**Universidade de São Paulo**

**Faculdade de Filosofia, Letras e Ciências Humanas**

**Departamento de Ciência Política**

**Métodos Quantitativos e Técnicas em Ciência Política**

1º semestre / 2024

**Lab 7. Inferência Estatística**

**Questão 1**

Em Junho/2014, um jornal de grande circulação reportou os resultados de uma pesquisa de opinião que perguntava: “De tudo que você sabe sobre a presidenta Dilma Rousseff, ela possui a honestidade e integridade que você espera em um presidente?”. A pesquisa entrevistou 518 adultos e 233 responderam “sim”.

1. Encontre uma estimativa do ponto para a porcentagem de adultos que julgam que a presidenta possui honestidade e integridade que se espera de um presidente. Encontre o erro padrão para essa média.
2. Para $α=5\%$, estabeleça um intervalo de confiança para a estimativa indicada. O que aconteceria com o intervalo estimado se o utilizássemos $α=1\%$ ? E se a amostra aumentasse de tamanho?

**Questão 2**

A publicação, em 2006, *Attitudes towards European Union Elargement* dos estados do Eurobarometer declara:

“Os leitores são lembrados de que os resultados do levantamento de dados são *estimativas*, a precisão delas depende do tamanho da amostra e do percentual observado. Com amostras de aproximadamente 1.000 entrevistas, os percentuais reais variam dentro dos seguintes limites de confiança:



1. Explique como eles obtiveram 3,0 pontos como limite para 40% ou 60% de média observada;
2. Explique por que a margem de erro (limites) difere para os percentuais observados.
3. Explique por que a precisão é a mesma para um percentual observado em particular e para 100 menos aquele valor (por exemplo, para 40% e 60%);
4. Explique por que é mais difícil estimar a proporção populacional quando ela está próxima a 50% do que quando ela está próxima a 0 ou a 100%.
5. Nas pesquisas de opinião no Brasil, como as de intenção de votos, os resultados são divulgados como se os percentuais expressos estivessem associados a uma mesma margem de erro (por exemplo, 2 pontos para mais ou para menos). Tal informação está correta? Explique.

**Questão 3**

Vamos supor que queremos descobrir a média de uma população qualquer e assumimos que este valor seja igual a $μ\_{0}$. Em nosso processo de inferência, aleatoriamente, selecionamos uma amostra de tamanho *n* e, em nossa amostra, encontramos como valor da média amostral o resultado igual a $\overbar{y\_{0}}$.

1. Qual a probabilidade de encontrarmos exatamente o valor $μ\_{0}$ em nossa amostra? Por que isso ocorre?
2. Quais podem ser as causas da diferença entre $μ\_{0}$ e $\overbar{y\_{0}}$?

Vamos aceitar o pressuposto de que todo o procedimento amostral foi correto. Estamos, então, diante da seguinte questão: há razões para que a diferença entre $μ\_{0}$ e $\overbar{y\_{0}}$ seja tal que nos leve a acreditar que o parâmetro populacional não seja mais $μ\_{0}$? Em outras palavras, o que aprendemos com essa amostra que colhemos a respeito do parâmetro populacional?

Para respondermos a essas perguntas, realizaremos um teste de hipóteses. Para isto, é necessário estabelecermos um critério de comparação do valor encontrado em nossa amostra com um valor que acreditamos ser o correto para o parâmetro populacional. As razões pelas quais assumimos que o parâmetro populacional é um valor e não outro depende da conveniência do trabalho em questão. No caso, vamos utilizar a informação de referência $μ\_{0}$. Estabelecemos assim, em geral, um par de hipóteses que segue o seguinte formato:

$$H\_{0}: μ=μ\_{0}$$

$$H\_{A}: μ\ne μ\_{0}$$

Ao fazermos um teste de hipóteses, há três formas de proceder que nos conduzem ao mesmo resultado. Vamos considerar estas três abordagens.

A primeira delas está em calcular o valor do escore-Z (ou *Z-score*) para a estimativa amostral.

1. Apresente a fórmula do *Z-score* e dê um significado para seu resultado;

A partir deste resultado, temos condições de nos posicionar diante das hipóteses enunciadas acima.

1. Para um determinado nível de significância α, qual a regra adotamos para o teste de hipóteses?

O segundo procedimento possível é o de estabelecermos um intervalo a partir do qual aceitamos a variação em torno da média populacional.

1. Em decorrência do que aceitamos a existência desta variação?
2. Como construímos este intervalo? E como tomamos uma decisão frente ao par de hipóteses acima indicado?

Por fim, temos um terceiro procedimento. Note que nos dois anteriores, de maneira mais ou menos direta, dependemos de um valor para um nível de significância α. Não conseguimos nos livrar dele, mas podemos evidenciar quão próximo estamos desse número arbitrário. Para isso, calculamos o *p-valor*.

1. O que significa o *p-valor*?
2. Como utilizamos o seu resultado para nos posicionarmos diante do par de hipóteses enunciadas?

