

Simulação - TLC

Marina Muradian

Aula 2 - Marketing Analytics

População

Vamos supor que sabemos que em certa cidade a probabilidade de uma pessoa ter a intenção de comprar certo produto é de 45%.

Como a nossa variável de interesse X só pode assumir os valores 1 (com probabilidade de 45%) e 0 (com probabilidade de 55%), vamos definir a variável `pop`:

```
pop <- c(0,1)
```

`n = 2`

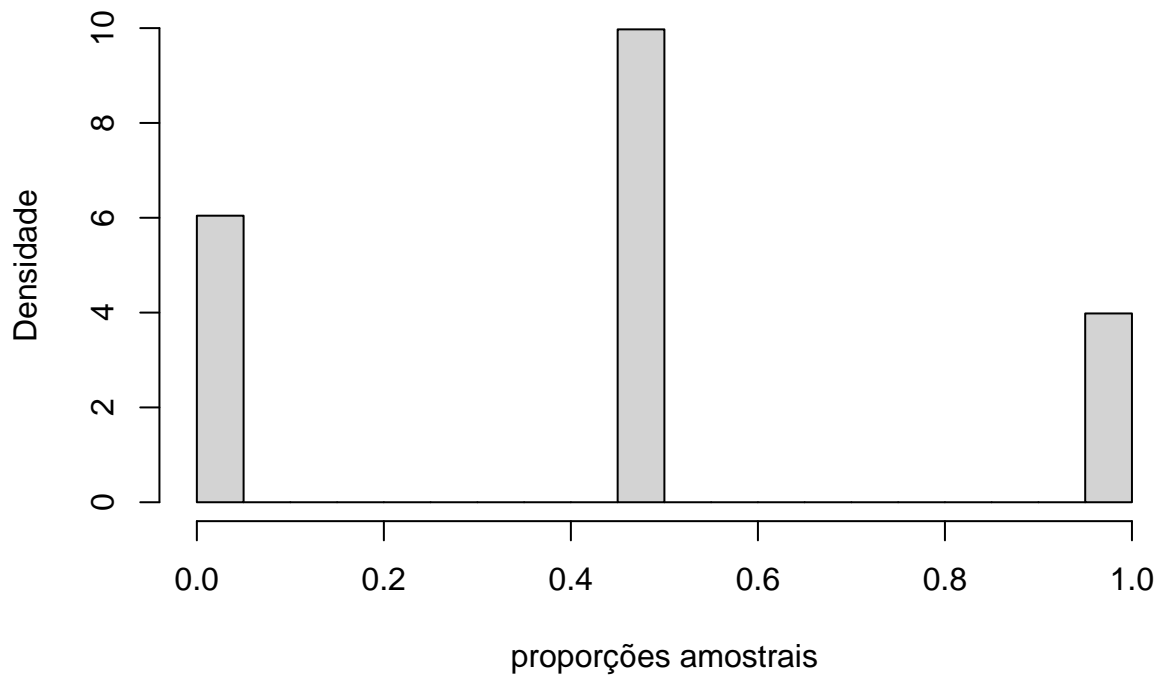
Vamos começar nossa simulação, gerando 10000 amostras de tamanho (n) 2 desta população:

```
n2 <- matrix(
  sample(pop, size = 10000*2,
         replace = TRUE,
         prob = c(0.55, 0.45)),
  ncol = 2)
```

Agora, vamos calcular a média destas 10000 amostras e fazer um histograma para visualizá-las:

```
p_hat2 <- rowMeans(n2)
hist(p_hat2,
     probability = TRUE,
     main = "Histograma das médias das amostras quando n = 2",
     ylab = "Densidade",
     xlab = "proporções amostrais")
```

Histograma das médias das amostras quando $n = 2$



Vamos repetir essa simulação para $n = 20$, $n = 200$ e $n = 2000$ e verificar o que podemos visualizar no histograma das 10000 médias amostrais obtidas:

$n = 20$

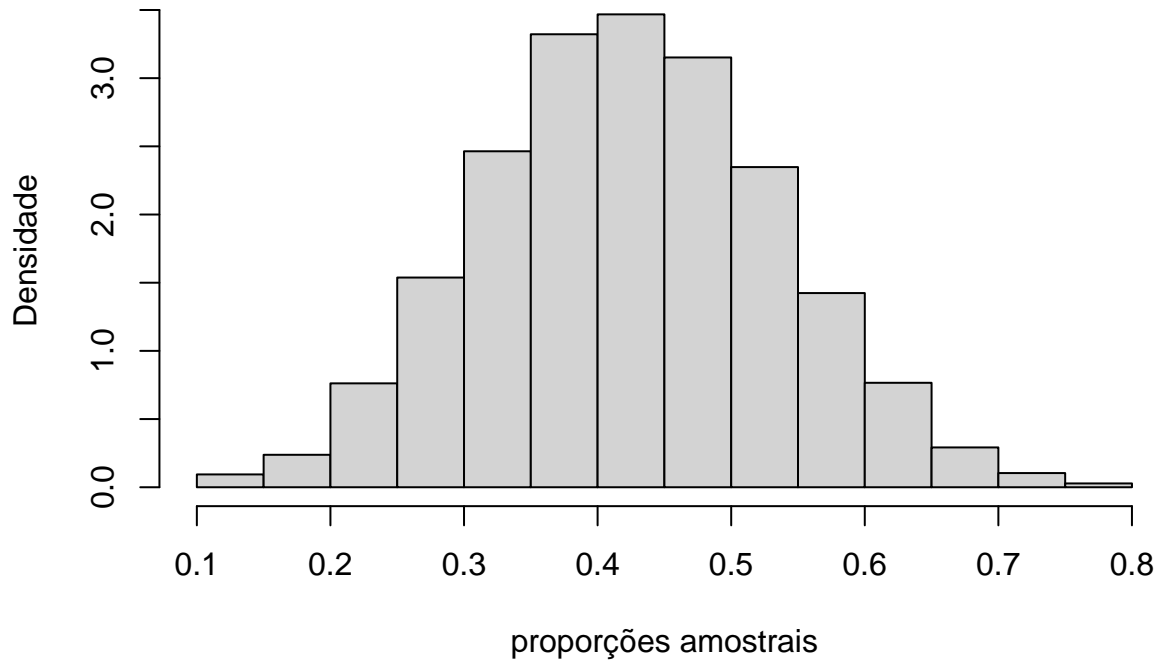
Vamos gerar 10000 amostras de tamanho (n) 20 desta população:

```
n20 <- matrix(  
  sample(pop,  
    size = 10000*20,  
    replace = TRUE,  
    prob = c(0.55, 0.45)),  
  ncol = 20)
```

Agora, vamos calcular a média destas 10000 amostras e fazer um histograma para visualizá-las:

```
p_hat20 <- rowMeans(n20)  
hist(p_hat20,  
  probability = TRUE,  
  main = "Histograma das médias das amostras quando  $n = 20$ ",  
  ylab = "Densidade",  
  xlab = "proporções amostrais")
```

Histograma das médias das amostras quando $n = 20$



$n = 200$

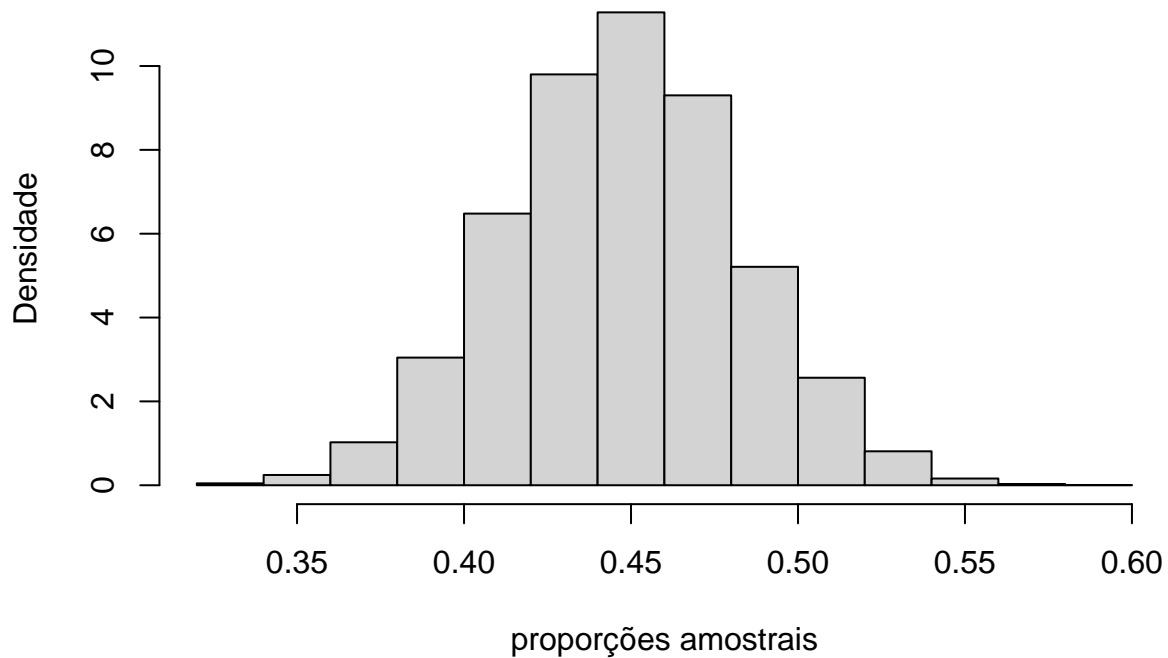
Vamos gerar 10000 amostras de tamanho (n) 200 desta população:

```
n200 <- matrix(  
  sample(pop,  
    size = 10000*200,  
    replace = TRUE,  
    prob = c(0.55, 0.45)),  
  ncol = 200)
```

Agora, vamos calcular a média destas 10000 amostras e fazer um histograma para visualizá-las:

```
p_hat200 <- rowMeans(n200)  
hist(p_hat200,  
  probability = TRUE,  
  main = "Histograma das médias das amostras quando  $n = 200$ ",  
  ylab = "Densidade",  
  xlab = "proporções amostrais")
```

Histograma das médias das amostras quando $n = 200$



n = 2000

Vamos gerar 10000 amostras de tamanho (n) 2000 desta população:

```
n2000 <- matrix(
  sample(pop,
    size = 10000*2000,
    replace = TRUE,
    prob = c(0.55, 0.45)),
  ncol = 2000)
```

Agora, vamos calcular a média destas 10000 amostras e fazer um histograma para visualizá-las:

```
p_hat2000 <- rowMeans(n2000)
hist(p_hat2000,
  probability = TRUE,
  main = "Histograma das médias das amostras quando n = 2000",
  ylab = "Densidade",
  xlab = "proporções amostrais")
```

Histograma das médias das amostras quando $n = 2000$

