Random Variables

Raúl López

2024-02-22

Introduction

Key Points

- Random variables are numeric outcomes resulting from random processes.
- Statistical inference offers a framework for quantifying uncertainty due to randomness.

Model a random variable with `sample()` function.

```
set.seed(1)
# define random variable x to be 1 if blue, 0 otherwise
beads <- rep(c("red", "blue"), times = c(2, 3))
x <- ifelse(sample(beads, 1) == "blue", 1, 0)
print(x)
```

```
## [1] 0
```

Sampling Models

Key Points

- A sampling model models the random behavior of a process as the sampling of draws from an urn.
- The probability distribution of a random variable is the probability of the observed value falling in any given interval.
- We can define a CDF F(a)=Pr(S<=a) to answer questions related to the probability of S being in any interval.
- The average of many draws of a random variable is called its expected value.
- The standard deviation of many draws of a random variable is called its standard error.

We use sampling model to run a Monte Carlo simulation and use the results to estimate the probability of an event. In this example, we have a casino with 1000 players of roulette each day. We have 38 pockets (18 red, 18 black and 2 green). The player bet 1$ and the events are:

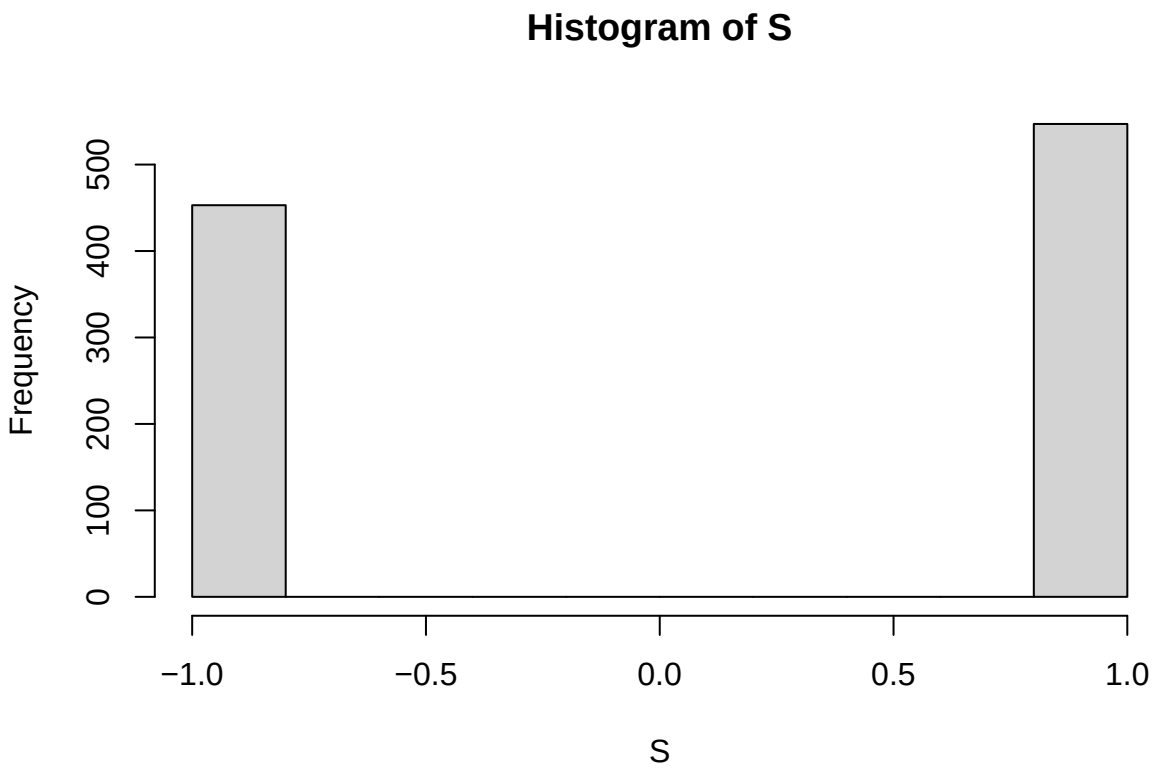
- red: the casino losses 1$
- black and green: the casino wins 1$

The model of this random variable is:

```
red <- 18
black <- 18
green <- 2
n <- 1000

S <- sample(c(1,-1), # options are +1 and -1
       n, # players are 1000
       replace = TRUE,
```

```
       prob = c(20/38, # probability of win
                18/38)) # probability of lose
hist(S) # distribution
```




```
sum(S) # total gains
```

```
## [1] 94
```

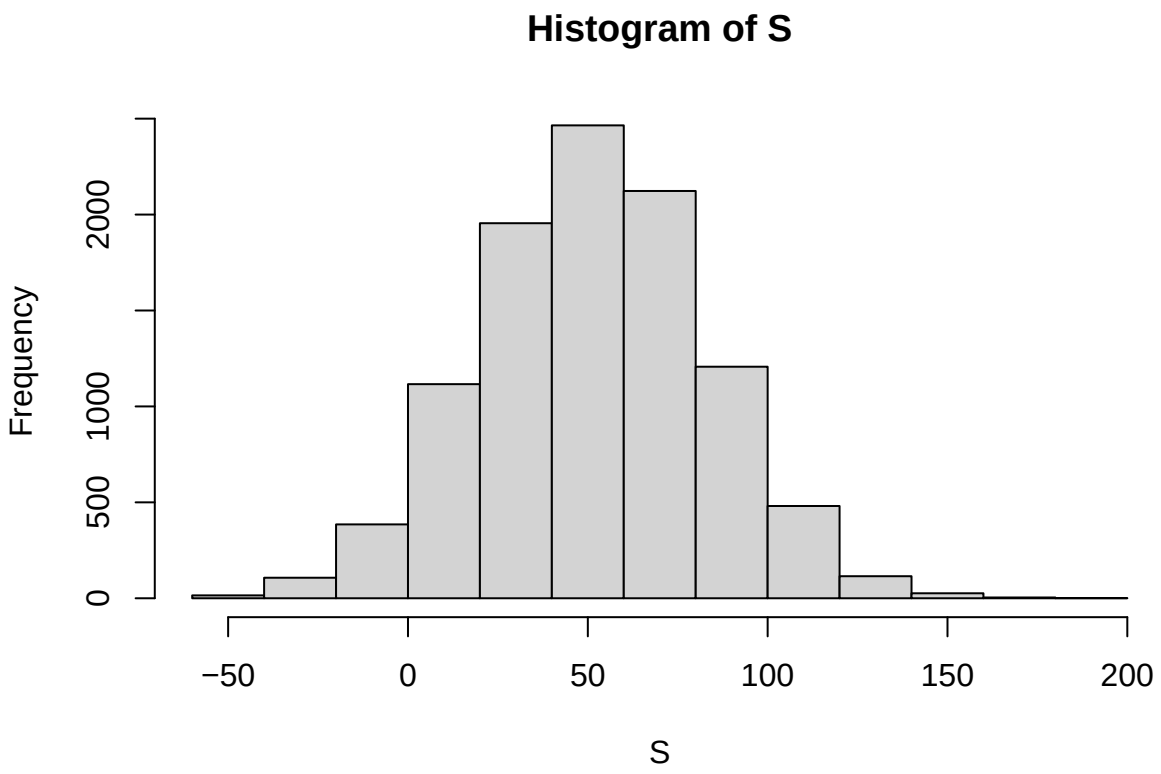
```
mean(S) # mean gain
```

```
## [1] 0.094
```

We use the sampling model to run a Monte Carlo simulation

```
n <- 1000     # number of roulette players
B <- 10000     # number of Monte Carlo experiments
S <- replicate(B, {
    X <- sample(c(-1,1), n, replace = TRUE, prob = c(9/19, 10/19))    # simulate 1000 roulette spins
    sum(X)    # determine total profit
})

hist(S) # distribution of results
```




```
mean(S) # mean results
```

```
## [1] 52.5392
```

```
mean(S < 0)    # probability of the casino losing money
```

```
## [1] 0.0442
```