# Assigment 3

Aaron Mulvey

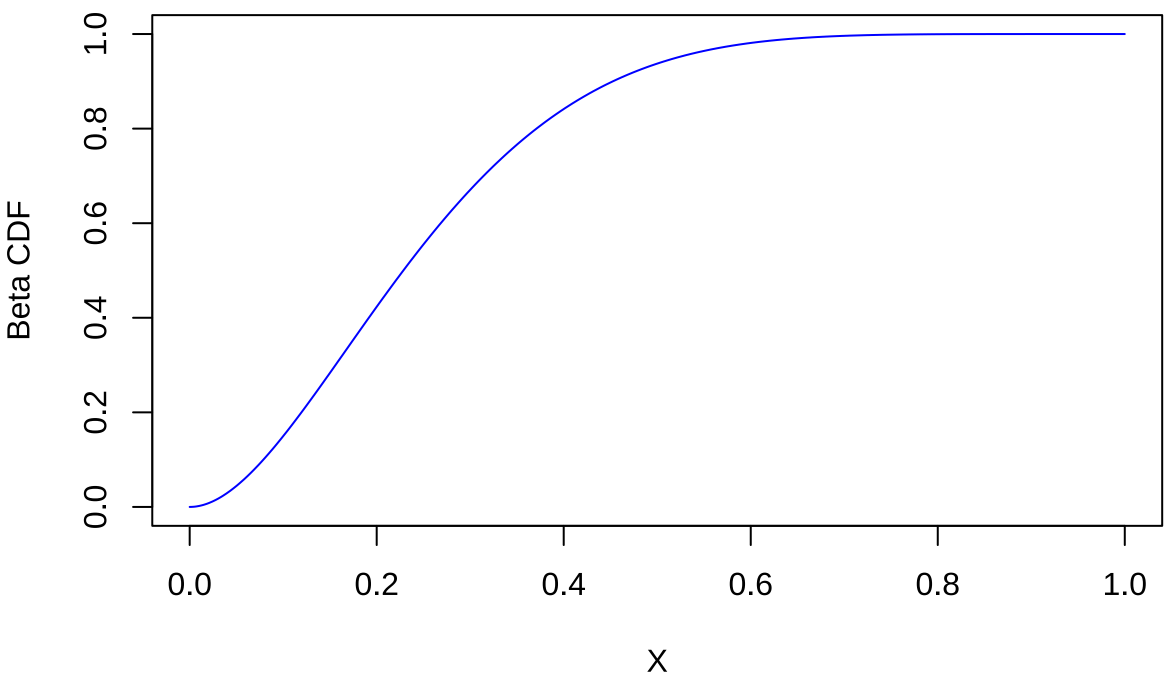
10/7/2022

##Question 1 (Median)

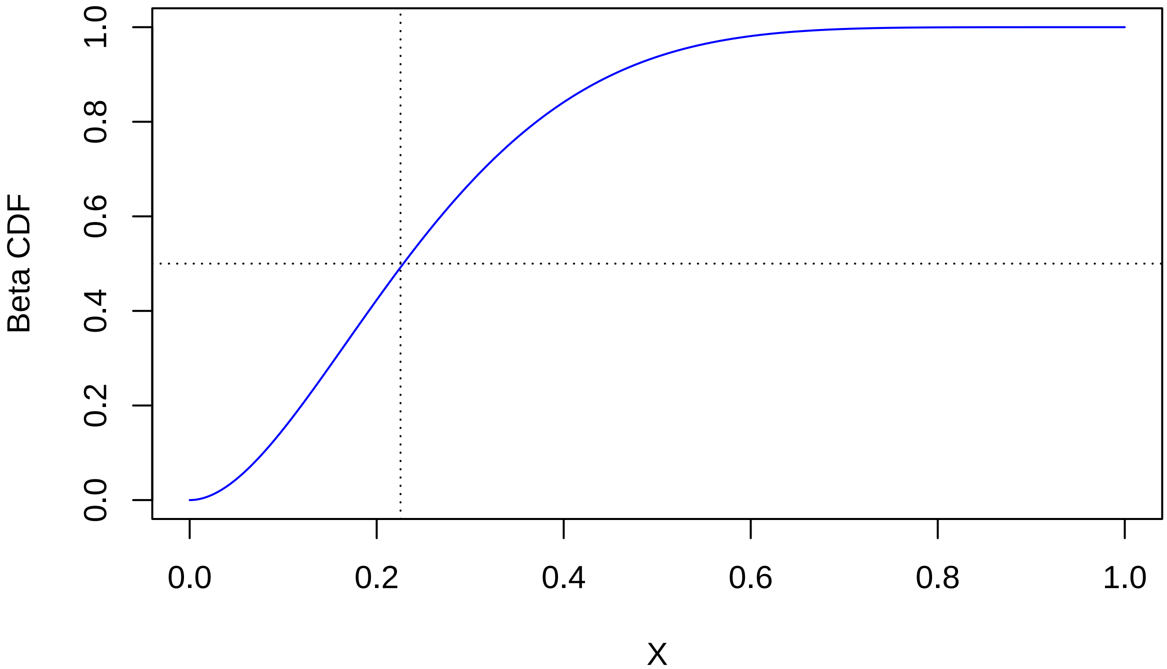
In the last homework assignment, you calculated the expectation and the variance of the Beta distribution with parameters a=6 and b=2. This time around, find the median. Give a numerical solution without using pbeta and check graphically if your solution makes sense. Tip: Use uniroot in R.

I am going to plot the CFD of the beta distribution.

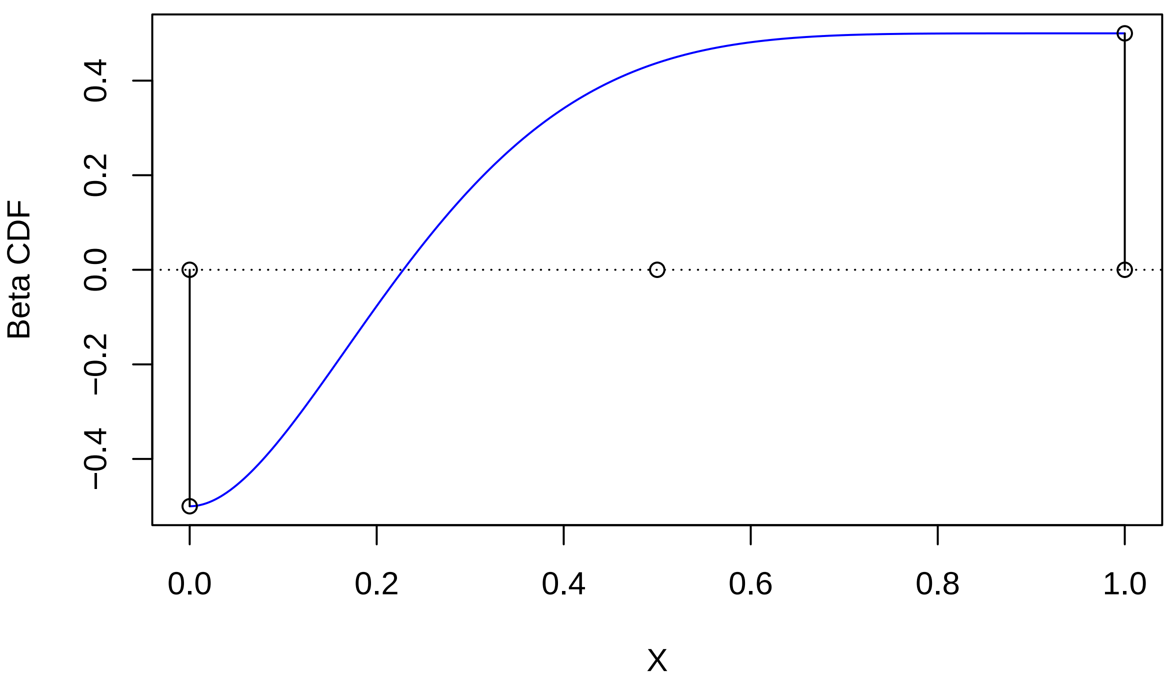
```
alpha = 6
beta = 2

secuencia = seq(0, 1, by = 0.001)

plot(secuencia, pbeta(secuencia, 2,6), xlab="X",
     ylab = "Beta CDF", type = "l", col = "blue")
```



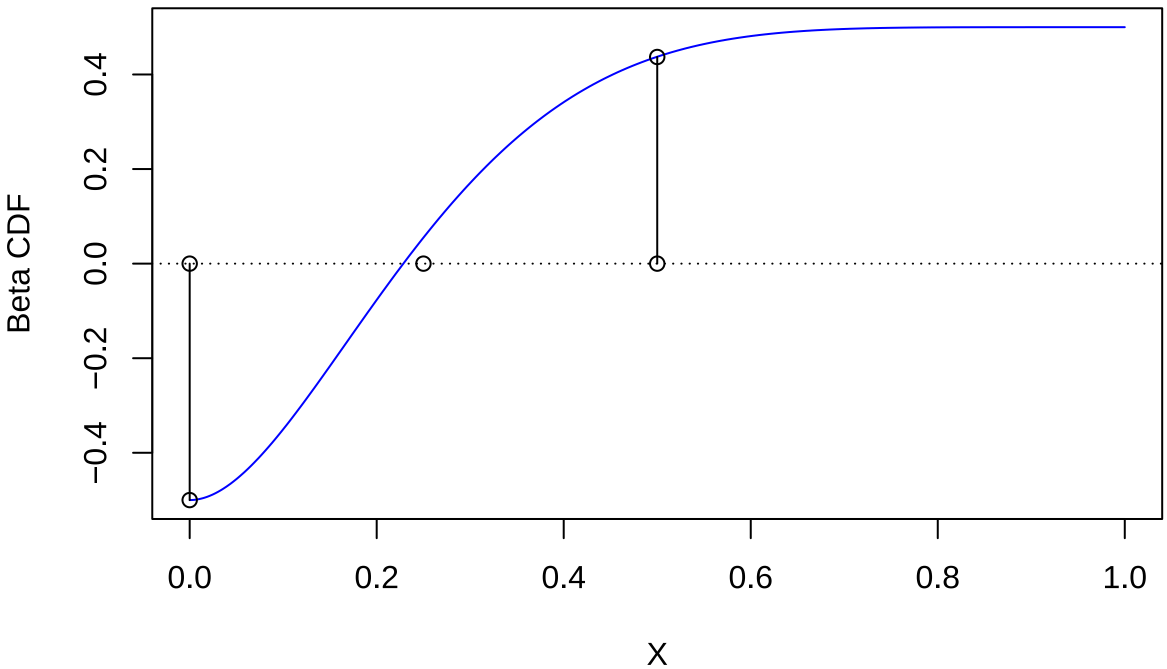

The median of the beta distribution is a real number that F(X < m) = 0.5 and F(X > m) = 0.5, so I will add some lines for see where the median should be.




Now I want to get that point in the function, but my function is behind 0 so I will subtract 0.5 so that I can get root.




Next I will use bisection method for getting the roots of my function.




If I continue like this I will reach the root therefore my solution.

But we cannot use pbeta so I will create my own method that will integrate the beta distribution for the given x so that I am not using pbeta

First I am going to create a function that integrates the beta function for a given X.

```
BetaDistribution <- function(x, alpha, beta){
  return (1 - integrate(dbeta, lower=0, upper=x, alpha, beta)$value)
}
```

Now I am going to use a bisection method for geting the root

```
median <- function(alpha, beta) {
  return(uniroot(function(x) BetaDistribution(x,alpha,beta) - 0.5, lower = 0, upper = 1)$root)
}

median(alpha, beta)
```

```
## [1] 0.7714982
```

I saw an aproximation for the median that is:

```
aproximation <- function(alpha,beta){
    return ((alpha - 1/3) / (alpha + beta-2/3))
}

aproximation(alpha, beta)
```

```
## [1] 0.7727273
```

Now I will calculate the error of the aproximation

```
sqrt((aproximation(alpha, beta) - median(alpha, beta))^2)
```

```
## [1] 0.001229036
```

##Question 2

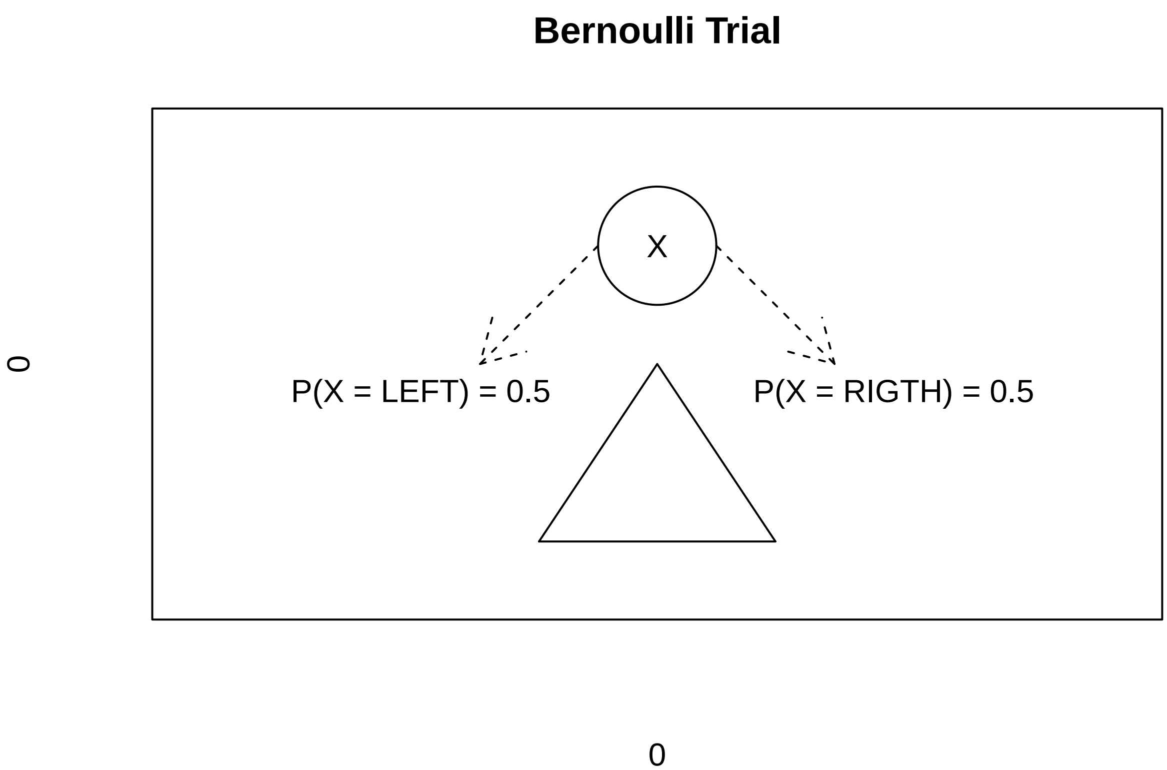
Answer the question in topic 8 on slide 14: Can you use the central limit theorem to explain why the histogram produced should look approximately Normal? Give a mathematical explanation and provide the parameters of the Normal distribution.

Can you use the central limit theorem to explain why the histogram produced should look approximately Normal?

When one ball reach the triangle, we can see that the ball has two possibilities, either go to the left or go to the right. So the ball has to take one direction (D), and the probability of taking one direction or the other is a Bernoulli trial.

$$D = \begin{cases} 0 & \text{if } x = \text{Left} \\ 1 & \text{if } x = \text{Rigth} \end{cases}, D \sim Bernoulli(P = 0.5)$$

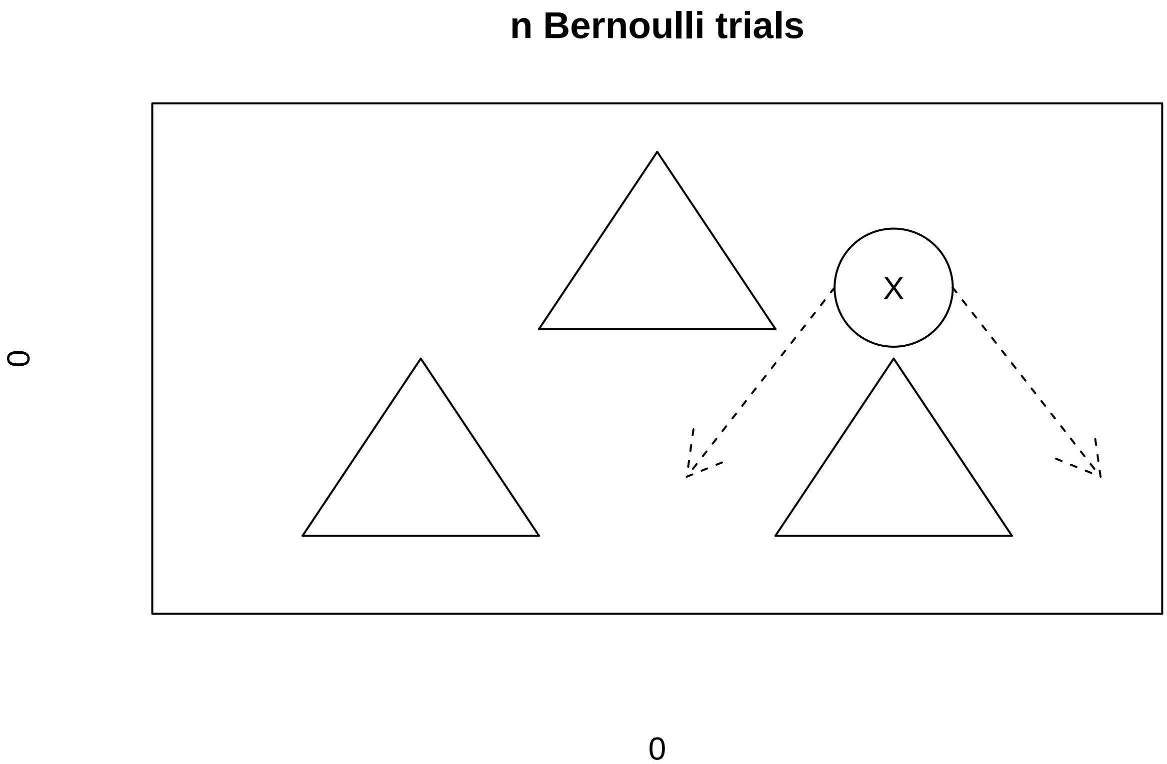
```
t <- seq(0, 2 * 3.1416, length.out = 100)
a <- 0
b <- 1
radio <- 0.5
x <- a + cos(t)*radio
y <- b + sin(t)*radio

plot(0, 0, asp = 1, type = "n", xlim = c(-2, 2), ylim = c(-2, 2), xaxt = "n", yaxt = "n")
title("Bernoulli Trial")



segments(-1, -1.5, 1,-1.5)
segments(-1, -1.5, 0, 0)
segments(0, 0, 1,-1.5)
arrows(0.5, 1, 1.5, 0, lty = 2)
arrows(-0.5, 1, -1.5, 0, lty = 2)
lines(x, y)
text(x = 0, y = 1, label='X')
text(-2, -0.25, label='P(X = LEFT) = 0.5')
text(2, -0.25, label='P(X = RIGTH) = 0.5')
```




At the end the ball will finish in one box or in the other one depending the direction that takes.

Now lets assume that we have n levels of triangles, so the ball can reach different boxes. The total number of boxes that the ball can reach is n + 1.

```
t <- seq(0, 2 * 3.1416, length.out = 100)
a <- 2
b <- 0.6
radio <- 0.5
x <- a + cos(t)*radio
y <- b + sin(t)*radio

plot(0, 0, asp = 1, type = "n", xlim = c(-2, 2), ylim = c(-2, 2), xaxt = "n", yaxt = "n")
title("n Bernoulli trials")
lines(x, y)
text(x = a, y = b, label='X')

arrows(2.5, 0.6, 3.75, -1, lty = 2)
arrows(1.5, 0.6, 0.25, -1, lty = 2)

segments(-1, 0.25, 1,0.25)
segments(-1, 0.25, 0, 1.75)
segments(1, 0.25, 0, 1.75)

segments(-3, -1.5, -1,-1.5)
segments(-3, -1.5, -2, 0)
segments(-1, -1.5, -2, 0)

segments(3, -1.5, 1,-1.5)
segments(3, -1.5, 2, 0)
segments(1, -1.5, 2, 0)
```




As you can see there are 2 levels so n = 2 and the possible boxes where the ball can go is n + 1 = 3.

The box in which that the ball will end is a sequence of n Bernoulli trials. We can name the boxes like this:

$$Box = \text{0, 1,...,n}$$

So for know in which box is the ball what we have to check is the sum of all the individual Bernoulli trials:

$$Ball_{Box} = \text{0, 1, 1, 1, 0,..., n}$$

We can say that the box in which the ball will end is a random variable X, that is distributed as a binomial distribution. Where the n represents the set of boxes, p the probability that is 0.5 and x select a box.

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

So the set of boxes are distributed as:

$$X \sim Binomial(n, p)$$

##Central limit theorem

Let

$$X_1, X_2, X_3, ..., X_n$$

Be random independent variables identically distributed with

$$E[X_i] = \mu \text{ and } Var(X_i) = \sigma^2 < \infty$$

We say that,