by team.