

Análise estatística simples

Embora pudéssemos aplicar métodos estatísticos complexos para analisar a probabilidade de tal resultado, há uma simplificação válida e que dá uma probabilidade mínima de tal evento ocorrer.

Vamos considerar que, como não há correlação entre o modelo de urna e o voto do eleitor, em uma dada unidade da federação a chance de um candidato ter mais votos nas urnas antigas é de 50%. Ou seja: em um dado estado, as probabilidades de ele ter mais votos nas antigas ou nas 2020 são iguais.

Qual é então a probabilidade de que o candidato do PT tenha, por puro acaso, tido mais votos nas urnas antigas em 20 de 27 unidades da federação?

A probabilidade de um evento com chance de 50% ocorrer exatamente 20 vezes em 27 tentativas pode ser calculada utilizando a **distribuição binomial**.

1. **Modelo:** Considera-se que as 27 tentativas são independentes e cada uma possui probabilidade $p = 0,5$ de sucesso. Nesse caso, a variável aleatória X , que conta o número de sucessos, segue a distribuição binomial $X \sim \text{Bin}(n = 27, p = 0,5)$.
2. **Fórmula:** A probabilidade de ocorrer exatamente k sucessos em n tentativas é dada por

$$\Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Aplicando os valores $n = 27$, $k = 20$ e $p = 0,5$, obtém-se:

$$\Pr(X = 20) = \binom{27}{20} (0,5)^{20} (0,5)^7 = \binom{27}{20} (0,5)^{27}.$$

3. **Combinações:** O termo $\binom{27}{20}$ representa o número de maneiras de escolher quais 20 tentativas, entre as 27, resultarão em sucesso. Por simetria da fórmula combinatória, tem-se $\binom{27}{20} = \binom{27}{7} = 888\,030$.
4. **Cálculo numérico:**

$$\Pr(X = 20) = 888\,030 \times (0,5)^{27} \approx 0,006616339.$$

5. **Interpretação:** A probabilidade é de aproximadamente **0,6616%**, o que significa que, em média, esse resultado ocorrerá em cerca de 1 a cada 151 conjuntos de 27 tentativas.

Essa análise é uma simplificação e ignora o fato de que as diferenças médias de percentual de votos para o candidato do PT, exibidos no último gráfico, são muito maiores nos 20 casos em que o candidato é favorecido nas urnas antigas do que nos outros 7 casos.

O código para gerar esse relatório diretamente a partir dos dados do TSE está em <https://github.com/sairutsa/analise-eleicoes-2022>