

# Roteiro de aula Prática - Amostragem e Teorema do Limite Central

Prof. Leonardo Carvalho

2022-11-06

## TEORIA NECESSÁRIA

### Calculando o tamanho amostral em pop. finitas.

Considerando uma população com Tamanho da população = 5400 ou seja:  $N = 5400$  e uma taxa de erro assumida de 5%

Então:

- Código em R:

```
N = 5400
populacao = rnorm(5400, mean = 50, sd = 5)
```

1. Passo: Otenha o  $n_0$ .

- Código em R:

```
n0 = 1/((0.05)^2)
n0
```

```
## [1] 400
```

2. Passo: De posse desse valor, obtenha o valor do tamanho amostral  $n$ .

- Código em R:

```
n = (N*n0)/(N+n0)
n
```

```
## [1] 372.4138
```

3. Passo: Retire a amostra  $n$  calculada da população  $N$  com a função *sample*.

- Código em R:

```
amostra = sample(populacao, 372)
amostra
```

```
## [1] 54.31278 43.35777 51.99667 43.32833 44.50014 57.08790 45.19152 52.97843
## [9] 54.13590 52.05443 50.41547 44.97177 48.77960 47.27087 63.97476 60.11716
## [17] 57.40285 57.37145 50.13339 39.36330 48.19920 55.48940 40.88965 46.05325
## [25] 48.08397 52.99480 48.36053 41.76543 43.13639 48.69324 49.65650 41.55374
## [33] 47.60233 54.00246 57.40127 52.08038 46.75323 44.48417 45.98559 50.11655
## [41] 49.81622 46.56477 53.45846 44.22771 42.01970 51.73289 50.71716 47.79719
## [49] 55.55624 53.76758 42.10212 46.19007 41.02717 48.40921 48.48785 47.58657
## [57] 56.23768 46.69224 45.21123 49.75740 52.06475 46.81050 50.53667 50.60310
## [65] 49.94369 56.31997 47.92000 53.33982 48.33653 51.71203 47.93580 49.49253
## [73] 50.65787 47.95947 59.01539 51.51140 58.20537 52.20438 45.95411 54.21538
## [81] 42.89834 52.29922 57.00702 50.79220 40.45815 55.18077 54.94539 48.70965
## [89] 44.38702 59.29851 52.53265 36.25453 52.40484 51.13434 48.56057 50.65974
## [97] 52.48088 50.07648 54.61723 58.12977 42.84541 47.02528 49.01924 55.18063
## [105] 54.75032 48.97423 43.97262 59.77036 62.02080 49.51058 46.52039 45.45698
## [113] 42.73858 50.75304 49.73653 54.41358 47.69366 50.38337 45.42105 44.59647
## [121] 40.04567 45.48169 39.72378 42.94490 53.97975 53.03237 47.36415 49.17800
## [129] 51.97411 41.51134 45.49032 47.43666 56.60171 54.74508 53.83766 53.56305
## [137] 52.62941 52.96890 48.27679 49.65262 49.05860 53.98466 50.51545 51.13964
## [145] 49.78473 49.27462 46.79149 45.77183 46.09575 50.50315 44.51057 56.86690
## [153] 53.36827 49.10302 60.78892 44.76098 49.62907 50.12994 54.28134 52.25479
## [161] 52.96344 55.03382 50.13538 52.07814 46.59990 53.29167 46.71717 50.52834
## [169] 49.00007 53.32258 53.87166 52.39576 55.94384 43.58319 44.29925 47.69999
## [177] 41.62304 49.56177 44.69766 61.19061 48.91412 49.72871 57.45038 51.03830
## [185] 56.78862 52.19344 41.55680 41.58027 47.71537 58.02396 45.65846 47.15481
## [193] 48.31043 44.77630 49.86164 50.12417 43.88533 51.98558 49.55510 50.41490
## [201] 46.16225 48.21956 56.47441 56.92558 45.01860 55.30764 54.02238 46.81689
## [209] 48.95725 51.45301 50.82019 46.08884 48.72379 44.97855 48.94332 64.07600
## [217] 45.18658 41.27206 47.46884 43.94237 57.23935 57.87456 51.77216 42.36438
## [225] 46.44514 56.79911 49.92936 57.67703 44.88374 49.91611 50.62274 44.78739
## [233] 48.22255 50.51982 51.90457 41.70569 48.64968 43.77238 52.90612 41.39721
## [241] 55.90844 42.26150 58.17023 51.66245 51.00441 56.99852 59.41101 55.61435
## [249] 47.58734 56.98983 58.39928 49.47039 55.03588 50.28359 55.05470 55.69931
## [257] 43.28913 46.73094 52.75024 50.28790 51.51083 48.93991 54.06178 50.08611
## [265] 48.65502 53.39083 44.32270 52.15349 45.08510 53.83894 45.84533 44.09358
## [273] 48.30791 55.72950 52.46823 47.38491 49.39198 53.46387 55.07362 51.35832
## [281] 47.05354 48.74530 47.78902 53.62937 52.43697 50.21939 44.76917 55.05930
## [289] 49.35861 49.60022 42.22151 51.12627 45.37353 60.64256 51.35763 56.90910
## [297] 52.76585 48.48622 45.96438 48.03138 48.48251 47.30573 55.78825 53.51748
## [305] 44.57280 52.94374 47.88052 53.08274 49.41403 64.35717 53.46571 47.75471
## [313] 54.61502 51.87236 50.60128 49.24369 48.70380 50.04042 46.33661 47.24393
## [321] 49.10266 49.51394 48.38618 58.63379 49.47361 55.62961 57.37493 53.66295
## [329] 48.79235 54.29258 54.67931 53.74599 45.15033 46.97543 51.86834 43.11009
## [337] 50.48455 37.57222 52.87218 49.05601 42.39415 50.30076 44.16418 51.39256
## [345] 50.00631 45.82045 43.09473 50.76609 48.26778 50.14054 51.04635 40.97099
## [353] 55.79906 52.26865 48.78627 51.48059 51.77522 47.44924 60.13206 52.98030
## [361] 51.03498 55.99643 56.96719 46.75466 42.68710 48.72568 53.41656 53.72333
## [369] 50.08372 48.64368 45.44480 42.57030
```

Após compreender como descobrir o tamanho da TOTAL da amostra podemos facilmente calcular os tamanhos amostrais a serem obtidos nas situações de **amostragem simples** e **amostragem estratificada**.

Considerando um tamanho de amostra calculado  $n = 56$  e uma população dividida em 4 estratos  $k = 4$ .

#### Primeiro caso - Os estratos tem tamanhos N iguais:

- Se  $N_1 = N_2 = NK$ , então:
  - Número de indivíduos por estrato:  $n_K = \frac{n}{K}$
- Código em R:

```
n = 56
k = 4
n_final = n/k
n_final
```

```
## [1] 14
```

Portanto, o tamanho da amostra para cada estrato é  $n = 14$

#### Segundo caso - Os estratos tem tamanhos de N diferentes(40, 80 e 120):

- Se  $N_1$  é diferente de  $N_2, \dots, N_K$ , então:
  - Número de indivíduos por estrato:  $n_K = \frac{nN_k}{N}$

Tendo em vista que o tamanho total da população é a soma dos tamanhos de cada estrato, então:

- $N \text{ total} = 40 + 80 + 120 = 240$

#### Podemos então calcular o tamanho amostral para cada estrato:

- Código em R:

```
n = 56
N_total = 240

n1 = (n*40)/N_total
n2 = (n*80)/N_total
n3 = (n*120)/N_total
n1
```

```
## [1] 9.333333
```

```
n2
```

```
## [1] 18.66667
```

n3

## [1] 28

Portanto os tamanhos das amostras para os estratos 1, 2 e 3 são  $n_1 = 9$ ,  $n_2 = 18$  e  $n_3 = 28$ , respectivamente.

## PRÁTICA

### Desenho do problema:

Considere que você é o pesquisador responsável por estudar uma espécie de copépodes (*Undinula vulgaris*) distribuídos ao longo do litoral paraense em três faixas distintas e esparçadas espacialmente (Faixa 1: 3000 indivíduos, Faixa 2: 5000 indivíduos, Faixa 3: 2000 indivíduos) de acordo com a figura:

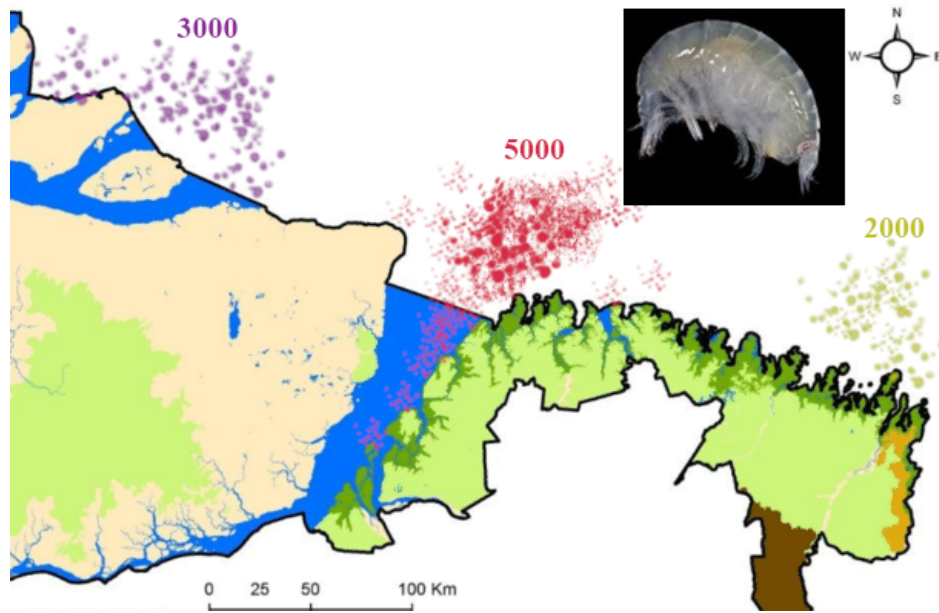


Figure 1: Distribuição hipotética de copépodes na região costeira do estado do Pará

A tarefa é estimar o tamanho médio da espécie e sua variação na região. Para tal você precisa:

1. Descobrir o tamanho amostral para as coletas.
2. Tomar decisão sobre o tipo de população
3. Calcular a média e o desvio padrão da(s) amostra(s).

- Dicas:

- Utilizar as expressões adequadas para a obtenção do  $n$  amostral
- Obter as amostras utilizando a função `sample` do *R*
- Usar as funções `mean` e `sd` do *R*

Simulação do arquivo de dados em *R*:

- Código em *R*:

```
AREA1 = data.frame("Comprimento_area1" = rnorm(3000, mean = 3, sd = 0.2))  
AREA2 = data.frame("Comprimento_area2" = rnorm(5000, mean = 2, sd = 0.2))  
AREA3 = data.frame("Comprimento_area3" = rnorm(2000, mean = 2, sd = 0.2))
```