# Lecture 1: Intro to simulation

## Warm-up question

**Problem:** 10 people are at a party, and all of them are wearing hats. They each place their hat in a pile; when they leave, they choose a hat at random. What is the probability at least one person selected the correct hat?

**Question:** Work with your neighbor to discuss the following question:

• Without calculating probabilities, how could you design an experiment to estimate this probability?

Designing an experiment step 1; need 10 nats! (each belongs to one person) person 1 person 2 3 ... Person 10 Mct 1 1 Hat 2 1 3 ... Hat b 17 Step 2: randomly "snuffle" nats / sampling randomly assign a hat to each person Step 3: who got their original hat? Step 4; Repect many times! for loop!

# Step 1: representing the hats

```
1 hats <- (1:10) L rector 1,2,3,...,10

rector 2

hats <- (1:10) L rector 1,2,3,...,10

side in the left hand side

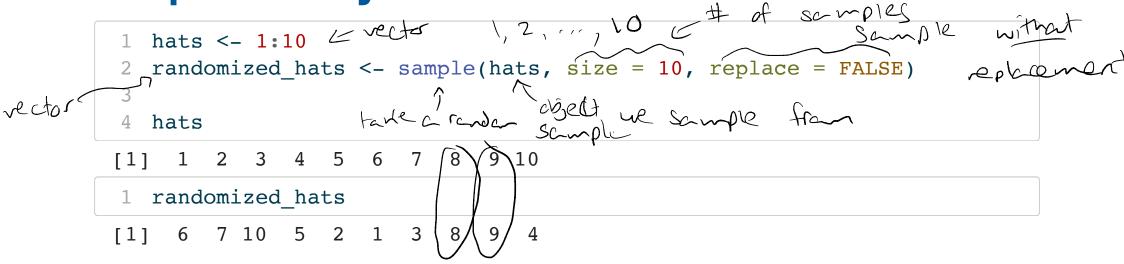
hats [1] 1 2 3 4 5 6 7 8 9 10

1 hats [3]

[1] 3 1 third entry
```

- hats is a vector, containing the numbers 1 to 10
- entries in a vector are accessed by their index

Step 2: everyone draws a random hat



- The sample function creates a random sample from a vector
- How many people selected their original hat?

# Step 3: check who got their original hat

```
1 hats <- 1:10
 2 randomized hats <- sample(hats, size = 10, replace = FALSE)</pre>
 1 hats
[1] 1 2 3 4 5 6 7
                          9\10
 1 randomized hats
             5 2 1 3 \8/\ 9/
[1] 6 7 10
 1 hats == tandomized_hats == tst for equality
[1] FALSE FALSE FALSE FALSE FALSE (
                                           TRUE TRUE FALSE
 1 # TRUE is 1, FALSE is 0
                                    how many people got their original nat?
 2 sum(hats == randomized hats)
[1]/2
 1 # did at least one person get their hat?
 2 sum(hats == randomized hats) > 0
             (ct least are
                              person got their original
[1] TRUE
```

### Code so far

```
1 hats <- 1:10
2 randomized_hats <- sample(hats, size = 10, replace = FALSE)

1 sum(hats == randomized_hats) > 0

[1] TRUE
```

• Is this a good estimate of the probability?

No! Need to reject many times

# Step 4: iteration

#### A for loop repeats code many times:

```
(run code)

= 2

(run code)

(run code)

:
:
```

# Step 4: iteration

results 2- repland, usin) NA NA NA CIODOO

#### A for loop repeats code many times:

```
1 nsim <- 10000 # number of simulations
        2 hats <- 1:10
                         (ncts
        3 results <- rep(NA, nsim) # vector to store results</pre>
                          repeat NA usin times"
results
        5 for(i in 1:nsim){
            randomized hats <- sample(hats, size = 10, replace = FALSE)
            results[i] <- sum(hats == randomized_hats) > 0
(Store
                                    crech weter each person got their
       10 head(results)
                                                prob. = proportion of TRUE
               TRUE TRUE FALSE
       [1] FALSE
                               TRUE
                                     TRUE
          NANA NA .... NA
                                                         in esults verter
            i=\ '
                (check)
           FALSE NA NA NA
            PALSE TRUE NA NA
            1~ 10000
```

# Step 4: iteration

A for loop repeats code many times:

• What if I wanted to repeat the simulation, with a different number of people?

# Removing magic numbers

#### Without magic numbers:

```
1 nsim <- 10000 # number of simulations
 2 M <- 10 # number of people
 3 hats <-(1:M)
  results <- rep(NA, nsim) # vector to store results
   for(i in 1:nsim){
     randomized hats <- sample(hats,
                                replace = FALSE)
10
     results[i] <- sum(hats ==
11
                         randomized hats) > 0
12 }
13
  mean(results)
```

[1] 0.6269

• Why did I get different results?

Final code

setting a seed: each time I wn a process who random' functions, reproducibility I get the same esuit

```
1 set.seed(3) # set a seed for reproducibility
 3 M <- 10 # number of people at the party
 4 hats <- 1:M # numbered hats
 5 nsim <- 10000 # number of simulations
 6 results <- rep(NA, nsim) # vector to store the results
   for(i in 1:nsim){
     # hats are randomly assigned to each person
10
     randomized hats <- sample(hats, M, replace = F)
11
12
     # did at least one person get their hat back?
     results[i] <- sum(randomized hats == hats) > 0
13
14 }
15
16 mean(results)
```

## Summary of coding practices

- avoid magic numbers
- set a seed for reproducibility
- use meaningful names
- add comments

## Class activity

Work with a neighbor on the class activity (link below and on the course website):

https://sta279s24.github.io/class\_activities/ca\_lecture\_1.html