

# Лабораторная работа 1

Могильников Дмитрий

2022-11-10

Для каждого задания:

1. Вычислите для каждого набора основные описательные статистики: среднее, медиану, моду, дисперсию, стандартное отклонение.
2. Постройте для каждого набора гистограммы с двумя разными интервалами разбиения, постройте графики функции плотности.

## Задание 1

Смоделируйте три набора случайных данных объема 200, с нормальным распределением, параметры  $\mu$  и  $\sigma$  задайте самостоятельно

```
#Создадим функцию для вычисления моды распределения:
```

```
getmode <- function(v) {  
  uniqv <- unique(v)  
  uniqv[which.max(tabulate(match(v, uniqv)))]  
}
```

```
#Задаем нормальное распределение для трех наборов данных объемом 200:
```

```
norm_dist1 <- rnorm(200, 5, 1)  
norm_dist2 <- rnorm(200, 10, 10)  
norm_dist3 <- rnorm(200, 0, 2)
```

```
#Вычислим среднее для первого распределения:
```

```
mean(norm_dist1)
```

```
## [1] 4.93298
```

```
#Вычислим медиану для первого распределения:
```

```
median(norm_dist1)
```

```
## [1] 4.977806
```

```
#Вычислим моду для первого распределения:
```

```
getmode(norm_dist1)
```

```
## [1] 6.057726
```

```
#Вычислим дисперсию для первого распределения:
```

```
var(norm_dist1)
```

```
## [1] 1.009087
```

```
#Вычислим стандартное отклонение для первого распределения:
```

```
sd(norm_dist1)
```

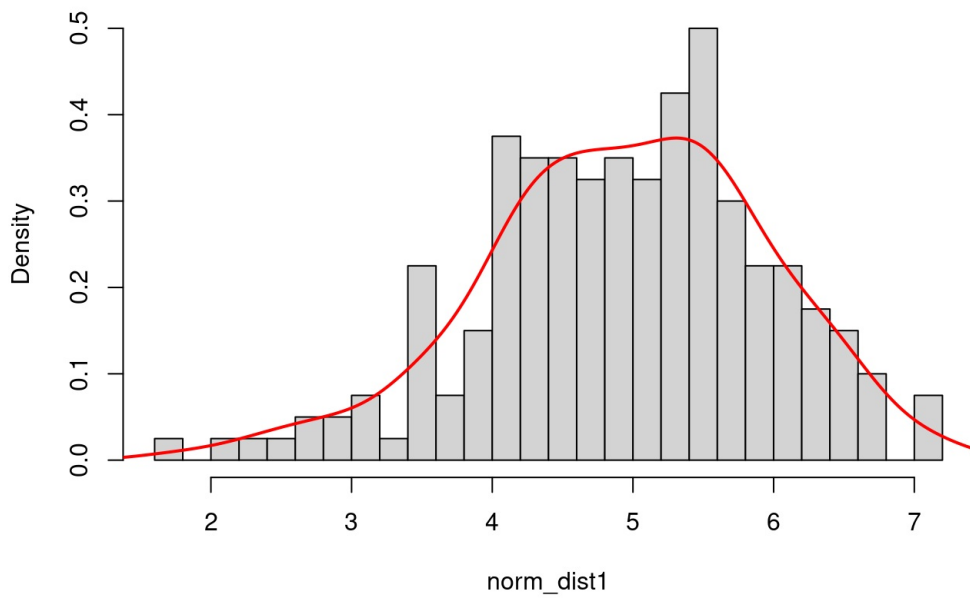
```
## [1] 1.004533
```

```
#Построим гистограммы с двумя разными интервалами разбиения и графики функции плотности:
```

```
#Строим гистограмму для первого распределения с разбиениями на 20 и 10 столбцов:
```

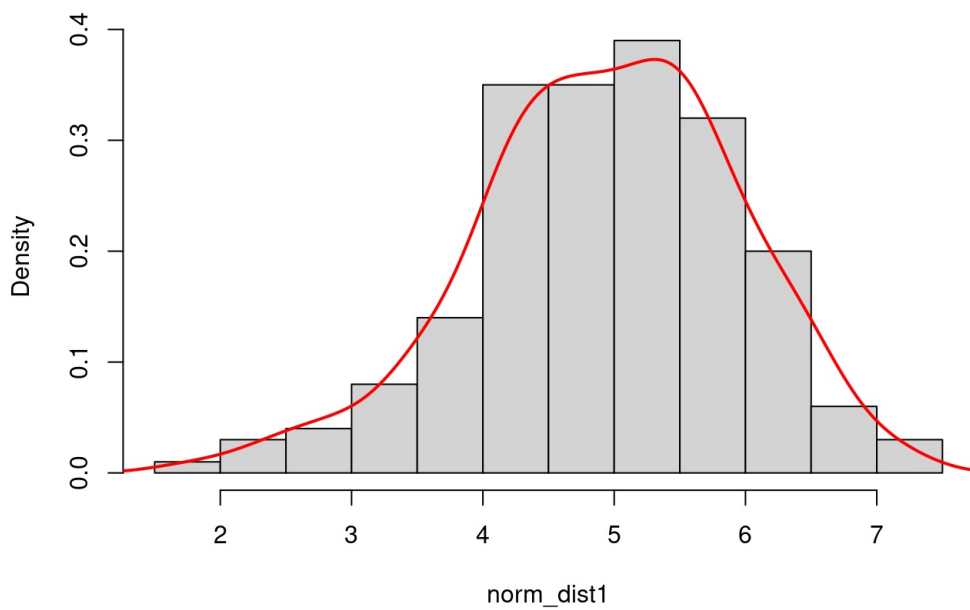
```
hist(norm_dist1, breaks = 20, freq = 0)  
lines(density(norm_dist1), lwd = 2, col = 'Red')
```

Histogram of norm\_dist1



```
hist(norm_dist1, breaks = 10, freq = 0)
lines(density(norm_dist1), lwd = 2, col = 'Red')
```

Histogram of norm\_dist1



```
#Для получения параметров для второго и третьего распределения воспользуемся функцией summary:
#Второе распределение
summary(norm_dist2)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -16.640   4.291  11.304   11.352  17.307   35.848
```

```
var(norm_dist2)
```

```
## [1] 91.76033
```

```
sd(norm_dist2)
```

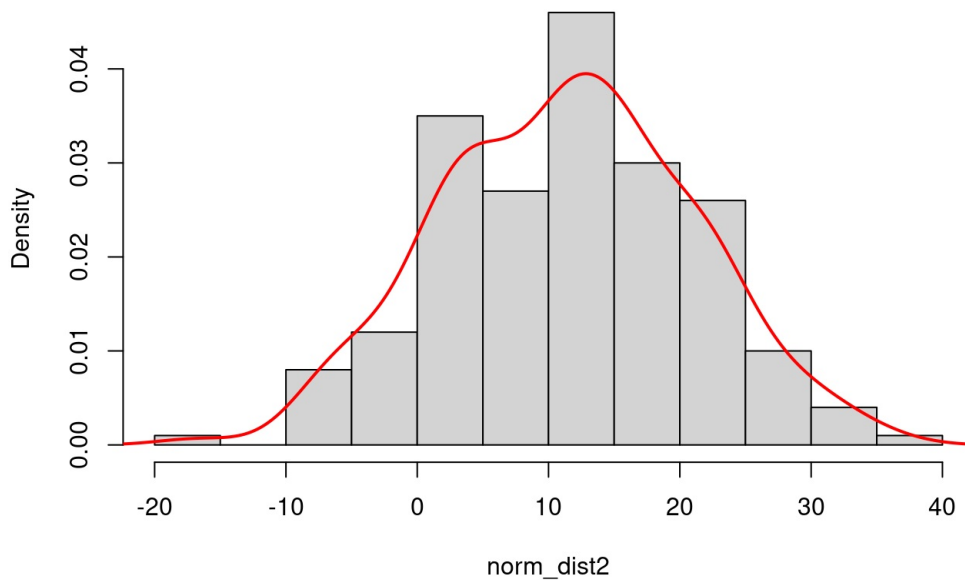
```
## [1] 9.579161
```

```
getmode(norm_dist2)
```

```
## [1] 3.241551
```

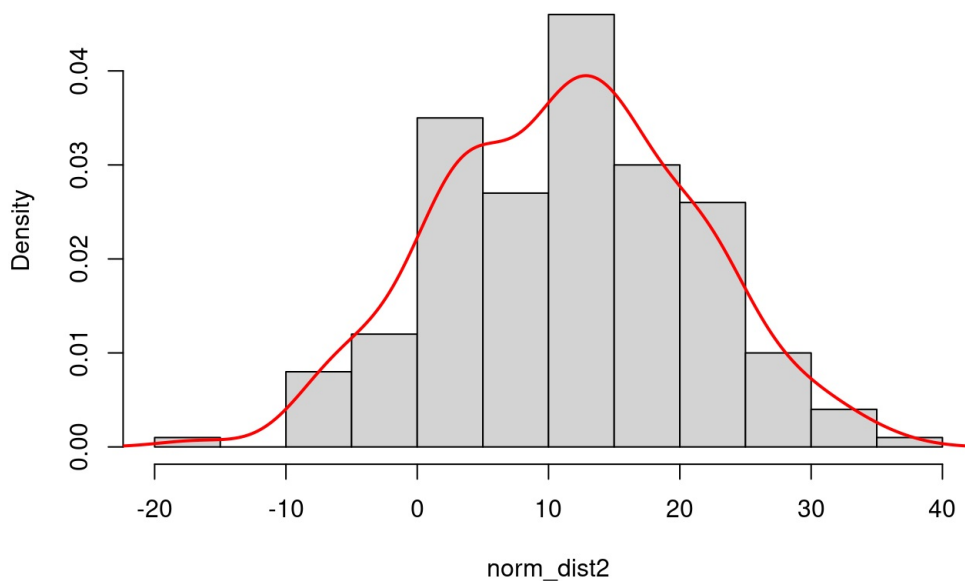
```
#Гистограммы с двумя разными интервалами разбиения (15, 8) и графики функции плотности для второго нормального распределения  
hist(norm_dist2, breaks = 15, freq = 0)  
lines(density(norm_dist2), lwd = 2, col = 'Red')
```

**Histogram of norm\_dist2**



```
hist(norm_dist2, breaks = 8, freq = 0)  
lines(density(norm_dist2), lwd = 2, col = 'Red')
```

**Histogram of norm\_dist2**



```
#Третье распределение  
summary(norm_dist3)
```

```
##      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.     
## -5.98121 -1.39608  0.33873  0.01184  1.43973  4.46375
```

```
var(norm_dist3)
```

```
## [1] 4.12638
```

```
sd(norm_dist3)
```

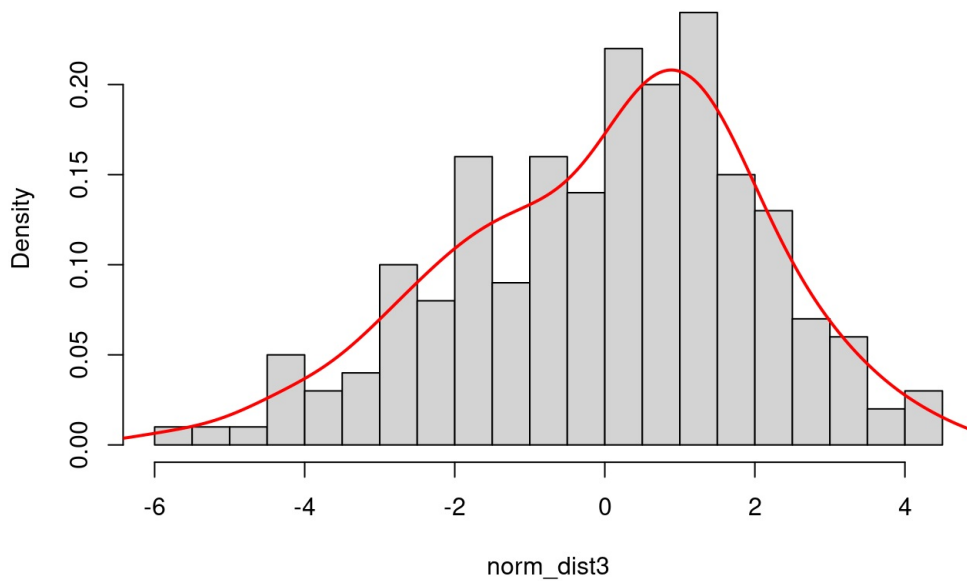
```
## [1] 2.031349
```

```
getmode(norm_dist3)
```

```
## [1] -2.313734
```

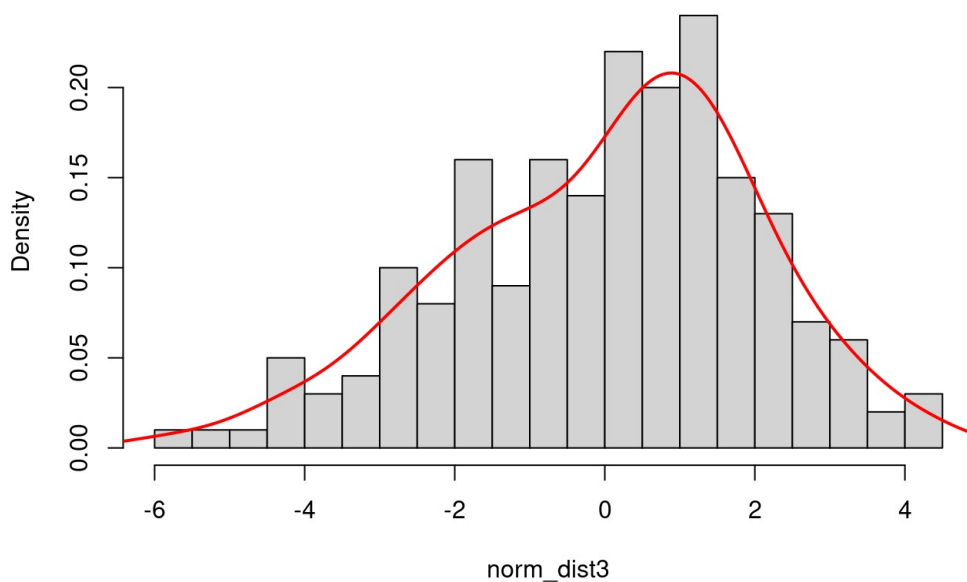
```
#Гистограммы с двумя разными интервалами разбиения(25, 15) и графики функции плотности для третьего нормального р  
аспределения  
hist(norm_dist3, breaks = 25, freq = 0)  
lines(density(norm_dist3), lwd = 2, col = 'Red')
```

**Histogram of norm\_dist3**



```
hist(norm_dist3, breaks = 15, freq = 0)  
lines(density(norm_dist3), lwd = 2, col = 'Red')
```

**Histogram of norm\_dist3**



## Задание 2

Смоделируйте три набора случайных данных объема 200 с распределением Пуассона, параметр lambda задайте самостоятельно

```
#задаем распределение Пуассона для трех наборов данных объемом 200:
```

```
pois_dist1 <- rpois(200, 1)
pois_dist2 <- rpois(200, 4)
pois_dist3 <- rpois(200, 10)
```

```
#Получим параметры распределения:
```

```
#Для первого распределения
```

```
summary(pois_dist1)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.00    0.00    1.00    1.01   2.00    5.00
```

```
var(pois_dist1)
```

```
## [1] 0.9647236
```

```
sd(pois_dist1)
```

```
## [1] 0.9822035
```

```
getmode(pois_dist1)
```

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

```
#Для второго распределения
```

```
summary(pois_dist2)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.000   3.000   4.000   4.055   5.000   11.000
```

```
var(pois_dist2)
```

```
## [1] 3.881382
```

```
sd(pois_dist2)
```

```
## [1] 1.970122
```

```
getmode(pois_dist2)
```

```
## [1] 3
```

```
#Для третьего распределения
```

```
summary(pois_dist3)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      4.000   7.000  10.000   9.975  12.000  19.000
```

```
var(pois_dist3)
```

```
## [1] 11.64259
```

```
sd(pois_dist3)
```

```
## [1] 3.412124
```

```
getmode(pois_dist3)
```

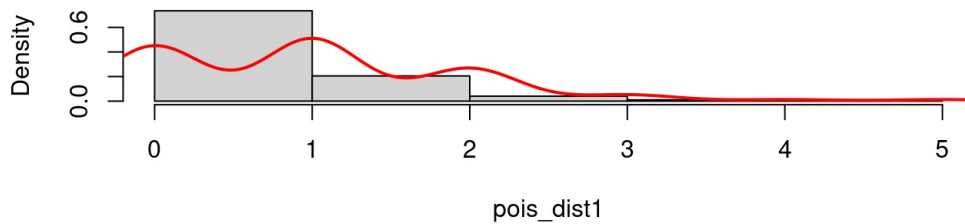
```
## [1] 10
```

*#Для отображения двух гистограм на одном графике используем:*  
`par(mfrow = c(2, 1))`

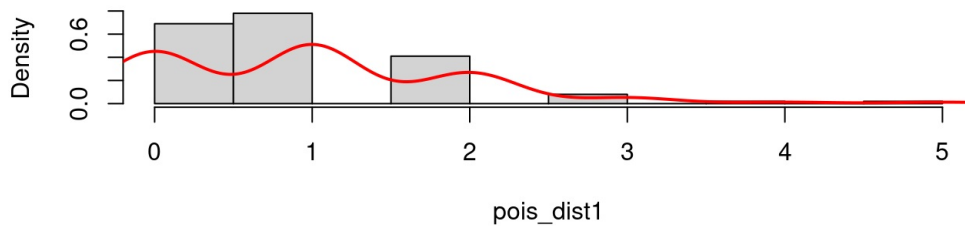
*#Строим гистограммы для распределений с разбиениями для первого (5, 10) для второго (6, 10) и третьего (8, 15) столбцов:*

```
hist(pois_dist1, breaks = 5, freq = 0)  
lines(density(pois_dist1), lwd = 2, col = 'Red')  
hist(pois_dist1, breaks = 10, freq = 0)  
lines(density(pois_dist1), lwd = 2, col = 'Red')
```

**Histogram of pois\_dist1**

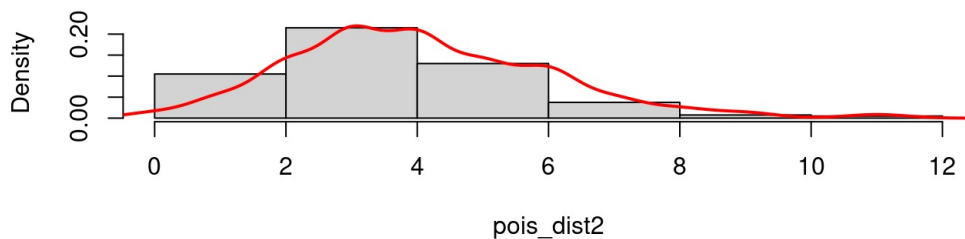


**Histogram of pois\_dist1**

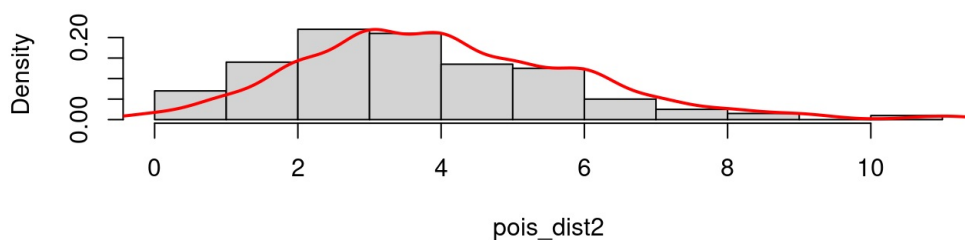


```
hist(pois_dist2, breaks = 6, freq = 0)  
lines(density(pois_dist2), lwd = 2, col = 'Red')  
hist(pois_dist2, breaks = 10, freq = 0)  
lines(density(pois_dist2), lwd = 2, col = 'Red')
```

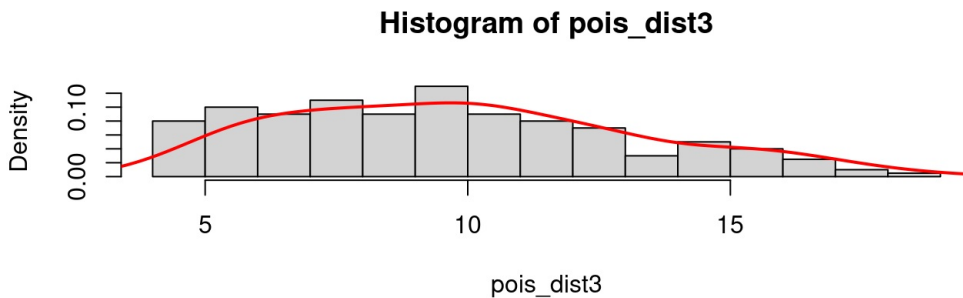
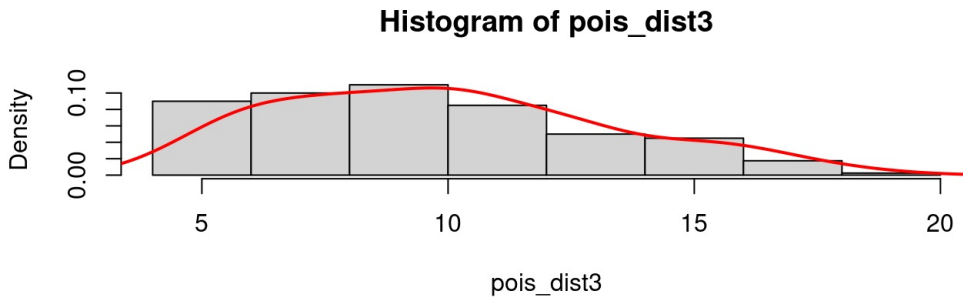
**Histogram of pois\_dist2**



**Histogram of pois\_dist2**



```
hist(pois_dist3, breaks = 8, freq = 0)
lines(density(pois_dist3), lwd = 2, col = 'Red')
hist(pois_dist3, breaks = 15, freq = 0)
lines(density(pois_dist3), lwd = 2, col = 'Red')
```



## Задание 3

Смоделируйте три набора случайных данных объема 200 с биномиальным распределением, параметры size, prob задайте самостоятельно

```
#задаем биномиальное распределение для трех наборов данных объемом 200:
binom_dist1 <- rbinom(200, 20, 0.5)
binom_dist2 <- rbinom(200, 20, 0.75)
binom_dist3 <- rbinom(200, 40, 0.5)
```

```
#Получим параметры распределения:
#Для первого:
summary(binom_dist1)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      4.00   8.00   10.00   9.81  11.00   16.00
```

```
var(binom_dist1)
```

```
## [1] 5.772764
```

```
sd(binom_dist1)
```

```
## [1] 2.402658
```

```
getmode(binom_dist1)
```

```
## [1] 9
```

```
#Для второго:
summary(binom_dist2)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      9.00  14.00  15.00  14.89  16.00   19.00
```

```
var(binom_dist2)
```

```
## [1] 3.61598
```

```
sd(binom_dist2)
```

```
## [1] 1.901573
```

```
getmode(binom_dist2)
```

```
## [1] 16
```

```
#Для третьего:  
summary(binom_dist3)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     
##      11.0   17.0   20.0   19.7   22.0   28.0
```

```
var(binom_dist3)
```

```
## [1] 10.86631
```

```
sd(binom_dist3)
```

```
## [1] 3.296408
```

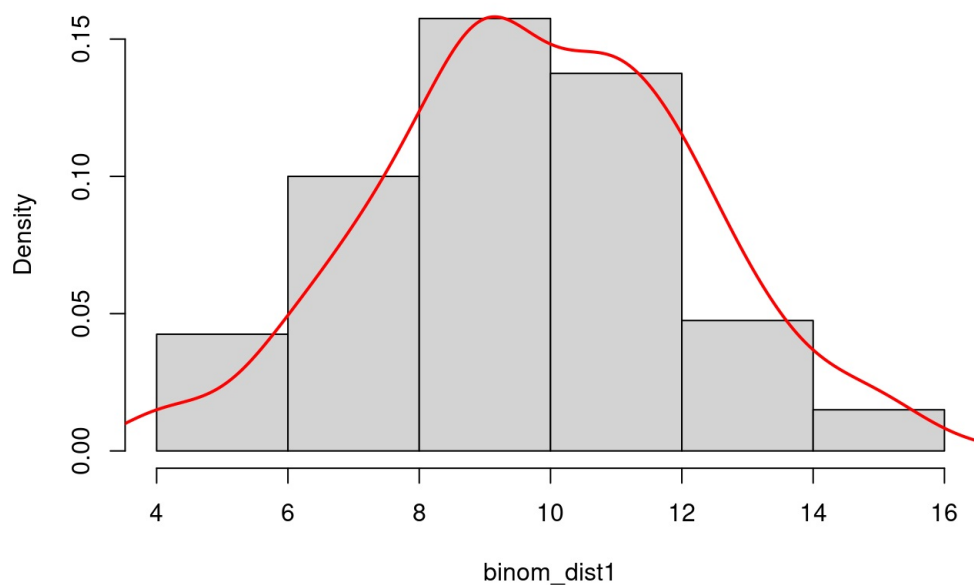
```
getmode(binom_dist3)
```

```
## [1] 22
```

```
#Строим гистограмму для распределений с разбиениями для первого (5, 10) для второго (10, 8) и третьего (8, 15) столбцов:
```

```
hist(binom_dist1, breaks = 5, freq = 0)  
lines(density(binom_dist1), lwd = 2, col = 'Red')
```

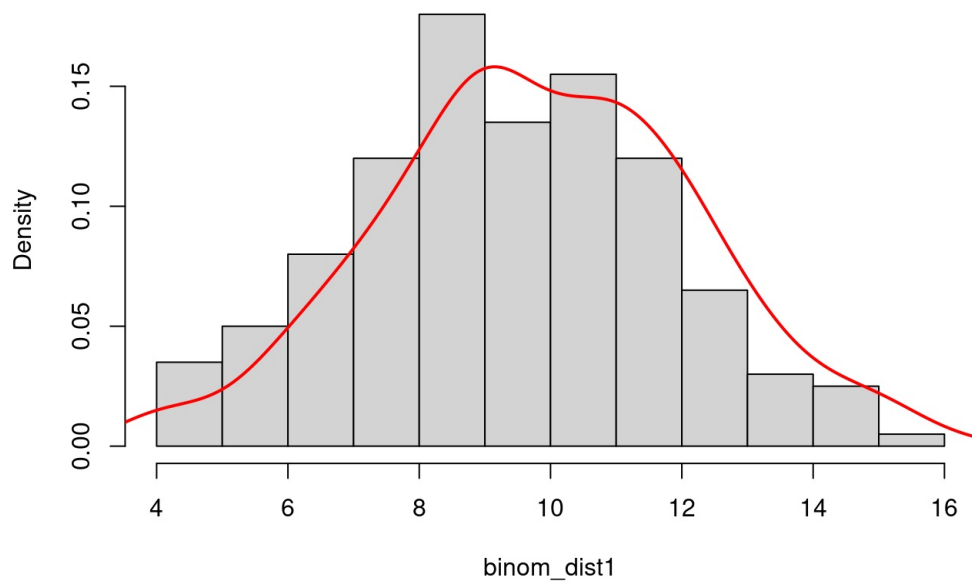
**Histogram of binom\_dist1**



```
hist(binom_dist1, breaks = 10, freq = 0)  
lines(density(binom_dist1), lwd = 2, col = 'Red')
```

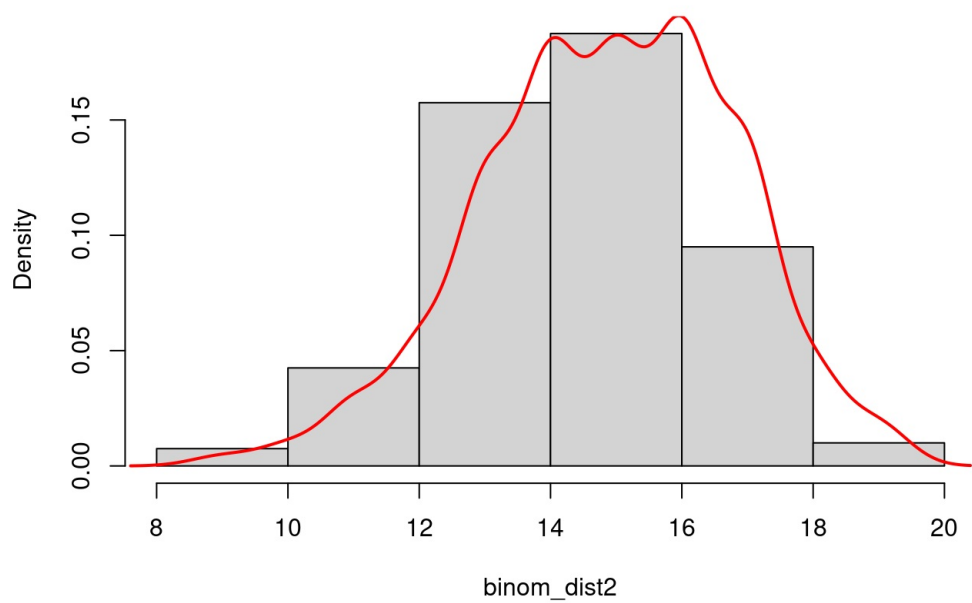


**Histogram of binom\_dist1**



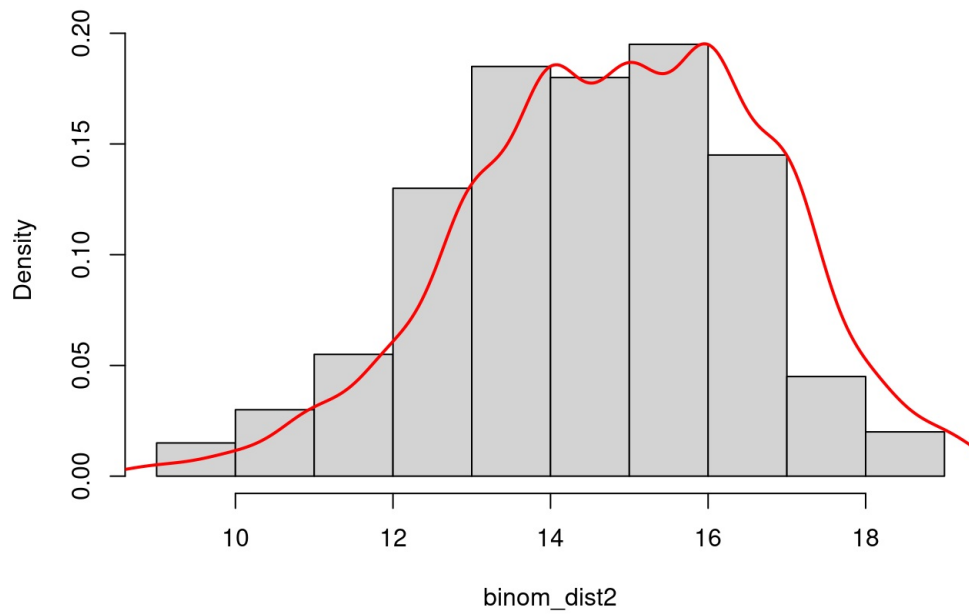
```
hist(binom_dist2, breaks = 6, freq = 0)
lines(density(binom_dist2), lwd = 2, col = 'Red')
```

**Histogram of binom\_dist2**



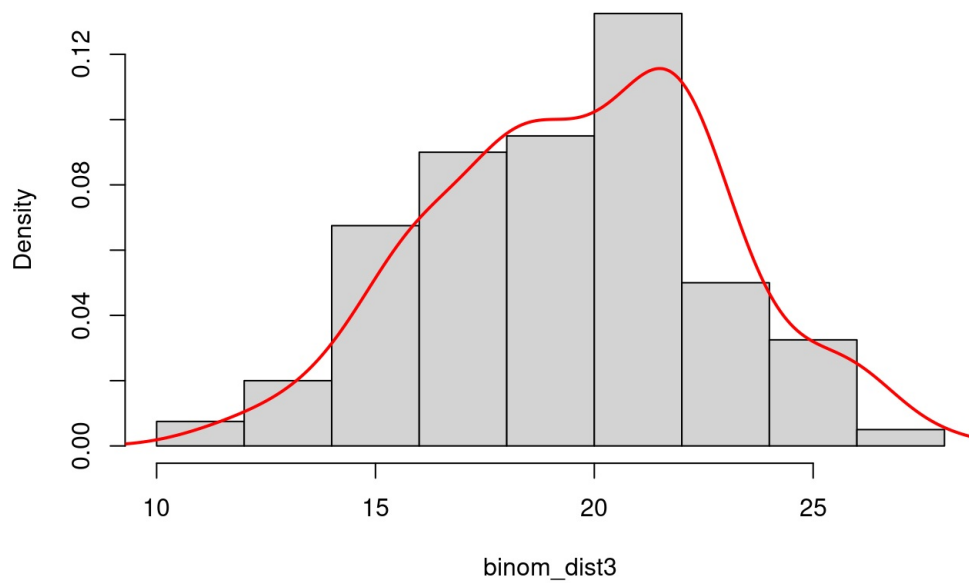
```
hist(binom_dist2, breaks = 10, freq = 0)
lines(density(binom_dist2), lwd = 2, col = 'Red')
```

**Histogram of binom\_dist2**



```
hist(binom_dist3, breaks = 8, freq = 0)
lines(density(binom_dist3), lwd = 2, col = 'Red')
```

**Histogram of binom\_dist3**



```
hist(binom_dist3, breaks = 15, freq = 0)
lines(density(binom_dist3), lwd = 2, col = 'Red')
```

Histogram of binom\_dist3

