

## 1. APRESENTAÇÃO E INTRODUÇÃO À ESTATÍSTICA

### Estatística Básica – algumas considerações:

- A estatística está associada na elaboração, análise e comunicação de todo tipo de pesquisa científica de qualquer área.
- Também é muito utilizada na divulgação destas políticas junto à sociedade.
- É utilizada de forma intencional ou não para justificar ou induzir a conclusões muitas vezes equivocadas.
- Uma sólida compreensão do método estatístico, suas características e principalmente suas limitações, é fundamental para uma análise crítica responsável da massa de informações dessa natureza existente no mundo atual.

### b) Apresentação do Plano de Ensino

#### b.1. Objetivos do curso:

- b.1.1. Estatística como ferramenta de trabalho;
- b.1.2. Conhecimento básico das fases da análise estatística;
- b.1.3. Compreender o processo de construção de indicadores;
- b.1.4. Conhecer as ferramentas essenciais à análise estatística (pesquisa);

#### b.2. Estrutura do curso: aulas expositivas e aulas práticas (tutoriais e exercícios);

- b.2.1. Privilegiar conhecimento conceitual com cálculos em Planilha Eletrônica;
- b.2.2. Abordagem de interpretação de conceitos e uso de ferramentas matemáticas;

#### b.3. Bibliografia: indicada no plano de ensino e este Resumo.

### c) Introdução à Estatística

#### c.1. Importância do estudo da Estatística na formação profissional:

- c.1.1. O raciocínio estatístico é muito utilizado no governo e na administração;
- c.1.2. O conhecimento estatístico serve para tomar decisões fundamentadas;
- c.1.3. Utilização em outras disciplinas do curso;
- c.1.4. Interpretação de resultados publicados em jornais, revistas profissionais e artigos científico.
- c.1.5. Usar a interpretação estatística nos artigos da imprensa e no cotidiano.
- c.1.6. Exemplos da “estatística” na mídia em geral

##### c.1.6.1. Aumento de 100% no consumo de frutas por crianças:

- como foi feita a pesquisa – onde foi feita – quais são os números brutos

##### c.1.6.2. Pesquisa comprova efeitos benéficos do vinho – contestações

- variáveis de confusão (o vinho, o estilo de vida, a alimentação, etc.)

#### c.2. Técnicas e métodos envolvendo planejamento, coleta de dados, inferência

(previsões), processamento, análise e disseminação das informações obtidas;

#### c.3. Técnicas para abordarmos cientificamente situações de incerteza – busca evidências quantitativas de fenômenos em geral;

#### c.4. Cuidados: “se torturarmos apropriadamente os números, eles confessam qualquer coisa” – “Como dizer mentiras usando a Estatística”;

- c.4.1. Importância de conhecimento da Estatística para o correto entendimento do significado e limitações das “estatísticas”;

#### c.5. Desenvolvimento de raciocínio abstrato, avaliação crítica, associação, dedução e síntese; domínio de ferramentas básicas de computação (Excel / outros);

- c.5.1. Os cálculos serão implementados no Microsoft Excel.

- c.5.1.1. Softwares apropriados fazem o trabalho pesado.
- c.5.2. O objetivo é chegar a uma compreensão suficiente para entender cada uma das etapas do método estatístico, o processo como um todo, suas limitações, e possibilitar a comunicação com um profissional da Estatística.
- c.6. A Estatística estuda fenômenos com um conjunto de muitos indivíduos, com pelo menos uma característica comum;
  - c.6.1. “Coleção de métodos para planejar, obter e organizar dados, resumi-los, interpretá-los e deles extrair conclusões”.
- c.7. Definição, segundo Antonio A. Crespo: “Estatística é uma parte da matemática aplicada que fornece métodos para a coleta, a organização, a descrição, a análise e a interpretação de dados quantitativos e qualitativos, e a utilização desses dados para a tomada de decisão.”
- c.8. Outras definições: Uma técnica para:
  - c.8.1. lidar com a coleta, análise, interpretação e apresentação de dados;
  - c.8.2. extrair significados a partir de dados;
  - c.8.3. lidar com as incertezas e planejar o futuro com segurança razoável;
  - c.8.4. para fazer inferências sobre o desconhecido;
  - c.8.5. É a arte de aprender a partir de dados (Ross, 2010).

#### **d) Estatística Aplicada**

- d.1. Termo “estatística” vem de “status”, em latim “Estado”, por ser feito pelos governantes para auxiliar a condução do Estado;
- d.2. Antiguidade: registro de número de habitantes, de nascimentos, de óbitos, estimativas de riqueza individual e social; distribuição de terras ao povo, cobrança de impostos, levantamentos em geral (censo – estatística);
- d.3. Idade média: registro de dados referentes à população (batizados, casamentos, funerais), ao comércio, produção de bens e alimentos (em 1085, Guilherme da Inglaterra realizou levantamento de informações sobre proprietários de terras, empregados e animais com objetivo de cobrança de impostos);
- d.4. Renascimento: desenvolvimento das ciências em geral, incluindo a Matemática. No século XVIII tornou-se campo específico do conhecimento sob o nome de Estatística (por Godofredo Achenwall) – desenvolvimento dos cálculos de probabilidades;
- d.5. Tecnologia: desenvolvimento de processos e técnicas avançadas e utilização de computadores e supercomputadores;
- d.6. Exemplos atuais: situação social de comunidades em geral, censo IBGE, pesquisa eleitoral, caracterização geográfica de epidemias, etc. Os resultados são utilizados como subsídios para elaboração de políticas e intervenções do Estado
- d.7. Pesquisa científica: investigação de fenômenos em geral, buscando tendências reveladas pelos dados obtidos.

#### **e) Censo / Amostragem**

- e.1. Censo: quando todos os indivíduos da população são estudados;
- e.2. Amostragem: quando parte da população é estudada.

## 2. O MÉTODO ESTATÍSTICO, SUAS CARACTERÍSTICAS E LIMITAÇÕES

### 1. O Método Estatístico

- 1.1. Refere-se a todo o processo de utilizar o método científico para responder questões e tomar decisões.
- 1.2. Envolve projetar estudos, coletar dados “bons”, descrever os dados com números e gráficos, analisar os dados, elaborar conclusões, eventualmente tomar decisões.
- 1.3. “Deixar os dados falarem por si próprios” – “Ouvir o que os dados relatam”
  - 1.3.1. Mas os dados só falam claramente quando são organizados, resumidos, apresentados, e os auxiliamos falar propondo questões apropriadas.

### 2. Etapas de um Estudo

- 2.1. Definição do problema (formulação, meta, revisão literatura, tipo de estudo)
- 2.2. Planejamento (o que, onde, como, quando, quem, quanto)
- 2.3. Coleta / tabulação dos dados
- 2.4. Descrição dos dados (resumo e apresentação – tabelas e gráficos)
- 2.5. Análise dos dados (técnicas de análise estatística)
- 2.6. Interpretação dos dados (extrair conclusões)

### 3. Tipos de Estudo

- 3.1. Observacional: os dados são coletados de indivíduos (pessoas ou coisas) de maneira tal que não os afeta. Por exemplo, questionários apresentados a indivíduos selecionados “de forma amostral” de uma população de interesse.
  - 3.1.1. Podem ser conduzidos de diversas formas: e-mail, web sites, telefone, etc.
  - 3.1.2. Se conduzidos não apropriadamente, podem resultar em informação “enviesada” (tendenciosa). Por exemplo, indução por escolha não apropriada das palavras ou formulação nas questões.
  - 3.1.3. Podem relatar relações entre as variáveis encontradas. Não podem estabelecer relações de causa e efeito. Por exemplo, pessoas que relatam tomar mais de duas latas de coca-cola no dia por relatam dormir menos horas por noite. Não se pode concluir que é a coca-cola que faz perder o sono. Outras variáveis não estudadas podem ser responsáveis pelo efeito (festas).
- 3.2. Experimento: impõe um ou mais tratamentos aos indivíduos (pessoas ou coisas), de tal modo que uma clara comparação pode ser feita (entre o antes e o depois do tratamento). Por exemplo, um estudo sobre o efeito da dosagem de um anti-hipertensivo na pressão sanguínea: um grupo recebe dose de 10 mg, outro recebe dose de 20 mg, outro 30 mg, etc. Um grupo “de controle” recebe placebo.
  - 3.2.1. Devem ser conduzidos em ambiente controlado e projetados para minimizar vieses que possam ocorrer.
  - 3.2.2. Possíveis problemas: pesquisadores sabem quem se submeteu a cada tratamento, alguma característica não levada em conta pode afetar o resultado (peso dos indivíduos quando estudando a dose do medicamento).
  - 3.2.3. Quando bem projetado e conduzido cuidadosamente, podem-se estabelecer relações de causa e efeito;
- 3.3. Dados publicados: estudos sobre dados já publicados, disponíveis em diversas fontes de dados, por exemplo, IBGE; INEP, etc.

#### 4. Coleta de Dados

- 4.1. Tendo definido os objetivos do estudo, escolhido as variáveis a serem analisadas, definido o tipo de estudo (observacional ou experimental), chega-se à escolha da amostra – processo de amostragem.
- 4.2. Selecionando uma boa amostra: “Entra lixo, sai lixo” – se a amostra é tendenciosa, os resultados também serão tendenciosos. É necessário garantir a “aleatoriedade da seleção” dos indivíduos (pessoas ou coisas) que comporão a amostra.
  - 4.2.1. Pesquisas em web sites, em programas de TV (ligue e dê sua opinião) resultam de amostra “auto-selecionada” e não têm validade científica.
- 4.3. Diversos métodos são utilizados para selecionar amostras aleatórias: aleatória simples, sistemática, estratificada, por conglomerados (aula de amostragem).
- 4.4. Evitando viés (tendência) em seus dados: a amostra deve ser representativa da população que se quer estudar. Alguns possíveis equívocos:
  - 4.4.1. pesquisa de satisfação no trabalho, por telefonema em residência entre 09:00 e 17:00 – não atinge aqueles que trabalham durante o dia fora de casa, aqueles que não tem telefone, etc.
  - 4.4.2. questionários muito extensos: a partir de um certo ponto as respostas se tornam cada vez menos confiáveis – respostas sem maior reflexão.
- 4.5. A coleta de dados propriamente dita: momento crucial do processo. Deve-se garantir que os dados sejam cuidadosamente coletados, buscando-se minimizar o máximo possível as falhas (respostas incompletas, leituras apressadas, etc.).

#### 5. Descrição dos Dados

- 5.1. Em geral, a coleta de dados resulta em grande volume de dados. A descrição desses dados objetiva visualizá-los de forma global e resumida. É comumente realizada de duas formas: visual (gráficos) e numérica.
- 5.2. Estatística Descritiva: números que descrevem importantes aspectos do conjunto de dados. Em geral são acompanhados de gráficos ou mapas.
  - 5.2.1. Se os dados são categóricos (grupos como sexo, partido político, etc.), são comumente resumidos pelo número de indivíduos em cada grupo (frequência), ou porcentagem de indivíduos em cada grupo (frequência relativa).
  - 5.2.2. Se os dados são numéricos (medidas: peso, altura, idade, etc.), além dos agrupamentos acima (faixa de peso, idade, etc.), mais aspectos podem representar os dados: tendência central (média, mediana), medidas de dispersão (variância, desvio padrão, extensão), e em certos casos, a relação entre duas variáveis (peso e altura).
  - 5.2.3. Obs: em alguns casos, números são utilizados para representar variáveis categóricas (sexo M=1, sexo F=2) – mas isso não significa que os dados são numéricos e não faz sentido calcular média desses dados.
- 5.3. Gráficos e mapas: resumem os dados em uma forma visual. Exemplos: distribuição por gênero, uso de celular por faixa de idade, pesquisas de opinião em geral.
  - 5.3.1. Dados categóricos: principalmente pizza e barras
  - 5.3.2. Dados numéricos: principalmente histograma e linha

#### 6. Análise dos dados

- 6.1. Análise Estatística: existem diversos tipos; é fundamental escolher a apropriada.
  - 6.1.1. Escolha do método não apropriado leva a resultados incorretos

- 6.1.2. Cenários onde um número fixo de tentativas independentes resulta em sucesso ou fracasso utilizam a distribuição binomial;
- 6.1.3. Cenários onde a distribuição segue uma curva com formato de sino usam a distribuição normal;
- 6.2. Intervalos de Confiança: são intervalos em torno da média amostral onde se pode ter uma confiança determinada (95%, por exemplo) de se encontrar a média da população;
- 6.3. Teste de Hipóteses: testa se a diferença das médias de duas amostras podem indicar diferença de média entre duas populações (se a amostra é pequena -  $n < 30$  - e a distribuição é “normal”, utiliza-se o teste t);
- 6.4. Correlação: examina a relação entre duas variáveis numéricas (peso e altura) utilizando correlação e regressão linear simples; se as variáveis são categóricas (sexo, cidade onde nasceu, etc.) utiliza-se outro tipo de análise.

## **7. Interpretação dos dados (extrair conclusões)**

- 7.1. O que os resultados significam: atribuição do pesquisador;
- 7.2. Equívocos comuns:
  - 7.2.1. Generalizar resultados que são válidos em condições particulares
  - 7.2.2. Resultados obtidos com amostras tendenciosas;
- 7.3. Para se chegar a conclusões sólidas, é crucial ter conhecimento de todas as etapas do método estatístico, seu significado, suas limitações e sua validade.

## **8. Conceitos Básicos**

- 8.1. Ramos da Estatística:
  - 8.1.1. Estatística Descritiva: a observação de fenômenos de mesma natureza, a coleta de dados numéricos relativos a esses fenômenos, a organização e a classificação desses dados observados e a sua apresentação através de gráficos e tabelas, além da descrição desses dados através do cálculo de coeficientes;
  - 8.1.2. Estatística Inferencial ou Dedutiva: processo de elaboração de conclusões baseadas em informações contidas em uma amostra. A análise e interpretação de dados amostrais que possibilita conclusões sobre a população. Esse processo de generalização está associado a uma margem de incerteza.
    - 8.1.2.1. Inferência estatística: técnicas para chegar a conclusões sobre uma população baseando-se na informação contida na amostra;
- 8.2. População: conjunto de “indivíduos” que apresentam algumas características em comum, sobre os quais desejamos obter alguma informação particular;
  - 8.2.1. Indivíduo: são os objetos descritos por um conjunto de dados. Podem ser pessoas, animais ou coisas;
- 8.3. Amostra: subconjunto da população, definidos segundo algum critério estabelecido pelo observador, sobre o qual se conduz algum estudo com objetivo de se fazer inferências estatísticas sobre toda a população;
- 8.4. Parâmetro: uma medida numérica que descreve uma característica da população
- 8.5. Estatística: uma medida numérica que descreve uma característica da amostra
- 8.6. Variável: uma característica de interesse de elementos individuais de uma população ou amostra que será analisada (Indicadores definidos pelo observador, utilizados para se “medir” alguma característica da amostra); podem ser Quantitativas e Qualitativas

- 8.6.1. Observação: é o valor de uma variável para um elemento particular da amostra
- 8.6.2. Conjunto de dados: é o conjunto das observações de uma variável para todos os elementos da amostra.
- 8.7. Medidas Descritivas: “ouvir o que os dados falam”
  - 8.7.1. Medidas de tendência central: média, mediana, moda (ou modo)
  - 8.7.2. Medidas de dispersão: extensão, variância, desvio padrão
  - 8.7.3. Medidas de variabilidade: percentil, decil, quartil
  - 8.7.4. Medidas de assimetria: curtose, assimetria
- 8.8. Teste de hipóteses: teste t, teste z; erros tipo I e tipo II
- 8.9. Inferência: erro padrão, intervalo de confiança
- 8.10. Regressão e Correlação: medidas de associação entre dois parâmetros
- 8.11. Estatística não paramétrica: a variável ou os dados amostrais não são “normais”, ou são ordinais ou nominais.

