Foundations of Probability in R

2023-02-05

Contents

The binomial distribution

Simulating coin flips

In these exercises, you'll practice using the rbinom() function, which generates random "flips" that are either 1 ("heads") or 0 ("tails").

```
# Generate 10 separate random flips with probability .3
rbinom(10, 1, p = 0.3)
```

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

Simulating draws from a binomial

In the last exercise, you simulated 10 separate coin flips, each with a 30% chance of heads. Thus, with rbinom(10, 1, .3) you ended up with 10 outcomes that were either 0 ("tails") or 1 ("heads").

But by changing the second argument of rbinom() (currently 1), you can flip multiple coins within each draw. Thus, each outcome will end up being a number between 0 and 10, showing the number of flips that were heads in that trial.

```
# Generate 100 occurrences of flipping 10 coins, each with 30% probability

rbinom(100, 10, p = 0.3)
```

```
##   [1] 2 2 4 2 3 3 4 3 5 4 1 5 3 2 5 5 1 3 2 5 3 8 4 4 2 3 4 1 3 1 7 3 1 2 2 2 1
##  [38] 4 4 3 1 6 2 3 7 6 0 5 1 3 2 3 3 5 2 2 4 4 4 1 4 4 3 5 3 3 4 2 3 4 3 2 3 3
##  [75] 2 5 5 2 0 5 4 6 5 1 1 4 3 1 4 3 1 3 5 2 4 1 1 1 6 0
```

Calculating density of a binomial

If you flip 10 coins each with a 30% probability of coming up heads, what is the probability exactly 2 of them are heads?

- Answer the above question using the dbinom() function. This function takes almost the same arguments as rbinom(). The second and third arguments are size and prob, but now the first argument is x instead of n. Use x to specify where you want to evaluate the binomial density.
- Confirm your answer using the rbinom() function by creating a simulation of 10,000 trials. Put this all on one line by wrapping the mean() function around the rbinom() function.

```r
# Calculate the probability that 2 are heads using dbinom

dbinom(2, 10, .3)
```

```
## [1] 0.2334744
```

```r
# Confirm your answer with a simulation using rbinom

#flips <- rbinom(10000, 10, .3)

mean(rbinom(10000, 10, .3) == 2)
```

```
## [1] 0.2354
```

Calculating cumulative density of a binomial

If you flip ten coins that each have a 30% probability of heads, what is the probability at least five are heads?

- Answer the above question using the pbinom() function. (Note that you can compute the probability that the number of heads is less than or equal to 4, then take 1 - that probability).
- Confirm your answer with a simulation of 10,000 trials by finding the number of trials that result in 5 or more heads.

```r
# Calculate the probability that at least five coins are heads

1 - pbinom(4, 10, .3)
```

```
## [1] 0.1502683
```

```r
# Confirm your answer with a simulation of 10,000 trials

mean(rbinom(10000, 10, .3) >=  5)
```

```
## [1] 0.1541
```

Varying the number of trials

In the last exercise you tried flipping ten coins with a 30% probability of heads to find the probability at least five are heads. You found that the exact answer was 1 - pbinom(4, 10, .3) = 0.1502683, then confirmed with 10,000 simulated trials.

- Did you need all 10,000 trials to get an accurate answer? Would your answer have been more accurate with more trials?

```r
# Here is how you computed the answer in the last problem
mean(rbinom(10000, 10, .3) >= 5)
```

```
## [1] 0.1571
```

```r
# Try now with 100, 1000, 10,000, and 100,000 trials

mean(rbinom(100, 10, .3) >= 5)
```

```
## [1] 0.18
```

```r
mean(rbinom(1000, 10, .3) >= 5)
```

```
## [1] 0.151
```

```r
mean(rbinom(10000, 10, .3) >= 5)
```

```
## [1] 0.147
```

```r
mean(rbinom(100000, 10, .3) >= 5)
```

```
## [1] 0.14989
```