

**Universidade Federal do Rio Grande do Sul**

**MAT02036 - Amostragem 2**

**Seminário 1 - Relatório**

# **Ativações de Tela por Alunos da Estatística**

**Grupo 7: Amostragem Sistemática - Média**

**Autores:** Davi Augusto, Gabriel Netto e João Arend

**Data:** 27 de Nov de 2025

# Contents

<b>1</b>	<b>Motivação</b>	<b>2</b>
<b>2</b>	<b>Objetivos</b>	<b>2</b>
<b>3</b>	<b>Métodos</b>	<b>2</b>
3.1	População alvo . . . . .	2
3.2	População de pesquisa . . . . .	2
3.3	Unidade elementar . . . . .	2
3.4	Unidade amostral . . . . .	2
3.5	Sistema de referência . . . . .	3
3.6	Forma de coleta dos dados . . . . .	3
3.7	Característica de interesse e parâmetro . . . . .	3
3.8	Plano amostral e sua implementação . . . . .	3
3.9	Estimador . . . . .	3
3.9.1	Estimador Pontual da Média . . . . .	3
3.9.2	Estimador Pontual da Variância da Média . . . . .	4
3.10	Inferência (estimação ou teste) . . . . .	4
3.11	Nível de confiança ou significância . . . . .	4
3.12	Software Utilizado . . . . .	4
<b>4</b>	<b>Resultados</b>	<b>4</b>
4.1	Tabela Descritiva . . . . .	5
4.2	Histograma . . . . .	5
4.3	Estimadores . . . . .	6
4.4	Intervalo de Confiança ( $\alpha = 0,10$ ) . . . . .	6
<b>5</b>	<b>Sugestão tamanho de amostra</b>	<b>7</b>
<b>6</b>	<b>Considerações finais</b>	<b>8</b>
6.1	Limitações . . . . .	8
6.1.1	Erros não amostrais a se considerar: . . . . .	8
<b>7</b>	<b>Apêndices</b>	<b>8</b>
7.1	Sorteio . . . . .	8
7.2	Questionário . . . . .	10
7.2.1	Ativações de Tela por Alunos de Estatística do Semestre 2025/2 . . . .	10
7.2.2	Tutorial . . . . .	10
7.2.3	Perguntas . . . . .	12

# 1 Motivação

Contexto fictício elaborado pelo grupo:

**Solicitante:** Diretório acadêmico da Estatística.

**Necessidade:** O diretório está elaborando um relatório sobre a saúde digital e o bem-estar dos estudantes do curso de estatística. Uma das métricas de interesse é a “distração por smartphone”, que pode impactar nos estudos, saúde mental e de maneira mais geral na qualidade de vida do estudante. Para avaliar essa questão, o número de desbloqueios de tela foi uma variável escolhida.

**Questão de Pesquisa:** “Qual o número médio de desbloqueios de tela de celular por dia entre os alunos de graduação em Estatística da UFRGS?”

## 2 Objetivos

O objetivo do estudo é estimar o número de desbloqueios médio dos alunos de estatística da UFRGS através de amostragem sistemática.

## 3 Métodos

O método de seleção probabilística adotado foi a Amostragem Sistemática. Esse método garante uma cobertura uniforme da lista de matriculados, selecionando os elementos da população a partir de um intervalo fixo após um ponto de partida aleatório.

### 3.1 População alvo

Todos os alunos de graduação do curso de Estatística da UFRGS.

### 3.2 População de pesquisa

Todos os alunos de graduação com matrícula ativa no curso de Estatística no semestre letivo 2025/2.

### 3.3 Unidade elementar

Aluno de graduação do curso de Estatística da UFRGS.

### 3.4 Unidade amostral

Aluno com matrícula ativa no curso de Estatística no semestre letivo 2025/2.

### 3.5 Sistema de referência

O ordenamento de alunos de graduação com matrícula ativa no curso de Estatística no semestre 2025/2. Essa lista é fornecida pela própria UFRGS a qualquer discente através do Portal do Aluno.

### 3.6 Forma de coleta dos dados

Os dados foram coletados através de um questionário elaborado no *Google Forms*. Disponível no link: Questionário

Ao final deste documento também especificamos como ficou o questionário caso tenha algum problema com o link externo.

O link do questionário foi enviado para os alunos sorteados através do *whatsapp*.

### 3.7 Característica de interesse e parâmetro

Número médio de desbloqueios de tela do smartphone por dia dos alunos de graduação na estatística.

### 3.8 Plano amostral e sua implementação

O plano amostral aplicado foi a amostragem sistemática. Os alunos foram organizados em uma lista com base no ordenamento do curso de estatística no semestre 2025/02.

A implementação se deu por sorteio realizado no *R*, cujo o código está disponível ao final deste documento. Pós sorteio o grupo foi responsável por contatar os alunos sorteados para enviar o questionário.

### 3.9 Estimador

#### 3.9.1 Estimador Pontual da Média

A média populacional ( $\bar{Y}$ ) foi estimada através da média amostral simples ( $\bar{y}$ ), calculada sobre o conjunto de respondentes efetivos:

$$\hat{Y}_{AS} = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

Onde:

- $y_i$ : Valor observado (número de desbloqueios) para o  $i$ -ésimo aluno;
- $n$ : Tamanho da amostra efetiva (número de respondentes,  $n = 21$ ).

### 3.9.2 Estimador Pontual da Variância da Média

Para mensurar a precisão da estimativa e construir o Intervalo de Confiança, utilizou-se o estimador da variância da Amostragem Aleatória Simples (AAS) sem reposição, ajustado para população finita:

$$\hat{V}(\bar{y}_{AS}) \approx \left(1 - \frac{n}{N}\right) \frac{s^2}{n}$$

Onde:

- $s^2$ : Variância amostral dos dados coletados ( $n = 21$ );
- $\left(1 - \frac{n}{N}\right)$ : Fator de correção para populações finitas (fpc).

Ps: Assume-se que o ordenamento não apresenta correlação com nossa variável de interesse, permitindo o uso da fórmula da variância da AAS como aproximação conservadora.

### 3.10 Inferência (estimação ou teste)

A abordagem inferencial adotada neste estudo é a Estimação por Intervalo. Considerando nossa questão de pesquisa (“Qual o número médio de desbloqueios de tela de celular por dia”) o nosso objetivo estatístico não é testar uma hipótese pré-definida, mas sim estimar o parâmetro populacional desconhecido ( $\bar{Y}$ ) com uma margem de erro controlada.

Dessa forma, os resultados serão apresentados através de um Intervalo de Confiança (IC). Esta técnica permite fornecer um conjunto de valores plausíveis para a média de desbloqueios de toda a graduação em Estatística, considerando a variabilidade observada na amostra e o erro amostral inerente ao processo.

A construção do intervalo baseia-se na distribuição  $t$  de Student, adequada para o tamanho de amostra obtido e para situações onde a variância populacional é desconhecida.

### 3.11 Nível de confiança ou significância

Para a construção do intervalo, definiu-se um Nível de Confiança de 90% ( $1 - \alpha = 0,90$ ).

### 3.12 Software Utilizado

Todas as análises e códigos disponíveis foram executados em R, tendo como referência a versão mais recente disponível na plataforma PositCloud (4.5.1).

## 4 Resultados

A análise dos dados foi realizada com base nas respostas válidas de  $n = 21$  estudantes, a partir de uma população total de  $N = 192$  alunos matriculados.

Seguem as seguintes estatísticas descritivas relativas a resposta dos estudantes sobre o número de desbloqueios de tela no dia anterior:

## 4.1 Tabela Descritiva

```
library(dplyr)
library(readr)
library(readxl)
library(knitr)

df <- read_xlsx("data/respostas.xlsx")

colnames(df) <- c("ativacoes", "percepcao")

n <- nrow(df)
N <- 192

resumo_descritivo <- df %>%
  summarise(
    Media      = mean(ativacoes),
    DesvioPadrao = sd(ativacoes),
    Variancia   = var(ativacoes),
    Minimo      = min(ativacoes),
    Maximo      = max(ativacoes),
    n           = n()
  )

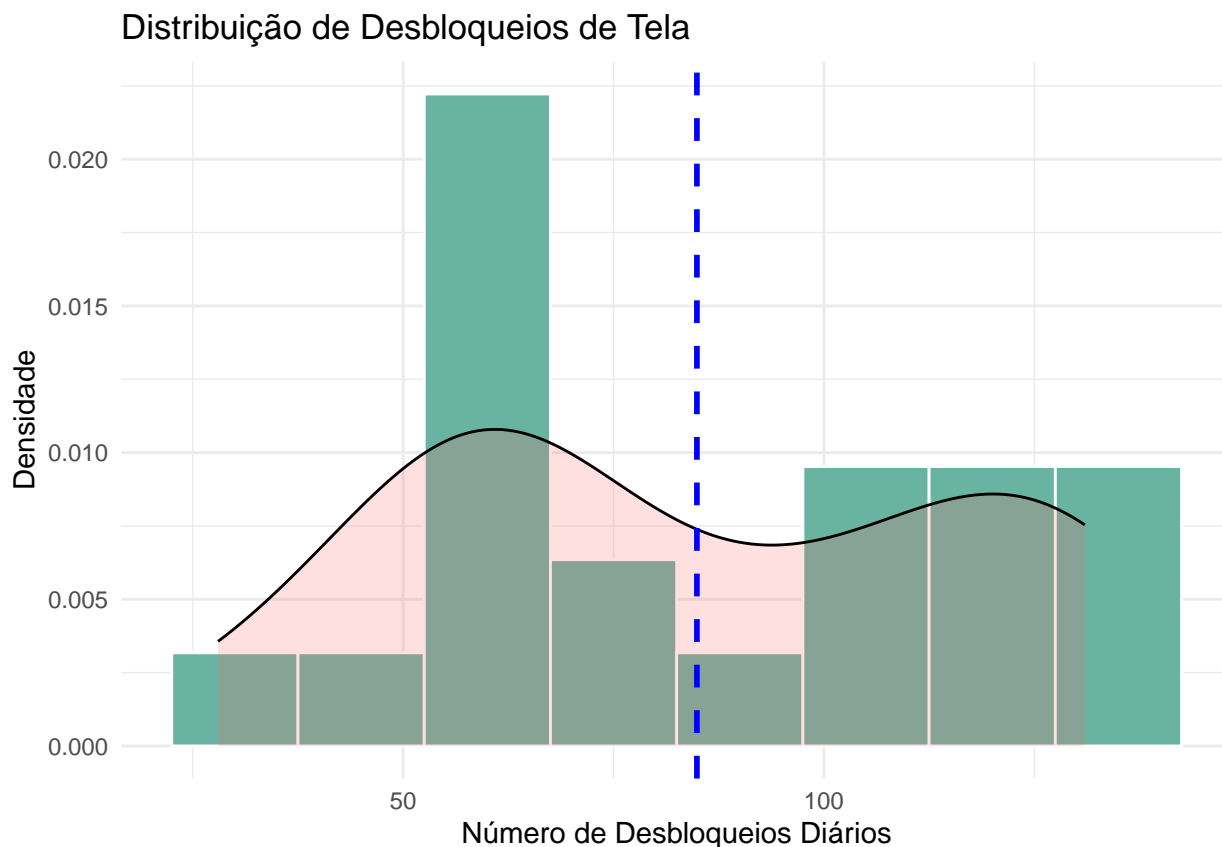
kable(resumo_descritivo, booktabs = TRUE , row.names = FALSE)
```

Media	DesvioPadrao	Variancia	Minimo	Maximo	n
84.90476	33.23388	1104.49	28	131	21

## 4.2 Histograma

```
library(ggplot2)

ggplot(df, aes(x = ativacoes)) +
  geom_histogram(aes(y = after_stat(density)), binwidth = 15, fill = "#69b3a2", color = "blue") +
  geom_density(alpha = 0.2, fill = "#FF6666") +
  labs(
    title = "Distribuição de Desbloqueios de Tela",
    x = "Número de Desbloqueios Diários",
    y = "Densidade"
  ) +
  theme_minimal() +
  geom_vline(aes(xintercept = mean(ativacoes)), color="blue", linetype="dashed", linewidth=1)
```



A linha azul representa a média.

### 4.3 Estimadores

Encontrando os **estimadores** mencionados anteriormente:

```
bar_y <- sum(df$ativacoes) / n
```

$$\hat{Y}_{AS} = 84.9047619$$

```
# Var(bar_y) = (1 - n/N) * (s^2 / n)
```

```
var_bar_y <- (1 - n/N) * ((var(df$ativacoes)) / n)
```

$$\hat{V}(\bar{y}_{sy}) \approx 46.8422$$

### 4.4 Intervalo de Confiança ( $\alpha = 0,10$ )

```
erro_padrao <- sqrt(var_bar_y)
```

```
# Intervalo de Confiança (90%)
```

```
# Usando t-student com n-1 graus de liberdade
```

```
alpha <- 0.10
```

```
t_critico <- qt(1 - alpha/2, df = n - 1)
```

```
margem_erro <- t_critico * erro_padrao
ic_inferior <- bar_y - margem_erro
ic_superior <- bar_y + margem_erro
```

Portando temos uma *margem de erro* de 11.8042103.

Chegado em um *intervalo de confiança*:

IC = [73.1; 96.71].

Explorando o comportamento da variável de interesse, o número de desbloqueios de tela por dia dentro de nossa amostra, observamos que os dados tem uma grande amplitude entre os estudantes, variando de no mínimo 28 para no máximo 131.

## 5 Sugestão tamanho de amostra

Consideramos para sugestão de tamanho de amostra os critérios:

1. O Nível de Confiança ( $Z$ ): Para uma amostra futura diminuiríamos o valor de  $\alpha$  para o tradicional 0.05. Portanto consideramos  $Z = 1.96$ .
2. Erro aceitável ( $E$ ): Para um estudo futuro definimos arbitrariamente que queremos um erro máximo de 10.

A partir dessa definição usamos a seguinte fórmula para sugestão:

$$n_0 = \left( \frac{Z \cdot s}{E} \right)^2$$

E o seguinte ajuste para população finita:

$$n_{final} = \frac{n_0}{1 + \frac{n_0}{N}}$$

Portanto:

```
Z <- 1.96
E <- 10
s <- sd(df$ativacoes)
N <- 192

n0 <- ((Z * s) / E)^2
n_final <- (n0 / (1 + (n0 / N)))
```

chegando na sugestão de  $n = 35$  para um futuro estudo.



## 6 Considerações finais

### 6.1 Limitações

- **Taxa de não resposta:** Apesar de considerarmos o plano amostral com  $n_1 = 20$  inicialmente, fizemos o sorteio com  $n_2 = 40$ , conforme já mencionado. Considerando apenas esse  $n_2$  obtivemos uma taxa de não resposta de 47,5%.

#### 6.1.1 Erros não amostrais a se considerar:

- **Erro de entendimento:** Houve a tentativa do grupo de amenizar esse fator disponibilizando no início do questionário um tutorial que possibilitasse direcionar corretamente o respondente.
- **Erro de mensuração:** Aqui considera-se “chutes” e omissões por parte dos respondentes.

## 7 Apêndices

### 7.1 Sorteio

Segue o sorteio realizado a partir do ordenamento do curso de estatística no semestre 2025/2:

Importando ordenamento 2025/02, disponível no portal de serviços da universidade:

```
ord <- read.csv('data/ordenamento_2025.csv', sep = ',')
```

Definindo parâmetros da amostragem sistemática;

Inicialmente escolhemos um  $n$  arbitrário = 20. Portanto, para seleção da amostra, conforme foi recomendado em aula aumentamos o  $n$  para nos prevenirmos de uma taxa de resposta baixa.

```
# Tamanho da população ---  
N <- nrow(ord)  
  
# Tamanho da amostra ---  
n <- 40  
  
# Intervalo de seleção ---  
k <- floor(N/n)
```

Sorteando os registros:

```
set.seed(42)  
s <- sample(1:k, 1)  
  
ind <- seq(from = s, to = N, by = k)
```

```
amostra <- ord[ind, c(2,3)]
amostra
```

##	Etap	Aluno
## 1	8	LISIANE BUENO PEITZ
## 5	8	LUIZA DALMOLIN BENEDUZI
## 9	8	MARIA DA GRACA GOMES
## 13	8	GABRIEL DORNELLES DE CASTRO
## 17	8	AMANDA RAMOS MACEDO
## 21	8	PIETRA SEHN ASSMUS
## 25	6	JULIA CAUDURO BACKES GEHLEN
## 29	6	EDUARDO AUGUSTO TRAVI GARCEZ
## 33	6	Arthur Sandi Bauermann
## 37	6	Davi Augusto Farinela da Silva
## 41	6	Sofia Gattermann Santana
## 45	6	Mariana Pereira Ferreira Ança
## 49	5	RAÍSSA MAGNAN SCORSATTO
## 53	5	LORENZO MORAES COPETTI
## 57	4	Felipe Kauer Padilha
## 61	4	DIOGO VIERI BOLZAN
## 65	4	ANDREWS ALEXANDER PEREIRA SILVEIRA
## 69	4	Gustavo Ubatuba Farias
## 73	4	BEATRIZ BACK
## 77	3	JOAO EDUARDO VENTURINI FLORES
## 81	3	ARTHUR OPITZ VIEIRA DA CUNHA
## 85	3	LEONARDO LUIS MENEGON
## 89	2	DIOGO ISRAEL SCHWANCK
## 93	2	LÉO CIZILIÃO ROCHA DA SILVA
## 97	2	BRENDA OBERDIEK
## 101	2	LEONARDO RUBERT POHLMANN
## 105	2	HELENA STREIT
## 109	2	KATIA GONCALVES DOS SANTOS
## 113	2	LUCAS SUERTEGARAY NORONHA
## 117	2	VICTOR FERNANDO DUTRA GOMES
## 121	2	Ismail Kenzo Hirakata Musa
## 125	2	VINÍCIOS DOS SANTOS PEREIRA
## 129	2	MARINA RAE L MARQUES
## 133	2	Diego Costa Kenne da Silva
## 137	2	ALESSANDRA BELLÓ SOARES
## 141	2	TAÍS LOUREIRO BELLINI
## 145	1	ILENICE GICÉLIA TROJAHN
## 149	1	JOSEMARA DE PAULA ROCHA
## 153	1	Paula Gabriela da Silva Pires
## 157	1	Fabiano Mello de Souza

## 161	1	DIOGO SEVERO MACHADO
## 165	1	JÚLIA RUVINSKI LEÃO
## 169	1	SANDRO GILBERT IVANISKI DOS SANTOS
## 173	1	WILLIAM SOARES SILVEIRA
## 177	1	CHAIANI DOS ANJOS GONCALVES
## 181	1	Victor dos Santos Ferri
## 185	1	LUCIANO GOULART RIOS
## 189	1	FABIO ROCHA DE SOUZA

Essa é a lista de pessoas a serem contatadas para reponder o forms.

O número de alunos selecionados foi um pouco maior que 40 pelo arredondamento em `k <- floor(N/n)`.

## 7.2 Questionário

Segue o questionário elaborado pelo grupo e disponibilizado aos estudantes sorteados:

### 7.2.1 Ativações de Tela por Alunos de Estatística do Semestre 2025/2

Este formulário faz parte de um trabalho da disciplina de Amostragem 2 da UFRGS do grupo composto pela Davi Augusto F. da Silva, Gabriel Netto e João Vitor Arend.

Estamos conduzindo um estudo focado no bem-estar digital dos estudantes. Nosso objetivo principal é estimar o “número médio de desbloqueios de tela de celular por dia”, e sua ajuda é fundamental. Para selecionar os participantes, utilizamos um método de Amostragem Sistemática, e você foi um dos sorteados!

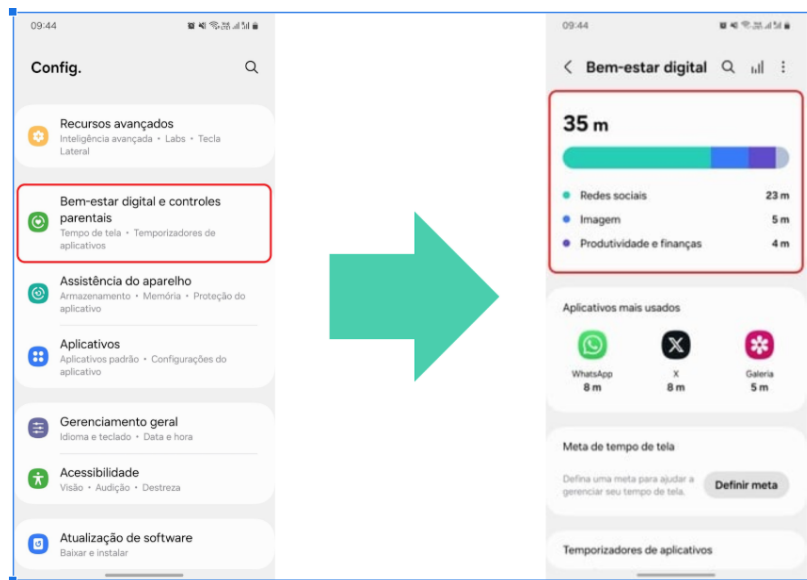
Pedimos que siga as instruções abaixo para responder. É rápido (menos de 2 minutos) e sua resposta é totalmente anônima, sendo crucial para o sucesso da nossa análise.

A participação é voluntária, anônima e todas as informações fornecidas serão tratadas com sigilo, em conformidade com a Lei Geral de Proteção de Dados Pessoais (LGPD) e as diretrizes da UFRGS disponíveis em: <https://www.ufrgs.br/proprivacidade/>.

### 7.2.2 Tutorial

**Como acessar essa informação (Android):**

- 1) Abra as “Configurações” role para baixo e toque no menu de “Bem-estar digital e controles parentais”
- 2) Uma tela com seus dados de uso preliminares será mostrada. Toque sobre a barra para acessar mais informações. (Toque no ícone de gráfico no topo da tela para ver o resumo de uso da última semana.)

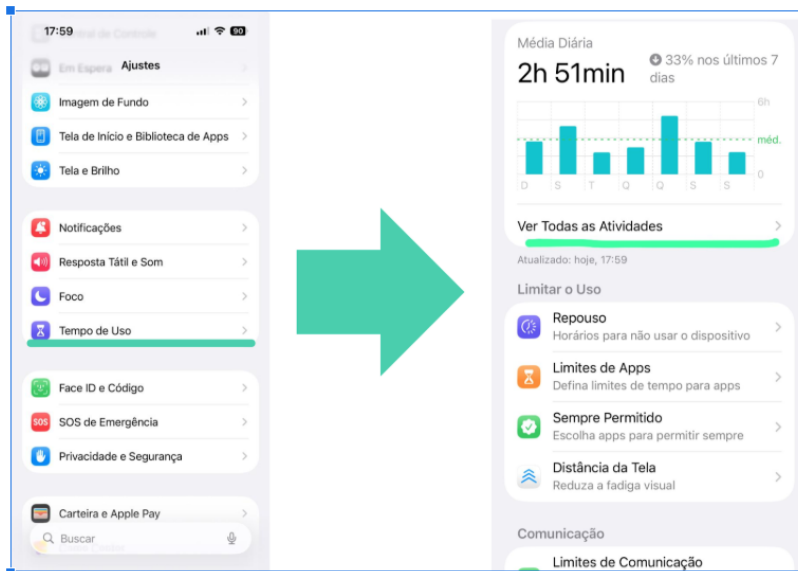


- 3) Na nova tela, arraste para baixo para observar o seu número de ativações/desbloqueios de tela no dia.



### Como acessar informação (IOS)

- 1) Abra o aplicativo Ajustes no seu iPhone. Role para baixo e toque em Tempo de Uso. (Se for a primeira vez que você acessa, talvez precise tocar em Ativar Tempo de Uso.)
- 2) Na tela principal do Tempo de Uso, para um relatório mais detalhado, toque em Ver Todas as Atividades.



- 3) Na nova tela, arraste para baixo para observar o seu número de ativações/desbloqueios de tela no dia.



### 7.2.3 Perguntas

1. Quantas ativações você teve no dia de ontem? (**Coloque apenas o número**)
2. Você considera esse número de desbloqueios/ativações alto?

	1	2	3	4	5	
Muito baixo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito alto

---