Unidade 1 - Introdução e conceitos iniciais de estatística

Visão geral

Se baseia na coleta de dados e análise desses para estabelecer previsões

Definições:

Dados: consistem em informações provenientes deobservações, contaens,medicaçõesou respostas

Estatística: é a ciência que trata da coleta, organizção, análise e interpretação dos dados para a tomada de decisões. Derivada do latim "status", define um estado, situação.

População: é a coleta de todos os resualtdos, respostas medições ou contagens consideradas de interesse.

Amostra: subconjunto da população deve ser representativa de uma população demodo que seus dados possam ser usados para tirar conclusões sobre aquela população

Conjuntos de dados:

Parâmetro é a descrição numérica de uma característica populacional.

Estatística é a descrição numérica de uma característica amostral

A estatística amostral pode diferir de uma amostra para a outra, enquanto um parâmetro populacional é constante para uma população

Ramos da estatística:

Estatística descritiva é o ramo da estatística que envolve a organização, o resumo e a representação dos dados.

Estatística inferencial é o ramo da estatística que envolve o uso de uma amostra para chegar a conclusões sobre uma população. Uma ferramenta básica no estudo da estatística inferencial é a probabilidade.

Classificação de dados

Tipos de dados:

Dados qualitativos: consistem atributos, rótulos ou entradas não numéricas

Dados quantitativos: consistem em medidas numéricas ou contagens

Níveis de mensuração:

Referentes á dados qualitativos:

Nível **nominal** de mensuração: são apenas dados qualitativos. São **categorizados usando nomes, rótulos ou qualidades**. Não podem ser manipulados em resultados matemáticos

Nível **ordinal** de mensuração: são qualitativos ou quantitativos. Dados nesse nível **podem ser postos em ordem ou classificados, mas as diferenças entre as entradas de dados não tem sentido matemátic**o

Referentes á dados quantitativos:

Nível de mensuração **intervalar**: podem ser o**rdenados e é possível calcular diferenças que tenham sentido matemático entre as entradas de dados**.Nessa mensuração um **zero represanta uma posição** em determinada escala.

Nível de mensuração de **razão**: são **similares aos dados do nível intervalar com apropriedade adicional de que um registro zero representa um zero natural.** Assim pode ser formar uma razão entre dois valores de modo que eles sejam expressos como múltiplos um do outro.

Nível de 
mensuração 
Nominal 
Ordinal 
Intervalar 
Razão 
Categorizar 
os dados 
Sim 
Sim 
Sim 
Sim 
Ordenar os 
dados 
Não 
Sim 
Sim 
Sim 
Determinar se um 
Subtrair os 
dados 
Não 
Não 
Sim 
Sim 
dado é múltiplo 
do outro 
Não 
Não 
Não 
Sim 

Planejamento de um estudo estatístico

Passo a passo:

1. Identificar as variáveis de interesse e a população do estudo
2. Desenvolver um plano detalhado para a coleta de dados. Caso seja utilizada uma amostra deve garantir que esta é representativa da população
3. Coletar dados
4. Descreve-los com técnicas da estatística descritiva
5. Interpretar os dados e tomar decisões sobre a população usando estatística inferencial
6. Identifique quaisquer erros possívies

Estudo observacional:

O pesquisador observa e mede as características de interesse de parte de uma população, mas não muda as condições existentes.

Experimento:

Um tratamento é aplicado em umaparte dapopulçaõ, chamada de grupo de tratamento. Outro grupo,o grupodecontrole, não recebe tratamento. As respostas são comparadas e estudadas.

Coleta de dados

Simulação:

Uso de um modelo matemático ou físico para reproduzir as condições de uma situação ou processo

Pesquisa:

Investigação de uma ou mais características de uma população - mais frequentemente, as pesquisas são conduzidas por pessoas por meio de entrevistas

Planejamento Experimental

Elementos chave: controle, aleatorização e repetição

Experimento cego:

Técnicana qual o indivíduo não sabe seestá recebendo um tratamento ou um placebo.

Existe uma variação desse experimento chamada duplo-cego, nela nem a pessoa nem o aplicador sabe se o elemento aplicado será placebo ou não, assim o controle éobrigtóriamente feito por uma terceira pessoa que informa o pesquisador após a conclusão da pesquisa.

Falhas de experimento:

Variável de confusão: ocorre qunado o pesquisador não pode distiguir um ou mais fatoes que causaram osefeitos provocados sobre a variável em estudo, gerando confusão

Efeito placebo: ocorre quando um indivíduo reage favoravelmente a um tratamento quando na verdade, recebeu um placebo

Aleatorização:

É o processo de se designar indivíduos aleatóriamente para diferentes grupos de tratamento

Planejamento completamente aleatorizado:

Os indivíduos são designados para diferentes grupos de tratamento por meio de seleção aleatória.

Planejamento em blocos aleatorizados:

O pesquisador separa os indivíduos com características similares em blocos e então dentro de cada bloco eles são aleatóriamente designados para groups decontrole ou tratamento

Replicação:

É a repetição de um experimento sob condições iguais ou semelhantes.

Por exemplo se um experimento que trata de uma amostra mas ela não ésuficiente para chegar a uma conclusão verídica, o experimento deverá ser replicado com uma amostra maior.

Técnicas de amostragem

Censo:

Contagem/ medição de toda a população. A realização de um senso fornece informações completas, mas é frequentemente caro e difícil de realizar.

Amostragem:

É uma contagem de parte da população, mais comunmente usada nos estudos estatíscos. Para coleta rdados não viesados, o pesquisador deve se assegurar que a amostra é de fato representativa da população.

Amostragem aleatória:

todosos elementos de uma população tem chances iguais de serem selecionados

Amostragem aleatória simples:

Cada amostra possível de mesmo tamanho tem a mesma chance de ser selecionada

Amostragem estratificada:

elementos de uma população são divididos em dois ou mais subconjuntos, chamdos estratos, que compartilham uma característica similar. Assim uma amostra é selecionada aleatoriamente de cada um dos estratos. O uso de uma amostra estratificada assegura que cada segmento da populçaõ esta representando.

◎ 00 ◎ 0 0 ◎ 〇 0 ◎ 〇 
コ 野 ぎ 一 xa 
Gru を 一 】 
-1 média 
02 宿 卩 
n を 三 
02 宿 3 】 

Amostragem por conglomerado:

Utilizda quando recai subgrupos que ocorrem naturalmente, cada um tendo característics similares. Dividir a população em grupos chamados conglomerados, e selecionar todos os elementos em um ou mais conglomerados sorteados.

Ao usar essa forma de amostragem é preciso assegurar que todos os conglomerados tenham caracteristicas similares

Amostragem por conglomerado — Zonas de códigos postais no 
Condado de West Ridge. 
Zona I 
Zona 2 
Zona 3 
Znna 4 

Amostragem por conveniência:

Consiste somente em membros da população que são faceís de contatar. É uma tecnica tendenciosa deve ser evitada