Jim Albert Method

This uses Albert's package on hitting data, BApredict

```r
suppressPackageStartupMessages(library(BApredict))
## hits modified from original Albert program so it doesn't need
## to go out to Sports Illustrated except once a day.
set.seed(42)  # to get consistent results from run to run
source("pred-hr.R")
load("~/Documents/Programming_Languages/R/baseball/hitdata.RData")
model <- pred_hr(hits, "Stanton", standings, 10000)
current_hrs <- model$current_HR
homers <- tibble(hrs = model$future_HR)
ggplot(homers, aes(x = hrs)) + geom_histogram(bins = 20)
```

```r
print(paste("Probability of tying or breaking the record: 
\[1\] "Probability of tying or breaking the record: 0.2359")
```

```r
# [1] "Probability of tying or breaking the record: 0.2359"
```
print(paste("Average number of home runs Stanton will hit =", 
round(mean(homers$hrs), 3)))

## [1] "Average number of home runs Stanton will hit = 0.928"

print(paste("Standard deviation: ", round(sd(homers$hrs), 3)))

## [1] "Standard deviation: 0.939"

### Bob Carpenter (Stan Group) Method

ab <- 583
hr <- 59
games <- 156
phrs <- 
  sum(rbinom(1e5, 
    rpois(1e5, 10 * ab / games), 
    rbeta(1e5, 1 + hr, 1 + ab - hr)) 
  >= (61 - hr)) / 1e5

print(paste("Probability of tying or breaking the record: ", 
  round(phrs, 3)))

## [1] "Probability of tying or breaking the record: 0.888"

### Breaking down the at bats in the Carpenter Model

new_abs <- rpois(1e5, 10 * ab / games)

print(paste("Estimated number of remaining at bats: ", 
  round(mean(new_abs), 2)))

## [1] "Estimated number of remaining at bats: 37.38"

summary(new_abs)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
## 13.000 33.000  37.000  37.378 41.000  67.000

suppressPackageStartupMessages(library(ggpubr))
nab <- tibble(new_abs)
gghistogram(nab, x = "new_abs", binwidth = 1, add = "mean", 
  add_density = TRUE, ggtheme = theme_gray(), fill = "gray", 
  title = "Distribution of New At Bats for Stanton", 
  xlab = "New At Bats")
Revised Probability of Carpenter Model with 4 at bats per game

```r
at_bats_game <- 4
phrs2 <- sum(rbinom(1e5, 
    at_bats_game * (162 - games), 
    rbeta(1e5, 1 + hr, 1 + ab - hr)) 
  >=(61 - hr)) / 1e5
print(paste("Revised probability of tying or breaking the record:" ,
            round(phrs2, 3)))
```

## [1] "Revised probability of tying or breaking the record: 0.714"

Further Revision Using Stanton’s Existing At Bats and Home Runs

```r
phrs3 <- sum(rbinom(1e5, 
    at_bats_game * (162 - games), 
    hr / ab) 
  >=(61 - hr)) / 1e5
print(paste("Revised probability of tying or breaking the record:" ,
            round(phrs3, 3)))
```

## [1] "Revised probability of tying or breaking the record: 0.715"