S&DS 238 Problem Set #3

Shaun Radgowski

September 25, 2020

Problem 1

The probability of getting X successes from n independent trials, each with probability θ, is in this case just the probability distribution for a binomial.

$$P\{X\} = \binom{n}{X}\theta^X(1-\theta)^{n-X} = \frac{n!}{X!(n-X)!}\theta^X(1-\theta)^{n-X}$$

To find the value of θ that maximizes this function, we take the derivative with respect to θ and set it equal to 0.

$$\tfrac{dP}{d\theta} = \tfrac{n!}{X!(n-X)!}(X(1-\theta)^{n-X}\theta^{X-1} - (n-X)(1-\theta)^{n-X-1}\theta^X) = 0$$

$$X(1-\theta)^{n-X}\theta^{X-1} = (n-X)(1-\theta)^{n-X-1}\theta^X$$

$$X(1-\theta) = (n-X)\theta, \quad X - X\theta = \theta n - \theta X$$

$$\theta = \tfrac{X}{n}$$

This value of theta maximizes the probability function because the probability function is concave down. We know this because this is a binomial distribution, which is bell-shaped with lower probabilities at the edges and one single global maximum point in the middle, so the single extreme point (where the derivative equals 0) must be a global maximum point instead of a local maximum. This can be confirmed by noting that the second derivative $l''(\theta) = -\frac{X}{\theta^2} - \frac{n-X}{(1-\theta)^2}$ is always negative, signaling that the function is concave down and the local extreme point is indeed a maximum point.

Problem 2

(2a)

The probability of team 1 winning the series on game k is the same as the probablity of team 1 winning any 3 of games 1 through $k-1$ times the probability of winning the next game (which would be game k). The probability of winning 3 of games 1 through $k-1$ is thus a binomial distribution for $k-1$ games at probability p each:

$$P\{X=3\} = \binom{k-1}{3}p^3(1-p)^{k-4}$$

Multiplying by the probability of winning the next game (game k) thus yields the probability that team 1 wins the series on game k. This is only possible on for rounds 4-7, which are the only possible values of k.

$$P\{X=3\}p = \binom{k-1}{3}p^4(1-p)^{k-4} \quad \text{for } k = 4,5,6,7$$

(2b)

The total probability that team 1 wins the world series is just the sum of their probabilities of winning across rounds 4-7. The winner will need to win on one of those four rounds, so the sum of those four games' probabilities equals the total probability of winning because all of the options for winning are disjoing (one team cannot win on multiple rounds). For $p = 0.6$, the total probability that team 1 wins is thus $0.1296 + 0.20736 + 0.20736 + 0.165888 =$ **0.710208**.

(2c)

```
winning_probability <- function(p){
  games <- 3:6
  sum(dbinom(3, games, p) * p)
}

winning_probability(0.6)
```

```
## [1] 0.710208
```

(2d)

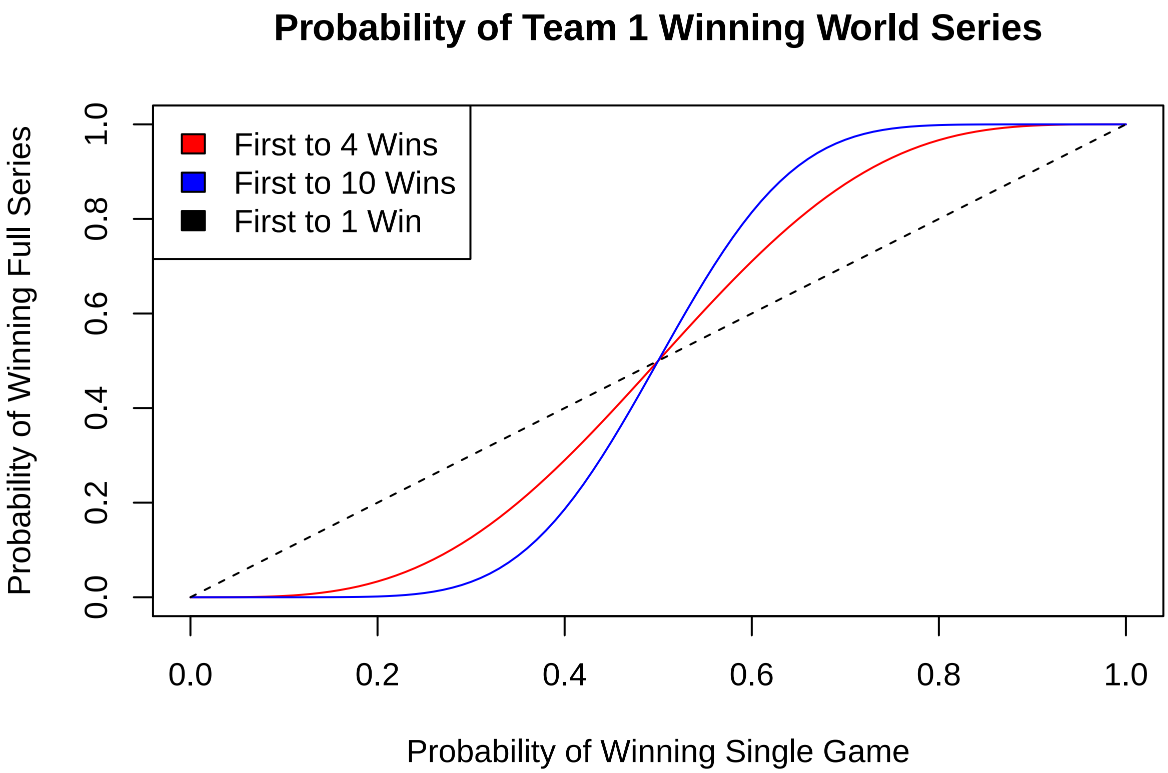
```
winning_probability_new <- function(p){
  games <- 9:18
  sum(dbinom(9, games, p) * p)
}

winning_probability_new(0.6)
```

```
## [1] 0.813908
```

(2e)

```
x <- seq(0, 1, by=0.01)
y1 <- seq(0, 1, by=0.01)
y2 <- seq(0, 1, by=0.01)
for (i in 1:101) {
  y1[i] = winning_probability(x[i])
  y2[i] = winning_probability_new(x[i])
}
plot(x, y1, main="Probability of Team 1 Winning World Series", col="red",
     ylab="Probability of Winning Full Series", xlab="Probability of Winning Single Game", type="l")
lines(x, y2, col="blue")
lines(x, x, col="black", lty=2)
legend("topleft", c("First to 4 Wins", "First to 10 Wins", "First to 1 Win"),
       fill=c("red", "blue", "black"))
```



This graph suggests that giving a team more chances to succeed (first to 10 wins being more chances than first to 4 wins, which is in turn more chances than first to 1 win) heightens whatever advantage or disadvantage they already possess. If they have a high likelihood of winning one game, they'll have an even higher likelihood of winning the tournament if the rules are first to four wins or first to ten wins. This makes sense qualitatively, as a higher sample size of games will allow the true ratio of victories to more closely approximate the two teams' skill levels by the law of large numbers. If one team has a 90% chance of winning a single game, they could still have one fluke bad game if they only play once. But if they play 19 games, the chance that they'll have enough flukes to lose the tournament is much lower.

Problem 3

This is equivalent to saying that $P(A|R) = 0.6$. According to Bayes' Rule:

$$P(A|R) = \frac{P(R|A)P(A)}{P(R)}, \quad P(A^C|R) = \frac{P(R|A^C)P(A^C)}{P(R)}$$

$$0.6 = \frac{0.8P(A)}{P(R)}, \quad (1-0.6) = \frac{0.5(1-P(A))}{P(R)}$$

$$P(A) = 0.75P(R), \quad P(R) = 1.25 - 1.25P(A)$$

$$P(A) = 0.75(1.25 - 1.25P(A)) = 0.9375 - 0.9375P(A)$$

$$P(A) = \frac{0.9375}{1.9375} \approx 0.484$$

Problem 4

We'll first discretize the parameter space for (v_F, v_B) to consist of the 1001x1001 grid of points given by $\{(\theta_F, \theta_B)\}$ where each can take on any discrete value in the set $\{0, 0.001, 0.002, \ldots, 1\}$.

We'll assume each point in this parameter space (θ_F, θ_B) has an equal probability of $\frac{1}{1001^2}$. Given the actual proportions v_F and v_B, we can assume that $X_F \sim \text{Bin}(237, v_F)$, $X_B \sim \text{Bin}(215, v_F)$. The sum of the (scaled) posterior probabilities where $v_B > v_F$ is thus our desired probability.

```r
post <- function(f, b) {
  probf <- dbinom(8, 237, f)
  probb <- dbinom(11, 215, b)
  return(probf * probb)
}

total <- 0
x <- 0
for (i in 1:1000) {
  for (j in 1:1000) {
    total <- total + post(i*0.001, j*0.001)
    if (j > i) {
      x <- x + post(i*0.001, j*0.001)
    }
  }
}

x/total
```

```
## [1] 0.8112291
```