**Universidade de São Paulo**

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

**Departamento de Ciência Política**

FLS 5028 – Métodos Quantitativos e Técnicas de Pesquisa em Ciência Política

FLP 0406 – Métodos e Técnicas de Pesquisa em Ciência Política

1º Semestre de 2024

Prof. Dr. Glauco Peres da Silva

Laboratório 2 - Probabilidades

No laboratório de hoje, focaremos em aprender diferentes formas de pensar como a probabilidade pode ser medida. São formas diferentes que, mesmo respeitando a Teoria da Probabilidade, criam processos distintos de pensar as chances de ocorrência de um evento. Os exercícios que estão em destaque devem ser feitos apenas se houver tempo para tal.

**1ª abordagem**

Para a primeira dessas situações, considere o lançamento de um dado. Vamos a partir dele criar a seguinte variável .

1. Qual o espaço amostral desta variável?
2. Esboce um histograma para a variável *X*.
3. Identifique e .
4. Discuta com seus colegas de grupo qual o procedimento adotado para encontrar as probabilidades do item anterior. Qual a relação entre as considerações feitas para a resposta ao item c) e o real lançamento de um dado?

\*\*\*

Um armário contém 6 pares de meias pretas, 8 pares de meias azuis, 5 pares de meias verdes e um par de meia listradas. Dois dos pares de meias pretas e um par de meia azul estão furadas. Qual a probabilidade de que:

a) Uma única meia escolhida ao acaso seja azul?

b) Uma única meia escolhida ao acaso seja branca ou preta?

c) Um par de meia escolhido ao acaso tenha exatamente uma meia verde e uma meia com furo?

\*\*\*

**2ª abordagem**

Considere a seguinte situação real: foram realizados 2715 sorteios da Mega Sena organizada pela Caixa Econômica Federal. Abaixo está um gráfico que mostra a quantidade de vezes nestes sorteios em que cada um dos 60 números foi selecionado.

O número 10 foi sorteado 318 vezes e o 53 apareceu 311, sendo os dois mais vezes selecionados. Já o 21 o 26 e o 55, aparecerem, respectivamente, em 227, 228 e 230 sorteios apenas.

1. Determine e .
2. Qual a diferença fundamental entre a probabilidade calculada agora com a dos dados no exercício anterior?
3. Vendo a variação dos números tal como acontece no gráfico acima,

\*\*\*

Qual seria a sua estratégia numa eventual aposta para o próximo sorteio da Mega Sena: você escolheria os números que saíram mais vezes (já indico, na ordem: 10, 53, 5, 23, 34 e 37), nos que saíram menos vezes (aqui estão, na ordem: 21, 26, 55, 15, 22 e 48) ou adotaria ainda outra estratégia?

\*\*\*

**3ª abordagem**

Por fim, a terceira forma de identificar a probabilidade pode ser entendida a partir do cálculo de chances da vitória de um time de futebol em uma partida qualquer. Assuma que ocorrerá no próximo final de semana um jogo entre o time A e o time B. Nossas dúvidas estão sobre como pensamos as probabilidades dos resultados deste jogo.

1. Se considerássemos os três resultados possíveis, qual a chance de vitória do time B?
2. Se historicamente sabemos que times que possuem o mando de jogo vencem em 50% das vezes, que são derrotados em torno de 22% das vezes, e que o mando de jogo é do time A, como isso mudaria a sua resposta no item anterior?
3. Como usar essas informações, se você sabe que o melhor jogador do time B estará fora e, além do mais, o time está em uma fase ruim, com o seu técnico à beira da demissão?

\*\*\*\*\*\*\*\*

A primeira abordagem é chamada de *frequentista*. Avaliamos a probabilidade de um evento qualquer ocorrer a partir da sua frequência a partir do espaço amostral considerado. Assim, temos que:

É importante destacar que não foi necessário observar *X*. A probabilidade é calculada a partir da posição relativa de cada um dos valores de X dentro do espaço amostral. Assim, se um valor específico ocorre com frequência igual a ¼, dizemos que .

A segunda abordagem é chamada de verossimilhança. Neste caso, a probabilidade de um evento qualquer também é a frequência relativa do evento. Mas aqui ela não é esperada, mas sim observada. Ou seja, o evento que produz X foi observado uma quantidade n de vezes e partir dele, se extrai a frequência relativa de X. Assim:

A probabilidade é estimada a partir de um processo chamado de *máximo-verossimilhança*. Ele depende do que foi observado. Por isso, dizemos que há um parâmetro que condiciona as ocorrências de X e acaba por interferir na frequência de ocorrência de cada um de seus valores.

Por fim, a terceira abordagem é a bayesiana. Sua lógica é a inversa da anterior. Suponha que existam determinadas probabilidades de que X assuma cada um de seus valores no espaço amostral. Os dados que coletamos a respeito do comportamento de X nos dão informações a respeito daquelas probabilidades. Neste caso, temos:

O que estamos estressando aqui é o fato de que os eventos observados nos informam sobre algo desconhecido, que é a probabilidade de X assumir um valor específico. É como se não tivéssemos, necessariamente, informações suficientes para encontrar o real valor, e a partir das observações que fazemos nos aproximamos dela. Se retomamos o exemplo acima, é como se reconhecêssemos que existe a probabilidade do time A vencer, mas não teremos observações suficientes para encontrar isso, já que A e B não se enfrentarão nas mesmas condições número elevado de vezes. Lançamos mão de outras informações que nos ajudam a encontrar esse valor.