

Zi Xuan Li

Professor Vahdani

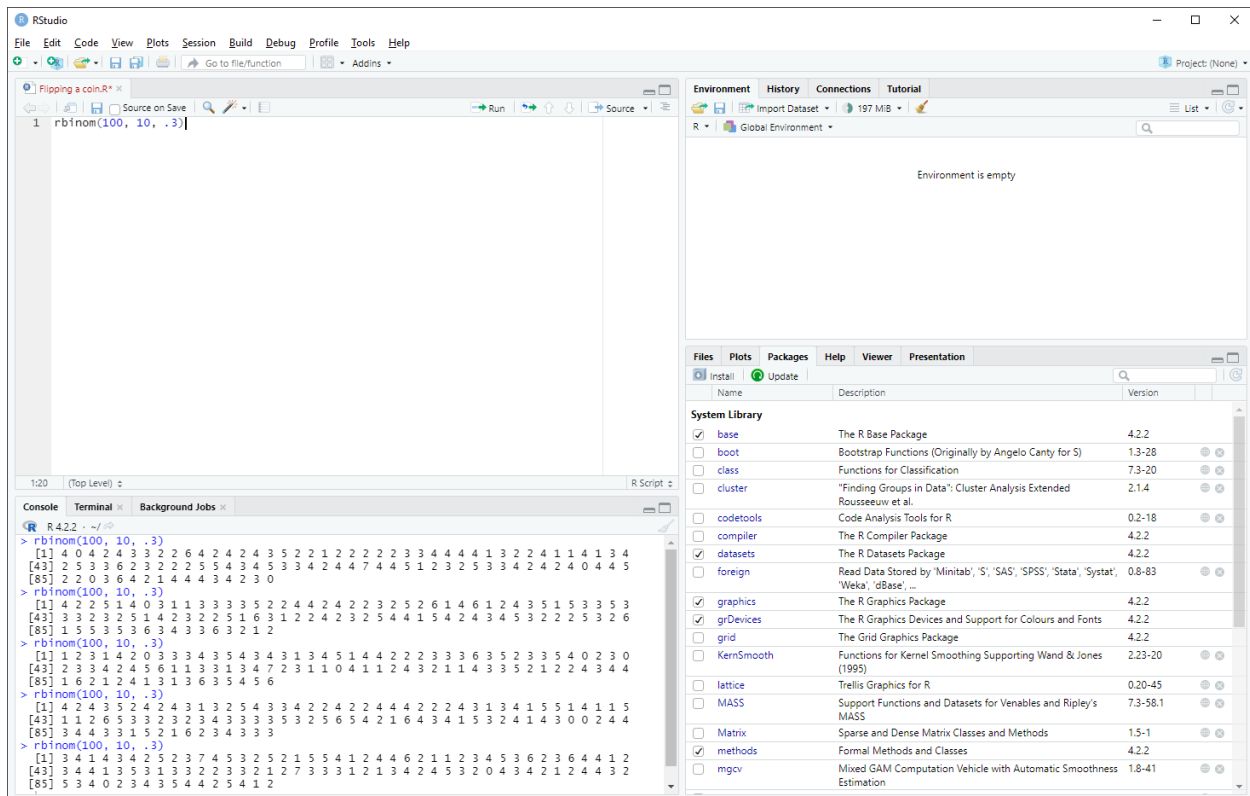
CSC 21700

December 18th, 2022

R Assignment #1

Exercise 1:

Generate 100 experiments of flipping 10 coins, each with 30% probability.



The screenshot shows the RStudio interface. The script editor contains the following code:

```
1 rbinom(100, 10, .3)
```

The console shows the output of the code, which is a 100x10 matrix of random numbers (0s and 1s) generated by the `rbinom` function. The output is displayed in a grid format, with the first 10 rows shown in the console and the rest in the background jobs pane.

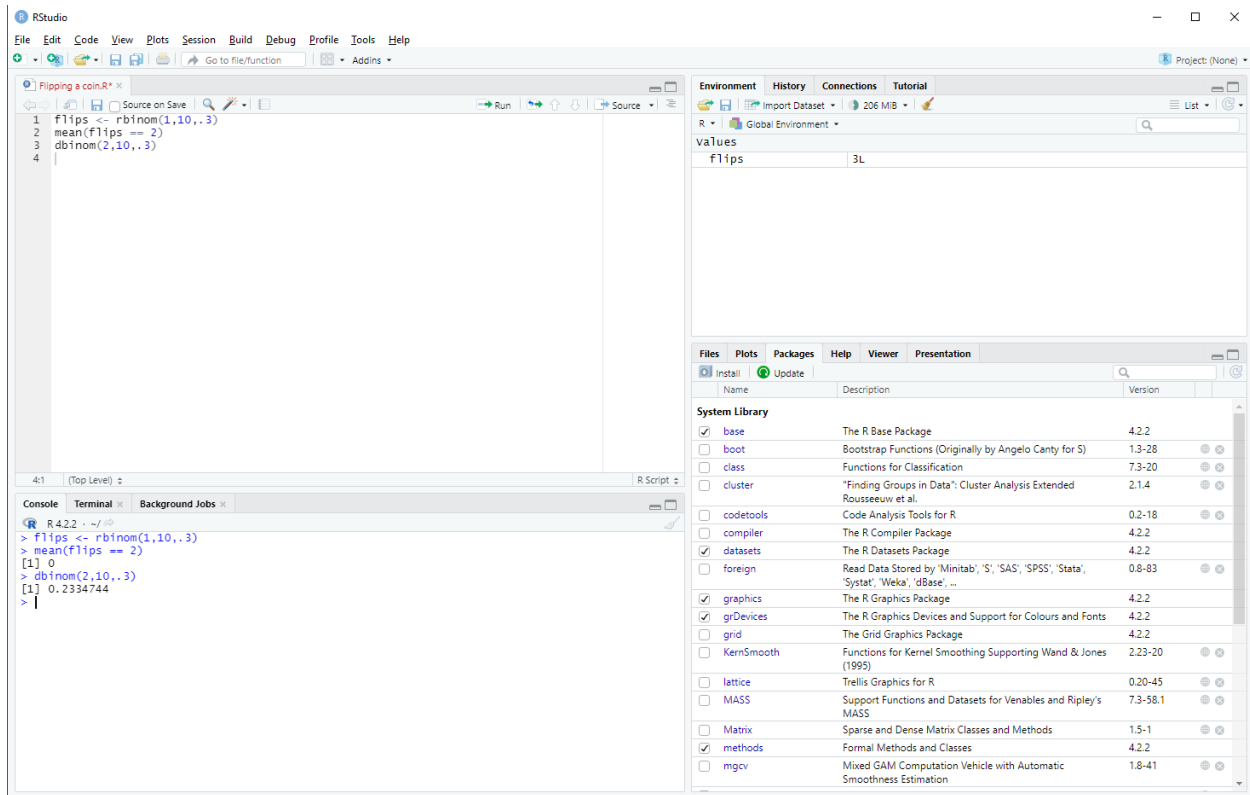
The Environment pane shows the Global Environment, which is empty. The Packages pane shows the installed packages, including the base package and various data science packages.

What is the most common number why?

The most common number is 3 because the program is using unfair coins where each toss has a probability of 0.3 of landing on heads.

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 calculations

The result from the simulation when the number of experiments was set to 1 was 0, while the exact calculation is 0.2334744.

Exercise 3:

Part a) use 10000 experiments and report the result.

The screenshot shows the RStudio interface. The script editor contains the following code:

```
1 flips <- rbinom(10000,10,.3)
2 mean(flips == 2)
3 dbinom(2,10,.3)
4
```

The console output shows the results of the execution:

```
R 4.2.2 ~ ~/R
> mean(flips == 2)
[1] 0.2385
> dbinom(2,10,.3)
[1] 0.2334744
>
```

The Environment pane shows the variable 'flips' as an integer vector of length 10,000.

| Name | Description | Version |
|------------|--|----------|
| base | The R Base Package | 4.2.2 |
| boot | Bootstrap Functions (Originally by Angelo Canty for S) | 1.3-28 |
| class | Functions for Classification | 7.3-20 |
| cluster | "Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al. | 2.1.4 |
| codetools | Code Analysis Tools for R | 0.2-18 |
| compiler | The R Compiler Package | 4.2.2 |
| datasets | The R Datasets Package | 4.2.2 |
| foreign | Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ... | 0.8-83 |
| graphics | The R Graphics Package | 4.2.2 |
| grDevices | The R Graphics Devices and Support for Colours and Fonts | 4.2.2 |
| grid | The Grid Graphics Package | 4.2.2 |
| KernSmooth | Functions for Kernel Smoothing Supporting Wand & Jones (1995) | 2.23-20 |
| lattice | Trellis Graphics for R | 0.20-45 |
| MASS | Support Functions and Datasets for Venables and Ripley's MASS | 7.3-58.1 |
| Matrix | Sparse and Dense Matrix Classes and Methods | 1.5-1 |
| methods | Formal Methods and Classes | 4.2.2 |
| mgcv | Mixed GAM Computation Vehicle with Automatic Smoothness Estimation | 1.8-41 |

Part b) use 100000000 experiments and report the result.

The screenshot shows the RStudio interface. The script editor contains the following code:

```
1 flips <- rbinom(100000000,10,.3)
2 mean(flips == 2)
3 dbinom(2,10,.3)
4
5
```

The console output shows the results of the execution:

```
R 4.2.2 ~ ~/R
> flips <- rbinom(100000000,10,.3)
> mean(flips == 2)
[1] 0.2334797
> dbinom(2,10,.3)
[1] 0.2334744
>
```

The Environment pane shows the variable 'flips' as a large integer vector of length 100,000,000, occupying 400 MB.

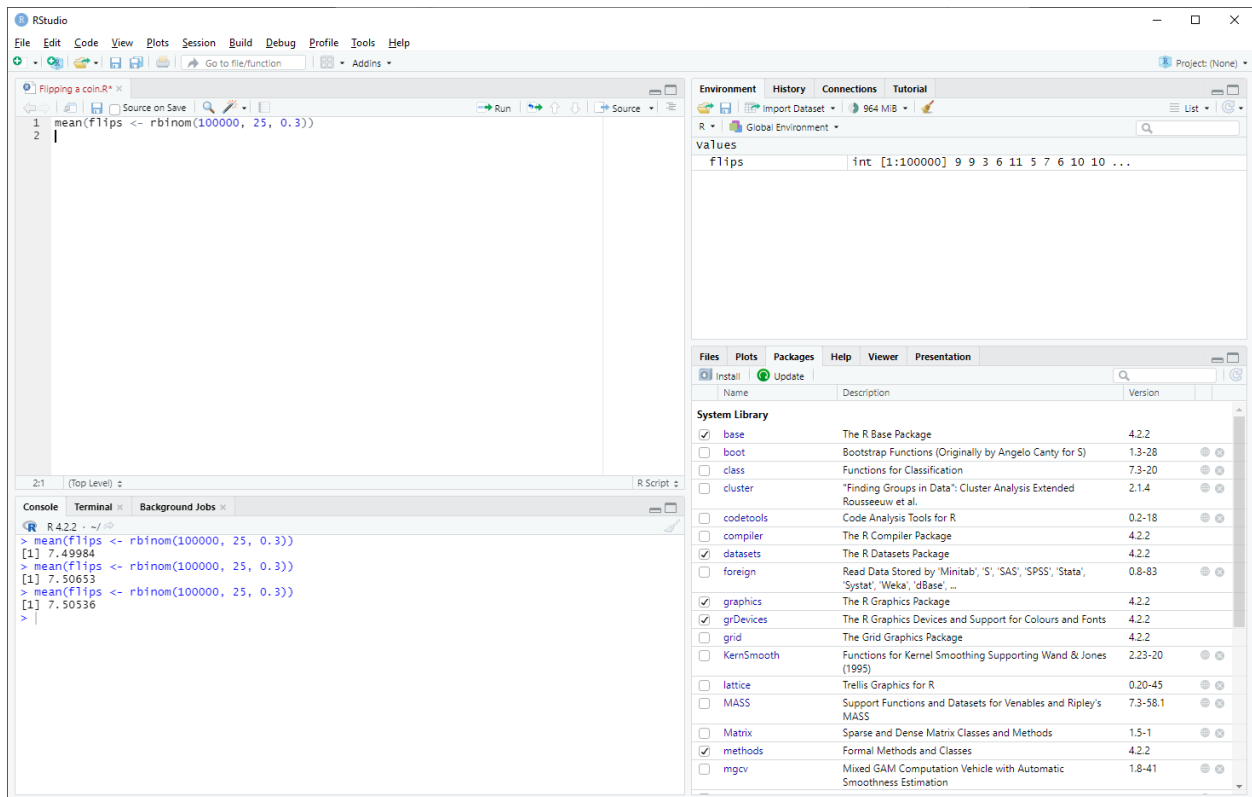
| Name | Description | Version |
|------------|--|----------|
| base | The R Base Package | 4.2.2 |
| boot | Bootstrap Functions (Originally by Angelo Canty for S) | 1.3-28 |
| class | Functions for Classification | 7.3-20 |
| cluster | "Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al. | 2.1.4 |
| codetools | Code Analysis Tools for R | 0.2-18 |
| compiler | The R Compiler Package | 4.2.2 |
| datasets | The R Datasets Package | 4.2.2 |
| foreign | Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ... | 0.8-83 |
| graphics | The R Graphics Package | 4.2.2 |
| grDevices | The R Graphics Devices and Support for Colours and Fonts | 4.2.2 |
| grid | The Grid Graphics Package | 4.2.2 |
| KernSmooth | Functions for Kernel Smoothing Supporting Wand & Jones (1995) | 2.23-20 |
| lattice | Trellis Graphics for R | 0.20-45 |
| MASS | Support Functions and Datasets for Venables and Ripley's MASS | 7.3-58.1 |
| Matrix | Sparse and Dense Matrix Classes and Methods | 1.5-1 |
| methods | Formal Methods and Classes | 4.2.2 |
| mgcv | Mixed GAM Computation Vehicle with Automatic Smoothness Estimation | 1.8-41 |

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

After comparing the results of part a & b with their exact calculations respectively. I concluded that the more experiments that were performed, the more accurate the probability of seeing 2 heads when flipping 10 fair coins became.

Exercise 4:

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



The screenshot shows the RStudio interface. The script editor contains the following code:

```
1 mean(flips <- rbinom(100000, 25, 0.3))
2
```

The console output shows the result of the simulation:

```
R 4.2.2 ~ ./
> mean(flips <- rbinom(100000, 25, 0.3))
[1] 7.49984
> mean(flips <- rbinom(100000, 25, 0.3))
[1] 7.50653
> mean(flips <- rbinom(100000, 25, 0.3))
[1] 7.50536
>
```

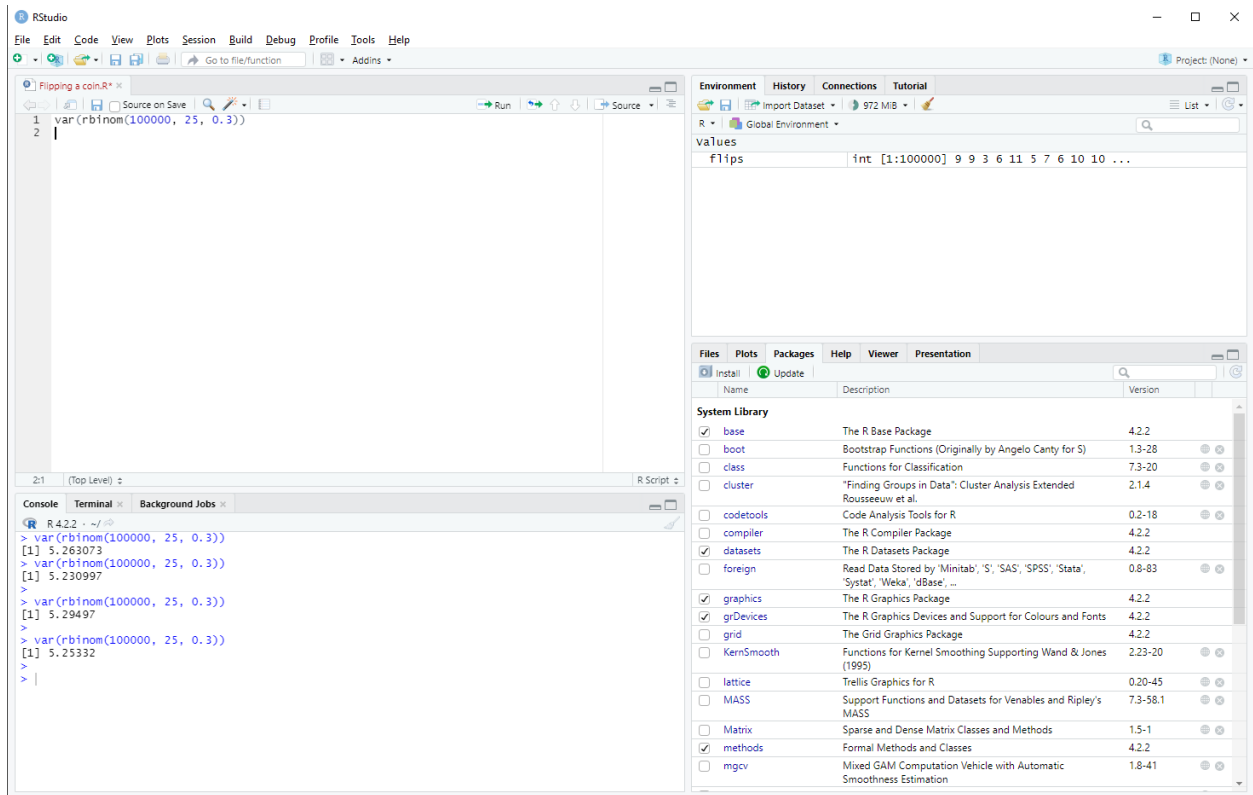
The Environment pane shows a variable named 'flips' of type 'int' with dimensions [1:100000]. The Packages pane shows the installed R packages, including the base library and various data science packages.

Compare your simulation with the exact calculation.

If $X \sim \text{Binomial}(25, 0.3)$, then $E[X]$ is 7.5. The simulation result is 7.49984 and the exact calculation of $E[X]$ is 7.5.

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.

If $X \sim \text{Binomial}(25, 0.3)$, then $\text{Var}[X]$ is 5.25. The simulation result is 5.263073 and the exact calculation of $\text{Var}[X]$ is 5.25.