

Estimativa de uma proporção

Eduardo Cunha
Raul Aparício Gonçalves

SITUAÇÃO PROBLEMA | – Sondagem Eleitoral –

As eleições livres são um direito adquirido pelos portugueses após o processo de democratização do nosso país, isto é, após o dia 25 de Abril de 1974. Mas, além de um direito, moralmente é também um dever cívico, embora em Portugal não seja obrigatório.

Chegada a altura das eleições, e com maior destaque nas semanas anteriores ao ato eleitoral, muitas são as sondagens realizadas e publicadas! Vários são os *media* que se associam a instituições e empresas com reconhecimento científico para terem as mais credíveis sondagens. Também os partidos elaboram as suas próprias sondagens.

Nesta atividade vamos, também nós, fazer uma sondagem, não “à boca das urnas” mas após “a contagem dos votos”. Mas para que servirá uma sondagem se já sabemos o resultado eleitoral? O nosso intuito será apenas científico, isto é, estudar o processo associado a uma sondagem: a inferência estatística – estimar uma proporção.

Consideremos, então, os resultados das Eleições Autárquicas de 2017 de uma pequena freguesia transmontana, a União de freguesias de Ala e Vilarinho do Monte. Em 2017 esta freguesia, do concelho de Macedo de Cavaleiros, tinha 665 eleitores recenseados.

O facto de conhecermos o resultado eleitoral será importante para a construção da nossa população e, dessa forma, para a obtenção das amostras e do cálculo das respetivas estatísticas. E, muito importante, para avaliar os resultados de cada sondagem efetuada.

Os dados da votação de 2017 encontram-se no ficheiro **inferencia_sondagem.tns**, que deves abrir e usar nesta atividade, e estão guardados nas variáveis, do tipo lista, **eleitores** e **votacao**.

1. Usando as listas **eleitores** e **votacao** podes obter o parâmetro proporção (percentagem) para cada partido e para os votos brancos, nulos e para a abstenção.

Na folha de cálculo da página 1.3 do ficheiro tns insere as seguintes funções da TI-Nspire CX:

- célula A2: =**dim**(eleitores)
- célula C2: =**dim**(votantes)
- célula D2: =c2/a2*100.0

	A	B	C	D
=				
1	Eleitores...		Votantes...	
2				
3	PSD+CD, PS		PCP+PE...BE	
4				
5				
A2				

Estimativa de uma proporção

Eduardo Cunha
Raul Aparício Gonçalves

Assim, obterás na célula A2 o número de eleitores, na célula C2 o número de votantes e na D2 a percentagem de votantes.

Para obteres o número de votos de cada partido, brancos e nulos, e as respetivas percentagens, vamos inserir fórmulas adequadas nas linhas 4 e 5, respetivamente. Por exemplo, para obter os dados para o partido constante na célula A3:

- célula A4: =countif(votacao,A3) - contará o número de votos igual à célula A3
- célula A5: =A4/C2*100.0 - percentagem de votos do partido respetivo (A3)

Copia para o teu caderno a tabela anterior e os valores calculados.

2. Usa, agora, a página 1.4 do teu ficheiro tns para realizares a tua primeira sondagem eleitoral. Para isso irás de precisar de uma nova função da TI-Nspire CX, a função que permite obter uma amostra aleatória de uma dada dimensão partindo de uma conhecida população:

- **randSamp**(Lista, #Ensaios [,Reposição])

Lista – população, nesta atividade é a lista **eleitores**

#Ensaios – dimensão da amostra, começaremos por dimensão 50

Reposição=0 com reposição (predefinição), eleitores repetidos

Reposição=1 sem reposição, não há repetição de eleitores na amostra.

Assim, na linha de inserção da expressão geradora da lista **amostra** (coluna A da Folha de Cálculo), coloca a expressão:

- **randsamp**(eleitores,50,1)

Obterás na lista **amostra**, coluna A, uma amostra de dimensão 50 da população **eleitores**.

Usando a função **countif()** obtém na coluna C o número de eleitores da amostra que se abstiveram

(na célula C1 inserir “=countif(amostra, B1)), e que votaram em cada um dos 4 partidos que foram a eleição (células C2, C3, C4 e C5). Na coluna D obtém as respetivas percentagens (proporções), sendo que na percentagem de votos de cada partido é apenas considerado os votos efetivos, por isso à dimensão da amostra é necessário retirar as abstenções.

Compara os resultados obtidos na tua primeira sondagem com os respetivos parâmetros. Consideras que esta tua sondagem seria uma boa sondagem? E a dos teus colegas?

Para obteres uma nova sondagem, com uma nova amostra da mesma dimensão, bastará usares o atalho que permite repetir os cálculos, clicando sucessivamente as teclas **ctrl** + **R**.

	A amostra	B	C	D
=				
1		Abstenção		
2		PSD+CDS		
3		PS		
4		PCP+PEV		
5		BE		
A	amostra			

Estimativa de uma proporção

Eduardo Cunha
Raul Aparício Gonçalves

3. Sabe-se que o intervalo de confiança a 95% para a proporção populacional p (parâmetro a estimar) é dado por $\left[\hat{p} - 1,96 \times \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n}}, \hat{p} + 1,96 \times \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n}} \right]$, em que \hat{p} é a proporção da amostra (estimativa) e n a dimensão da amostra.

Determina o intervalo de confiança da tua amostra para a estimativa de cada um dos partidos.

Para tal utiliza novamente as ferramentas da tecnologia TI-Nspire CX, na folha de cálculo da página 1.3 insere na coluna F a fórmula que te permite obter o limite inferior do intervalo de confiança, na coluna G obtém o limite superior.

Por exemplo, para a coligação “PSD+CDS”, cuja estimativa \hat{p} está calculada na célula d2, inserindo as fórmulas:

- célula F2: = d2 – 1.96 * sqrt(d2*(1 – d2) / dim(amostra))
- célula G2: = d2 + 1.96 * sqrt(d2*(1 – d2) / dim(amostra))

Na coluna H deverás obter a amplitude do intervalo de confiança, efetuando a diferença dos seus extremos. Regista o intervalo de confiança que obtiveste e compara-o com o dos teus colegas. Que observas?!

4. Mas afinal o que significará a expressão “um intervalo de confiança a 95% para a proporção populacional p ”? Será uma probabilidade?! Se sim, de quê?!

Vejamos o seu significado recorrendo à realização de várias sondagens partindo de amostras, de uma dada dimensão, geradas aleatoriamente e analisando os respetivos intervalos de confiança.

Para tal, usa a página 2.1 do teu ficheiro tns, onde já se encontra construído um simulador de sondagens para a nossa população de eleitores.

Na folha de cálculo, ecrã ao lado, é gerada automaticamente uma amostra com dimensão indicada na célula B2 (está 50, mas pode ser alterada), e obtida uma estimativa (na célula B6) para a proporção de votos do partido indicado na célula B4. Nas colunas C e D encontram-se os vários intervalos de confiança das sucessivas “sondagens” eleitorais.

Clica, sucessivamente, nas teclas **ctrl** + **R** para obteres mais amostras e respetivos intervalos de confiança. Em conjunto com os teus colegas, obtém um grande número de experiências, superior a 100. Qual a percentagem de intervalos de confiança que contêm o parâmetro que te encontras a estimar?! Que traduzirá esse valor?