THE UNIVERSITY OF TEXAS AT AUSTIN
McCOMBS SCHOOL OF BUSINESS

Simulation 1

Lecture 27

STA 371G

What is simulation?

- Sometimes (e.g., using decision trees), it's possible for us to work out the expected outcome for a situation analytically.

What is simulation?

- Sometimes (e.g., using decision trees), it’s possible for us to work out the expected outcome for a situation analytically.
- But decision trees only tell us about averages (expected value), and given only one “bite at the apple,” our actual outcome may not be close to the average.

What is simulation?

- Sometimes (e.g., using decision trees), it's possible for us to work out the expected outcome for a situation analytically.
- But decision trees only tell us about averages (expected value), and given only one "bite at the apple," our actual outcome may not be close to the average.
- And what do we do when faced with a scenario that can't be modeled by decision trees, e.g., predicting the value of a portfolio at retirement age?

What is simulation?

- Sometimes (e.g., using decision trees), it's possible for us to work out the expected outcome for a situation analytically.
- But decision trees only tell us about averages (expected value), and given only one "bite at the apple," our actual outcome may not be close to the average.
- And what do we do when faced with a scenario that can't be modeled by decision trees, e.g., predicting the value of a portfolio at retirement age?
- Simulation lets us get a picture of the full distribution of possible outcomes in just about any problem scenario.

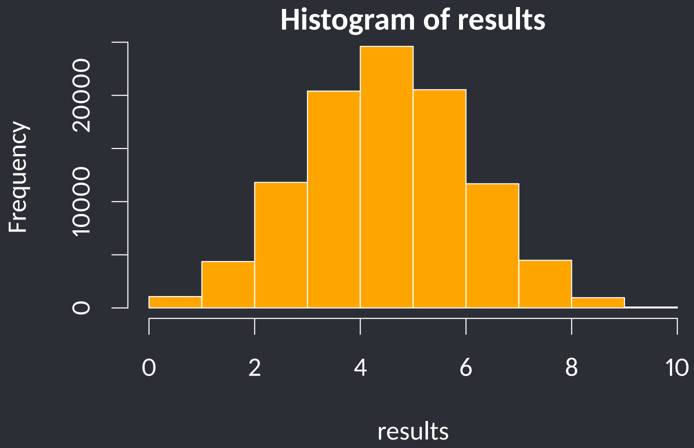
What is simulation?

- Sometimes (e.g., using decision trees), it’s possible for us to work out the expected outcome for a situation analytically.
- But decision trees only tell us about averages (expected value), and given only one “bite at the apple,” our actual outcome may not be close to the average.
- And what do we do when faced with a scenario that can’t be modeled by decision trees, e.g., predicting the value of a portfolio at retirement age?
- Simulation lets us get a picture of the full distribution of possible outcomes in just about any problem scenario.
- We’ll start with some silly examples today and then delve into business applications next time.

Example 1: Coin flipping

Suppose we flip a coin 10 times. What will the distribution of the number of heads look like?

```r
results <- replicate(100000, {
  flips <- sample(c(0, 1), 10, replace=T)
  sum(flips)
})
hist(results, breaks=10, col='orange')
```



Theory tells us that the outcome of flipping a coin 10 times and counting heads should be a Binomial distribution with expected value $np = 10 \cdot 0.5 = 5$ and SD $\sqrt{np(1-p)} = \sqrt{10 \cdot 0.5 \cdot 0.5} \approx 1.58$.

We can compare our simulated results against the theoretical results:

```
mean(results)

[1] 5.00154

sd(results)

[1] 1.581519
```


LET'S
MAKE
A
DEAL


1
2
3


GSN
CADILLAC


ZONK!!

Example 3: Monty Hall and *Let's Make a Deal*

- Monty Hall used to host the game show *Let's Make a Deal*.

Example 3: Monty Hall and *Let's Make a Deal*

- Monty Hall used to host the game show *Let's Make a Deal*.
- There are three doors: two contain a goat, and one contains...

Example 3: Monty Hall and *Let’s Make a Deal*

- Monty Hall used to host the game show *Let’s Make a Deal*.
- There are three doors: two contain a goat, and one contains...

Example 3: Monty Hall and *Let's Make a Deal*

- Monty Hall used to host the game show *Let's Make a Deal*.
- There are three doors: two contain a goat, and one contains...a new car!
- You (the contestant) select one of three doors.

Example 3: Monty Hall and *Let's Make a Deal*

- Monty Hall used to host the game show *Let's Make a Deal*.
- There are three doors: two contain a goat, and one contains…a new car!
- You (the contestant) select one of three doors.
- Without revealing what is behind the selected door, Monty opens one of the other two doors to reveal a goat.

Example 3: Monty Hall and *Let's Make a Deal*

- Monty Hall used to host the game show *Let's Make a Deal*.
- There are three doors: two contain a goat, and one contains…a new car!
- You (the contestant) select one of three doors.
- Without revealing what is behind the selected door, Monty opens one of the other two doors to reveal a goat.
- Monty then gives you a choice: keep your original door, or switch to the other (unopened) door.

Example 2: Monty Hall and *Let's Make a Deal*

Do you have a better chance of getting the car by switching, or by keeping your original selection—or does it not matter?

```r
doors <- c(1, 2, 3)
results <- replicate(100000, {
  # Randomly select my door and the door with the car.
  car.door <- sample(doors, 1)
  my.door <- sample(doors, 1)
  # If I chose the door with the car, he randomly opens one
  #   of the other two doors.
  # If I chose another door, he opens the remaining door.
  if (car.door == my.door) {
    monty.opens.door <- sample(setdiff(doors, my.door), 1)
  } else {
    monty.opens.door <- setdiff(doors, c(my.door, car.door))
  }
  # Switch doors: select the door that Monty did not open.
  my.door <- setdiff(doors, c(monty.opens.door, my.door))
  my.door == car.door
})
sum(results) / 100000

[1] 0.66824
```

```r
doors <- c(1, 2, 3)
results <- replicate(100000, {
  # Randomly select my door and the door with the car.
  car.door <- sample(doors, 1)
  my.door <- sample(doors, 1)
  # If I chose the door with the car, he randomly opens one
  #   of the other two doors.
  # If I chose another door, he opens the remaining door.
  if (car.door == my.door) {
    monty.opens.door <- sample(setdiff(doors, my.door), 1)
  } else {
    monty.opens.door <- setdiff(doors, c(my.door, car.door))
  }
  # Don't switch; just check to see if I won the car.
  my.door == car.door
})
sum(results) / 100000
```

```
[1] 0.3306
```

Simulating random variables

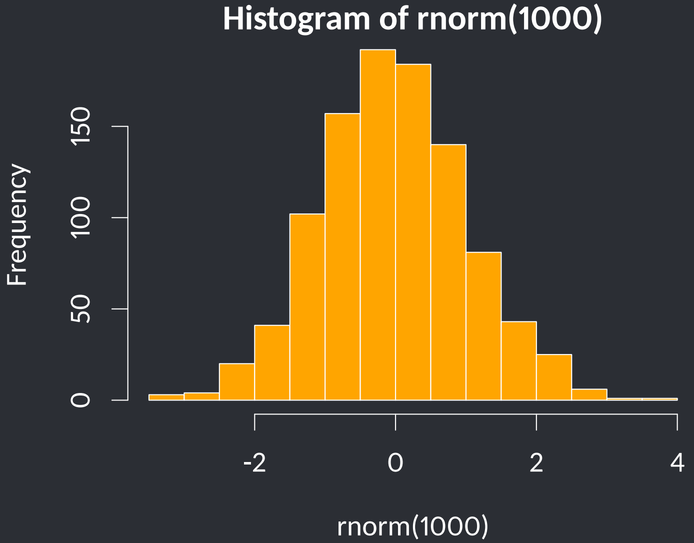
We have seen how the `sample` command can be used to draw from a set of alternatives with equal probability (e.g., flipping a coin). The `rnorm` command can be used to draw randomly from a normal distribution. Let's create 10 random heights, with mean 68 (inches) and SD 4.

```
rnorm(10, 68, 4)

 [1] 68.57607 63.99855 70.53150 65.56610 66.38570
 [6] 69.82433 67.64435 69.96917 72.92480 70.23245
```

Let's check that `rnorm` works as advertised!

```
hist(rnorm(1000), col='orange')
```



Example 3: Will I get an A?

- Let's say I'm taking a class with two midterms (each 25% weight in final grade) and a final exam (50% weight in final grade).

Example 3: Will I get an A?

- Let's say I'm taking a class with two midterms (each 25% weight in final grade) and a final exam (50% weight in final grade).
- I just got my score on the first midterm (75%).

Example 3: Will I get an A?

- Let's say I'm taking a class with two midterms (each 25% weight in final grade) and a final exam (50% weight in final grade).
- I just got my score on the first midterm (75%).
- I want to know how likely it is that I can get 90% or above on my final grade.

Example 3: Will I get an A?

- Let's say I'm taking a class with two midterms (each 25% weight in final grade) and a final exam (50% weight in final grade).
- I just got my score on the first midterm (75%).
- I want to know how likely it is that I can get 90% or above on my final grade.
- This is hard!

Example 3: Will I get an A?

Let’s start by making some assumptions:

- I think I can improve on the second midterm, and then even more on the final.

Example 3: Will I get an A?

Let’s start by making some assumptions:

- I think I can improve on the second midterm, and then even more on the final.
- I’ll model my Midterm 2 grade as a normal distribution.

Example 3: Will I get an A?

Let’s start by making some assumptions:

- I think I can improve on the second midterm, and then even more on the final.
- I’ll model my Midterm 2 grade as a normal distribution.
- My best guess is that I’ll get an 80% on Midterm 2, and I’m 95% sure it will be between 70% and 90%

Example 3: Will I get an A?

Let's start by making some assumptions:

- I think I can improve on the second midterm, and then even more on the final.
- I'll model my Midterm 2 grade as a normal distribution.
- My best guess is that I'll get an 80% on Midterm 2, and I'm 95% sure it will be between 70% and 90%
- So my Midterm 2 grade should be simulated as a normal distribution with mean 80 and SD 5 (since 95% of a normal distribution is roughly ± 2 SD from the mean).

Example 3: Will I get an A?

Let's start by making some assumptions:

- I think I can improve on the second midterm, and then even more on the final.
- I'll model my Midterm 2 grade as a normal distribution.
- My best guess is that I'll get an 80% on Midterm 2, and I'm 95% sure it will be between 70% and 90%
- So my Midterm 2 grade should be simulated as a normal distribution with mean 80 and SD 5 (since 95% of a normal distribution is roughly ± 2 SD from the mean).
- I think I can improve more on the final; my best guess is that I'll get a 90%, and I'm 95% sure I'll get between 80% and 100%.

Example 3: Will I get an A?

- For each run, we will:

Example 3: Will I get an A?

- For each run, we will:
  1. Randomly draw a Midterm 2 score from its normal distribution, and a Final Exam score from its normal distribution.

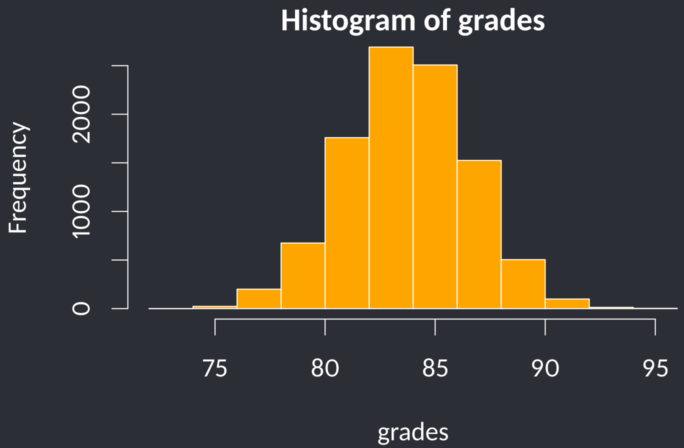
Example 3: Will I get an A?

- For each run, we will:
    1. Randomly draw a Midterm 2 score from its normal distribution, and a Final Exam score from its normal distribution.
    2. Calculate a final score for the course, and see if it's over 90%.

Example 3: Will I get an A?

- For each run, we will:
    1. Randomly draw a Midterm 2 score from its normal distribution, and a Final Exam score from its normal distribution.
    2. Calculate a final score for the course, and see if it's over 90%.
- Then we will count the percentage of runs where we got 90%+ for the course. That will be our estimate of the probability of getting an A.

```r
grades <- replicate(10000, {
  midterm1 <- 75
  midterm2 <- rnorm(1, mean=80, sd=5)
  final.exam <- rnorm(1, mean=90, sd=5)
  0.25*midterm1 + 0.25*midterm2 + 0.5*final.exam
})
hist(grades, col='orange')
```



```r
runs <- replicate(10000, {
  midterm1 <- 75
  midterm2 <- rnorm(1, mean=80, sd=5)
  final.exam <- rnorm(1, mean=90, sd=5)
  0.25*midterm1 + 0.25*midterm2 + 0.5*final.exam >= 90
})
sum(runs) / 10000
```

```
[1] 0.0115
```

```r
runs <- replicate(10000, {
  midterm1 <- 75
  midterm2 <- rnorm(1, mean=80, sd=5)
  final.exam <- rnorm(1, mean=90, sd=5)
  0.25*midterm1 + 0.25*midterm2 + 0.5*final.exam >= 90
})
sum(runs) / 10000

[1] 0.0115
```

There's only about a 1.15% chance that I'll get an A.