WU #3 - Simulating CIs

Math 154 - Fall 2017 - Jo Hardin

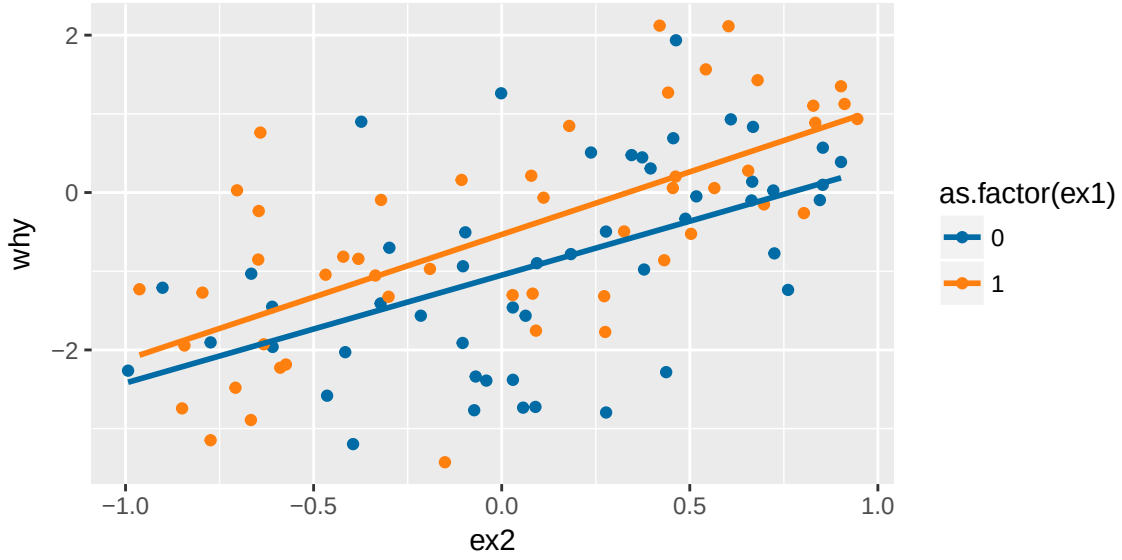
Name: _____

Let's say that you want to check the mathematical formula for creating CIs for the slope coefficient. In particular, you want to know if the procedure you use actually does capture the population parameter 95% of the time. You are simulating from the following population model:

$$Y = -1 + 0.5X_1 + 1.5X_2 + \epsilon, \quad \epsilon \sim N(0,1)$$

Consider:

```
n <- 100
x1 <- rep(c(0,1), each=n/2)
x2 <- runif(n, min=-1, max=1)
```



Write down the steps for the simulation that would assess whether the CI procedure is valid for the parameter on X_2. (No R code here, use words to enumerate the process.)

Solution:

```
CI <- lm(y~x1+x2) %>% tidy(conf.int=TRUE)

CI
```

```
##          term   estimate std.error statistic      p.value
## 1 (Intercept) -0.7301088 0.1429774 -5.106462 1.638299e-06
## 2          x1  0.2206139 0.2028775  1.087424 2.795441e-01
## 3          x2  1.4315449 0.1794914  7.975564 3.024212e-12
##     conf.low  conf.high
## 1 -1.0138795 -0.4463382
## 2 -0.1820417  0.6232695
## 3  1.0753042  1.7877855
```

```
between(beta2, CI[3,6], CI[3,7])
```

```
## [1] TRUE
```

```
beta2.in <- c()
beta0 <- -1; beta1 <- 0.5; beta2 <- 1.5
n<- 100; reps=10000

set.seed(4747)
for(i in 1:reps){
  x1 <- rep(c(0,1), each=n/2)
  x2 <- runif(n, min=-1, max=1)
  y <- beta0 + beta1*x1 + beta2*x2 + rnorm(n, mean=0, sd = 1)

  CI <- lm(y~x1+x2) %>% tidy(conf.int=TRUE)

  beta2.in <- c(beta2.in, between(beta2, CI[3,6], CI[3,7]))
}
sum(beta2.in)/reps
```

```
## [1] 0.952
```