The World Series

Xueyuan Li

2022-09-18

Introduciton

The world series is a best-of-7 match-up (i.e., the first team to win 4 out of 7 games wins the championship) between the champions of the American and National Leagues of Major League Baseball.

Today we are going to talk about the some questions about the probability of winning in World Series.

Background

Setup:

1. Suppose that the Braves and the Yankees are teams competing in the World Series.
2. Suppose that in any given game, the probability that the Braves win is PB and the probability that the Yankees win is PY=1-PB.

We use *negative Binomial* cumulative probability: *pnbinom(ML, size, prob)*

ML= max number of losses

size= number of successful trials

prob= probability of success in each trial.

For example:

(1) There are 7 games in total, so team B needs to win at least 4 of these games to win the world series. In other words, team B can not lose more than 3 games to win the world series.

(2) So we assume the team B wins the 7th game, then they can only lose up to 3 games in the first 6 games, which also means they need to win a minimum of 3 games.

(3) the probability distribution of the number of failures that appear will be a negative binomial distribution.

Here is an example Table: (w=wins, f=false)

One	Two	Three	Four	Five	Six	Seven
w	w	w	f	f	f	w
w	w	f	w	f	f	w
w	f	w	f	w	f	w

One	Two	Three	Four	Five	Six	Seven
...	...	...	...	...	...	w

Questions and Solutions:

```
PB=0.55
PB_w=pnbinom(3,4,PB)
PB_w
```

1. What is the probability that the Braves win the World Series given that PB = 0.55?

```
## [1] 0.6082878
```

```
#dnbinom(0,4,0.55)+dnbinom(1,4,0.55)+dnbinom(2,4,0.55)+dnbinom(3,4,0.55)
```

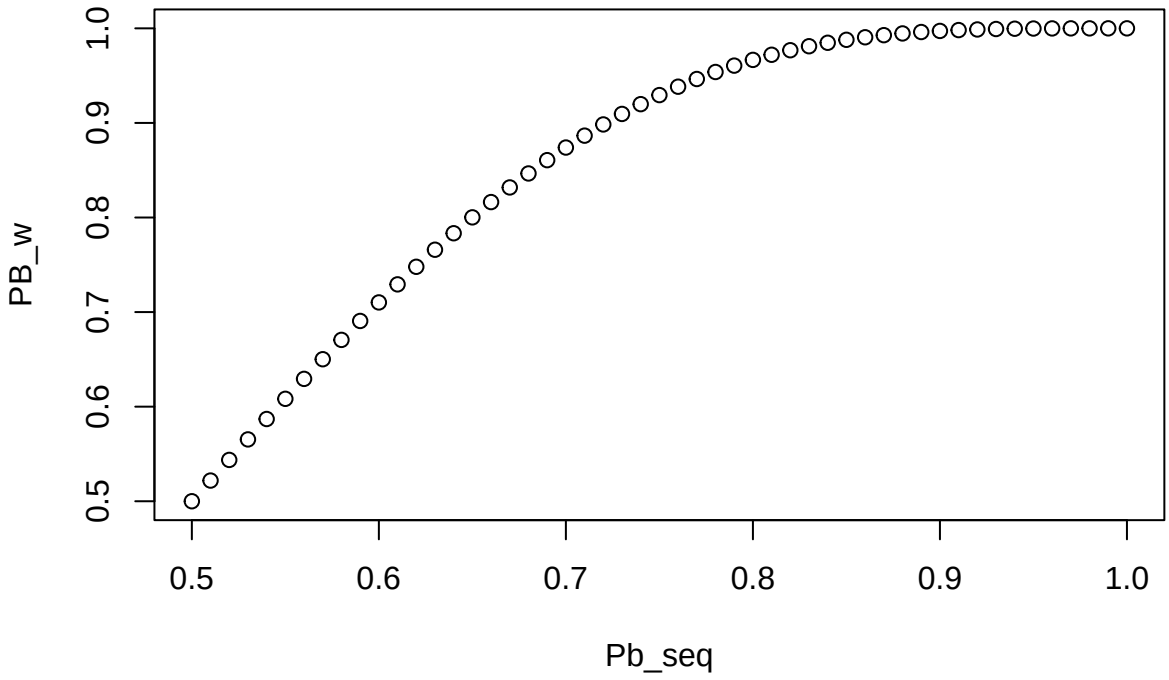
Solution: we assume the probability of Braves win in each games is 0.55, then we need to compute the probability that the Braves win the World Series using *negative Binomial*. We set the max loss number is 3, number of successful trials is 4. Then we get the probility that the Braves win the World Series is 0.6082878.

```
Pb_seq=seq(0.5,1.0,by=0.01)  #seq(min,max,by=increase)
PB_w= pnbinom(3,4,Pb_seq)
PB_w
```

2. What is the probability that the Braves win the World Series given that PB = x? This will be a figure (see below) with PB on the x-axis and P(Braves win World Series) on the y-axis.

```
##  [1] 0.5000000 0.5218663 0.5436801 0.5653893 0.5869421 0.6082878 0.6293763
##  [8] 0.6501589 0.6705884 0.6906193 0.7102080 0.7293131 0.7478954 0.7659184
## [15] 0.7833483 0.8001543 0.8163083 0.8317859 0.8465656 0.8606298 0.8739640
## [22] 0.8865576 0.8984038 0.9094991 0.9198442 0.9294434 0.9383045 0.9464394
## [29] 0.9538632 0.9605947 0.9666560 0.9720724 0.9768724 0.9810869 0.9847497
## [36] 0.9878968 0.9905661 0.9927972 0.9946307 0.9961084 0.9972720 0.9981634
## [43] 0.9988237 0.9992928 0.9996085 0.9998064 0.9999187 0.9999736 0.9999947
## [50] 0.9999997 1.0000000
```

```
plot(Pb_seq,PB_w)
```




```
#x-axis: the probability of Braves wins each game
#y-axis: the probability of Braves wins World Series
```

Solution: We assume that the probability of Braves win each game is from 0.5 to 1.0, the growth interval is 0.01, for example PB can be 0.51,0.52,0.53... Then we compute the probability that the Braves win the World Series also using *negative Binomial*. Finally we use graph to show the relationship between PB and probability of Braves win the World Series.

x-axis: the probability of Braves wins each game

y-axis: the probability of Braves wins World Series

We can find that as the probability of Braves wins each game gradually increases, the probability of the Braves win the World Series also gradually increases.

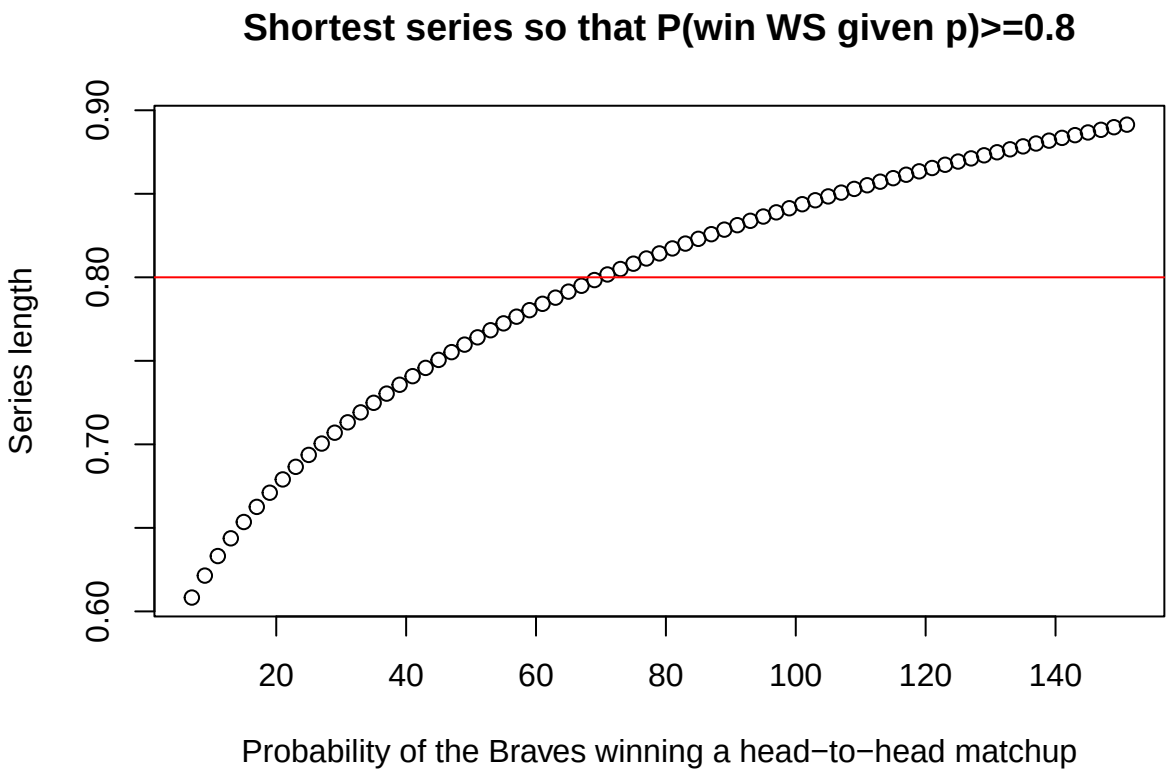
3, Suppose one could change the world series to be best-of-9 or some other best-of-X series. What is the shortest series length so that P(Braves win WS| PB=0.55)>=0.8

```
num_games=seq(7,151,by=2)
Pb=0.55
#you need to win more than num_games/2 games
num_wins= ceiling(num_games/2)
#use ceiling function to get integer which much closer to the num_games, but more than num_games
```

```
num_loss= num_games - num_wins
P_B_WS= pnbinom(num_loss,num_wins,Pb)

#Draw the line with Pn win=0.8, The intersection points means when Pn win=0.8, the number of games they
plot(num_games, P_B_WS,main = "Shortest series so that P(win WS given p)>=0.8",xlab="Probability of the


num_games[which(P_B_WS>=0.8)]
```




```
## numeric(0)
```

```
#They need to play at least 71 games to make Pb win >=0.8
```

Solution: We assume the num of games is a uniformly increasing sequence with an appreciation of 2, for example it can be 7,9,11,13... And we need to find the shortest series length(num of games) so that the probability of the Braves win the World Series given that PB = 0.55 is more than 0.8.

We use graph show the relationship between the num_games(x-axis) and the probability of the Braves win the World Series(y-axis). And we also add a line that show the probability of the Braves win the World Series is 0.8.

So we can see a intersection point which means when the the probability of the Braves win the World Series is 0.8, the shortest series length is 71.

```
PB=0.6
num_games=seq(7,151,by=2)
#num_games
num_wins= ceiling(num_games/2)
num_loss= num_games - num_wins
P_B_WS= pnbinom(num_loss,num_wins,PB)
#plot(num_games, P_B_WS)+ abline(h=0.8, col="red")
aa<-num_games[which(P_B_WS>=0.8)]
data.frame(PB,min(aa))
```

4, What is the shortest series length so that P(Braves win WS| PB=x)>=0.8? This will be a figure (see below) with PB on the x-axis and series length is the y-axis.

```
##    PB min.aa.
## 1 0.6      17
```

```
PB=0.65
num_games=seq(7,151,by=2)
#num_games
num_wins= ceiling(num_games/2)
num_loss= num_games - num_wins
P_B_WS= pnbinom(num_loss,num_wins,PB)
#plot(num_games, P_B_WS)+ abline(h=0.8, col="red")
aa<-num_games[which(P_B_WS>=0.8)]
data.frame(PB,min(aa))
```

```
##     PB min.aa.
## 1 0.65       7
```

```
PB=0.7
num_games=seq(7,151,by=2)
#num_games
num_wins= ceiling(num_games/2)
num_loss= num_games - num_wins
P_B_WS= pnbinom(num_loss,num_wins,PB)
#plot(num_games, P_B_WS)+ abline(h=0.8, col="red")
aa<-num_games[which(P_B_WS>=0.8)]
data.frame(PB,min(aa))
```

```
##    PB min.aa.
## 1 0.7       7
```

```
PB=1
num_games=seq(7,151,by=2)
#num_games
num_wins= ceiling(num_games/2)
num_loss= num_games - num_wins
P_B_WS= pnbinom(num_loss,num_wins,PB)
#plot(num_games, P_B_WS)+ abline(h=0.8, col="red")
aa<-num_games[which(P_B_WS>=0.8)]
data.frame(PB,min(aa))
```

```
##   PB min.aa.
## 1  1       7
```

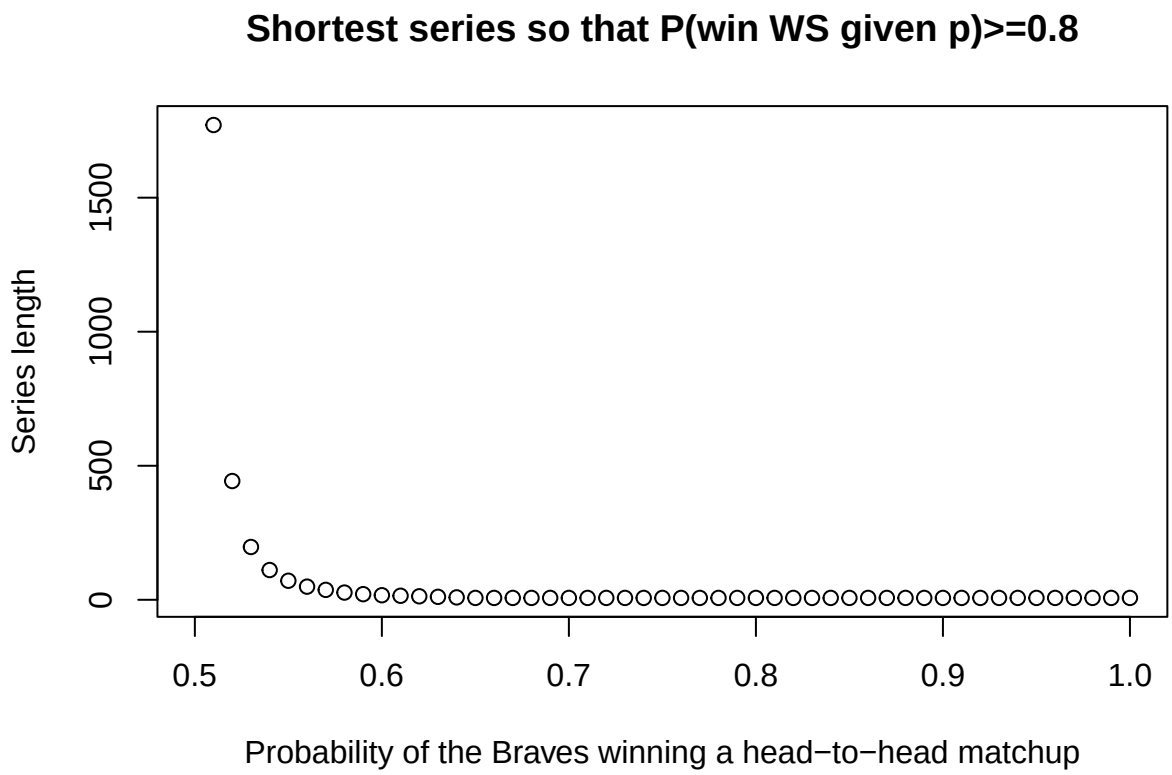
when pb=0.55 then number of games=71 when pb=0.6 then the number of games=17 when pb=0.65, then the number of games=7 when pb=0.7, then the number of games=7 ... when pb=1.0, them the number of games=7

```
x=seq(0.5,1,by=0.01)
y=rep(NA,length(x))

series <- function(x){
  num_games= seq(7,1773,by=2)
  num_wins= ceiling(num_games/2)
  num_loss= num_games - num_wins
  P_B_WS= pnbinom(num_loss,num_wins,x)
 return(num_games[which(P_B_WS>=0.8)][1])
}


for (i in 2:51){
 y[i]= series(x[i])
}

plot(x,y,main = "Shortest series so that P(win WS given p)>=0.8",xlab="Probability of the Braves winning
```



Solution:

I assume the probability of Braves win in each games is from 0.5 to 1.0. We set a function to loop through each PB. and get every the shortest series length.

Then I use graph show the relationship between sequence of PB (x-axis) and the shortest number of game they need to play to win the WS(y-axis).

I can easily find that as the probability of Braves wins each game gradually increases, the shortest number of games they need to play declines until equal to 7.

```
#P(B win WS in 7 games| Pb=0.55)
Pb=0.55
PB_w1=pnbinom(3,4,Pb)
PB_w1
```

5, Calculate P(PB=0.55|Braves win WS in 7 games) under the assumption that either PB=0.55 OR PB=0.45, Explain your solution.

```
## [1] 0.6082878
```

```
#P(B win WS in 7 games| Pb=0.45)
Pb=0.45
PB_w2=pnbinom(3,4,Pb)
PB_w2
```

```
## [1] 0.3917122
```

```
#P(B win WS in 7 games) -> denominator
sum=PB_w1+PB_w2

#P(PB = 0.55|Braves win World Series in 7 games)
PB_wws1=PB_w1*0.5/sum
#P(PB = 0.45|Braves win World Series in 7 games)
PB_wws2=PB_w2*0.5/sum

#show the results
data.frame(sum,PB_w1,PB_wws1,PB_w2,PB_wws2)
```

```
##   sum     PB_w1   PB_wws1     PB_w2   PB_wws2
## 1   1 0.6082878 0.3041439 0.3917122 0.1958561
```

Solution:

We make an assumpttion that either PB = 0.55 or PB = 0.45, so P(PB=0.55)=P(PB=0.45)=0.5

P(PB = 0.55|Braves win World Series in 7 games) means: Given Braves win WS in 7 games, what the probability of PB=0.55

For each Pb we need to compute P(B win WS in 7 games| Pb), them add them together :

P(B win WS in 7 games| Pb=0.55) + P(B win WS in 7 games| Pb=0.45) = P(B win WS in 7 games)

So I finally compute the results of :

P(PB = 0.55|Braves win World Series in 7 games)= P(B win WS in 7 games| Pb=0.55)*P(PB=0.55) / P(B win WS in 7 games)=0.3041439

P(PB = 0.45|Braves win World Series in 7 games)= P(B win WS in 7 games| Pb=0.45)*PP(PB=0.45) / P(B win WS in 7 games)=0.1958561

Conclusion

In summary, we talk about 5 questions about the probability of winning in World Series. We use the rules of probability and discrete probability functions to answer the questions, especially using *negative Binomial* cumulative probability to compute the probability of Brave wins in 7 games.