Discrete Probability Distributions

Contents

Visualizing Discrete Distributions

In this problem we will draw plots for some of the discrete probability distributions that we have studied in class

Binomial distribution.

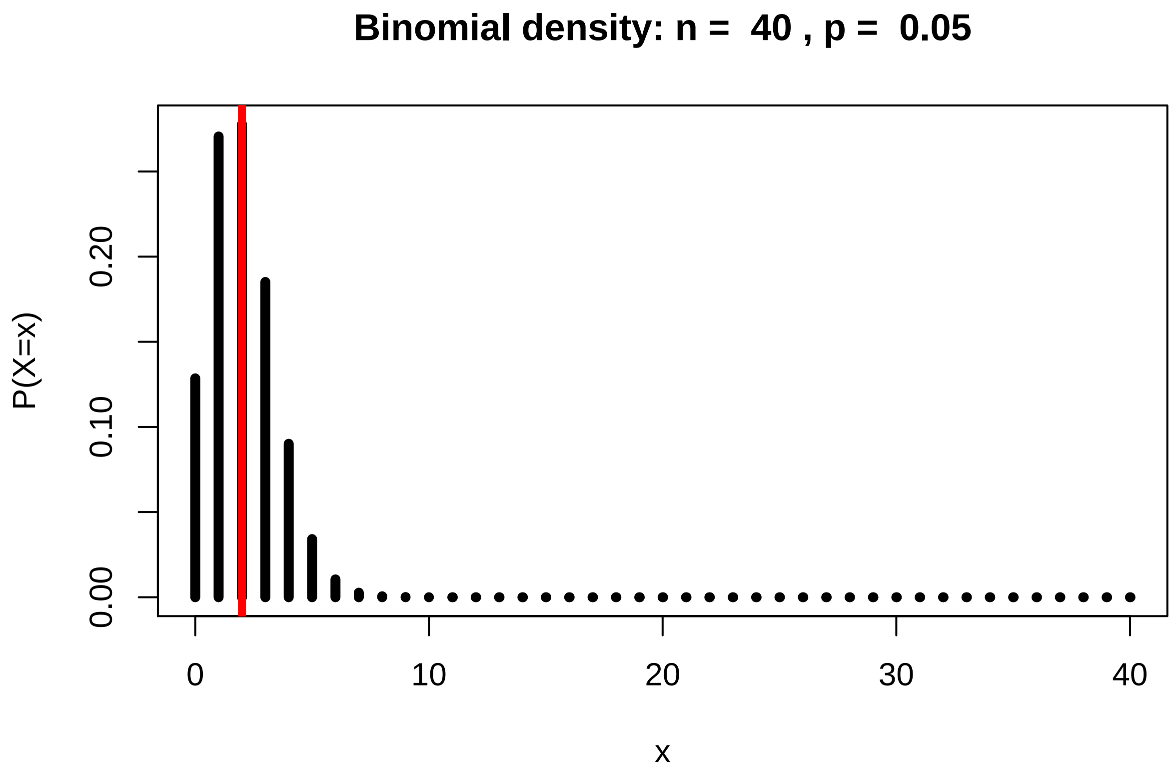
We will plot the density function for the binomial distribution Bin(n, p). Note: 1) The values for this random variable are 0, 1, 2, . . . , n. 2) The density plot will have a bar of height P(X=k), at the point 'k' on the x-axis. 3) In the plot include a vertical line at the expected value of Bin(n,p).

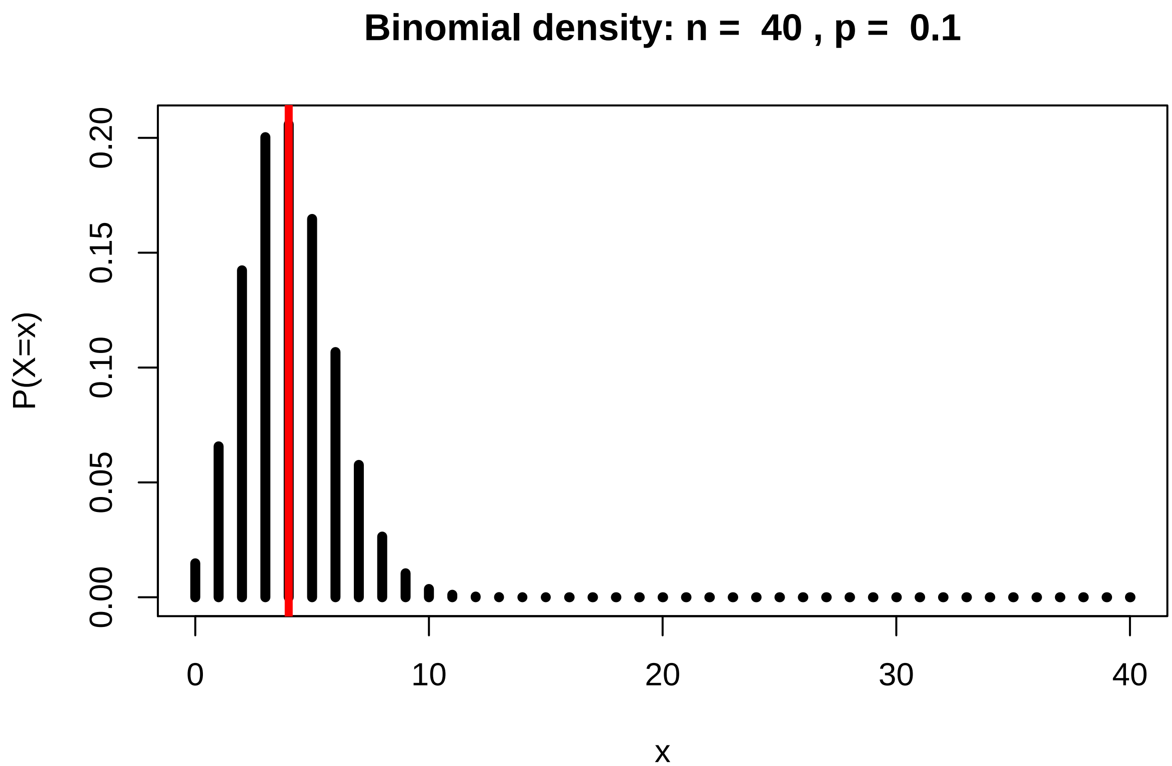
Write a function plot_binom, that takes input values: n and p, and returns the density plot of Bin(n,p).

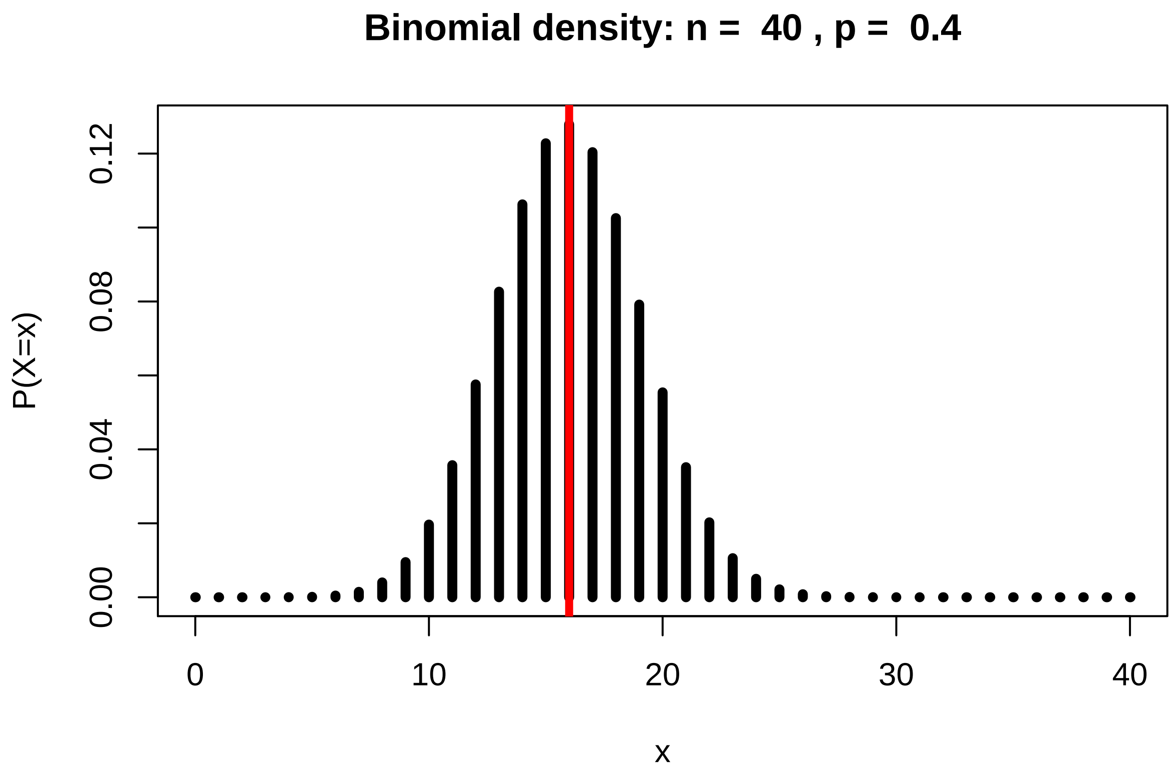
```
plot_binom <- function(n, p){
  x <- 0:n
  binom <- dbinom(x, n, p)
  mu_x <- n * p
  plot(x, binom, type = "h", lwd = 5,
       main = paste("Binomial density: n = ", n, ", p = ", p),
       xlab = "x", ylab = "P(X=x)")
  abline(v = mu_x, col = 'red', lwd = 4)
}
```

Fix n = 40. Compute plots for the following values of p: 0.05, 0.1, 0.4, 0.6, 0.9, 0.95. Use the command "par(mfrow=c(3,2))" to have all the plots on the same frame.

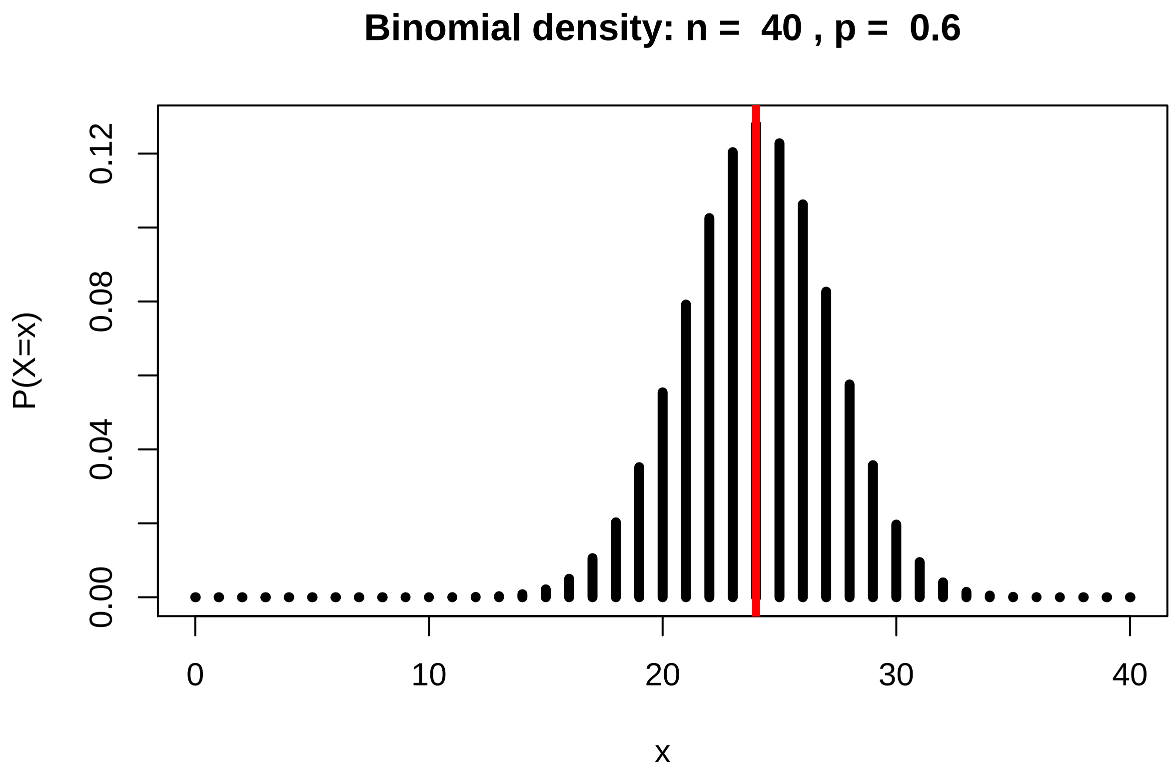
```
n <- 40
for (p in c(0.05, 0.1, 0.4, 0.6, 0.9, 0.95)) {
  plot_binom(n, p)
}
```

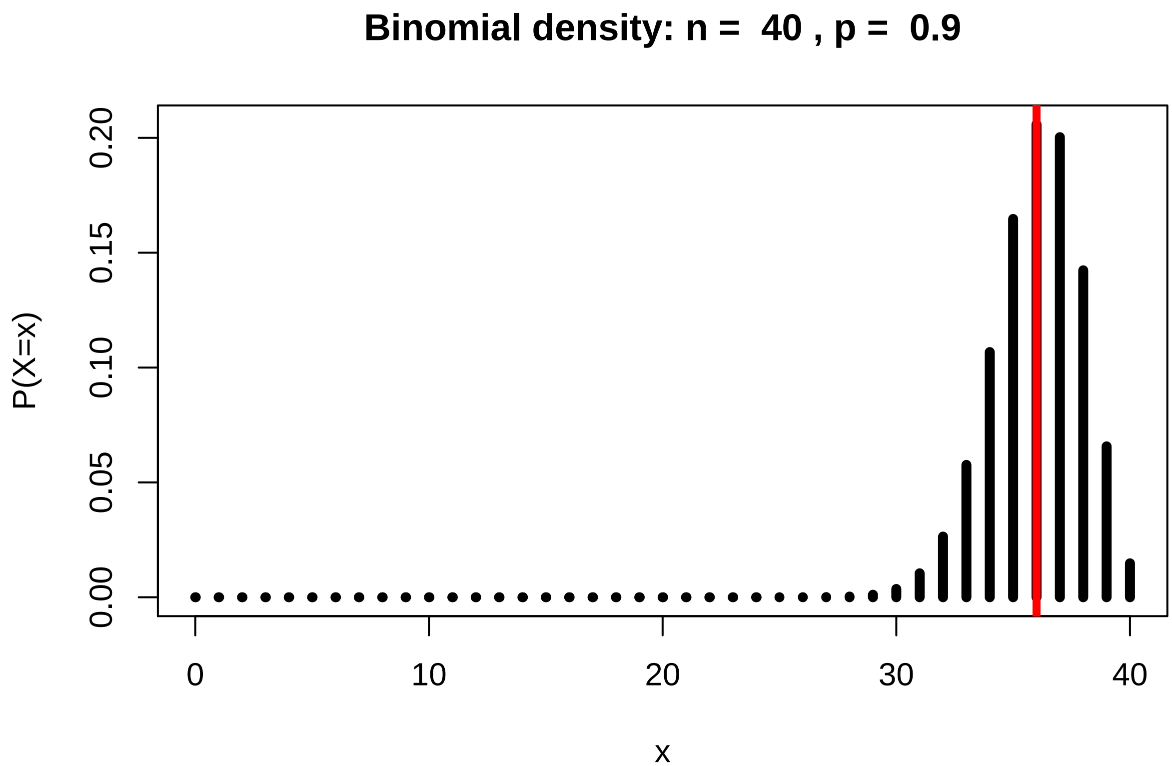

Binomial density: n = 40 , p = 0.05
P(X=x)
0.00
0.10
0.20
0
10
20
30
40
x

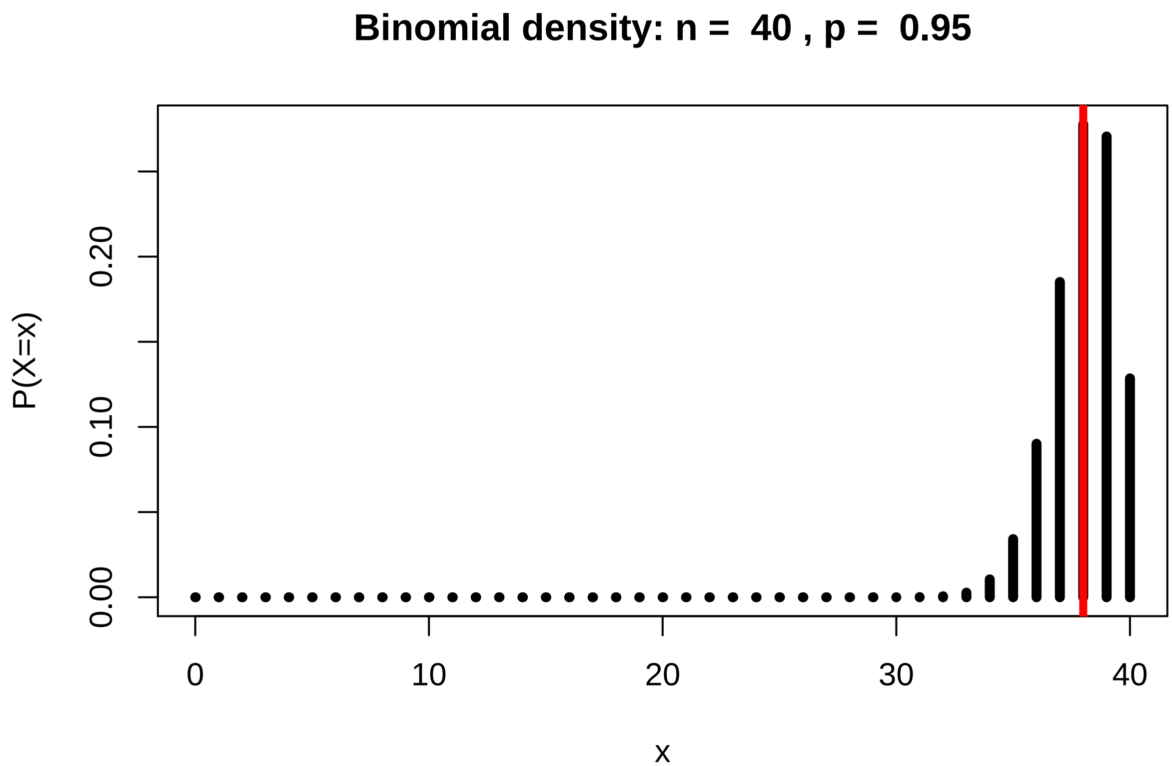

Binomial density: n = 40 , p = 0.1
P(X=x)
0.00
0.05
0.10
0.15
0.20
0
10
20
30
40
x



Binomial density: n = 40 , p = 0.4
P(X=x)
0.00
0.04
0.08
0.12
0
10
20
30
40
x


Binomial density: n = 40 , p = 0.6
P(X=x)
0.00
0.04
0.08
0.12
0
10
20
30
40
x


Binomial density: n = 40 , p = 0.9
P(X=x)
0.00
0.05
0.10
0.15
0.20
0
10
20
30
40
x



Write at least two observations that you can note from these plots. Consider skewness and symmetry. 1) The distribution seems to reach perfect symmetry at $p = 0.5$. 2) When $p < 0.5$, the distribution is left-skewed, and when $p > 0.5$, the distribution is right-skewed. The farther p is from 0.5, the more skewed the distribution is. 3) $p = q$ and $p = 1 - q$ yield the same plot, but mirrored.

Poisson Distribution.

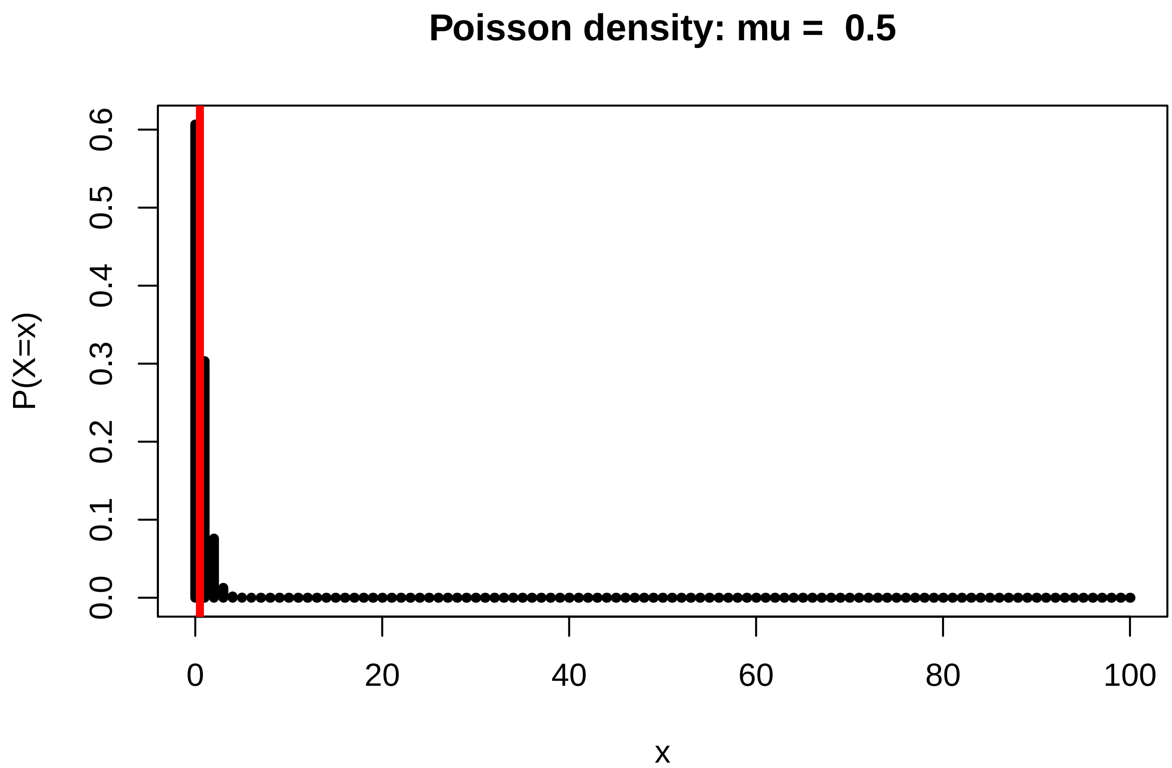
We will plot the density function for the Poison distribution Pois(mu). Note: 1) The values for this random variable are: 0, 1, 2, 3, ... 2) The density plot will have a bar of height P(X=k), at the point 'k' on the x-axis. 3) Since most of the densities will be concentrated at lower values of k, we will fix a large enough value of n, say n = 100, when drawing the density plots. 3) In the plot include a vertical line at the expected value of Pois(mu).

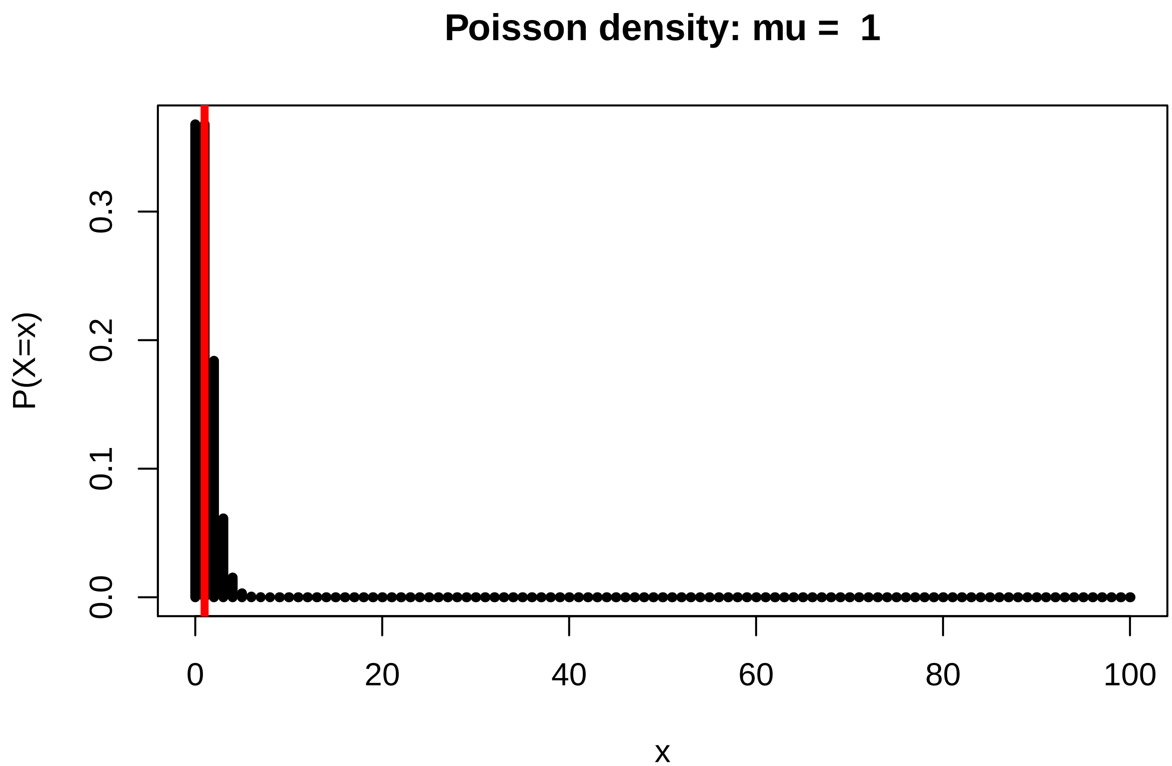
Write a function plot_pois, that takes input values: mu, and returns the density plot of Pois(mu).

```
n <- 100
plot_pois <- function(mu){
  x <- 0:n
  pois <- dpois(x, lambda = mu)
  plot(x, pois, type = "h", lwd = 5,
       main = paste("Poisson density: mu = ", mu),
       xlab = "x", ylab = "P(X=x)")
  abline(v = mu, col = 'red', lwd = 4)
}
```

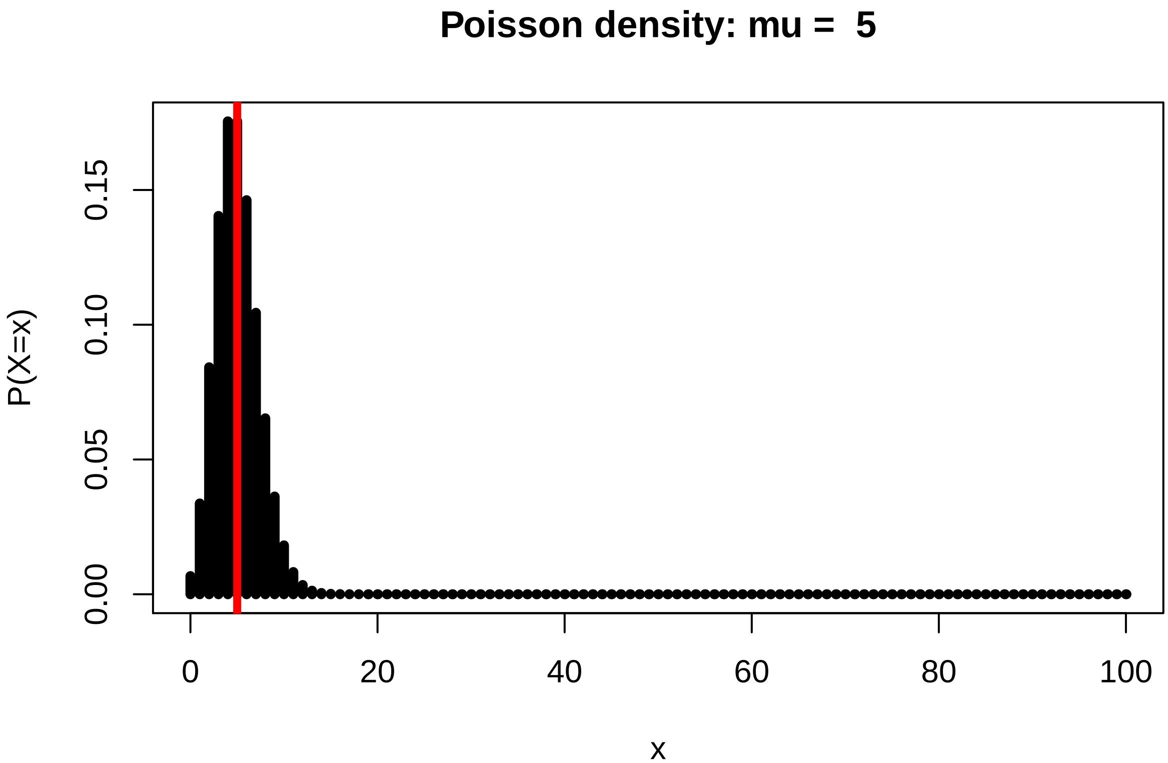
For the following values of mu compute the plots for the Poisson Density: mu: 0.5, 1, 5, 8, 20, 50

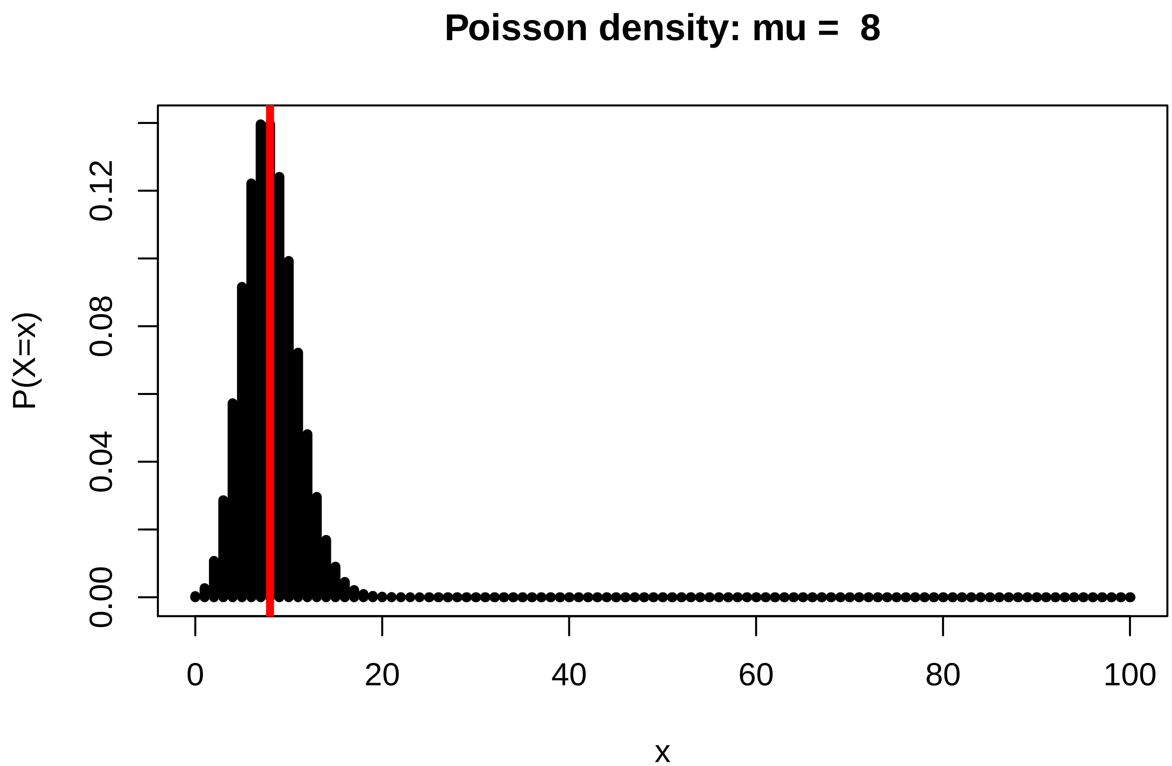
```r
for (mu in c(0.5, 1, 5, 8, 20, 50)) {
  plot_pois(mu)
}
```

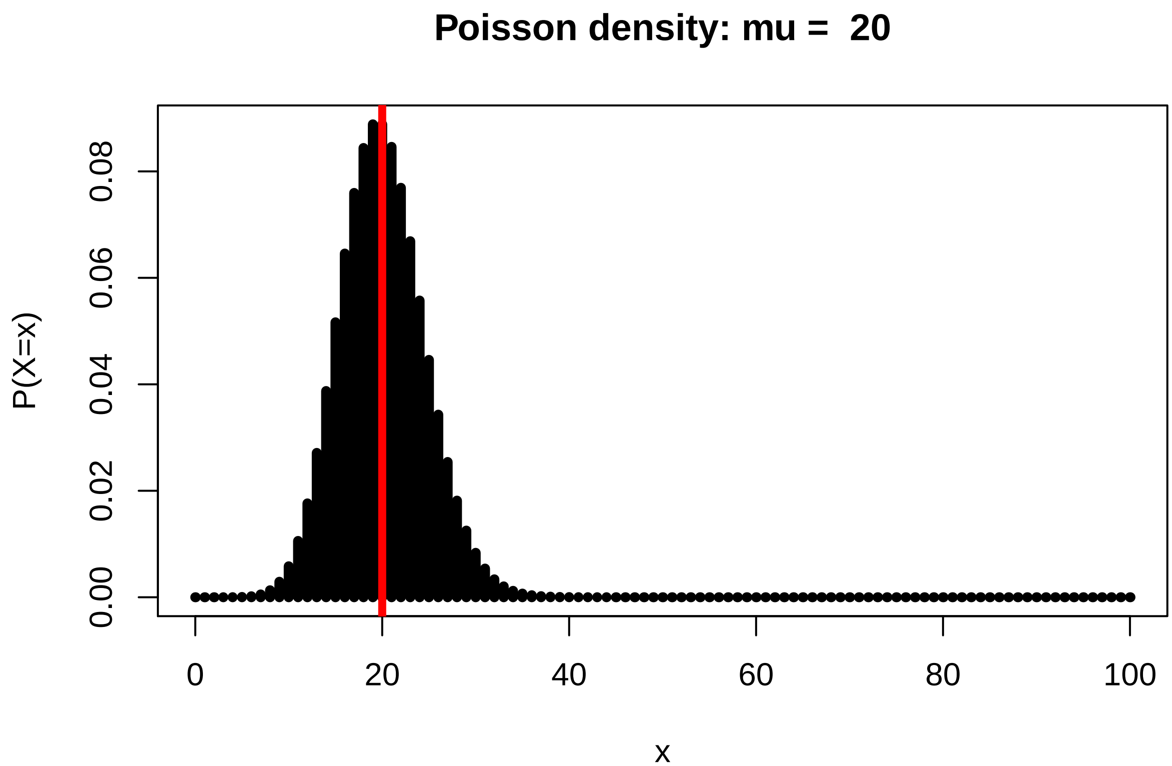


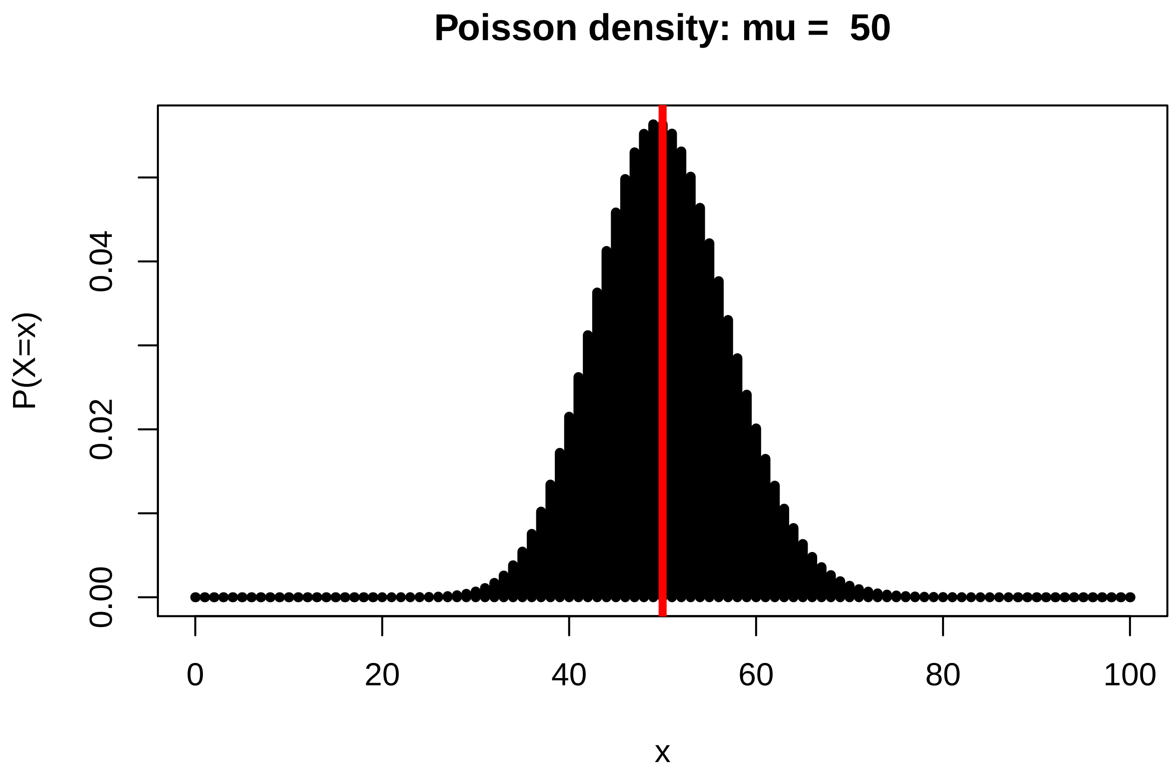


Poisson density: mu = 1
P(X=x)
0.0
0.1
0.2
0.3
0
20
40
60
80
100
x


Poisson density: mu = 5
P(X=x)
0.00
0.05
0.10
0.15
0
20
40
60
80
100
x


Poisson density: mu = 8
P(X=x)
0.00
0.04
0.08
0.12
0
20
40
60
80
100
x

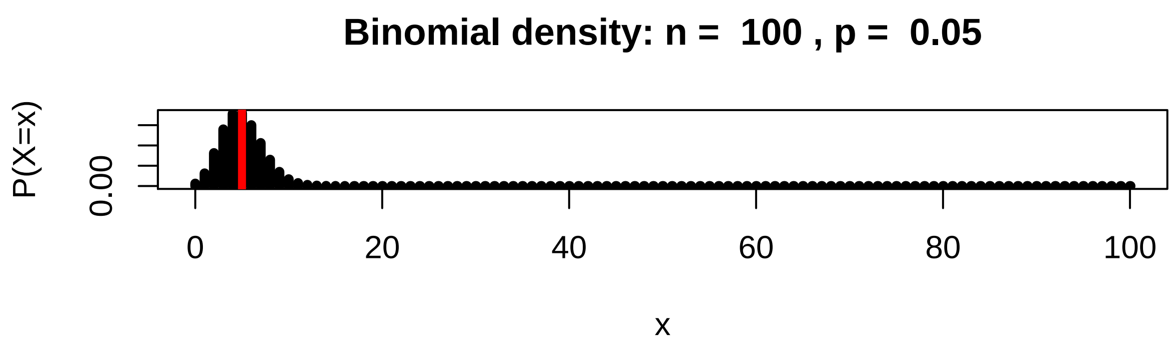

Poisson density: mu = 20
P(X=x)
0.00
0.02
0.04
0.06
0.08
0
20
40
60
80
100
x

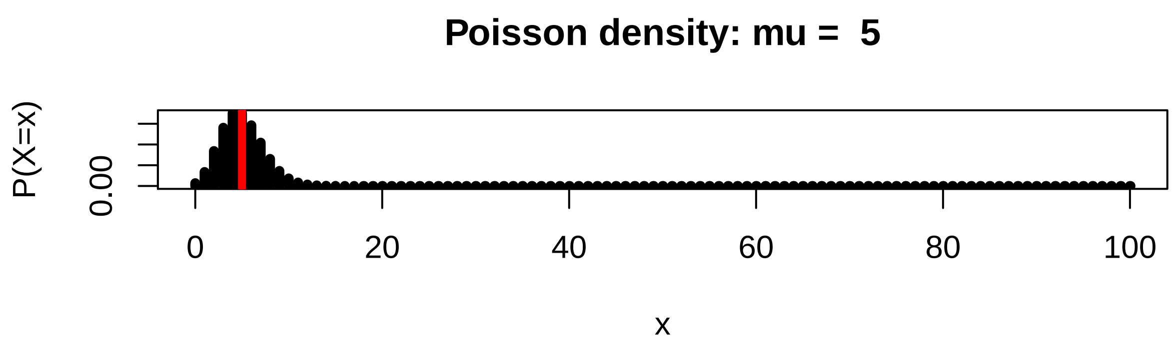


What observations can you make about the density plots of the Poisson distribution for large values of mu? Answer: 1) Poisson distribution is not symmetrical, although it is close to being symmetrical the larger μ is. 2) Poisson distribution is left-skewed, and it's more skewed the smaller μ is. 3) For large μ, the variance is higher, and the maximum value is lower.

Now use your plot functions to evaluate the following two plots on the same frame (one below the other): plot_binom(100, 0.05) plot_pois(5)

```
par(mfrow = c(2,1))
plot_binom(100, 0.05)
plot_pois(5)
```





What observations can you make? Answer: The two plots are very similar. This confirms the proposition that the binomial distribution in limit approaches the Poisson distribution given that np approaches μ.

Plots From HW5

Recall problem from HW5. \ Yulia owns a coffee distribution company. A very large batch of bags of coffee beans has arrived. She will accept the batch only if the proportion of underweight bags is at most 0.10. (Hint: Use the Binomial distribution with appropriate parameters for this problem)

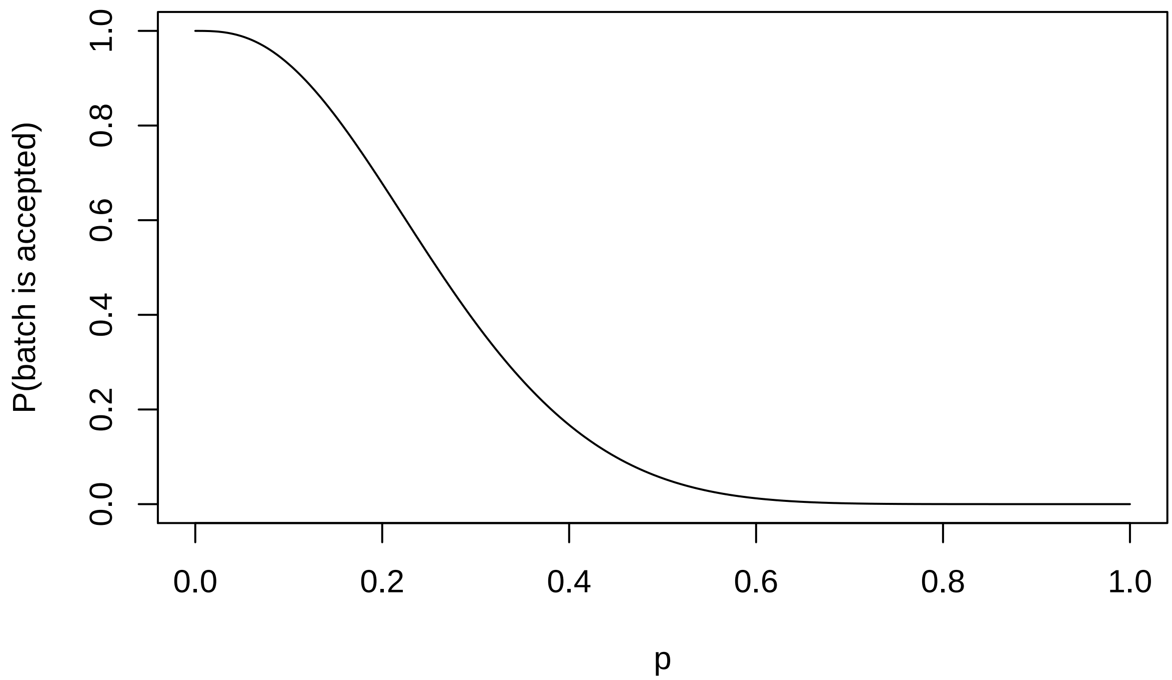
Case 1:

Yulia decides to randomly select 10 bags of coffee, and will accept the batch only if the number of underweight bags in the sample is at most 2. \ Let p denote the actual proportion underweight bags in the batch. Calculate P(batch is accepted) as a function of p, call this function 'at_most2' and plot it (domain is (0,1)).

```
yulia <- function(n, m, p) {
  y <- 0
  for (i in 0:m)
    y <- y + choose(n, i) * p^i * (1-p)^(n-i)
  y
}

at_most2 <- function(p) yulia(10, 2, p)
x <- 0:1000 / 1000
y <- at_most2(x)
plot(x, y, type = "l", lwd = 1,
```

```
           xlab = "p", ylab = "P(batch is accepted)")
```

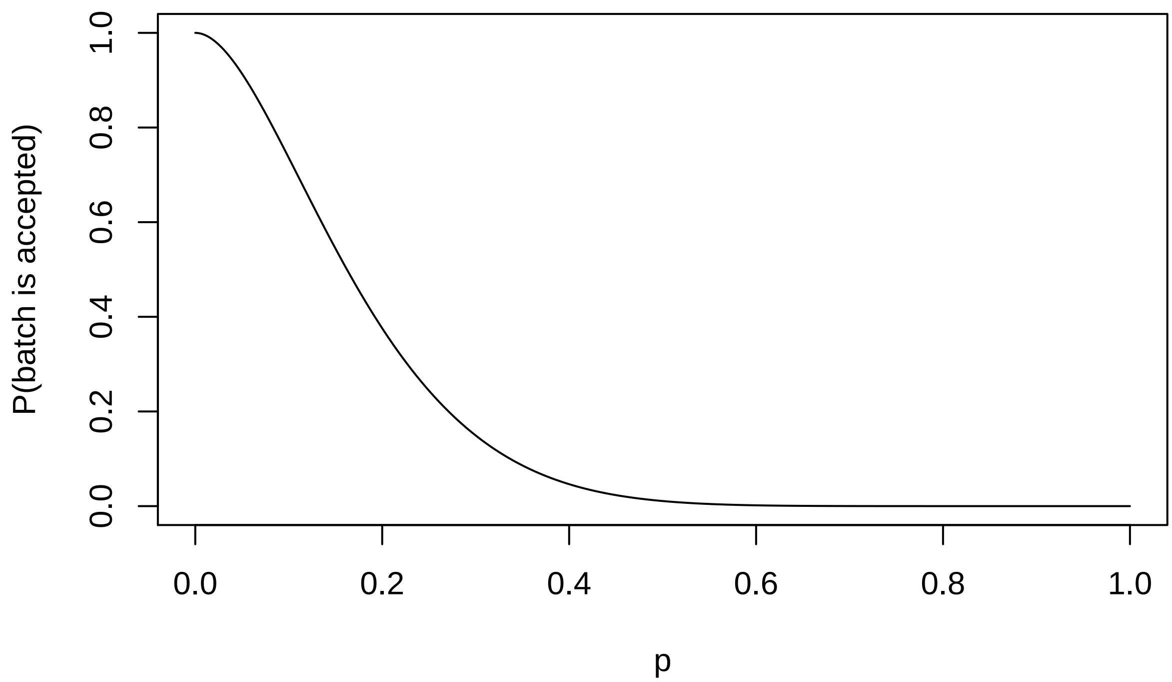


Case 2:

Now suppose Yulia decides to randomly select 10 bags of coffee, and will accept the batch only if the number of underweight bags in the sample is at most 1. \

Let p denote the actual proportion underweight bags in the batch. Calculate $P(\text{batch is accepted})$ as a function of p, call this function 'at_most1' and plot it (domain is (0,1)).

```
at_most1 <- function(p) yulia(10, 1, p)
x <- 0:1000 / 1000
y <- at_most1(x)
plot(x, y, type = "l", lwd = 1,
           xlab = "p", ylab = "P(batch is accepted)")
```

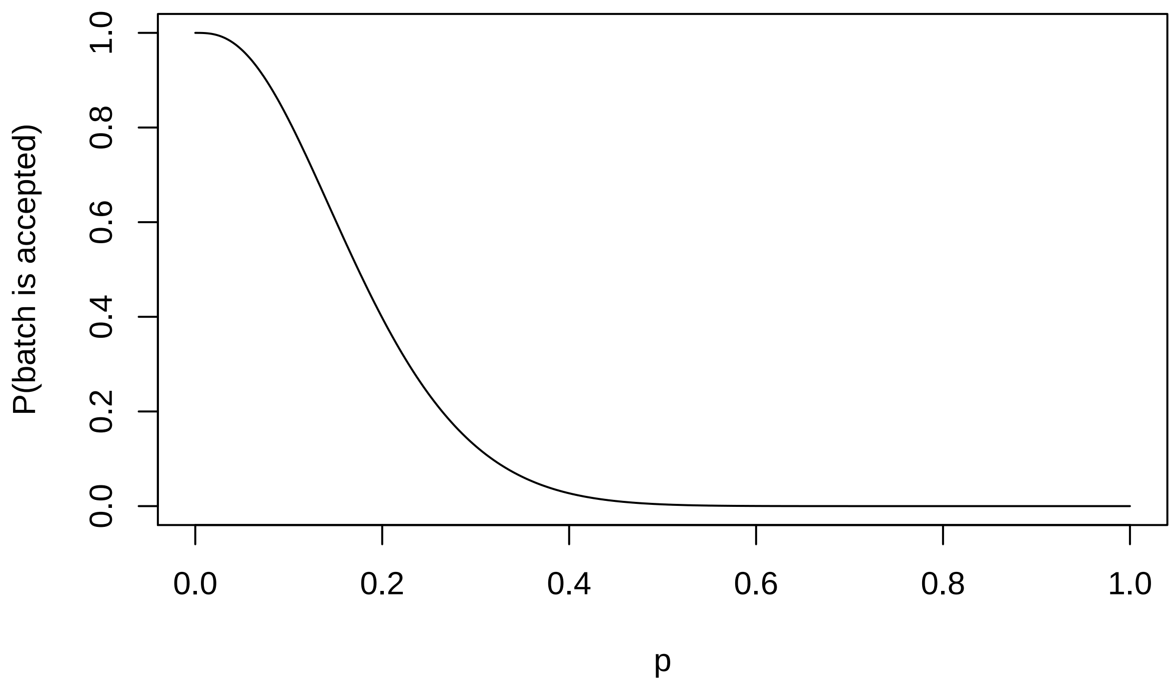


Case 3:

Finally, suppose Yulia decides to randomly select 15 bags of coffee, and will accept the batch only if the number of underweight bags in the sample is at most 2. \

Let p denote the actual proportion underweight bags in the batch. Calculate $P(\text{batch is accepted})$ as a function of p, call this function 'at_most2_15bags' and plot it (domain is (0,1)).

```
at_most2_15bags <- function(p) yulia(15, 2, p)
x <- 0:1000 / 1000
y <- at_most2_15bags(x)
plot(x, y, type = "l", lwd = 1,
           xlab = "p", ylab = "P(batch is accepted)")
```


P(batch is accepted)
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
p