ST 540 hw4

Robin Baldeo

February 3, 2019

library(purrr)  
library(tidyr)  
library(dplyr)  
library(knitr)

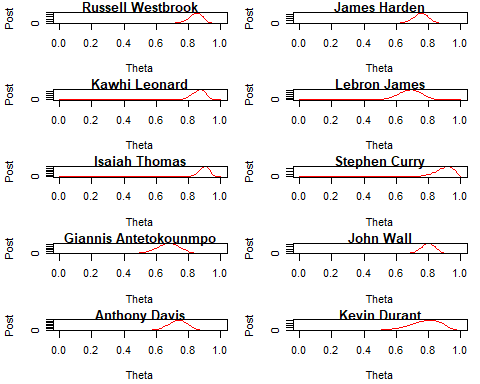
#raw data for study  
player<- c("Russell Westbrook","James Harden", "Kawhi Leonard", "Lebron James", "Isaiah Thomas", "Stephen Curry", "Giannis Antetokounmpo", "John Wall", "Anthony Davis", "Kevin Durant")  
overall<- c(.845, .847, .880, .674, .909, .898, .770, .801, .802, .875)  
cMakes<- c(64, 72, 55, 27, 75, 24, 28, 66, 40, 13)  
cAtt<- c(75,95,63, 39, 83, 26, 41, 82, 54, 16)  
id<- c(1:10)  
  
#construct list of all vectors  
data<- list( "player" = player, "overall"= overall, "cMakes" = cMakes, "cAtt"= cAtt, "id" = id)

## (A)

My prior is a Beta(1,1) because we are supposed to use a uniform prior with a binominal likelihood. Since this is a conjugate prior the result of this is a Beta(y + 1, n-y + 1) posterior.

## (B)

#theta  
m<- seq(0,1, by = .01)  
  
#function for posterior  
betaPost<-function( x,n){  
 a= x +1  
 b = n-x + 1  
 cbind("x" = m, "y" = dbeta(m, a ,b ))%>%return()  
}  
   
#create a list of all the posteriror to be plotted  
clutch<- pmap(data, ~betaPost( ..3, ..4))  
  
#plotting posteriror  
op <- par(pty="m", mfrow=c(5, 2), mar=c(4.2, 4.2, 1, 1))  
  
 for(i in 1:length(player)){  
 plot(x = clutch[[i]][,1],y = clutch[[i]][,2], type = "l",col = "red", xlab = "Theta", ylab= "Post", main = player[i])  
 }



par(op)

## (C)

#function for summary  
postSum<-function( x,n, nam, over){  
 a= x +1  
 b = n-x + 1  
 #mean  
 pMean=a/(a+b)  
 #std  
 pStd = sqrt((a\*b)/(((a+b)\*\*2)\*(a+b+1)))  
 #credible interval  
 ci = qbeta(c(.025, .975), a,b)  
 #lower bound credible interval  
 lowCi=ci[1]  
 #upper bound credible interval  
 hiCi=ci[2]  
 #portion less than .05  
 less= pbeta(.5,a,b)  
 cbind("name" = nam, "overAll"= over, "pMean" = pMean, "pStd" = pStd, "lowCi"= lowCi, "hiCi"= hiCi, "less"= less)%>%  
 as.data.frame%>%  
 return()  
}  
  
#creating dataframe with summary   
df\_c<- pmap\_df(data, ~postSum(..3,..4, ..5, ..2))  
  
#adding players names   
df\_c$name<- player[df\_c$name]  
  
df\_c%>%kable(caption = "Summary Of Player's Posterior")

Summary Of Player’s Posterior

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| name | overAll | pMean | pStd | lowCi | hiCi | less |
| Russell Westbrook | 0.845 | 0.8441558 | 0.0410685 | 0.7557660 | 0.9156623 | 0.0000000 |
| James Harden | 0.847 | 0.7525773 | 0.0435895 | 0.6625065 | 0.8327883 | 0.0000002 |
| Kawhi Leonard | 0.880 | 0.8615385 | 0.0425138 | 0.7684737 | 0.9336259 | 0.0000000 |
| Lebron James | 0.674 | 0.6829268 | 0.0718029 | 0.5346837 | 0.8142710 | 0.0082945 |
| Isaiah Thomas | 0.909 | 0.8941176 | 0.0331787 | 0.8209403 | 0.9498226 | 0.0000000 |
| Stephen Curry | 0.898 | 0.8928571 | 0.0574346 | 0.7571017 | 0.9764725 | 0.0000028 |
| Giannis Antetokounmpo | 0.770 | 0.6744186 | 0.0706429 | 0.5291394 | 0.8043320 | 0.0097602 |
| John Wall | 0.801 | 0.7976190 | 0.0435786 | 0.7059140 | 0.8759195 | 0.0000000 |
| Anthony Davis | 0.802 | 0.7321429 | 0.0586560 | 0.6099716 | 0.8386204 | 0.0001776 |
| Kevin Durant | 0.875 | 0.7777778 | 0.0953772 | 0.5656821 | 0.9318923 | 0.0063629 |

## (D)

#checking to see if the overall matches the posterior  
# df\_d<- pmap\_df(data, ~{cbind("name" = ..5, "pBelow" = pbeta(..2,..3+ 1,..4 - ..3 + 1), "pAbove" = 1- pbeta(..2,..3+ 1,..4 - ..3 + 1))}%>%as.data.frame)  
  
df\_d<- pmap\_df(data, ~{cbind("name" = ..5, "pBelow" = pbeta(..2,..3,..4 ), "pAbove" = 1- pbeta(..2,..3,..4 - ..3))}%>%as.data.frame)  
  
  
#adding names and orginal overall to table  
df\_d$overAll<- overall[df\_d$name]  
df\_d$name<- player[df\_d$name]  
  
  
df\_d[-3]%>%  
 arrange(pBelow)%>%   
 kable(caption = "%Percentage above and below Overall")

%Percentage above and below Overall

|  |  |  |
| --- | --- | --- |
| name | pBelow | overAll |
| Lebron James | 0.9999947 | 0.674 |
| Kevin Durant | 1.0000000 | 0.875 |
| Giannis Antetokounmpo | 1.0000000 | 0.770 |
| Stephen Curry | 1.0000000 | 0.898 |
| Anthony Davis | 1.0000000 | 0.802 |
| Russell Westbrook | 1.0000000 | 0.845 |
| James Harden | 1.0000000 | 0.847 |
| Kawhi Leonard | 1.0000000 | 0.880 |
| Isaiah Thomas | 1.0000000 | 0.909 |
| John Wall | 1.0000000 | 0.801 |

There is a high probability the overall matches the clutch for James Harden. With Kawhi Leonard, Isaiah Thomas, Kevin Durant, Anthony Davis, and Giannis Antetokounmpo the pBelow gives probability that the overall scores are less than the clutch percentage. While for Lebron, Stephen and Russel their probability suggest that clutch percentage should be greater than the overall.

## (E)

I don’t think the data is sensitive to this prior since we are told that we are dealing with a large sample. So given the large sample size our posterior mean gets closer to the sample proportion. Therefore I think the prior has minimal effect.

## MCMC

#seed so experiment can be reproduced  
set.seed(100)  
  
#function that allows for random draws  
mcmc<- function(x,n){  
 a= x +1  
 b = n-x + 1  
 rbeta(100000, a,b)%>% return()  
}  
  
# Using the data data frame with raw information, random draws were made for each of the 10 players  
df\_e<- pmap(data, ~mcmc(..3,..4))  
  
#a matrix to hold all the comparison scores for each player  
l<-matrix(nrow= 10, ncol = 10)  
#comparison   
for(i in 1:10){  
 var<- double(10)  
 for(j in 1:10){  
 #all players were compared to each other and   
 #scores were recorded  
 l[i,j]= (mean(df\_e[[i]]> df\_e[[j]]))  
 }  
}  
  
#transofm each player results so it can be analysed row wise  
tr<- t(l)  
  
#max score of each player was determined and the player was recorded  
rank<- as.integer(rep(0,10))  
per<- as.double(rep(0, 10))  
for(i in 1:10){  
 rank[i]= tr[i,]%>%which.max()  
 per[i] = tr[i,]%>%max()  
   
}  
  
#summary of best player against each player  
results<- data.frame("Player" = player, "Best Score" = player[rank], "Percentage" = per)  
results%>%kable(caption = "Best Player for one on one comparison")

Best Player for one on one comparison

|  |  |  |
| --- | --- | --- |
| Player | Best.Score | Percentage |
| Russell Westbrook | Isaiah Thomas | 0.83010 |
| James Harden | Isaiah Thomas | 0.99498 |
| Kawhi Leonard | Isaiah Thomas | 0.72593 |
| Lebron James | Isaiah Thomas | 0.99777 |
| Isaiah Thomas | Stephen Curry | 0.53190 |
| Stephen Curry | Isaiah Thomas | 0.46810 |
| Giannis Antetokounmpo | Isaiah Thomas | 0.99877 |
| John Wall | Isaiah Thomas | 0.96284 |
| Anthony Davis | Isaiah Thomas | 0.99341 |
| Kevin Durant | Isaiah Thomas | 0.88361 |

Using simulation we see that Isaiah Thomas has the best free throw when compared to all players except Stephen Curry. Stephen Curry is the only player to beat Isaiah Thomas, therefore Stephen Curry is the best player.