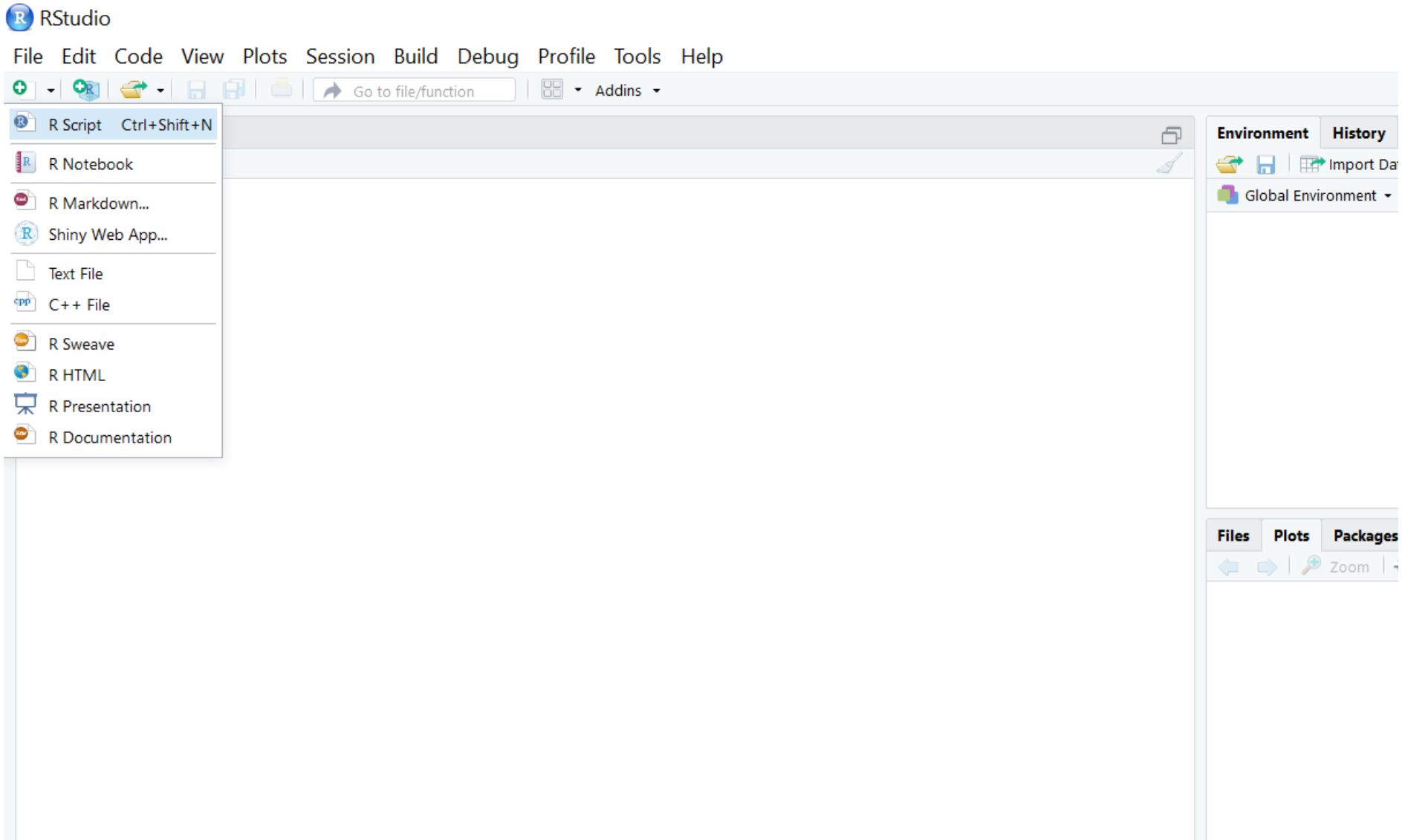
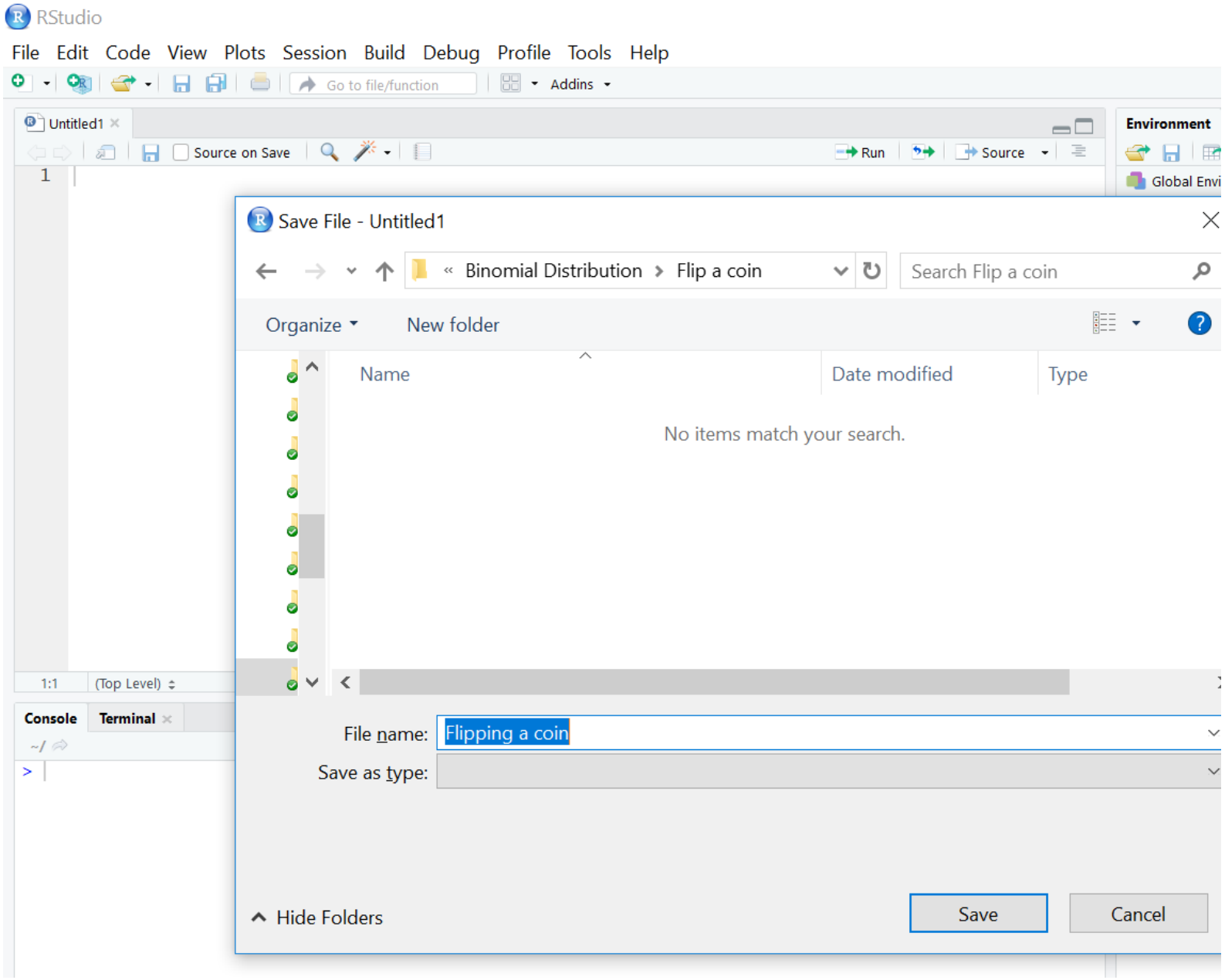
Binomial Distribution

- Open Rstudio. Click on the plus button under File on the right upper side. Then click on “R Script” to create a new R script.



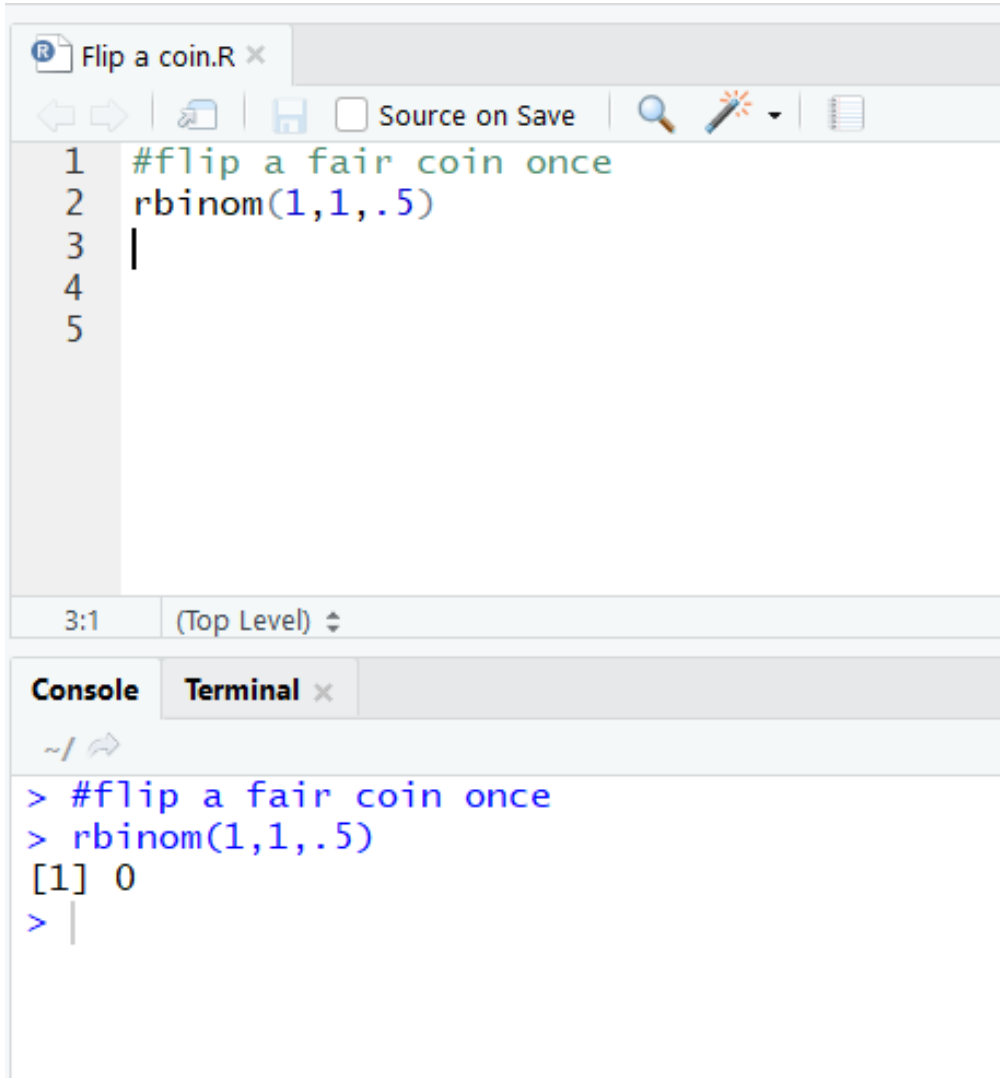
- Press ctrl+ S to save the file and rename it.



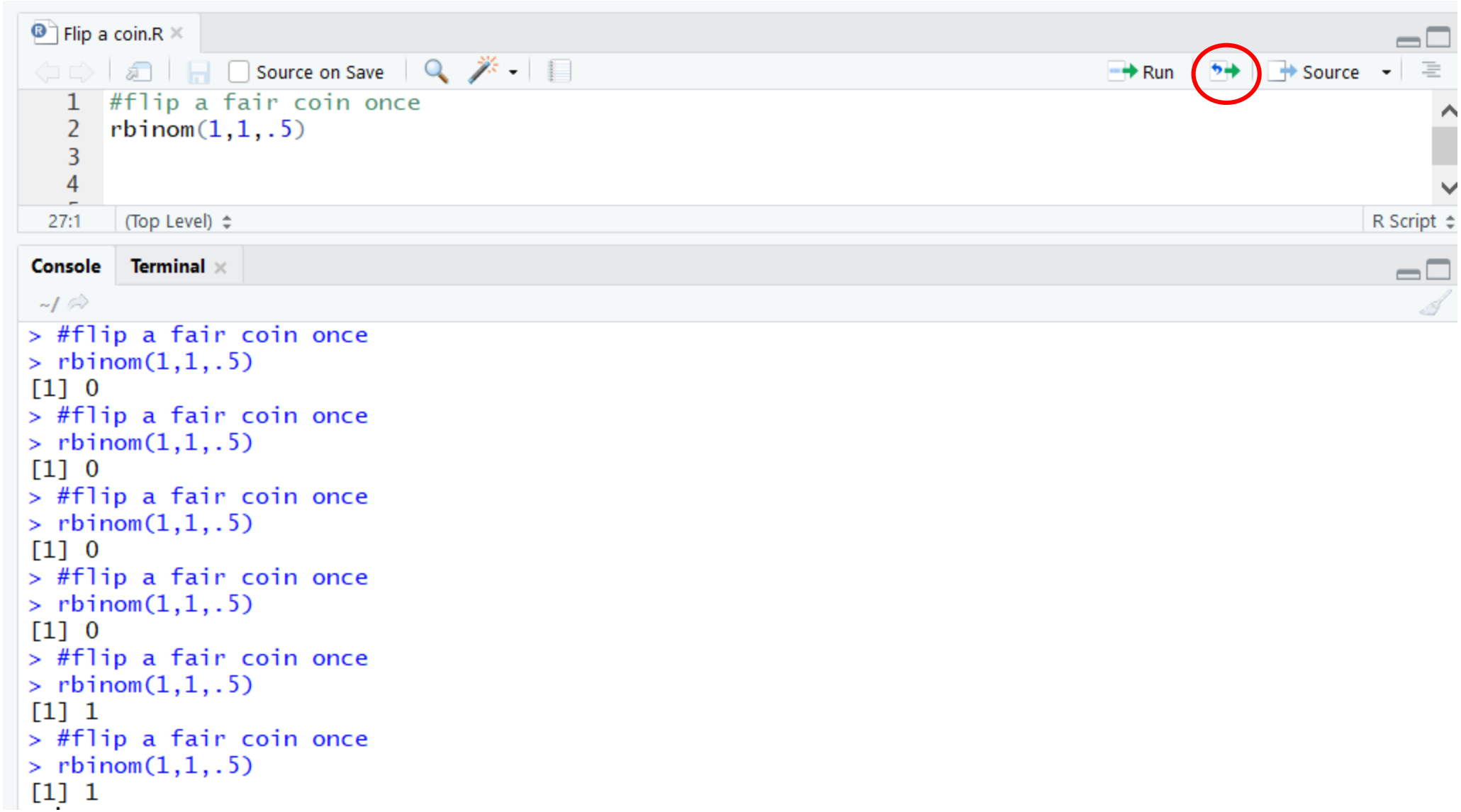
Flipping a coin in R

- Flipping a fair coin once, P(Head)=P(Tail)= 0.5
- Function: rbinom(1,1,.5)
- We expect the result to be 1 or 0.
- 0 represents tail and 1 represents Head.
- In R type rbinom(1,1,.5) and click on Run.

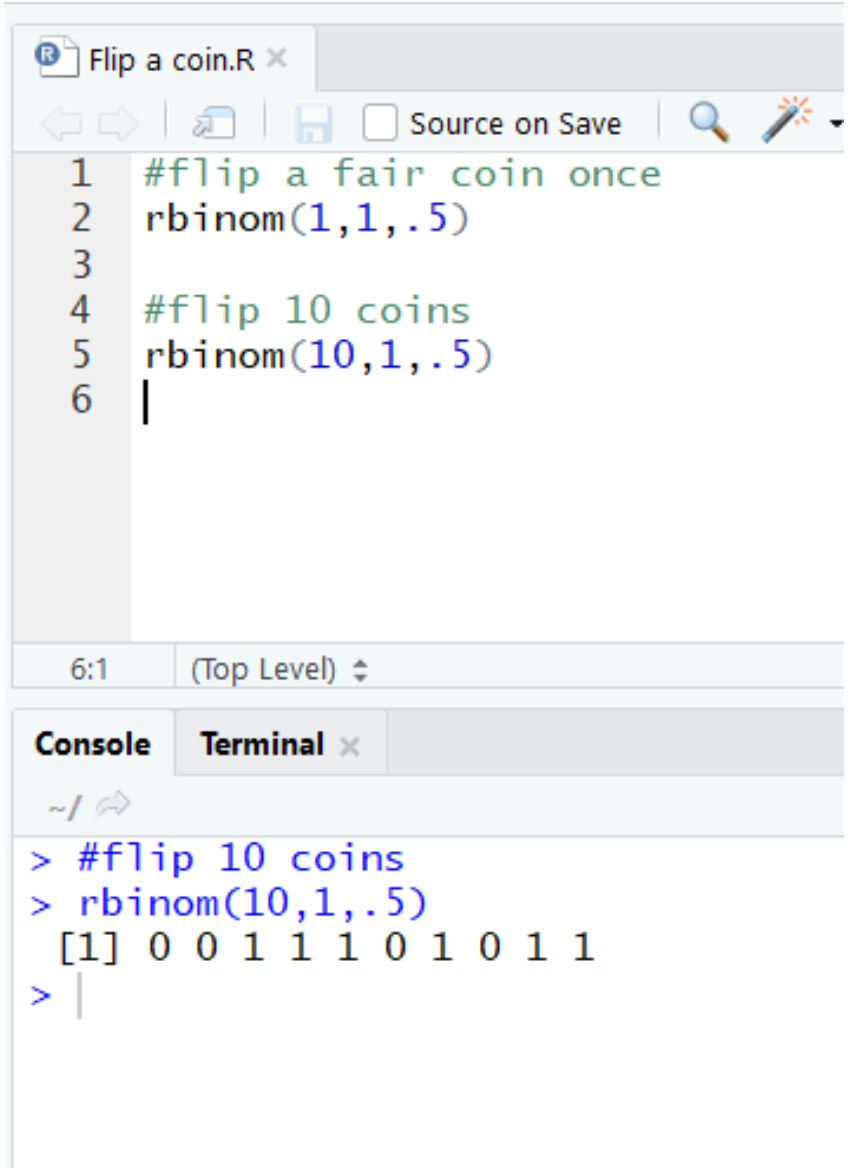
As you can see the result is tail (0).



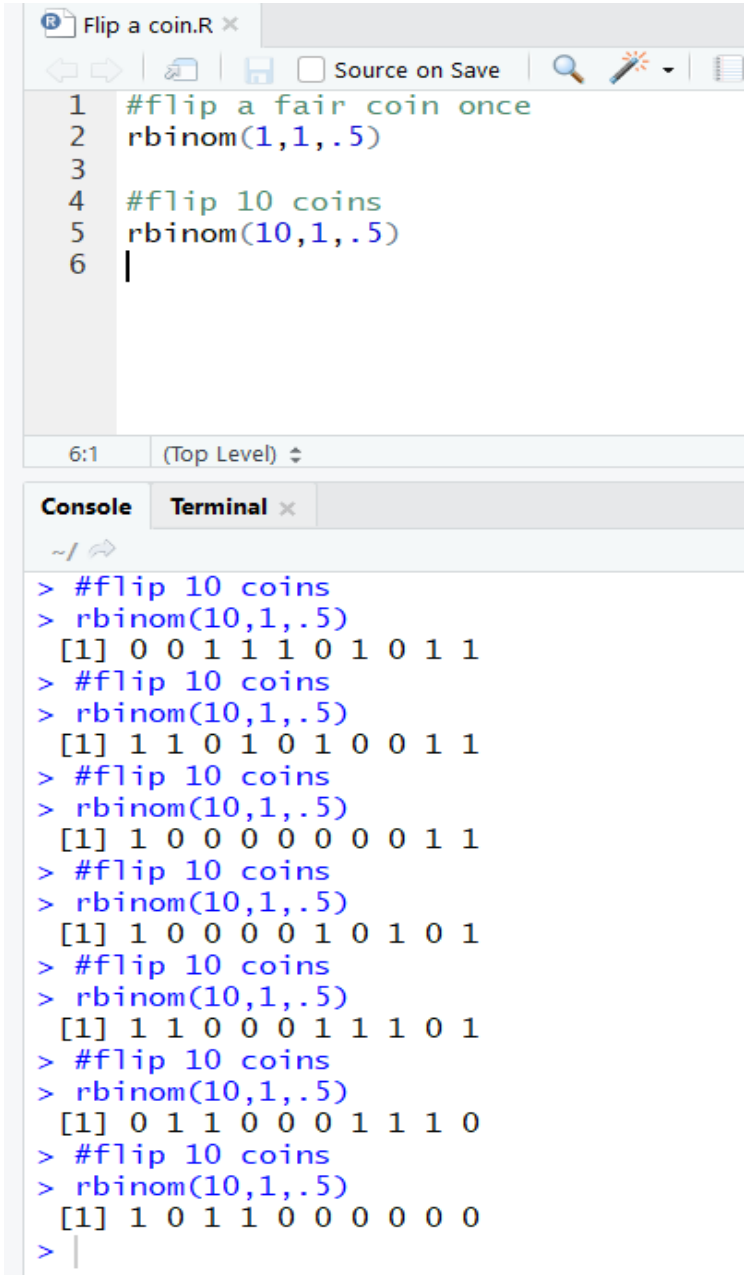
- Click on the button between Run and Source few times to re-run rbinom(1,1,.5) few times.



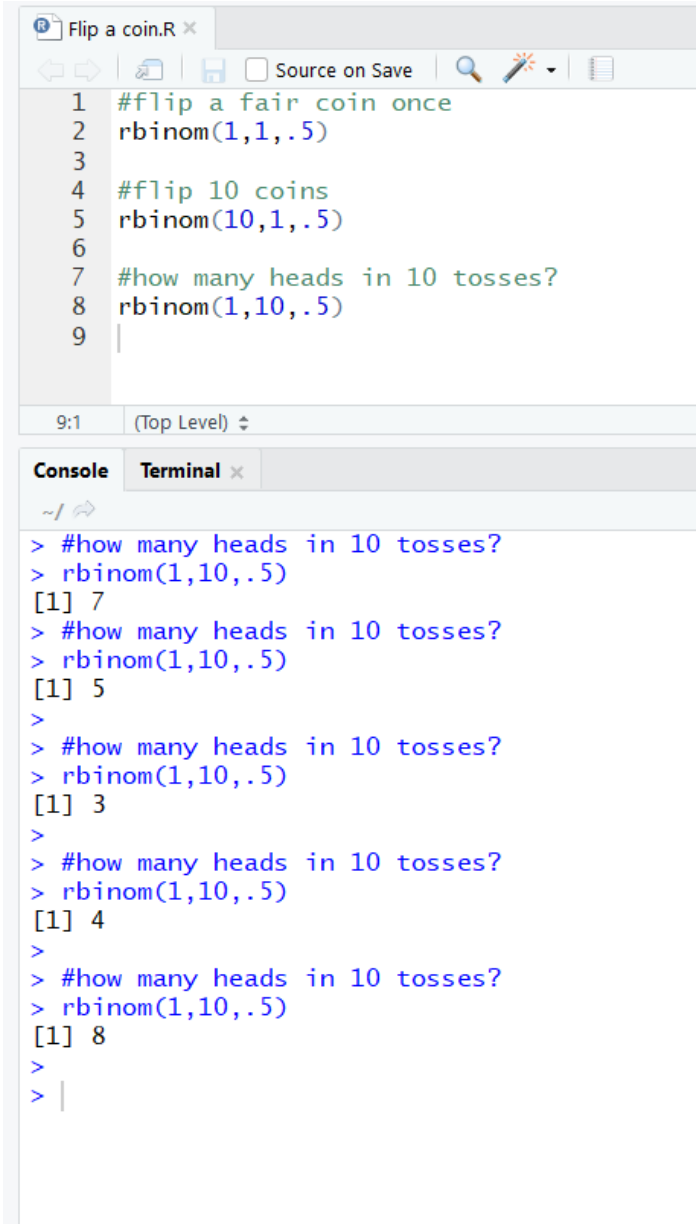
- Now using rbinom(10,1,.5), flip a coin 10 times.



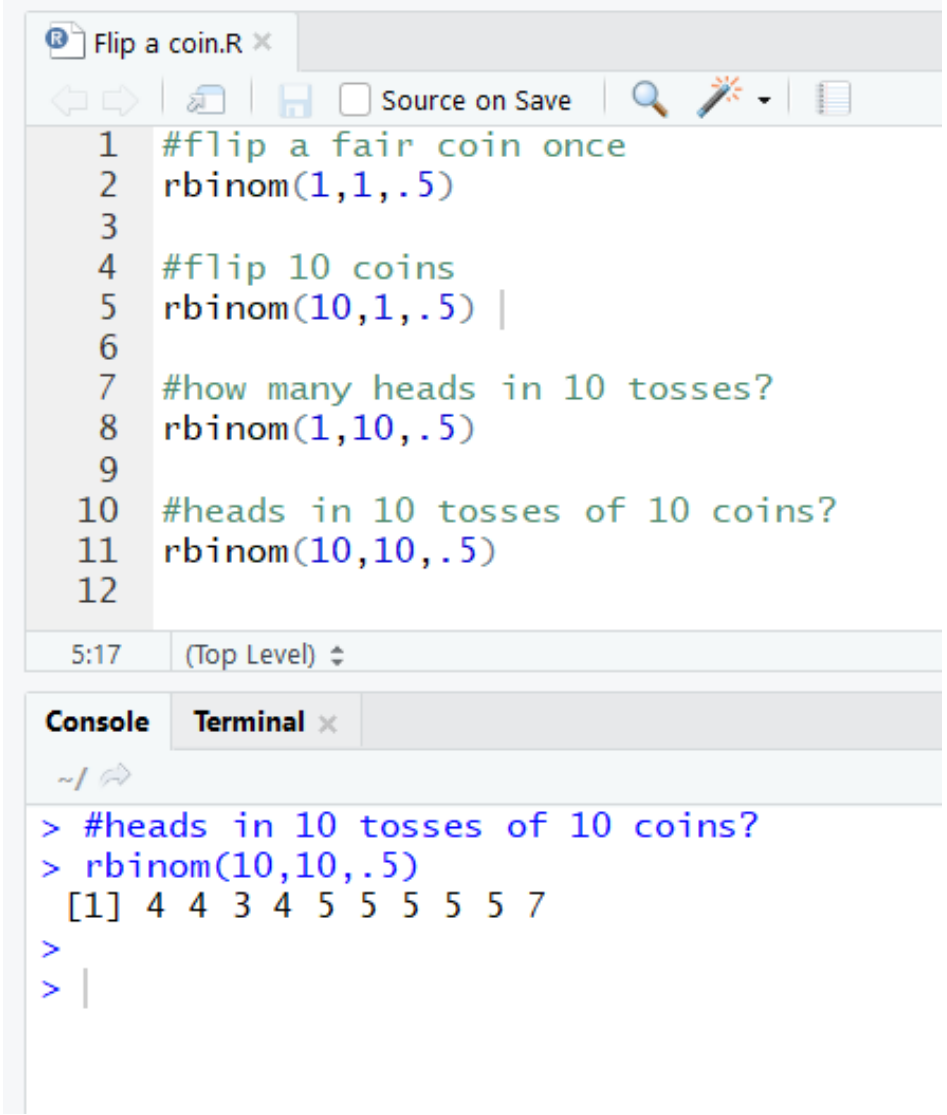
- Re-run the code to see different outcomes.



- rbinom(1,10, 0.5) will return number of heads in 10 tosses.



- rbinom(10,10, 0.5) will toss 10 coins each 10 times and return number of heads out of 10 tosses for each coin.



- Re-run rbinom(10,10, 0.5).

```
10  #heads in 10 tosses of 10 coins?
11  rbinom(10,10,.5)
12
```

12:1 (Top Level)

Console Terminal

```
> #heads in 10 tosses of 10 coins?
> rbinom(10,10,.5)
 [1] 4 4 3 4 5 5 5 5 5 7
>
> #heads in 10 tosses of 10 coins?
> rbinom(10,10,.5)
 [1] 4 1 5 5 4 6 6 6 7 7
> #heads in 10 tosses of 10 coins?
> rbinom(10,10,.5)
 [1] 5 7 7 4 8 6 3 5 4 2
> #heads in 10 tosses of 10 coins?
> rbinom(10,10,.5)
 [1] 4 5 5 7 5 3 5 6 6 6
> #heads in 10 tosses of 10 coins?
> rbinom(10,10,.5)
 [1] 4 7 6 6 7 5 3 6 7 8
>
```

- rbinom(10,10, 0.8) will return number of heads in 10 tosses of 10 **unfair** coins.
- Each toss is a head with probability 0.8.

```
12
13  #heads in 10 tosses of 10 unfair coins
14  rbinom(10,10,.8)
15
16
```

15:1 (Top Level)

Console Terminal

```
~/
> #heads in 10 tosses of 10 unfair coins
> rbinom(10,10,.8)
[1]  7  9  9  8  9  7 10 10 10  8
> #heads in 10 tosses of 10 unfair coins
> rbinom(10,10,.8)
[1]  9  8  8  8  7  6  5  7 10  9
> #heads in 10 tosses of 10 unfair coins
> rbinom(10,10,.8)
[1] 9 8 9 6 7 7 6 8 7 8
> #heads in 10 tosses of 10 unfair coins
> rbinom(10,10,.8)
[1]  8  5  9  8  9 10  8  7  9  9
> #heads in 10 tosses of 10 unfair coins
> rbinom(10,10,.8)
[1] 7 8 9 7 7 9 8 8 8 6
> #heads in 10 tosses of 10 unfair coins
> rbinom(10,10,.8)
[1]  9  8 10  6  8  7  8  9  7  8
>
```

What is the most common number? Why?

- rbinom(10,10, 0.2) will return number of heads in 10 tosses of 10 **unfair** coins.
- Each toss is a head with probability 0.2.

```
15
16  #heads in 10 tosses of 10 unfair coins
17  rbinom(10,10,.2)
18
19
```

18:1 (Top Level)

onsole Terminal

```
 #heads in 10 tosses of 10 unfair coins
 rbinom(10,10,.2)
[1] 2 2 1 2 3 2 1 1 1 2
 #heads in 10 tosses of 10 unfair coins
 rbinom(10,10,.2)
[1] 3 0 3 2 0 2 2 4 1 3
 #heads in 10 tosses of 10 unfair coins
 rbinom(10,10,.2)
[1] 1 3 4 4 2 0 2 1 0 3
 #heads in 10 tosses of 10 unfair coins
 rbinom(10,10,.2)
[1] 5 3 4 1 3 0 4 2 1 2
 #heads in 10 tosses of 10 unfair coins
 rbinom(10,10,.2)
[1] 2 4 1 2 0 3 1 4 1 2
 #heads in 10 tosses of 10 unfair coins
 rbinom(10,10,.2)
[1] 0 2 4 2 3 0 4 0 0 1
 #heads in 10 tosses of 10 unfair coins
 rbinom(10,10,.2)
[1] 1 2 0 1 2 2 4 0 1 3
```

What is the most common number? Why?

More generally,

The first argument of rbinom() is number of experiments.

The second argument is the number of coins flips.

The third argument is the probability of a 1 (“heads”).

For example, rbinom(100000,10,.5) is flipping 10 fair coins, and repeating the experiment 100,000 times.

Exercise 1

- Generate 100 experiments of flipping 10 coins, each with 30% probability.
- What is the most common number? Why?

- Binomial Distribution has two parameters:

  - $X \sim Binomial\ (size, p)$

  - Size= number of coin flips
  - p= the probability of seeing one head in a coin flip
  - Random variable X denotes number of heads.

- We flip a fair coins 10 times. What is the probability of seeing 5 heads?
  - $X \sim Binomial\ (10, .5)$
  - $\Pr(X = 5)$?

- We flip a fair coins 10 times. What is the probability of seeing 5 heads?

Simulation:

- Repeat this experiment 100,000 times: “number of draws=100,000”
- flips <- rbinom(100000,10,.5)
- flips contains 100000 numbers, each between 0 and 10 (number of heads).
- mean(flips == 5), returns percentage of number “5” among 100000 numbers.

The result is 0.24769.

```
21
22  # 100,000 experiments, in each experiment there are 10 coin flips, the coin is fair
23  flips <- rbinom(100000,10,.5)
24
25  # percentage of 'seeing 5 heads' among 100,000 outcomes
26  mean(flips == 5)
27
28
```

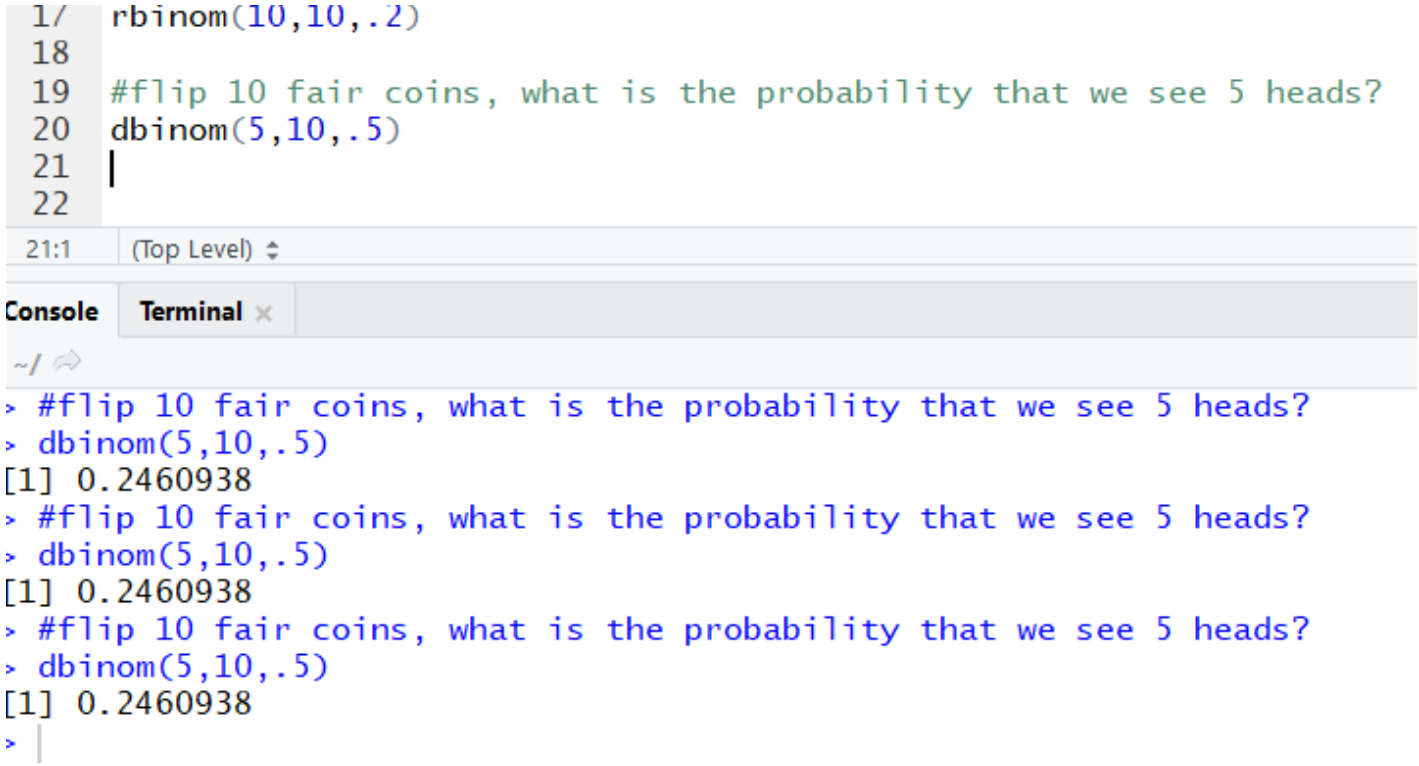
27:1 (Top Level)

Console Terminal

~/

```
> # 100,000 experiments, in each experiment there are 10 coin flips, the coin is fair
> flips <- rbinom(100000,10,.5)
> # percentage of 'seeing 5 heads' among 100,000 outcomes
> mean(flips == 5)
[1] 0.24769
>
```

- dbinom(5,10,.5) returns probability of seeing 5 heads out of 10 tosses, for a fair coin using exact calculation.
- Note that if you re-run it, you will get the same result.
- As you can see, the result of exact calculation is 0.2460938 which is very close to the result of our simulation 0.24769



If $X \sim Binomial\ (10, .5)$, then

dbinom(k,10,.5) returns $\Pr(X = k) = p_X(k)$

```
25  #flip 10 fair coins, what is the probability that we see 5 heads?
26  dbinom(5,10,.5)
27
28  #flip 10 fair coins, what is the probability that we see 6 heads?
29  dbinom(6,10,.5)
30
31  #flip 10 fair coins, what is the probability that we see 4 heads?
32  dbinom(7,10,.5)
33
```

33:1 (Top Level)

Console Terminal

~/

```
> #flip 10 fair coins, what is the probability that we see 5 heads?
> dbinom(5,10,.5)
[1] 0.2460938
> #flip 10 fair coins, what is the probability that we see 6 heads?
> dbinom(6,10,.5)
[1] 0.2050781
> #flip 10 fair coins, what is the probability that we see 4 heads?
> dbinom(7,10,.5)
[1] 0.1171875
>
```

Exercise 2

- If you flip 10 coins each with a 30% probability of coming up heads, what is the probability exactly 2 of them are heads?
- Compare your simulation with the exact calculation.

Exercise 3

- For exercise 2,

- Part a) use 10000 experiments and report the result.
- Part b) use 100000000 experiments and report the result.

- Compare the result of part a and part b, with the exact calculation. What is your conclusion?

If $X \sim Binomial\ (10, .5)$, what is the $\mathrm{E}[X]$? using calculation $\mathrm{E}[X] = 5$.

- Simulation: run the experiment 100,000 times.
- flips <- rbinom (100000, 10, .5)
- mean (flips): the average number of heads

```
33
34  #average number of heads
35  mean(flips <- rbinom (100000, 10, .5 ))
36
37
38
```

36:1 (Top Level)

Console Terminal

```
~/
> #average number of heads
> mean(flips <- rbinom (100000, 10, .5 ))
[1] 4.99987
> #average number of heads
> mean(flips <- rbinom (100000, 10, .5 ))
[1] 4.99843
> #average number of heads
> mean(flips <- rbinom (100000, 10, .5 ))
[1] 5.00363
> #average number of heads
> mean(flips <- rbinom (100000, 10, .5 ))
[1] 4.99523
>
```

Result of simulation is close to 5

If $X \sim Binomial\ (100, .2)$, what is the $\mathrm{E}[X]$? using calculation $\mathrm{E}[X] = 20$.

- Simulation: run the experiment 100,000 times.
- flips <- rbinom (100000, 100, .2)
- mean (flips): the average number of heads

```
36
37  #average number of heads
38  mean(flips <- rbinom (100000, 100, .2 ))
39
40
```

39:1 (Top Level)

Console Terminal

~/

```
> #average number of heads
> mean(flips <- rbinom (100000, 100, .2 ))
[1] 20.02572
> #average number of heads
> mean(flips <- rbinom (100000, 100, .2 ))
[1] 20.01465
> #average number of heads
> mean(flips <- rbinom (100000, 100, .2 ))
[1] 19.99774
> #average number of heads
> mean(flips <- rbinom (100000, 100, .2 ))
[1] 20.01099
> #average number of heads
> mean(flips <- rbinom (100000, 100, .2 ))
[1] 20.00108
>
```

Result of simulation is close to 20

Exercise 4

- What is the expected value of a binomial distribution where 25 coins are flipped, each having a 30% chance of heads?

- Compare your simulation with the exact calculation.

If $X \sim Binomial\ (10, .5)$, what is the $\mathrm{Var}[X]$? using calculation $\mathrm{Var}[X] = 2.5$.

- Simulation: run the experiment 100,000 times.
- X <- rbinom (100000, 10, .5)
- var(X): the variance

```
39
40  var(rbinom (100000, 10, .5 ))
41
42
```

40:1 (Top Level)

Console Terminal

~/

```
> var(rbinom (100000, 10, .5 ))
[1] 2.487262
> var(rbinom (100000, 10, .5 ))
[1] 2.511887
> var(rbinom (100000, 10, .5 ))
[1] 2.490653
> var(rbinom (100000, 10, .5 ))
[1] 2.511053
>
```

Result of simulation is close to 2.5

If $X \sim Binomial\ (100, .2)$, what is the $\mathrm{Var}[X]$? using calculation $\mathrm{Var}[X] = 16$.

- Simulation: run the experiment 100,000 times.
- X <- rbinom (100000, 100, .2)
- var(X): the variance

```
41
42  var(rbinom (100000, 100, .2 ))
43
44
```

43:1 (Top Level)

Console Terminal

~/

```
> var(rbinom (100000, 100, .2 ))
[1] 16.04723
> var(rbinom (100000, 100, .2 ))
[1] 15.99349
> var(rbinom (100000, 100, .2 ))
[1] 16.0699
> var(rbinom (100000, 100, .2 ))
[1] 15.99154
> var(rbinom (100000, 100, .2 ))
[1] 15.96807
> var(rbinom (100000, 100, .2 ))
[1] 16.02688
>
```

Result of simulation is close to 16

Exercise 5

- What is the variance of a binomial distribution where 25 coins are flipped, each having a 30% chance of heads?

- Compare your simulation with the exact calculation.