Revisão de Inferência Estatística

Parte 1

Você quer estudar uma População?

Qual População?

Eleitores

Lâmpadas

Queijos

Azeite de Oliva

Pais de alunos do Ensino Público Básico

Pacientes internados

Crianças

Qual característica dessa População você quer estudar?

Lâmpadas:
tempo de vida útil?

Eleitores:
em quem vão votar?

Pais de alunos do Ensino Público Básico: qual sua opinião sobre a volta às aulas?

Crianças: quantas tiveram sarampo na infância?

Poderíamos testar todas as lâmpadas para determinarmos sua vida útil?

Conseguiríamos entrar em contado com os 100 milhões de eleitores brasileiros para saber em quem vão votar?

É possível perguntar a opinião de todos os pais?

Então surge naturalmente a ideia de Amostra: um subconjunto da população.

Pegaríamos um lote de determinada Lâmpada.

Amostra é formada por n indivíduos da População

Os Institutos de Pesquisas consultam em torno de 3000 mil pessoas, por pesquisa, na época de Eleições.

Amostra precisa ser formada por:

- indivíduos independentes que não se relacionam

Consultar um grupo de pais para inferir a opinião da maioria.

Como modelamos matematicamente/estatisticamente essas situações?

Queremos criar uma modelagem que seja a mais geral possível. Que possamos usar para qualquer amostra.

O conceito de Variável Aleatória vai nos ajudar muito nessa tarefa.

Mas o que é uma Variável Aleatória?

Variável Aleatória?

Sim! É uma Função com um Domínio bem particular.

Seu Domínio é Espaço Amostral.

Espaço Amostral

Definição 1.1.1 *(Espaço Amostral) Vamos chamar de Espaço Amostral o conjunto de todos os possíveis resultados de um experimento aleatório.*

Usaremos a letra grega Ω *como notação para o Espaço Amostral.*

Quando falamos em *Experimento Aleatório* estamos nos referindo a um experimento que utiliza um mecanismo que gera resultados aleatórios. É um dos conceitos mais básicos em probabilidade e estatística. Entre esses experimentos, os mais comuns e conhecidos são bingo, loteria, lançamento de dados, lançamento de uma moeda.

Variável Aleatória

Definição 1.1.1 *Uma variável aleatória é uma função que associa um número real a cada elemento de* Ω.

$$X : \Omega \to \mathbb{R}.$$

Associamos uma variável aleatória X_i a cada elemento da amostra de tamanho n:

$$X_1, X_2, \ldots, X_n$$

Pais de alunos do Ensino Público Básico

Qual sua opinião sobre a volta às aulas?

- Há quase 3 milhões de alunos nas Escolas Estaduais do Estado de São Paulo.
- Digamos que sejam consultados 5000 pais .

Associamos uma variável aleatória X_i a cada pai da amostra de pais de tamanho $n = 5000$:

$$X_1, X_2, \ldots, X_{5000}$$

Uma variável aleatória X com distribuição de probabilidade de Bernoulli é usada para modelar experimentos aleatórios com respostas "sim"e "não".

Ela Atribui valor 1 ao "sim"e 0 ao "não", com probabilidades p e $1-p$, respectivamente.

Por exemplo, podemos considerar o lançamento de uma moeda, no qual sai cara com probabilidade p e sai coroa com probabilidade $1-p$.

Então seja X uma v.a de Bernoulli, temos que X é igual a 1 se sai cara e igual a 0 se sai coroa:

$$X = \begin{cases} 1 & \text{se sai cara} \\ 0 & \text{se sai coroa.} \end{cases}$$

Vamos considerar que os pais (pai e mãe de um aluno) respondam "sim" com probabilidade p desconhecida.

p corresponde à proporção de pais que concordam com a volta às aulas presenciais.

Não sabemos p. Como estimá-lo?

Podemos somar os valores coletados (1´s e 0´s), e dividir por 5000, ou seja, calcular a média aritmética.

Deve ser observado que o valor dessa média depende da amostra de pais sorteada. Essa média é chamada de Média Amostral.

Esse método chama-se Estimação Pontual.

A probabilidade com que uma variável aleatória discreta, X, assume cada um dos seus valores é chamada de Distribuição de Probabilidade de X.

Voltando ao exemplo de variável aleatória de Bernoulli ,

Por exemplo, podemos considerar o lançamento de uma moeda, no qual sai cara com probabilidade p e sai coroa com probabilidade $1 - p$.

Então seja X uma v.a de Bernoulli, temos que X é igual a 1 se sai cara e igual a 0 se sai coroa:

$$X = \begin{cases} 1 & \text{se sai cara} \\ 0 & \text{se sai coroa.} \end{cases}$$

Temos que P(X=1)=p e P(X=0)=1-p, que é a distribuição de probabilidade de X.

$P(X=1)=p$ e $P(X=0)=1-p$

Apareceu algo novo aqui! P(...)?

Probabilidade é sempre aplicada a Eventos do Espaço Amostral

Eventos?

São subconjuntos do Espaço Amostral.

$P(X=1)=p$ e $P(X=0)=1-p$

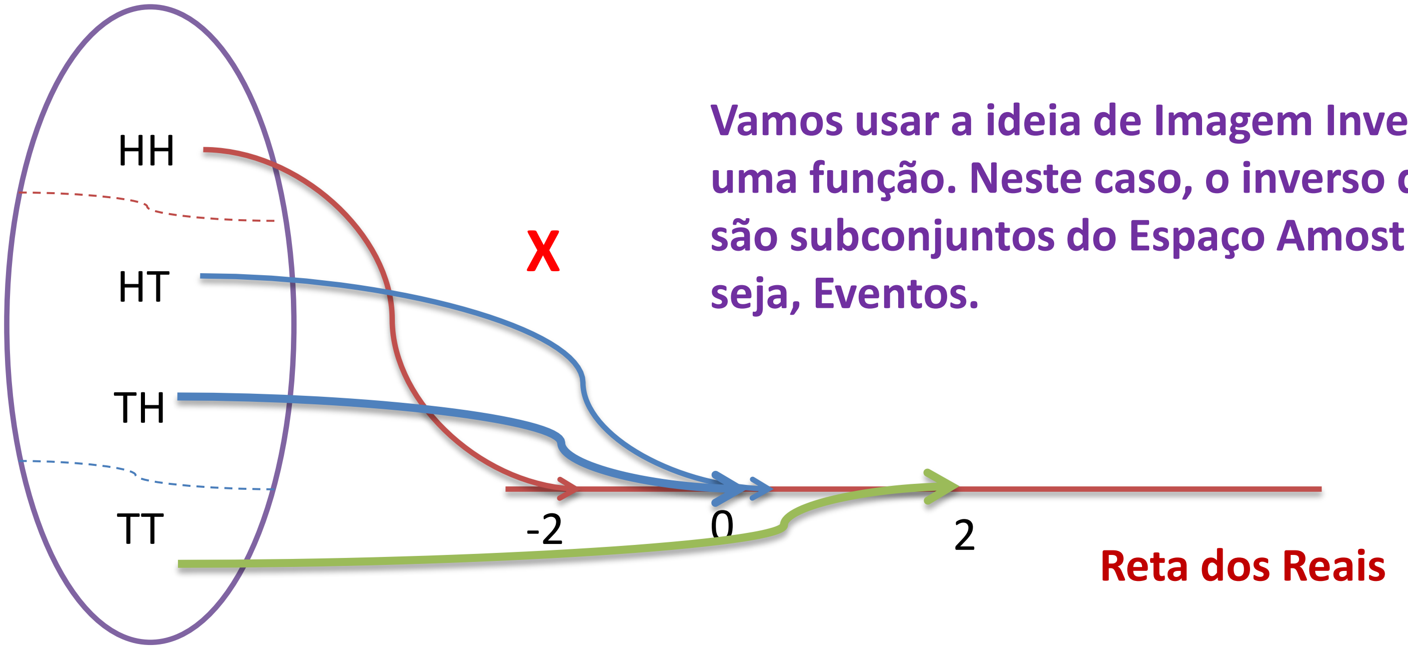
Mas o que significa $P(X=1)$? Onde está o evento?

$P(X=1)$ é uma notação para $P(\{X=1\})$.

E o que significa $\{X=1\}$?

Agora temos que conhecer melhor o Conceito de Variável Aleatória.

- X assume -2 se saírem duas caras: X(HH)=-2
- X assume 0 se sair apenas 1 cara: X(HT)=X(TH)= 0
- X assume 2 se não saírem caras: X(TT)= 2



Vamos usar a ideia de Imagem Inversa de uma função. Neste caso, o inverso de X são subconjuntos do Espaço Amostral, ou seja, Eventos.

Espaço Amostral

{X=-2} é o evento que contém os elementos do Espaço Amostral que foram levados em "-2" pela X. E temos o mesmo para {X=0} e {X=2}.

{X=-2} ={HH},   {X=0}={HT, TH}   e   {X=2}={TT}

Se a moeda for honesta, podemos calcular a distribuição de Probabilidade de X:

- **P(X=-2) =P({X=-2})=P({HH})=1/4**
- **P(X=0) =P({X=0})=P({HT,TH})=1/2**
- **P(X=2) =P({X=2})=P({TT})=1/4**

Quer saber mais sobre Variáveis Aleatórias? Clique no link abaixo

https://drive.google.com/file/d/1SyJ6WEGQY_jJFG4wPauj90hk1b6qCA9i/view?usp=sharing

Vai nos ser extremamente útil a Esperança de uma variável aleatória

A distribuição de massa de probabilidade de uma variável aleatória X nos fornece uma série de números, os quais são as probabilidades de ocorrência de cada um dos possíveis valores de X. Seria interessante resumir essa informação em um único número representativo. Para encontrar esse número introduzimos o conceito de esperança de X, valor esperado ou média. A esperança de X, é a média ponderada dos valores assumidos por X, onde a ponderação dos valores é feita pelos $p_X(x)'s$.

Definição 1.3.1 *(Esperança de uma v.a X ou seu valor esperado, ou sua média).*

O valor esperado de uma variável aleatória X, com distribuição de massa de probabilidade p_X, é definido por:

$$\mathbb{E}(X) = \sum_x x p_X(x).$$

A Esperança da Variável Aleatória do exemplo anterior é calculada como segue abaixo:

E(X)= (-2). P(X=-2) + (0) .P(X=0)+ (2). P(X=2)=0

Também será importante sabermos calcular a Variância de uma variável aleatória

Um outro importante número associado à variável aleatória X é a variância, que é denotada por $Var(X)$ e é definida como valor esperado da v.a $(X - \mathbb{E}(X))^2$, isto é,

$$Var(X) = \mathbb{E}[(X - \mathbb{E}(X))^2].$$

Como $(X - \mathbb{E}(X))^2$ só assume valores positivos, observamos que $Var(X)$ é sempre um número positivo.

Que pode ser mais facilmente calculada como: $Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2.$

É mais razoável usar o desvio padrão:

$$\sigma_X = \sqrt{Var(X)}.$$

O desvio-padrão é muitas vezes mais fácil de ser interpretrado, porque ele apresenta as mesmas unidades de X. Por exemplo, se a unidade de medida de X for metro, então a unidade do desvio-padrão também será metro, enquanto que a unidade de medida da variância será metro quadrado.

Quer saber mais sobre Esperança e Variancia? https://drive.google.com/file/d/1rjOFQ0SyzwfAmo-_ReXlXn2amVGvfRAE/view?usp=sharing

No caso de X ter distribuição de Bernoulli,

P(X=1)=p e P(X=0)=1-p

Dizemos que p é o parâmetro desconhecido da distribuição de probabilidade de Bernoulli, que está relacionado com sua Esperança:

E(X)=p.

Por isso, a média amostral é um estimador natural para p.

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}.$$

Quer saber mais? Clique em:

https://drive.google.com/file/d/14_UY61vQ4cgCZutysI21acj5CEPDTupQ/view?usp=sharing

População

Característica de Interesse:
X com distribuição de probabilidade conhecida, mas com parâmetro(s) desconhecido(s)

Como estimar esse parâmetro?

- Tomamos uma Amostra da População.
- Usamos uma variável aleatória chamada estimador, que é função da amostra.

O Estimador fornece uma Estimativa para o parâmetro desconhecido.

Podemos usar outros métodos para nos aproximarmos do valor verdadeiro de p:

- **Construir um intervalo de Confiança**.

Margem de erro: conceito relacionado com Intervalo de Confiança. Ver link abaixo:
https://brasilescola.uol.com.br/matematica/o-que-margem-erro-uma-pesquisa.htm

- **Usar teste de Hipóteses para p.**