StarsGroup Exercise

Serge Nakache

August 14, 2020

Problem

Two players A and B play against each other in a tennis match. The match is played at the best of three sets. So there are two scenarios: a three-set match or a two-set match. Which scenario should I bet on to win money, knowing that you have no information about the two players.?

Probability of each scenario, given p

We define p as the probability that player A wins a set against B

- **Two sets match scenario** :
  - A wins the first, A wins the second : $P(AA) = p^2$
  - $P(BB) = (1-p)^2$
  - $P(Scenario2S) = p^2 + (1-p)^2$
- **Three sets match scenario** :
  - Since events are independant : $P(ABA) = P(BAA) = p^2(1-p)$
  - $P(BAB) = P(ABB) = p(1-p)^2$
  - $P(Scenario3S) = 2p - 2p^2$

Most probable scenario

Now we want to know what the most likely scenario is. One way to do this is to calculate the quotient `Q = P(Scenario 2 sets) / P(Scenario 3 sets)`. If this quantity is greater than 1, then the 2-set match is the most likely scenario. If this quantity is less than 1, then the 3-set match is the most likely scenario. Knowing that we know nothing about the players, we assume that p follows a uniform law on [0,1].

$$Q = \frac{\int_P(Scenario2S)}{\int_P(Scenario3S)} = \frac{\int_0^1 p^2 + (1-p)^2 dp}{\int_0^1 2p - 2p^2 dp} = 2$$

Given that Q>1, we can say that the most probable event is the scenario of a match in two sets. So we want to bet on the **two sets match scenario**.
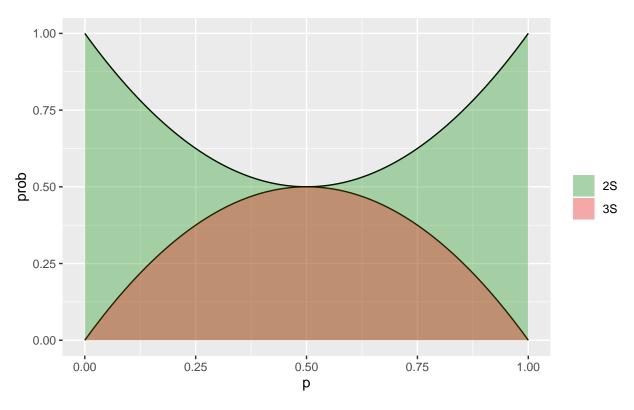

Vizualisation

```
# Build probabilites vector and p
p = seq(0,1,0.01)
two_sets = p^2+(1-p)^2
three_sets = 2*p - 2*p^2
```

1

```
# Needed dataframe as input for ggplot function
a = rep(x = "2S",length(two_sets))
b = rep(x = "3S",length(three_sets))
groupv = c(a,b)
df = data.frame(p = c(p,p), prob = c(two_sets,three_sets), group=groupv)
```

```
ggplot(df, aes(x=p, y=prob, group=groupv, fill=groupv)) +
  geom_line(size=.5) +
  geom_ribbon(data=subset(df,p>=0 & p<=1),aes(x=p,ymax=prob),ymin=0,alpha=0.3) +
  scale_fill_manual(name='', values=c("2S" = "green4", "3S" = "red"))+
  ggtitle(label = "Probability of the two scenarios given p",subtitle = " ")
```


Probability of the two scenarios given p

As you can see the green area representing *2 sets match scenario* is much larger than the red area *3 sets match scenario*. Scenario 2 sets is therefore much more likely. Green area is supposed to be twice the red area.