

# MATH 4780 (MSSC 5780) In-class Activity 1

September 28, 2023

David and Rachel, please present your work on Sept 19. Remember to ask undergraduate students two Non-yes-no questions. You, as a teacher, try to get students understand your work.

## Programming simulation (David)

Write a function to

1. Simulate  $n$  data points from the model

$$y = \beta_0 + \beta_1 x + \epsilon$$

with data points  $x$  uniformly sampled from the range  $(0, 100)$  and  $\epsilon$  drawn independently from  $N(0, \sigma^2)$ .

2. Fit a linear regression to the data.
3. What happens to the parameter estimates and uncertainties when the sample size is increased?

Your function should take arguments  $\beta_0, \beta_1, n, \sigma$ , and it should return the data, and print out the fitted regression. The template of the function is

```
function_name <- function(beta0, beta1, n, sigma) {  
  
  ## generate data  
  
  ## fit regression  
  
  ## show the output  
  
  ## return the data
```

```
    return(your_data)
}
```

## No-intercept Model (Rachel)

Generate the data using the following code

```
x <- runif(100)
y <- 5 * x + rnorm(100) ## beta_0 = 0, beta_1 = 5
```

Here you can see the true intercept is zero and slope is five.

1. Suppose we don't know both  $\beta_0$  and  $\beta_1$ . Fit the linear regression to the data.
2. Suppose we do know  $\beta_0 = 0$  and fit the no-intercept regression  $y = \beta_1 x + \epsilon$  to the data.
3. Compare the confidence interval for  $\beta_1$  from the model with and without intercept. Is one always shorter than the other? Discuss it.