

Indícios de Inconsistência nas Pesquisas Eleitorais

Argumentos Estatísticos

Luiz Carlos da Rocha
Pres. Comissão de Fiscalização do CONFE
Rio de Janeiro, 24/06/2022

Regra de Inconsistência

Se duas pesquisas eleitorais, ambas com confiabilidade de 95% e margens de erro $\epsilon_1\%$ e $\epsilon_2\%$ respectivamente divulgarem prévias eleitorais de um mesmo candidato, com entrevistas realizadas numa mesma semana e na mesma UF, inclusive Brasil, com diferença de intenção de votos superior à soma $\epsilon_1\% + \epsilon_2\%$, então há indício de inconsistência.

Argumentação estatística

Sejam as prévias percentuais (%) divulgadas por duas pesquisas eleitorais sujeitas às condições de Confronto:
i) mesmo cargo eletivo numa mesma área de pesquisa (inclusive Brasil); ii) mesma lista de candidatos ao cargo eletivo; iii) eleitores consultados numa mesma semana ou em períodos contíguos. Considerem-se ainda os dados divulgados juntamente com os resultados das pesquisas; os tamanhos das amostras n_1 e n_2 e as confiabilidades dadas por: nível de confiança 95% e margens de erro $\epsilon_1\%$ e $\epsilon_2\%$.

Na argumentação que se segue, a prévia porcentual (%) é dividida por 100, denominada de razão de votos, significando o total da intenção de votos declarada na pesquisa dividida pelo número de eleitores entrevistados, a margem de erro também é dividida por 100.

Para cada candidato, em qualquer das duas listas, a razão de votos β da pesquisa se comporta com distribuição Normal de média μ , onde μ significa o verdadeiro valor da razão de votos do candidato na área pesquisada. De acordo com as três condições de Confronto, o valor μ é o mesmo nas duas pesquisas, isto é, os resultados das pesquisas traduzem medidas estatísticas independentes de uma mesma situação. A variância de votos β depende do próprio μ e do tamanho da amostra da pesquisa, sua expressão é $\mu\sigma/n$.

Considere um candidato qualquer C da lista de nomes apresentada aos entrevistados e designe-se por β_{c1} e β_{c2} as variáveis que traduzem a razão de votos do candidato em cada pesquisa. De acordo com a teoria estatística e com as hipóteses de Confronto, as variáveis se comportam com distribuição Normal independentes de parâmetros; $E[\beta_{c1}] = E[\beta_{c2}] = \mu$, $V[\beta_{c1}] = \mu\sigma/n_1$, $V[\beta_{c2}] = \mu\sigma/n_2$, logo a diferença $\beta_{c1} - \beta_{c2}$ também terá comportamento Normal com $E[\beta_{c1} - \beta_{c2}] = 0$ e $V[\beta_{c1} - \beta_{c2}] = \mu\sigma/n_1 + \mu\sigma/n_2$. Então;

$$\text{Prob}\{\beta_{c1} - \beta_{c2} > (\epsilon_1 + \epsilon_2)\} = \text{Prob}\{N(0, [\mu\sigma/n_1 + \mu\sigma/n_2]^{1/2}) > (\epsilon_1 + \epsilon_2)\} = \text{Prob}\{N(0,1) > (\epsilon_1 + \epsilon_2) / (\mu\sigma/n_1 + \mu\sigma/n_2)^{1/2}\}$$

Substituindo os tamanhos das amostras n_1 e n_2 na expressão da **Prob** pela relação $n = \mu\sigma [1,96/\epsilon]^2$, obtém-se o resultado exclusivamente em função dos parâmetros de confiabilidade das pesquisas;

$$\text{Prob}\{\beta_{c1} - \beta_{c2} > \epsilon_1 + \epsilon_2\} = \text{Prob}\{N(0,1) > 1,96(\epsilon_1 + \epsilon_2) / (\epsilon_1^2 + \epsilon_2^2)^{1/2}\} \quad (*)$$

A **Prob**. da expressão acima quando calculada com valores das margens praticadas pelas empresas de pesquisa (2,0% a 3,5%) assume valor aproximado e constante igual a 0,3%. Porém, a diferença entre as variáveis pode assumir valor positivo ou negativo, dependendo da ordem das parcelas, ou seja, do critério adotado para designar a primeira e a segunda parcela. No caso da aplicação da Regra de Inconsistência, subentende-se que a diferença é igual “a maior prévia menos a menor”, logo será sempre positiva. Tal critério implica que a probabilidade final da regra será igual ao dobro do valor calculado pela expressão (*), ou seja, igual a 0,6%. Portanto, a desigualdade $\beta_{c1} - \beta_{c2} > \epsilon_1 + \epsilon_2$ é esperada ocorrer aleatoriamente somente 6 vezes dentre 1.000 pesquisas.

Logo, se duas pesquisas atenderem as condições de Confronto e a diferença entre as respectivas prévias para um mesmo candidato C for superior à soma das suas margens de erro $\epsilon_1 + \epsilon_2$, então será levantada a hipótese de Inconsistência, pois provavelmente há discrepâncias entre as metodologias ou nos procedimentos operacionais de levantamento das informações.