

Análise Inferencial de Dados

Construção e Interpretação de Intervalos de Confiança:

O cálculo e a interpretação de intervalos de confiança permitem estimar, com determinado nível de certeza, onde está o valor real de uma métrica, como a média ou a proporção, com base em dados amostrais. Essa abordagem ajuda a avaliar a precisão das estimativas e entender a variabilidade dos dados, fornecendo uma base mais sólida para análises e tomadas de decisão.

Indicadores utilizados para realizar o intervalo de confiança:
SELIC

IPCA

INADIMPLENCIA (pessoas físicas de São Paulo)

```
# i Use `print(n = ...)` to see more
> mean(dadosipca$SELIC)
[1] 0.8541502
> mean(dadosipca$IPCA)
[1] 0.4600791
> mean(dadosipca$Inadimplencia)
[1] 4.024941
> |
```

As médias de cada indicador foram calculadas.

```
> sd(dadosipca$SELIC)
[1] 0.3091318
> sd(dadosipca$IPCA)
[1] 0.3254362
> sd(dadosipca$Inadimplencia)
[1] 0.8893526
> |
```

O desvio padrão foi calculado.

Values	
mean_inad	4.02494071146245
mean_ipca	0.460079051383399
mean_selic	0.854150197628458
sd_inad	0.889352636219338
sd_ipca	0.325436205553117
sd_selic	0.309131799423229

Média e desvio padrão de cada indicador (visualizando os dados)

```
> #verificando
> print(erropadrao_selic)
[1] 0.01943495
> print(erropadrao_ipca)
[1] 0.02046
> print(erropadrao_inad)
[1] 0.05591312
> |
```

```
erropadrao_selic <- sd_selic / sqrt(n)
erropadrao_ipca <- sd_ipca / sqrt(n)
erropadrao_inad <- sd_inad/sqrt(n)
|
#verificando
```

Erro padrão calculado, podemos encontrar o erro dividindo o desvio padrão pela raiz quadrada do tamanho da amostra.

```
#agora vamos escolher
#O valor Z represent

z <- 1.96|

# intervalo de confi
```

Foi usado o valor 1,96 porque estamos trabalhando com o intervalo de confiança de 95%.

```

52
53 #intervalo de confiança para selic
54 ic_selic_inf <- mean_selic - z * erropadrao_selic
55 ic_selic_sup <- mean_selic + z * erropadrao_selic
56
57 #intervalo de confiança para ipca
58 ic_ipca_inf <- mean_ipca - z * erropadrao_ipca
59 ic_ipca_sup <- mean_ipca + z * erropadrao_ipca
60
61 # intervalo de confiança para inadimplencia
62 ic_inad_inf <- mean_inad - z * erropadrao_inad
63 ic_inad_sup <- mean_inad + z * erropadrao_inad
64
65

```

Após esse processo, podemos realizar o cálculo do intervalo de confiança dos indicadores, determinando os limites inferiores e superiores.

```

> cat("IC da SELIC (95%):", ic_selic_inf, "a", ic_selic_sup, "\n")
IC da SELIC (95%): 0.8160577 a 0.8922427
> cat("IC do IPCA (95%):", ic_ipca_inf, "a", ic_ipca_sup, "\n")
IC do IPCA (95%): 0.4199775 a 0.5001806
> cat("IC da Inadimplência (95%):", ic_inad_inf, "a", ic_inad_sup, "\n")
IC da Inadimplência (95%): 3.915351 a 4.13453
> |

```

- **IC da SELIC (95%):** de **0.8161** a **0.8922** no período de 2004 até 2025
- **IC do IPCA (95%):** de **0.4199** a **0.5001** no período de 2004 até 2025
- **IC da Inadimplência de São Paulo para pessoas físicas (95%):** de **3.9154** a **4.1345** no período de 2004 até 2025

Análise sobre o intervalo de confiança:

SELIC: De acordo com os dados apresentados, pode-se concluir que os dados da Selic entre 2004 até 2025 apresentam uma média perto de 0,8541 e uma variação moderada ($CV=36\%$), como indicado pelo desvio padrão, que é igual a 0,3091. A partir da amostra calculada, estima-se, com 95% de confiança, que o intervalo calculado compreenda a média real da população.

IPCA: De acordo com os dados apresentados, pode-se concluir que entre 2004 até 2025 os dados do IPCA apresentam uma média perto de 0,46 e uma variação maior que a da Selic ($CV=70\%$), como indicado pelo desvio padrão de 0,3254. A partir da amostra calculada, estima-se, com 95% de confiança, que o intervalo calculado compreenda a média real da população.

INADIMPLÊNCIA: De acordo com os dados apresentados, pode-se concluir que entre 2004 e 2025 o indicador Inadimplência apresenta uma média perto de 4.024 e uma variação considerada baixa ($CV=22\%$), com o desvio padrão de 0,8893. A partir da amostra calculada, estima-se, com 95% de confiança que o intervalo calculado compreenda a média real da população.

Diante dos dados analisados é possível concluir que o indicador com maior taxa de variabilidade ao longo do período de 2004 a 2025 é o IPCA. Isso acontece, pois, seu coeficiente de variação é mais elevado quando comparado aos outros indicadores, ou seja, isso indica que nesse período analisado esse indicador teve maior instabilidade de seus valores.